

حکمرانی هوش مصنوعی (۲): مفاهیم، ابعاد و مؤلفه‌ها



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

شماره مسلسل: ۲۰۳۴۰
کد موضوعی: ۳۵۰



مرکز پژوهش‌های
مجلس شورای اسلامی

تاریخ انتشار:
۱۴۰۳/۱۰/۱۸

عنوان گزارش:

حکمرانی هوش مصنوعی (۲): مفاهیم، ابعاد و مؤلفه‌ها

نوع گزارش: طرح و لایحه □، نظارتی □، راهبردی ■

نام دفتر:

مطالعات حکمرانی (گروه سیاست‌پژوهی و آزمایشگاه حکمرانی)

تهیه و تدوین کنندگان:

مصطفی محمودی، ایمان اکبری، محمدمهدی مهربان (گروه سیاست‌پژوهی و آزمایشگاه حکمرانی)

مدیر مطالعه:

محمدمهدی مهربان هلان، حسین بابایی مجرد

اظهار نظر کننده:

یحیی مرتب (دفتر مطالعات مدیریت)، سهیلا خردمند نیا (دفتر مطالعات صنعت، معدن و انرژی)

اظهار نظر کننده خارج از مرکز:

محمدمهدی قائمی نیا (استادیار دانشکده حکمرانی دانشگاه تهران)

ناظر علمی:

مهدی عبدالحمید

گرافیک و صفحه آرایی:

انسبه بهاء بزرگی
ساجده زارع مرزی

ویراستار ادبی:

سیده مرضیه موسوی راد

واژه‌های کلیدی:

۱. حکمرانی هوش مصنوعی
۲. هوش مصنوعی
۳. قانون نرم
۴. حکمرانی داده
۵. تنظیم‌گری

تاریخ شروع مطالعه:

۱۴۰۲/۱۲/۰۱



فهرست مطالب

چکیده.....	۶
خلاصه مدیریتی.....	۷
۱. مقدمه.....	۹
۲. پیشینه.....	۱۰
۳. چرا حکمرانی هوش مصنوعی؟.....	۱۲
۴. چیستی حکمرانی هوش مصنوعی.....	۱۲
۵. نگاهی به وضعیت کنونی حکمرانی هوش مصنوعی در جهان.....	۱۶
۶. نگاهی آماری به مقوله حکمرانی هوش مصنوعی در جهان.....	۲۴
۷. الزامات طراحی مدل حکمرانی هوش مصنوعی.....	۲۷
۸. جمع بندی و توصیه های سیاستی.....	۳۰
منابع و مأخذ.....	۳۴

فهرست جداول

جدول ۱. تعاریف حکمرانی هوش مصنوعی و مضامین آن در ادبیات حکمرانی هوش مصنوعی.....	۱۴
جدول ۲. مؤلفه های اصلی AIACT.....	۱۸
جدول ۳. جمع بندی وضعیت کنونی حکمرانی هوش مصنوعی در جهان.....	۲۲
جدول ۴. حکمرانی هوش مصنوعی در چند کشور منتخب در سال ۲۰۲۲.....	۲۵

فهرست شکل ها

شکل ۱. تعداد لوايح مربوط به هوش مصنوعی به تفکیک کشور ها، ۲۰۲۲-۲۰۱۶.....	۲۴
شکل ۲. نمودار تعداد لوايح مربوط به هوش مصنوعی که در کشورهای منتخب به تصویب رسیده است ۲۰۱۶-۲۰۲۲.....	۲۵
شکل ۳. تعداد ذکر کلمات «هوش مصنوعی» در رویه های قانونی به تفکیک کشور، ۲۰۲۲-۲۰۱۶.....	۲۶
شکل ۴. کشورهای دارای سندهای راهبردی حکمرانی هوش مصنوعی.....	۲۷
شکل ۵. مدل لایه ای برای حکمرانی هوش مصنوعی.....	۲۸



حکمرانی هوش مصنوعی (۲): مفاهیم، ابعاد و مؤلفه‌ها

چکیده



حکمرانی هوش مصنوعی به عنوان بستر راهبری توسعه هوش مصنوعی، یک مفهوم پویا و چندوجهی بوده و نقش مهمی در ساخت آینده جوامع مدرن ایفا خواهد کرد. حکمرانی هوش مصنوعی شامل مجموعه‌ای از فرایندها، استانداردها، قوانین و خط‌مشی‌ها می‌شود که به منظور راهبری، مدیریت، کنترل بهینه، توسعه و کاربست فناوری هوش مصنوعی و پیامدهای مرتبط با آن طراحی می‌شود. نتیجه آن اقداماتی است که هدف آن تضمین ایمنی، عدالت، شفافیت و اخلاق در توسعه، اجرا و استفاده از سیستم‌ها و فناوری‌های مبتنی بر هوش مصنوعی است. این اقدامات شامل: تعیین استانداردها و دستورالعمل‌ها، مدیریت داده‌ها، حفظ حریم خصوصی، تعیین مسئولیت‌ها و حقوق، مدیریت و نظارت بر الگوریتم‌ها و تضمین توسعه مداوم و پایداری استفاده از هوش مصنوعی برای همه سطوح فنی و دانشی جامعه است.

این گزارش مفاهیم مرتبط با حکمرانی هوش مصنوعی را تبیین کرده است. در این گزارش تلاش شده است ابعاد مختلف حکمرانی هوش مصنوعی مورد بررسی و تحلیل قرار گیرد. در این راستا ابتدا مفاهیم اساسی این حوزه از جمله تعاریف و اهمیت حکمرانی هوش مصنوعی مورد بررسی قرار گرفته است؛ سپس به بررسی چالش‌های موجود در این زمینه پرداخته شده که شامل مسائل فنی، ذی‌نفعان و قواعد کلان (تنظیم‌گری) می‌شود. بررسی آمارهای بین‌المللی نیز نشان می‌دهد که توجه به اهمیت حکمرانی هوش مصنوعی به سرعت در حال گسترش است و تأثیرات آن بر جوامع قابل توجه است. از آنجاکه هوش مصنوعی در سال‌های آتی یک فناوری تحولی، راهبردی در توسعه ابعاد مختلف زندگی اجتماعی خواهد بود، ضرورت پرداختن به حکمرانی هوش مصنوعی در کشور را دوچندان می‌سازد.



■ بیان / شرح مسئله

استفاده فزاینده از هوش مصنوعی در زمینه‌های گوناگون نظیر بهداشت، آموزش، امنیت و خدمات عمومی، اهمیت ایجاد زیرساخت‌های حقوقی و اجرایی مناسب برای مدیریت این فناوری را دوچندان کرده است. کشورهای در حال توسعه همچون ایران، که با چالش‌هایی نظیر محدودیت زیرساخت‌های فنی، نبود نظام جامع حکمرانی داده و عدم آگاهی عمومی مواجه هستند، نیازمند ایجاد چارچوبی منسجم برای هدایت این فناوری‌اند. عدم مدیریت مؤثر بر توسعه و کاربست هوش مصنوعی ممکن است به تشدید نابرابری‌ها، نقض حریم خصوصی، تهدیدات امنیتی و کاهش اعتماد عمومی بینجامد.

■ نقطه نظرات / یافته‌های کلیدی

حکمرانی هوش مصنوعی به عنوان چارچوبی برای مدیریت و کنترل این فناوری، بر اصولی نظیر ایمنی، عدالت، شفافیت، و احترام به حقوق بشر استوار است. این مفهوم تنها به وضع مقررات محدود نمی‌شود، بلکه شامل سازوکارهایی برای ارتقای فرهنگ‌سازی، تقویت دانش عمومی و توانمندسازی ذی‌نفعان نیز می‌شود. علاوه بر این، حکمرانی هوش مصنوعی باید بتواند تعادلی میان نوآوری و رعایت ملاحظات اجتماعی ایجاد کند تا به ارتقای کیفیت زندگی کمک نماید و هم‌زمان از بروز خطرات و پیامدهای ناخواسته جلوگیری کند. با توجه به این مهم و ابلاغ سند ملی هوش مصنوعی که بیانگر اهمیت راهبردی آن است، پرداخت به مفهوم حکمرانی هوش مصنوعی با هدف طراحی قواعد کلان کنشگری ذی‌نفعان این حوزه امری حیاتی است. قواعدی که به‌وضوح اهداف، مسئولیت‌ها و راهبردهای لازم جهت توسعه مؤثر و مطلوب این فناوری را تبیین کند و چابکی لازم برای مواجهه با تعارضات این حوزه را داشته باشند. از وجوه دیگر پرداخت راهبردی به مفهوم حکمرانی هوش مصنوعی، توانمندسازی مدیران و کارکنان بخش دولتی و خصوصی با هدف ارتقای سطح آگاهی و میزان پذیرش فناوری‌های مبتنی بر هوش مصنوعی است که خود از چالش‌های قابل‌توجه پیش‌رو کشورهاست. در همین راستا توصیه‌های زیر پیشنهاد می‌شود:

■ پیشنهاد راهکارهای تقنینی، نظارتی یا سیاستی

👉 **حکمرانی داده:** حکمرانی هوش مصنوعی، پیش از هر چیز، مستلزم طراحی نظام حکمرانی داده است. چراکه داده، به‌مثابه مواد خام در توسعه هوش مصنوعی است. در بند «۱» ماده (۵) سند ملی هوش مصنوعی (راهبردها و اقدامات ملی)، به «لزوم تدوین نظام حکمرانی داده، به‌منظور تسهیل در بهره‌برداری انتفاعی و غیرانتفاعی از داده‌های تولیدی کشور از طریق اصلاح قوانین و مقررات موجود و به رسمیت شناختن مالکیت و حق بهره‌برداری انتفاعی هر شخص یا نهاد بر داده‌های تولیدی خود با حفظ ملاحظات امنیت ملی، رعایت حریم خصوصی و اخلاق عمومی» اشاره شده است.

👉 **تدوین نظام نوآورانه تسهیلگر-تنظیم‌گر جامع و چندلایه:** با توجه به ابلاغ سند ملی هوش مصنوعی و هدف‌گذاری‌های کلان آن، نیل به اهداف سند مستلزم طراحی نظامی است که شامل مجموعه‌ای از رهنمایی‌ها، قوانین و مقررات باشد که تمامی جنبه‌های توسعه و استفاده از هوش مصنوعی را پوشش دهد. این نظام باید به‌گونه‌ای طراحی شود که ضمن تضمین سرمایه‌گذاری استراتژیک، قابلیت انعطاف‌پذیری داشته باشد و بتواند با تغییرات سریع فناوری همگام شود و نیازهای مختلف جامعه را برآورده کند.

👉 **ایجاد سازوکارهای نظارت و ارزیابی:** با توجه به بند «۶» ماده (۴) سند ملی هوش مصنوعی (سیاست‌های راهبردی)، که به ضرورت نظارت و مراقبت بر تصمیمات و خودمختاری هوش مصنوعی و جلوگیری از رخ‌نمانگاری و اعتبارسنجی اجتماعی بدون رعایت حقوق افراد و قوانین کشور تأکید دارد و در راستای اطمینان از رعایت قوانین و مقررات مربوط به حکمرانی هوش مصنوعی، ایجاد سازوکارهایی برای نظارت و ارزیابی منظم، با توجه به اینکه مقوله «نظارت» از کلان‌کارکردهای مجلس شورای اسلامی است، ضروری است. این سازوکارها باید طبق بند «۳-۱» «راهبردها و اقدامات ملی» سند ملی هوش مصنوعی، شامل «تدوین استانداردهای مربوط به مالکیت فکری تولیدات مبتنی بر هوش مصنوعی و به اشتراک‌گذاری داده و همچنین ایمنی و امنیت سیستم‌ها و خدمات هوش مصنوعی» باشد.



- ✓ **توسعه زیرساخت‌های فنی، فرهنگی و ترویجی:** بند «۲» ماده (۵) سند ملی هوش مصنوعی با عنوان «آموزش و پژوهش»، به اهمیت توجه زیرساختی و ساختاری در توسعه منابع انسانی در این حوزه پرداخته است. در این راستا به‌منظور ارتقای نیروی انسانی مؤثر و توسعه هوش مصنوعی در کشور، لازم است که زیرساخت‌های فنی، فرهنگی و ترویجی مناسب ازجمله با بسترسازی تأسیس و گسترش اندیشکده‌ها، کانون‌های فکری و حلقه‌های میانی در حوزه هوش مصنوعی ایجاد شود.
- ✓ **ارتقای همکاری بین‌المللی:** بند «۶» «راهنماها و اقدامات ملی» سند ملی هوش مصنوعی، به «تعاملات بین‌المللی» اشاره دارد. حکمرانی هوش مصنوعی یک مسئله جهانی بوده و نیازمند همکاری و تعامل بین کشورهاست. با توجه به گزاره‌های بند «۶»، ایجاد کنسرسیوم‌های منطقه‌ای و بین‌المللی با هدف تبادل دانش و فناوری و تدوین استانداردهای مشترک برای حکمرانی هوش مصنوعی، می‌تواند به بهبود کیفیت و کارایی حکمرانی این فناوری کمک کرده و از ایجاد انحصار و یکجانبه‌گرایی در حکمرانی جهانی آن پیشگیری کند. این همکاری‌ها می‌توانند به کشورهای مختلف امکان دهند تا با چالش‌های مشابه به‌طور مؤثرتر مقابله کرده و از تجربیات یکدیگر بهره‌برداری کنند.
- ✓ **توسعه مکانیسم‌های مشارکت عمومی:** «تقویت مردم‌محوری و روحیه تعاون در توسعه و بهره‌برداری از هوش مصنوعی»، از موارد مورد توجه سند ملی هوش مصنوعی در ماده (۴) یعنی «سیاست‌های راهبردی» است. مشارکت مردم، نخبگان و ذی‌نفعان مختلف در فرایند تصمیم‌گیری و اجرای خط‌مشی‌های مربوط به حکمرانی هوش مصنوعی بسیار ضروری و ارزشمند است. بنابراین لازم است که مجلس و مرکز پژوهش‌ها، با توجه به ظرفیت‌های ارتباطی و اندیشکده‌های حکمرانی و قانونگذاری استانی خود، سازوکارهایی برای بهره‌گیری نظرات و نگرش‌های گوناگون جامعه و ایجاد فضایی برای ارتباط مستقیم بین حاکمیت و شهروندان و بسط حکمرانی مردمی در توسعه هوش مصنوعی فراهم کنند.
- ✓ **تولید ادبیات حقوقی:** در زمینه مسائل حقوقی، یکی از چالش‌های عمده کشور عدم وجود کلیدواژه‌های تخصصی فارسی ناظر به مفاهیم فنی - حقوقی در حوزه هوش مصنوعی است. از این‌رو پیشنهاد می‌شود برای تبیین مفهومی و ارائه این کلیدواژه‌های فارسی، کمیته‌ای میان‌رشته‌ای با مشارکت بخش‌های مختلف ذی‌نفع در مرکز پژوهش‌های مجلس در این زمینه تشکیل شود.

۱. مقدمه

مفهوم هوش مصنوعی به تلاش‌هایی در توسعه کارکردهایی رایانه‌مبنا جهت تقلید هوش انسانی با بهره‌مندی از توانایی تفسیر و یادگیری از داده‌ها، می‌پردازد. برای نمونه شبکه‌های عصبی مصنوعی برای درک و تقلید از فعالیت نورون‌ها در مغز انسان طراحی شد؛ بعدها شبکه‌های عصبی پیچیده‌تری در اواخر قرن بیستم توسعه یافت و توانمندی‌هایی مانند تشخیص تصویر و گفتار را در این سیستم‌ها ایجاد کرد [۱].

این تلاش‌ها منجر به معرفی مفهوم هوش مصنوعی به‌عنوان برنامه‌های کامپیوتری (یا ماشین‌هایی) شد که می‌توانند وظایف از پیش تعریف شده را با سرعت و دقت بسیار بالاتر انجام دهند. در موج اخیر پیشرفت‌های هوش مصنوعی که با پیشرفت در تجزیه و تحلیل کلان‌داده تسهیل شده است، قابلیت‌های هوش مصنوعی به برنامه‌های رایانه‌ای که می‌توانند حجم عظیمی از داده‌ها را یاد گرفته و بدون راهنمایی انسان تصمیم بگیرند - که معمولاً به‌عنوان الگوریتم‌های یادگیری ماشینی نامیده می‌شوند - گسترش یافته است [۱].

برخلاف الگوریتم‌های ابتدایی هوش مصنوعی که بر قوانین از پیش برنامه‌ریزی شده برای اجرای وظایف تکراری تکیه می‌کردند، الگوریتم‌های یادگیری ماشینی با محاسبات آماری و قوانینی درباره نحوه یادگیری از داده‌ها طراحی شده‌اند که شامل «استدلال استنتاجی»، «ادراک»، «طبقه‌بندی» و «بهینه‌سازی» جهت تقلید اعمال انسانی است [۳] و [۴]. این اعمال می‌تواند شامل پیش‌بینی، درک، مواجهه، توصیه و مواردی از این دست که متأثر از هوش انسانی هستند، باشد. ظرفیت‌های حاصل از فناوری هوش مصنوعی در تحول ابعاد تعامل حاکمیت‌ها و شهروندان قابل توجه است. از جمله ظرفیت پیشرفت در پزشکی و بهداشت عمومی، حمل‌ونقل عمومی، انرژی، آموزش، علم، امنیت، رشد اقتصادی و پایداری محیط زیست است [۵]. با این حال، خطرات آن نیز قابل توجه هستند و چالش‌های شدیدی را برای حکمرانی ایجاد می‌کنند. با رصد الگوی زمانی مواجهه با فناوری‌های نوین می‌توان پیش‌بینی کرد که احتمالاً در مواجهه به هوش مصنوعی، سیاستمداران، خط‌مشی‌گذاران و سایر ذی‌نفعان عمدتاً بر خطرات و آسیب‌های فناوری‌های مبتنی بر هوش مصنوعی تمرکز می‌کنند [۴]. در مورد تأثیر فناوری‌های دیجیتال بر جامعه، چالش‌های مربوط به هوش مصنوعی، سیستم‌های مستقل و الگوریتم‌ها، اغلب در قالب فهرستی از موضوعات اساسی که شامل ملاحظات حکمرانی، سیاسی، حقوقی و اخلاقی است مورد بحث قرار می‌گیرند [۵].

