Submit Date: March 23, 2025
Revise Date: May 2, 2025
Accept Date: May 18, 2025
Publish Date: May 23, 2025

The Encyclopedia of Comparative Jurisprudence and Law

Domains of Artificial Intelligence Utilization in Fundamental Human Rights in Light of Examining Its Legitimacy

Abbas Salahshour[1], **Ahmad Reza Tavakoli***[2], **Mohammadali Heidari**[2]

1. Ph.D. Student, Department of Theology, Najafabad Branch, Islamic Azad University, Najafabad , Iran.
2. Assistant Professor, Department of Theology, Najafabad Branch, Islamic Azad University, Najafabad, Iran.

* Corresponding Author's Email: Tavakolimady @gmail.com

ABSTRACT

Today, the manifestations of human rights in artificial intelligence and the aspects of its legitimacy represent a critical challenge centered on how to address related gaps, risks, and their influence on the principles of human rights. Key concerns in this domain include algorithmic transparency, cybersecurity vulnerabilities, injustice, bias and discrimination, privacy issues, data protection, liability for harm, and the absence of accountability, all of which have been examined using a library-based research method. This study focuses on the concept of "vulnerability" as a foundation for consolidating an understanding of the domain of artificial intelligence and fundamental human rights, which are frequently at the heart of concern. It employs this concept to explore the process of legitimizing the application of such technology. The findings indicate that, despite the progress made in the legal and regulatory governance of artificial intelligence and the acknowledgment that this domain requires continuous evaluation and agility in approach, the discussion advances the argument that, given the profound impacts of AI technologies—particularly on vulnerable individuals and groups and their human rights—it is crucial to evaluate their legitimacy through the lens of ethics before introducing them into spheres that affect individual rights.

Keywords: *Artificial Intelligence; Data Protection; Fundamental Rights; Legitimacy*

How to cite: Salahshour, A., Tavakoli, A. R., & Heidari, M. (2025). Domains of Artificial Intelligence Utilization in Fundamental Human Rights in Light of Examining Its Legitimacy. *The Encyclopedia of Comparative Jurisprudence and Law, 3*(1), 1-16.

تاریخ ارسال: ۳ فروردین ۱۴۰۴
تاریخ بازنگری: ۱۲ اردیبهشت ۱۴۰۴
تاریخ پذیرش: ۲۸ اردیبهشت ۱۴۰۴
تاریخ چاپ: ۲ خرداد ۱۴۰۴

دانشنامه فقه و حقوق تطبیقی

حوزه‌های بهره مندی هوش مصنوعی در حقوق بنیادین افراد در پرتو بررسی جهات مشروعیت آن

عباس سلحشور[۱]، احمد رضا توکلی [۲*]، محمدعلی حیدری[۲]

۱. دانشجوی دکتری گروه الهیات، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران.

۲. استادیار، گروه الهیات، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران.

* پست الکترونیک نویسنده مسئول: Tavakolimady @gmail.com

چکیده

امروزه جلوه‌های حقوق بشری در هوش مصنوعی و جهات مشروعیت آن، چالشی است که با نحوه پرداختن به آنها، شکاف‌ها و چالش ها، و تأثیرگذاری بر اصول حقوق بشر متمرکز می‌شود. چنین موضوعاتی عبارتند از: شفافیت الگوریتمی، آسیب پذیری‌های امنیت سایبری، بی عدالتی، سوگیری و تبعیض، مسائل مربوط به حریم خصوصی و حفاظت از داده ها، مسئولیت آسیب و عدم پاسخگویی که با روشی کتابخانه‌ای بررسی شده است. این پژوهش با تمکر بر «آسیب‌پذیری» برای تثبیت درک حوزه هوش مصنوعی و حقوق بنیادین افراد که اغلب مورد نگرانی است بهره برده و آن را در راستای بررسی مشروعیت بخشی به این تکنولوژی استفاده می‌کند. یافته‌ها نشان می‌دهد با توجه به کارهای مطلوب انجام شده در حوزه قانون و نظارت بر هوش مصنوعی و اذعان به اینکه این حوزه نیاز به ارزیابی مداوم و چابکی در رویکرد دارد، این بحث را پیش می‌برد که با توجه به شدت تأثیرات فناوری‌های هوش مصنوعی، به ویژه بر افراد و گروه‌های آسیب پذیر و حقوق بشر آن‌ها حائز اهمیت بوده و قبل از وارد نمودن آن‌ها بر عرصه و حضور حقوق افراد باید مشروعیت آن‌ها را با محک اخلاق سنجید.

کلیدواژگان: *هوش مصنوعی؛ حفاظت از داده‌ها؛ حقوق بنیادین؛ مشروعیت*

نحوه استناددهی: سلحشور، عباس.، توکلی، احمدرضا.، و حیدری، محمدعلی. (۱۴۰۴). حوزه‌های بهره مندی هوش مصنوعی در حقوق بنیادین افراد در پرتو بررسی جهات مشروعیت آن. *دانشنامه فقه و حقوق تطبیقی*، ۳(۱)، ۱۶-۱.

مقدمه

هوش مصنوعی[1] و پیامدهای آن برای حقوق اساسی بدون شک یکی از چالش‌های اصلی جامعه ما است. در واقع، در حالی که تاکنون توجه زیادی به پتانسیل این علم جدید برای حمایت از رشد اقتصادی شده است، این سوال که چگونه ممکن است بر حقوق اساسی تأثیر بگذارد کمتر مورد بررسی قرار گرفته است. بر اساس این فرض، آژانس حقوق بنیادین[2] اتحادیه اروپا گزارشی تهیه کرده است که خطرات مرتبط با استفاده از هوش مصنوعی را با توجه به چهار حوزه مختلف (ازجمله مصالح اجتماعی، سیاست پیشگویانه، خدمات بهداشتی و تبلیغات هدفمند) شناسایی می‌کند. دامنه کامل حقوق بنیادین را در رابطه با هوش مصنوعی، همانطور که در منشور و معاهدات ذکر شده است، بسته به زمینه‌ای که هوش مصنوعی در آن استفاده می‌شود، قابل اجرا هستند. تحقیقات نشان داده است که آگاهی کامل در خصوص تمام حقوق بنیادین که سیستم‌های هوشمند بر آن تأثیر می‌گذارند وجود ندارد. با این حال اطمینان از یک سیستم کنترلی مؤثر و قابل اعتماد برای نظارت بر این عرصه چالشی است که در صورت لزوم، مدیریت هرگونه تأثیر منفی سیستم‌های هوش مصنوعی بر حقوق بنیادین افراد، با استفاده از رویکردهای حفاظتی (مانند مقدمات حفاظت از داده) را باید مد نظر قرار دهند.

۱- هوش مصنوعی و جغرافیای مفهومی آن در حوزه اخلاق و تکنولوژی

دامنه یک نهاد قانونی که بر هوش مصنوعی نظارت می‌کند تا حد زیادی با تعریف اصطلاح هوش مصنوعی تعیین می‌شود. علیرغم کاربرد گسترده آن، هیچ تعریف روشن و قابل قبول جهانی از این بستر وجود ندارد. با این حال برخی هوش مصنوعی را شاخه‌ای از علوم رایانه می‌دانند و بنابراین «ویژگی‌های هوش مصنوعی را از طریق ترکیب هوش انسانی با تکنولوژی مطالعه می‌کند» (Stone et al., 2022).

بحث گسترده‌ای در مورد اینکه چگونه جوامع می‌توانند مزایای هوش مصنوعی را در عین شناسایی، رسیدگی و کاهش معایب، حفظ کنند، وجود دارد. در اینجا، رویکردی فراگیر برای ترویج استفاده از هوش مصنوعی برای شکوفایی انسان مطرح شده است. چنین رویکردی مستلزم توجه به جنبه‌های مختلف، به ویژه احترام به حقوق بشر و ارزش‌های اخلاقی است. این امر مستلزم آموزش و افزایش آگاهی است که به فعالیت‌های مسئولانه در سطح فردی، حرفه‌ای و شرکتی تبدیل می‌شود (Stone et al., 2022).

