University of Tehran Press

Publice Law Studies Quarterly

Homepage: http://jplsq.ut.ac.ir

Online ISSN: 2423-8139

Volume: --, **Issue:** --
-- --

A Critical Study of the EU Proposal on Harmonization of Artificial Intelligence Act, Form the Perspective of, Combating Cyber Terrorism

Najmeh Razmkhah[1]✉

1. Assistant Prof., Department of Law, Faculty of Law and Social Science, Payame Noor University, Tehran, Iran. Email: razmkhah@pnu.ac.ir

Article Info

Article Type:
Research Article

Pages: 1-22

Received:
2022/05/13
Received in revised form:
2022/09/10
Accepted:
2022/10/24
Published online:
----/--/--

Keywords:
European Union, Cyber Terrorism, Right to Privacy, Rule of Non-Discrimination, Artificial Intelligence Act (AIA).

Abstract

Recourse to the capabilities of artificial intelligence has undoubtedly affected human life. Activists in this field always talk about the effective role of artificial intelligence in ensuring human welfare and security. Countering terrorist acts in cyberspace is one of the issues that have been raised under the title of advantages of artificial intelligence. But the use of this technology, despite its various benefits, has raised serious concerns among human rights defenders for violating some fundamental human rights principles. Concerns of this kind and the lack of comprehensive and enforceable regulations in the field of monitoring and controlling the effects of using artificial intelligence, led the European Union to present a draft law on artificial intelligence (AIA). But given some ambiguities and shortcomings in the content of the draft, have the authors of this document been able to take into account the various dimensions of the human rights challenges posed by the use of artificial intelligence? The present article, with a critical look, seeks to find the answer to this question, by using the analytical-descriptive method.

How To Cite

Razmkhah, Najmeh (2023). A Critical Study of the EU Proposal on Harmonization of Artificial Intelligence Act, Form the Perspective of, Combating Cyber Terrorism. *Public Law Studies Quarterly*, -- (--), 1-27. DOI: https://doi.com/10.22059/JPLSQ.2022.343006.3086

DOI

10.22059/JPLSQ.2022.343006.3086

Publisher

University of Tehran Press.

فصلنامه مطالعات حقوق عمومی

Homepage: http://jplsq.ut.ac.ir

انتشارات دانشگاه تهران

شاپا الکترونیکی: ۸۱۳۹-۲۴۲۳

دوره: -- شماره: --

نقدی بر پیش‌نویس قانون اتحادیهٔ اروپا در همسان‌سازی قوانین حاکم بر هوش مصنوعی، از منظر مقابله با تروریسم سایبری

نجمه رزمخواه[1]✉

۱. استادیار، گروه حقوق، دانشکدهٔ حقوق و علوم اجتماعی، دانشگاه پیام نور، تهران، ایران. رایانامه: razmkhah@pnu.ac.ir

اطلاعات مقاله

نوع مقاله: پژوهشی

صفحات: ۱-۲۲

تاریخ دریافت:
۱۴۰۱/۰۲/۲۳

تاریخ بازنگری:
۱۴۰۱/۰۶/۱۹

تاریخ پذیرش:
۱۴۰۱/۰۸/۰۲

تاریخ انتشار برخط:
----/--/--

کلیدواژه‌ها:
اتحادیهٔ اروپا، تروریسم سایبری، حق بر حریم خصوصی، قاعدهٔ منع تبعیض، قانون هوش مصنوعی.

چکیده

استفاده از قابلیت‌های هوش مصنوعی، بی‌شک زیست بشری را تحت تأثیر خود قرار داده است. فعالان این حوزه، همواره از نقش مؤثر هوش مصنوعی در تأمین رفاه و امنیت بشری صحبت می‌کنند. مقابله با اقدامات تروریستی در فضای مجازی از جمله مواردی است که تحت عنوان مزایای ناشی از به‌کارگیری هوش مصنوعی مطرح شده است. اما توسل به این فناوری با وجود مزایای مختلف، به طرح نگرانی‌های جدی از سوی حامیان حقوق بشر، به‌دلیل نقض برخی از اصول بنیادین حقوق بشری انجامیده است. نگرانی‌هایی از این دست و نبود مقررات جامع و لازم‌الاجرا در زمینهٔ نظارت و کنترل آثار ناشی از کاربرد هوش مصنوعی، اتحادیهٔ اروپا را بر آن داشت تا با هدف قانونمندسازی فعالیت‌های مرتبط با هوش مصنوعی در بخش‌های مختلف، از جمله مقابله با اقدامات تروریستی، پیش‌نویس همسان‌سازی قوانین حاکم بر هوش مصنوعی را ارائه دهد. اما آیا با توجه به وجود برخی نواقص در محتویات پیش‌نویس، تهیه‌کنندگان این سند، توانسته‌اند به‌خوبی ابعاد مختلف چالش‌های حقوق بشری ناشی از کاربرد هوش مصنوعی را مدنظر قرار دهند؟ پژوهش حاضر با رویکردی انتقادی و با استفاده از شیوهٔ تحلیلی- توصیفی در پی یافتن پاسخی برای این پرسش است.

استناد

رزمخواه، نجمه (۱۴۰۲). نقدی بر پیش‌نویس قانون اتحادیهٔ اروپا در همسان‌سازی قوانین حاکم بر هوش مصنوعی، از منظر مقابله با تروریسم سایبری. *مطالعات حقوق عمومی*، -- (--)، ۱-۲۷.
DOI: https://doi.com/10.22059/JPLSQ.2022.343006.3086

DOI

10.22059/JPLSQ.2022.343006.3086

ناشر

مؤسسه انتشارات دانشگاه تهران.

۱. مقدمه

هوش مصنوعی، به‌عنوان یک فناوری نوین، ابعاد مختلف زیست بشری از اقتصاد تا امنیت را تحت تـأثیر خود قرار داده است. فعالان حوزهٔ هوش مصنوعی، همواره از قابلیت‌های این فناوری و نقش مـؤثر آن در مقابله با مشکلات و معضلات جوامع امروزی، صحبت می‌کنند (Spike Back, 2018: 23).

افزایش اقدامات تروریستی در فضای مجازی و استفاده از فضای مجازی به‌منظور تأمین منابع مالی و همچنین جذب افراد در گروه‌های تروریستی، امروزه به‌نحو روزافزونی در حال افزایش است. در این زمینه طراحان و تهیه‌کنندگان سیستم‌های هوش مصنوعی مدعی‌اند که می‌توان به‌منظور پـیش‌بینـی اقـدامات تروریستی، تشخیص اطلاعات جعلی و همچنین، شناسایی افراد آسیب‌پذیر در برابر افراط‌گرایی، با اتکا به قابلیت‌های هوش مصنوعی، گام مؤثری در جهت مقابله با فعالیت‌های تروریستی برداشت. از سوی دیگر، به‌دلیل مشکلات موجود در حوزهٔ جمع‌آوری اطلاعات و طراحی الگوریتم‌های هـوش مصـنوعی و امکـان وجود سوگیری، نگرانی‌هایی از سوی فعالان مدنی از منظر نقض برخی از اصول بنیادین حقـوق بشـری، مثل حق بر حریم خصوصی و منع اعمال رفتارهای تبعیض‌آمیز مطرح شده است.

نگرانی‌های مطرح‌شده در خصوص عملکرد هوش مصنوعی و همچنین نبود نظام حقوقی لازم‌الاجـرا و مدون در زمینهٔ کنترل و نظارت بر فرایندهای مرتبط با طراحی و ارائهٔ سیسـتم‌هـای هـوش مصـنوعی، کمیسیون اروپایی را بر آن داشت که در آوریل ۲۰۲۱، پیش‌نویسی را با عنوان «قـانون اتحادیـهٔ اروپـا در همسان‌سازی قوانین حاکم بر هوش مصنوعی»[1]، بـا هـدف ایجـاد نظـام حقـوقی واحـد و هماهنـگ، در خصوص فرایند تولید، عرضه، به‌کارگیری و استفاده از سیستم‌های هوش مصنوعی در قارهٔ اروپا ارائه دهد. در صورت تصویب، این مقررات در چارچوب نظام حقوقی ۲۷ کشور عضو اتحادیه لازم‌الاجرا می‌شود و با توجه به مفاد بند ۱ مادهٔ ۲ پیش‌نویس، به عرضه‌کنندگان و ارائه‌دهندگان سیستم‌های هوش مصنوعی به اتحادیهٔ اروپا، که در کشورهای ثالث (خارج از قلمرو اتحادیه) مستقرند، نیز تسری می‌یابد. با تصویب ایـن قانون، اتحادیهٔ اروپا قصد دارد از توسعهٔ هوش مصنوعی قابل اعتماد و اخلاق‌مدار حمایت کنـد و آن‌طـور که شورای اروپایی مدعی است، رهبری جهان را در این زمینه به‌دست گیرد.

اما با نگاهی به متن پیش‌نویس، به‌نظر می‌آید در برخی موارد، با ابهام‌ها و ایرادهایی همراه است. مواردی که به طرح سؤال‌ها و ابهام‌های مختلفی منجر می‌شود، از جمله اینکه آیا تهیه‌کنندگان این سند توانسـته‌انـد به‌خوبی ابعاد مختلف چالش‌های حقوق بشری ناشی از به‌کارگیری هوش مصنوعی را با توجه به نگرانی‌های

1. Harmonized Rules on Artificial Intelligence (ARTIFICIAL INTELLIGENCE ACT)

مطرح‌شده در این زمینه، مدنظر قرار دهند؟ آیا با توجه به آسیب‌هایی که ممکن است در اثر توسل به هوش مصنوعی، به اشخاص وارد شود، نظام حقوقی مشخصی را در خصوص دادخواهی و رسیدگی به دعاوی افراد آسیب‌دیده پیش‌بینی کرده‌اند؟ پاسخگویی به این قبیل پرسش‌ها و چالش‌ها، هدفی است که مقالهٔ حاضر با استفاده از شیوهٔ تحلیلی – توصیفی، دنبال می‌کند. به همین منظور ابتدا به قابلیت‌های هوش مصنوعی در مواجهه با تروریسم سایبری پرداخته می‌شود، سپس چالش‌های موجود در این زمینه بیان می‌شود و در ادامه با نگاهی انتقادی، مفاد پیش‌نویس قانون هوش مصنوعی اتحادیهٔ اروپا، بررسی و سپس پیشنهادهایی به‌منظور رفع ایرادهای وارده، بر متن پیش‌نویس از منظر حقوقی مطرح می‌شود.

۲. کلیات و مفاهیم

در ابتدا و پیش از ورود به بحث اصلی مقاله، با توجه به اهمیت جایگاه برخی از واژه‌ها و عبارات، در مباحث بعدی و همچنین با توجه به تخصصی بودن آنها، به‌اختصار توضیحاتی در خصوص مفهومشان ارائه شده است.