هدف کلی این گزارش، تشریح مفهوم حکمرانی بر هوش مصنوعی به تناسب کارکردها و چالش‌های این فناوری است. درک مشترک از مفاهیم هوش مصنوعی برای خط‌مشی‌گذاران، مدیران دولتی و شهروندان، هم‌زمان با درهم‌آمیختگی مفهوم هوش مصنوعی و امور حاکمیتی بسیار مهم است. با کاوش در مفهوم و چارچوب حکمرانی بر هوش مصنوعی، هدف ما روشن کردن ابعاد این مفهوم و کمک به بازتولید و توسعه کارآمد حکمرانی بر این حوزه از علم و فناوری است. با بررسی انتقادی وضعیت حکمرانی بر هوش مصنوعی در جهان، این گزارش به دنبال تبیین ارتباط با حکمرانی بر هوش مصنوعی و توسعه پایدار هوش مصنوعی مسئول و مؤثر است.



۲-۱. سوابق مطالعاتی (ترجیحاً در مرکز)

گزارش‌های متعددی در مرکز پژوهش‌های مجلس با کلیدواژه هوش مصنوعی تدوین شده که شامل بررسی ابعاد گوناگون این فناوری است. برای مثال سلسله گزارش‌های «هوش مصنوعی در جهان» [۶] کشورهایی را به صورت مجزا مورد بررسی قرار داده است. سه گزارش نیز تحت عناوین: «چالش‌ها و راهکارهای تحقق حکمرانی هوشمند» [۷]، «تصمیم‌گیری مبتنی بر داده: رویکردی جهت تحول حکمرانی پارلمانی» [۸] و «نگاشت نهادی و تقسیم‌کار ملی در حوزه توسعه هوش مصنوعی و حکمرانی داده‌محور» [۹] در سال ۱۴۰۲ در مرکز به چاپ رسید که هر کدام به نوعی به مقوله حکمرانی هوش مصنوعی و رویکرد داده‌مبنا اشاره داشتند. دفتر مطالعات حکمرانی در گزارش دیگری با عنوان «بررسی لایحه برنامه هفتم توسعه (۸۸): توسعه پایدار هوش مصنوعی در کشور» [۱۰] آورده است که برنامه هفتم توسعه به عنوان ریل‌گذار پیشرفت کشور در پنج سال آینده، نقش مهمی در توسعه پایدار فناوری‌های نوین در کشور خواهد داشت. از جمله فناوری‌هایی که طی سال‌های اجرای برنامه هفتم توسعه نقشی راهبردی در پیشرفت کشور خواهد داشت، فناوری هوش مصنوعی است.

نظر به اهمیت بحث هوش مصنوعی در کشور و تأکیدات مقام معظم رهبری در سال ۱۴۰۳ مبنی بر آن، با تکیه بر گزارش‌های پیشین و مفاهیم بحث شده، در گزارش حاضر به وجوه گوناگون مفهوم حکمرانی هوش مصنوعی با هدف واکاوی ابعاد و وجوه مختلف فناوری هوش مصنوعی در جامعه با یک نگاه کل‌نگر و ارائه یک تصویر کلی از آن، پرداخته شده است. لازمه این امر، ارائه پیشنهادها و عملیاتی برای کشور، بررسی تجربیات بین‌المللی و کشورهای پیشرو در این حوزه و در مرحله بعد پیشنهاد مدلی جهت مواجهه با هوش مصنوعی در کشور و طراحی نظام حکمرانی هوش مصنوعی مطلوب در کشور است که در این گزارش مدنظر خواهد بود.

۲-۲. سوابق تقنینی به همراه آسیب‌شناسی^۵

در ارتباط با هوش مصنوعی در کشور، ماده (۶۵) قانون برنامه هفتم بند «ج» دولت را مکلف به نوشتن برنامه ملی هوش مصنوعی کرده است. ضمناً سند ملی توسعه هوش مصنوعی ابلاغی سال ۱۴۰۳ حاوی موادی مرتبط با این گزارش است. ضروری است این سند که ایجاد خطوط راهنما و چارچوب‌های کلی توسعه هوش مصنوعی در کشور را مدنظر داشته است، در شیوه مواجهه با این فناوری و توسعه و جهت‌دهی حکمرانی هوش مصنوعی در کشور، مدنظر باشد. بنابراین کلیات این سند در ادامه بررسی خواهد شد. در ماده (۲)، اصول و مبانی ارزش حاکم بر این سند را شامل: اصل اعتقادی و ارزش‌های اسلامی، توجه به ابعاد وجودی انسان مبتنی بر فلسفه و معارف اسلامی، تأمین آرمان‌ها و ارزش‌های جامعه اسلامی، خوداتکایی، کرامت و عدالت، استقلال و امنیت، آینده‌نگری، امنیت داده‌ها، انحصارزدایی، جلوگیری از سلطه هوش مصنوعی بر انسان‌ها و تقویت ارتباطات انسانی و مواردی از این دست می‌داند.

در ماده (۳)، درخصوص اهداف کلا بر: ۱. تربیت سرمایه انسانی، ۲. ارتقا و تأمین زیرساخت‌ها، ۳. جهش تولیدات فکری، علمی و فناورانه، ۴. ارتقای نوآوری برپایه فناوری‌های بدیع با اثرگذاری اجتماعی و اقتصادی بالا، ۵. ارتقای رقابت‌پذیری اقتصادی کشور، ۶.

۱. شماره مسلسل: ۱۸۹۵۴

۲. شماره مسلسل: ۱۸۸۶۲

۳. شماره مسلسل: ۱۹۴۳۴

۴. شماره مسلسل: ۱۹۳۹۵

۵. شامل قوانین مصوب مجلس شورای اسلامی و نتایج و ارزیابی از قوانین مصوب و مصوبات شورای عالی (مانند شورای عالی انقلاب فرهنگی و مصوبات هیئت‌وزیران)، ضروری است درخصوص بررسی سوابق در بخش بین‌الملل؛ عنوان موارد بررسی شده، سال اجرای آنها و مهم‌ترین یافته، در قالب یک جدول ذکر شود. ۶ ج - دولت مکلف است در راستای حمایت از توسعه زیست‌بوم تحول‌آفرین هوش مصنوعی قابل اعتماد و پایدار و به‌منظور تعیین چهارچوب و ساز و کار تعامل تمامی ذی‌نفعان، فراهم نمودن دانش و زیرساخت‌های دانش فنی، اجتماعی، اخلاقی و حقوقی، ترویج و افزایش آگاهی در مورد کارکردهای هوش مصنوعی در زمینه‌های مختلف و خطرات بالقوه آن حداکثر ظرف شش‌ماه از لازم‌الاجرا شدن این قانون نسبت به اجرای «برنامه ملی توسعه هوش مصنوعی» با رعایت سیاست‌های کلی نظام، مصوبات شورای عالی انقلاب فرهنگی و سند راهبردی جمهوری اسلامی ایران در فضای مجازی اقدام قانونی به عمل آورد.

ارتقای تعاملات آموزشی، فناوریانه و اقتصادی بین‌المللی و ۷. ارتقای کیفیت حکمرانی تأکید شده است. همچنین شاخص‌هایی جهت ارزیابی کلان هر کدام از اهداف مذکور در بازه زمانی پنج‌ساله ذکر شده است.

ماده (۴)، مرتبط با سیاست‌های راهبردی را می‌توان به صورت کلی در این حوزه‌ها دسته‌بندی کرد: ۱. تحریک رشد اقتصادی و اشتغال و ارزش‌دهی به فعالیت‌های بنگاه‌های اقتصادی مرتبط با هوش مصنوعی و ایجاد زنجیره کامل ایده تا ثروت در این زمینه، ۲. پیشرفت در راستای ارزش‌های اسلامی و تقویت مردم‌محوری و تعاون، روابط اجتماعی، عدالت و قدرت ملی، ۳. تأکید بر حقوق افراد و قوانین کشور و مسئولیت‌پذیری و صیانت از اخلاق، کرامت انسانی، اعتماد به نفس، حافظه انسانی، خلاقیت و حفظ حریم خصوصی در توسعه این فناوری، ۴. توسعه کاربردهایی با فراگیری و اثرگذاری بالا و رافع چالش‌های ملی و خدمات عمومی، ۵. فرصت‌آفرینی برای زیست‌بوم هوش مصنوعی در تعاملات بین‌المللی.

در ماده (۵)، راهبردها و اقدامات ملی مدنظر قرار گرفته: ۱. زیرساخت‌های حکمرانی (فنی، قانونی، تنظیم‌گری و استانداردسازی)، ۲. آموزش و پرورش (مهارت‌افزایی و تربیت نیروی انسانی)، ۳. زیرساخت داده و پردازش (ارتقای زیرساخت‌های نگهداری، امنیت، تاب‌آوری و تسهیل دسترسی به داده و امکانات پردازشی)، ۴. توسعه محصولات فناوریانه و تجاری‌سازی، ۵. توسعه بازار در زیست‌بوم هوش مصنوعی، ۶. تعاملات بین‌المللی در حوزه هوش مصنوعی، ۷. ترویج و فرهنگ‌سازی (ارتقای سواد عمومی و تناسب فرهنگی در استفاده صحیح از این فناوری) مورد تأکید قرار گرفته است.

ماده (۶)، به بیان اولویت‌های کشور در محورهای آموزش و پرورش، بهداشت و درمان، حکمرانی دولتی و خدمات عمومی، دفاعی، امنیتی و انتظامی، صنایع و انرژی، محیط زیست و کشاورزی، زبان و ادبیات فارسی، فرهنگ تمدن و سبک زندگی اسلامی-ایرانی، حقوقی و قضایی و رسانه و فضای مجازی پرداخته و در هر محور اولویت‌ها را ذیل دو اولویت الف و ب مشخص کرده است. این ماده سازمان هوش مصنوعی را موظف به تهیه برنامه‌های ملی کاربرست هوش مصنوعی در این حوزه‌ها با همکاری دستگاه‌های مرتبط ظرف مدت حداکثر ۶ ماه کرده است.

در ماده (۷)، جهت تدبیر سازوکار اجرا برای این سند، نظارت کلان و سیاستگذاری بر آن را برعهده شورای عالی انقلاب فرهنگی قرار داده و ستاد علم، فناوری و نوآوری این شورا را مسئول رصد اجرای این سند و بازنگری‌های لازم قرار داده است.

همچنین این ماده شورای ملی راهبری هوش مصنوعی را جهت برنامه‌ریزی، راهبری و هماهنگی اجرای این سند با ترکیب مشخص از دستگاه‌های مختلف با حضور رئیس سازمان ملی هوش مصنوعی به‌عنوان دبیر شورا، تعیین کرده است. به‌علاوه سازمان ملی هوش مصنوعی ذیل ریاست‌جمهور، مأمور به سازمان‌دهی و توسعه زیست‌بوم هوش مصنوعی در کشور شده است.

در نهایت دستگاه‌های اجرایی مسئول اجرای این سند، موظف به برنامه‌ریزی جهت لحاظ قرار دادن اقدامات موضوع مواد (۵) و (۶) در اولویت سالیانه و تدبیر منابع و ظرفیت‌های لازم در راستای تحقق اهداف این سند شده‌اند.

به‌نظر می‌رسد این سند به‌رغم جامعیت و تلاش در راستای شمولیت موضوعات مختلف تا حدودی دچار عدم انسجام و پراکندگی محتوایی شده است. برای نمونه برخی از عناوین ذیل مواد مختلف، با یکدیگر هم‌پوشانی داشته و برخی بسیار کلی و عمومی و برخی بسیار جزئی و پروژه‌ای، به بیان موارد پرداخته‌اند. حال آنکه از یک سند ملی، بیان اهداف و خط‌مشی‌ها به شکلی نظام‌مند و ساختاریافته با در نظرگیری ملاحظات و ظرفیت‌های موجود کشور در کنار تعیین اولویت‌های ملی براساس این ملاحظات و ظرفیت‌ها در بازه زمانی مشخص، انتظار می‌رفت. پرداختن به انبوهی از مسائل، دغدغه‌ها و اولویت‌ها که در سطح بین‌المللی مطرح شده هر کدام در جای خود صحیح و قابل بررسی است، بدون تعیین اولویت‌های اصلی و تناسب با شرایط کشور ظرفیت‌ها و تهدیدهای پیش رو و پیشنهاد حدود چهل مورد اولویت در حوزه‌های مختلف که غالباً یا کلی‌گویی و یا بسیار موردی و پروژه‌ای مطرح شده است، شایسته سند ملی هوش مصنوعی کشور نیست.

بااین حال کشور پس از عبور از چالش و کشمکش تدوین سند میان دستگاه‌های مختلف، اکنون صاحب یک سند ملی بوده که ضروری است با ملاحظه انتقادهای وارده، دستگاه‌های مختلف اجرایی و تقنینی براساس این سند، اکنون به عرصه اقدام و عمل وارد شوند و در چارچوب‌های کلان مؤلفه‌های حکمرانی هوش مصنوعی لازم است برنامه‌ریزی‌های عملیاتی خود را آغاز کنند پیش از آنکه کشور در تحولات لحظه‌شمار این حوزه دچار عقب‌ماندگی غیرقابل جبران شود. در ضمن ضروری است باید با توجه به تحولات سریع این فناوری چه در کشور و چه در سطح جهان و منطقه، امکان بازنگری و به‌روزرسانی این سند، برای جبران خلأها و نقص‌ها فراهم باشد.



۳. چرا حکمرانی هوش مصنوعی؟

امروزه اهمیت توجه به فناوری هوش مصنوعی و توسعه آن در کشور، بر کسی پوشیده نیست. رهبر معظم انقلاب پس از تأکید بر قرارگیری کشور در ۱۰ کشور برتر هوش مصنوعی، در اولین جلسه با هیئت دولت چهاردهم پس از ذکر مواردی از به کارگیری هوش مصنوعی در کشور فرمودند که در مسئله هوش مصنوعی بهره‌بردار بودن امتیاز نیست، این فناوری لایه‌های عمیقی دارد باید بر آن لایه‌ها مسلط شد و این این لایه‌ها دست دیگران است. ایشان پس از اهمیت تسریع در تحقق این اهداف پیش از ایجاد سازمان‌های بین‌المللی بازدارنده از توسعه این فناوری در سایر کشورها فرمودند: «خودتان باید برسید به فناوری‌های عمیق و ژرف این مسئله، لایه‌های زیرساختی. لایه‌های زیرساختی هوش مصنوعی را باید در کشور دنبال کنید»^۱. توسعه پایدار هوش مصنوعی و دستیابی به لایه‌های زیرساختی مورد اشاره، جز با وجود یک نظام ارزیابی، برنامه‌ریزی، اجرا و رصد به همراه تدبیر سازوکارهای نهادی مورد نیاز، در کشور ممکن نخواهد بود. به علاوه، با توجه به ظرفیت‌ها و توانمندی‌های هوش مصنوعی، کاربست هوش مصنوعی بدون نظارت و راهبری صحیح، می‌تواند آسیب‌ها و خطرات متعددی از جمله آسیب‌های اجتماعی و اخلاقی قابل توجهی ایجاد کند. مجموع این تدابیر در مواجهه با هوش مصنوعی از جمله برنامه‌ریزی، رصد و ارزیابی، اجرا و تدبیر سازوکار و فرایند، نهاد و تقسیم وظایف، قوانین و مقررات، اصول و چارچوب‌ها و... را حکمرانی هوش مصنوعی می‌نامیم. هدف حکمرانی هوش مصنوعی، بسترسازی و تسهیل توسعه این فناوری با ایجاد توازن بین نوآوری و ملاحظات قانونی، شرعی و اخلاقی به نحوی است که اطمینان حاصل شود که در کنار بهره‌مندی حداکثری از توانمندی‌های روزافزون این فناوری، سیستم‌های هوش مصنوعی، کرامت و حقوق اساسی انسان و جوامع را نقض نمی‌کنند.

۴. چستی حکمرانی هوش مصنوعی

حکمرانی هوش مصنوعی به چارچوبی اشاره دارد که تضمین می‌کند ابزارها و سیستم‌های هوش مصنوعی ایمن و اخلاقی هستند و در ادامه نیز بر همین منوال باقی خواهد ماند. حکمرانی هوش مصنوعی می‌تواند شامل قانون و مقررده گذاری، استانداردسازی، تنظیم‌گری، خط‌مشی‌گذاری، نهادسازی، حمایت، تسهیلگری و بسترسازی، بازدارندگی و جرم‌انگاری و... باشد که می‌تواند مسیر کاربست، تحقیقات و توسعه هوش مصنوعی را برای اطمینان از رعایت عدالت، آزادی، کرامت، انصاف و احترام به نسبت به حقوق شهروندان، هدایت کند [۱۱].

همچنین حکمرانی هوش مصنوعی شامل مکانیسم‌های نظارتی می‌شود که خطراتی مانند تعصب یا جانبداری، نقض حریم خصوصی، تشدید نابرابری‌ها، نقض آزادی‌های مدنی و سوءاستفاده را مورد بررسی قرار می‌دهد و درعین حال، نوآوری و اعتماد را تقویت می‌کند. یک رویکرد اخلاقی برای حکمرانی هوش مصنوعی، مستلزم مشارکت طیف گسترده‌ای از ذی‌نفعان، از جمله توسعه‌دهندگان، کاربران، خط‌مشی‌گذاران و اخلاق‌گرایان هوش مصنوعی است تا اطمینان حاصل شود که سیستم‌های مرتبط با هوش مصنوعی برای همسویی با ارزش‌های جامعه توسعه یافته و استفاده می‌شوند [۱۲].

همچنین حکمرانی هوش مصنوعی به نقص‌های ذاتی ناشی از عنصر انسانی در ایجاد و نگهداری هوش مصنوعی می‌پردازد. از آنجاکه هوش مصنوعی یک محصول مهندسی شده و حاصل یادگیری ماشینی بوده که توسط افراد ایجاد شده، مستعد سوگیری‌ها و خطاهای انسانی نیز است. حکمرانی هوش مصنوعی یک رویکرد ساختاریافته برای کاهش این خطرات ارائه کرده و تضمین می‌کند که الگوریتم‌های یادگیری ماشین، نظارت، ارزیابی و به‌روزرسانی می‌شوند تا از تصمیم‌های ناقص یا مضر جلوگیری شود [۱۲]. با توجه به رشد سریع و گسترش روزافزون فناوری‌های مبتنی بر هوش مصنوعی، پرداختن به حکمرانی بر این فناوری‌ها، اهمیتی دوچندان یافته است. به دلیل پیچیدگی‌ها و تأثیرات گسترده هوش مصنوعی بر ابعاد مختلف زندگی بشر، ضروری است که مدل‌های حکمرانی

۱. بیانات مقام معظم رهبری در تاریخ ۱۴۰۳/۶/۶.

