تحلیل حقوقی نقش هوش‌مصنوعی در امتیازدهی اجتماعی مبتنی بر تبعیض الگوریتمی با تأکید بر قانون هوش‌مصنوعی اتحادیه اروپا

چکیده

تحولات شتابان در عرصهٔ هوش‌مصنوعی، الگوهای سنتی تصمیم‌گیری را دگرگون ساخته و ابعاد نوینی از تعامل انسان و داده را پیش‌روی حقوق‌دانان و سیاست‌گذاران نهاده است. در این میان، امتیازدهی اجتماعی به‌عنوان سازوکاری نوظهور که بر مبنای داده‌های رفتاری و ویژگی‌های شخصی افراد، آنان را ارزیابی و طبقه‌بندی می‌کند، به یکی از مناقشه‌برانگیزترین چالش‌های حقوقی روز بدل‌شده‌است. این فرایند، هرچند می‌تواند کارآمدی تصمیمات اداری و تخصیص منابع را افزایش‌دهد، اما در سطح حقوق بنیادین با مفاهیمی چون برابری، کرامت انسانی و حق حریم‌خصوصی در تعارض قرارمی‌گیرد و زمینه‌ساز تبعیض‌های سیستماتیک و نقض حقوق اشخاص می‌شود.

پژوهش حاضر با روش توصیفی ـ تحلیلی، به واکاوی مبانی، چالش‌ها و سازوکارهای حقوقی امتیازدهی اجتماعی مبتنی بر تبعیض الگوریتمی به ویژه در قانون هوش‌مصنوعی اتحادیه اروپا پرداخته‌است. یافته‌ها نشان می‌دهد که اتحادیه اروپا با اتخاذ رویکردی ریسک‌محور و مرحله‌ای، میان نظام‌های ممنوع و پرخطر تمایز نهاده و صرفاً اشکالی از امتیازدهی اجتماعی را که منجر به رفتار ناموجه یا نامتناسب با اشخاص شود، منع‌کرده‌است؛ درحالیکه سایر کاربردها را مشروط به رعایت الزامات شفافیت، توضیح‌پذیری و نظارت انسانی مجازدانسته‌است. در مقابل، نظام حقوقی ایران علی‌رغم برخورداری از اصول کلی منع تبعیض و حمایت از داده‌های شخصی، هنوز فاقد چارچوبی منسجم برای تنظیم و نظارت بر سامانه‌های امتیازدهی مبتنی بر هوش‌مصنوعی است. بر این اساس، مقاله حاضر بر ضرورت ایجاد چارچوبی بومی و قابل‌اجرا برای تنظیم‌گری امتیازدهی اجتماعی تأکید دارد؛ چارچوبی که با بهره‌گیری از ظرفیت مقررات موجود ازجمله لایحه حمایت از داده‌های شخصی و با پیش‌بینی الزامات متناسب با سطح خطر سامانه‌ها، از جمله شفافیت، ارزیابی سوگیری، حق توضیح، اعتراض، بازبینی انسانی، مستعارسازی و ناشناس‌سازی داده‌ها و سازوکارهای مؤثر نظارت و رسیدگی، از حقوق بنیادین اشخاص در برابر تبعیض‌های الگوریتمی حمایت کند.

کلید واژگان: امتیازدهی اجتماعی، تبعیض سیستماتیک (الگوریتمی)، سوءگیری الگوریتمی، تعصب.

Legal analysis of the role of artificial intelligence in social scoring based on algorithmic discrimination with emphasis on the European Union's artificial intelligence law

Abstract

Rapid advancements in artificial intelligence are fundamentally transforming traditional decision-making paradigms and presenting jurists, legal scholars, and policymakers with unprecedented, complex dimensions of human-data interaction. Among these multifaceted technological developments, the implementation of social scoring mechanisms has emerged as a particularly contentious legal challenge. This sophisticated mechanism, which systematically evaluates and classifies individuals based on their aggregate behavioral data and diverse personal characteristics, has rapidly become one of the most controversial contemporary legal issues globally. Although this algorithmic process can potentially enhance the technical efficiency of administrative decisions and the allocation of public resources, it fundamentally conflicts with established core legal principles, such as equality, inherent human dignity, and the fundamental right to privacy, thereby paving the way for systemic discrimination and the widespread infringement of individual rights.

Employing a rigorous descriptive-analytical research methodology, this article meticulously scrutinizes the theoretical foundations, emerging challenges, and existing legal frameworks of social scoring regarding algorithmic discrimination, with a particular focus on the European Union's recent Artificial Intelligence Act. The findings clearly indicate that the European Union, by successfully adopting a comprehensive risk-based and graduated regulatory approach, effectively distinguishes between prohibited and high-risk systems. It explicitly bans only those specific forms of social scoring that lead to unjustified or disproportionate treatment of individuals, while permitting other legitimate applications strictly conditional upon adherence to stringent requirements for transparency, explainability, and meaningful human oversight. In contrast, Iran's legal system, despite being grounded in general constitutional principles prohibiting discrimination and protecting personal data, still lacks a coherent, specialized framework for regulating and monitoring AI-based scoring systems. Accordingly, this article strongly emphasizes the urgent need to establish a context-specific and enforceable framework for regulating social scoring; a framework that draws on existing regulations, including the Bill on the Protection of Personal Data, and establishes requirements proportionate to the specific level of risk posed, including transparency, bias assessment, the right to explanation, human review, and effective oversight mechanisms to protect rights.

Keywords: Social Scoring, Systemic (Algorithmic) Discrimination, Algorithmic Bias, Prejudice.

مقدمه

تحولات فناوری‌های دیجیتال در دهه‌های اخیر، محیط دیجیتال را از فضایی مکمل در کنار محیط واقعی به بستری مؤثر در زندگی فردی و اجتماعی تبدیل کرده است؛ به‌گونه‌ای که گسترش این فناوری‌ها، نظام‌های حقوقی را نیز با ضرورت بازنگری در قواعد موجود و پیش‌بینی سازوکارهای حقوقی متناسب با تحولات فناورانه مواجه ساخته است (Sadeghi & Malekzadeh Ghaleh, 2026, p. 282).از جمله فناوری‌های عام‌الشمول و تأثیرگذار بر نسل‌های چندگانه حقوق بشر، فناوری هوش‌مصنوعی[1] است؛ برای مثال «اعلامیهٔ اروپایی حقوق دیجیتال و اصول برای دههٔ دیجیتال» کمیسیون اروپا، در بخش سوم با عنوان آزادی انتخاب، تصریح‌می‌کند که اشخاص باید امکان بهره‌گیری آگاهانه از فناوری‌های دیجیتال از جمله هوش‌مصنوعی را داشته‌باشند و درعین‌حال، سازوکارهای لازم برای صیانت از سلامت، امنیت و حقوق بنیادین آنان در برابر پیامدهای احتمالی این فناوری‌ها پیش‌بینی شود. کنار هم آمدنِ تأکید بر «مزایا» و «مخاطرات» در این سند نشان می‌دهد که نوآوری‌هایی نظیر هوش مصنوعی، به‌رغم ظرفیت‌های قابل توجه، بالقوه می‌توانند منشأ تهدید نیز باشند (Mostafavie Ardabili et al., 2022, p. 49).

با گسترش کاربردهای هوش مصنوعی در فرایندهای تصمیم‌گیری، ادبیات حقوقی این حوزه به‌تدریج از بررسی ظرفیت‌های فنی این فناوری به سوی مطالعه آثار آن بر حقوق بنیادین افراد، از جمله حریم خصوصی، برابری، منع تبعیض، کرامت انسانی و خودمختاری فردی، حرکت کرده است. در این مطالعات، مشروعیت استفاده از هوش مصنوعی با توجه به پیامدهای آن برای حقوق و آزادی‌های اشخاص مورد ارزیابی قرار گرفته است. از جمله مهم‌ترین چالش‌های شناسایی‌شده در این زمینه می‌توان به شفافیت الگوریتمی، سوگیری و تبعیض، حفاظت از داده‌های شخصی، مسئولیت‌پذیری و پاسخگویی اشاره کرد. بر این اساس، استفاده از سامانه‌های هوش مصنوعی در حوزه‌هایی که می‌تواند بر وضعیت حقوقی یا اجتماعی افراد اثرگذار باشد، مستلزم ارزیابی مستمر مشروعیت و آثار آن بر حقوق بنیادین است (Salahshour et al., 2025).

سامانه‌های الگوریتمی برای عملکرد صحیح، یادگیری و تولید خروجی‌های دقیق، به حجم مناسبی از داده‌های باکیفیت، متنوع و قابل اعتماد وابسته‌اند؛ در غیر این صورت، احتمال بروز خطا و ارائه تصمیمات نادرست افزایش می‌یابد (European Commission, 2021, p. 44). بااین‌حال، حتی زمانی که چنین سیستم‌هایی به داده‌های کافی و مطلوب نیز دسترسی داشته‌باشند، نمی‌توان از وقوع خطاهای تحلیلی یا پیامدهای تبعیض‌آمیز در فرآیند پردازش داده‌ها جلوگیری کامل کرد. از منظر نظری، اتکای این سامانه‌ها به ساختارهای منطقی و ریاضی سبب می‌شود که برخلاف انسان، فاقد تمایلات شخصی و گرایش‌های عاطفی باشند و در نتیجه، بالقوه توانایی ارائه تصمیمات بی‌طرفانه‌تر و مبتنی بر شاخص‌های عینی را داشته‌باشند (Montaigne Institut, 2020, pp. 17-18). در مقابل، ویژگی‌های فنی همین مدل‌ها ممکن است به شکل‌گیری نوعی تبعیض سیستماتیک منجرشود که از آن با عنوان «تبعیض الگوریتمی» یا تصمیم‌گیری خودکار یادمی‌شود. با توجه به تعهد دولت‌ها نسبت به تحقق برابری و مقابله با اشکال مختلف تبعیض، گسترش سوگیری‌های ناشی از تصمیمات الگوریتمی به یکی از چالش‌های مهم سیاست‌گذاری عمومی بدل‌شده و نهادهای بین‌المللی به‌طور مکرر نسبت به آثار این نوع تبعیض هشدارداده‌اند (Kaye, 2018, pp. 36-38).