۲. ۱. هوش مصنوعی[1]

فناوری هوش مصنوعی با توسعه و پیشرفت ابزارهایی ارتباط دارد که می‌توانند برخی از قابلیت‌های انسانی را مثل برنامه‌ریزی و یادگیری را انجام دهند. گاهی اوقات به‌اشتباه، هوش مصنوعی به‌جای یادگیری ماشین[2] استفاده می‌شود، درحالی‌که یادگیری ماشین، از زیرشاخه‌های وسیع و پرکاربرد هوش مصنوعی است که از طریق آن با استفاده از تعداد زیادی از داده‌های[3] جمع‌آوری‌شده و پردازش آنها در قالب الگوریتم‌هایی[4] شبیه به مغز انسان، رایانه توانایی اتخاذ تصمیم‌های منطقی، مانند آنچه را که انسان به‌طور طبیعی انجام می‌دهد، به‌دست می‌آورد (مازاریان، ۱۳۹۸: ۱۶۵). نوع تکامل‌یافتهٔ یادگیری ماشین را در اصطلاح یادگیری عمیق[5] می‌نامند که از ساختار و عملکرد مغز انسان یعنی اتصال سلول‌های عصبی

1. Artificial Intelligence
2. Machine Learning
3. Data
۴. الگوریتم مجموعه‌ای از دستورالعمل‌هاست که به رایانه می‌گوید چگونه مجموعه‌ای از داده‌های مربوط به جهان را به اطلاعات مفید تبدیل کند. به‌عبارت دیگر، الگوریتم فرایندی است که رایانه از آن برای تبدیل داده‌های ورودی به داده‌های خروجی استفاده می‌کند.
5. Deep Learning

الهام گرفته شده است، به‌نحوی که از یک شبکهٔ عصبی قابل برنامه‌ریزی استفاده می‌کند تا بدون کمک گرفتن از هوش بشری قادر به اتخاذ تصمیم‌های دقیق باشد. در مقایسه با یادگیری ماشین، الگوریتم‌های یادگیری عمیق به حجم بیشتری از داده‌ها نیاز دارند، تا بتوانند ماهیت ارتباط داده‌ها با یکدیگر را درک کنند و امکان حل مسائل پیچیده‌تری را برای رایانه فراهم آورند (Hassani, 2020: 144).

۲. ۲. تروریسم سایبری[1]

اصطلاح تروریسم سایبری که اولین بار در دههٔ ۱۹۸۰ میلادی توسط بری کالین[2] کارشناس سازمان امنیت ایالات متحدهٔ آمریکا، از تلفیق دو عبارت سایبر و تروریسم، ابداع شد، به یکی از اشکال نوین تروریسم بین‌المللی اشاره دارد. مارک پُلیت[3] تروریسم سایبری را حملهٔ عامدانه و مخرب با اهداف سیاسی، علیه مدیریت سامانه‌های اطلاعاتی تعریف می‌کند (Veerasamy, 2009: 10). به‌عبارت دیگر، تروریسم سایبری را می‌توان اختلال از طریق دسترسی غیرمجاز به رایانه، انتشار ویروس، بمب‌های ایمیلی و غیره با هدف آسیب رساندن، تعریف کرد (قاسمی و باقرزاده، ۱۳۹۴: ۲۲۸).

تروریسم سایبری، سعی دارد با ایجاد اختلال یا خرابکاری در سامانه‌ها، شبکه‌ها یا مؤلفه‌های اطلاعاتی رایانه‌ای و ایجاد رعب و وحشت در جوامع و همچنین، ایجاد صدمه و خسارت فیزیکی یا اطلاعاتی به اموال و اشخاص و در نهایت، تضعیف دولت یا نهادهای دولتی، در جهت نیل به اهداف سیاسی، اجتماعی و ایدئولوژیکی موردنظر خود گام بردارد (سلیمی، ۱۳۹۸: ۲۲۴). برای مثال می‌توان به حملات تروریستی ۲۰۰۷، به شبکهٔ بانکی کشور استونی و حملات ۲۰۱۷ به شبکهٔ برق‌رسانی اوکراین اشاره کرد (Shackelford, 2009: 193).

۲. ۳. پیش‌نویس قانون اتحادیهٔ اروپا در همسان‌سازی قوانین حاکم بر هوش مصنوعی

کمیسیون اروپایی در آوریل ۲۰۲۱، پیش‌نویس قانونی را با عنوان همسان‌سازی قوانین حاکم بر هوش مصنوعی[4] ارائه کرد. در صورت تصویب این پیش‌نویس توسط پارلمان و شورای اروپایی، مقررات آن در چارچوب نظام حقوقی کشورهای عضو اتحادیه لازم‌الاجرا خواهد شد، مقرراتی که همزمان طراحان،

1. Cyber terrorism
2. Barry Callin
3. Mark Pollitt
4. European Commission Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonized Rules on Artificial Intelligence Act (AIA).

تولیدکنندگان، عرضه‌کنندگان و کاربران هوش مصنوعی را حتی فراتر از قارهٔ اروپا تحت پوشش خود قرار می‌دهد. این طرح که در قالب ۱۲ بخش و ۸۵ ماده آماده شده[۱]، به‌عنوان اولین ابتکار جهانی در خصوص قانونمندسازی فناوری هوش مصنوعی معرفی شده و قرار است در نیمهٔ دوم سال ۲۰۲۲ میلادی به تصویب برسد، در صورت تصویب، مقررات آن در چارچوب نظام حقوقی ۲۷ کشور عضو اتحادیه لازم‌الاجرا می‌شود و با توجه به مفاد بند ۱ مادهٔ ۲ پیش‌نویس، به عرضه‌کنندگان و ارائه‌دهندگان سیستم‌های هوش مصنوعی به اتحادیهٔ اروپا، که در کشورهای ثالث (خارج از قلمرو اتحادیه)، مستقرند نیز، تسری می‌یابد. بر اساس مادهٔ ۱، هدف اصلی از ارائهٔ پیش‌نویس، برقراری نظام حقوقی متحدالشکل و جامع در جهت نظارت بر سازوکارهای مرتبط با طراحی، تولید، ارائه و استفاده از سیستم‌های هوش مصنوعی، بیان شده است. به‌نحوی که ضمن استفادهٔ هرچه بیشتر از قابلیت‌های هوش مصنوعی، با در نظر گرفتن چالش‌های موجود، احترام به اصول بنیادین حقوق بشر و اصول اخلاقی به‌عنوان محور اصلی تحقیقات و به‌طور کلی، تمام فعالیت‌های مرتبط با این حوزه از فناوری، در نظر گرفته شود. با تصویب این قانون، اتحادیهٔ اروپا قصد دارد از توسعهٔ هوش مصنوعی قابل اعتماد و اخلاق‌مدار حمایت کند و آن‌طور که شورای اروپا مدعی است، رهبری جهان را در این زمینه به‌دست گیرد.

اما اینکه آیا در عمل این اتفاق خواهد افتاد یا نه، نیازمند بررسی و مداقهٔ بیشتر در متن پیش‌نویس ارائه‌شده از سوی کمیسیون اروپایی است. البته پیش از بررسی متن پیش‌نویس، با توجه به استقبال دولتمردان از قابلیت‌های هوش مصنوعی به‌منظور حل معضلات جامعهٔ بشری، از جمله مقابله با فعالیت‌های تروریستی، بهتر دیده شد ضمن بیان نقاط قوت هوش مصنوعی، نقاط ضعف و مشکلاتی که می‌تواند در این زمینه گریبانگیر جوامع شود، نیز مطرح شود، به‌نظر می‌رسد بیان این قبیل مشکلات می‌تواند مبنای مناسبی به‌منظور بررسی و نقد پیش‌نویس از جنبه‌های مختلف، به‌خصوص بحث مقابله با تروریسم باشد.

۳. قابلیت‌های هوش مصنوعی در مقابله با تروریسم سایبری

تهدیدات بالقوهٔ گروه‌های تروریستی و همچنین آشنایی گستردهٔ این قبیل گروه‌ها با فضای مجازی، فعالان حوزهٔ هوش مصنوعی را بر آن داشت تا با توسل به هوش مصنوعی به‌منظور مبارزه علیه تروریسم، تلاش کنند. موضوع مهمی که دبیر کل سازمان ملل متحد، آنتونی گوترش[۲] نیز در چارچوب استراتژی دبیر کل برای

---

۱. کمیسیون اروپایی به‌منظور تبیین بهتر مفاد پیش‌نویس، یادداشتی توضیحی را در قالب ۸۹ بند قبل از مقدمهٔ پیش‌نویس ارائه داد ه است که می‌تواند به‌عنوان ابزار کمکی در تفسیر مقررات پیش‌نویس مورد توجه قرار گیرد.

2. António Guterres

فناوری‌های جدید[1] به آن اشاره کرد: «اگر هوش مصنوعی به‌درستی کنترل شود و ارزش‌ها و تعهدات تعیین‌شده در منشور ملل متحد و اعلامیهٔ جهانی حقوق بشر را ارج نهد، می‌تواند نقش مؤثری در تحقق توسعهٔ پایدار، از طریق پایان دادن به فقر، حفاظت از سیارهٔ زمین و تضمین صلح و رفاه برای همگان، داشته باشد». به همین نحو، هوش مصنوعی می‌تواند ابزاری قدرتمند در مبارزه با تروریسم سایبری باشد.

۳. ۱. فراهم کردن امکان پیش‌بینی اقدامات تروریستی

سازمان‌های امنیتی با به‌کارگیری قابلیت‌های هوش مصنوعی، می‌توانند از طریق پیش‌بینی حملات تروریستی و اتخاذ اقدامات پیشگیرانه یا هشدار به مراجع ذی‌صلاح، مانع از بروز خسارت‌های جبران‌ناپذیر ناشی از این اقدامات شوند. برای مثال می‌توان به یکی از مدل‌های یادگیری ماشین، به نام استون[2] اشاره کرد. این الگوریتم به‌نحوی طراحی شده است که می‌تواند پیش‌بینی کند در صورت بازداشت هریک از اعضای شبکهٔ تروریستی کدام عضو جانشین او خواهد شد و همچنین، شبکه چگونه خود را بازسازی و احیا خواهد کرد (Karabiyik, 2016: 56). همچنین می‌توان به اینسیکت اینتلیجنس[3] به‌عنوان یک شرکت نوپای فعال در حوزهٔ هوش مصنوعی اشاره کرد که با توسل به مدل‌های مختلف یادگیری ماشین، قابلیت پیش‌بینی تهدیدات آنلاین در رسانه‌های اجتماعی، از جمله فعالیت‌های تروریستی را به‌دست آورده است.

۳. ۲. شناسایی افراد آسیب‌پذیر در برابر افراط‌گرایی

مجریان قانون و سازمان‌های ضد تروریسم با توسل به تکنیک‌های یادگیری ماشین می‌توانند با شناسایی کلمات و عبارات کلیدی که اغلب توسط افراط‌گرایان تروریست استفاده می‌شود، احتمال گرایش آنها به سمت اقدامات تروریستی را بررسی و پیش‌بینی کنند. برای مثال می‌توان به پروژهٔ ۲۰۲۰ اتحادیهٔ اروپا با عنوان سیستم تشخیص و هشدار زودهنگام محتوای تروریستی برخط (هشدار قرمز)[4]، اشاره کرد. این پروژه از جمله ابزارهایی است که هدف اصلی آن شناسایی و تشخیص گرایش به افراط‌گرایی در مراحل اولیه و مقدماتی است.[5]

---

1. Secretary-General's Strategy on New Technologies
2. STONE
3. INSIKT Intelligence
4. Early Detection and Alert System for Online Terrorist Content (RED-Alert).
۵. در این زمینه می‌توان به سیستم نظارت دولت آلمان بر فعالیت‌های افراط‌گرایی با عنوان اختصاری (MOTRA) اشاره کرد، که در واقع ابزار نظارتی جامعی برای تجزیه‌وتحلیل داده‌های جمع‌آوری‌شده به‌منظور ارزیابی تحولات اجتماعی رخ‌داده در جوامع با قابلیت بالای پذیرش افراط‌گرایی است.