متناسبی طراحی و پیاده‌سازی شوند. این مدل‌ها نه تنها باید بتوانند از مخاطرات احتمالی پیشگیری کنند، بلکه باید به گونه‌ای طراحی شوند که از مزایای بالقوه هوش مصنوعی به نفع جامعه بهره‌برداری کنند. بنابراین ضروری است در ابتدا، ابعاد و مؤلفه‌های این حکمرانی شناسایی شود.

۴-۱. ابعاد و مؤلفه‌های اساسی حکمرانی هوش مصنوعی

یک مطالعه سیاستی درباره نقشه راه حوزه خط‌مشی هوش مصنوعی که در سال ۲۰۱۷ انجام شده نشان می‌دهد مسائل این حوزه در پنج محور کلان قرار می‌گیرد [۸]:

- ✓ **عدالت و برابری:**^۱ تا چه حد می‌توان سیستم‌های هوش مصنوعی را طوری طراحی و اجرا کرد که ارزش‌های انسانی مانند انصاف، مسئولیت‌پذیری و شفافیت را منعکس کرده و از نابرابری‌ها و سوگیری‌ها جلوگیری کند؟
- ✓ **استفاده از زور و اجبار:**^۲ از آنجا که سیستم‌های مبتنی بر هوش مصنوعی اکنون در تصمیم‌گیری در مورد استفاده از زور دخیل هستند، برای مثال در مورد سلاح‌های خودمختار چه مقدار کنترل انسانی لازم یا مورد نیاز است؟ چه کسی مسئولیت عواقب بر هوش مصنوعی را برعهده دارد؟
- ✓ **ایمنی و گواهینامه:**^۳ به‌ویژه در جایی که سیستم‌های مبتنی بر هوش مصنوعی جلوه فیزیکی دارند، چگونه آستانه‌های ایمنی را تعریف و تأیید می‌کنیم. برای مثال، از طریق تنظیم استاندارد و صدور گواهینامه؟
- ✓ **حریم خصوصی:**^۴ از آنجا که هوش مصنوعی توسط داده‌های آنلاین تغذیه می‌شوند، پیامدهای حریم خصوصی و تهدیدات حریم خصوصی جدید فناوری‌های نسل بعدی چیست؟ برای مثال، از نظر نظارت دولت یا نفوذ شرکت بر مشتریان؟
- ✓ **جابه‌جایی نیروی کار و مالیات:**^۵ ماشین‌های مبتنی بر هوش مصنوعی تا چه اندازه جایگزین کارهایی می‌شوند که قبلاً توسط انسان انجام شده است؟ آیا معنای کار را تغییر می‌دهند؟ اگر ربات‌ها مالیات نپردازند، هوش مصنوعی چه تأثیری بر امور مالی عمومی دارد؟ در ادبیات این حوزه، تعداد کمی از مقالات به‌صراحت حکمرانی هوش مصنوعی را تعریف می‌کنند [۱۹]. تعریف حکمرانی هوش مصنوعی چالش‌برانگیز است. زیرا فقدان اجماع آکادمیک در مورد تعریف هوش مصنوعی و عناصر تشکیل‌دهنده آن وجود دارد [۲۰] و [۲۱].



جدول ۱. تعاریف حکمرانی هوش مصنوعی و مضامین آن در ادبیات حکمرانی هوش مصنوعی [۲۲]

مأخذ	تعریف حکمرانی هوش مصنوعی
[۲۱]	حکمرانی هوش مصنوعی با حفظ تعادل بین تحقق منافع و بهینه‌سازی سطوح ریسک و استفاده از منابع، به دنبال ایجاد ارزش بهینه در این حوزه است. هدف اصلی اجرای حکمرانی هوش مصنوعی این است که اطمینان حاصل شود که سازمان‌های دولتی قادر به انجام مدیریت ریسک، ارائه ارزش، همسویی استراتژیک مدیریت منابع و مدیریت عملکرد هستند تا برای همه ذی‌نفعان ارزش ایجاد کنند.
[۲۳]	از منظر ایمنی هوش مصنوعی، حکمرانی سیستم‌های هوش مصنوعی مستلزم حل چالش همسویی ارزش‌هاست که هدف آن پیاده‌سازی سیستم‌های هوش مصنوعی است تا با ارزش‌های اخلاقی انسانی همسو باشند.
[۲۴]	حکمرانی هوش مصنوعی را می‌توان به‌عنوان ابزارها، راه‌حل‌ها و اهرم‌های مختلفی که بر توسعه برنامه‌های مبتنی بر هوش مصنوعی تأثیر می‌گذارند، توصیف کرد. برخی از نمونه‌ها عبارتند از: ارتقای چارچوب‌های هنجاری، اخلاقیات و ارزش‌ها (که ممکن است به شکل خودتنظیمی از سوی شرکت‌های پیشرو فناوری که روی پروژه‌های خاصی کار می‌کنند ارائه شود)، تحقیق در مورد آثار، پیامدها و راه‌حل‌های ممکن برای استفاده از هوش مصنوعی، افزایش آگاهی برای ذی‌نفعان، تولید راه‌حل‌های فنی که با مسائل خاصی که توسط فناوری هوش مصنوعی مطرح می‌شود (مانند تفسیرپذیری الگوریتمی و توضیح‌پذیری که توانایی درک دقیق نحوه تصمیم‌گیری یک الگوریتم است) و اجرای اقدامات قانونی و ایجاد نهادهای رسمی نظارتی که صلاحیت اداره فناوری‌ها و زمینه‌های مرتبط با هوش مصنوعی را دارند.
[۲۵]	مجموعه رو به رشدی از ادبیات که پرسش‌های مربوط به هوش مصنوعی و چارچوب‌های اخلاقی، قوانین حاکم بر تأثیر هوش مصنوعی و رباتیک، رویکردهای فنی مانند ارزیابی تأثیر الگوریتمی و ایجاد قابلیت اعتماد از طریق اعتبارسنجی سیستم را پوشش می‌دهد. این سه نیروی راهنما در حکمرانی هوش مصنوعی (قانون، اخلاق و فناوری) می‌توانند مکمل یکدیگر باشند.
[۲۶]	«حکمرانی هوش مصنوعی شامل ملاحظات سیاسی، نظامی، اقتصادی، حاکمیتی و اخلاقی و ابعاد پیچیده است که هوش مصنوعی پیشرفته در جامعه ایجاد می‌کند.» حکمرانی هوش مصنوعی را می‌توان به اجزای دیگر یعنی «چشم‌انداز فنی، حکومت آرمانی و خط‌مشی هوش مصنوعی» تفکیک کرد.

الیمن و کستر^۱ در سال ۲۰۱۹ دیدگاه متفاوتی را اتخاذ کردند. آنها با بیان اینکه حکمرانی هوش مصنوعی بر پرداختن به الزامات اخلاقی جامعه، با تأکید ویژه بر همسویی ارزش‌های انسانی با اجرای هوش مصنوعی متمرکز است، بر (اهمیت) ارزش‌های اجتماعی و اخلاقی تأکید کردند [۲۳]. زادجالی^۲ در سال ۲۰۲۰ بر نقش دولت‌ها در رابطه با مفهوم‌پردازی حکمرانی هوش مصنوعی تأکید کرد و اظهار داشت که این شامل ظرفیت سازمان‌های حاکمیتی برای ایجاد ارزش برای همه ذی‌نفعان از طریق عملکردهای نظیر همسویی ارزش و مدیریت عملکرد است [۲۷].

سه موضوع تکرارشونده از ادبیات حوزه حکمرانی هوش مصنوعی استخراج شده است که می‌توان آنها را جزو مؤلفه‌های حکمرانی هوش مصنوعی دانست که عبارتند از:

1. Aliman and Kester
2. Ali Zadjali

۱ حکمرانی فناوری

موضوع حکمرانی فناوری، متمرکز بر داده‌ها و الگوریتم‌هایی است که پایه‌های فناوری‌های هوش مصنوعی هستند و همچنین چالش‌های حاکمیتی مرتبط با ویژگی‌های فنی سیستم‌های هوش مصنوعی را مورد بحث قرار می‌دهد. ادبیات بررسی شده شامل مطالعات متعددی است که چارچوب‌هایی را برای حکمرانی هوش مصنوعی با محوریت داده‌ها و الگوریتم‌ها توسعه می‌دهند و از حوزه‌هایی مانند ایمنی هوش مصنوعی و نوآوری مسئولانه استفاده می‌کنند [۲۸] و [۲۹].

دیدگاه فناوری شامل سه سطح به هم پیوسته است: داده‌ها، الگوریتم‌ها و سیستم‌های هوش مصنوعی. سیستم‌های هوش مصنوعی با بهره‌گیری از الگوریتم‌ها بر اساس داده‌ها یاد می‌گیرند و تطبیق می‌یابند. بنابراین جای تعجب نیست که «حکمرانی داده» به عنوان یک ستون کلیدی «حکمرانی هوش مصنوعی» در نظر گرفته شود [۳۰]. مطالعاتی که بر حکمرانی داده متمرکز شده‌اند، موضوعاتی مانند حریم خصوصی داده‌ها، حفاظت قانونی و یکپارچگی را مورد بحث قرار داده‌اند که جزء لاینفک ابتکارات عملی حکمرانی هوش مصنوعی هستند [۳۱]. علاوه بر این، نقش داده‌های آموزشی به صراحت در ادبیات مورد بحث قرار می‌گیرد تا از عملکرد مناسب سیستم‌های هوش مصنوعی اطمینان حاصل شود.

برای مثال، استفاده از مجموعه داده‌های آموزشی متعدد به طور بالقوه می‌تواند پاسخ‌گویی را بهبود بخشد و داده‌های موجود در رسانه‌های اجتماعی، گزارش‌های مربوط به اقدامات بشردوستانه و تخلفات اخلاقی توسط نهادهای عمومی را می‌توان برای تسهیل خودآموزی سیستم‌های هوش مصنوعی برای رفتار اخلاقی مورد استفاده قرار داد [۳۲].

شفافیت، توضیح‌پذیری و غیرقابل بررسی بودن الگوریتم‌ها، چالش‌های مهمی در حکمرانی هوش مصنوعی هستند. از این رو، الگوریتم‌ها و داده‌های مورد استفاده برای توسعه و آموزش مدل‌های هوش مصنوعی در مطالعات بررسی شده، با تأکید بر پاسخ‌گویی، شفافیت و توضیح‌پذیری مدل‌ها و الگوریتم‌های هوش مصنوعی مورد بحث قرار گرفته‌اند [۳۳]. علاوه بر این، شفافیت الگوریتم‌ها و سیستم‌های هوش مصنوعی در اصول، دستورالعمل‌های نظارتی و اسناد موجود، وضعیت حیاتی دارد. مشکلات «جعبه سیاه» و کدورت توسط محققان به عنوان چالش‌های حکمرانی هوش مصنوعی درک می‌شوند که نیاز به راه‌حل‌های فنی دارند [۲۹] و [۳۴].

علاوه بر شفافیت، حکمرانی هوش مصنوعی با مسائل مربوط به عدالت و آسیب‌های احتمالی نیز سروکار دارد و الگوریتم‌های هوش مصنوعی به طور بالقوه می‌توانند پیامدهای نامطلوب اجتماعی متعددی ایجاد کنند [۳۰].

بنابراین نگرانی‌هایی در مورد سوءاستفاده و تبعیض احتمالی مرتبط با سیستم‌های هوش مصنوعی مطرح می‌شود. موضوعات مرتبط با «سوگیری»، یکی دیگر از نگرانی‌های کلیدی در ادبیات حکمرانی هوش مصنوعی است [۳۳]. برای مثال، فرض بر این است که سوگیری در هوش مصنوعی می‌تواند به دلیل طراحی، کد یا داده‌های آن، ایجاد شود و هر منبع سوگیری نیاز به یک رویکرد کاهش یا اصلاح متمایز دارد.

چنین تلاش‌هایی در مقابل حکمرانی هوش مصنوعی می‌تواند قابلیت اطمینان در سیستم‌های هوش مصنوعی را تأیید کند و منافع ذی‌نفعان (مانند مصرف‌کنندگان و شهروندان) را که تحت تأثیر آن قرار دارند، متعادل کند [۲۹]. در حالی که این تصور رایج است که سیستم‌های هوش مصنوعی جعبه‌های سیاه غیرقابل کاوش هستند، کرول^۱ (۲۰۱۸) پیشنهاد کرد که این سیستم‌ها اساساً قابل درک هستند و تصور فعلی غیرقابل تشخیص بودن هوش مصنوعی می‌تواند ناشی از پویایی قدرت موجود در این حوزه باشد [۳۵].

۲ ذی‌نفعان

ذی‌نفعان به عنوان موضوع دوم، به بازیگران و نقش‌های درگیر در توسعه، استقرار و اداره سیستم‌های هوش مصنوعی اشاره دارد. در لایه کلان، ذی‌نفعان حکمرانی هوش مصنوعی، وزارتخانه‌ها، سازمان‌ها و ادارات هستند که باید با توجه به اهداف برنامه‌های توسعه ملی، ظرفیت پذیرش ابزارهای مبتنی بر هوش مصنوعی را در خود ایجاد کرده، فراگردهای دولتی را بهبود بخشند و با ارتقای توانمندی امنیتی، حافظ امنیت اطلاعات و حریم خصوصی شهروندان باشند. از طرف دیگر قوه مقننه موظف است که با ایجاد یک نظام پویا با هدف تسهیلگری و بازتوزیع صحیح این حوزه، به صورت مستمر و دائمی، قواعد و قوانین زمین بازی حکمرانی هوش مصنوعی را با هدف بهبود مستمر این حوزه، تغییر دهد.

1. Kroll



هنگام بحث در مورد پویایی ذی‌نفعان حکمرانی هوش مصنوعی، مطالعات بررسی شده نشان می‌دهد که تخصص فنی و قدرت عمدتاً در میان تعداد کمی از بازیگران، به‌ویژه در شرکت‌های فناوری بزرگ با مجموعه داده‌های گسترده یافت می‌شود [۳۴]. عدم تعادل قدرت بین بازیگرانی که مشخصات سیستم‌های هوش مصنوعی را تعیین می‌کنند (یعنی مقررات، هنجارهای سازمانی و شرکت‌ها) و متخصصانی که آنها را اجرا می‌کنند، وجود دارد. این موضوع می‌تواند به تخصیص نامتعادل مسئولیت‌های حکمرانی هوش مصنوعی بین ذی‌نفعان تبدیل شود که در آن ممکن است به نفع سازمان‌های خصوصی، به‌ویژه سازمان‌هایی که در اقتصادهای توسعه یافته‌تر مستقر هستند، اعمال شود [۳۶].

علاوه بر توزیع نابرابر دانش فنی و قدرت، سلسله‌مراتب دیگری نیز بین ذی‌نفعان مختلف حکمرانی هوش مصنوعی وجود دارد [۲۸] و [۳۶]. این سلسله‌مراتب شامل میزان نفوذ، مسئولیت‌ها و سطوح پاسخ‌گویی متفاوت آنهاست. برای مثال، نهادهای بزرگ‌تر مانند دولت‌ها معمولاً با ایجاد و اصلاح دستورالعمل‌های موجود، نقش نظارتی قوی‌تری در حکمرانی هوش مصنوعی ایفا می‌کنند [۳۴].

۳۳ مقررات و فرایندها

موضوع سوم به قانون و مقررگذاری و تنظیم‌گری در حکمرانی هوش مصنوعی اشاره دارد. مقررات موجود برای حکمرانی هوش مصنوعی شامل قوانین سخت (قانون الزام‌آور) و رویکردهای حکمرانی نرم، از جمله استانداردها، گواهی‌ها و ممیزی‌هاست [۳۷]. قوانین سخت، مانند قوانین ضد تبعیض و مقررات حفاظت از داده‌های عمومی اتحادیه اروپا، پایه‌های نظارتی ضروری حکمرانی هوش مصنوعی هستند [۳۸].

مکانیسم‌های حکمرانی نرم (مانند استانداردها)، درجات کمتری از رسمیت نهادی و انعطاف‌پذیری بیشتری نسبت به قوانین سخت دارند. استانداردسازی اجرای عملی حکمرانی هوش مصنوعی در مراحل شکل‌گیری خود است [۳۹]. سازمان بین‌المللی استاندارد ISO و IEEE دو نهاد مهم استانداردسازی در حکمرانی هوش مصنوعی هستند. بر این اساس، ایزو، یک گروه کاری برای قابلیت اطمینان هوش مصنوعی راه‌اندازی کرده است تا به شفافیت و قابلیت ردیابی هوش مصنوعی رسیدگی کند. به‌طور مشابه، IEEE سری استانداردهای P7000 را ایجاد کرده که P7001 مربوط به استانداردسازی شفافیت سیستم‌های مستقل و تعیین دلایل تصمیم‌گیری آن است [۳۹]. چنین استانداردهایی می‌توانند توسط ذی‌نفعان گوناگون برای توسعه سطوح شفافیت قابل اندازه‌گیری استفاده شوند.

حکمرانی نرم همچنین شامل اصولی مانند استحکام فنی، امنیت داده‌ها، حریم خصوصی، قابلیت اطمینان، شفافیت، انصاف و مسئولیت‌پذیری است [۱].

در ادامه جهت ادراک بهتر مفاهیم فوق، به‌صورت خلاصه، به مواردی از «حکمرانی هوش مصنوعی» در جهان اشاره می‌شود.

۵. نگاهی به وضعیت کنونی حکمرانی هوش مصنوعی در جهان

در این بخش به‌صورت گذرا، رویکردها و فعالیت‌های پنج مورد مطالعاتی در حوزه حکمرانی هوش مصنوعی بیان می‌شود که شامل: اتحادیه اروپا، آمریکا، چین، هند و سنگاپور است.