یکی از چهره‌های تبعیض الگوریتمی، امتیازدهی اجتماعی است یعنی ارزیابی افراد بر اساس مجموعه‌داده‌های متنوع به‌منظور تولید امتیازات مؤثر بر تصمیمات حیاتی در حوزه‌هایی چون اعتبارسنجی، خدمات اجتماعی و اشتغال. با ظهور هوش‌مصنوعی، اتکاء به سامانه‌های امتیازدهی (چه از سوی نهادهای عمومی و چه خصوصی) به‌شکل فزاینده‌ای برجسته‌تر شده‌است که نگرانی‌های جدی پیرامون حقوق بنیادین برمی‌انگیزد، به‌ویژه در حوزه‌های حفاظت از داده‌ها، منع تبعیض و امکان کنترل اجتماعی، به‌خصوص در ارتباط با داده‌های زیستی حساس. شهروندان اغلب بی‌اطلاع از این فرآیند، برای تعیین صلاحیت اعتباری، ارزیابی عملکرد شغلی، سنجش ایمنی رانندگی، احتمال ارتکاب اقدامات تروریستی یا تقلب اجتماعی و ... امتیازدهی می‌شوند. (Hurley & Adebayo, 2016, p. 148; Weinder, 2017, p. 213; Sánchez-Monedero & Dencik, 2019, p. 14; Dubois, 2021, pp. 120-123; Comel, 2025, p. 3). این امتیازها که مدعی پیش‌بینی رفتار افراد هستند، در تصمیم‌گیری‌های کلیدی درباره آن‌ها

---

1. Artificial Intelligence (AI)

بکارمی‌روند. شرکت‌ها و نهادهای عمومی یا خود این امتیازها را تولید می‌کنند یا آن‌ها را از واسطه‌های داده و مؤسسات اعتبارسنجی خریداری می‌کنند؛ نهادهایی که حجم عظیمی از داده‌ها را از منابع مختلف گردآوری و تحلیل می‌نمایند.(Genicot, 2025, p. 2)

لذا یکی از مهم‌ترین پیامدهای تصمیم‌گیری الگوریتمی، احتمال شکل‌گیری یا تشدید تبعیض است. در ادبیات حقوق عمومی ایران، تبعیض الگوریتمی به‌عنوان یکی از مسائل ناشی از استفاده از الگوریتم‌ها و هوش مصنوعی در تصمیم‌گیری مورد بررسی قرار گرفته و علل آن در دو سطح مرتبط با داده‌های ورودی و مرتبط با طراحی و عملکرد سامانه مورد توجه قرار گرفته است. همچنین، با توجه به فقدان قواعد خاص و جامع برای حمایت از اشخاص در برابر چنین تصمیماتی در نظام حقوقی ایران، استفاده از مجموعه‌ای از راهکارهای حقوقی و تجربیات نظام‌های پیشرو برای کاهش این خلأ مورد توجه قرار گرفته است (Ansari, 2022).

در سطح بین‌المللی نیز تنظیم رابطه میان هوش مصنوعی و حقوق بشر، به سمت ایجاد چارچوب‌هایی برای کنترل خطرات ناشی از چرخه حیات سامانه‌های هوش مصنوعی حرکت کرده است. کنوانسیون چارچوب شورای اروپا در مورد هوش مصنوعی، حقوق بشر، دموکراسی و حاکمیت قانون، بر ضرورت سازگاری فعالیت‌های مربوط به چرخه حیات سامانه‌های هوش مصنوعی با حقوق بشر، دموکراسی و حاکمیت قانون تأکید دارد و خطر ایجاد یا تشدید نابرابری و تبعیض در حوزه‌های دیجیتال را مورد توجه قرار می‌دهد. همچنین، نگرانی از استفاده سرکوبگرانه از سامانه‌های هوش مصنوعی، نظارت خودسرانه یا غیرقانونی و لطمه به حریم خصوصی و خودمختاری فردی، از جمله مبانی توجه به تنظیم‌گری این فناوری در این چارچوب است (Moshrefian, 2025). اهمیت این رویکرد برای موضوع حاضر در آن است که ارزیابی حقوقی سامانه‌های هوش مصنوعی را از سطح کارآمدی فنی فراتر برده و آن را با الزامات حمایت از حقوق بشر و جلوگیری از تبعیض پیوند می‌دهد.

در ارتباط مستقیم‌تر با پدیده امتیازدهی، استفاده از تحلیل داده برای دسته‌بندی، ارزیابی و پیش‌بینی وضعیت افراد و جمعیت‌ها، زمینه شکل‌گیری آنچه در ادبیات «جامعه امتیازدهی» نامیده شده است را فراهم کرده است. در بررسی‌های مربوط به حکمرانی داده‌محور در بریتانیا، استفاده از تحلیل داده در خدمات عمومی برای دسته‌بندی، ارزیابی و پیش‌بینی در سطح فردی و جمعیتی مورد مطالعه قرار گرفته و نشان داده شده است که چنین سازوکارهایی صرفاً ابزارهای فنی برای مدیریت بهتر جمعیت نیستند، بلکه می‌توانند بر رابطه دولت و شهروند و ابعاد سیاسی و اجتماعی حکمرانی داده‌محور اثرگذار باشند (Dencik et al., 2019). از این منظر، امتیازدهی اجتماعی را می‌توان در چارچوب گسترده‌تر داده‌ای‌شدن حکمرانی بررسی کرد که در آن داده‌های مربوط به افراد برای ارزیابی، پیش‌بینی و طبقه‌بندی آنان به کار گرفته می‌شود.

در مجموع، مطالعات پیشین هر یک بخشی از ابعاد حقوقی و فنی هوش مصنوعی، تبعیض الگوریتمی، حفاظت از داده‌های شخصی، شفافیت تصمیم‌گیری و امتیازدهی داده‌محور را بررسی کرده‌اند؛ بااین‌حال، پیوند میان امتیازدهی اجتماعی و تبعیض الگوریتمی، به‌ویژه از منظر تأثیر آن بر حقوق بنیادین اشخاص، همچنان نیازمند تحلیل منسجم است. همچنین، در پژوهش‌های موجود کمتر به تفکیک دقیق امتیازدهی اجتماعی از سایر کاربردهای ارزیابی الگوریتمی و جایگاه آن در قانون هوش مصنوعی اتحادیه اروپا، به‌ویژه ممنوعیت مقرر در ماده ۵ این قانون، پرداخته‌شده‌است. این خلأ در نظام حقوقی ایران اهمیت بیشتری دارد؛ زیرا در نبود چارچوب حقوقی جامع برای تنظیم فرایندهای تصمیم‌گیری خودکار و مقابله با تبعیض الگوریتمی، امکان نقض حریم خصوصی، اصل برابری و سایر حقوق بنیادین افزایش‌می‌یابد. بر این اساس، مقاله حاضر ابتدا هوش مصنوعی و امتیازدهی اجتماعی را به‌عنوان محورهای اصلی پژوهش مفهوم‌شناسی می‌کند(شمارهٔ ۱)؛ سپس جایگاه امتیازدهی اجتماعی مبتنی بر تبعیض الگوریتمی را در قانون هوش مصنوعی اتحادیه اروپا تحلیل می‌نماید (شمارهٔ ۲)؛ و در نهایت، راهکارهای مقابله با امتیازدهی اجتماعی مبتنی بر سوگیری الگوریتمی را برای پیشنهاد به قانونگذار ایران بررسی‌می‌کند (شمارهٔ ۳).

۱. مفهوم‌شناسی

۱/۱. هوش‌مصنوعی

رد پای هوش‌مصنوعی را می‌توان از دهه ۱۹۶۰ پیداکرد ولی در سال ۲۰۰۶ با مقالهٔ «الگوریتم یادگیری سریع برای شبکه‌های باور عمیق» که ماشین‌های محدودشده بولتزمن[۱] را دوباره معرفی‌کرد، شروع به توسعه به چیزی شبیه به شکل فعلی خود کرد.

هوش‌مصنوعی شاخه‌ای از علوم رایانه است که با خودکارسازی رفتارهای هوشمندانه سروکاردارد (Luger; Stubblefield, 1993, p. 1). در تعریفی دیگر هوش‌مصنوعی نرم‌افزار یا سخت‌افزار الگوریتم طراحی‌شده‌ای دانسته‌شده که با توجه به محیط، داده‌ها را اکتساب‌کرده و با تفسیر و پردازش ساختاریافته اطلاعات، بهترین تصمیم‌ها و اقدامات را برای رسیدن به نتیجه مطلوب اتخاذمی‌کند. (European Commission, 2019, p. 3)

بر اساس «قانون هوش‌مصنوعی اتحادیه اروپا»[۲]، «هوش‌مصنوعی» سیستمی مبتنی بر ماشین با سطوح مختلفی از خودمختاری است که ممکن است پس از استقرار، تطبیق‌پذیری نشان‌دهد، و برای اهداف صریح یا ضمنی، از ورودی‌های دریافتی، نتیجه‌گیری می‌کند و خروجی‌هایی مانند پیش‌بینی‌ها، محتوا، توصیه‌ها یا تصمیمات مؤثر بر محیط‌های فیزیکی یا مجازی، تولیدمی‌کند.

در ماده ۱ «سند ملی هوش‌مصنوعی ایران» هوش‌مصنوعی به توانایی ماشین برای انجام عملکردهای خودکار و نظام‌مند از جمله یادگیری، درک، استنتاج، حل مسئله، پیش‌بینی، تصمیم‌گیری و اقدام از طریق بکارگیری دانش و اطلاعات و پردازش داده گفته‌می‌شود که منشأ اثرگذاری‌های گسترده بر انسان و روابط انسانی در محیط فیزیکی یا مجازی و همچنین بازتاب‌های زیست‌محیطی است. هوش‌مصنوعی ماهیتی داده‌ای، شبکه ای، الگوریتمی، خوشه‌ای، لایه‌ای و یکپارچه، مبتنی بر منطق‌های کلاسیک و سایر منطق‌های نوین دارد.