۳. ۳. شناسایی و تشخیص اطلاعات جعلی و نادرست منتشرشده توسط تروریست‌ها در فضای مجازی

جعل یا تحریف اطلاعات، این قابلیت را دارد که در گسترش رفتارهای خشونت‌آمیز نقش داشته باشد. شناسایی حساب‌های کاربری جعلی می‌تواند گام مؤثری برای مقابله با انتشار اطلاعات نادرست باشد. سازمان اطلاعات و امنیت بریتانیا[1] با به‌کارگیری قابلیت‌های هوش مصنوعی توانسته است حساب‌های جعلی را شناسایی کند (Smith, 2021:1). به همین ترتیب در سریلانکا، از یادگیری ماشین برای شناسایی و تشخیص اطلاعات نادرست استفاده کرده‌اند.

۳. ۴. کنترل و حذف خودکار محتوا

یکی از رایج‌ترین اقداماتی که توسط رسانه‌های اجتماعی با هدف مقابله با تروریست‌ها و افراط‌گرایان در فضای مجازی انجام می‌گیرد، توسل به حذف محتویات توهین‌آمیز و تهاجمی[2] است. به این منظور، عقاید توهین‌آمیز و تهاجمی حذف شده و وب‌سایت‌های اعلام‌کنندهٔ این نظرها نیز مسدود می‌شوند. برای مثال فیس‌بوک برای اولویت‌بندی محتوا به یادگیری ماشین متوسل شده است، بدین‌ترتیب که پست‌هایی که خط‌مشی‌های شرکت را نقض می‌کنند، به‌وسیلهٔ فیلترهای یادگیری ماشین علامت‌گذاری می‌شوند (Saltman, 2020: 2). وزارت کشور بریتانیا نیز در سال ۲۰۱۸، اعلام کرد که به‌دنبال توسعهٔ فناوری جدیدی است که بتواند با بهره‌گیری از یادگیری ماشین، محتوای صوتی و تصویری موجود در فضای مجازی را با هدف بررسی تبلیغات گروه‌های تروریستی مثل داعش تجزیه‌وتحلیل کند.[3]

۳. ۵. افزایش توانایی سازمان‌های ضدتروریست در پردازش اطلاعات مرتبط با گروه‌های تروریستی

تجزیه‌وتحلیل داده‌های جمع‌آوری‌شده در خصوص حملات تروریستی نقش مهمی در زمینهٔ شناسایی ارتباطات و نحوهٔ برنامه‌ریزی عملیات تروریستی دارد. اما در اغلب مواقع با توجه به حجم و سرعت رو به رشد داده‌های جمع‌آوری‌شده، این اقدام به‌سختی میسر است. به‌نحوی که تحلیل دقیق آنها گاه امکان‌پذیر نیست. اما هوش مصنوعی می‌تواند در این خصوص نقش مهمی در افزایش توانایی مقامات دولتی و محلی به‌منظور پردازش مقادیر زیادی از داده‌ها به شیوه‌ای مؤثر و با کمترین میزان منابع انسانی و مالی، ایفا کند.[4]

---

1. British Intelligence and Security Organization
2. deplatforming
3. Home Office & The Rt Hon Amber Rudd. (Feb. 13, 2018). New Technology revealed to help fight terrorist content online. GOV.UK. Accessible at https://www.gov.uk/government/news/new-technology-revealed-to-help-fight-terrorist-content-online.
4. Cisco Annual Internet Report (2018–2023) White Paper, Accessible at

۴. هوش مصنوعی و چالش‌های پیش‌رو در فرایند مقابله با تروریسم سایبری

قابلیت‌های هوش مصنوعی برای مقابله با تهدیدهای تروریسم سایبری، قابل چشم‌پوشی نیست، اما در کنار این جذابیت‌ها، توسل به هوش مصنوعی بی‌شک با خود چالش‌هایی نیز به‌همراه خواهد داشت.

۴. ۱. چالش‌های فنی

۴. ۱. ۱. تشخیص نادرست

با توسل به هوش مصنوعی، محققان توانایی شناسایی مظنونان به اقدامات تروریستی را کسب کرده‌اند، اما با توجه به احتمال وجود خطا در نتایج، این امکان وجود دارد که اشخاص بی‌گناه به‌اشتباه تروریست معرفی شوند و برعکس افراد تروریست، از گردونهٔ تشخیص به‌راحتی فرار کنند. به‌منظور کاهش این قبیل اشتباهات، طراحان هوش مصنوعی سعی کرده‌اند الگوریتم‌ها را به نحوی تعریف کنند که سطح تشخیص سیستم را در جهت شناسایی هرچه بیشتر مظنونان ارتقا دهند، در نگاه اول، این رویکرد به نظر مناسب می‌رسد، اما می‌تواند به اعمال رفتارهای تبعیض‌آمیز علیه افراد بی‌گناه منجر شده و با حداقل معیارهای تعریف‌شده از سوی سیستم، تعداد بیشتری از اشخاص به‌عنوان تروریست معرفی شوند (Macdonald, 2019: 190).

۴. ۱. ۲. وجود سوگیری در داده‌ها

رعایت انصاف و عدالت از اصول کلیدی در فرایند استفاده از فناوری هوش مصنوعی است. اما سوگیری در جمع‌آوری داده‌ها و طراحی آنها، به طرح نگرانی‌هایی در زمینهٔ توسل به هوش مصنوعی منجر شده است. برای مثال طبق تحقیقات یک مؤسسهٔ مستقل حقوقی مشخص شد که الگوریتم‌های مورد استفاده توسط قضات تصمیم‌گیرنده در حوزهٔ آزادی مشروط در ایالات متحدهٔ آمریکا، به ضرر گروه‌های نژادی آفریقایی‌تبار، سوگیری داشته و احتمال ارتکاب مجدد جنایت از سوی آنها را بسیار بیشتر در نظر گرفته است (Angwin, 2016: 1). در واقع الگوریتم‌های هوش مصنوعی با توجه به ماهیت داده‌های جمع‌آوری‌شده و دستکاری مغرضانهٔ آنها، می‌تواند به‌نوعی نمایانگر و مجری تعصبات موجود در جوامع به ضرر گروه یا طبقهٔ خاصی از افراد باشد.

---

https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf.

۴. ۱. ۳. ابهام و عدم شفافیت در نحوهٔ تصمیم‌گیری

توضیح‌پذیری و شفافیت، از جمله اصول اساسی و مهم در فرایند استفادهٔ مسئولانه از هوش مصنوعی است. به‌عبارت دیگر، تصمیمات اتخاذشده بر اساس الگوریتم‌های از قبل طراحی‌شده، باید توسط کاربران نهایی قابل درک باشد. اما در عمل سیستم‌های مبتنی بر هوش مصنوعی، مانند یک جعبهٔ سیاه، طراحی شده و به‌هیچ‌وجه نحوهٔ ایجاد خروجی با توجه به داده‌های ورودی برای کاربران انسانی قابل درک نیست. به‌عبارت دیگر، این رویه امکان ردیابی و بررسی نتایج و علت اتخاذ تصمیمات، به‌ویژه زمانی‌که از این فناوری در محاکمهٔ اشخاص استفاده می‌شود را با مشکل مواجه می‌سازد (Kaur, 2020: 3).

۴. ۱. ۴. نقش شایان توجه بخش خصوصی در فرایندهای مرتبط با طراحی و توسعهٔ هوش مصنوعی

نقش بخش خصوصی در مقایسه با ارگان‌های دولتی، به‌دلیل گسترده‌تر بودن منابع و امکانات در دسترس آن، در فرایند مقابله با تروریسم سایبری رو به افزایش است. در نتیجه، نهادهای دولتی مجبور می‌شوند برای تکمیل تلاش‌های خود، از بخش خصوصی کمک بخواهند. اما ماهیت روابط میان بخش خصوصی و دولتی در زمینهٔ مقابله با تروریسم، به‌دلیل تفاوت در اهداف و نوع فعالیت‌ها، موضوع چندان ساده‌ای نیست. برای مثال حذف شواهد مرتبط با اقدامات تروریستی توسط شرکت‌های خصوصی فعال در حوزهٔ هوش مصنوعی (Macdonald, 2019: 190)، عدم بایگانی صحیح مدارک و نیاز به حکم دادگاه برای دریافت اسناد و شواهد، به بروز شک و تردید و عدم اطمینان در میان مجریان قانون در خصوص امکان توسل به آنها در خلال فرایند تعقیب قانونی تروریست‌های سایبری منجر شده است. نیاز به تخصص فنی لازم برای توسعه، نگهداری و به‌کارگیری برنامه‌های کاربردی هوش مصنوعی نیز چالش بزرگی است که افزایش جذب کارشناسان مجرب هوش مصنوعی از سوی شرکت‌های بزرگ غیردولتی، به آن دامن می‌زند.

۴. ۲. چالش‌های حقوقی

توسل به هوش مصنوعی در راستای مقابله با تروریسم سایبری، با خود نگرانی‌های جدی از منظر امکان نقض حقوق بنیادین بشری به‌همراه داشته است. در همین زمینه مجمع عمومی سازمان ملل متحد در قطعنامهٔ ۶۲/۱۵۹ با عنوان حمایت از حقوق بشر و آزادی‌های اساسی (A/RES/62/159, 2007)، ضمن تأکید بر تعهدات حقوق بشری دولت‌ها از آنها خواسته است تا در فرایند مقابله با تروریسم، مانع از سوءاستفادهٔ بازیگران غیردولتی شده و اجازه ندهند به بهانهٔ مقابله با تروریسم، اصول بنیادین حقوق بشری نادیده گرفته شود. همچنین شورای حقوق بشر در قطعنامهٔ ترویج، حمایت و برخورداری از حقوق

بشر در اینترنت، بر اهمیت احترام به اصول بنیادین حقوق بشر در فضای مجازی مانند دنیای واقعی تأکید دارد (A/HRC/20/L.13, 2012). به همین دلیل به‌درستی ادعا می‌شود که مجریان قانون در فرایند مقابله با تروریسم، باید اطمینان حاصل کنند که هوش مصنوعی را به ابزاری برای سوءاستفاده و اتخاذ اقدامات تبعیض‌آمیز تبدیل نکنند. به‌دلیل ویژگی‌های ذاتی هوش مصنوعی، در عمل دامنهٔ گسترده‌ای از اصول بنیادین حقوق بشری، می‌تواند تحت تأثیر استفادهٔ نادرست از این فناوری قرار گیرد، هرچند حق بر حریم خصوصی، برابری و منع تبعیض با توجه به ماهیت خاصی که دارند، می‌توانند بیشترین میزان آسیب را متحمل شوند (Mckendrick, 2019: 3).

۴. ۲. ۱. نقض حق بر حریم خصوصی

این حق در چارچوب مادهٔ۲۰ اعلامیهٔ جهانی حقوق بشر، مواد ۹ و ۱۷ میثاق بین‌المللی حقوق مدنی و سیاسی و همچنین در مادهٔ ۸ کنوانسیون اروپایی حقوق بشر و مقررات اتحادیهٔ اروپا در حمایت از داده‌ها[1] نیز مورد تأکید قرار گرفته است. این مقررات که در ۱۲ آوریل ۲۰۱۶ به تصویب پارلمان اروپا رسید و از ۲۵ می ۲۰۱۸ در چارچوب نظام حقوقی تمام کشورهای عضو اتحادیهٔ اروپا لازم‌الاجرا شد، به‌دنبال ایجاد یک نظام حقوقی هماهنگ در سرتاسر اتحادیه، به‌منظور حمایت از حریم خصوصی اشخاص و ایجاد تحول در نحوهٔ تعامل سازمان‌ها و شرکت‌ها با اطلاعات و داده‌های شخصی کاربران است، به‌نحوی که پردازش اطلاعات شخصی، فقط در چارچوب مقررات حاضر یا با اجازهٔ صاحب داده‌ها میسر شود. با توجه به اهمیت حفاظت از حریم خصوصی افراد، با رویکردی متفاوت از قبل، مواد ۱۲ تا ۲۳ مقررات، به شناسایی و اجرای این حق، اختصاص یافته است.