از آنجاکه بخش‌های گوناگون حاکمیت نیز با مقوله هوش مصنوعی سروکار دارند، آنها نیز با ابهامات اخلاقی و قانونی جدیدی مواجه خواهند بود. اندیشکده‌ها، سازمان‌های تحقیقاتی و سایر گروه‌ها در حال ارائه توصیه‌هایی به خط‌مشی‌گذاران در مورد چگونگی اطمینان از استفاده مسئولانه و اخلاقی از هوش مصنوعی هستند. قانون مهم حفاظت از داده‌های عمومی اتحادیه اروپا^۱ که در سال ۲۰۱۸ اجرایی شد، نمونه‌ای از این تلاش‌های آغازین بود و به قانون جامع AI ACT اتحادیه اروپا منتهی شد. برخی از کشورها، اصولی را برای هوش مصنوعی تدوین کرده‌اند، درحالی‌که برخی دیگر در حال تهیه پیش‌نویس قوانین و مقررات هستند. کاسینز

1. General Data Protection Regulation (GDPR)

نیومن می‌گوید: «برنامه‌های خاص کاربردی مبتنی بر هوش مصنوعی وجود دارد که ممکن است به‌قدری برای گروه‌هایی از مردم مضر باشد که فرایندهای خودتنظیمی صنعت کافی نباشد. دولت‌ها در حال بیدار شدن از این واقعیت هستند که باید در مورد این موضوع به روشی کاملاً جامع فکر کنند». براندون پرسل، تحلیلگر اصلی در **Forrester Research**، می‌افزاید: «قانونگذاران باید به فناوری‌ها برسند» و تکنولوژی با سرعت سرسام‌آوری در حال حرکت است. در این بخش از گزارش به تفکیک مواردی از مواجهه کشورها با بحث حکمرانی هوش مصنوعی بیان شده است. پرداخت به مقوله حکمرانی هوش مصنوعی، به مانند خود هوش مصنوعی، به‌صورت مستمر و دائمی در حال توسعه و پیشرفت است. درحالی‌که برخی از کشورها، لوایح و قوانین سیاستی در حوزه حکمرانی هوش مصنوعی تدوین کرده‌اند، بخش قابل توجهی از کشورهای جهان، همچنان در مرحله پژوهش و مطالعه این مهم هستند.

۱-۵. اروپا

اتحادیه اروپا، خود را به‌عنوان یک پیشرو جهانی در بحث حکمرانی هوش مصنوعی معرفی کرده است. مقررات عمومی حفاظت از داده‌ها در سال ۲۰۱۸ اجرایی شد و دامنه آن به برخی از سازمان‌های خارج از اتحادیه اروپا نیز گسترش یافت [۴۱]. برخی از کشورهای اروپایی نیز در حال بررسی مقررات خود هستند و گام‌های اولیه را برای اعمال سیاست‌ها برداشته‌اند. استفاده از هوش مصنوعی در اتحادیه اروپا توسط قانون اقدام هوش مصنوعی^۱، اولین قانون جامع هوش مصنوعی در جهان، سامان‌دهی شده است.^۲ اتحادیه اروپا به‌عنوان بخشی از استراتژی دیجیتالی خود می‌خواهد هوش مصنوعی را تنظیم‌گری کند تا شرایط بهتری را برای توسعه و استفاده از این فناوری نوآورانه تضمین کند.

مجموعه قوانین **AI Act**، به دنبال توسعه ایمن هوش مصنوعی است و بر خطرات مرتبط با کاربست چنین فناوری‌ای تأکید می‌کند. اتحادیه اروپا در آوریل سال ۲۰۲۱ پیش‌نویس قانون هوش مصنوعی (**AI Act**) را ارائه کرد. این قانون، اولین تلاش جامع در سطح جهانی برای تنظیم‌گری هوش مصنوعی بوده و هدف آن اطمینان از ایمنی و احترام به حقوق اساسی افراد است. این قانون با هدف تقویت زیرساخت‌های یک زیست‌بوم قابل اعتماد برای توسعه هوش مصنوعی مسئول، طراحی و تصویب شده است. این چارچوب باید تضمین کند که سامانه‌های هوش مصنوعی داخل اتحادیه اروپا با امنیت و رعایت قوانین موجود در مورد حقوق اساسی و ارزش‌های اتحادیه سازگار باشند؛ همچنین تضمین حقوقی را برای تسهیل سرمایه‌گذاری و نوآوری در زمینه هوش مصنوعی فراهم کند. قانون **AI Act**، از دوازده محور اصلی تشکیل شده که به شرح زیر است:

1. AI Act

2. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regula>



جدول ۲. مؤلفه‌های اصلی AIAC [۳۶]

عناوین	شرح
موضوع و محدوده	این مورد، محدوده اعمال قوانین جدید را مشخص می‌کند و عرضه در بازار، راه‌اندازی و استفاده از سیستم‌های هوش مصنوعی را پوشش می‌دهد.
لیست انواع هوش مصنوعی ممنوعه	تمام سامانه‌های مبتنی بر هوش مصنوعی را فهرست کرده و توصیف می‌کند که برای مثال با نقض برخی از حقوق اساسی، به‌عنوان ناقض ارزش‌های اتحادیه اروپا، غیرقابل قبول تلقی می‌شوند.
سامانه‌های هوش مصنوعی پرخطر	حاوی قوانین خاصی برای سامانه‌هایی است که ریسک بالایی ایجاد می‌کنند. طبقه‌بندی یک سامانه هوش مصنوعی به‌عنوان پرخطر، براساس هدف مورد نظر سامانه هوش مصنوعی است.
تعهدات شفافیت	درباره تعهدات شفافیت سامانه‌هایی است که مستقیماً با انسان‌ها تعامل دارند که برای مواردی مانند تشخیص احساسات یا تعیین ارتباطات اجتماعی براساس داده‌های بیومتریک یا تولید داده استفاده می‌شوند.
حمایت از نوآوری	به هدف ایجاد یک چارچوب قانونی که مانع نوآوری نمی‌شود کمک می‌کند.
سیستم‌های حکمرانی	■ تنظیم‌گری سامانه‌های حکمرانی، در سطح اتحادیه ملی و اروپا، ■ تسهیل‌گری در بحث نظارت کمیسیون اروپا و مقامات ملی، ■ تعیین تعهدات نظارت و گزارش برای ارائه‌دهندگان هوش مصنوعی.
تعهدات داوطلبانه	چارچوبی برای تشویق ارائه‌دهندگان سامانه‌های هوش مصنوعی غیرخطرناک به‌منظور اعمال داوطلبانه الزامات سیستم‌های پرخطر.
جریمه‌ها، به‌روزرسانی‌ها، تفویض قدرت	■ قوانینی را برای تبادل اطلاعات به‌دست آمده تعیین می‌کند. ■ قوانینی را برای اعمال تفویض اختیارات اجرایی تعیین می‌کند. ■ اینکه شامل تعهدی برای کمیسیون اروپا جهت ارزیابی منظم نیاز به‌روزرسانی و تهیه گزارش‌های زمان‌بندی شده در این موارد است.

مأخذ: پارلمان اروپا، ۲۰۲۳.

۲-۵. آمریکا

دفتر سیاست علم و فناوری کاخ سفید در زمان رئیس‌جمهور اسبق باراک اوباما مجموعه‌ای از کارگاه‌های آموزشی در مورد هوش مصنوعی برگزار کرد که منجر به انتشار دو مقاله در سال ۲۰۱۶ شد: «برنامه استراتژیک تحقیق و توسعه هوش مصنوعی ملی» و «آماده‌سازی برای آینده هوش مصنوعی».^۱

■ برنامه استراتژیک تحقیق و توسعه هوش مصنوعی ملی آمریکا

■ هدف:

هدف از برنامه استراتژیک تحقیق و توسعه ملی هوش مصنوعی، هدایت سرمایه‌گذاری در تحقیق و توسعه هوش مصنوعی است تا اطمینان حاصل شود که آمریکا در این زمینه یک رهبر جهانی باقی می‌ماند. این سند بر روی تحقیقات بلندمدت هوش مصنوعی تمرکز دارد که ممکن است فوراً برای صنعت سودآور نباشد، بنابراین نیاز به حمایت دولت دارد.

■ مباحث کلیدی:

این طرح، هفت اولویت استراتژیک را برای تحقیقات هوش مصنوعی با بودجه فدرال ترسیم می‌کند:

۱ **سرمایه‌گذاری بلندمدت در تحقیقات هوش مصنوعی:** سرمایه‌گذاری مستمر در تحقیقات اساسی هوش مصنوعی برای ایجاد پیشرفت‌های آینده،

۲ **همکاری انسان و هوش مصنوعی:** توسعه سیستم‌های هوش مصنوعی که به‌جای جایگزینی انسان‌ها، به‌طور مؤثر با انسان‌ها

1. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

همکاری می کنند،

۳ پیامدهای اخلاقی، حقوقی و اجتماعی: درک و رسیدگی به خطرات بالقوه هوش مصنوعی و همسویی توسعه هوش مصنوعی با ارزش های اجتماعی،

۴ ایمنی و امنیت هوش مصنوعی: تضمین ایمنی و قابل اعتماد بودن سیستم های هوش مصنوعی قبل از اینکه به طور گسترده به کار گرفته شوند،

۵ دیتابیس های اشتراکی عمومی: ایجاد مجموعه داده ها^۱ و محیط های آزمایشی^۲ با کیفیت بالا و در دسترس برای پیشبرد تحقیقات هوش مصنوعی،

۶ استانداردها و شاخص های هوش مصنوعی: توسعه شاخص هایی برای ارزیابی پیشرفت هوش مصنوعی و هدایت مسیرهای تحقیقاتی،

۷ نیروی کار ملی تحقیق و توسعه هوش مصنوعی: ایجاد و حفظ نیروی کار تحقیقاتی قوی هوش مصنوعی برای رسیدگی به چالش های ذکر شده در طرح.

■ نتایج مورد انتظار:

- ✓ ایجاد شکوفایی اقتصادی،
- ✓ بهبود کیفیت زندگی و آموزش،
- ✓ ارتقای امنیت ملی و داخلی.

■ سند آماده سازی برای آینده هوش مصنوعی آمریکا

این سند که توسط شورای ملی علم و فناوری (NSTC) تهیه شده، مزایا و خطرات بالقوه فناوری های هوش مصنوعی (AI) را تشریح کرده و توصیه هایی را برای دولت آمریکا پیشنهاد می کند. همچنین پیشنهادهایی برای آمادگی در برابر نقش مهمی که هوش مصنوعی در آینده ایفا می کند، ارائه کرده است. این سند، چندین مزیت بالقوه هوش مصنوعی را برجسته می کند، از جمله:

- ✓ بهبود خدمات عمومی: هوش مصنوعی به طور بالقوه می تواند کارایی خدمات دولتی را با خودکار کردن وظایف و فعال کردن تصمیم گیری مبتنی بر داده بهبود بخشد.
- ✓ رشد اقتصادی: هوش مصنوعی پتانسیل بی نظیری در افزایش بهره وری و ایجاد فرصت های اقتصادی جدید دارد.
- ✓ پیشرفت در حل چالش های اجتماعی: هوش مصنوعی می تواند برای مقابله با چالش های اجتماعی گوناگون مانند مراقبت های بهداشتی، آموزش عمومی و تغییرات آب و هوایی مورد استفاده بیشتری قرار گیرد.

■ خطرات بالقوه هوش مصنوعی:

- ✓ جابه جایی شغل: خودکارسازی ناشی از کاربست هوش مصنوعی می تواند منجر به از دست دادن مشاغل در بخش های خاصی از نیروی کار شود.
 - ✓ جانبداری و انصاف: اگر سیستم های هوش مصنوعی به دقت طراحی و اجرا نشوند، می توانند تعصبات اجتماعی موجود را تداوم یا تقویت کنند.
 - ✓ ایمنی و امنیت: اطمینان از ایمنی و امنیت سیستم های مبتنی بر هوش مصنوعی، به ویژه در برنامه های کاربردی عمومی، یک چالش حیاتی است.
- این سند چندین توصیه به دولت آمریکا برای رسیدگی به خطرات و فرصت های بالقوه هوش مصنوعی ارائه می کند، از جمله:
- ✓ سرمایه گذاری در تحقیق و توسعه هوش مصنوعی،

1. Data Base
2. testing Environments



- ✓ توسعه نیروی کار ماهر،
- ✓ رسیدگی به نگرانی‌های اخلاقی،
- ✓ ترویج همکاری بین‌المللی.

به‌طور کلی، این سند بر اهمیت آماده‌سازی برای آینده هوش مصنوعی به‌گونه‌ای تأکید می‌کند که مزایای بالقوه آن را به حداکثر برساند و خطرات بالقوه آن را کاهش دهد.

۳-۵. چین

در چند سال گذشته، چین قوانین و مقررات ملی الزام‌آوری را در مورد هوش مصنوعی (AI) وضع کرده است. این مقررات، توصیه‌هایی را نیز در جهت تنظیم‌گری انتشار محتوا (اعم از متون، تصاویر و ویدئوهای تولید شده مصنوعی و سیستم‌های هوش مصنوعی مانند چت جی‌پی‌تی) ارائه می‌کنند. این قوانین، الزامات جدیدی را برای نحوه ساخت و استقرار الگوریتم‌ها و همچنین برای اطلاعاتی که توسعه‌دهندگان هوش مصنوعی باید برای دولت و مردم افشا کنند، ایجاد می‌کند. این اقدامات، زمینه‌های فکری و بوروکراتیک را برای یک قانون جامع ملی هوش مصنوعی ایجاد می‌کند که چین احتمالاً در سال‌های آینده منتشر خواهد کرد که احتمالاً یک پیشرفت بالقوه مهم برای حکمرانی جهانی هوش مصنوعی خواهد بود.

■ مقررات کلیدی:

- ✓ **الگوریتم‌های توصیه‌گر (۲۰۲۱):** این مقررات، الگوریتم‌هایی را که برای انتشار محتوا به‌صورت آنلاین استفاده می‌شوند، هدف قرار می‌دهند و الزاماتی را برای شفافیت، کنترل کاربر و حفاظت از اطلاعات، مشمول مدیریت الگوریتمی می‌دانند.^۱
- ✓ **سنتز عمیق یا ترکیب‌بندی عمیق Deep Synthesis (۲۰۲۲):** این مقررات بر محتوای تولید شده مصنوعی (برای مثال دیپ فیک) تمرکز دارد که برای شناسایی چنین محتوایی نیاز به برچسب زدن واضح^۲ دارد.
- ✓ **هوش مصنوعی مولد (پیش‌نویس ۲۰۲۳):** این پیش‌نویس مقررات به سیستم‌هایی مانند چت جی‌پی‌تی می‌پردازد و الزام می‌کند که داده‌های آموزشی و خروجی‌های مدل‌های زبانی، «درست و دقیق» باشند. همچنین یک سیستم رجیستری دولتی برای الگوریتم‌ها معرفی می‌کند.^۴

■ انگیزه‌ها:

- ✓ **کنترل اطلاعات:** هدف اصلی این مقررات تقویت کنترل دولت چین بر اطلاعات آنلاین است.
- ✓ **ثبات اجتماعی:** از اهداف این مقررات کاهش خطرات بالقوه برای ثبات اجتماعی، مانند انتشار اطلاعات نادرست یا محتوای مضر است.
- ✓ **توسعه اقتصادی:** چین به دنبال تقویت توسعه مسئولانه صنعت هوش مصنوعی داخلی خود و درعین‌حال شکل دادن به حکمرانی جهانی هوش مصنوعی است.

■ نظام بوروکراتیک حکمرانی هوش مصنوعی چین:

- ✓ **اداره فضای سایبری چین: CAC** ^۶ این اداره نقش پیشرو در تهیه پیش‌نویس مقررات هوش مصنوعی ایفا می‌کند که نشان‌دهنده تمرکز آن بر کنترل محتوای آنلاین است.
- ✓ **وزارت علوم و فناوری: MOST** ^۵ این وزارتخانه به‌طور فزاینده‌ای در مدیریت هوش مصنوعی درگیر می‌شود، به‌ویژه با تغییر

1. <https://www.china-briefing.com/news/china-passes-sweeping-recommendation-algorithm-regulations-effect-march-1-2022/>
 2. clear labeling
 3. <https://www.china-briefing.com/news/china-to-regulate-deep-synthesis-deep-fake-technolo>
 4. <https://www.globalpolicywatch.com/2023/04/china-proposes-draft-measures-to-regulate-generative-ai/>
 5. Ministry of Science and Technology (MOST)
 6. Cyberspace Administration of China (CAC)

تمرکز به سمت تنظیم‌گری تحقیقات زیربنایی هوش مصنوعی.

❖ **اندیشکده‌ها و محققان:** مؤسساتی مانند آکادمی فناوری ارتباطات اطلاعات چین^۱ و مؤسسه مدیریت بین‌المللی هوش مصنوعی دانشگاه Tsinghua^۲ به شکل‌دهی چشم‌انداز فکری حکمرانی هوش مصنوعی چین کمک می‌کنند.

۴-۵. هند

دولت فدرال هند، در ژوئن سال ۲۰۱۸ یک استراتژی ملی هوش مصنوعی را به تصویب رساند الزام می‌کند که یک مرکز مطالعاتی در زمینه پایداری فناوری برای رسیدگی به مسائل مربوط به اخلاق، حریم خصوصی و موارد دیگر راه‌اندازی شود.^۳ استراتژی ملی هند برای حکمرانی هوش مصنوعی با تأکید بر مفهوم «هوش مصنوعی برای همه»^۴ به دنبال هدایت مسیر توسعه و پذیرش هوش مصنوعی در کشور است. هدف این سند راهبردی استفاده از هوش مصنوعی برای رشد فراگیر و در ادامه نیز قرار دادن هند به‌عنوان یک بازیگر جهانی در بحث حکمرانی هوش مصنوعی است. این سند، پنج بخش دارای اولویت را (که هوش مصنوعی در آن دارای پتانسیل تأثیر اجتماعی و اقتصادی قابل‌توجهی در هند است) مورد هدف قرار داده است:

❖ **مراقبت‌های بهداشتی:** افزایش دسترسی به مراقبت‌های بهداشتی مقرون‌به‌صرفه، به‌ویژه در مناطق دور افتاده و بهبود تشخیص و درمان بیماری،

❖ **کشاورزی:** افزایش بازده محصولات، بهبود استفاده از منابع و تضمین رشد درآمد کشاورزان،

❖ **آموزش:** ارائه تجربیات یادگیری شخصی، پر کردن شکاف کیفیت و بهبود دسترسی به آموزش، به‌ویژه در مناطق محروم،

❖ **شهرها و زیرساخت‌های هوشمند:** پرداختن به چالش‌های شهرنشینی، از طریق مدیریت کارآمد ترافیک، دفع زباله و بهینه‌سازی منابع،

❖ **تحرك و حمل‌ونقل هوشمند:** توسعه سیستم‌های حمل‌ونقل ایمن‌تر، کارآمدتر و کاهش ازدحام.