۱/۲. امتیازدهی اجتماعی

اگر هر کنش یک فرد در طول زندگی بتواند به یک عدد تقلیل‌یابد، این عدد می‌تواند بر فعالیت‌های آیندهٔ فرد و تعاملات وی با افراد و نهادها اثرگذار باشد؛ به این پدیده «امتیازدهی»[۳] گفته‌می‌شود؛ یعنی ارزیابی افراد بر پایهٔ مجموعه داده‌های متنوع (مانند ویژگی‌ها و رفتارهای کیفی انسان‌ها ازجمله تحصیلات و دارایی و پیشینهٔ فرهنگی) و در نهایت، تبدیل به داده‌های کمی و تولید عددی که فرآیند تصمیم‌گیری‌های مهم (مانند اعتبارسنجی مالی نظیر تأیید وام، دسترسی به خدمات اجتماعی مانند مسکن یا قراردادهای برق، یا استخدام و ارزیابی قابلیت‌های شغلی) را تحت‌تأثیر قرارمی‌دهد تا امکان پیش‌بینی کنش‌های آتی و اخذ تصمیمات ارزیابی‌محور را فراهم‌آورند. (Dencick et al., 2018,p. 10)

امتیاز یک شهروند می‌تواند بر اساس فعالیت‌های روزمره او تعیین شود؛ مثلاً رعایت پروتکل‌های بهداشت عمومی مانند قرنطینه یا واکسیناسیون ممکن‌است باعث افزایش امتیاز فرد شود. شهروندان، اغلب بدون اطلاع خودشان، امتیازدهی می‌شوند تا مشخص‌شود که آیا واجد شرایط اعتباری هستند یا خیر (مثلاً در بیمه‌ها)، کارگران خوب، رانندگان ایمن یا تروریست‌های بالقوه هستند یا نه (شناسایی ریسک‌ها)، احتمال ارتکاب کلاهبرداری اجتماعی دارند و... . (Dencick et al., 2018,p. 10) در این سازوکار، ابتدا داده‌های مربوط به موضوع (فرد یا گروه) گردآوری شده و سپس از طریق الگوریتم تفسیرمی‌شود تا قضاوت، رتبه یا اندازه‌گیری مشخصی حاصل‌شود و بسته به میزان انطباق با اهداف تعریف‌شده، مشوق‌ها یا محرومیت‌هایی را درپی‌داشته‌باشد(Backer, 2017). بنابراین، هرچه فرد با انتظارات رمزگذاری‌شده در الگوریتم تطابق بیشتری داشته باشد، امتیاز او بالاتر و احتمال دریافت پاداش افزایش‌می‌یابد. مثلاً در فرآیند استخدام، سیستم‌های پیگیری درخواست‌های شغلی، از الگوریتم‌های امتیازدهی برای ارزیابی رزومه‌ها بر اساس معیارهایی چون تحصیلات، تجربه و مهارت‌ها استفاده‌می‌کنند. امتیازهای بالاتر، اولویت بررسی را افزایش‌می‌دهند و روند غربالگری اولیه را برای شرکت‌ها و داوطلبان تسهیل‌می‌کنند.(HiPeople, 2023)

---

[۱]. الگوریتمی است که می‌تواند با بازسازی ورودی، به‌طور خودکار الگوهای ذاتی در داده‌ها را تشخیص‌دهد.

[2] . AI Act

زین‌پس در این مقاله «قانون هوش‌مصنوعی» خوانده‌می‌شود.

[۳]. Scoring

موارد ذیل را می‌توان عناصر کلیدی سامانه‌های امتیازدهی دانست:

۱. ارزیابی احتمال رفتارهای فردی یا گروهی؛ ۲. اتکاء بر داده‌ها و تولید نتایج کمی برای تحلیل‌های مقایسه‌ای؛ ۳. تأثیر در پیش‌بینی آینده، به‌مثابه ابزاری در حمایت از تصمیم‌گیری از طریق ارزیابی عینی.(Dencik et al., 209, p.24)

این قابلیت فناورانه، طیف رو به گسترشی از فعالیت‌ها و رفتارهای انسانی را به داده‌های قابل‌کمی‌سازی تبدیل‌کرده و هدف نهایی آن، استنتاج احتمالات، پیش‌بینی رفتار انسانی و اتخاذ تصمیماتی آگاهانه بر اساس این پیش‌بینی‌هاست. از این‌رو، روش‌های امتیازدهی را می‌توان عناصر اساسی «پارادایم داده‌محور شدن»[1] دانست؛ زیرا این روش‌ها، کنش‌های آنلاین و آفلاین افراد را به داده‌هایی قابل اندازه‌گیری تبدیل و پدیدهٔ «جامعهٔ امتیازدهی‌شده» را روزبه‌روز آشکارتر می‌کنند. این روند با ظهور سامانه‌های تصمیم‌گیری خودکار[2] تسهیل‌شده‌است؛ سامانه‌هایی که در آن‌ها، اختیار تصمیم‌گیری در ابتدا به فرد یا نهاد حقوقی دیگری واگذارمی‌شود تا با استفاده از مدل‌های اجرایی خودکار، اقدامات مشخصی را انجام‌دهد. این سامانه‌ها با امکان پردازش سریع و مؤثر حجم عظیمی از داده‌ها، ارزیابی‌ها را تسریع می‌بخشند. بااینحال، این گرایش نشانگر اتکای روزافزون به شاخص‌های قابل‌کمی‌سازی برای تصمیم‌گیری‌ها و تعاملات اجتماعی است که می‌تواند پیامدهایی جدی در خصوص حقوق بنیادین افراد و آثار منفی متعددی به همراه داشته‌باشد که درقسمت‌های بعدی تحلیل خواهند شد.(Comel, 2025, p. 6)

امروزه امتیازها بطور فزاینده‌ای با استفاده از هوش‌مصنوعی تولیدمی‌شوند. امتیازدهی اجتماعی در قانون جدید هوش‌مصنوعی اتحادیه اروپا مورد اشاره قرارگرفته‌است اما در سابقهٔ قاعده‌گذاری و ادبیات پژوهشی کشور ما تا حدی زیادی مغفول‌مانده‌است. اگرچه در این قانون تکنیک‌های امتیازدهی اجتماعی بعنوان یکی از شیوه‌های هوش‌مصنوعی مفصلاً تنظیم‌نشده‌اند و به شکلی نسبتاً کلی نوشته‌شده‌است، اما با الهام از چین، در ماده ۵ بعنوان یکی از معدود کاربردهای ممنوعهٔ هوش‌مصنوعی ذکرشده‌اند(Genicot, 2025, p. 2)؛ چین در سال ۲۰۱۴ قصد خود را برای راه‌اندازی یک پروژه دولتی بزرگ («نظام اعتبار اجتماعی») با هدف رتبه‌بندی شهروندان چینی بر اساس رفتار خوب آنها، با استفاده از فناوری‌های دیجیتال و هوش‌مصنوعی اعلام‌کرد؛ نظامی به‌منظور نظارت بر شهروندان، نهادها و کسب‌وکارها، از طریق یک چارچوب پیچیده نظارت و ارزیابی، همراه با سازوکارهای تشویق و تنبیه(Wang et al., 2023, pp. 451-455). اما در اروپا، افراد توسط نهادهای دولتی و خصوصی در زمینه‌های مختلف، مانند ارزیابی اعتبار، نظارت بر بهره‌وری کارکنان، تشخیص کلاهبرداری اجتماعی یا خطرات تروریستی و غیره، امتیازدهی می‌شوند.(Genicot, 2025, p. 2)

بنظر می‌رسد که در قلمرو اتحادیه اروپا، تعریف قطعی، جامع و مورد توافق همگانی از اصطلاح «امتیازدهی» وجودندارد و صرفاً انواع یا شیوه‌های امتیازدهی از هم تفکیک‌شده‌است. بااینحال، در برخی مقررات ملی، مثلاً در آلمان، چنین تعریفی وجوددارد: استفاده از یک مقدار احتمالی دربارهٔ اقدامات آتی یک فرد که در فرآیند تصمیم‌گیری بکار گرفته‌می‌شود.(Comel, 2025, p. 4)

طبق بند توضیحی ۳۱ آیین‌نامه هوش‌مصنوعی[3] هم، امتیازدهی اجتماعی به فرآیندی اشاره‌دارد که طی آن، سامانه‌های هوش‌مصنوعی اقدام به ارزیابی یا طبقه‌بندی افراد یا گروه‌ها می‌نمایند؛ این ارزیابی با تحلیل داده‌های مختلف پیرامون تعاملات اجتماعی، رفتارها و ویژگی‌های شخصیتی افراد در زمینه‌های مختلف انجام‌می‌پذیرد، اعم از اطلاعات «شناخته‌شده[4]، استنباط‌شده [5]یا پیش‌بینی‌شده[6]»، در بازه‌های زمانی مشخص. امتیاز بالاتر می‌تواند موجب ایجاد فرصت‌ها و امتیازهای ویژه گردد و رتبه پایین‌تر ممکن‌است به طرد اجتماعی منجرشود.(Comel, 2025, p. 9)

---

1. Datafication
2. ADM
3. AIA
4. Known
5. Inferred
6. Predicted

در «مقررهٔ عمومی حفاظت از داده‌ها» اتحادیه اروپا (جی‌دی‌پی‌آر)[1]، مفهوم «امتیازدهی» به‌طور مستقیم نه در مواد قانونی و نه در شروح مقدماتی مورد اشاره قرار نگرفته‌است، ولی سامانه‌های امتیازدهی مبتنی‌بر هوش‌مصنوعی (به‌ویژه در حوزه اعتبار) تنظیم و تا حدی تعریف‌شده‌است که رأی اخیر و مهم دیوان اروپایی دادگستری[2] این امر را نشان‌می‌دهد. پرونده در مورد دعوایی میان یک فرد و شرکت خصوصی شوفا[3] در آلمان بود؛ شرکتی که از طریق روش‌های آماری-ریاضی، اطلاعاتی را در خصوص اعتبارسنجی مصرف‌کنندگان به شرکای قراردادی خود ارائه‌می‌داد. در مانحن‌فیه، امتیاز اعتباری صادرشده توسط شوفا برای متقاضی، مبنای رد درخواست اعتبار او توسط مؤسسه طرف قرارداد قرارگرفته‌بود. متقاضی با استناد به حق دسترسی مندرج در قسمت (h) بند یک مادهٔ ۱۵[4]، خواستار دسترسی به داده‌های شخصی خود و حذف اطلاعات نادرست احتمالی شد. شوفا در پاسخ، امتیاز اعتباری و مروری کلی از روش‌های محاسبه را ارائه‌داد؛ اما با تکیه بر اصل محرمانگی تجاری، از افشای عوامل مدنظر در محاسبه یا وزن آن‌ها خودداری و تصریح‌نمود که اطلاعات صرفاً در اختیار شرکای قراردادی قرارگرفته و تصمیم نهایی توسط ایشان اتخاذشده‌است.

دادگاه آلمان پرونده را برای تفسیر بند ۱ ماده ۲۲ مقرره، به دیوان اروپایی دادگستری ارجاع‌نمود. طبق این بند، «شخص موضوع داده‌ها باید حق داشته‌باشد که مشمول تصمیمی که صرفاً مبتنی بر پردازش خودکار، از جمله پروفایل‌سازی، است و آثار حقوقی بر او دارد یا بطور مشابه او را بطور قابل‌توجهی تحت‌تأثیر قرارمی‌دهد، قرارنگیرد.»