حمایت از حریم خصوصی به‌عنوان یک حق اساسی، امری ضروری برای حمایت از کرامت بشری است. موضوعی که لزوم احترام به اطلاعات شخصی افراد را به‌عنوان یک حق بنیادین مطرح می‌سازد. همان‌طورکه گزارشگر ویژهٔ ترویج و حمایت از آزادی بیان و عقیده، فرانک لارو[2]، بیان می‌کند: «حق داشتن حریم خصوصی، به مفهوم بهره‌مندی از محدودهٔ شخصی و خصوصی[3] است، محدوده‌ای که هر فرد در آن از خودمختاری، آزادی و تعامل، بدون مداخلهٔ دولت و سایر اشخاص، برخوردار بوده و شخصاً در مورد فاش شدن اطلاعات شخصی و نحوهٔ استفاده از آنها تصمیم می‌گیرد» (A/HRC/23/40, 2013).

1. General Data Protection Regulation (GDPR)
2. Frank La Rue
3. Private Sphere

از سوی دیگر، سیستم‌های هوش مصنوعی اغلب از طریق دسترسی، تجزیه‌وتحلیل داده‌های از قبل جمع‌آوری‌شده طراحی می‌شوند؛ اقدامی که در تقابل با لزوم حفاظت از حریم خصوصی و داده‌های شخصی قرار دارد. تجزیه‌وتحلیل داده‌ها با استفاده از سیستم‌های هوش مصنوعی ممکن است به فاش شدن اطلاعات خصوصی افراد منجر شود. برای مثال مدل‌هایی از یادگیری ماشین، طراحی شده‌اند که می‌توانند با توجه به موقعیت مکانی فرد از طریق تلفن همراه، سن، جنسیت، شغل و حتی وضعیت تأهل وی را مشخص کنند (Bellovin, 2014: 555). به همین دلیل، مجمع عمومی سازمان ملل متحد با صدور قطعنامه‌ای با عنوان حق بر حریم خصوصی در عصر دیجیتال، جمع‌آوری داده‌ها و اطلاعات شخصی را به‌عنوان عملی غیرقانونی و خودسرانه توصیف و این نوع اقدامات مداخله‌جویانه را نقض حریم خصوصی انسان‌ها و خلاف اصول دموکراتیک معرفی کرده است (A/RES/68/167, 2013).

۴. ۲. ۲. نقض قاعدهٔ منع تبعیض

در کنار نقض حریم خصوصی اشخاص با توسل به هوش مصنوعی، نباید انجام رفتارهای متعصبانه به‌دلیل وجود سوگیری در الگوریتم‌های هوش مصنوعی و نقض قاعدهٔ ممنوعیت تبعیض را نادیده گرفت. قاعده‌ای که به‌دنبال تضمین ممنوعیت رفتارهای تبعیض‌آمیز بر اساس ویژگی‌های نژادی، جنسیتی، قومیتی و مذهبی است. در نتیجه هدف قرار دادن افراد از جمله گروه‌های آسیب‌پذیر جامعه در خلال عملیات هوشمند مقابله با تروریسم، صرفاً بر مبنای ویژگی‌های فوق، اقدامی تبعیض‌آمیز و غیرمنصفانه بوده و ناقض اصول بنیادین حقوق بشری است.

طبق مادهٔ ۲۶ میثاق بین‌المللی حقوق مدنی و سیاسی، تمامی اشخاص در مقابل قانون متساوی هستند و بدون هیچ‌گونه تبعیض استحقاق دریافت حمایت بالسویه از سوی قانون را دارند. از این لحاظ قانون باید ضمن ممنوعیت اقدامات تبعیض‌آمیز، حمایت از تمامی اشخاص، فارغ از نژاد، رنگ، جنسیت، مذهب، عقیده را تضمین کند. در قالب مادهٔ ۲۷ نیز، بر تعهد دولت‌ها به حمایت از اقلیت‌های نژادی، مذهبی یا زبانی و خودداری از هرگونه رفتار تبعیض‌آمیز علیه آنها تأکید شده است. اما با توجه به عملکرد سیستم‌های هوش مصنوعی، نگرانی‌های زیادی در این زمینه مطرح شده است. در عمل، مدل‌های هوش مصنوعی با هدف طبقه‌بندی و جداسازی افراد، طراحی و تولید شده‌اند؛ موضوعی که می‌تواند هنگام مواجهه با طبقات مختلف اجتماع، نتایج تبعیض‌آمیزی را به‌همراه داشته باشد (Land, 2017: 232). برای مثال می‌توان به سیستم‌های هوشمند فعال در دادگاه‌های ایالات متحدهٔ آمریکا اشاره کرد که برای تسهیل اتخاذ تصمیم، به‌منظور صدور قرار وثیقه یا صدور حکم بازداشت متهمین، به قضات آمریکایی

کمک می‌کنند. پس از بررسی نتایج حاصل از تصمیمات اتخاذشده با توسل به این سیستم، مشخص شد که متهمان سیاه‌پوست در مقایسه با سفیدپوستان، به نسبت بیشتری در طبقهٔ متهمان پرخطر دسته‌بندی شده‌اند، در نتیجه مدت بازداشت آنها و مبلغ وثیقهٔ تعیین‌شده برای آزادی موقتشان، به‌طور چشمگیری افزایش یافته است (Flores, 2018: 9).

شبیه به این نوع رفتار در نرم‌افزارهای تشخیص چهره نیز رؤیت شده است، به‌طوری‌که تشخیص چهرهٔ افراد با پوست تیره نرخ بالاتری از خطا را در مقایسه با سفیدپوستان به خود اختصاص داده است. شناسایی نادرست می‌تواند عواقب ناخوشایندی را برای رنگین‌پوستان با خود به‌همراه داشته باشد. در سال ۲۰۱۸، به‌منظور بررسی ابعاد مختلف مشکلات ناشی از تشخیص نادرست، نرم‌افزار تشخیص چهرهٔ آمازون، توسط اتحادیهٔ آزادی‌های مدنی آمریکا[1] بررسی شد. به این منظور، ابتدا چهرهٔ تمام ۵۳۵ عضو کنگرهٔ ایالات متحده اسکن شد، اطلاعات هیچ‌یک از نمایندگان در پایگاه دادهٔ مورد استفاده در طراح سیستم تشخیص چهرهٔ آمازون وجود نداشت، با این حال، ۲۸ مورد سابقهٔ غیرواقعی در سیستم تشخیص چهره از نمایندگان، یافت شد. که ۵۸ درصد از این موارد به نمایندگان رنگین‌پوست مربوط می‌شد (Brandom, 2018: 3). مواردی از این نوع، گواهی بر این ادعاست که سیستم‌های تشخیص چهرهٔ آمازون نیز مانند سایر سیستم‌ها، نسبت به رنگین‌پوستان، سوگیری دارند.

۵. ایرادات وارد بر مقررات پیش‌نویس همسان‌سازی قوانین حاکم بر هوش مصنوعی

وجود چالش‌های حقوقی و فنی علیه هوش مصنوعی و همچنین فقدان نظام حقوقی جامع و متحدالشکل در زمینهٔ نظارت بر فرایند طراحی، تولید، عرضه و استفاده از هوش مصنوعی، بدون شک، توسل به این فناوری نوین را در جنبه‌های مختلف زندگی بشری با مشکلات عدیده‌ای مواجه خواهد ساخت (Fishman, 2019: 88)؛ موضوعی که از نگاه اتحادیهٔ اروپا دور نماند و به اقدام در جهت تدوین و تصویب یک سند حقوقی، با هدف نظم‌دهی و کنترل فرایندهای مرتبط با سیستم‌های هوش مصنوعی، منجر شد. در این زمینه کمیسیون اروپایی در آوریل ۲۰۲۱، پیش‌نویس قانونی را با عنوان همسان‌سازی قوانین حاکم بر هوش مصنوعی، به‌عنوان پاسخی به نگرانی‌های فعالان حقوق بشر و مهر تأییدی بر تعهدات بین‌المللی دولت‌ها در راستای تأمین و اجرای حقوق بنیادین بشر ارائه کرد. اما آیا تهیه‌کنندگان پیش‌نویس، در اجرای رسالت خود موفق بوده‌اند؟ آیا توانسته‌اند به‌خوبی تعهدات حقوق بشری بازیگران

1. American Civil Liberties Union

اصلی دنیای هوش مصنوعی را منعکس نمایند؟ بررسی و مطالعهٔ متن پیش‌نویس نشان می‌دهد که به‌رغم وجود نقاط قوت و توجه به جنبه‌های مختلف مسائل مرتبط با سیستم‌های هوش مصنوعی، از جمله توجه به دو موضوع مهم، یعنی لزوم حفاظت از داده‌ها و اطلاعات شخصی، به‌منظور ارج نهادن به حق بنیادین اشخاص در بهره‌مندی از حریم خصوصی[۱] و همچنین خودداری از اعمال رفتارهای تبعیض‌آمیز در فرایند توسل به هوش مصنوعی[۲]، هنوز در برخی از بخش‌ها، کاستی‌ها و ایراداتی وجود دارد، که در ادامه به آنها پرداخته خواهد شد. در این زمینه ابتدا ایرادات کلی موجود در متن پیش‌نویس با توجه به چالش‌های مطرح‌شده در بخش‌های پیشین، بیان می‌شود، سپس به‌صورت اختصاصی، موضوع مقابله با اقدامات تروریستی با استفاده از هوش مصنوعی در پرتو مقررات حاضر، بررسی خواهد شد.

۵. ۱. ایرادات کلی

۵. ۱. ۱. توجه ناکافی به اصول اخلاقی

با وجود تأکید بر لزوم رعایت اصول اخلاقی در سایر اسناد مرتبط با نظام‌مند کردن هوش مصنوعی که قبل از پیش‌نویس ارائه شده‌اند[۳] و البته تأکید کمیتهٔ اقتصادی، اجتماعی اتحادیهٔ اروپا مبنی بر لزوم تمرکز

---

۱. مطابق بند ۱ از مادهٔ ۵۴ با عنوان پردازش اطلاعات شخصی به‌منظور توسعهٔ سیستم‌های هوش مصنوعی در جهت رفاه عمومی: «داده‌ها و اطلاعات شخصی که طبق قانون به‌منظور توسعه و ایجاد سیستم‌های هوش مصنوعی جدید با کمک به مقامات ذی‌صلاح در جهت پیشگیری، تحقیق، کشف یا تعقیب مجرمان، مقابله با تهدیدات امنیتی، پیشگیری و کنترل بیماری‌ها، حفاظت از محیط‌زیست و غیره..... جمع‌آوری شده‌اند، باید طبق یک سازوکار قانونی، فقط در دسترس اشخاص معین و قابل اعتماد بوده، به‌صورت ایزوله و در محیطی امن مورد پردازش قرار گرفته و بلافاصله بعد از پایان طراحی، حذف گردند. انتقال داده‌ها صرفاً با مجوز قانونی و بدون ایجاد هرگونه خدشه به حریم خصوصی اشخاص انجام گردد و هر زمان در طول پردازش مشخص شد که احتمال نقض حقوق بنیادین صاحبان اطلاعات وجود دارد، بلافاصله باید عملیات پردازش متوقف گردد».