■ چالش‌های پذیرش هوش مصنوعی در هند:

❖ **فقدان تخصص هوش مصنوعی:** سرمایه انسانی محدود در تحقیق و توسعه هوش مصنوعی اصلی و کاربردی،

❖ **دسترسی به داده‌ها:** عدم وجود مجموعه داده‌های ساختاریافته و با کیفیت بالا و مکانیسم‌هایی برای به اشتراک‌گذاری امن داده‌ها،

❖ **هزینه‌های بالای منابع:** محدودیت‌های مالی و شکاف‌هایی در حوزه آگاهی از فناوری، مانع پذیرش هوش مصنوعی می‌شود،

❖ **حریم خصوصی و امنیت:** نگرانی در مورد حریم خصوصی داده‌ها و سوءاستفاده احتمالی از فناوری‌های هوش مصنوعی،

❖ **همکاری:** نیاز به مشارکت قوی‌تر بین دانشگاه، صنعت و دولت برای هدایت نوآوری هوش مصنوعی.

۵-۵. سنگاپور

دولت فدرال سنگاپور، چارچوبی برای حکمرانی هوش مصنوعی در ژانویه سال ۲۰۱۹ منتشر کرد و راهنمایی‌هایی را به بخش خصوصی در مورد چگونگی رسیدگی به مسائل اخلاقی ناشی از استفاده از هوش مصنوعی ارائه داد.^۵ هدف سند استراتژی ملی هوش مصنوعی سنگاپور، تبدیل این کشور به‌عنوان یک بازیگر جهانی در توسعه و استقرار راه‌حل‌های هوش مصنوعی است. این استراتژی بر رویکردهای کلان کل کشور تأثیر دارد و به دنبال تغییر اقتصاد، جامعه و دولت سنگاپور از طریق پذیرش هوش مصنوعی^۶ است.

1. <http://www.caict.ac.cn/english/>

2. <https://aiig.tsinghua.edu.cn/en/>

3. <https://www.niti.gov.in/>

4. AI for All

5. <https://www.opengovasia.com/singapore-releases-first-artificial-intelligence-ai-governance-framework-in-asia/>

6. AI Adoption



■ چشم‌انداز سند:

۱. سنگاپور به یک مرکز جهانی برای توسعه، آزمایش، استقرار و مقیاس‌بندی راه‌حل‌های هوش مصنوعی، از جمله رسیدگی به چالش‌های حکمرانی هوش مصنوعی تبدیل شود،
۲. هوش مصنوعی برای ایجاد ارزش اقتصادی و بهبود زندگی شهروندان سنگاپور استفاده شود،
۳. سنگاپوری‌ها به مهارت‌های لازم هوش مصنوعی مجهز شوند و مزایای بالقوه هوش مصنوعی را درک کنند.

■ پروژه‌های ملی هوش مصنوعی (برای هدایت پذیرش هوش مصنوعی و نشان دادن پتانسیل تحول‌آفرین):

۱. برنامه‌ریزی هوشمند حمل‌ونقل: بهینه‌سازی جابه‌جایی بار برای بهبود کارایی و کاهش تراکم ترافیک،
۲. خدمات شهری یکپارچه و کارآمد: افزایش پاسخ‌گویی و قابلیت اطمینان خدمات شهری به کمک هوش مصنوعی،
۳. پیش‌بینی و مدیریت بیماری‌ها: استفاده از هوش مصنوعی برای پیشگیری و مدیریت بهتر بیماری‌های همه‌گیر،
۴. آموزش شخصی‌شده از طریق یادگیری و ارزیابی تطبیقی: شخصی‌سازی تجربیات یادگیری و بهبود نتایج آموزشی،
۵. عملیات پاکسازی و امنیت مرزها: تقویت امنیت مرزی و درعین‌حال، بهبود تجربه مسافران.

۵-۶. جمع‌بندی

جدول زیر خلاصه‌ای از رویکردهای کلیدی، فعالیت‌ها و قواعد اصلی و اهداف و نتایج مورد انتظار از حکمرانی هوش مصنوعی در پنج کشور/ منطقه مختلف را نشان می‌دهد. هر کشور/ منطقه با توجه به نیازها و اولویت‌های خود، رویکردها و مقررات متفاوتی را برای مدیریت و استفاده مسئولانه از هوش مصنوعی اتخاذ کرده است.

جدول ۳. جمع‌بندی وضعیت کنونی حکمرانی هوش مصنوعی در جهان

کشور/ منطقه	رویکرد حکمرانی هوش مصنوعی	فعالیت‌ها و قواعد کلیدی	اهداف و نتایج مورد انتظار	نقاط قوت و ضعف
اتحادیه اروپا	- پیشرو در تنظیم‌گری جامع هوش مصنوعی، - تأکید بر ایمنی و حقوق اساسی، - ایجاد یک چارچوب قانونی برای نوآوری و مسئولیت‌پذیری.	- GDPR (۲۰۱۸): مقررات عمومی حفاظت از داده‌ها، - AI ACT (۲۰۲۱): اولین قانون جامع هوش مصنوعی در جهان، - تعیین مقررات برای هوش مصنوعی پرخطر، تعهدات شفافیت، حمایت از نوآوری و سیستم‌های حکمرانی.	- اطمینان از ایمنی و رعایت حقوق اساسی، - تقویت زیرساخت‌های قانونی برای نوآوری و سرمایه‌گذاری، - ایجاد یک زیست‌بوم قابل اعتماد برای توسعه هوش مصنوعی.	نقاط قوت: چارچوب قانونی جامع و جهانی برای هوش مصنوعی، تمرکز بر حقوق اساسی و ایمنی، پیشرو در تدوین مقررات. نقاط ضعف: فرایندهای پیچیده بوروکراتیک که ممکن است نوآوری را کندتر کند، هزینه‌های بالای پیاده‌سازی و پایبندی به قوانین برای کسب‌وکارها.

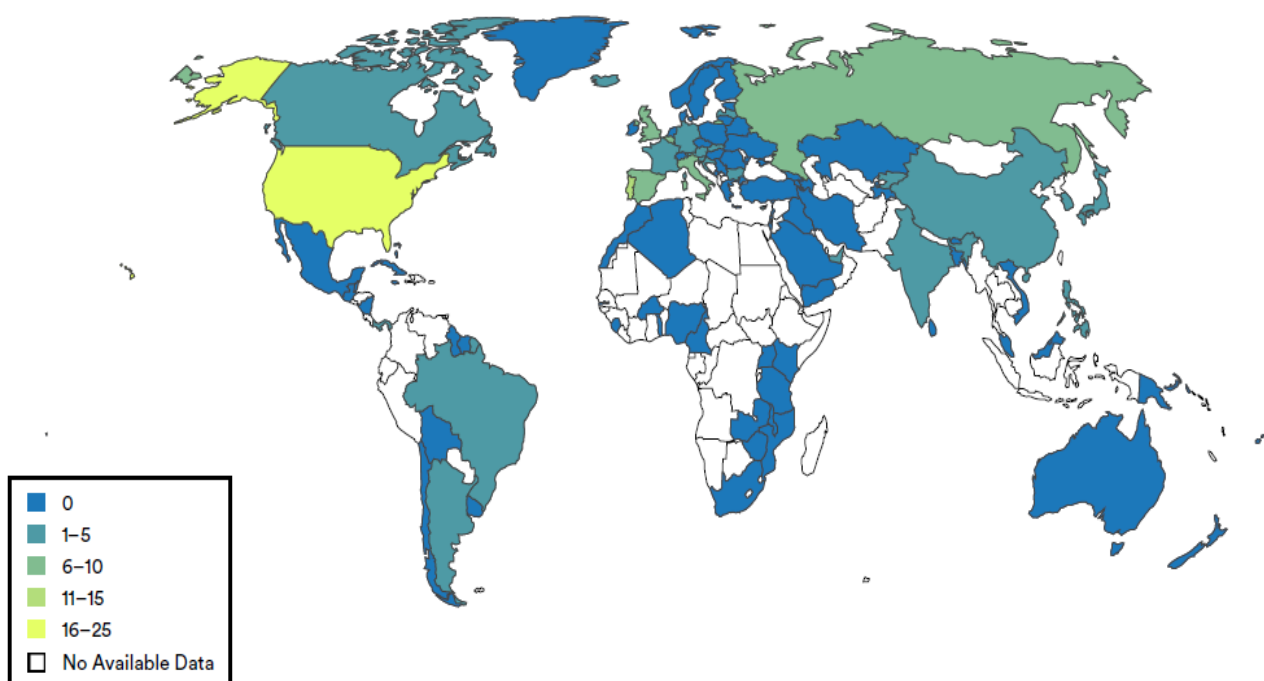
کشور منطقه	رویکرد حکمرانی هوش مصنوعی	فعالیت‌ها و قواعد کلیدی	اهداف و نتایج مورد انتظار	نقاط قوت و ضعف
آمریکا	- تمرکز بر تحقیق و توسعه بلندمدت، - تأکید بر همکاری انسان و هوش مصنوعی، - توجه به پیامدهای اخلاقی و اجتماعی.	- سند استراتژیک تحقیق و توسعه هوش مصنوعی ملی (۲۰۱۶): هدایت سرمایه‌گذاری در تحقیق و توسعه هوش مصنوعی، - آماده‌سازی برای آینده هوش مصنوعی: تشریح مزایا و خطرات بالقوه هوش مصنوعی و توصیه‌هایی به دولت برای آماده‌سازی.	- ایجاد شکوفایی اقتصادی و بهبود کیفیت زندگی، - تقویت امنیت ملی و داخلی، - افزایش بهره‌وری و ایجاد فرصت‌های اقتصادی جدید.	نقاط قوت: رهبری در تحقیق و توسعه، تمرکز بر نوآوری و رشد اقتصادی، همکاری با بخش خصوصی برای توسعه فناوری. نقاط ضعف: عدم وجود چارچوب قانونی جامع و الزام‌آور، کندی در تنظیم‌گری نسبت به سرعت پیشرفت فناوری، وابستگی به شرکت‌های بزرگ فناوری.
چین	- تنظیم‌گری دقیق با تأکید بر کنترل اطلاعات و ثبات اجتماعی، - تدوین مقررات برای الگوریتم‌ها و محتوای تولید شده مصنوعی.	- مقررات الگوریتم‌های توصیه‌گر: الزامات شفافیت و حفاظت از اطلاعات، - ستیز عمیق (۲۰۲۲): مقررات برای محتوای تولید شده مصنوعی مانند دیپ فیک، - هوش مصنوعی مولد (۲۰۲۳): الزام به دقت و صحت داده‌های آموزشی و خروجی‌های مدل‌های زبانی.	- کنترل اطلاعات و کاهش خطرات برای ثبات اجتماعی، - تقویت توسعه مسئولانه صنعت هوش مصنوعی داخلی، - شکل‌دهی حکمرانی جهانی هوش مصنوعی.	نقاط قوت: توانایی در تنظیم‌گری سریع و دقیق، کنترل قوی بر اطلاعات و محتوا، توسعه سریع صنعت هوش مصنوعی داخلی. نقاط ضعف: محدودیت‌های ناشی از کنترل شدید دولت، نگرانی‌های مربوط به حقوق بشر و حریم خصوصی، ریسک کاهش نوآوری به دلیل قوانین سختگیرانه.
هند	- تأکید بر مفهوم «هوش مصنوعی برای همه»، - استفاده از هوش مصنوعی برای رشد فراگیر و توسعه پایدار.	- استراتژی ملی هوش مصنوعی (۲۰۱۸): راه‌اندازی مرکز مطالعاتی برای مسائل اخلاقی و حریم خصوصی، - هدف‌گذاری در حوزه‌های اولویت‌دار مانند بهداشت، کشاورزی، آموزش، شهرها و زیرساخت‌های هوشمند و حمل‌ونقل.	- افزایش دسترسی به خدمات بهداشتی و آموزشی، - بهبود بازده کشاورزی و مدیریت منابع، - توسعه سیستم‌های حمل‌ونقل ایمن‌تر و کارآمدتر.	نقاط قوت: تمرکز بر استفاده گسترده و اجتماعی از هوش مصنوعی، تلاش برای رشد فراگیر و کاهش نابرابری‌ها، توسعه نیروی کار متخصص. نقاط ضعف: کمبود زیرساخت‌های داده و منابع مالی، محدودیت‌های آموزشی و تخصصی، چالش‌های حریم خصوصی و امنیت داده‌ها.
سنگاپور	- ایجاد چارچوبی برای حکمرانی هوش مصنوعی با تأکید بر مسائل اخلاقی، - تبدیل شدن به یک مرکز جهانی برای توسعه و استقرار راه‌حل‌های هوش مصنوعی	- چارچوب حکمرانی هوش مصنوعی (۲۰۱۹): ارائه راهنمایی به بخش خصوصی برای رسیدگی به مسائل اخلاقی، - پروژه‌های ملی هوش مصنوعی برای بهینه‌سازی حمل‌ونقل، خدمات شهری، پیش‌بینی بیماری‌ها، آموزش شخصی‌سازی شده و تقویت امنیت مرزی.	- بهینه‌سازی جابه‌جایی بار و کاهش ترافیک، - بهبود مدیریت بیماری‌ها و تجربیات یادگیری، - تقویت امنیت مرزی و افزایش بهره‌وری در خدمات شهری.	نقاط قوت: رویکرد پیشرفته و جامع به حکمرانی هوش مصنوعی، تمرکز بر آموزش و آگاهی‌بخشی، زیرساخت‌های قوی دیجیتال. نقاط ضعف: محدودیت در مقیاس و منابع، نیاز به همکاری بین‌المللی برای گسترش و توسعه بیشتر، وابستگی به فناوری‌های خارجی برای برخی از بخش‌های کلیدی.

۶. نگاهی آماری به مقوله حکمرانی هوش مصنوعی در جهان



در این بخش، پس از بررسی تعدادی از مهم‌ترین بازیگران و فعالان حوزه حکمرانی هوش مصنوعی، در راستای تکمیل موردکاوی حکمرانی هوش مصنوعی در سطح جهانی، آمارهای بین‌المللی این حوزه به تفکیک‌های ذیل به نمایش درآمده است. شاخص هوش مصنوعی^۱ دانشگاه استنفورد، تجزیه و تحلیلی از قوانین تصویب شده توسط نهادهای قانونگذاری در ۱۲۷ کشور انجام داد که حاوی کلمه «هوش مصنوعی» از سال ۲۰۱۶ تا ۲۰۲۲ است. از ۱۲۷ کشور مورد تجزیه و تحلیل، از سال ۲۰۱۶، ۳۱ کشور حداقل یک لایحه^۲ در حوزه هوش مصنوعی به تصویب رسانده‌اند و مجموعاً ۱۲۳ لایحه مربوط به هوش مصنوعی را تصویب کرده‌اند. شکل زیر نشان می‌دهد که از سال ۲۰۱۶ تا ۲۰۲۲، تعداد کل لوایح مربوط به هوش مصنوعی افزایش شدیدی داشته و به ۳۷ لایحه تصویب شده در سال ۲۰۲۲ رسیده است.

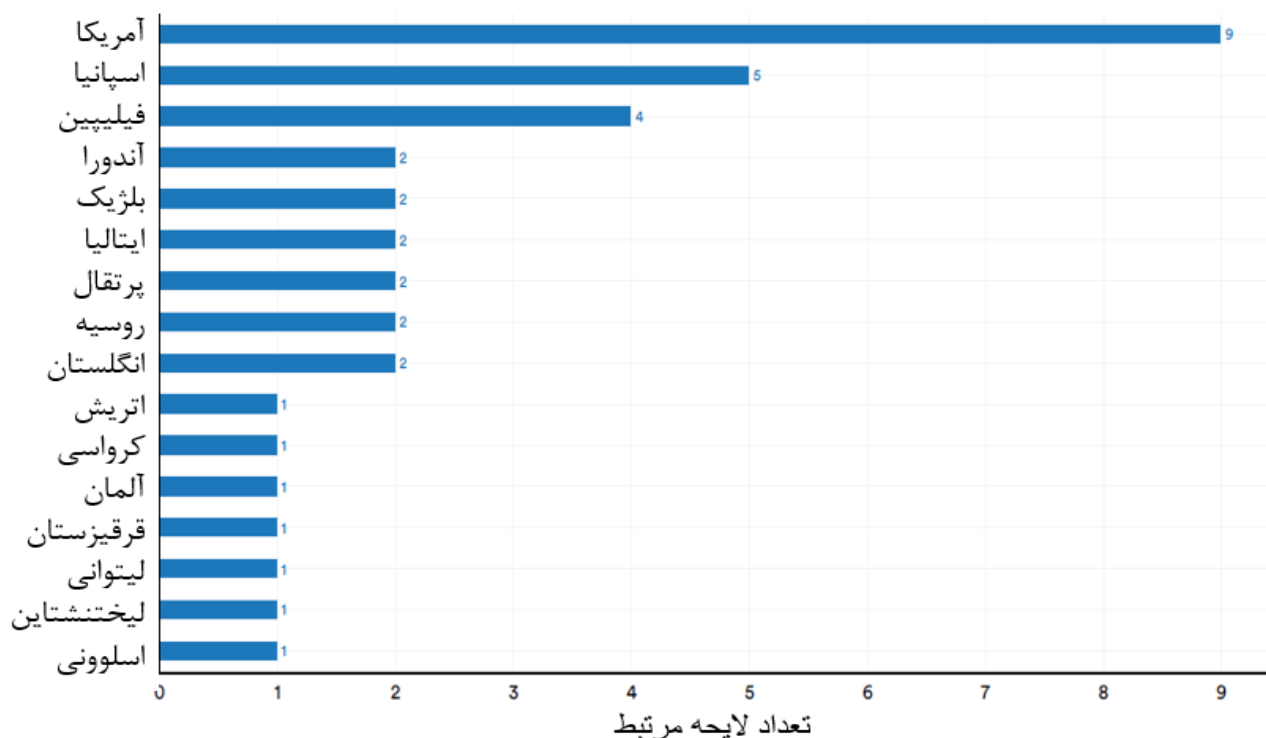
شکل ۱. تعداد لوایح مربوط به هوش مصنوعی به تفکیک کشورها، ۲۰۱۶-۲۰۲۲ [۳۷]



شکل زیر نشان‌دهنده تعداد قوانینی هستند که در سال ۲۰۲۲ در مورد هوش مصنوعی به تصویب رسیدند. آمریکا با ۹ قانون در صدر فهرست قرار گرفت و پس از آن اسپانیا و فیلیپین که به ترتیب ۵ و ۴ قانون تصویب کردند.

1. AI Index
2. Bill

شکل ۲. نمودار تعداد لایحه مربوط به هوش مصنوعی که در کشورهای منتخب به تصویب رسیده است ۲۰۱۶-۲۰۲۲ [۳۷]



جدول زیر مؤلفه‌هایی از حکمرانی هوش مصنوعی را در کشورهای مختلف بررسی کرده است.