این امر هوش انسانی را به یک انتخاب طبیعی برای محک زدن پیشرفت هوش مصنوعی تبدیل می‌کند. حتی ممکن است به عنوان یک قانون اساسی پیشنهاد شود که هر فعالیتی که رایانه‌ها قادر به انجام آن هستند در صورت انجام با منطق انسانی، به عنوان نمونه‌ای از هوشمندی تلقی شوند. البته ناگفته نماند که پنین تطبیقی با هر میزان از توانایی ذهنی انسان فقط شرط کافی بوده نه یک شرط ضروری است. با این توصیف اگر چه سیستم‌های زیادی از جمله برنامه ریزی روزانه ورود و خروج هزاران پرواز در یک فرودگاه وجود دارند که حداقل از نظر سرعت از هوش انسانی فراتر می‌روند، اما هنوز نمی‌توان به صورت متقن اذعان کرد این عملکرد از هوش انسانی فراتر بوده و تحکم دارد. چرا که هنوز نمی‌توان چالش اخلاق، احساس و.... که سنخ تصمیم با هوش انسانی را در بر می‌گیرد را در متن هوش مصنوعی بیان کرد. بنابراین همواره این پرسش وجود دارد که در حیطه هوش است چه مسأله‌ای موضوعیت می‌یابد. برخی پیشنهاد می‌کنند که هوش مصنوعی به یک ماشین برای انجام وظایف شناختی مرتبط با ذهن انسان، از جمله «درک، استدلال، یادگیری، تعامل با محیط، حل مسئله، تصمیم‌گیری و حتی نشان دادن خلاقیت» نیاز دارد (Rai &

[1] - Artificial Intelligence (AI)

[2] - Agency for Fundamental Rights (FRA)

Constantinides, 2019). از طرفی نیز سازمان همکاری اقتصادی و توسعه[1] پیشنهاد می‌کند که یک سیستم هوش مصنوعی می‌تواند «پیش‌بینی‌ها، توصیه‌ها یا تصمیم‌گیری‌هایی را انجام دهد که بر محیط‌های واقعی یا مجازی تأثیر می‌گذارند» (OECD، ۲۰۱۹، ص ۷).

۲- ظهور فنآوری و حوزه‌های نمود آن در حقوق بنیادین افراد با محک اخلاق

اصول و ارزش‌هایی که بر رفتار این نوع هوش حاکم است و عملکرد آن را کنترل می‌کند، ایمنی و توانایی تشخیص درست و غلط را تضمین می‌کند. در مقابل سعی می‌شود اخلاق بخشی که فلسفه عصر معاصر را تشکیل می‌دهد را تحت عنوان تخصص "تکنولوژی-اخلاق" در کنه و ذات هوش مصنوعی درونی کند و بر مسائل اخلاقی مربوط به ربات‌ها و اشکال مختلف هوش مصنوعی متمرکز شود (Ghaly, 2022). بنابراین هوش مصنوعی می‌تواند روش‌های جدیدی از تفکر را فعال کند که به طراحی چشم اندازهای آینده محور با مزایای اخلاقی و سایر مزایای بالقوه انسان مدار متعدد متمرکز شود.

این چشم اندازها با مزایای بسیاری تحقیقات و سیاست‌های هوش مصنوعی را هدایت می‌کند. در مقابل نگرانی‌هایی است که هوش مصنوعی می‌تواند مسائل اخلاقی را مطرح کند و منجر به نقض حقوق انسان مدار در حوزه بشری شود. بسیاری از این نگرانی‌ها در ویژگی‌های خاص هوش مصنوعی ریشه دارند. چرا که بیشتر سیستم‌های هوش مصنوعی فعلی به مجموعه داده‌های گسترده برای اهداف آموزشی و اعتبارسنجی نیاز دارند. در جایی که این مجموعه داده‌ها حاوی داده‌های شخصی هستند، سؤالاتی در مورد حفظ حریم خصوصی و حفاظت از داده‌ها مطرح می‌شود.[2] قابلیت‌های فنی جدید همچنین چالش‌های جدید حفاظت از داده‌ها را ایجاد می‌کند (Veale et al., 2018). این موارد اغلب با نگرانی‌های مربوط به ایجاد اطمینان، ایمنی و امنیت مرتبط هستند. در پرتو این نگرش باید اذعان داشت هوش مصنوعی نگرانی متعدد اخلاقی و حقوق بشری را مطرح می‌کند. این امر می‌تواند منجر به سوگیری‌ها و تبعیض شود، نگرانی‌های مربوط به حریم خصوصی و امنیتی را افزایش داده، نابرابری‌های اقتصادی و سایر موارد را تشدید کند. برایند این آشفتگی این مهم است که به عدم تعادل سیاسی و قدرت منجر شده و تعامل، افکار و زندگی انسان‌ها را تحت تأثیر قرار دهد.

علاوه بر این نگرانی‌ها که تأثیر هوش مصنوعی بر سلوک بشری را در بر گرفته و متمرکز است، استفاده گسترده از هوش مصنوعی در جامعه مدرن می‌تواند پیامدهای اجتماعی مشکل‌ساز به‌ویژه برای اقلیت‌ها و گروه‌های آسیب‌پذیر نیز به همراه داشته باشد. مسأله‌ای که مشروعیت آن را هم از نظر اخلاقی و هم عرفی و نیز دینی مورد سوال قرار می‌دهد. (رودریگز، ۲۰۲۰). چالش جدی در این زمینه این است که مزایای اقتصادی وعده داده شده توسط هوش مصنوعی ممکن است نابرابری‌های موجود را تشدید کند و سوالات عدالت و توزیع را مطرح کرده و بر میزان مشروعیت آن در جامعه و در نزد افراد عامه و تحت نفوذ این فنآوری مورد نقد و پرسشگری قرار گیرد. چالشی جدی که عدم توجه به آن نه تنها اخلاقاً که از منظر دینی و عرفی نیز مردود می‌باشد.

همچنین این نگرانی‌ها در مورد تأثیر هوش مصنوعی بر اشتغال نیز وجود دارد. به ویژه (کاپلان و هن‌لین، ۲۰۱۹ ویلکاکس، ۲۰۲۰)، امکان نظارت کارگران و استثمار عمومی اکثر افراد توسط سازمان‌هایی که قادر به استفاده از قابلیت‌های هوش مصنوعی هستند (Coeckelbergh, 2020).

---

[1] - Organization for Economic Co-operation and Development (OECD)

[2] - EDPS. Artificial intelligence, robotics, privacy and data protection (Room document for the 38th international conference of data protection and privacy commissioners). Marrakech: EDPS; 2016 Retrieved from https://edps.europa.eu/sites/edp/files/publication/16- 10- 19_marrakesh_ai_paper_en.pdf.

۳- اقتضائات مشروعیت هوش مصنوعی در بازخورد با حقوق بنیادین امروزی بشری

۳-۱- حق بر تکریم

برای تعیین پیامدهای حقوق بشر تصمیم‌گیری هوش مصنوعی، ارزیابی حقوق بشر باید شامل ارزیابی انطباق سیستم‌های هوش مصنوعی با حقوق اساسی از جمله حق بر تکریم حقوق انسانی که از جمله مولفه‌ها و قواعد حقوق بشر می‌باشد، باشد. چنین ارزیابی خطراتی حق بر تکریم انسان را در بر می‌گیرد که برایند آن شامل احترام به کرامت انسانی، آزادی افراد، احترام به دموکراسی، عدالت و حاکمیت قانون، برابری، عدم تبعیض و همبستگی و حقوق شهروندی است. با این حال به طور خاص، حتی اگر امروزه حق بر تکریم انسان به دلیل توسعه سریع فناوری‌های جدید به چالش کشیده شده است، معنای کلی آن دست نخورده باقی مانده است. یعنی هر انسانی دارای یک «ارزش ذاتی» است که هرگز نباید توسط دیگران یا موضوع جدیدی در معرض خطر یا سرکوب قرار گیرد. تجاوز به اصولی که تکریم انسان را به محاق می‌برد، اعمالی «ناسازگار با حیثیت و حیثیت انسان» است. چنین اعمالی در سطح بین‌المللی محکوم می‌شوند، زیرا این اعمال مضر، «نفی تکریم انسانی» هستند و «تأثیر منفی» بر حیثیت و تمامیت اخلاقی دارند. از این نظر، حق بر تکریم انسان در چارچوب‌های قانونی، منطقه‌ای و بین‌المللی محافظت می‌شود و می‌تواند نقشی محوری در محدود کردن میزان تفویض اختیار به سیستم‌های هوش مصنوعی داشته باشد، همانطور که استفاده از آن در موارد مختلف نشان می‌دهد (Commissioner for Human Rights, 2019).

حق بر تکریم انسانی در جامعه فی الواقع یک اصل راهنما در حقوق بین‌الملل است و به عنوان یک «حق مادر» (به عنوان یک منبع حقوق و به عنوان صلاحیت ذاتی) و به عنوان منبع تعهد عمل می‌کند. به ویژه، کرامت فرد امروزه به عنوان «اصل راهنمای اساسی حقوق بین‌الملل حقوق بشر» در نظر گرفته می‌شود. یک «اصول اصلی حقوق بشر در مورد حق بر تکریم» وجود دارد. به همین دلیل، اسناد حقوق بشر نه تنها به طور یکسان، به تعبیر منشور سازمان ملل متحد، «ایمان به کرامت» شخص انسان را «دوباره» تأیید می‌کند، بلکه کرامت انسانی را به‌عنوان شالوده چهار مفهوم مختلف نیز به رسمیت می‌شناسد.