۲. در قالب بند دوم از بخش اول یادداشت‌های توضیحی کمیسیون اروپا آمده است: «پیش‌نویس حاضر بیانگر الزامات خاصی است که هدفشان به حداقل رساندن خطر تبعیض الگوریتمی است، به‌ویژه در رابطه با طراحی مجموعه داده‌های مورد استفاده برای توسعهٔ سیستم‌های هوش مصنوعی، مدیریت خطر، مستندسازی و نظارت انسانی در سراسر چرخهٔ تولید سیستم‌های هوش مصنوعی ». مؤید این ادعا، بند ۱ مادهٔ ۱۱پیش‌نویس است، که از لزوم ارائهٔ مستندات فنی قبل از عرضه و راه‌اندازی سیستم‌های پرخطر هوش مصنوعی در بازار صحبت می‌کند. مستنداتی که طبق بند ۳ ضمیمهٔ شمارهٔ ۴ باید حاوی اطلاعاتی، در خصوص نظارت بر عملکرد سیستم‌های هوش مصنوعی، به‌ویژه در خصوص قابلیت‌ها و محدودیت‌هایش در عملکرد مناسب و پیامدهای ناخواستهٔ قابل پیش‌بینی و عواقب ناخوشایندش برای حقوق اساسی انسان و اعمال رفتارهای تبعیض‌آمیز، با توجه به هدف موردنظر از طراحی و توسل به سیستم‌های هوش مصنوعی، باشد.

۳. برای مثال می‌توان به پیش‌نویس توصیه‌های یونسکو در مورد اخلاق هوش مصنوعی اشاره کرد.

بر اخلاق، در عمل اثری از لزوم رعایت اصول اخلاقی و تعهدات ناشی از آن، در مفاد متن ۸۵ ماده‌ای پیش‌نویس دیده نمی‌شود. در قالب یادداشت‌های توضیحی[1] که در ۸۹ بند و قبل از مقدمهٔ پیش‌نویس مطرح شده، ۱۱ مرتبه به واژهٔ اخلاق اشاره شده است، از جمله بند (۱)۱: «کمیسیون قوانینی را برای ایجاد یک رویکرد هماهنگ اروپایی در مورد پیامدهای انسانی و اخلاقی هوش مصنوعی ارائه خواهد کرد». همچنین در قسمت دیگری از این بند اشعار می‌دارد که مقررات این پیش‌نویس «از هدف اتحادیه مبنی بر رهبری جهانی در توسعهٔ هوش مصنوعی امن، قابل اعتماد و اخلاقی، پشتیبانی می‌کند و حمایت از اصول اخلاقی را چنان‌که به‌طور خاص توسط پارلمان اروپا درخواست شده است، تضمین می‌نماید». اما در متن پیش‌نویس به‌صراحت و به‌صورت مستقیم عنوان یا بخش خاصی به موضوع تعهدات اخلاقی فعالان حوزهٔ هوش مصنوعی اختصاص نیافته است. این بی‌توجهی در حالی رخ داده که اخلاق نقطهٔ ثقل بیشتر انتقادات بیان‌شده علیه هوش مصنوعی است. چراکه استفاده از سیستم‌های هوش مصنوعی بدون توجه به اصول اخلاقی، می‌تواند به تبعیض، انحصار، محرومیت، از بین رفتن تکثر فرهنگی و اقتدارگرایی فنی منجر شود و علاوه‌بر این، اتحادیهٔ اروپا در پی آن است که با تصویب این مقررات، رهبری جهان در توسعهٔ هوش مصنوعی امن و اخلاق‌مدار را در اختیار بگیرد (Likov, 2021: 166).

۵. ۱. ۲. مبهم بودن برخی از واژه‌ها و عبارت‌ها

۵. ۱. ۲. ۱. مفهوم عبارت هوش مصنوعی

طبق بند ۱ مادهٔ ۳ پیش‌نویس: «هوش مصنوعی، عبارت است از نرم‌افزاری که با توسل به روش‌ها و رویکردهای مطرح‌شده در پیوست اول متن پیش‌نویس، طراحی و توسعه یافته است و می‌تواند برای نیل به مجموعه‌ای از اهداف خاص تعیین‌شده توسط انسان، خروجی‌هایی چون محتوا، پیش‌بینی‌ها، توصیه‌ها یا تصمیماتی را تولید کند که بر محیط‌هایی که با آنها در تعامل است، تأثیر می‌گذارند». بندهای ۱ تا ۳ پیوست شمارهٔ ۱ به روش‌ها و رویکردهای موردنظر در مادهٔ ۳ پرداخته و به مواردی از جمله یادگیری ماشین، یادگیری تحت نظارت، با استفاده از طیف گسترده‌ای از روش‌ها مثل یادگیری عمیق، رویکردهای منطقی، از جمله بازنمایی دانش، برنامه‌نویسی استقرایی (منطقی)، پایگاه‌های دانش، موتورهای استنتاج و قیاس،

1. Explanatory Memorandum

استدلال (نمادین)، سیستم‌های خبره و همچنین رویکردهای آماری، تخمین بیزی[1]، اشاره کرده است. به‌زعم فعالان حوزهٔ حقوق بشر، این تعریف بیش‌ازحد فنی بوده و به احتمال زیاد، به‌سرعت منسوخ خواهد شد. چراکه در عمل نهادهای قانونی، هنگام استناد به مقررات پیش‌نویس و اجرای تعهدات ناشی از آن و رسیدگی به موارد نقض‌شده توسط فعالان و متخصصان هوش مصنوعی، با مفاهیم بسیار تخصصی مواجه خواهند شد، مسئله‌ای که می‌تواند فرایند رسیدگی به تخلفات و البته جبران خسارت‌های ناشی از آن را با مشکل مواجه سازد. معضلی که به‌طور قطع گریبانگیر اشخاصی خواهد شد که در اثر به‌کارگیری هوش مصنوعی متضرر شده و برای جبران زیان‌های وارده متوسل به طرح دعوی و دادخواهی می‌شوند.

۵. ۱. ۲. ۲. مفهوم عبارت به اندازهٔ کافی[2]

طبق مادهٔ ۱۳ پیش‌نویس، با عنوان شفافیت و ارائهٔ اطلاعات به کاربران، سیستم‌های هوش مصنوعی پرخطر باید به‌گونه‌ای طراحی و توسعه داده شوند تا اطمینان حاصل شود که عملکرد آنها به اندازهٔ کافی شفاف است، به نحوی که کاربران بتوانند خروجی سیستم را تفسیر کرده و به نحو مناسب از آن استفاده کنند. این عبارت در پاراگراف ۴۴ از یادداشت‌های توضیحی کمیسیون نیز استفاده شده است: «فرایند طراحی، ارزیابی و بررسی داده‌ها با توجه به هدف موردنظر از توسل به سیستم هوش مصنوعی، باید به اندازهٔ کافی، شفاف، عاری از خطا و کامل باشد». فارغ از اینکه چنین انتظاراتی، علی‌رغم مطلوب بودن، بیش‌ازحد ایده‌آل به‌نظر می‌رسند، به نحوی که به‌زعم برخی از محققین، به‌ندرت امکان برآورده شدنشان به‌صورت کامل میسر خواهد شد، عبارت به «اندازهٔ کافی» نیز به‌نوعی مبهم به‌نظر می‌رسد. بر اساس چه معیار یا استانداردی می‌توان کافی بودن یا نبودن را بررسی و اعلام کرد که اقدامات صورت گرفته، کفایت لازم را دارند؟

۵. ۱. ۳. اتخاذ رویکرد مبتنی بر خطر به‌منظور تقسیم‌بندی سیستم‌های هوش مصنوعی (رویکردی غیرمنسجم و غیرمنعطف)

کمیسیون اروپا در فرایند تشخیص هوش مصنوعی قابل اعتماد، رویکرد مبتنی بر خطر را انتخاب کرده و در مواد ۵ تا۵۱ از بخش‌های دوم و سوم پیش‌نویس، سیستم‌های هوش مصنوعی را به چهار دسته یا

---

۱. رویکرد یا روش بیزی، یکی از شیوه‌های استنباط آماری است که امکان تجزیه‌وتحلیل داده‌ها را فراهم می‌سازد. در این روش، استنباط هم بر اساس نمونه‌های تصادفی و هم بر اساس اطلاعات پیشین پی‌ریزی می‌شود. به این ترتیب، احتمال رخداد یک فرضیه را با توجه به شواهد و اطلاعات قبلی محاسبه می‌کند و سپس تصمیم‌سازی صورت می‌گیرد.

2. Sufficiently

طبقه (غیرقابل قبول[1]، پرخطر[2]، کم‌خطر[3] و بی‌خطر[4]) تقسیم‌بندی کرده است. بر این اساس هرچه میزان خطرهای ناشی از توسل به سیستم هوش مصنوعی بیشتر باشد، طراحان، تأمین‌کنندگان، ارائه‌دهندگان و کاربران سیستم‌های هوش مصنوعی باید شرایط و تعهدات سخت‌تری را بپذیرند و رعایت کنند. دستهٔ اول در هرم خطر، سیستم‌های هوش مصنوعی با خطر غیرقابل قبول هستند (بخش ۲، مادهٔ ۵)، چراکه به‌زعم کمیسیون، استفاده از این دسته از سیستم‌های هوش مصنوعی، نقض اصول بنیادین حقوق بشر را به‌دنبال خواهد داشت، ازاین‌رو استفاده از آنها ممنوع است. از جمله سیستم‌های هوش مصنوعی که از آسیب‌پذیری کودکان یا افراد دارای معلولیت سوءاستفاده می‌کنند، مانند اسباب‌بازی‌هایی که کودک را وادار به انجام رفتارهایی می‌کند که می‌تواند آسیب‌های جسمی یا روحی برای او به دنبال داشته باشد. اما چه میزان از آسیب‌های جسمی یا روحی می‌تواند به ممنوعیت توسل به این‌گونه از سیستم‌های هوش مصنوعی منجر شود؟ همچنین با توجه به پیچیدگی الگوریتم‌های به‌کاررفته در سیستم‌های هوش مصنوعی، مثل پلتفرم‌های اشتراک‌گذاری تصاویر و فیلم‌ها، که می‌توانند کاربران را به سمت اطلاعات نادرست و افراطی سوق دهند (Amershi, 2019: 13)، آیا امکان تعیین ارتباط مستقیم میان این قبیل سیستم‌ها و آسیب‌های روحی وارده به کاربران و در نتیجه، اعمال ممنوعیت‌های پیش‌بینی‌شده در مقررات، وجود دارد؟ همچنین توسل به سیستم‌های هوش مصنوعی طراحی شده در جهت رتبه‌بندی اشخاص بر اساس رفتارهای اجتماعی و مشخصات فردی، توسط نهادهای دولتی به‌منظور بهره‌مندی از مزایای اجتماعی نیز، ممنوع شده است، این ممنوعیت، شامل استفاده از سیستم‌های احراز هویت بیومتریک[5]، از جمله دوربین‌های تشخیص چهره در فضای عمومی به‌منظور اجرای قانون نیز می‌شود، مگر آن‌که به‌منظور جست‌وجوی قربانیان جنایت، جلوگیری از تهدیدهای قریب‌الوقوع، شناسایی جنایتکاران و تروریست‌ها استفاده شود.[6] هرچند فعالان مدنی معتقدند که مفاد این بخش از مادهٔ ۵ بحث‌برانگیز است .به‌زعم آنها، استفاده از دوربین‌های هوشمند در فضای عمومی باید مشمول

1.Unacceptable Risk
2. High Risk
3. Low Risk
4. Minimal Risk

۵. Biometric Identification Real-time: از فناوری بیومتریک برای احراز هویت اشخاص استفاده می‌کند، در واقع توسل به هوش مصنوعی و بیومتریک در کنار یکدیگر می‌تواند به ایجاد مدل‌های امنیتی پویا منجر شود. برای مثال هوش مصنوعی، چهرهٔ فرد را بر اساس تصویرهای ثبت‌شده شناسایی کرده و از بیومتریک‌های سه‌بعدی برای موفقیت در احراز هویت چهرهٔ افراد استفاده می‌کند.