دیوان اروپایی دادگستری سه شرط اساسی را برای اجرای این بند لازم دانست: باید تصمیمی وجودداشته‌باشد که ۱. به‌طور چشمگیر بر فرد اثر بگذارد؛ ۲. منحصراً براساس پردازش خودکار، از جمله پروفایل‌سازی، اتخاذشده‌باشد؛ ۳. پیامد حقوقی یا تأثیر قابل‌توجه مشابهی برای فرد داشته‌باشد. ازنظر دادگاه مفهوم «تصمیم» دامنهٔ وسیعی دارد و شامل رد درخواست‌های اعتباری آنلاین یا فرآیندهای استخدامی بدون مشارکت انسانی نیز می‌شود؛ اقدام شوفا با تعریف «مقررهٔ عمومی حفاظت از داده‌ها» از «پروفایل‌سازی» نیز منطبق‌بوده و شرط دوم را محقق‌می‌سازد. درخصوص شرط سوم نیز، امتیاز به‌طور چشمگیری در رفتار نهاد تصمیم‌گیرنده مؤثراست؛ مثلاً تصمیم بانک برای اعطای وام، به‌واسطهٔ همین مقدار احتمالی اتخاذمی‌گردد و از این‌رو، تأثیر قابل‌ملاحظه‌ای بر فرد موضوع داده‌ها دارد. براساس این تفسیر، رویه‌های امتیازدهی مبتنی‌بر هوش‌مصنوعی ذیل ماده ۲۲ «مقررهٔ عمومی حفاظت از داده‌ها» قرار می‌گیرند. بنابراین، هرچند در متن مقرره صراحتاً بیان‌نشده، براساس رأی، این سند، سامانه‌های امتیازدهی مبتنی‌بر هوش‌مصنوعی را تنظیم و تا حدی تعریف‌نموده‌است، مشروط بر آن‌که چنین سامانه‌هایی شامل پردازش خودکار و پروفایل‌سازی باشند. (Comel, 2025, pp. 16-17)

امتیازدهی اجتماعی، هم نتیجهٔ تبعیض (از نوع الگوریتمی) است (الف) و هم مولد تبعیض (ب) و البته علاوه بر تبعیض گاه موجب نقض سایر حقوق بنیادین نیز می‌شود (ج):

الف) ارتباط امتیازدهی اجتماعی و سوگیری الگوریتمی

پدیدهٔ امتیازدهی اجتماعی زمانی به تبعیض منتهی‌می‌شود که سازوکارهای تصمیم‌گیری بر پایهٔ الگوهای سوگیری سیستماتیک یا الگوریتمی عمل‌کنند. برای فهم صحیح تبعیض الگوریتمی، نخست باید مفهوم «الگوریتم» و ماهیت «تصمیمات الگوریتمی» روشن گردد (Ansari, 2022, p. 148). الگوریتم در ادبیات علمی تعاریف متفاوتی دارد. بر اساس یکی از تعاریف، الگوریتم مجموعه‌ای از دستورالعمل‌های

---

[1]. General Data Protection Regulation (GDPR)

[2]. ECJ; Case C-634/21 SCHUFA Holding (Scoring) [2023] ECLI:EU:C:2023:957

[3]. SCHUFA Holding AG

[4]. ماده ۱۵: «شخص موضوع داده‌ها حق‌دارد از کنترل‌کننده تأییدبگیرد که آیا داده‌های شخصی مربوط به او پردازش‌می‌شود یا خیر، و درصورت وجود چنین موردی، به داده‌های شخصی و اطلاعات زیر دسترسی‌داشته‌باشد:
(h) وجود تصمیم‌گیری خودکار، از جمله پروفایل‌سازی، که در ماده ۲۲(۱) و (۴) به آن اشاره‌شده است و حداقل در آن موارد، اطلاعات معناداری در مورد منطق مربوطه، و همچنین اهمیت و پیامدهای پیش‌بینی‌شده چنین پردازشی برای شخص موضوع داده.»

ساخت‌یافته است که مسئله‌ای را صورت‌بندی و مراحل لازم برای حل آن را تعیین‌می‌کند؛ به‌گونه‌ای که داده‌های ورودی از طریق محاسبات مشخص به خروجی مطلوب تبدیل‌شوند (Committee of Experts on Internet Intermediaries, 2018, p. 5).

در گزارش دولت فرانسه به کنگره نیز الگوریتم چنین توصیف‌شده‌است: مجموعه‌ای از مراحل محدودکه مسئله و فازهای حل آن مسئله را روشن‌می‌سازند. این تعریف دامنه‌ای گسترده دارد و انواع توالی‌های عملیاتی منجر به تولید یا پردازش اطلاعات را دربر می‌گیرد (Lucchesi & Chignard, 2019, p. 1).

از منظر ماهوی، تمایز قائل‌شدن میان افراد براساس ویژگی‌های معین، یکی از کارویژه‌های بنیادین سامانه‌های الگوریتمی است. (Xenidis; Senden, 2020, p. 5). اما مدل‌های هوش‌مصنوعی مولد، زمانیکه بر روی داده‌های معیوب، ناقص یا غیرمعمول آموزش داده‌می‌شوند، نتایجی را ایجادمی‌کنند که از نظر سیستمی تعصب و سوگیری دارند (سوگیری الگوریتمی) که بدون کنترل و نظارت، گسترش‌می‌یابد و بر تصمیم گیرندگانِ متکی بر این نتایج تأثیرمی‌گذارد و بطور بالقوه منجر به تبعیض می‌شود و بار حقوقی ایجادمی‌کند. (Moore, 2023, paras 5-6)

سوگیری الگوریتمی بعنوان یکی از ریسک‌های سیستماتیک در فصل ۳ «قانون هوش‌مصنوعی» تعریف‌شده‌است: «ریسک سیستماتیک خطری است که به مدل‌های هوش‌مصنوعی همه‌منظوره با قابلیت‌های تأثیرگذار بالا اختصاص‌دارد؛ مدلی که به‌سبب گستره نفوذ یا به‌دلیل آثار منفی بالفعل یا قابل پیش‌بینی آن بر سلامت عمومی، ایمنی، امنیت عمومی، حقوق بنیادین یا کلیت جامعه، می‌تواند تأثیر چشمگیری بر بازار داخلی اتحادیه اروپا داشته باشد و در مقیاسی وسیع در سراسر زنجیره ارزش گسترش یابد..»

ب) ارتباط امتیازدهی اجتماعی و تبعیض

فناوری‌هایی چون هوش‌مصنوعی، دغدغه‌های مهمی را، به‌ویژه در حوزه‌های حریم خصوصی، آزادی بیان، کنترل اجتماعی، و بخصوص حق برابری و منع تبعیض، پدیدمی‌آورد.(McLean et al., 2021, p. 649)

در بسیاری از موارد، ماهیت تبعیض‌آمیز داده‌های ورودی، مستقیماً به تولید خروجی تبعیض‌آمیز در فرایندهای تصمیم‌گیری منجرمی‌شود. به بیان دیگر، هنگامی که الگوریتم‌ها بر مبنای داده‌های یاددهندهٔ آلوده به سوگیری آموزش می‌بینند، خروجی‌های هوش‌مصنوعی نیز می‌تواند نسبت به اشخاص یا گروه‌ها واجد آثار تبعیض‌آمیز باشد. بخشی از این داده‌ها ممکن است ذاتاً حامل نوعی سوگیری باشند و بخشی دیگر ناشی از رفتارها و تصمیمات تبعیض‌آمیز انسان‌ها شکل گرفته‌اند. در حالت اخیر، حتی اگر تلاش‌هایی برای گزینش داده‌های خنثی و عاری از سوگیری صورت گیرد، داده‌ها همچنان بازتاب‌دهندهٔ مجموعه‌ای از تبعیض‌های رسمی و غیررسمی موجود در ساختارهای اجتماعی خواهند بود. در چنین زمینه‌هایی، بازتولید نتایج تبعیض‌آمیز از سوی سیستم‌های الگوریتمی عملاً گریزناپذیر است (Cofone, 2019, pp. 1404-1406).. مثلاً در ۱۹۸۰، یک دانشکده پزشکی در انگلستان برای فرآیند مرتب‌سازی متقاضیان پذیرش از سامانه‌ای رایانه‌ای بهره گرفت که بر اساس داده‌های تاریخی مربوط به پذیرش‌های پیشین آموزش‌دیده‌بود. این سامانه به دلیل سوگیری‌های ضمنی موجود در داده‌های یاددهنده، به‌طور غیرمستقیم علیه زنان و دارندگان سابقه مهاجرت تبعیض قائل‌شد. مطابق گزارش مجله پزشکی انگلیس، این الگوریتم منشأ سوگیری جدید نبود، بلکه صرفاً بازتاب‌دهنده تبعیض‌های نهادینه‌شده موجود بود. ظاهراً در سال‌های تولید داده‌های یاددهنده، افرادی که دانشجویان را انتخاب‌می‌کردند، نسبت به زنان و مهاجران، سوگیری داشته‌اند.(Zuiderveen Borgesius, 2018, p. 11)

پس هوش‌مصنوعی با استفاده از مکانیسم امتیازدهی اجتماعی بر مبنای الگوریتم‌ها، از یکسو می‌تواند در کاهش انواع مختلف تبعیض‌ها مؤثر باشد و از سوی دیگر خودش منجر به تبعیض شود؛ تبعیض‌های اخیر به «تبعیض‌های سیستماتیک» مشهورند که انواع مختلفی دارند:

الف) تبعیض مستقیم و غیرمستقیم

تبعیض غیرمستقیم زمانی رخ‌می‌دهد که هوش‌مصنوعی معیاری یا عملی به‌ظاهر خنثی را بکارمی‌گیرد که بطور نامتناسبی به یک گروه مانند زنان یا افراد با سن، مذهب یا منشأ خاص، آسیب‌می‌رساند.