جدول ۴. حکمرانی هوش مصنوعی در چند کشور منتخب در سال ۲۰۲۲ [۳۷]

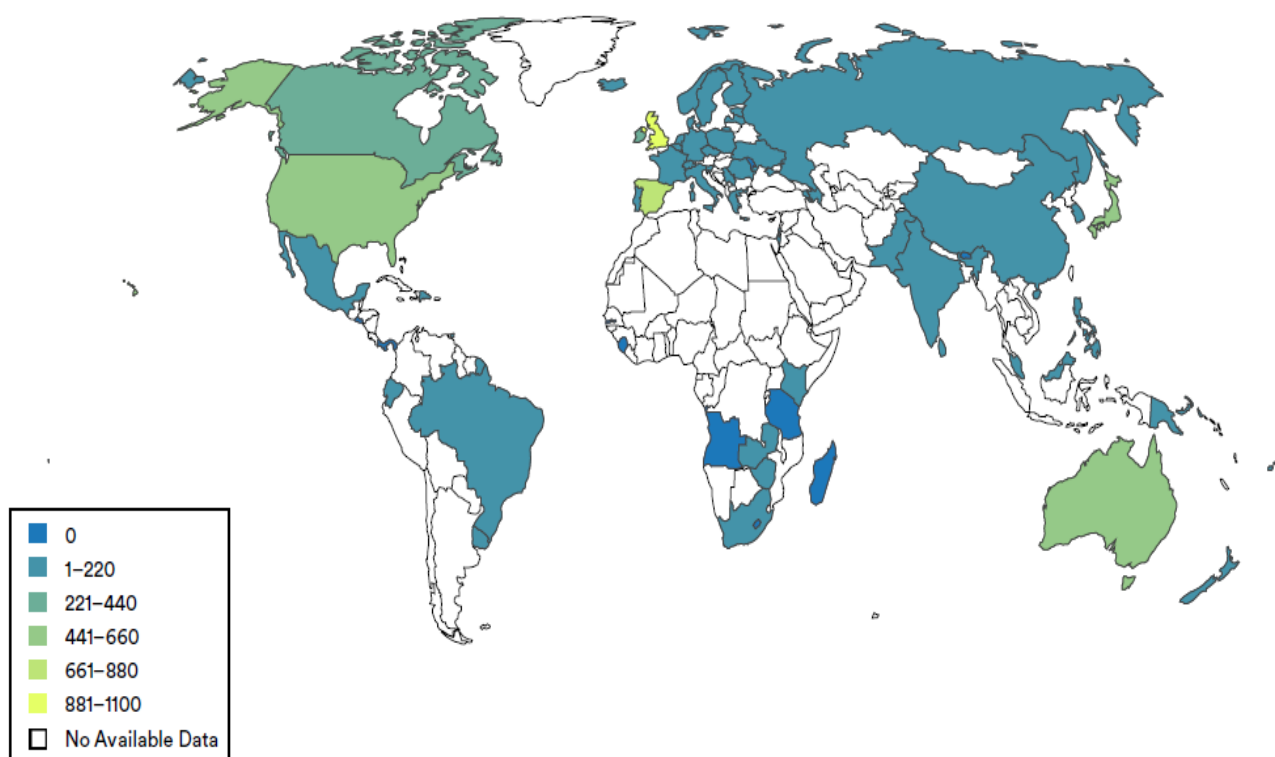
کشور	نام لایحه	تشریح
جمهوری قرقیزستان	درباره پارک منابع خلاق	این سند، وضعیت قانونی، مدیریت و رویه‌های عملیاتی پارک منابع خلاق را تعیین می‌کند که برای تسریع توسعه منابع خلاق از جمله هوش مصنوعی ایجاد شده است.
لتونی	اصلاحات قانون امنیت ملی	مفاد این قانون، محدودیت‌هایی را برای شرکت‌های تجاری، انجمن‌ها و بنیادهای مهم برای امنیت ملی، از جمله یک شرکت تجاری که هوش مصنوعی را توسعه می‌دهد، ایجاد می‌کند.
فیلیپین	قانون کمیسیون دوم کنگره در مورد آموزش	کنگره طبق مفاد این قانون، یک کمیسیون خاص برای بررسی، ارزیابی وضعیت آموزش فیلیپین ایجاد می‌کند. این قانون خواستار اصلاحاتی برای مقابله با چالش‌های جدید آموزش و پرورش ناشی از انقلاب صنعتی چهارم است که تا حدی با توسعه سریع هوش مصنوعی مشخص می‌شود. چالش‌های استفاده از هوش مصنوعی در آموزش در فیلیپین شامل موارد زیر است: - نیاز به یک زیرساخت فناوری قوی و منابع محاسباتی کافی همسو با چارچوب‌های خط‌مشی کلان ملی، - رسیدگی به نگرانی‌های مربوط به حریم خصوصی داده‌ها، - شکاف دیجیتال، - نیاز به آموزش و توسعه مستمر اساتید.



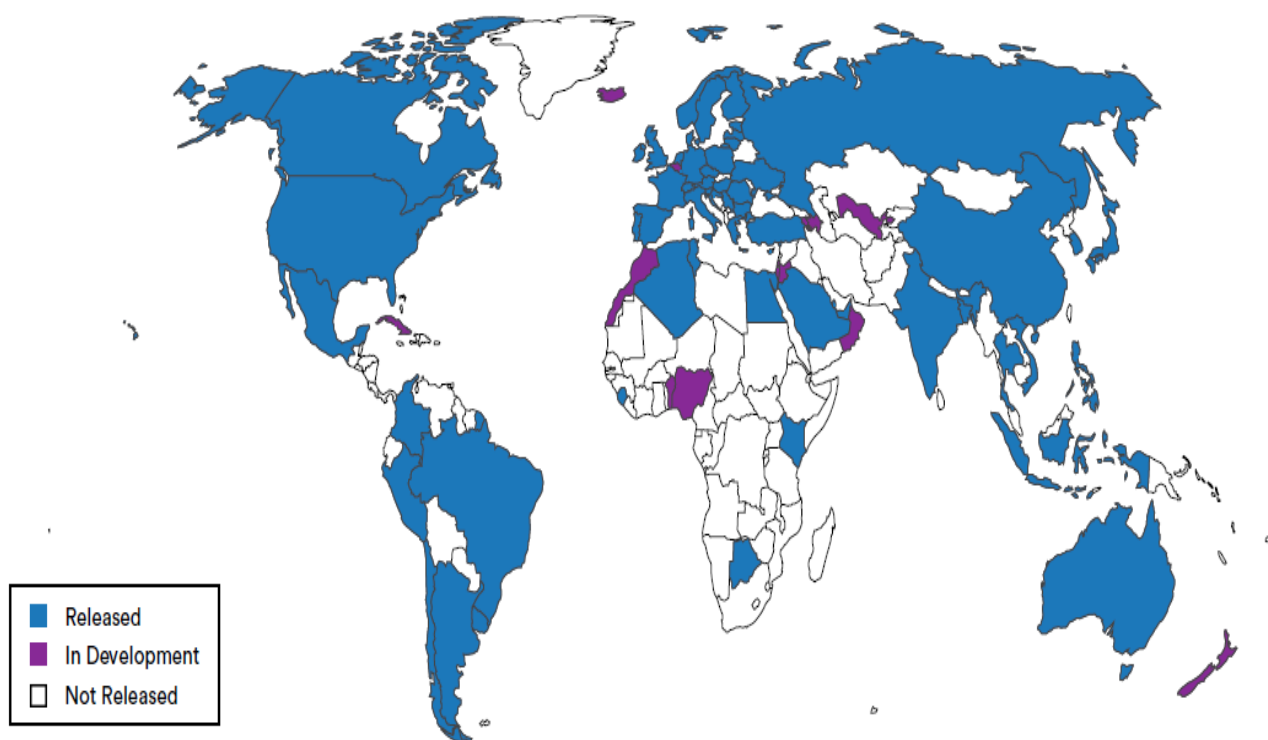
کشور	نام لایحه	تشریح
اسپانیا	حق رفتار برابر و عدم تبعیض	یک مفاد این قانون مقرر می‌دارد که الگوریتم‌های هوش مصنوعی درگیر در تصمیم‌گیری‌های ادارات دولتی، هر زمان که از نظر فنی امکان‌پذیر باشد، معیارهای به حداقل رساندن سوگیری، شفافیت و پاسخ‌گویی را در نظر می‌گیرند.
آمریکا	قانون آموزش هوش مصنوعی	این لایحه، اداره مدیریت و بودجه را ملزم می‌کند تا یک برنامه آموزشی مختص هوش مصنوعی برای جذب نیروی کار سازمان‌های اجرایی (برای مثال، کسانی که مسئول مدیریت برنامه یا تدارکات هستند) را طراحی کند. هدف برنامه این است که اطمینان حاصل شود که نیروی کار، از توانایی‌ها و خطرات مرتبط با هوش مصنوعی آگاهی دارد.

شکل زیر تعداد کل موارد ذکر شده «هوش مصنوعی» در رویه‌های قانونی کشورها در هفت سال گذشته را نشان می‌دهد. از ۸۱ کشور در نظر گرفته شده، ۶۲ کشور حداقل یک بار از لفظ «هوش مصنوعی» استفاده کرده‌اند و بریتانیا با ۱۰۹۲ بار استفاده، بر این لیست غالب است و پس از آن اسپانیا (۸۳۲)، آمریکا (۶۲۶)، ژاپن (۵۱۱) و هنگ کنگ (۴۷۸) قرار دارند.

شکل ۳. تعداد ذکر کلمات «هوش مصنوعی» در رویه‌های قانونی به تفکیک کشور، ۲۰۲۲-۲۰۱۶ (همان)



شکل ۴. کشورهای دارای سندهای راهبردی حکمرانی هوش مصنوعی (همان)



۷. الزامات طراحی مدل حکمرانی هوش مصنوعی



در این بخش، پس از مطالعه مفهوم حکمرانی هوش مصنوعی و موردکاوی آن، با توجه به اهمیت و لزوم وجود رویکرد بومی و منطقه‌ای به حکمرانی هوش مصنوعی، به بررسی الزامات طراحی مدل حکمرانی هوش مصنوعی بررسی شده است. هنگام بررسی مدل‌های حکمرانی هوش مصنوعی با رویکردی آینده‌نگر که به موضوعاتی مانند عدالت، امنیت، حریم خصوصی، جابه‌جایی نیروی کار و مالیات می‌پردازند، لازم است فراتر از این موضوعات رایج رفته و به چالش‌های ساختاری و الزامات مهم‌تری که در برخورد با فناوری‌های هوش مصنوعی وجود دارند، توجه کرد. در ادامه، سه چالش اساسی بررسی می‌شود که باید در طراحی مدل‌های حکمرانی هوش مصنوعی مدنظر قرار گیرند [۱۱]. این سه چالش کلیدی در طراحی این مدل‌های حکمرانی بررسی می‌شود: عدم تقارن دانشی، یافتن اجماع هنجاری و عدم تطبیق‌پذیری دولت.

■ **عدم تقارن دانشی:** در حالی که هوش مصنوعی این ظرفیت را دارد که زندگی میلیاردها نفر را تغییر دهد، تنها تعداد کمی از متخصصان، تکنیک‌های اساسی را درک می‌کنند. سیستم‌های مبتنی بر هوش مصنوعی اغلب غیرقابل تشخیص هستند و گاهی اوقات منجر به عدم تقارن دانشی گسترده بین توسعه‌دهندگان این سیستم‌ها و سایر ذی‌نفعان، از جمله مصرف‌کنندگان و خط‌مشی‌گذاران می‌شود. یک سیستم حکمرانی مؤثر برای هوش مصنوعی باید مکانیسم‌هایی را با هدف بهبود درک جمعی ما از پدیده هوش مصنوعی در مظاهر مختلف و زمینه‌های کاربردی آن ارائه دهد.



■ **یافتن اجماع هنجاری:**^۱ بحث‌های سیاستی و حکمرانی فعلی عمدتاً بر روی خطرات و چالش‌های مرتبط با هوش مصنوعی متمرکز است، اما هوش مصنوعی همچنین مزایای بالقوه فوق‌العاده‌ای را برای جامعه ارائه می‌دهد. یک مدل حکمرانی باید فضاهایی را برای تجزیه و تحلیل هزینه-فایده و ایجاد اجماع هنجاری در بین ذی‌نفعان مختلف باز کند، به‌ویژه در مواردی که مبادلات در طراحی سیستم‌های هوش مصنوعی دخیل هستند. یک مدل حکمرانی مطلوب همچنین باید با تفاوت‌های هنجاری بین زمینه‌ها و جغرافیاها سروکار داشته باشد و قابلیت همکاری بین چارچوب‌ها و رویکردهای مختلف را فراهم کند [۱۳].

■ **عدم تطابق دولت:**^۲ حتی در مواردی که درک مشترکی از فناوری‌های هوش مصنوعی، تکنیک‌های اساسی و اجماع اجتماعی در مورد آنچه مطلوب است یا نیست، وجود دارد، طراحی ابزار مؤثر، کارآمد و مشروع (استراتژی‌ها، رویکردها، ابزارها و غیره) برای حل و فصل مسائل اساسی ذکر شده با توجه به شرایط عدم قطعیت و پیچیدگی در اکوسیستم هوش مصنوعی، چالش‌برانگیز است [۱۴]. در میان تمام بازیگران، توجه به این نکته حائز اهمیت است که هر مدل نوظهوری باید در چارچوب‌های نهادی موجود قوانین و سیاست‌های قابل اجرا، به‌ویژه قانون اساسی، سیاست‌های کلی و یا حتی حقوق اساسی و ارزش‌های سازگار با قانون اساسی قرار گیرد و با آن تعامل داشته باشد. زیرا توسعه و استقرار حکمرانی هوش مصنوعی در خلأ صورت نمی‌گیرد [۱۵]. پس از بررسی چالش‌ها، به سراغ مطالعه یکی از مدل‌های قابل توجه حکمرانی هوش مصنوعی، یعنی «مدل لایه‌ای»^۳ می‌رویم که توسط دانشمندان دانشگاه هاروارد در سال ۲۰۱۷ مطرح شده است.

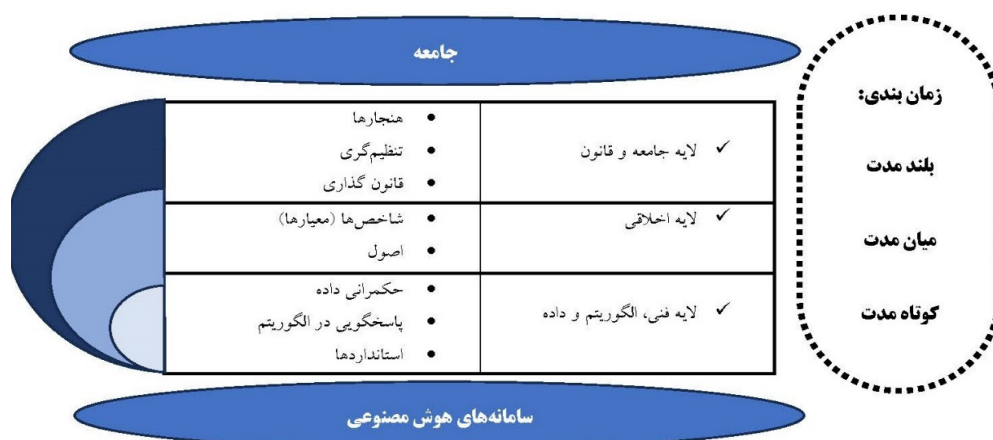
■ **۷-۱. مدل لایه‌ای حکمرانی هوش مصنوعی**^۴ مقیاس، ناهمگونی، پیچیدگی و میزان استقلال تکنولوژیک سیستم‌های هوش مصنوعی نیازمند تفکر جدیدی در مورد سیاست، قانون و مقررات است. گسر و آلمیدا، در مقاله‌ای مدل لایه‌ای حکمرانی هوش مصنوعی را ارائه و تلاش کرده‌اند ماهیت پیچیده حکمرانی هوش مصنوعی را با استفاده از یک مدل تحلیلی با سه لایه به تصویر بکشند. لایه‌های تعاملی از بالا به پایین، به شرح زیر است:

✓ اجتماعی و حقوقی،

✓ اخلاقی،

✓ مبانی فنی که از لایه‌های اخلاقی و اجتماعی حمایت می‌کنند.

شکل ۵. مدل لایه‌ای برای حکمرانی هوش مصنوعی [۴۲]



1. Finding Normative Consensus
2. Government Mismatches
3. A Layered Model for AI Governance
4. <https://dash.harvard.edu/handle/1/34390353>

شکل ۵، نمایشی از مدل حکمرانی لایه‌ای را نشان می‌دهد. ابزارهای ترسیم شده روی لایه‌ها، به تناسب هر بازه زمانی قابل توسعه هستند. در کوتاه‌مدت، پیشنهادها و حاکمیتی می‌تواند بر توسعه استانداردها و اصول الگوریتم‌های هوش مصنوعی متمرکز شود. در میان مدت و بلندمدت، دولت‌ها می‌توانند روی قوانین خاصی برای تنظیم تعامل ذی‌نفعان حوزه هوش مصنوعی کار کنند. این مدل می‌تواند یک اکتشاف مفید باشد که نشان می‌دهد چگونه اصول، خط‌مشی‌ها، هنجارها و قوانین در پاسخ به چالش‌ها و فرصت‌های مبتنی بر هوش مصنوعی می‌توانند ترکیب شوند و با هم در لایه‌های مختلف کار کنند.

■ لایه فنی، الگوریتم و داده

لایه فنی، اساس اکوسیستم حکمرانی هوش مصنوعی را تشکیل می‌دهد. سیستم‌های مبتنی بر هوش مصنوعی و سیستم‌های خودکار چه فیزیکی باشند مانند خودروهای خودران و ربات‌های صنعتی و چه نرم‌افزاری مانند سیستم‌های عدالت کیفری، تشخیص پزشکی یا دستیاران هوشمند، به داده‌ها و الگوریتم‌ها وابسته هستند [۱۶]. برای الگوریتم‌هایی که تأثیرات اجتماعی دارند، اصولی مانند مسئولیت‌پذیری، قابلیت توضیح، دقت، قابلیت حسابرسی و انصاف پیشنهاد شده‌اند [۱۷]. همچنین فرایند جمع‌آوری، استفاده و مدیریت داده‌ها توسط این الگوریتم‌ها که به عنوان حکمرانی داده شناخته می‌شود، باید بر اساس اصولی باشد که عدالت را تقویت کند و از تبعیض بر اساس نژاد، رنگ، ملیت، مذهب، جنسیت، ناتوانی یا وضعیت خانوادگی جلوگیری کند.