۶. مادهٔ ۵ بند ۱ (د).

ممنوعیت‌های گسترده‌تری شود، چراکه می‌تواند به سوءاستفاده و اعمال رفتارهای تبعیض‌آمیز علیه برخی اقشار جامعه، به‌خصوص اقشار آسیب‌پذیر منجر شود. علاوه‌بر این ممنوعیت مندرج در مادهٔ ۵ شامل مواردی که از دوربین‌های بیومتریک در جهت اهداف دیگری مثل کنترل جمعیت یا نظارت بر سلامت عمومی به‌کار می‌روند، نمی‌شود، چالش دیگری که می‌تواند، حق شهروندان برای ناشناس ماندن در ملأ عام را نقض کند (Veale, 2021: 102).

بخش سوم (مواد ۶ تا ۵۱)، به‌عنوان بخش اساسی پیش‌نویس به سیستم‌های پرخطر اختصاص یافته است. در این زمینه پیش‌نویس خواهان اعمال مقررات سختگیرانه و منسجم در جهت کنترل آثار خطرناک سیستم‌های هوش مصنوعی شده است که استفاده از آنها می‌تواند آسیب‌های جدی به سلامتی، ایمنی و حقوق بنیادین بشر وارد کند. در واقع این طبقه از سیستم‌های هوش مصنوعی می‌توانند آثار بالقوه مخربی را بر منافع شخصی افراد داشته باشند. به همین دلیل، پیش از عرضه یا استفاده باید به‌طور کامل ارزیابی شده و مشخصاتشان در پایگاه اطلاعاتی که تحت نظارت کمیسیون اروپاست، ثبت شوند. علاوه‌بر این، تعهد به ارائهٔ مستندات فنی دقیق، ثبت سوابق فعالیت و وجود سطح مناسبی از نظارت انسانی، ازجمله دیگر تعهدات لازم‌الاجرایی است که در مواد ۸ تا ۱۰ پیش‌نویس برای عرضه‌کنندگان این طبقه از سیستم‌های هوش مصنوعی پیش‌بینی شده است. از جمله سیستم‌های هوش مصنوعی طبقه‌بندی‌شده در این دسته می‌توان به نرم‌افزارهای هوشمند بررسی رزومهٔ اشخاص به‌منظور طی کردن مراحل استخدام، ربات‌های جراح، سیستم‌های هوشمند فعال در مدیریت زیرساخت‌های حیاتی مثل آب و برق، کنترل مرزها و مدیریت مهاجرت، اشاره کرد، فهرست مشخصی از این سیستم‌ها در پیوست شمارهٔ ۳ به متن پیش‌نویس اضافه شده است. بدون شک، توسل به شیوه‌های سختگیرانه در جهت کنترل و کاهش آثار سوء ناشی از هوش مصنوعی، می‌تواند گام مؤثری در جهت ترویج احترام به حقوق بنیادین بشر و حفاظت از بشریت در برابر این فناوری باشد. اما سؤالی که در اینجا مطرح می‌شود، آن است که، هزینهٔ انطباق با چنین مجموعه‌ای از تعهدات سختگیرانه چقدر خواهد بود؟ نتایج یک تحقیق انجام‌شده به درخواست کمیسیون اروپا نشان می‌دهد که هزینه تخمینی برای تبعیت از تعهدات مذکور، نزدیک به ۱۰۰۰۰ یورو خواهد بود، البته در صورت به‌کارگیری نیروی انسانی اضافی در جهت افزایش میزان نظارت انسانی، این مبلغ به ۳۰۰۰۰ یورو افزایش خواهد یافت (Renda, 2021: 125). آیا در عمل امکان تأمین مالی این مقدار هزینه برای تمامی نهادها و شرکت‌های فعال در حوزهٔ هوش مصنوعی وجود دارد؟ موضوعی که می‌تواند به‌نوعی انحصارطلبی در دنیای هوش مصنوعی، تبدیل شود. هرچند به‌طور کلی، فعالان مدنی معتقدند که اتخاذ رویکرد مبتنی بر خطر و طبقه‌بندی هوش مصنوعی بر مبنای خطرهای از

پیش تعیین‌شده، نمی‌تواند راهکار چندان موفقی باشد، چراکه در عمل، با توجه به نقش مهم نحوهٔ استفاده از هوش مصنوعی در تعیین میزان خطر، در اغلب موارد، تعیین میزان دقیق آثار سوء و عواقب منفی ناشی از آن به‌درستی ممکن نخواهد بود. علاوه‌بر این، مفهوم عبارت پرخطر به‌عنوان یکی از بخش‌های کلیدی مقررات، با ابهام همراه است (Floridi, 2021: 220). به این ترتیب که از یک طرف برخی سیستم‌های هوش مصنوعی، پرخطرند، زیرا بخش‌های حیاتی و مهم جامعه به عملکرد صحیح آنها بستگی دارد، برای مثال یک اتومبیل خودران وسیلهٔ خوبی است که اگر به‌درستی کار نکند، می‌تواند بسیار خطرناک باشد. از سوی دیگر، برخی سیستم‌های هوش مصنوعی وجود دارند که استفادهٔ غیراخلاقی از آنها خطرناک است و می‌تواند به بروز مشکلات گسترده‌ای منجر شود، مانند سوءاستفاده از سیستم‌های بیومتریک تشخیص هویت، به‌عنوان یک فناوری نظارتی، با هدف اجرای قانون که استفاده از آن در چارچوب شق «د» از بند ۱ مادهٔ ۵ پیش‌نویس ممنوع شده است. برخلاف اتومبیل خودران، این نوع سیستم، اگر مورد استفاده قرار بگیرد، و به‌درستی و بدون اشکال فنی کار کند، خطرناک است. حال اگر تفاوت بین این دو مفهوم مختلف از عبارت پرخطر تشخیص داده نشود- یک سیستم پرخطر است چون به‌درستی کار نمی‌کند، اما سیستم دیگر پرخطر است، چون به‌درستی کار می‌کند- به بروز سردرگمی منجر خواهد شد. وجود ابهام‌هایی از این دست، در مفهوم عبارت‌ها و عدم اطمینان در خصوص ماهیت ویژه و منحصربه‌فرد خطرهای موردنظر در متن پیش‌نویس، امکان انطباق فرایند طراحی و توسعهٔ هوش مصنوعی با مقررات پیش‌بینی‌شده در جهت ارزیابی سیستم‌های مزبور را تضعیف می‌کند (Floridi, 2020: 370)، موضوع حیاتی که باید از سوی تهیه‌کنندگان پیش‌نویس مدنظر قرار گیرد.

نکتهٔ دیگری که در بحث طبقه‌بندی، توجه منتقدان را به خود جلب کرده است، پیش‌بینی امکان به‌روزرسانی سیستم‌های طبقهٔ پرخطر بر اساس پیوست شمارهٔ ۳ است. طبق بند ۲ از مادهٔ ۶ پیش‌نویس: «علاوه‌بر سیستم‌های هوش مصنوعی پرخطر ذکرشده در بند ۱، سیستم‌های هوش مصنوعی مذکور در پیوست ۳ نیز پرخطر تلقی می‌شوند». به‌عبارت دیگر، سازوکاری در پیش‌نویس تدارک دیده شده است تا بتوان با گذر زمان و بررسی عملکرد سیستم‌های هوش مصنوعی، موارد جدیدتری را به فهرست پرخطرها اضافه کرد. اما جالب اینجاست که چنین سازوکاری برای سیستم‌های غیرقابل قبول (ممنوعه) که در مادهٔ ۵، ذیل بخش دوم از متن پیش‌نویس با عبارت روش‌های ممنوعهٔ استفاده از هوش مصنوعی[1] مطرح شده‌اند[2] و همچنین سیستم‌های کم‌خطر (ذیل بخش چهارم از متن پیش‌نویس با عنوان تعهد به

1. Prohibited Artificial Intelligence Practices

۲. بر اساس بند ۱ مادهٔ ۵ استفاده از هوش مصنوعی در جهت انجام اقدامات زیر ممنوع است: «استفاده از هوش مصنوعی

شفاف‌سازی (عملکرد) برخی از سیستم‌های هوش مصنوعی)[1] پیش‌بینی نشده است.[2] البته شایان ذکر است که پیوست شمارهٔ ۳ نیز به‌دلیل انعطاف‌ناپذیری و ضعف در قابلیت پاسخگویی به نیازهای آتی مورد انتقاد است، چراکه به‌صورت پیش‌فرض، هشت عنوان (شناسایی بیومتریک و طبقه‌بندی اشخاص حقیقی، مدیریت و بهره‌برداری از زیرساخت‌های حیاتی، آموزش و پرورش و حرفه‌آموزی، استخدام، مدیریت کارگران و دسترسی به خوداشتغالی، دسترسی و بهره‌مندی از خدمات ضروری خصوصی و عمومی، اجرای قانون، مدیریت مهاجرت، پناهندگی و کنترل مرز و بالاخره، اجرای عدالت) تعیین شده و اضافه کردن سیستم‌های هوش مصنوعی جدید به طبقهٔ پرخطرها، فقط در چارچوب این هشت عنوان میسر شده است.

۵. ۱. ۳. عدم پیش‌بینی حق دریافت خسارت توسط اشخاص آسیب‌دیده

در متن پیش‌نویس، صحبتی از نحوهٔ جبران خسارت‌های جسمی یا روحی وارده به افراد در اثر استفاده از سیستم‌های هوش مصنوعی نشده است. به‌عبارت دیگر، هیچ‌گونه مقرراتی که به‌صراحت سازوکارهای مرتبط با اقامهٔ دعوا توسط افراد یا جامعهٔ مدنی، علیه سیستم‌های پرخطر به‌منظور جبران آسیب‌های وارده را پیش‌بینی و طراحی کرده باشد، وجود ندارد. هرچند مادهٔ ۷ پیش‌نویس اشاره‌هایی به موضوع جبران خسارت‌های وارده توسط سیستم‌های پرخطر داشته است، اما نه صحبتی از سازوکارهای مرتبط با شیوهٔ جبران در میان است و نه به حق افراد به اقامهٔ دعوا به‌منظور دریافت غرامت توجه شده است.[3]

به‌نحوی که به ایجاد آسیب‌های روحی و جسمی علیه افراد منجر گردد و یا به سوءاستفاده یا تحریف رفتارهای افراد آسیب‌پذیر منجر گردد، یا به انجام رفتارهای نامطلوب و ناموجه علیه برخی از افراد منجر گردد و غیره...».

1. Transparency Obligations for Certain AI Systems

۲. در خصوص سیستم‌های هوش مصنوعی کم‌خطر ذکر این نکته ضروری است که در متن پیش‌نویس به‌صراحت چنین عنوانی آورده نشده است. بلکه هنگام طبقه‌بندی سیستم‌های هوش مصنوعی، از عبارت‌های ممنوعه و پرخطر در قالب بخش‌های شمارهٔ دو و سه استفاده شده و بلافاصله در بخش چهارم، تعهد به شفاف‌سازی در خصوص برخی از سیستم‌های هوش مصنوعی آورده شده است. سیستم‌هایی که استفاده از آنها در تعامل با افراد ممنوع نشده و پرخطر نیستند، اما در مواردی ممکن است مشکلاتی را برای افراد ایجاد کنند، ازاین‌رو کمیسیون از لزوم شفاف‌سازی صحبت کرده و از ارائه‌دهندگان چنین سیستم‌هایی خواسته است تا به کاربران اطلاع داده و شفاف‌سازی کنند. برای مثال اگر کاربران با ربات چت آنلاین در ارتباط‌اند، باید به آنها اطلاع داده شود که در حال صحبت با یک ربات هستند تا بتوانند آگاهانه تصمیم بگیرند که آیا ادامه بدهند یا خیر.