تبعیض غیرمستقیم از طریق الگوریتم‌ها ممکن‌است به دلیل انتخاب معیارهای جایگزین[1] توسط برنامه‌نویسان رخ‌دهد. معیارهایی که به نظر بی‌طرفانه و منطقی هستند، اما در واقعیت با ویژگی‌های گروه‌های اجتماعی خاصی مانند نژاد، جنسیت یا سن ارتباط قوی دارند و نابرابری‌های موجود در جامعه را بازتولید می‌کنند. برای مثال نرم‌افزار COMPAS که برای پیش‌بینی ریسک ارتکاب مجدد جرم توسط دادگاه‌ها استفاده می‌شود، معیارهایی مانند «نداشتن دیپلم دبیرستان» یا «تعداد دستگیری‌های گذشته» را در نظر می‌گیرد. از آنجا که در برخی مناطق آمریکا، گروه‌های اقلیت به دلیل سوگیری‌های تاریخی، بیشتر تحت نظارت پلیس هستند و لذا بیشتر دستگیر شده‌اند، این داده‌ها در واقع معیارهای جایگزین برای نژاد می‌شوند. در نتیجه، الگوریتم بطور غیرمستقیم، علیه این گروه‌ها تبعیض قائل‌می‌شود، حتی اگر خود الگوریتم بطور مستقیم از نژاد بعنوان یک فاکتور استفاده‌نکند.(Cossette-Lefebvre & Maclure, 2022, p. 6)

در مقابل، تبعیض مستقیم مواردی را شامل‌می‌شود که تصمیم براساس ویژگی خاص فردی گرفته‌می‌شود، درحالیکه آن ویژگی نباید در تصمیم‌گیری تأثیرگذار باشد. مثلاً ردکردن استخدام فردی بر اساس نژاد، قومیت، رنگ پوست، مذهب، جنسیت، سن یا معلولیت جسمی/ذهنی (Cossette-Lefebvre & Maclure, 2022, p. 2)

ب) تبعیض براساس موضوع

تبعیض الگوریتمی براساس موضوع می‌تواند اقسام مختلفی داشته‌باشد:

- **تبعیض بر اساس وضعیت اقتصادی:** ممنوعیت تبعیض بر مبنای عواملی چون رنگ پوست، نژاد و خاستگاه اجتماعی، بطور سنتی و کلاسیک خیلی وقت است که به رسمیت شناخته‌شده و در معاهدات حقوق بشری آمده‌است. ولی سوگیری الگوریتمیِ مبنای امتیازدهی اجتماعی، ممکن‌است براساس عواملی مانند وضعیت اقتصادی، منجر به تبعیض شود که در فهرست علل مذکور نیست. ازاین‌رو زمینه‌های تبعیض الگوریتمی گسترده‌تر است. مثلاً پژوهشی در سال ۲۰۱۵ نشان‌داد که بسیاری از الگوریتم‌های مورد استفاده از سوی شرکت‌های بیمه در آمریکا برای تعیین مبلغ بیمه خودرو، بیش از اینکه به سوابق راننده در رانندگی توجه‌کنند، به ارزیابی وضعیت اعتبارسنجی (مالی) اتکادارند. در نتیجه، مثلاً در فلوریدا، فردی با سابقه رانندگی خوب و بدون نمره منفی اما با نمره اعتبار مالی ضعیف، در مقایسه با راننده‌ای با سابقه محکومیت رانندگی در حال مستی ولی دارای نمره اعتبارمالی خوب، باید مبلغ بیشتری برای بیمه خودرو بپردازد.(Rovatsos et al., 2019, p. 19)

- **تبعیض براساس رویکردهای جنسیتی:** برخی ارزیابی‌های میدانی نشان‌می‌دهند که الگوریتم‌ها چگونه در کشورهای مختلف منشأ انواع تصمیمات تبعیض‌آمیز براساس جنسیت شده‌اند. مثلاً در فرانسه، استفاده از الگوریتم برای تعیین قیمت بصورت شخصی‌سازی‌شده[2]، می‌تواند با افزایش قیمت کالاهای مربوط به قاعدگی زنان، برای آنان تبعیض ایجادکند. یا در الگوریتم‌های کاریابی، مشاهده‌شده‌است که این سامانه‌ها در بلژیک هنگام شناسایی فرصت‌های شغلی، زنان را نسبت به مردان در موقعیتی نابرابر قرارداده‌اند و در کرواسی نیز در ارائه فرصت‌های کاری برای زنانی که دارای دو فرزند و بیشتر هستند، رفتاری تبعیض‌آمیز از خود نشان‌داده‌اند (Gerards & Xenidis, 2021, p. 86).

- **تبعیض براساس اعتقادات مذهبی:** حق آزادی مذهب در میثاق بین‌المللی حقوق مدنی و سیاسی شناسایی‌شده و براساس آن هرکس حق آزادی فکر و وجدان و مذهب دارد و می‌تواند به‌طور فردی یا جمعی، علنی یا در خفا، در اعمال مذهبی شرکت‌کند. (Mostafavi Ardebili et al., 2023, p. 89)

اما پژوهش‌ها نشان داده‌اند که GPT-3، به‌عنوان یکی از پیشرفته‌ترین مدل‌های زبانی، دارای سوگیری مستمر و پایدار علیه مسلمانان از حیث پیوند آنان با خشونت است که شدت آن حتی در مقایسه با سایر گروه‌های دینی بارزتر است. مثلاً در الگوریتم این مدل زبانی، واژهٔ «Muslim» در ۲۳٪ موارد با «terrorist» قیاس‌گردیده، درحالیکه واژهٔ «Jewish» صرفاً در ۵٪ موارد با «money» مرتبط‌شده‌است. همچنین، حتی با

---

1. Proxies
2. Personalized price

به‌کارگیری مداخلات مثبت در متن ورودی، تکمیل‌های خشونت‌آمیز مرتبط با «Muslims» به‌طور نسبی کاهش‌یافته، اما همچنان در سطحی بالاتر از سایر گروه‌های دینی باقی‌مانده‌است. (Abid, Farooqi, & Zou, 2021, p. 1)

یا مثلاً ابتکار موسوم به «Extreme Vetting Initiative» در آمریکا با هدف ارزیابی سطح ریسک متقاضیان ویزا از طریق نظارت بر فعالیت‌های آنان در شبکه‌های اجتماعی، به‌طور نامتناسب افراد متعلق به کشورهای با اکثریت مسلمان را هدف قرارمی‌داد و در نتیجه با انتقادات گسترده‌ای در خصوص سوگیری الگوریتمی و نقض آزادی‌های مدنی مواجه‌شد (https://www.accessnow.org).

- **تبعیض شغلی:** حق برخورد عادلانه در مسائل مربوط به کار از اصول اولیه برابری میان شهروندان است. مفهوم فرصت‌های شغلی برابر، در مقررات هر کشور وجوددارد و تضمین‌می‌کند که هیچکس در فرایند جست‌وجوی شغل و پس از اشتغال مورد تبعیض قرارنمی‌گیرد. «کنوانسیون تبعیض (حرفه و اشتغال)» ۱۹۵۸ شماره (۱۱۱)[1] تبعیض را به منزله هر تمایز، محرومیت یا اولویت ایجادشده براساس نژاد، رنگ، جنس، مذهب و عقیده سیاسی تعریف‌می‌کند که بر برابری فرصت در اشتغال اثرمی‌گذارد.

یکی از حوزه‌های مهم کارکرد هوش‌مصنوعی بکارگیری این فناوری در راستای تحقق برابری در محیط کار است. تحقیقات تجربی متعددی در زمینه تأثیر هوش‌مصنوعی در محیط کار انجام‌شده‌است؛ چراکه هرگونه تأثیر این هوش در این محیط مستقیماً ارتقا یا تنزل حقوق شناخته‌شده بشری را به دنبال خواهدداشت؛ هوش‌مصنوعی معمولاً در دو مرحله به کارفرمایان و کارکنان در فرآیند جذب نیروی کار، برابری در کار، بهبود شرایط محیط کار و مدیریت منابع انسانی کمک می‌کند: ۱) فرآیند جذب و استخدام: هوش‌مصنوعی می‌تواند رزومه افراد را بررسی و نامزدهای مناسب برای استخدام و گزینش در یک شغل را فهرست کند. هوش‌مصنوعی قادراست از طریق حالت صوتی- تصویری و انتخاب کلمات نامزد اشتغال، گفتار و زبان بدن را ارزیابی و ویژگی‌های نامزدها را تجزیه و تحلیل کند و بدین شکل نابرابری در استخدام را کاهش‌دهد. مثلاً تکستیو[2]، یک شرکت کاریابی هوش‌مصنوعی‌محور است که توصیفات شغلی را تجزیه و تحلیل و نظارت‌می‌کند و برای جایگزینی نامزدهای منفعل، اطمینان از تعادل مبتنی بر جنس ، ازبین‌بردن تعصبات ناخودآگاه مبتنی بر جنس و هدف قراردادن واجد شرایط‌ترین نامزدها فعالیت می‌کند(Brennan et al., 2018). ۲) پس از مرحلهٔ استخدام و در جریان مدیریت منابع انسانی: هوش‌مصنوعی تعصب و تبعیض در محیط کار را شناسایی‌می‌کند و پیشنهاداتی برای کاهش و اصلاح آن می‌دهد برنامه‌های مبتنی بر هوش‌مصنوعی، داده‌های عملکرد کارکنان را بررسی، کارکنان خوب را شناسایی و کارکنانی که نیاز به تغییر موقعیت دارند را معرفی‌می‌کنند.(Bhardwaj et al., 2020)

اما در کنار این تأثیرات مثبتی که هوش‌مصنوعی می‌تواند در ارتباط مسائل مختلف حوزهٔ اشتغال داشته‌باشد، مخاطراتی نیز در همین زمینهٔ تبعیض درپی‌دارد به نحویکه می‌توان آن را به مثابهٔ یک شمشیر دو لبه در نظر گرفت: ابزاری که هم می‌تواند از تبعیض کاری جلوگیری‌کند یا آن را کاهش‌دهد، و از سوی دیگر خود می‌تواند به تبعیض کاری منجرشود. طبق «قانون هوش‌مصنوعی»، سیستم‌های هوش‌مصنوعی مورد استفاده در استخدام و مدیریت کارگران، به‌دلیل اثرگذاری گسترده بر تصمیم‌گیری‌های موثر بر استخدام و انتخاب افراد، شرایط کار، نظارت یا ارزیابی افراد و ارتقا و خاتمه روابط کاری، تخصیص وظایف بر اساس رفتار فردی، باید بعنوان سیستم‌های پرخطر طبقه‌بندی شوند چون می‌توانند تأثیر قابل‌توجهی بر آینده شغلی و معیشت کارکنان و کارگران داشته‌باشند. چنین سیستم‌هایی ممکن‌است الگوهای تاریخی تبعیض را تداوم بخشند، مثلاً علیه زنان، معلولان، یا افراد دارای ریشه‌های نژادی، قومی، گروه‌های سنی یا گرایشات جنسی خاص. (European Union, 2024, Art. 57)

تبعیض براساس موضوع می‌تواند انواع مختلف دیگری داشته باشد که قابل احصاء نیست مثلاً در مسئله مهاجرت و پناهندگی، جاییکه هوش‌مصنوعی با امتیازدهی نادرست به برخی گروه‌های پناهندگان، منجر به تبعیض شود.(Batte, 2025, p. 3)

---

[1] . Discrimination (Employment and Occupation) Convention, 1958 (No. 111) https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C111

[2] . https://textio.com

ج) ارتباط امتیازدهی اجتماعی و نقض دیگر حقوق بنیادین

امتیازدهی اجتماعی به‌طور نظام‌مند، گروه‌ها یا افراد خاصی را در موقعیتی نامناسب و گاه تبعیض‌آمیز قرار می‌دهد؛ (Ferrer et al., 2021, p. 75) . اما غیر از خودِ تبعیض و نقض اصل برابری، چنین نظامی ممکن‌است چندین نقض بالقوه حقوق بنیادین را درپی داشته‌باشد:

– حق حریم خصوصی و حفاظت از داده‌های شخصی به دلیل افشای عمومی تعاملات و دیدگاه‌های شخصی افراد، نقض‌می‌شود؛ امری که افراد را در معرض نظارت و احتمال سوءاستفاده از اطلاعات محرمانه‌شان قرار می‌دهد.[1]

– تشویق نظام به همسان‌سازی و تنبیه اندیشهٔ مخالف، موجب لطمه به آزادی بیان می‌گردد و اثر خفقان‌آور بر آزادی‌های مدنی دارد؛ زیرا افراد ممکن‌است از ابراز دیدگاه‌های متفاوت بازداشته‌شوند.