کرامت انسانی، در ابتدا، «پایه آزادی، عدالت و صلح در جهان است» که برای اولین بار توسط اعلامیه جهانی حقوق بشر اعلام شده است. علاوه بر این، همه افراد بشر مطابق با ماده ۱این اعلامیه از «برابری در کرامت» برخوردارند و حق دارند «توسعه معنوی» خود را در بازخورد با حق بر تکریم و تقدس انسانی، دنبال کنند. به همین دلیل، «پیشرفت و توسعه فنی و تکنولوژیک دراجتماع» باید مبتنی بر احترام به کرامت باشد. بنابراین، یک شناخت کلی وجود دارد که حق بر تکریم مبنای مفاهیم و اصول «آزادی» است. حتی زمانی که یک سند بین المللی در مورد این اصل سکوت می‌کند، حق بر تکریم مبتنی بر کرامت انسانی همچنان برای استقلال شخصی به عنوان یک هدف اساسی «اهمیت محوری» در حوزه‌های مختلف از جمله هوش مصنوعی را القاء می‌کند. به طور کلی توافق شده است که احترام به کرامت انسانی «یک قانون اساسی و قابل اجرا جهانی است» (Ginevra Le, 2022). ملاحظات کارایی و اثربخشی که توسل به فناوری‌های هوش مصنوعی را توجیه می‌کند، همچنان در معرض محدودیت‌های قانونی مهم است و ممکن است به طور هنجاری برای اطمینان از حفاظت از حق بر تکریم اشخاص، عملکردی ناکافی داشته باشد. کرامت انسانی مستلزم این است که مردم شایسته رفتار با احترام باشند. سیستم‌های هوش مصنوعی باید به گونه‌ای طراحی و راه‌اندازی شوند که از هجمه نابجا به انسان‌ها در مقابل سلامت جسمی و روانی و همچنین از احساس هویت فرهنگی آن‌ها محافظت شود (Richardson et al., 2019).

در وهله دیگر، هنگامی که سیستم‌های نظارت تصویری یا سایر سیستم‌های نظارتی در مناطقی نصب می‌شوند که به جای آن حریم خصوصی بالا مورد انتظار است، مانند اتاق‌های رختکن یا توالت‌ها، به دلیل شرمندگی شخصی که ممکن است ایجاد کند، به حق بر تکریم انسان آسیب وارد می‌شود. سوم نیز استفاده از داده‌های حساس می‌تواند کرامت انسانی را به خطر بیندازد، زیرا در موارد درخواست اطلاعات تهاجمی توسط کارفرمایان به رسمیت شناخته شده است.

۳-۲- اصل برتری انسان در تقابل با هوش مصنوعی مبنای مشروعیت ماهیت هوش مصنوعی

اساساً یکی از مراتبی که بر محدودیت هوش مصنوعی در مقابل هوش انسانی قرار می‌گیرد، اصل برتری انسان و هوش انسانی در مقابل هوش مصنوعی به عنوان یک ماشین هوشمند می‌باشد. شاید این جمله بیشتر از آنچه فنی باشد، دارای فلسفه حقوق بشری بوده و مبنای آن با احساس و حق و امتیاز طبیعی انسان در جامعه گره خورده باشد. با این حال، این طیف با توسعه هوش عمومی مصنوعی که به عنوان ابر هوش نیز شناخته می‌شود، بیشتر گسترش خواهد یافت. همانطور که در کتاب پایان کار جرمی ریفکین، پایان کار کشاورزان، کارگران یقه آبی و کارگران خدماتی به دلیل انقلاب‌های صنعتی اول، دوم و سوم پیش‌بینی می‌شود، پس از انقلاب صنعتی چهارم با هوش مصنوعی، روبات‌ها، بیوتکنولوژی، نانوتکنولوژی و وسایل نقلیه خودران، را می‌توان برای جابجایی شغل‌های عمومی انسانی توسعه داده و پیش بینی کرد. این مسئله معمای بزرگی برای پاسخگویی در جامعه معاصر بوده که چارچوب هوش مصنوعی را به عنوان مبنایی مادی می‌سازد. موضوعی که تنها با مادیت صرف عجین شده و بعد مادی انسان را در مقابل بعد اخلاقی به چالش می‌کشد.

۴- زمینه‌های کاربرد هوش مصنوعی در حقوق بنیادین بشر

این واقعیتی است که تاکنون توجه زیادی به پتانسیل این علم برای حمایت از رشد اقتصادی شده است، اما این پرسش که چگونه ممکن است بر حقوق اساسی تأثیر گذاشته و طیف وسیعی از حقوق بنیادین اشخاص را در بر گیرد، کمتر مورد بررسی قرار گرفته است. در ذیل سعی می‌شود این مهم در قالب حفاظت از داده ها، اصل عدم تبعیض در دسترسی به این علم و دیگر اصول بنیادین را به صورت وجیز مورد بررسی قرار دهیم.

۴-۱- حفاظت از داده ها

حق احترام به زندگی خصوصی و حفاظت از داده‌های شخصی (ماده ۷ و ۸ منشور حقوق بشر) هسته اصلی بحث‌های حقوق اساسی در مورد استفاده از هوش مصنوعی است. حق احترام به زندگی خصوصی و حفاظت از داده‌های شخصی در حالی که ارتباط نزدیکی با هم دارند، حقوق مجزا و مستقلی هستند. آن‌ها به عنوان حق "کلاسیک" برای حفاظت از حریم خصوصی و یک حق "مدرن" تر، حق حفاظت از داده‌ها توصیف شده اند. در این راستا هر دو تلاش می‌کنند از ارزش‌های مشابه، یعنی خودمختاری و کرامت انسانی افراد، محافظت کرده و به آن‌ها حوزه شخصی اعطا کنند که در آن بتوانند آزادانه شخصیت خود را توسعه داده، فکر کنند و عقایدشان را شکل دهند. بنابراین، آن‌ها پیش نیاز اساسی برای اعمال سایر حقوق اساسی، مانند آزادی اندیشه، وجدان و مذهب (ماده ۱۰ منشور)، آزادی بیان و اطلاعات (ماده ۱۱ منشور)، و آزادی اجتماعات و تشکل‌ها (ماده ۱۲ منشور) را تشکیل می‌دهند (Heleen, 2020).

در بازخورد با موضوع باید اذعان کرد؛ همانطور که بشر به ادغام هوش مصنوعی در جنبه‌های مختلف زندگی خود ادامه می‌دهد، واضح است که در بازخورد با این رویکرد حفظ حریم خصوصی و ملاحظات اخلاقی اهمیت فزاینده‌ای پیدا کند. مزایای بالقوه هوش مصنوعی بسیار گسترده است، اما خطرات مرتبط با استفاده از آن نیز متنوع، جدید و بسیار زیاد است. حاکمیت‌ها (قوا در این

مجال قوه قضاییه) به عنوان یک عامل در بهره از مزایای هوش مصنوعی، باید رویکردی فعال برای مقابله با چالش‌های فوق داشته باشد تا از حریم خصوصی افراد محافظت کرده و این اطمینان خاطر را به مخاطبان (دادجویان و دادخوهان) دهد که هوش مصنوعی به صورت اخلاقی و مسئولانه استفاده می‌شود. سازمان‌ها، نهادها، مراجع قانونی و قضایی و... که از هوش مصنوعی استفاده می‌کنند باید حفظ حریم خصوصی و ملاحظات اخلاقی را در طراحی و اجرای سیستم‌های هوش مصنوعی خود در اولویت قرار دهند. این شامل شفافیت در مورد جمع‌آوری و استفاده از داده‌ها، تضمین امنیت داده‌ها، ممیزی منظم برای تعصب و تبعیض، و طراحی سیستم‌های هوش مصنوعی است که به اصول اخلاقی پایبند می‌باشد. مراجعی که این ملاحظات را در اولویت قرار می‌دهند، به احتمال زیاد با مخاطبان خود اعتماد سازی کرده، از آسیب به شهرت خود جلوگیری می‌کنند و روابط قوی تری با مراجعین خود خواهند داشت (Lilian & Michael, 2017).