■ لایه اخلاقی

در بالای لایه فنی، می‌توان به نگرانی‌های اخلاقی پرداخت که برای انواع برنامه‌ها و سیستم‌های هوش مصنوعی اعمال می‌شود. یکی از منابع مهم برای توسعه چنین اصول اخلاقی، اصول مشترکی است که در قوانین بین‌المللی و ملی به آنها تصریح شده است. نمونه دیگری از ظهور هنجارهای اخلاق هوش مصنوعی، اصول کلی انجمن مهندسان برق و الکترونیک (IEEE) برای هوش مصنوعی و سیستم‌های مستقل است [۱۶].

اقدامات الگوریتم‌ها را می‌توان بر اساس معیارها و اصول اخلاقی مورد ارزیابی قرار داد. برای مثال اگر یک برنامه هوش مصنوعی داده‌های یک شرکت بیمه را تحلیل کرده و از گروه خاصی از افراد، بر اساس عواملی مانند جنسیت یا سن، حق بیمه بیشتری را به تبعیض دریافت کند، چنین تصمیم‌گیری برخلاف اصل اخلاقی رفتار برابر و منصفانه است.

■ لایه جامعه و قانون (اجتماعی و حقوقی)

لایه اجتماعی و قانونی می‌تواند به فرایند ایجاد نهادها و تخصیص مسئولیت‌ها برای تنظیم سیستم‌های هوش مصنوعی و خودمختار بپردازد. برای مثال، متیو شرر^۱ [۱۷] یک نهاد سیاستگذاری را توصیف می‌کند که قدرت تعریف هوش مصنوعی را دارد، استثناهایی ایجاد می‌کند که اجازه می‌دهد تحقیقات هوش مصنوعی در محیط‌های خاصی بدون مسئولیت دقیق محققان انجام شود و فرایند صدور گواهی هوش مصنوعی را ایجاد کند. یک نقطه شروع برای هنجارهای خاص با هدف تنظیم هوش مصنوعی، می‌تواند اصول و معیارهایی باشد که از لایه‌های اخلاقی و فنی، علاوه بر چارچوب‌های قانونی ملی و بین‌المللی از قبل موجود و کلی‌تر، بیرون می‌آیند. پیاده‌سازی ساختارهای حاکمیتی برای هوش مصنوعی و سیستم‌های تصمیم‌گیری الگوریتمی می‌تواند در چندین لایه رخ دهد و شامل رویکردهای ترکیبی باشد. فرایندهای حکمرانی می‌تواند از راه‌حل‌های بازارمحور تا ساختارهای مبتنی بر دولت متغیر باشد و می‌تواند در سطح ملی یا بین‌المللی اعمال شود. در سطح منطقه‌ای، یک نمونه مناسب، مقررات عمومی حفاظت از داده‌ها (GDPR) است؛ یک مقررات گسترده و پیچیده با هدف تقویت و یکسان‌سازی حفاظت از داده‌ها برای همه افراد در اتحادیه اروپا. این قانون یک «حق توضیح» ارائه می‌کند که شرکت‌ها را موظف می‌کند که هدف یک الگوریتم و نوع داده‌هایی را و هنگام تصمیم‌گیری خودکار استفاده می‌کند، توضیح دهند [۱۸].

در زمینه مسائل حقوقی یکی از چالش‌های عمده کشور، عدم وجود کلیدواژه‌های تخصصی فارسی ناظر به مفاهیم فنی - حقوقی است. این مسئله هرگونه الگوبرداری از قوانین سایر کشورها و قانونگذاری تخصصی در این حوزه را با مشکل و ابهام مواجه خواهد کرد. برای

1. Matthew Scherer



مثال بسیاری از واژگان تخصصی قانون اقدام هوش مصنوعی اتحادیه اروپا مبتنی بر اصطلاحاتی است که پیش از آن در قوانین قبلی آنها از جمله مقررات عمومی حفاظت از داده‌ها تولید شده است و متأسفانه هیچ قانونی در ایران این گونه مفاهیم را پوشش نداده است. این مسئله می‌تواند حتی پس از قانونگذاری، به چالشی برای نهادهای داوری و قضایی تبدیل شود و اعمال عدالت را تضعیف کند.

۸. جمع‌بندی و توصیه‌های سیاستی

هوش مصنوعی به‌عنوان یکی از پیشرفته‌ترین و تحول‌سازترین فناوری‌های عصر حاضر، در حال ایجاد تغییرات بنیادین در بسیاری از جنبه‌های زندگی بشر است. این فناوری نه تنها می‌تواند بهره‌وری و کارایی را در حوزه‌های مختلف از جمله صنعت، بهداشت، حمل‌ونقل و آموزش بهبود بخشد، بلکه فرصت‌های جدیدی را برای نوآوری، رشد اقتصادی و توسعه اجتماعی فراهم می‌آورد. با این حال، گسترش هوش مصنوعی چالش‌های پیچیده‌ای را در زمینه‌های اخلاقی، قانونی، فرهنگی و اجتماعی به همراه دارد که نیازمند تدوین و اجرای چارچوب‌های حکمرانی مناسب برای مدیریت و نظارت بر این فناوری است. در این راستا، کشورها باید به‌طور فعال به طراحی، تدوین و پیاده‌سازی این چارچوب‌ها بپردازند تا بتوانند از مزایای هوش مصنوعی بهره‌برداری کرده و هم‌زمان با مدیریت مخاطرات آن، از اثرات منفی احتمالی آن جلوگیری کنند.

بررسی وضعیت کنونی حکمرانی هوش مصنوعی در جهان نشان می‌دهد که کشورهای مختلف رویکردهای متنوعی را برای پرداخت به این فناوری اتخاذ کرده‌اند. برای مثال، اتحادیه اروپا با تصویب قانون جامع هوش مصنوعی (AI Act)، تلاش دارد تا یک چارچوب قانونی منسجم و قابل اجرا برای توسعه و استفاده مسئولانه از سیستم‌های هوش مصنوعی ایجاد کند. این قانون بر اهمیت تضمین ایمنی، شفافیت و احترام به حقوق اساسی افراد تأکید دارد و هدف آن فراهم‌سازی شرایطی برای توسعه پایدار و مسئولانه هوش مصنوعی است.

در آمریکا، رویکرد حکمرانی بیشتر بر تقویت تحقیق و توسعه و ایجاد نیروی کار ماهر متمرکز شده است. سند استراتژیک تحقیق و توسعه هوش مصنوعی ملی آمریکا، بر نیاز به تحقیقات بلندمدت و ایجاد اصول و استانداردهایی برای تضمین امنیت سیستم‌های هوش مصنوعی تأکید دارد. این سند، همچنین بر اهمیت آمادگی برای بهره‌برداری از فرصت‌های اقتصادی و مقابله با چالش‌های اخلاقی و اجتماعی ناشی از هوش مصنوعی تأکید می‌کند.

در آسیا، کشورهایی مانند چین و هند نیز در حال تدوین و اجرای برنامه‌های ملی هوش مصنوعی با تمرکز بر توسعه اقتصادی، اخلاق، حریم خصوصی و تدوین چارچوب‌هایی برای تنظیم‌گری هوش مصنوعی هستند. چین با وضع مقررات ملی الزام‌آور برای هوش مصنوعی، تلاش دارد تا نه تنها کنترل دقیق‌تری بر توسعه داخلی این فناوری داشته باشد، بلکه به‌طور فعال در تدوین استانداردهای جهانی برای حکمرانی هوش مصنوعی مشارکت کند. هند نیز با تأکید بر رویکرد «هوش مصنوعی برای همه»، به دنبال بهره‌گیری از پتانسیل هوش مصنوعی برای رشد فراگیر و توسعه پایدار آن در سطح ملی است.

این نمونه‌ها نشان می‌دهند که حکمرانی هوش مصنوعی یک مسئله جهانی و چندجانبه بوده که نیازمند همکاری و هماهنگی بین‌المللی برای تبادل دانش، تدوین استانداردهای مشترک و تقویت زیرساخت‌های نظارتی است. کشورها باید از تجربیات یکدیگر در زمینه حکمرانی هوش مصنوعی بهره‌برداری کرده تا بتوانند با چالش‌های مشابه به‌طور مؤثر مقابله کنند.

تدوین مدل‌های حکمرانی مؤثر برای هوش مصنوعی نیازمند توجه به چالش‌های ساختاری و الزامات اساسی در این حوزه است. یکی از چالش‌های اصلی در این زمینه، **عدم تقارن دانشی** میان توسعه‌دهندگان فناوری و سایر ذی‌نفعان از جمله خط‌مشی‌گذاران و مصرف‌کنندگان است. این عدم تقارن می‌تواند منجر به سوء تفاهم‌ها، کاهش اعتماد عمومی و مقاومت در برابر پذیرش فناوری‌های هوش مصنوعی شود. برای حل این مسئله، مدل‌های حکمرانی باید مکانیسم‌هایی را برای بهبود درک عمومی و افزایش آگاهی در مورد هوش مصنوعی و کاربردهای آن ارائه دهند.

چالش دیگر، **یافتن اجماع هنجاری** میان ذی‌نفعان مختلف است. مدل‌های حکمرانی باید فضای لازم را برای تجزیه و تحلیل هزینه-فایده و ایجاد اجماع هنجاری میان ذی‌نفعان فراهم کنند. این مدل‌ها باید بتوانند با تفاوت‌های فرهنگی، اجتماعی و جغرافیایی

سازگاری داشته باشند و قابلیت همکاری بین چارچوب‌ها و رویکردهای مختلف را فراهم کنند.

علاوه بر این، **عدم تطابق دولت‌ها** با تغییرات سریع و پیچیدگی‌های فناوری‌های هوش مصنوعی یکی دیگر از موانع مهم است. دولت‌ها باید به سرعت بتوانند با تغییرات و نوآوری‌های جدید سازگار شوند و از چارچوب‌های تنظیم‌گری مناسب برای مدیریت این فناوری استفاده کنند. این امر نیازمند ایجاد زیرساخت‌های قانونی و نهادی است که بتواند با تحولات سریع در حوزه هوش مصنوعی همگام شود و پاسخ‌گوی نیازهای روزافزون جامعه باشد.

۸-۱. توصیه‌ها و پیشنهادهای سیاستی

با توجه به تحلیل‌های انجام‌شده و بررسی وضعیت کنونی حکمرانی هوش مصنوعی در جهان، چندین توصیه سیاستی برای بهبود حکمرانی هوش مصنوعی در کشور قابل ارائه است:

👉 **حکمرانی داده:** حکمرانی هوش مصنوعی، پیش از هر چیز، مستلزم طراحی نظام حکمرانی داده است. چراکه داده، به مثابه مواد خام در توسعه هوش مصنوعی است. در بند «۱» ماده (۵) سند ملی هوش مصنوعی (راهندها و اقدامات ملی)، به «لزوم تدوین نظام حکمرانی داده، به منظور تسهیل در بهره‌برداری انتفاعی و غیرانتفاعی از داده‌های تولیدی کشور از طریق اصلاح قوانین و مقررات موجود و به رسمیت شناختن مالکیت و حق بهره‌برداری انتفاعی هر شخص یا نهاد بر داده‌های تولیدی خود با حفظ ملاحظات امنیت ملی، رعایت حریم خصوصی و اخلاق عمومی» اشاره شده است. مطالعه عمیق مبانی حقوقی حکمرانی داده، در عین توجه به مبانی فنی و زیرساختی، از اهم نیازهای خط‌مشی‌پژوهان و خط‌مشی‌گذاران این حوزه است. چراکه این مباحث بنیادین، تکیه‌گاه طراحی نظام حکمرانی هوش مصنوعی در کشور است.

👉 **طراحی نظام نوآورانه تسهیلگر-تنظیم‌گر در قالب قوانین و مقررات:** با توجه به ابلاغ سند ملی هوش مصنوعی و هدف‌گذاری‌های کلان آن، نیل به اهداف سند مستلزم طراحی یک نظام نوآوری فناورانه پویا و چابک در این حوزه خاص است تا ضمن برآورده کردن انتظارات واحدهای توسعه‌دهنده این فناوری، ضامن حفظ امنیت ملی، حفاظت از حقوق شهروندان و تکامل فناوری باشد. این نظام باید به صورت هم‌زمان تسهیلگری و تنظیم‌گری را تضمین کرده و اطمینان دهد که به وسیله قوانین و مقررات، مانع نوآوری نمی‌شود، بلکه صرفاً ریسک ناشی از چالش‌های حیاتی مانند حفظ حریم خصوصی داده‌ها، پاسخ‌گویی الگوریتمی و شفافیت را مدیریت می‌کند. از این رو از مهم‌ترین گام‌های حکمرانی هوش مصنوعی در کشور، طراحی و تدوین یک نظام تسهیلگر-تنظیم‌گر انعطاف‌پذیر و پویاست که کاربست هوش مصنوعی را تسهیل کرده و شتاب دهد. این نظام باید شامل قواعد و چارچوب‌هایی برای سرمایه‌گذاری استراتژیک، ایجاد صندوق‌های سرمایه‌گذاری خطرپذیر تخصصی، مشوق‌های مالیاتی، حفظ ایمنی و حریم خصوصی، تضمین عدالت و تعادل اجتماعی، توسعه مداوم زیرساخت‌ها، ایجاد بازارهای رقابتی داخلی و حضور در بازارهای بین‌المللی باشد. علاوه بر این، نظارت فعال بر اجرای این قواعد و ارائه تنظیمات به‌روز و سازگار با پیشرفت‌های فناوری نیز حیاتی است. همچنین مطابق با بند «۱-۵» از ماده (۵) سند ملی هوش مصنوعی، «ایجاد محیط‌های حقیقی و مجازی آزمون تنظیم‌گری به منظور فراهم‌آوری زیرساخت بهره‌برداری آزمایشی محصولات و خدمات حوزه هوش مصنوعی در مقیاس محدود و قابل کنترل» (سندباکس تنظیم‌گری هوش مصنوعی)، از راهنمدها و اقدامات ملی مورد انتظار است.

👉 **ایجاد سازوکارهای نظارت و ارزیابی:** با توجه به بند «۶» ماده (۴) سند ملی هوش مصنوعی (سیاست‌های راهبردی)، که به ضرورت نظارت و مراقبت بر تصمیمات و خودمختاری هوش مصنوعی و جلوگیری از رخنه‌انگاری و اعتبارسنجی اجتماعی بدون رعایت حقوق افراد و قوانین کشور تأکید دارد و در راستای اطمینان از رعایت قوانین و مقررات مربوط به حکمرانی هوش مصنوعی، ایجاد سازوکارهایی برای نظارت و ارزیابی منظم، با توجه به اینکه مقوله «نظارت» از کلان‌کارکردهای مجلس شورای اسلامی است، ضروری است. این سازوکارها باید طبق بند «۱-۳» «راهنمدها و اقدامات ملی» سند ملی هوش مصنوعی، شامل «تدوین استانداردهای مربوط به مالکیت فکری تولیدات مبتنی بر هوش مصنوعی و به اشتراک‌گذاری داده و همچنین ایمنی و امنیت سیستم‌ها و خدمات هوش مصنوعی» باشد. به منظور پشتیبانی علمی از نظارت مجلس در این حوزه و تضمین عدالت، آزادی‌ها و حقوق اساسی مردم، انتظار



می‌رود مرکز پژوهش‌های مجلس به یکی از کانون‌های مرجع مطالعات و تجمیع ایده‌های کلان حکمرانی هوش مصنوعی تبدیل شود. **توسعه زیرساخت‌های فنی، فرهنگی و ترویجی:** بند «۲» ماده (۵) سند ملی هوش مصنوعی با عنوان «آموزش و پژوهش»، به اهمیت توجه زیرساختی و ساختاری در توسعه منابع انسانی در این حوزه پرداخته است. در این راستا به منظور ارتقای نیروی انسانی مؤثر و توسعه هوش مصنوعی در کشور، لازم است که زیرساخت‌های فنی، فرهنگی و ترویجی مناسب از جمله با بسترسازی تأسیس و گسترش اندیشکده‌ها، کانون‌های فکری و حلقه‌های میانی در حوزه هوش مصنوعی ایجاد شود. این موضوع شامل آموزش متخصصان در زمینه هوش مصنوعی، توسعه تکنولوژی محلی و ایجاد بسترهای مشارکتی بین حاکمیت، صنعت و دانشگاه‌ها برای تحقق اهداف حکمرانی هوش مصنوعی است که هماهنگی نهادی میان بخش‌های مختلف آموزشی، فرهنگی و تحقیقاتی را می‌طلبد. همچنین ایجاد شبکه‌ای منسجم از کنشگران رسانه‌ای فعال در این عرصه می‌تواند به ترویج مناسب فرهنگ استفاده از هوش مصنوعی و پذیرش اجتماعی آن در یک بازه زمانی مناسب کمک کرده و از ایجاد نگرانی‌های ضدامنیتی برای پشتیبانی از توسعه این فناوری پیشگیری کند.

ارتقای همکاری بین‌المللی: بند «۶» «راهنمادهای و اقدامات ملی» سند ملی هوش مصنوعی، به «تعاملات بین‌المللی» اشاره دارد. حکمرانی هوش مصنوعی یک مسئله جهانی بوده و نیازمند همکاری و تعامل بین کشورهاست. با توجه به گزاره‌های بند «۶»، برنامه‌های تبادل دانش و تجربه، تدوین پیمان‌ها و توافق‌نامه‌های بین‌المللی برای حفظ حقوق و حریم خصوصی به‌خصوص با کشورهای همسایه و مشارکت فعال در تدوین استانداردهای جهانی برای حکمرانی هوش مصنوعی با هدف توسعه عادلانه فناوری‌های مبتنی بر هوش مصنوعی، تضمین دسترسی‌پذیری عادلانه، تضمین حقوق جوامع و... از جمله اقدامات مؤثر در این زمینه هستند. در این راستا پیشنهاد ایجاد همسویی میان کشورهای در حال توسعه در قالب‌هایی مانند بریکس، شانگهای و عدم‌تعهد با هدف توسعه پایدار جهانی و مخالفت با یکجانبه‌گرایی، قدمی بزرگ برای پیشگیری از ایجاد انحصار حقوقی برای فناوری هوش مصنوعی در کشورهای بلوک ناتو است که باید با اولویت دنبال شود. این همکاری‌های بین‌المللی خواهد توانست چارچوب حقوقی کشور را در این حوزه، پیش از تدوین چارچوب‌های حقوقی بین‌المللی و تحمیل آن بر نظام حقوقی کشور، ساختاردهی کند و در مرحله تدوین چارچوب‌های حقوقی بین‌المللی، همکاری با بازیگران جهانی با اهداف مشترک، مانع از تحمیل خواسته‌ها و اهداف کشورهای پیشرو در این حوزه و اعمال محدودیت‌های بازدارنده بر کشور شده یا حداقل پیشرفت‌های ضروری پیش از تحمیل این فرایند جهانی، در کشور به‌دست آمده باشد. زیرا رصدها نشان می‌دهد آمریکا و کشورهای اروپایی اولین گام‌ها برای تدوین یک کنوانسیون بین‌المللی را برداشته‌اند و کوتاهی در این زمینه می‌تواند منجر به بروز چالش‌های بلندمدت برای کشور شود.