– نقض اصل کرامت و خودمختاری: این پدیده در ادبیات علمی با عنوان «جبرباوری داده‌ای»[2] شناخته می‌شود. افراد بر اساس استنباط‌های آماری استخراج‌شده از داده‌هایشان ارزیابی می‌شوند که ممکن‌است خودمختاری ایشان در ارائه‌ی تصویر از خویشتن را تضعیف‌کند. در نتیجه، این فرآیندها مفاهیم پیچیده را به ساختارهای خشک و غیرقابل انعطاف بدل‌می‌سازند و با حذف سیالیت و امکان مناقشه‌گری، موجب خدشه به خودمختاری فردی و زمینه‌ساز تبعیض‌های عمیق‌تر می‌گردند(Mann & Mtzner, 2019). نهادهای متعددی نگرانی‌های خویش را نسبت به این موضوع ابرازداشته‌اند؛ از جمله، یونسکو در «توصیه‌نامه اخلاق در هوش‌مصنوعی»، تصریح‌نموده که سامانه‌های هوش‌مصنوعی نباید برای امتیازدهی اجتماعی یا نظارت گسترده مورد استفاده قرار گیرند؛ یا «مرجع حفاظت داده‌های ایتالیا» که نسبت به پیامدهای حقوقی منفی احتمالی ناشی از نظام «شهروندی مبتنی بر امتیاز» بر حقوق و آزادی‌های افراد، به‌ویژه اقشار آسیب‌پذیر، هشدارداده‌است.(Garante per la Protezione dei Dati Personali, 2022)

اصل برائت، حقوق مصرف‌کننده و ... هم از دیگر مواردی هستند که ممکن‌است از طریق امتیازدهی اجتماعی نقض‌شوند. بخصوص وقتی فرآیندها غالباً به‌صورت غیرشفاف عمل می‌کنند و افراد یا از تحت‌پروفایل قرارگرفتن خود آگاه‌نیستند یا درک کاملی از سازوکارهای دخیل ندارند. (Article 29 Data Protection Working Party, 2018)

۲) جایگاه **امتیازدهی اجتماعی در قانون هوش‌مصنوعی اتحادیه اروپا**

براساس مقدمه ۳۱ «قانون هوش‌مصنوعی»، سیستم‌های هوش‌مصنوعی ارائه دهندهٔ امتیازدهی اجتماعی، ممکن‌است:

۱. منجر به نتایج تبعیض‌آمیز و محرومیت و به‌حاشیه‌راندن گروه‌های خاصی شوند؛

۲. حق کرامت و عدم تبعیض و برابری و عدالت را نقض‌کنند؛

---

[1]. مفهوم جامعه پانوپتیکون (Panopticon Society) در این زمینه قابل تأمل است. پانوپتیکون یک طرح معماری برای زندان است که توسط جرمی بنتام ارائه شد. در این طراحی، یک برج مرکزی وجود دارد که نگهبان در آن مستقر می‌شود و می‌تواند بدون آنکه خود دیده شود، تمامی زندانیان را در سلول‌های پیرامون زیر نظر بگیرد. اصل کلیدی در این سیستم، ایجاد «احساس نظارت دائمی» است. ازآنجاکه زندانیان هرگز مطمئن‌نیستند که در لحظه تحت‌نظر هستند یا خیر، این عدم قطعیت آنان را وادارمی‌کند تا دائماً رفتار خود را نظم بخشیده و کنترل کنند، گویی که همواره ناظری در حال مشاهده آنان است. این مکانیسم، هزینه‌های نظارت را به حداقل و کنترل را به حداکثر می‌رساند، زیرا افراد به تدریج تبدیل به ناظران خود می‌شوند.
در جهان معاصر، فناوری‌های دیجیتال و سیستم‌های نظارتی، پانوپتیکون را از یک استعاره به واقعیتی اجتماعی تبدیل‌کرده‌اند. مثلاً در فیسبوک حرکات کاربران تحت‌نظر گرفته‌می‌شود یا در شهرهایی با دوربین‌های نظارتی فراوان مانند لندن و شهرهای چین که در ازای هر شش نفر یک دوربین مداربسته وجوددارد. این نظام نظارتی مستمر و فراگیر، به وضوح با حق بر حریم خصوصی که در اسناد بین‌المللی مانند «الگوی بین‌المللی حقوق مدنی و سیاسی» به رسمیت شناخته شده‌است، در تعارض است. هسته اصلی حق حریم خصوصی، حق فرد برای داشتن حوزه‌ای مستقل، فارغ از مداخله و نظارت دولتی و عمومی است. پانوپتیکون دیجیتال این حریم را بکلی ازبین‌می‌برد و فرد را در معرض مشاهده دائمی و بدون رضایت قرارمی‌دهد.

[2]. Data determinism

۳. اشخاص حقیقی یا گروه‌هایی از آن‌ها را براساس داده‌های متعدد مربوط به رفتار اجتماعی آن‌ها در زمینه‌های مختلف یا ویژگی‌های شخصی یا شخصیتی شناخته‌شده، استنباط‌شده یا پیش‌بینی‌شده در دوره‌های زمانی خاص، ارزیابی یا طبقه‌بندی می‌کنند؛

۴. امتیاز اجتماعی به‌دست‌آمده از چنین سیستم‌های هوش‌مصنوعی ممکن‌است منجر به رفتار مضر یا نامطلوب با اشخاص در زمینه‌های اجتماعی شود که با زمینه‌ای که داده‌ها در ابتدا در آن تولید یا جمع‌آوری شده‌اند، ارتباطی ندارند یا به رفتاری مضر که با شدت رفتار اجتماعی آن‌ها نامتناسب یا ناموجه است، منجرشود؛

۵. بنابراین، سیستم‌های هوش‌مصنوعی که مستلزم چنین شیوه‌های امتیازدهی غیرقابل قبولی هستند و منجر به چنین پیامدهای مضر یا نامطلوبی می‌شوند، باید ممنوع شوند؛

۶. این ممنوعیت نباید بر شیوه‌های ارزیابی قانونی اشخاص حقیقی که برای هدف خاصی مطابق با قوانین اتحادیه و ملی انجام می‌شوند، تأثیر بگذارد.(AI Act,2024)

بطور ویژه در اینمورد، ماده ۵ قانون، اقدامات ممنوعه در حوزه هوش‌مصنوعی را فهرست می‌کند. بند ۱ این ماده مقررمی‌دارد: «رویه‌های هوش‌مصنوعی ممنوع

شیوه‌های هوش‌مصنوعی زیر ممنوع خواهندبود:

(ج) عرضه در بازار، به‌کارگیری یا استفاده از سیستم‌های هوش‌مصنوعی برای ارزیابی یا طبقه‌بندی اشخاص حقیقی یا گروه‌هایی از افراد در یک دوره زمانی خاص براساس رفتار اجتماعی یا ویژگی‌های شخصی یا شخصیتی شناخته‌شده، استنباط‌شده یا پیش‌بینی‌شده آنها، بطوری که امتیاز اجتماعی منجر به یک یا هر دو مورد زیر شود:

۱) رفتار مضر یا نامطلوب با اشخاص حقیقی یا گروه‌هایی از افراد خاص در زمینه‌های اجتماعی که با زمینه‌هایی که داده‌ها در ابتدا در آنها تولید یا جمع‌آوری شده‌اند، نامرتبط هستند؛

۲) رفتار مضر یا نامطلوب با اشخاص حقیقی یا گروه‌هایی از افراد خاص که ناموجه یا نامتناسب با رفتار اجتماعی یا شدت آن است؛ European (Union, 2024)

تعریف امتیازدهی اجتماعی در این ماده به هیچ‌وجه واضح نیست. عبارت «ارزیابی یا طبقه‌بندی اشخاص حقیقی یا گروه‌هایی از افراد در یک دوره زمانی خاص براساس رفتار اجتماعی آنها» بسیار مبهم‌است و هدف قانون‌گذار معلوم‌نیست. این امر، چگونگی اعمال ممنوعیت امتیازدهی اجتماعی را دشوارمی‌کند. حتی ممکن‌است تصورشود که ماده ۵ قانون، امتیازدهی اجتماعی را کلاً ممنوع‌کرده‌است.

در مقایسه با سایر شیوه‌های هوش‌مصنوعیِ ممنوعه مانند شناسایی بیومتریک بلادرنگ، امتیازدهی اجتماعی موضوع بحث نسبتاً کمی بین قانونگذاران بوده‌است. تنها اصلاحیه مهمی که به پیشنهاد اولیه کمیسیون اضافه‌شده‌است، ممنوعیت امتیازدهی اجتماعی بطور کلی، چه توسط سازمان‌های دولتی و چه خصوصی، است درحالیکه کمیسیون این ممنوعیت را به مقامات دولتی محدودکرده‌بود. آنچه واضح‌بود این بود که اتحادیه اروپا می‌خواست خود را از چین متمایزکند، که در سال ۲۰۱۴ قصد خود را برای راه‌اندازی یک «سیستم دولتی اعتبار اجتماعی» اعلام کرد.

بنظرمی‌رسد که این ممنوعیت بگونه‌ای انعطاف‌پذیر تدوین‌شده‌است و بنابراین احتمالاً آن را به ابزاری مفید، مشابه و مکمل ماده ۲۲ مقررهٔ عمومی حفاظت از داده‌ها، برای محافظت از افراد در برابر اشکال خاصی از استفاده نامتناسب از امتیازدهی مبتنی بر هوش‌مصنوعی تبدیل می‌کند(Genicot, 2025, p. 1)مرز بین امتیازدهی اجتماعی و سیستم‌های امتیازدهی هوش‌مصنوعیِ پرخطر در «قانون هوش‌مصنوعی» به‌هیچ‌وجه مشخص‌نیست. تشخیص اینکه آیا یک سیستم بطور علّی به یک نتیجه خاص مرتبط است ممکن‌است دشوارباشد، به‌ویژه در مواردیکه زمینه‌های

اجتماعی با شرایطی که داده‌ها در آن تولید یا جمع‌آوری شده‌اند متفاوت است، یا زمانیکه رفتار حاصل ناموجه یا نامتناسب تلقی می‌شود. این ابهام ممکن‌است به نهادهای خصوصی یا عمومی اجازه‌دهد از این سیستم‌ها استفاده و استدلال کنند که امتیازدهی عامل تعیین‌کننده نبوده‌است. در واقع، ماده ۵(۱)(c) که شامل این ممنوعیت است، بطور گسترده تدوین‌شده و امکان کاربرد گسترده را فراهم‌می‌کند.