پس واضح است که با اولویت دادن به حریم خصوصی و اتخاذ سیاست‌های قوی حفاظت از داده ها، می‌توانیم اطمینان حاصل کرده که فناوری هوش مصنوعی به گونه‌ای توسعه می‌یابد و استفاده می‌شود که به حریم خصوصی افراد و سایر ملاحظات اخلاقی احترام می‌گذارد.

استفاده گسترده از فناوری‌های هوش مصنوعی ممکن است با ادامه توسعه فناوری‌ها، مسائل ناشناخته و نگرانی‌های جدیدی را در مورد حق احترام به زندگی خصوصی ایجاد کند. فناوری‌های مبتنی بر هوش مصنوعی ممکن است طرز فکر ما را در مورد حریم خصوصی تغییر دهند. ابزارهای الگوریتمی می‌توانند رفتار افراد را به روش‌های بی‌سابقه‌ای پیش‌بینی و آشکار کنند - بدون اینکه مردم حتی متوجه شوند که چنین اطلاعاتی را ارائه داده اند. به عنوان مثال، داده‌های شخصی به‌دست‌آمده از اینترنت ممکن است برای تبلیغات هدفمند مورد استفاده قرار گیرد و بسیاری از نگرانی‌های مربوط به حقوق اساسی را برانگیزد. مشکلات مرتبط با اشتراک‌گذاری داده‌های شخصی از طریق برنامه‌های تلفن هوشمند، به‌ویژه نگرانی‌های مهمی را ایجاد می‌کند. از جمله مصادیق مرتبط به اثرات مضر بالقوه که ممکن است حریم خصوصی اشخاص را در برگیرد؛ دستکاری و بهره‌برداری از آسیب‌پذیری‌های افراد، تبعیض، کاهش مسائل امنیتی و تقلب در هویت دیجیتال را می‌توان نام برد.

۴-۲- بررسی اصل عدم تبعیض در دسترسی به هوش مصنوعی

برابری انسان‌ها یکی از اصول اولیهٔحقوق بشر است. انسان‌های برابر نیازمندحقوق برابرند و به گونه‌ای آرمان خواهانه خواستار امکانات و فرصت‌هایی برابر در جامعه‌اند. حقوق بشر برای آنکه ناظر به حقوق همهٔ انسانها باشد باید بسیاری از تفاوتهای اجتماع انسانی را نادیده بگیرد و بسیاری از تبعیضات را محکوم کند. این بدان معنا است که قوانین و سیاست‌ها و برنامه‌ها نباید تبعیض آمیز باشد و همچنین مقامات دولتی نباید قوانین و سیاست‌ها و برنامه‌ها را به شیوه‌ای تبعیض آمیز یا خود سرانه اعمال یا اجرا کنند (Salari, 2018). لزوم توجه به برابری و منع تبعیض نه تنها در دنیای حقیقی بلکه در فضای الگوریتمی هوش مصنوعی نیز موضوعیت دارد و باید مورد توجه قرار گیرد. در یک تقسیم‌بندی مرسوم، تبعیض به دو صورت مستقیم و غیرمستقیم تبلور می‌یابد. تبعیض مستقیم به وضعیتی گفته می‌شود که در آن به وضوح با برخی از افراد یا گروه‌های دارای موقعیت مشابه رفتاری متفاوت صورت می‌گیرد. تبعیض غیرمستقیم نیز به وضعیتی گفته می‌شود که در آن یک سیاست یا رویهٔ ظاهراً خنثی افراد یا گروههایی را در مقایسه با دیگران در تنگنایی خاص قرار می‌دهد. صرفنظر از اینکه حسب رویه قصد تبعیض وجود دارد یاخیر، اگر در نهایت به ضرر یک فرد یا گروه گام برداشته شود این وضعیت مصداق مفهوم تبعیض غیرمستقیم قلمداد می‌شود.

انصاف و عدم تبعیض[1] در ورود و جایگاه شناسی هوش مصنوعی در فرایند دادرسی دارای چالش‌هایی است که که به طور مستقیم عدالت سایبورگی را تحت الشعاع قرار می‌دهد. با این حال قبل از مفهوم شناسی انصاف و عدم تبعیض در هوش مصنوعی باید به این مهم اشاره کرد که متغیرهای تاثیرگذار برای پردازش اطلاعات مربوط به حقوق افراد ممکن است به دلیل عدم واکنش سیستم در پردازش، این ریسک را ایجاد کند که برخی از داده‌ها در هنگام ترجمه به زبان الگوریتم از بین برود. همچنین کاملاً واضح است که غرض ورزی در ورود اطلاعات در تصمیم گیری‌های هوش مصنوعی تاثیر گذار می‌باشد. بنابراین انصاف و تبعیض در سیستم‌های الگوریتمی در سطح جهانی به‌عنوان موضوعات مهمی شناخته می‌شوند که باید آن را نه از منظر حقوقی که از نگاهی فنی مورد تجه قرار داد و نسبت به آن دغدغه داشت. امروزه اگر چه از دیدگاه نظارتی تلاش‌هایی با مفاهیم «درمان متفاوت» و «تأثیر متفاوت» در این عرصه صورت پذیرفته است. با این حال، نمی‌توان معادله‌ای بین هوش مصنوعی و انصاف ورزی تام در دادرسی برقرار کرد.

به هر حال از مثالهای بارز تبعیض در حوزۀ هوش مصنوعی و قانون می‌توان به کیس بریشا بوردن[2] و ورنون پراتر[3] اشاره کرد. بوردن دختر هجده سالۀ سیاه پوستی بود، بدون محکومیت قبلی و ورنون پراتر مرد سفیدپوست۴۱ ساله‌ای بود که قبالاً به سرقت مسلحانه متهم و به پنج سال زندان محکوم شده بود. هر دو در سال۲۰۱۴ به سرقت کالاهایی به ارزش حدوداً هشتاد دلار متهم شدند. الگوریتم هوش مصنوعی (پروفایل ارزیابی ریسک به منظور اعمال مجازاتهای جایگزین)[4] بوردن را پرخطر و دارای ریسک زیاد تکرار عمل مجرمانه در آینده رتبه بندی کرد، درحالیکه به آقای پراتر ریسک کم اختصاص داده شد. الگوریتم هوش مصنوعی به وضوح بین آن دو تبعیض قائل شد. ازاینرو ضروری است در طراحی الگوریتمهای یادگیری هوش مصنوعی از دسته بندی افراد یا گروهها اجتناب شود تا از هرگونه تبعیض جلوگیری گردد (Mostafavi Ardebili et al., 2022). بنابراین معیارهایی که از نظر قانونی مطابقت دارند، در عین حال به اندازه کافی ثابت هستند که بتوان کدگذاری شوند. این کاوش به صورت مستقل دشوار است، زیرا قانون و رویه قضایی انصاف و تبعیض را به‌عنوان مفاهیمی بنیادی و زمینه‌ای و در ضمن قوانین تعریف می‌کنند (Wachter et al., 2021). با این حال جوامع حقوقی و فنی می‌توانند با مشارکت و گفتگوی نزدیکتر در مورد چگونگی طراحی ملاحظاتی در این حوزه (تلاشی که در مقررات عمومی حفاظت از داده‌ها در اروپا) در مورد برابری و انصاف در هوش مصنوعی و حکمرانی سود ببرند. جامعه حقوقی می‌تواند از برخی از استراتژی‌های منسجم و ایستا ارائه شده توسط جامعه فنی برای اندازه گیری نابرابری و تعصب غیر شهودی بهره مند شود، در مقابل جامعه فنی نیز باید ماهیت زمینه‌ای و انعطاف پذیر برابری در قوانین را هنگام طراحی سیستم‌ها و مکانیسم‌های مرتبط به هوش مصنوعی بدون سوء گیری و تعصب در نظر بگیرد.