۳. بر اساس شق هشتم از بند ۲ مادهٔ ۷: «هنگام بررسی احتمال بروز آسیب علیه سلامتی و ایمنی افراد و یا نقض حقوق بنیادین آنها در اثر استفاده از سیستم‌های هوش مصنوعی پرخطر، کمیسیون باید، موارد زیر را در نظر بگیرد: .... لزوم

شرایطی که می‌تواند با توجه به ابهام‌ها و پیچیدگی‌های موجود در فرایند طراحی و تولید سیستم‌های هوش مصنوعی و همچنین به‌دلیل فقدان توازن قدرت میان تولیدکنندگان و به‌طور کلی فعالان حوزهٔ هوش مصنوعی و افراد آسیب‌دیده، مانعی بزرگ در فرایند رسیدگی حقوقی و جبران خسارت‌های وارده محسوب شود و به‌عنوان یک انحراف جدی از مقررات قانون عمومی حمایت از داده‌ها، مورد توجه قرار گیرد، مقرراتی که مجموعه‌ای از شرایط و امکانات را در جهت اقامهٔ دعوا توسط اشخاص یا وکلایشان برای جبران خسارت‌های وارده در اثر استفاده از هوش مصنوعی پیش‌بینی کرده است. به‌زعم کارشناسان حقوق فضای مجازی، چنین رویکردی نشان‌دهندهٔ ناتوانی و ضعف مقررات پیش‌نویس در محافظت از افراد در برابر آسیب‌های ناشی از به‌کارگیری سیستم‌های هوش مصنوعی پرخطر است (Roberts, 2021: 65).

۵. ۲. عدم صراحت در خصوص استفاده از هوش مصنوعی به‌منظور مواجهه با تروریسم سایبری

همان‌طورکه گفته شد، اتحادیهٔ اروپا قصد دارد با تصویب مقررات پیش‌نویس، رهبری جهان را در فرایند توسعهٔ هوش مصنوعی امن و مطمئن بر عهده گیرد. به‌عبارت دیگر به‌دنبال آن است که ضمن توسعه، تولید، عرضه و استفاده از قابلیت‌های هوش مصنوعی، سازوکارهایی ارائه دهد که امکان مقابله با چالش‌های موجود در این زمینه را به موازات بهره‌مندی از مزایا و منافع فراهم سازد. موضوع مهمی که بدون شک در بحث مواجهه با تروریسم سایبری با استفاده از امکانات هوش مصنوعی باید مدنظر قرار گیرد. اما سؤالی که در اینجا مطرح می‌شود، این است که آیا تهیه‌کنندگان پیش‌نویس، موضوع مقابله با فعالیت‌های تروریستی با استفاده از هوش مصنوعی را مدنظر قرار داده‌اند به‌عبارت دیگر، آیا مقررات پیش‌نویس در خصوص قانونمند کردن استفاده از هوش مصنوعی برای مقابله با اقدامات تروریستی، با توجه به چالش‌هایی که می‌تواند در این زمینه ایجاد شود، قابل اجراست؟

در پاسخ به این پرسش باید گفت در برخی از بخش‌های پیش‌نویس، از جمله مادهٔ ۵ به‌صورت مختصر اشاره‌ای به فعالیت‌های تروریستی شده است.[۱] اما این وضعیت موجب نمی‌شود که مادهٔ شمارهٔ ۲ با عنوان دامنهٔ[۲] اجرای مقررات، نادیده گرفته شود. در بخشی از این ماده، به اقداماتی که مشمول مقررات

اتخاذ اقدامات مؤثر در جهت جبران خسارت‌های وارده».

۱. مادهٔ ۵ به ممنوعیت‌های مطرح‌شده در استفاده از سیستم‌های هوش مصنوعی اختصاص یافته است. از جمله ممنوعیت استفاده از دوربین‌های امنیتی بیومتریک در فضای عمومی، اما ذیل شق ۴ بند ۱ این ماده، استفاده از دوربین‌های مزبور به‌منظور پیشگیری از حمله‌های تروریستی، مجاز شناخته شده است.

2. Scope

مزبور قرار نمی‌گیرند، اشاره شده است. طبق بند ۳ مادهٔ ۲: «مقررات پیش‌نویس در مورد سیستم‌های هوش مصنوعی که منحصراً برای اهداف نظامی توسعه‌یافته یا استفاده می‌شوند، اجرا نخواهد شد». به‌عبارت دیگر، استفاده از سیستم‌های هوش مصنوعی، در جهت تأمین اهداف نظامی، از شمول این مقررات مستثنا شده است.

اما اهداف نظامی[1] در چارچوب مقررات هوش مصنوعی چه معنایی دارد؟ به‌عبارت دیگر، اهداف نظامی دقیقاً چه مواردی را شامل می‌شود، آیا مقابله با تروریسم، هدف نظامی محسوب می‌شود که اگر چنین باشد، بحث هوش مصنوعی و مواجهه با تروریسم، از شمول مقررات پیش‌نویس، خارج خواهد بود.

در این زمینه با توجه به اینکه مصادیق اهداف نظامی در مادهٔ ۲ ارائه نشده است، به‌نظر می‌رسد مراجعه به بند ۱۲ یادداشت‌های توضیحی کمیسیون اروپا، که به موضوع اهداف نظامی پرداخته است، راهگشا باشد. بر اساس بند ۱۲: «سیستم‌های هوش مصنوعی که منحصراً برای (تأمین) اهداف نظامی، توسعه یافته یا استفاده می‌شوند، باید از شمول مقررات پیش‌نویس، مستثنا گردند، چنانچه این اهداف، در چارچوب بخش پنجم[2] معاهدهٔ اتحادیهٔ اروپا[3]، به‌منظور تأمین امنیت مشترک و اجرای سیاست‌های خارجی، تعیین شده باشند». هرچند در مقررات مطرح‌شده در چارچوب بخش پنجم نیز، بدون ذکر نامی از اهداف نظامی، صرفاً به صلاحیت‌های مرتبط با تصمیم‌ها[4]، عملیات[5]، تدارکات[6]، هزینه‌ها[7] و مشاوره‌هایی[8] که پیامدهای نظامی یا دفاعی دارند، اشاره شده است. از سوی دیگر، از کاربرد «یا» در بین واژه‌های «نظامی» و «دفاعی»، ممکن است، چنین استنباط شود که پیامدهای نظامی می‌تواند با پیامدهای دفاعی، تفاوت داشته باشد. به‌این ترتیب، به‌نظر می‌رسد که تهیه‌کنندگان پیش‌نویس، اهداف نظامی را شامل فعالیت‌های نظامی انجام‌گرفته طبق سیاست‌های دفاعی مشترک کشورهای عضو اتحادیه، در نظر گرفته‌اند. ازاین‌رو استفادهٔ غیرنظامی از هوش مصنوعی، با هدف انجام اقدامات دفاعی، مثل پیش‌بینی

---

1. Military Purposes

۲. بخش پنجم معاهدهٔ اتحادیهٔ اروپا با عنوان مقررات عمومی در سیاست خارجی و امنیت مشترک اتحادیه اروپا (مواد ۲۱-۴۶)، به‌صراحت بر لزوم همکاری کشورهای عضو اتحادیه به‌منظور نیل به گسترش دموکراسی، حاکمیت قانون، احترام به اصول بنیادین حقوق بشر، توسعهٔ روابط صلح‌آمیز با کشورهای ثالث، تقویت امنیت جهانی در پرتو اهداف و اصول منشور ملل متحد در داخل و خارج از مرزهای اتحادیه تأکید دارد.

3. Treaty on the European Union (TEU)
4. TEU, Art31(4)
5. TEU, Art41(2)
6. TEU, Art42(1,3,6)
7. TEU, Art45(1)
8. TEU, Art43(1)

عملیات تروریستی، شناسایی افراد آسیب‌پذیر در برابر افراط‌گرایی و همچنین حذف و کنترل اطلاعات نادرست و جعلی، که همگی از قابلیت‌های هوش مصنوعی در مقابله با تروریسم سایبری، هستند، از شمول مقررات پیش‌نویس، مستثنا نخواهد بود. البته همچنان در خصوص استفاده از قابلیت‌های هوش مصنوعی در عملیات نظامی با هدف دفاع از امنیت ملی یا بین‌المللی، مانند حمایت از کشورهای ثالث در مبارزه با گروه‌های شبه‌نظامی تروریست، در قلمرو داخلی‌شان، پاسخ مشخصی وجود ندارد. در این شرایط از یک سو عملیات نظامی انجام می‌گیرد، پس باید از شمول مقررات پیش‌نویس، استثنا شود، اما از سوی دیگر، مستلزم استفاده از نیروی مرگبار[1] نیست. علت وجود ابهام‌هایی از این نوع، عدم امکان تعیین مرز مشخص و دقیق میان اقدامات نظامی و دفاعی است. به‌عبارت دیگر هنگام طراحی و توسعهٔ سیستم‌های هوش مصنوعی، به این موضوع توجه نمی‌شود که از این سیستم در جهت شناسایی اقدامات تروریستی سازمان‌یافته استفاده خواهد شد، یا در انجام تحقیقات مرتبط با تهدید به بمب‌گذاری، چراکه در عمل، غیرممکن است، هنگام طراحی یک فناوری، از جمله هوش مصنوعی، با قطعیت اعلام کرد که منحصراً برای تأمین اهداف غیرنظامی و دفاعی طراحی و برنامه‌ریزی شده است یا خیر. به‌عنوان مثال، نرم‌افزار جاسوسی پگاسوس[2] در ظاهر فناوری هوشمندی است که منحصراً به‌منظور تأمین امنیت ملی، طراحی و توسعه یافته است. اما در عمل، نرم‌افزار جاسوسی است که از آن برای نفوذ به حریم شخصی افراد و سرقت اطلاعات کاربران و نقض حقوق بنیادین آنها استفاده می‌شود.[3]

علاوه‌بر موارد مذکور، به‌نظر می‌رسد استثنا کردن برخی از سیستم‌های هوش مصنوعی از شمول مقررات پیش‌نویس، با استناد بر اهداف امنیتی، آن‌طور که بند ۱۲ یادداشت‌های توضیحی، به آن پرداخته است، نتیجه‌ای جز استفادهٔ گسترده از سیستم‌های ممنوعهٔ هوش مصنوعی مثل دوربین‌های امنیتی بیومتریک، در فضای عمومی نخواهد داشت. بر اساس شق چهارم بند ۱ مادهٔ ۵ استفاده از این سیستم‌ها ممنوع است، مگر آن‌که استفاده از آنها برای پیشگیری از خطرات جانی قریب‌الوقوع و مقابله با حملات تروریستی، ضروری باشد. به‌عبارت دیگر، ممنوعیت‌های مطرح‌شده در مقررات پیش‌نویس، به‌راحتی و به بهانهٔ تأمین امنیت ملی، نادیده گرفته خواهند شد، چراکه در عمل، حتی با وجود محدودیت‌ها و ممنوعیت‌های پیش‌بینی‌شده، تعداد مظنونان احتمالی به ارتکاب اقدامات تروریستی، تقریباً همیشه به اندازهٔ کافی، زیاد خواهد بود، به‌نحوی که استفادهٔ مستمر از سیستم‌های هوش مصنوعی ممنوعه را توجیه کند.