در این راستا، همانطور که گذشت، ممنوعیت امتیازدهی اجتماعی ویژگی‌های مشترکی با ممنوعیت تصمیمات خودکار مندرج در ماده ۲۲ مقررهٔ عمومی حفاظت از داده‌ها دارد و تفسیر دیوان دادگستری اتحادیه اروپا[1] در پرونده شوفا می‌تواند بعنوان منبع الهامی برای پیش‌بینی دامنه ماده ۵(۱)(c) «قانون هوش‌مصنوعی» باشد.

توضیح اینکه ماهیت «اجتماعیِ» امتیاز (که آن را ممنوع‌می‌کند) یک ویژگی ذاتی امتیاز نیست، بلکه موضوعی وابسته به زمینه است. امتیاز می‌تواند در موردی قانونی باشد و در مورد دیگر بدلیل نحوه استفاده از آن (تکنیک‌های امتیازدهی و خطراتی که برای حقوق اساسی ایجاد می‌کند) بعنوان امتیاز اجتماعی واجد شرایط ممنوعیت باشد.(Comel, 2025, p. 25)

پس برای آنکه یک امتیاز مشمول ممنوعیت مندرج در ماده ۵ شود، تحقق یکی از دو شرط ضروری است: نخستین شرط به زمینه کاربردی مربوط می‌شود: شرط «نامرتبط‌بودن زمینه گردآوری داده با زمینه استفاده از امتیاز»؛ رفتار نامطلوب باید در زمینه‌های اجتماعی‌ای رخ داده باشد که با زمینه‌ای که داده‌ها در آن گردآوری یا تولید شده‌اند، نامرتبط باشد. این شرط به‌طور ضمنی در پی منع استفاده چندمنظوره از امتیاز است.

دومین شرط، شرط «هدف عمومی» بودن امتیازدهی اجتماعی است. در «راهنمای اخلاقی»، امتیازدهی اجتماعی به‌عنوان فرآیندی که «همه جنبه‌ها» را در بر می‌گیرد، توصیف شده است. به‌همین ترتیب، بند ۱۷ مقدمه پیش‌نویس اولیه قانون هوش‌مصنوعی کمیسیون اروپا فرض می‌کرد که امتیازدهی اجتماعی برای «اهداف عمومی» انجام می‌شود. هرچند این عبارت در نسخه نهایی حذف شده، اما جمله پایانی بند ۳۱ مقدمه فعلی تصریح می‌کند که ممنوعیت شامل امتیازدهی مشروع برای «هدف خاص» نمی‌شود؛ که به‌صورت قیاسی، امتیازهایی که برای هدف خاصی استفاده نمی‌شوند، باید ممنوع تلقی‌گردند.

می‌توان گفت شرط نخست گسترده‌تر از شرط دوم است. (Genicot, 2025, p. 15)

ماده ۵(۱)(c) سیستم‌های امتیازدهی را ممنوع‌نمی‌کند، بلکه فقط امتیازات اجتماعی را ممنوع‌می‌کند. این یافته، نتیجه‌گیری‌های پرونده شوفا در دیوان دادگستری اتحادیه اروپا را تکرارمی‌کند که در آن دادگاه اظهار داشت امتیاز فروخته‌شده توسط شوفا خود یک تصمیم خودکار بوده است زیرا نقش تعیین‌کننده‌ای در رد اعتبار متقاضی ایفاکرده‌است. حتی اگر سیستم امتیازدهی شوفا تصمیم‌گیری خودکار نباشد (زیرا امتیازات لزوماً به یک تصمیم منجرنمی‌شوند)، امتیاز شوفا در نهایت بعنوان یک تصمیم خودکار واجد شرایط شد (به دلیل تأثیر امتیاز بر رد نهایی اعتبار در این مورد). بطور مشابه، حتی اگر سیستم امتیازدهی شوفا (یا هر آژانس اعتبارسنجی مصرف‌کننده دیگر) یک سیستم امتیازدهی اجتماعی نباشد، یک امتیاز شوفا می‌تواند بعنوان یک امتیاز اجتماعی واجد شرایط شود اگر توسط شخص ثالثی در زمینه‌ای غیر از اعتبار استفاده‌شود.

علاوه بر این، ممنوعیت امتیازدهی اجتماعی ممکن‌است شامل امتیازاتی شود که برای یک هدف خاص و در یک زمینه محدود استفاده‌می‌شوند. در واقع، اگرچه بند ۳۱ بیان‌می‌کند که ارزیابی‌های قانونی افراد که برای یک هدف خاص انجام‌می‌شوند، تحت این ممنوعیت قرار نمی‌گیرند، اما طبق عبارت ماده ۵(۱)(c)(ii) یک امتیاز ممنوع است اگر پردازش ناشی از آن امتیاز، با توجه به رفتار اجتماعی یا شدت آن، نامتناسب یا ناموجه باشد. هیچ چیز در اینجا مانع از این نمی‌شود که امتیازی که برای یک هدف خاص استفاده می‌شود، اما نامتناسب یا ناموجه بنظرمی‌رسد، بعنوان یک امتیاز اجتماعی تلقی‌شود. با توجه به این موارد، ممنوعیت امتیازدهی اجتماعی بعنوان ابزاری مکمل برای ممنوعیت تصمیمات خودکار معرفی شده توسط ماده ۲۲ «مقررهٔ عمومی حفاظت از داده‌ها» بنظرمی‌رسد. (Genicot, 2025, p.18)

---

1. CJEU

برای درک نحوه مواجهه قانون هوش‌مصنوعی با تکنیک‌های نمره‌دهی، مرور «راهنمای اخلاقی برای هوش‌مصنوعی قابل اعتماد» (Ethics Guidelines) که در سال ۲۰۱۹ توسط گروه کارشناسی سطح‌بالای هوش‌مصنوعی وابسته به کمیسیون اروپا منتشرشد، مفیداست(High-Level Expert Group on Artificial Intelligence, 2019). این سند تأثیر قابل‌توجهی بر پیشنهاد مقررات کمیسیون داشته‌است. در این سند، تکنیک‌های نمره‌دهی در کنار سایر کاربردهای هوش‌مصنوعی که نگرانی‌های جدی ایجادمی‌کنند، موردبحث قرارگرفته‌اند و به‌عنوان تهدیدی برای اصول خودمختاری و عدم تبعیض توصیف‌شده‌اند. بااینحال، راهنمای اخلاقی میان «نمره‌دهی شهروندان - در مقیاس بزرگ یا کوچک - که اغلب در نمره‌دهی توصیفی و حوزه‌محور استفاده‌می‌شود» و «نمره‌دهی هنجاری شهروندان (ارزیابی کلی از شخصیت اخلاقی یا تمامیت اخلاقی) در همه جنبه‌ها و در مقیاس وسیع» تمایزقائل‌می‌شود؛ نوع دوم برای ارزش‌های دموکراتیک خطرناک است، به‌ویژه «زمانی که به‌صورت نامتناسب و بدون هدف مشروع مشخص و اعلام‌شده استفاده‌شود».

به‌نظر می‌رسد دو ویژگی نمره‌دهی هنجاری شهروندان را از سایر انواع نمره‌دهی متمایز می‌سازد: ۱) این نوع نمره‌دهی شامل ارزیابی کلی از شخصیت فرد است؛ ۲) این ارزیابی محدود به یک حوزه یا هدف خاص نیست، بلکه پیامدهایی برای تمام جنبه‌های زندگی دارد و به‌صورت نامتناسب استفادهمی‌شود. (Genicot, 2025, p.9)

به عنوان جمع بندی می‌توان گفت ممنوعیت امتیازدهی اجتماعی در ماده ۵ «قانون هوش مصنوعی» اتحادیه اروپا به معنای ممنوعیت مطلق تمامی اشکال امتیازدهی به افراد نیست. این ممنوعیت ناظر به نوع خاصی از امتیازدهی اجتماعی است که واجد شرایط مقرر در ماده ۵ بوده و در زمره کاربردهای ممنوع هوش مصنوعی قرار می‌گیرد. در مقابل، وجود سازوکارهای امتیازدهی در حوزه‌هایی مانند اعتبارسنجی مالی، ارزیابی کارکنان، بیمه و سایر زمینه‌های مشابه، به‌خودی‌خود موجب ممنوع بودن آنها تحت عنوان Social Scoring نمی‌شود. چنین کاربردهایی، حسب موضوع، کارکرد و میزان خطری که برای حقوق و منافع اشخاص ایجاد می‌کنند، ممکن است در چارچوب الزامات مربوط به سامانه‌های هوش مصنوعی پرخطر قرار گیرند (Genicot, 2025,p. 1-3).

از این منظر، باید میان «امتیازدهی اجتماعی» ممنوع و سایر اشکال امتیازدهی که ممکن است تحت نظام تنظیم‌گری مبتنی بر ریسک قرار گیرند، تفکیک کرد. امتیازدهی اجتماعی ممنوع، به دلیل ماهیت و آثار خاص خود، در سطح کاربردهای ممنوع قرار می‌گیرد؛ در حالی که امتیازدهی در حوزه‌هایی مانند اعتبار، استخدام یا بیمه، صرفاً به دلیل امتیازدهی بودن، ممنوع محسوب نمی‌شود و ممکن است در صورت احراز شرایط قانونی، در زمره سامانه‌های پرخطر قرار گیرد. بنابراین، «ممنوع» و «پرخطر» در ساختار قانون هوش مصنوعی اتحادیه اروپا دو سطح متفاوت از مداخله تنظیم‌گرانه‌اند و نمی‌توان یکی را معادل دیگری دانست. بر این اساس، برای تشخیص شمول ممنوعیت Social Scoring، صرف وجود یک سازوکار امتیازدهی کافی نیست؛ بلکه باید به معیارهای مقرر در ماده ۵ توجه کرد. بر اساس این معیارها، ممنوعیت ناظر به مواردی است که اشخاص یا گروه‌ها در یک دوره زمانی، بر مبنای رفتار اجتماعی یا ویژگی‌های شخصی یا شخصیتی شناخته‌شده، استنباط‌شده یا پیش‌بینی‌شده، ارزیابی یا طبقه‌بندی شوند و امتیاز حاصل، به رفتار زیان‌بار یا نامطلوب نسبت به آنان منتهی شود؛ به‌ویژه هنگامی که این رفتار نامطلوب در زمینه‌ای نامرتبط با زمینه جمع‌آوری داده‌ها اعمال شود یا نسبت به رفتار اجتماعی شخص، ناموجه یا نامتناسب باشد . از همین رو، تعمیم ممنوعیت مقرر برای Social Scoring به تمامی سازوکارهای امتیازدهی، موجب توسعه دامنه ممنوعیت فراتر از مفاد قانون خواهد شد.