به عنوان تکمله در مبحث تبعیض در هوش مصنوعی به عنوان یک چالش جدی می‌توان به موضوع انصاف در هوش مصنوعی نیز پرداخت که از جمله مولفه‌های مسئولیت اجتماعی هر شخص طبیعی و حقوقی قلمداد می‌شود و تا به صورت تام و کمال در فرد و عملکرد وی وجود نداشته باشد نمی‌توان گفت این شخص دارای مولفه‌های مختص مسئولیت اجتماعی می‌باشد. پس با توجه به اهمیت ان این موضوع در کنار عدم تبعیض باید در هوش مصنوعی به ویژه در فرایند دادرسی مدنی وجود داشته باشد. در تشریح بحث از منظری پیشینه‌ای باید عنوان کرد: انصاف در هوش مصنوعی از سال ۲۰۱۰ نگاه قابل توجهی را هم در تحقیقات و هم در صنعت به خود جلب کرده است. برای چندین دهه، محققان

[1]- fairness and non-discrimination
[2] - Brisha Borden
[3] - Vernon Prate
[4] - COMPAS

دریافتند که ارائه یک تعریف واحد از انصاف تا حدی چالش برانگیز است زیرا انصاف یک مفهوم اجتماعی و اخلاقی است. این مفهوم عمدتاً شخصی بوده لذا در بافت اجتماعی تغییر می‌کند و در طول زمان تکامل می‌یابد. پس انصاف را به یک هدف نسبتاً چالش برانگیز برای دستیابی به عمل تبدیل می‌کند. از آنجایی که مسئولیت اجتماعی در هوش مصنوعی (تشریح خواهد شد) یک فرآیند تصمیم گیری متناسب با ارزش‌های اجتماعی است، در اینجا یک تعریف عادلانه را در زمینه تصمیم گیری اتخاذ می‌کنیم: «انصاف عبارت است از فقدان هرگونه تعصب یا جانبداری نسبت به یک فرد یا گروه بر اساس خصوصیات ذاتی یا اکتسابی آنها» (Mehrabi et al., 2022).

البته باید توجه داشت که حتی یک سیستم هوش مصنوعی ایده‌آل «عادلانه» که در یک زمینه خاص تعریف شده، همچنان ممکن است منجر به تصمیم‌گیری‌های جانبدارانه نماید. زیرا کل فرآیند تصمیم‌گیری شامل عناصر متعددی مانند سیاست‌گذاران و محیط است. در حالی که تعیین مفهوم انصاف دشوار است، شناسایی ناعادلانه / تعصب / تبعیض ممکن است آسان تر باشد. در این راستا و در جهت بازشناسایی انصاف می‌توان شش نوع تبعیض بیان کرد که به خوبی موضوع را تشریح می‌کند.

تبعیض مستقیم ناشی از ویژگی‌های محافظت شده افراد است در حالی که تبعیض غیرمستقیم از ویژگی‌های به ظاهر عصبی و غیر محافظت شده است؛

تبعیض سیستمیک به سیاست‌هایی مربوط می‌شود که ممکن است تبعیض را علیه زیر گروه‌های جمعیت نشان دهد؛

تبعیض آماری زمانی رخ می‌دهد که تصمیم گیرندگان از آمار متوسط برای نشان دادن افراد استفاده کنند؛

بسته به اینکه آیا تفاوت بین گروه‌های مختلف قابل توجیه است یا نه، تبعیض قابل توضیح و غیرقابل توضیح بیشتری داریم (Cheng et al.). به هر حال انصاف در هوش مصنوعی نه تنها گزینه‌ای جهت فاصله یافتن از تبعیض در این حوزه می‌شود که مسئولیت اجتماعی هوش مصنوعی را نیز در کنار اخلاق گرایی و توجه به قانون مورد ارزیابی مسئولیت اجتماعی تعریف می‌کند.

۴-۳- اصل حق بر امنیت شخصی از مبادی اعلامیه حقوق بشر به عنوان اصل فقهی ضابطه هوش مصنوعی

در کنار حق بر امنیت شخصی که از جمله مبناهای اعلامیه حقوق بشر در قالب یک اصل فقهی است، وجود شفافیت برای اجرای هر چه بهتر و کاملتر مقررات موضوع مهمی است که در این زمینه قابل طرح می‌باشد. عدم شفافیت الگوریتمی موضوع مهمی است که در خط مقدم بحث‌های حقوقی در مورد هوش مصنوعی قرار دارد با توجه به تکثیر هوش مصنوعی در مناطق پرخطر تأکید می‌کند که "فشار برای طراحی و مدیریت هوش مصنوعی برای پاسخگویی، منصفانه و شفاف بودن در حال افزایش است." عدم شفافیت الگوریتمی مشکل ساز است. دسای و کرول (۲۰۱۷) نشان می‌دهند که چرا با استفاده از نمونه‌هایی از افرادی که از شغل محروم شده‌اند، وام‌ها را رد کرده‌اند، در لیست‌های پرواز ممنوع قرار داده‌اند یا از مزایای آن محروم شده‌اند بدون اینکه بدانند چرا؛ این اتفاق غیر منتظره از طریق تصمیم‌گیری برخی نرم‌افزارهای پردازش شده حادث می‌گردد.. اطلاعات در مورد عملکرد الگوریتم‌ها اغلب جهت دسترسی ضعیف هستند و این مشکل را تشدید می‌کند.

بنابراین افراد در مورد تصمیم گیری الگوریتمی «حق اطلاع داشتن» دارند. موضوعی که می‌توان در مواد ۱۳ و ۱۴ مقررات عمومی حفاظت از داده‌های اتحادیه اروپا آمده است.

مقررات عمومی مزبور شامل تعدادی اعلان فردی و حقوق دسترسی است. اعلان‌های مورد نظر هم در متن دستور العمل ایجاد شده و هم در ارتباط با الزامات رضایت آن در بازخورد با مواد دیگر تفسیر شده است. توضیح داده خواهد شد که چگونه این حقوق و الزامات در زمینه تصمیم گیری الگوریتمی و پروفایل‌های

مرتبط اعمال می‌شود. مقررات مزبور به طور سیستمی، اعلان‌ها و حقوق دسترسی عمومی را ایجاد می‌کند. پس از جمع‌آوری اطلاعات شخصی از یک فرد، یک شرکت باید هدف جمع‌آوری داده‌ها، گیرندگان داده‌ها و مدت زمان نگهداری داده‌ها را ارائه دهد. اگر یک شرکت اطلاعات شخصی را نه مستقیماً از یک فرد، بلکه از طرف دیگری به دست آورد، تقریباً اطلاعات یکسان باید افشا شود. علاوه بر این محدودیت، مقررات اروپایی مزبور شامل حقوق دسترسی مثبت برای افراد، از جمله دسترسی به منبع داده‌ها و یک کپی از خود داده‌ها نیز می‌باشد. حقوق دسترسی ممکن است در فواصل زمانی مورد استناد قرار گیرد و به افراد این امکان را می‌دهد که به طور مرتب اطلاعاتی را که یک شرکت در مورد آن‌ها دارد بررسی کنند. همچنین به نحوه انتقال اطلاعات می‌پردازد. بنابراین اطلاعات باید «به شکلی مختصر، شفاف، قابل درک به راحتی و به آسانی قابل دسترس باشد، همچنین با استفاده از زبانی روشن و ساده نیز منتقل شود (Kaminski, 2019).

۴-۴- شفافیت به عنوان اصل اساسی در هوش مصنوعی

حق توضیح به عنوان یک مکانیسم و ضابطه ایده آل برای افزایش پاسخگویی و شفافیت تصمیم سازی خودکار در نظر گرفته می‌شود. با این حال، دلایل متعددی وجود دارد که هم به وجود قانونی و هم در امکان‌پذیری چنین حقی شک ایجاد می‌شود. در مقابل پس‌زمینه رژیم پاسخگویی الگوریتمی تقویت‌شده در مقررات اروپایی، به طور واضح شفافیت یک اصل اساسی است که از جمله تمهیداتی که مقررات عمومی حفاظت از داده‌های اتحادیه اروپا بر آن توجه کرده «حق توضیح» می‌باشد. در واقع، می‌تواند برای مخاطبان شگفت‌انگیز باشد که چگونه بسیاری از موضوعات این مقررات باز به قوانین دولتی شباهت دارند. البته نه به دلایل سنتی و حقوقی که به طور ضمنی در راستای حفظ حریم خصوصی صورت می‌گرفت، بلکه این رویکرد اغلب از شفافیت به عنوان عنصری برای پاسخگویی استفاده می‌کند تا ثابت نماید سیستم‌ها منصفانه هستند. با این حال در بحث حق توضیح، محوریت شفافیت به موجب مقررات عمومی حفاظت از داده‌های اتحادیه اروپا از بین رفته است. بسیاری از محققین، به صورت قابل توجهی واکنش منفی و بدبینانه به چنین رویکردی نشان نداده‌اند. چه که معتقدند نوع افشاگری در مورد تصمیم گیری الگوریتمی تحت مقررات عمومی در پیش گفته مورد نیاز است.