توسعه مکانیسم‌های مشارکت عمومی: «تقویت مردم‌محوری و روحیه تعاون در توسعه و بهره‌برداری از هوش مصنوعی»، از موارد مورد توجه سند ملی هوش مصنوعی در ماده (۴) یعنی «سیاست‌های راهبردی» است. مشارکت مردم، نخبگان و ذی‌نفعان مختلف در فرایند تصمیم‌گیری و اجرای خط‌مشی‌های مربوط به حکمرانی هوش مصنوعی بسیار ضروری و ارزشمند است. بنابراین لازم است که مجلس و مرکز پژوهش‌ها، با توجه به ظرفیت‌های ارتباطی و اندیشکده‌های حکمرانی و قانونگذاری استانی خود، سازوکارهایی برای بهره‌گیری نظرات و نگرش‌های گوناگون جامعه و ایجاد فضایی برای ارتباط مستقیم بین حاکمیت و شهروندان و بسط حکمرانی مردمی در توسعه هوش مصنوعی فراهم کنند. خروجی این سازوکارها می‌تواند شامل مواردی مانند پیشنهادها، اجرایی، شناسایی چالش‌های عملیاتی در حوزه داده و الگوریتم‌ها، پیشنهاد راهکارهای تنظیم‌گرانه، پیشنهاد سازوکارهای تبدیلی و تعاملی میان ذی‌نفعان مختلف و مواردی از این دست توسط اشخاص و شرکت‌های بخش خصوصی باشد. این دسته از بازیگران به‌علت سروکار داشتن با مسائل اجرایی واقعی این حوزه و همچنین به‌روز و مجهز بودن به علم روز دنیا در حوزه فناوری‌های مرتبط، می‌توانند کمک شایانی کنند. این سازوکارهای مشارکتی همچنین خواهند توانست با ایجاد حس اهمیت‌دهی و شنیدن دغدغه‌ها توسط شهروندان از سمت حاکمیت، از نگرانی‌های ایشان و مقاومت‌های فردی و سازمانی در توسعه این فناوری در کشور، بکاهند.

تولید ادبیات حقوقی: در زمینه مسائل حقوقی، یکی از چالش‌های عمده کشور عدم وجود کلیدواژه‌های تخصصی فارسی ناظر به مفاهیم فنی - حقوقی در حوزه هوش مصنوعی است. از این رو پیشنهاد می‌شود برای تبیین مفهومی و ارائه این کلیدواژه‌های فارسی، کمیته‌ای میان‌رشته‌ای با مشارکت بخش‌های مختلف ذی‌نفع در مرکز پژوهش‌های مجلس در این زمینه تشکیل شود.

با توجه به سرعت پیشرفت و گسترش فناوری‌های هوش مصنوعی و تأثیرات گسترده آن بر جامعه، لزوم ایجاد یک نظام نوآوری فناورانه برای پرداختن به تسهیلگری و تنظیم‌گری مقوله هوش مصنوعی، به‌عنوان یکی از مؤلفه‌های اصلی حکمرانی هوش مصنوعی، بیش‌ازپیش احساس می‌شود. تسهیلگری و تنظیم‌گری مناسب می‌تواند به شتاب‌دهی، کنترل و نظارت بر توسعه و کاربردهای هوش مصنوعی کمک کرده و از بروز تخلفات و مخاطرات احتمالی جلوگیری کند. پیشنهاد می‌شود که یک چارچوب تسهیلگر-تنظیم‌گر جامع و چندلایه تدوین شود که شامل تمامی ذی‌نفعان و بازیگران مختلف باشد. این چارچوب باید به‌گونه‌ای طراحی شود که همگام با پیشرفت‌های فناوری، قابل تطبیق و به‌روزرسانی باشد و بتواند نیازهای جامعه را در تمامی سطوح برآورده کند. با این رویکرد، می‌توان به توسعه پایدار و مسئولانه هوش مصنوعی دست یافت و ایران را به‌عنوان یک کشور پیشرو در زمینه هوش مصنوعی در منطقه و جهان معرفی کرد. تسهیلگری و تنظیم‌گری جامع و مدبرانه نه‌تنها به حفاظت از حقوق شهروندان و جلوگیری از سوءاستفاده از فناوری کمک می‌کند، بلکه می‌تواند بستری مناسب برای نوآوری و رشد اقتصادی پایدار فراهم آورد.

- [1] A. J. Izenman, *Modern Multivariate Statistical Techniques*. New York, NY: Springer New York, 2008. doi: 10.1007/978-0-387-78189-1.
- [2] I. Linkov, B. Trump, K. Poinssatte-Jones, and M.-V. Florin, "Governance Strategies for a Sustainable Digital World," *Sustainability*, vol. 10, no. 2, p. 440, Feb. 2018, doi: 10.3390/su10020440.
- [3] C. O'neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown, 2017.
- [4] R. Brownsword and H. Somsen, "Law, innovation and technology: Before we fast forward-a forum for debate," *Law Innov Technol*, vol. 1, no. 1, pp. 1–73, 2009, doi: 10.1080/17579961.2009.11428364.
- [5] P. Stone et al., "Artificial Intelligence and Life in 2030: The One Hundred Year Study on Artificial Intelligence," Oct. 2022, Accessed: Feb. 23, 2024. <http://arxiv.org/abs/2211.06318>
- [۶] ج. کرمی، «هوش مصنوعی در جهان (۱) (فدراسیون روسیه)»، ماهنامه گزارش‌های کارشناسی مرکز پژوهش‌های مجلس شورای اسلامی، vol. ۲۶، no. ۷، p. ۲۰۱۸، https://report.mrc.ir/article_۷۴۱۷.html
- [۷] ح. نجفی رستاقی و م. عبدالحسین زاده، "چالش‌ها و راهکارهای تحقق حکمرانی هوشمند در کشور و ارائه توصیه‌های سیاستی برای مجلس شورای اسلامی"، (۱۸۹۵۴) ماهنامه گزارش‌های کارشناسی مرکز پژوهش‌های مجلس شورای اسلامی، vol. ۳۱، no. ۳، pp. ۱–۲۹، https://report.mrc.ir/article_۹۴۳۶.html
- [۸] ع. یوسفی، "تصمیم‌گیری مبتنی بر داده: رویکردی جهت تحول حکمرانی پارلمانی"، {۱۸۸۶۲} ماهنامه گزارش‌های کارشناسی مرکز پژوهش‌های مجلس شورای اسلامی، vol. ۳۱، no. ۱، pp. ۱–۳۱، ۲۰۲۳، https://report.mrc.ir/article_۹۳۵۹.html
- [۹] م. مهربان هلان، ع. یوسفی، ا. اکبری، و س. خردمندینا، "نگاشت نهادی و تقسیم کار ملی در حوزه توسعه هوش مصنوعی و حکمرانی داده محور"، (۱۹۴۳۴) ماهنامه گزارش‌های کارشناسی مرکز پژوهش‌های مجلس شورای اسلامی: vol. ۳۱، no. ۹، pp. ۱–۴۸، ۲۰۲۳، doi: https://report.mrc.ir/article_۹۸۴۷.html
- [۱۰] ا. اکبری، ع. یوسفی، و م. مهربان هلان، "بررسی لایحه برنامه هفتم توسعه (۸۸): توسعه پایدار هوش مصنوعی در کشور"، (۱۹۳۹۵) ماهنامه گزارش‌های کارشناسی مرکز پژوهش‌های مجلس شورای اسلامی، vol. 31، no. 8، pp. 1–38, 2023, https://report.mrc.ir/article_9802.html
- [11] U. Gasser and V. A. F. Almeida, "A Layered Model for AI Governance," *IEEE Internet Comput*, vol. 21, no. 6, pp. 58–62, Nov. 2017, doi: 10.1109/MIC.2017.4180835.
- [12] A. Daly et al., "Artificial intelligence governance and ethics: global perspectives," *arXiv preprint arXiv:1907.03848*, 2019.
- [13] R. Calo, "Artificial Intelligence Policy: A Roadmap," *SSRN Electronic Journal*, Sep. 2017, doi: 10.2139/SSRN.3015350.
- [14] I. Ulnicane, D. O. Eke, W. Knight, G. Ogoh, and B. C. Stahl, "Good governance as a response to discontents? Déjà vu, or lessons for AI from other emerging technologies," *Taylor & Francis Interdisciplinary Science Reviews*, 2021•Taylor & Francis, vol. 46, no. 1–2, pp. 71–93, 2021, doi: 10.1080/03080188.2020.1840220.
- [15] W. Wu, T. Huang, and K. Gong, "Ethical Principles and Governance Technology Development of AI in China," *Engineering*, vol. 6, no. 3, pp. 302–309, Mar. 2020, doi: 10.1016/J.ENG.2019.12.015.
- [16] H. Al Zadjali, "Building the right AI governance model in Oman," *ACM International Conference Proceeding Series*, pp. 116–119, Sep. 2020, doi: 10.1145/3428502.3428516.
- [17] N. M. Aliman and L. Kester, "Extending socio-Technological reality for ethics in artificial intelligent

- systems,” Proceedings - 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2019, pp. 275–282, Dec. 2019, doi: 10.1109/AIVR46125.2019.00064.
- [18] J. Butcher, I. B.-T. R. Journal, and undefined 2019, “What is the state of artificial intelligence governance globally?,” Taylor & FrancisJ Butcher, I BeridzeThe RUSI Journal, 2019•Taylor & Francis, vol. 164, no. 5–6, pp. 88–96, Sep. 2019, doi: 10.1080/03071847.2019.1694260.
- [19] C. C.-P. T. of the R. Society and undefined 2018, “Governing artificial intelligence: ethical, legal and technical opportunities and challenges,” royalsocietypublishing.orgC CathPhilosophical Transactions of the Royal Society A, 2018•royalsocietypublishing.org, vol. 376, no. 2133, Nov. 2018, doi: 10.1098/rsta.2018.0080.
- [20] B. Perry, R. U.-B. data and cognitive computing, and undefined 2019, “AI governance and the policymaking process: key considerations for reducing AI risk,” mdpi.comB Perry, R UukBig data and cognitive computing, 2019•mdpi.com, vol. 3, no. 2, pp. 1–17, Jun. 2019, doi: 10.3390/bdcc3020026.
- [21] H. Al Zadjali, “Building the right AI governance model in Oman,” ACM International Conference Proceeding Series, pp. 116–119, Sep. 2020, doi: 10.1145/3428502.3428516.
- [22] I. Ulnicane, W. Knight, T. Leach, B. C. Stahl, and W.-G. Wanjiku, “Framing governance for a contested emerging technology: insights from AI policy,” academic.oup.comI Ulnicane, W Knight, T Leach, BC Stahl, WG WanjikuPolicy and Society, 2021•academic.oup.com, 2021, doi: 10.1080/14494035.2020.1855800.
- [23] A. Buhmann, C. F.-T. in Society, and undefined 2021, “Towards a deliberative framework for responsible innovation in artificial intelligence,” Elsevier, Accessed: Feb. 23, 2024. <https://www.sciencedirect.com/science/article/pii/S0160791X20312781>
- [24] B. S. Barn, “Mapping the public debate on ethical concerns: algorithms in mainstream media,” Journal of Information, Communication and Ethics in Society, vol. 18, no. 1, pp. 38–53, Mar. 2020, doi: 10.1108/JICES-04-2019-0039.
- [25] C. D. Raab, “Information privacy, impact assessment, and the place of ethics *,” Computer Law and Security Review, vol. 37, Jul. 2020, doi: 10.1016/J.CLSR.2020.105404.
- [26] J. Buenfil, R. Arnold, B. Abruzzo, and C. Korpela, “Artificial Intelligence Ethics: Governance through Social Media,” 2019 IEEE International Symposium on Technologies for Homeland Security, HST 2019, Nov. 2019, doi: 10.1109/HST47167.2019.9032907.
- [27] R. J. D.-P. of the 20th A. International and undefined 2019, “The AI pandorica: linking ethically-challenged technical outputs to prospective policy approaches,” dl.acm.orgR J. DomanskiProceedings of the 20th Annual International Conference on Digital, 2019•dl.acm.org, pp. 409–416, Jun. 2019, doi: 10.1145/3325112.3325267.
- [28] B. Shneiderman, “Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems,” ACM Trans Interact Intell Syst, vol. 10, no. 4, Oct. 2020, doi: 10.1145/3419764.
- [29] J. K.-P. of the 2021 A. C. on Fairness and undefined 2021, “Outlining traceability: A principle for operationalizing accountability in computing systems,” dl.acm.orgJA KrollProceedings of the 2021 ACM Conference on Fairness, Accountability, and, 2021•dl.acm.org, pp. 758–771, Mar. 2021, doi: 10.1145/3442188.3445937.
- [30] W. Orr and J. L. Davis, “Attributions of ethical responsibility by Artificial Intelligence practitioners,” Inf Commun Soc, vol. 23, no. 5, pp. 719–735, Apr. 2020, doi: 10.1080/1369118X.2020.1713842.

- [31] L. Floridi et al., "An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," SpringerL Floridi, J Cowls, M Beltrametti, R Chatila, P Chazerand, V Dignum, C Luetge, R MadelinEthics, governance, and policies in artificial intelligence, 2021•Springer, vol. 28, no. 4, pp. 689–707, Dec. 2018, doi: 10.1007/s11023-018-9482-5.
- [32] H. S.-P. T. of the R. Society and undefined 2018, "Algorithmic accountability," royalsocietypublishing.orgH ShahPhilosophical Transactions of the Royal Society A, 2018 royalsocietypublishing.org, vol. 376, no. 2128, 2018, doi: 10.1098/rsta.2017.0362.
- [33] S. Larsson, F. H.-I. P. Review, and undefined 2020, "Transparency in artificial intelligence," policyreview.infoS Larsson, F HeintzInternet Policy Review, 2020 policyreview.info, vol. 9, no. 2, pp. 1–16, May 2020, doi: 10.14763/2020.2.1469.
- [34] P. de Almeida, C. dos Santos, J. F.-E. and Information, and undefined 2021, "Artificial intelligence regulation: a framework for governance," SpringerPGR de Almeida, CD dos Santos, JS FariasEthics and Information Technology, 2021•Springer, vol. 23, no. 3, pp. 505–525, Sep. 2021, doi: 10.1007/s10676-021-09593-z.
- [35] A. Daly et al., "Artificial Intelligence, Governance and Ethics: Global Perspectives," SSRN Electronic Journal, Jul. 2019, doi: 10.2139/SSRN.3414805.
- [36] C. van Noordt and G. Misuraca, "Exploratory Insights on Artificial Intelligence for Government in Europe," Soc Sci Comput Rev, vol. 40, no. 2, pp. 426–444, 2022, doi: 10.1177/0894439320980449.
- [37] N. Maslej et al., "Artificial intelligence index report 2023," arXiv preprint arXiv:2310.03715, 2023.
- [38] "J. Palfrey and U. Gasser, Interop: The Promise and... - Google Scholar." Accessed: Feb. 23, 2024. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=J.+Palfrey+and+U.+Gasser%2C+Interop%3A+The+Promise+and+Perils+of+Highly+Interconnected+Systems%2C+Basic+Books%2C+2012.&btnG=
- [39] C. Scott, "Regulation in the age of governance: The rise of the post-regulatory state," The Politics of Regulation: Institutions and Regulatory Reforms for the Age of Governance, pp. 145–174, 2004, doi: 10.4337/9781845420673.00016.
- [40] "U. Gasser, 'AI and the Law: Setting the Stage,'...-Google Scholar." Accessed: Feb. 23, 2024. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=U.+Gasser%2C+%E2%80%9CAI+and+the+Law%3A+Setting+the+Stage%2C%E2%80%9D+Medium%2C+26+June+2017%3B&btnG=
- [41] R. Chatila, K. Firth-Butterflied, ... J. H.-I. R. and, and undefined 2017, "The IEEE global initiative for ethical considerations in artificial intelligence and autonomous systems," academia.eduR Chatila, K Firth-Butterflied, JC Havens, K KarachaliosIEEE Robotics and Automation Magazine, 2017 academia.edu, Accessed: Feb. 23, 2024. [Online]. Available: <https://www.academia.edu/download/71373434/07886245.pdf>
- [42] S. Abiteboul, G. Miklau, J. Stoyanovich, and G. Weikum, "Data, responsibly (dagstuhl seminar 16291)," in Dagstuhl Reports, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.
- [43] M. U. Scherer, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies," SSRN Electronic Journal, Jun. 2015, doi: 10.2139/SSRN.2609777.
- [44] S. Wachter, B. Mittelstadt, L. F.-I. D. Privacy, and undefined 2017, "Why a right to explanation of automated decision-making does not exist in the general data protection regulation," academic.oup.comS Wachter, B Mittelstadt, L FloridiInternational Data Privacy Law, 2017 academic.oup.com, Accessed: Feb. 23, 2024. [Online]. Available: <https://academic.oup.com/idpl/article-pdf/doi/10.1093/idpl/ixp005/17932196/ixp005.pdf>

گزیده سیاستی

این گزارش با نگاهی نو به مفهوم حکمرانی هوش مصنوعی، پس از بررسی چالش‌ها و فرصت‌های کاربردی این فناوری، چارچوبی مفهومی جهت بهره‌مندی ایمن و مسئولانه از آن با هدف، تضمین عدالت و امنیت در کنار پیشرفت هوش مصنوعی برای شهروندان ارائه می‌دهد.



مرکز پژوهش‌های مجلس شورای اسلامی

تهران، خیابان پاسداران، روبروی پارک نیاوران (ضلع جنوبی، پلاک ۸۰۲)

تلفن: ۷۵۱۸۳۰۰۰ ☎ صندوق پستی: ۱۵۸۷۵-۵۸۵۵ ✉ پست الکترونیک: mrc@majles.ir

وبسایت: rc.majles.ir