---

1. Lethal Force
2. Pegasus
3. The Rise and Rise of Biometrics Mass Surveillance in the European Digital Rights, Chapter 4, May2020.

۶. نتیجه

مقابله با تروریسم در فضای مجازی و نقش مؤثر هوش مصنوعی در این فرایند همواره مورد توجه بوده است. اما فعالان حقوق بشر همواره نگران آن هستند که مبادا استفاده از این فناوری، به نقض حقوق بنیادین بشری منجر شود. وجود نگرانی‌هایی از این نوع، اتحادیهٔ اروپا را بر آن داشت که به‌منظور تضمین احترام به حقوق بشر، پیش‌نویس اتحادیهٔ اروپا در مورد همسان‌سازی قوانین حاکم بر هوش مصنوعی را ارائه دهد. پیش‌نویس از برخی جهات، گام مثبتی در جهت نظام‌مندسازی استفاده از این فناوری برداشته است. اما کاستی‌هایی نیز دیده می‌شود. مواردی که به‌نظر می‌رسد پیش از تصویب پیش‌نویس و تبدیل آن به قانون لازم‌الاجرا باید مدنظر قرار گیرند، در این زمینه پیشنهاد می‌شود:

– به‌منظور افزایش کارایی و اثربخشی سازوکار مرتبط با طبقه‌بندی سیستم‌های هوش مصنوعی و با توجه به اینکه خطرهای ناشی از سیستم‌های مزبور به‌طور کامل پیش‌بینی‌پذیر نیستند، در مورد سیستم‌های غیرقابل قبول و کم‌خطر نیز فهرستی مشابه فهرست ارائه‌شده در پیوست شمارهٔ ۳ در مورد سیستم‌های پرخطر، به‌منظور به‌روزرسانی موارد تحت پوشش این طبقه به متن پیش‌نویس اضافه شود.

– امکان به‌روزرسانی و تجدیدنظر در محتوای پیوست شمارهٔ ۳ پیش‌بینی شود. همچنین تعداد سرفصل‌های تعیین‌شده نیز افزایش یابد. به‌نظر می‌رسد بهتر است به هشت عنوان پیش‌بینی‌شده، سیستم‌هایی که با استفاده از داده‌های فیزیولوژیکی، رفتاری و بیومتریک، طراحی و تولید می‌شوند نیز اضافه شود.

– اصلاح مفاد بند ۱ مادهٔ ۵، در خصوص ممنوعیت استفاده از هوش مصنوعی، در مواردی که ممکن است رفتار افراد به‌نحوی تحریف شود که به ایجاد آسیب‌های جسمی و روانی علیه اشخاص، به‌خصوص گروه‌های آسیب‌پذیر منجر شود. مفاد این بخش از مادهٔ ۵ به‌نحوی طراحی شود که به ایجاد این شبهه منجر نشود که رفتار یک فرد می‌تواند به شکلی تحریف یا مورد سوءاستفاده قرار گیرد که آسیبی به او وارد نشود. از سوی دیگر، مفهوم عبارت آسیب‌های جسمی و روحی نیز توضیح داده شود، به‌نحوی که ابهام ایجادشده در میزان و نوع آن برطرف شود.

– لزوم پیش‌بینی سازوکار حقوقی مشخص به‌منظور جبران خسارت‌های وارده به اشخاص در اثر استفاده از سیستم‌های هوش مصنوعی پیشنهاد می‌شود. برای تسهیل این فرایند توصیه می‌شود در وهلهٔ اول، حق افراد به عدم مواجهه با سیستم‌های هوش مصنوعی غیرقابل قبول و حق دریافت توضیحات لازم در خصوص نحوهٔ اتخاذ تصمیم با استفاده از هوش مصنوعی، به زبان ساده و قابل فهم برای افرادی که تحت تأثیر تصمیم‌های مزبور قرار گرفته و آسیب دیده‌اند، در متن پیش‌نویس پیش‌بینی شود.

- جایگزینی عبارت اهداف نظامی در مادهٔ ۲ با عبارت عملیات دارای پیامد نظامی یا دفاعی همراه با ارائهٔ توضیحات تکمیلی در زمینهٔ مسائل تحت پوشش. همچنین هرگونه عملیات برنامه‌ریزی‌شده توسط کشورهای عضو اتحادیه در حوزهٔ سیاست‌های مشترک امنیتی و دفاعی، با توسل به سیستم‌های هوش مصنوعی، باید مسئولانه بوده و کاملاً مبتنی بر حقوق بشردوستانهٔ بین‌المللی و اصول بنیادین حقوق بشری باشد.

منابع

۱. فارسی

الف) مقالات

۱. قاسمی، غلام‌علی و باقرزاده، سجاد (۱۳۹۴). جایگاه حقوق بشر در مبارزه با سایبر تروریسم. *مجلهٔ حقوق بین‌المللی*، ۲، ۲۲۷-۲۵۴.

۲. مازاریان، علیرضا (۱۳۹۸). تحلیل انتقادی استدلال عدم تفاوت مربوط در دفاع از هوش مصنوعی. *پژوهش‌های فلسفهٔ کلامی*، ۷۹، ۱۶۵-۱۹۰.

۳. میربد، لیلا؛ سلیمی، صادق؛ نیاورانی، صابر و زمانی، سید قاسم (۱۳۹۸). تروریسم سایبری: نقض حقوق بشر و آزادی‌های بنیادین. *فصلنامهٔ حقوق پزشکی، ویژه‌نامهٔ حقوق بشر و حقوق شهروندی*، ۱۳، ۲۲۴-۲۴۰.

۲. انگلیسی

A) Articles

1. Amershi, S. (2019). Guidelines for human-AI interaction. *Microsoft Research*, 19, 1-13.
2. Angwin, J. (2016). Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks. *ProPublica*, 1-2.
3. Bellovin, S. M. (2014). When Enough is Enough: Location tracking, Mosaic Theory, and Machine Learning. *NYU Journal of Law and Liberty*, 8, 555-628.
4. Brandom, R. (2018). Amazon's facial recognition matched 28 members of Congress to criminal mugshots. *The Verge*, 1-4.
5. Fishman, B. (2019). Crossroads: Counter-Terrorism and the Internet. *Texas National Security Review*, 2, 2-100.
6. Flores, A. (2018). False Positives, False Negatives, and False Analyses: A Rejoinder to 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks. *Federal Probation*,vol. 80, 1-37.
7. Floridi, L. (2020). The fight for Digital Sovereignty: What it is, and Why it Matters, Especially for the EU.*Philosophy & Technology*, *33*, 369–378.
8. Floridi, L. (2021). The European Legislation on AI: a Brief Analysis of its Philosophical

Approach. *Philosophy & technology*, 34, 215–222 .
9. Ghasemi, G. A. & Bagherzadeh, S. (2015). The Position of Human Rights in the Fight against Cyber Terrorism", *International Law Review*, 2, 227-254 (In Persian).
10. Hassani, H. (2020). Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future*?. Journal of AI*, 1, 143-155.
11. Karabiyik, U. (2016). A Survey of Social Network Forensics. *Journal of Digital Forensics" , Security and Law*, 11, 55- 123.
12. Kaur, H. (2020). Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning. *CHI 2020 Paper*, 92, 1-14
14. Land, K. C. (2017). Automating Recidivism Risk Assessment: Should We Stay or Should We Go?. *Criminology & Public Policy*, 16, 231–233.
15. Lilkov, D. (2021). Regulating Artificial Intelligence in the EU: A Risky Game. *European View*, 0, 166–174.
16. Macdonald, S. (2019). Regulating terrorist content on social media: Automation and the rule of law. *International Journal of Law in Context*, 15, 183-197.
17. Mazarian, A. (2019). Critical Analysis against the Argument about the Difference in the Defense of Artificial Intelligence. *The Journal of Philosophical -Theological Research (JPTR)*, 79, 165-190 (In Persian).
18. McKendrick, K. (2019). Artificial Intelligence Prediction and Counterterrorism. The Royal Institute of International Affairs, 1, 1-3.
19. Mirbod, L., Salimi, S., Niavarani, S., & Zamani, S. G. (2019). Cyber Terrorism: Violation of Human Rights and Fundamental Freedoms. *Medical Law Journal*, 13 224-240 (In Persian).
20. Roberts, H. (2021). The Chinese Approach to Artificial Intelligence: An analysis of Policy, Ethics, and Regulation. *AI & SOCIETY*, 36, 59–77
21. Saltman, E. (2020). Countering Terrorism and Violent Extremism at Facebook: Technology, Expertise and Partnerships. *Observer Research Foundation*, 1-5.
22. Shackelford, SJ (2009). From Nuclear War to Net War: Analogizing Cyber Attacks in International Law. *Berkeley J Int'l Law*, 27, 192-251.
23. Smith, W. (2021). UK Intelligence Agency GCHQ Sets out AI Strategy and Ethics. *AI strategy*, 1, 1-3.
24. Spike Back, N. (2018). The Invention of Inductive Machines and the Artificial Intelligence Controversy. *Reseaux*, 5, 1-38.
25. Veale, M. (2021). Demystifying the Draft EU Artificial Intelligence Act — Analyzing the Good, the Bad, and the Unclear Elements of the Proposed Approach. *Computer Law Review International*, 22, 97-112.
26. Veerasamy, N. (2009). Conceptual High-level Framework of Cyberterrori. *International Journal of Information Warfare*, 8, 1-14.

C) Documents

27. Cisco Annual Internet Report (2018–2023), White Paper, 2020, Available at: https://www.cisco.com
28. Frank La Rue, Special Rapporteur on the Promotion and Protection of the Right to

Freedom of Opinion and Expression, United Nations General Assembly, A/HRC/17/27, 16 May 2011.
29. Genaral Data Protection Regulation (GDPR), European Union, 2018. Available at: https://european-union.europa.eu › index_en
30. International Covenant on Civil and Political Rights, United Nation General Assembly, 1966.
31. New technology revealed to help fight terrorist content online**,** Home Office & The Rt Hon Amber Rudd, 2018. Aavailable at: https://www.gov.uk/government/news/new-technology
32. Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonized Rules on Artificial Intelligence (ARTIFICIAL INTELLIGENCE ACT), European Commission, 2021. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT
33. Renda, Andrea, FCAI publishes progress report "Strengthening international cooperation on AI, CEPS, 2022. Available at: https://www.ceps.eu/ceps-news/fcai-publishes-progress-report
34. Resolution on the Promotion, Protection and Enjoyment of Human Rights on the Internet, Human Rights Council, A/HCR/RES/32/13, 18 July 2016.
35. Resolution on the Right to Privacy in the Digital Age, United Nation General Assembly Resolution71/199,26 February2014.
36. Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe, Final Report, European Union, 2021. Available at: https://op.europa.eu/en/publication.
37. The Rise and Rise of Biometrics Mass Surveillance in the European Digital Rights, Chapter 4, May2020. Available at: https://edri.org/wp-content/
38. Treaty on the Functioning of the European Union (TFEU), 2007. Available at: https://eur-lex.europa.eu/legal-content/EN
39. United Nations General Assembly Resolution on Protection of Human Rights and Fundamental Freedoms while Countering Terrorism, United Nations General Assembly, A/RES/62/159, 18 December 2007**.**
40. United Nations and the Rule of Law, Equality and Non-discrimination, 2004.
41. Using the Internet and Social Media for Counter Terrorism Investigation, INTERPOL & UNCCT, 2021. Available at: https://www.interpol.int/en/News-and-Events/News