بنابراین حق توضیح به عنوان یک مکانیسم امیدوارکننده در پیگیری گسترده توسط دولت و صنعت برای پاسخگویی و شفافیت در الگوریتم‌ها، هوش مصنوعی، روباتیک و سایر سیستم‌های خودکار تلقی می‌شود. موضوعی که تاحدودی می‌توان چالش حفظ حریم داده‌های خصوصی در پرونده‌های قضایی را در فرایند دادرسی با الگوریتم هخای هوش مصنوعی را مورد حفاظت و حمایت قرار می‌دهد. سیستم‌های خودکار می‌توانند اثرات ناخواسته و غیرمنتظره زیادی داشته باشند. ارزیابی عمومی در خصوص دلیل استفاده از مکانیسم‌های الگوریتمی پیچیده و غیرشفاف و وسعت و منبع این مشکلات اغلب دشوار است. تبلیغات قابل توجهی در مورد تأثیرات قدرتمند چنین حق قانونی قابل اجرایی برای افراد داده و اختلال در صنایع فشرده داده افزایش یافته است، که می‌تواند توضیح دهد که چگونه روش‌های خودکار پیچیده و شاید غیرقابل درک در عمل کار می‌کنند.

۵- آسیب پذیری‌های امنیت سایبری در هوش مصنوعی

همانطور که اسوبا و ولسر (۲۰۱۷) مسائل امنیتی مختلف و مرتبط با هوش مصنوعی را برجسته می‌کند، به تصمیم‌گیری کاملاً خودکار که منجر به خطاها و تلفات پرهزینه می‌شود، نیز مد نظر قرار می‌دهند. موضوعی که عرصه جدیدی از حمله را بر اساس «آسیب‌پذیری رژیم داده‌ها» باز می‌کند (Kaminski, 2019). اهمیت مسئله وقتی نمود می‌یابد که این عرصه به دلیل پتانسیل نهفته تأثیر نامطلوب بر حقوق اساسی شهروندان داشته و یا ممکن است ایجاد نماید. چنین موضوعاتی از آنجایی که زیرساخت‌های

حیاتی اطلاعات که حقوق آحاد جامعه را موضوع قرار می‌دهد، مورد آسیب‌هایی با تأثیرات شدید قرار می‌دهند. پس مهم هستند و تهدیدی برای زندگی و امنیت انسانی و به تبع آن دسترسی به منابع محسوب می‌شوند. آسیب‌پذیری امنیت سایبری نیز تهدید قابل‌توجهی هستند، زیرا اغلب پنهان می‌شوند و فقط تا دیروقت (پس از ایجاد آسیب) آشکار می‌شوند. به عبارت دیگر با عدم شفافیت حاکم بر این فضا و کارکرد هوش مصنوعی خلط شده و در نهایت بخش مهمی از حقوق افراد را تحت تأثیر قرار می‌دهد. با این حال راهبردها و ابزارهای مختلفی برای حل این موضوع استفاده یا پیشنهاد می‌شود. به عنوان مثال، قرار دادن مکانیسم‌های حفاظتی و بازیابی مطلوب در نهاد و متن آسیب‌ها؛ در نظر گرفتن و رسیدگی به آسیب‌پذیری‌ها در فرآیند طراحی؛ درگیر کردن تحلیلگران انسانی در تصمیم گیری‌های حیاتی هوش مصنوعی؛ استفاده از برنامه‌های مدیریت ریسک؛ و ارتقاء نرم افزارهای مختلف که الگوریتمهای هوش مصنوعی بر آن جریان داشته و اطلاعات و هر آنچه مرتبط با مسئله می‌باشد را تفسیر می‌کنند. بنابراین در مقابل هجمه‌های مختلف آسیب می‌توان راهکارهایی را نیز را ملحوظ نظر قرار داده و در راستای حقوق بنیادین افراد در همگرایی با هوش مصنوعی به کار بست.

۶- جبرگرایی تکنولوژیک با محوریت هوش مصنوعی و نظم حقوقی در حقوق بنیادین

هر چشم اندازی برای هوش مصنوعی باید شامل مقررات مناسب و موثر باشد، که با توجه به سرعت و غیرقابل پیش بینی بودن توسعه این فناوری ها، کاری بسیار دشوار است. از یک سو، مقررات بیش از حد با جزئیات ممکن است منجر به محدودیت در نوآوری توسط شرکت‌های فناوری شود، با این حال جبر توسعه تکنولوژی در به‌روزرسانی مجموعه‌ای از مقررات را برای قانون‌گذاران پس از پیشرفت فناوری دشوارتر می‌کند. از سوی دیگر، مقررات گسترده‌تر ممکن است حفره‌هایی ایجاد کند که شرکت‌ها از آن‌ها برای سودجویی به جای رویکردهای مبتنی بر حقوق بشر در صورت امکان استفاده کنند. چه جبر گرایی فناورانه را بپذیریم یا رد کنیم، واضح است که هوش مصنوعی حوزه‌ای است که قانونگذاران، به ویژه در جوامع دموکراتیک، ناگزیر از آن عقب می‌مانند. در اتحادیه اروپا، مقررات فوق العاده دقیق، دست و پا گیر و فراسرزمینی در دهه گذشته برای تقویت بنیان نظم حقوقی اروپا طراحی شده است تا بتواند در برابر چالش‌های عصر دیجیتال مقاومت کند. چارچوب اصلی این رویکرد قبلاً توسط مقررات عمومی حفاظت از داده‌ها (GDPR)و اخیراً توسط قانون بازارهای دیجیتال (DMA) و قانون خدمات دیجیتال (DSA) شکل گرفته است. باید دید که آیا و اگر چنین است، چگونه این چارچوب با طرح پیشنهادی مورد بحث برای قانون هوش مصنوعی (AIA) تکمیل خواهد شد. با استفاده از ابزارهای نظارتی، اتحادیه اروپا و کشورهای عضو آن در تلاش برای دستیابی به هدف توسعه و اجرای قوانینی متفکرانه، مؤثر و مترقی و در عین حال رعایت حقوق اساسی بشر و رفاه جوامع هستند. این اقدامات نشان دهنده یک سازش کلی است (Stéphanie Laulhé & Yulia, 2024).

جبرگرایی تکنولوژی را در بادی امر می‌توان در مصالحه بین الزامات اصول و هنجارهای حقوقی و آزادی نوآوری است. از یک سو، هر مقررات مناسب باید با هدف حمایت از حقوق اساسی بشر و منطبق با اطمینان قانونی باشد. از سوی دیگر، نباید شکاف‌ها و تضادهایی را که در آن فن‌آوری‌ها اجازه تکثیر بی‌کنترل می‌شوند و می‌تواند به طور قابل‌توجهی بر حقوق بشر، آزادی‌های اساسی و منافع مشروع آسیب وارد کرده و حتی چند برابر کند. علاوه بر این، به نظر می‌رسد نسخه‌های نهایی این اقدامات نه تنها بین پارلمان اروپا، شورا و کمیسیون اروپا، بلکه بین قانون‌گذارانی که منافع دولت‌ها و شهروندان آن‌ها را در مقابل کسب‌وکارهایی که نماینده صنعت هستند، به‌عنوان یک مصالحه تشکیل می‌دهد.

فن‌آوری‌هایی که از سوی کسب‌وکار با کمک لابی‌گری و نوآوری به پیش می‌آیند، ممکن است جزء تقریباً نامرئی در این مبادله باشند، لذا به اقداماتی دامن بزنند و به منسوخ شدن برخی از مقررات قبل از انتشار حتی برای چاپ کمک کنند. این به ویژه نشان دهنده چارچوب قانونی در مورد هوش مصنوعی است. در حالی که بحث‌های شدیدی در مورد اینکه آیا مدلی مبتنی بر تخصیص سطوح مختلف خطر به مقررات معطوف به هوش مصنوعی به اندازه کافی خوب است و آیا داشتن جزئیات فناوری در ضمیمه‌های این قانون درست است یا خیر، در جریان بوده است، مشکلات جدیدی از جمله فناوری‌های مبتنی بر مدل‌های بزرگ زبانی ظاهر می‌شوند. ما را به هوش مصنوعی مولد نزدیکتر می‌کند. توسعه سیستم‌های هوش مصنوعی احتمالاً ما را به روی آوردن به جبر فناوری در نسخه‌ی سخت آن، اگر نگوییم سخت، حداقل نرم، نزدیک‌تر می‌کند (Chalmers, 2019).

جبر تکنولوژیک مدعی است که تکنولوژی توسعه جامعه را تعیین می‌کند و در برخی مظاهر افراطی، این مفهوم فناوری را عاملی مستقل می‌داند. به طور کلی، این اصطلاح به این باور اشاره دارد که فناوری «یک نیروی کلیدی حاکم در جامعه» است (Smith, 1994).

این نیز دیدگاهی است که می‌تواند زمانی ارزشمند باشد که گرایش‌های شکل دهنده اجتماعی فناوری را در نظر بگیریم. علاوه بر این، جبر تکنولوژیکی توجه را به تأثیر فناوری در هر دو سطح کلان و خرد جلب می‌کند و پیشنهاد می‌کند که احتیاطات در مورد تعیین بیش از حد جدی گرفته شود. یکی از دلایل این واقعیت این است که بسیاری از مصنوعات و سیستم‌های تکنولوژیک مدرن به قدری پیچیده هستند که هیچ فرد یا گروهی از افراد به طور کلی از آن‌ها آگاهی ندارند یا طرح را به طور کامل نمی‌دانند، به این معنی که خطرات غیرقابل پیش‌بینی وجود دارد که پیامدهای فناوری افزایش می‌یابد. با توجه به افزودن یک بعد دیجیتالی به فعالیت‌های انسانی و حقوق بشر، همچنین به کارگیری گسترده الگوریتم‌های پیچیده‌تر، ممکن است یک رویکرد بسیار مفید باشد (Hallström, 2022).

در راستای هدف این پژوهش باید اذعان کرد جبر تکنولوژیکی را به عنوان روندی در نظر بگیریم که در آن فناوری تا حد زیادی جامعه مدرن را به طور کلی و نظم حقوقی اروپا را به طور خاص تعیین می‌کند. ما استدلال می‌کنیم که فن‌آوری‌ها قبلاً شروع به شکل‌دهی به نظم حقوقی اروپا به‌سوی نظم حقوقی دیجیتالی تازه‌ای کرده‌اند. به این ترتیب، فناوری‌های پیشرفت‌کننده AIS ممکن است ستون‌های اساسی این نظم را تغییر دهند، اگر به طور کلی آن‌ها را بیگانه نکنند، مگر اینکه این فناوری‌ها با طراحی یکپارچه شده اند، یعنی در مرحله تصور و در استفاده/تصفیه/ارتقاء بعدی آنها. نفوذ هدفمند و در عین حال فراگیر، نمایه سازی و دستکاری با کمک هوش مصنوعی می‌تواند دموکراسی را تضعیف کند (Stéphanie Laulhé & Yulia, 2024). تصمیم‌گیری زمانی که مبتنی بر توصیه‌های الگوریتمی باشد، بر اساس عدم وضوح و فرسایش بحث عمومی می‌تواند برای حاکمیت قانون و ارزش‌های دموکراتیک مضر باشد. اما شاید فوری ترین و آشکارا مخرب ترین اثر هوش مصنوعی برای حقوق اساسی انسان باشد.

نتیجه‌گیری

این پژوهش اگرچه مروری کلی از مسائل حقوقی بی‌شمار، شکاف‌ها و چالش‌ها و اصول متاثر حقوق بشر را که به هوش مصنوعی مرتبط هستند، ارائه می‌کند، اما در مقابل مسائل و چالش‌ها و موضوعات مد نظر در این حوزه کاملاً ابتدایی بوده و تنها اندکی از این حوزه گسترده را در بر می‌گیرد. با این حال سعی شده به‌عنوان مرجع مفید و پله‌ای برای محققانی که مطالعات بیشتری در مورد این موضوع انجام می‌دهند عمل کند. به‌ویژه که بحث مسائل حقوقی هوش مصنوعی را با آسیب‌پذیری مرتبط

می‌کند – بحثی که در سطوح مختلف بسیار مورد نیاز است. علاوه بر این، سه اقدام کلیدی را ارائه کرده ایم که باید برای محافظت از آحاد آسیب پذیر جامعه در مقابل هوش مصنوعی مد نظر قرار داد. پژوهش نشان می‌دهد که حوزه‌های بهره برداری از هوش مصنوعی به صورت ممتد با جلوه‌هایی از حقوق بنیادین تلاقی پیدا می‌کند. چرا که بسیاری از موضوعات که ممکن است حوزه فعالیت هوش مصنوعی باشد، پیامدهای گسترده اجتماعی و حقوق بشری داشته باشد. این حوزه‌ها بر طیفی از اصول حقوق بشر تأثیر می‌گذارند: حفاظت از داده‌ها، برابری، آزادی‌ها، استقلال انسانی و تعیین سرنوشت افراد، کرامت انسانی، امنیت انسانی، رضایت آگاهانه، صداقت، عدالت و برابری، عدم تبعیض، حریم خصوصی و حتی حق تعیین سرنوشت، مسائلی هستند که فعالیت هوش مصنوعی می‌تواند بر آن‌ها تأثیر گذاشته و در صورت عدم دقت نظر بر آن حتی تأثیرات منفی می‌گذارد. آسیب پذیری حقوق اساسی مردم در بازخورد با آنچه تکنولوژی نوین با موضوعیت هوش مصنوعی خوانده می‌شود، نمودار قالبی است که حتی ممکن است در برخی عرصه‌ها مشروعیت هوش مصنوعی را تحت تأثیر قرار دهد.

طرح بحث از تلاقی هوش مصنوعی و خلط این تکنولوژی با حقوق اساسی بشر یک طرف و مشروعیت آن در حوزه‌های مختلف کاری و حقوق بشری طرف مهم ماجرا می‌باشد. به هر حال از آنجایی که فناوری‌های هوش مصنوعی با حجم وسیعی از داده‌ها کار می‌کنند، اثرات متقاطع و چندگانه‌ای خواهند داشت. این اثرات مسائل حقوقی و حقوق بشری مرتبط با آن‌ها و تأثیرات روی افراد را تشدید می‌کند. اگر برنامه‌ها و رژیم‌های موضوع هوش مصنوعی را بدون توجه به تأثیرات بالقوه چنین فناوری‌هایی چه بر حقوق بشر، ارزش‌های اخلاقی و اجتماعی (یعنی طراحی، تجزیه و تحلیل، ارزیابی تأثیر از حریم خصوصی یا اخلاقیات) بررسی نماییم، می‌توان اثرات و نشانه‌های آن را به خوبی مشاهده نمود.

مشروعیت هوش مصنوعی به فراخور گروه‌ها و حقوق افراد قابل تفسیر می‌باشد. به عبارت دیگر گروه‌ها و جوامعی که بیشتر تحت تأثیر چنین مسائلی قرار می‌گیرند بسته به زمینه، کاربرد و استفاده از هوش مصنوعی، همانطور که در متن نیز بررسی شد، متفاوت خواهند بود. نیاز اساسی برای مقابله با عواملی که باعث آسیب‌پذیری می‌شوند با کاهش اثرات نامطلوب، توسعه ظرفیت‌ها برای انعطاف‌پذیری و مقابله با علل اصلی آسیب‌پذیری‌ها بر حقوق بنیادین افراد تنها با مشروعیت بخشی هوش مصنوعی به عنوان یک واقعیت نوین تکنولوژیک براورده می‌شود. چرا که با مشروعیت بخشی نظارت و مقررات زایی در این عرصه شکلی قانونی و روسمی به خود می‌گیرد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافعی وجود ندارد.

EXTENDED ABSTRACT

The increasing integration of artificial intelligence (AI) into various facets of contemporary life has prompted critical discussions regarding its impact on fundamental human rights. While AI offers transformative potential, particularly in fostering innovation and economic growth, its unchecked implementation raises substantial concerns about legality, ethics, and legitimacy. The conceptual scope of AI, however, remains contested. Although widely applied, no universally accepted definition exists, which complicates legal oversight. Some scholars view AI as a technological extension of human cognition, encompassing perception, reasoning, learning, interaction, and creativity (Rai & Constantinides, 2019; Stone et al., 2022). Others emphasize its capacity to influence both real and virtual environments by making predictions or decisions based on data processing. This interpretive flexibility

complicates the establishment of ethical and legal standards. Nevertheless, the intersection of AI and human dignity, especially through the lens of vulnerability, necessitates a comprehensive framework. As AI technologies gain autonomy and application in high-stakes domains such as healthcare, justice, and surveillance, concerns surrounding privacy, accountability, and bias intensify (Veale et al., 2018). These risks disproportionately affect marginalized groups, further entrenching systemic inequalities. Consequently, any legitimate deployment of AI must prioritize human rights through transparent algorithmic structures and ethical safeguards (Ghaly, 2022).

The legitimacy of AI must be interrogated in terms of its alignment with fundamental principles such as dignity, equality, security, and privacy. The right to human dignity, as enshrined in international charters, functions as the cornerstone of fundamental rights and offers a vital threshold for evaluating AI applications (Commissioner for Human Rights, 2019; Ginevra Le, 2022). Dignity implies intrinsic value and personal autonomy, which automated decision-making systems may compromise. For instance, video surveillance or employer data collection in sensitive contexts threatens personal integrity and psychological well-being (Richardson et al., 2019). Moreover, a tension exists between the efficiency-driven rationale of AI and the supremacy of human agency. While some advocate for a "human-in-the-loop" model, others warn against the gradual erosion of personal sovereignty as algorithmic governance becomes normalized. With developments in artificial general intelligence (AGI), this concern becomes more urgent. Scholars such as Rifkin have warned about the systemic displacement of labor across historical industrial revolutions—a trend that AI may intensify in the Fourth Industrial Revolution. This challenges ethical and philosophical constructs that underpin human superiority, autonomy, and labor rights. Therefore, the legitimacy of AI must rest not merely on performance metrics but on its ethical implications and the extent to which it honors the dignity and personhood of those it affects.

Furthermore, the integration of AI raises complex questions regarding data protection, non-discrimination, and equitable access. At the heart of privacy concerns are the twin rights to respect for private life and personal data protection, as articulated in Articles 7 and 8 of the Charter of Fundamental Rights of the European Union (Heleen, 2020). These rights are essential for safeguarding individual autonomy and enabling the free development of personality. Yet, AI systems often rely on vast datasets, raising the specter of invasive profiling, identity theft, and targeted manipulation. Transparent data usage protocols, consent frameworks, and regular audits are necessary to maintain public trust and ethical compliance (Lilian & Michael, 2017). Equally pressing is the challenge of algorithmic bias and discrimination. AI systems can inadvertently perpetuate existing inequalities when trained on skewed datasets. Case studies like that of Brisha Borden and Vernon Prater illustrate the disproportionate risk assessment AI systems impose on individuals based on race or background (Mostafavi Ardebili et al., 2022). These outcomes highlight the critical need for fairness, which encompasses both procedural transparency and outcome equity (Wachter et al., 2021). Fairness, however, remains a contested concept, shaped by cultural, social, and temporal contexts (Mehrabi et al., 2022). The absence of a universal definition complicates legal codification and judicial enforcement. Even ideal AI systems may yield biased decisions due to underlying systemic and policy variables. This underscores the necessity for interdisciplinary collaboration

between legal scholars and technical developers to align AI systems with human rights mandates.

In the context of personal security, transparency emerges as both a legal imperative and a functional necessity. The opacity of AI decision-making—referred to as the "black box" problem—presents significant legal and ethical challenges. Individuals affected by algorithmic outcomes, such as employment denial or loan rejection, often remain unaware of the logic behind these decisions (Kaminski, 2019). To mitigate this, European data protection regulations (e.g., GDPR Articles 13 and 14) mandate disclosure of data processing purposes, recipient categories, and storage durations. Yet, even these frameworks fall short of ensuring substantive understanding or meaningful consent. Calls for a "right to explanation" aim to bridge this gap by obligating AI developers to offer intelligible and accessible justifications for automated decisions. Nevertheless, critics argue that the technical complexity of AI systems often renders such rights illusory in practice. The need for transparency extends beyond individual rights—it is crucial for systemic accountability. Legal scholars and practitioners advocate for mechanisms that enable regular audits, stakeholder engagement, and redressal channels. Without these, the risk of arbitrary or discriminatory practices escalates. Transparency also intersects with cybersecurity concerns. As AI systems become integral to critical infrastructure, their susceptibility to data breaches and algorithmic manipulation poses grave threats to public safety (Kaminski, 2019). Robust cybersecurity measures, human oversight, and resilience protocols are essential to counteract these vulnerabilities.

The concept of technological determinism—or the idea that technological developments dictate social and legal structures—further complicates the debate over AI legitimacy. Some scholars argue that law often lags behind technological progress, leading to regulatory vacuums that jeopardize human rights (Stéphanie Laulhé & Yulia, 2024). The European Union has attempted to bridge this gap through a series of legislative instruments such as the General Data Protection Regulation (GDPR), Digital Markets Act (DMA), and the proposed Artificial Intelligence Act (AIA). These aim to create a harmonized legal framework that balances innovation with ethical safeguards. However, critics caution that overregulation may stifle innovation, while underregulation risks unchecked corporate dominance and human rights violations. The influence of powerful tech lobbyists may further dilute legal protections. Moreover, AI systems—particularly those built on large language models—raise new ethical dilemmas not fully contemplated in existing laws (Chalmers, 2019). These include the spread of misinformation, erosion of democratic deliberation, and algorithmic surveillance. Technological determinism, when left unchecked, could undermine democratic accountability and public trust. Acknowledging this, scholars propose embedding ethical norms at the design stage of AI systems—a principle known as "ethics by design" (Hallström, 2022). By integrating human rights considerations into the foundational architecture of AI technologies, societies can better navigate the complex interplay between innovation and legitimacy.

In conclusion, while this study offers an introductory survey of the intersection between AI and fundamental human rights, it reveals significant gaps in the ethical, legal, and institutional frameworks that govern this rapidly evolving domain. AI technologies pose multifaceted challenges that intersect with human dignity, data protection, equity, personal security, and democratic governance. Addressing these challenges demands a

paradigm shift—from reactive regulation to proactive ethical design, from siloed governance to interdisciplinary collaboration, and from mere compliance to sustained accountability. The legitimacy of AI hinges not on its technical sophistication but on its alignment with the values and rights that underpin democratic societies. Moving forward, lawmakers, technologists, and civil society must engage in continuous dialogue to ensure that AI serves the public good without compromising individual freedoms. This study advocates for a rights-based approach to AI governance that centers on the vulnerable, foregrounds ethical scrutiny, and upholds the principles of justice, equality, and human dignity.

References

Chalmers, A. W. (2019). Choosing lobbying sides: the general data protection regulation of the European Union. *J. Public Policy*(No39). https://doi.org/10.1017/S0143814X18000223

Cheng, Varshney, & Liu. Socially Responsible AI Algorithms: Issues, Purposes, and Challenges. *Journal of Artificial Intelligence Research*(N5).

Coeckelbergh, M. (2020). *AI Ethics*. https://doi.org/10.7551/mitpress/12549.001.0001

Commissioner for Human Rights. (2019). Unboxing Artificial Intelligence: 10 steps to protect Human Rights. In. European Office of the Commissioner for Human Rights.

Ghaly, M. (2022). For the Classical Beginnings of Self-Propelled Machines, an Article Published on the Internet, Aristotle's Politics. In.

Ginevra Le, M. (2022). AI vs Human Dignity: When Human Underperformance is Legally Required. *Groupe d'études géopolitiques*(4).

Hallström, J. (2022). Embodying the past, designing the future: technological determinism reconsidered in.

Heleen, L. J. (2020). An approach for a fundamental rights impact assessment to automated decision-making. *International Data Privacy Law*, *10*(1). https://doi.org/10.1093/idpl/ipz028

Kaminski, M. E. (2019). Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability. *University of Colorado Law School University of Colorado Law School*, *92*. https://doi.org/10.2139/ssrn.3351404

Lilian, E., & Michael, V. (2017). Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking for. *16 DUKE L. & TECH. REV. 17*, 44. https://doi.org/10.31228/osf.io/97upg

Mehrabi, N., Morstatter, F., Nripsuta, S., Kristina Lerman, A. U., & Galstyan, A. (2022). A Survey on Bias and Fairness in Machine Learning. https://doi.org/10.1145/3457607

Mostafavi Ardebili, M. M., Taghizadeh Ansari, M., & Rahmati Far, S. (2022). Functions and Requirements of Artificial Intelligence from the Perspective of Fair Trial. *Biannual Journal of Law and New Technologies*, *3*(6).

Rai, P., & Constantinides, S. (2019). Next-generation digital platforms: toward human-AI hybrids. *Mis Quart*, *43*(1).

Richardson, R., Schultz, J., & Crawford, K. (2019). 'Dirty data, bad predictions: how civil rights violations impact police data, predictive policing systems, and justice'. *NYULR*.

Salari, A. (2018). Equality and Non-Discrimination in the Human Rights System. *Legal Research Journal*, *17*(35).

Smith, M. (1994). Technological Determinism in American Culture. In *Does Technology Drive History? The Dilemma Of Technological Determinism*.

Stéphanie Laulhé, S., & Yulia, R. (2024). Challenges to Fundamental Human Rights in the age of Artificial Intelligence Systems: shaping the digital legal order while upholding Rule of Law principles and European values. *Era Forum*(No24). https://doi.org/10.1007/s12027-023-00777-2

Stone, P., Brooks, R., Brynjolfsson, E., & Calo, R. (2022). Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. *ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016*.

Veale, M., Binns, R., & Lilian, E. (2018). Algorithms that remember: model inversion attacks and data protection law. *Royal Society*, *376*(2133). https://doi.org/10.1098/rsta.2018.0083

Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*. https://doi.org/10.1016/j.clsr.2021.105567