ضمناً سیستم‌های امتیازدهی اجتماعی که از هوش‌مصنوعی استفاده می‌کنند، در صورت ایجاد آسیب‌های جدی می‌توانند تحت تعریف «حادثه جدی» در فصل ۳ «قانون هوش‌مصنوعی» قراربگیرند. مثلاً اگر این سیستم‌ها باعث محرومیت افراد از خدمات ضروری درمان یا مسکن شوند، ممکن‌است به آسیب جسمی یا روانی (بند الف) منجرگردند. یا اگر اختلالی در دسترسی به زیرساخت‌های حیاتی مانند حمل‌ونقل یا انرژی ایجاد کنند (بند ب)، یا حقوق اساسی افراد مانند حریم خصوصی و برابری را نقض نمایند (بند ج). حتی آسیب‌های مالی ناشی از امتیازدهی

ناعادلانه، مثل ازدست‌دادن خانه یا شغل (بند د)، نیز ممکن‌است بعنوان یک حادثه جدی درنظرگرفته‌شود. در نتیجه، اگر این سیستم‌ها با هوش‌مصنوعی اجرا شوند و پیامدهای خطرناکی داشته باشند، قانون اتحادیه اروپا می‌تواند برای مقابله با آن‌ها اقدام‌کند.[1]

۳. راهکارهای مقابله با امتیازدهی اجتماعی مبتنی بر سوگیری الگوریتمی

کشورها رویکردهای متنوعی برای مقابله با تبعیض‌های ناشی از تصمیمات الگوریتمی اتخاذ کرده‌اند. برخی کشورها، مانند آلمان، استفاده از تصمیمات خودکار را به‌طور کلی محدود کرده و صرفاً در موارد استثنایی مجاز می‌دانند؛ در فرانسه، میان انواع مختلف تصمیم‌گیری الگوریتمی خودکار تمایز قائل می‌شوند؛ بریتانیا، اعمال تصمیمات خودکار را مشروط به رعایت استانداردها و مقررات ویژه کرده است؛ و در ایالات متحده و کانادا، شفافیت فرایند اتخاذ تصمیمات الگوریتمی به‌عنوان یکی از محورهای اصلی نظارت مورد توجه قرار گرفته است. این رویکردها نشان می‌دهند که مقابله با آثار تبعیض‌آمیز تصمیم‌گیری الگوریتمی لزوماً از طریق ممنوعیت تمامی اشکال استفاده از الگوریتم‌ها صورت نمی‌گیرد، بلکه با توجه به نوع کاربرد و آثار آن، از محدودیت و نظارت بر تصمیم‌گیری خودکار تا الزامات شفافیت و پاسخگویی را دربرمی‌گیرد. از این منظر، این تجربه‌ها را می‌توان در دو رویکرد کلی نیز دسته‌بندی کرد: برخی کشورها، به‌ویژه در اروپا، از ظرفیت قوانین حفاظت از داده‌های شخصی برای کنترل تصمیمات الگوریتمی استفاده کرده‌اند، در حالی که برخی دیگر، مانند کانادا و ایالات متحده، با تصویب قوانین و مقررات خاص، کنترل و نظارت بر تصمیمات الگوریتمی را تقویت کرده‌اند (Ansari, 2022, pp. 161-162).

از طرف دیگر باید دانست در دنیا، همه اشکال رتبه‌بندی یا ارزیابی افراد، امتیازدهی اجتماعیِ ممنوعه محسوب نمی‌شوند. مطابق رویکرد مبتنی بر خطر در قانون هوش مصنوعی اتحادیه اروپا، استفاده از سامانه‌های هوش مصنوعی برای ارزیابی یا طبقه‌بندی اشخاص حقیقی بر پایه رفتار اجتماعی یا ویژگی‌های شخصیِ شناخته‌شده، استنباط‌شده یا پیش‌بینی‌شده، هنگامی ممنوع است که به رفتار زیان‌بار یا نامطلوب با اشخاص یا گروه‌ها بینجامد و این رفتار یا با زمینه‌ای که داده‌ها در آن جمع‌آوری شده‌اند نامرتبط باشد یا ناموجه و نامتناسب تلقی شود.

در مقابل، کاربردهایی مانند اعتبارسنجی مالی، ارزیابی اعتبار اشخاص، استخدام، ارزیابی کارکنان، پذیرش آموزشی و برخی خدمات عمومی، در صورت برخورداری از شرایط مقرر، لزوماً ممنوع نیستند؛ بلکه به دلیل احتمال تأثیر شدید بر حقوق و فرصت‌های افراد، در زمره سامانه‌های پرخطر قرار می‌گیرند. بنابراین، پاسخ حقوقی به این دسته از کاربردها ممنوعیت مطلق نیست، بلکه اعمال الزامات پیشینی و مستمر، از جمله ارزیابی خطر، کیفیت داده‌ها، شفافیت، نظارت انسانی، ثبت رویدادها، امکان اعتراض و ممیزی الگوریتمی است.

بر این مبنا، پیشنهادمی‌شود در نظام حقوقی ایران نیز میان امتیازدهی اجتماعیِ سرکوبگرانه (که موجب کنترل فراگیر، طرد اجتماعی یا محروم‌سازی ناموجه اشخاص می‌شود) و سامانه‌های امتیازدهیِ پرخطر تمایز نهاده شود. دسته نخست باید صریحاً ممنوع و دسته دوم مشروط به تضمین‌های حقوقی و نظارتی متناسب باشد. (European Union, 2024, Art. 5; Annex III)

الف) الزامات شفافیت و پاسخگویی و حق توضیح‌پذیری[2]

یکی از مشکلات سیستم‌های هوش‌مصنوعی، جعبه سیاه بودن و عدم شفافیت آن‌هاست که کشف تبعیض را دشوارمی‌سازد. شفاف‌کردن نحوه عملکرد این سیستم‌ها هم می‌تواند به کشف علل و نتایج تبعیض‌آمیز کمک‌کرده و هم مقابله با آن‌ها را تسهیل‌کند(Ansari, 2022, p. 167). لذا برای جلوگیری از تبعیض، هم داده‌های آموزشی و هم الگوریتم‌ها باید شفاف باشند. قابلیت توضیح و امکان اعتماد به مدل (بـا اقداماتی مانند شفافیت و پاسخگوکردن مدل، توسعه روش‌های اجتناب از سوگیری، اصلاح و خودکنترلی مدل) اهمیت زیادی دارد. (گزارش مرکز پژوهش‌های مجلس شورای اسلامی، ۱۴۰۳، ص۷) مدل‌های مولد اغلب در رفتار و خروجی‌های خود مبهم و غیرقابل پیش‌بینی هستند و ممکن‌است محتوای نادرست، نامناسب یا مضر برای کاربران یا جامعه تولیدکنند یا حقوق مالکیت معنوی یا ارزش‌های اخلاقی تولیدکنندگان یا مصرف‌کنندگان محتوای اصلی

---

[1]. فصل ۳: «حادثه جدی» به معنای حادثه یا نقص عملکرد یک سیستم هوش‌مصنوعی است که به طور مستقیم یا غیرمستقیم منجر به هر یک از موارد زیر می‌شود:
الف) مرگ یک شخص، یا آسیب جدی به سلامت یک شخص؛
ب) اختلال جدی و غیرقابل برگشت در مدیریت یا عملیات زیرساخت‌های حیاتی؛
ج) نقض تعهدات تحت قانون اتحادیه که برای حفاظت از حقوق بنیادین در نظر گرفته شده است؛
د) آسیب جدی به اموال یا محیط زیست.»

[2]. Explainability

را نقض‌کنند. بنابراین، باید اطمینان حاصل‌شود که آن‌ها برای اعمال خود شفاف و پاسخگو هستند و به حقوق ذی‌نفعان احترام‌می‌گذارند. آنها همچنین باید توضیحاتی برای خروجی‌ها و مکانیسم‌های اصلاح و کنترل خود ارائه‌دهند[1]. باید به کمک استفاده از داده‌های آموزشی حسابرسی شده، منابع متنوع و بهره‌گیری از تکنیک‌های کاهش سوگیری، الگوریتم‌ها را به‌نحوی توسعه‌داد که سوگیری‌ها را کاهش‌دهند.(گزارش مرکز پژوهش‌های مجلس شورای اسلامی، ۱۴۰۳، ص ۲۲)

این راهکار از سوی برخی دولت‌ها از جمله کانادا و آمریکا دنبال‌شده‌است. دولت‌ها می‌توانند با وضع مقررات، امکان ممیزی و تبیین تمامی سیستم‌های الگوریتمی را در مراحل طراحی و توسعه فراهم‌کنند. (ص ۱۶۷) توسعه‌دهندگان سیستم‌های هوش مصنوعی را می‌شود ملزم‌کرد بخشی از اطلاعات کلیدی مربوط به الگوریتم‌های خود را به‌صورت عمومی ارائه‌دهند؛ برای مثال، مشخص‌کنند که الگوریتم‌ها از چه متغیرهایی بهره‌گرفته و برای دستیابی به چه اهدافی برنامه‌ریزی شده‌اند. با این حال، انتشار حجم وسیعی از داده‌ها لزوماً موجب افزایش شفافیت نمی‌شود و هدف، افشای تمامی داده‌ها نیست؛ بلکه تمرکز بر ارائه داده‌هایی است که بتوانند به درک بهتر عملکرد و کارایی سیستم کمک کنند. (Ansari, 2022, pp. 167-168).

«قانون هوش‌مصنوعی» هم در مواد متعددی، این راهکار را مقررکرده‌است؛ مثلا در فصل ۳: به «دستورالعمل‌های استفاده» اشاره‌می‌کند یعنی «اطلاعاتی که توسط ارائه‌دهنده برای اطلاع‌رسانی به بهره‌بردار در خصوص هدف موردنظر و استفاده صحیح از یک سیستم هوش‌مصنوعی ارائه می‌شود.» همچنین در ماده ۱۳ (شفافیت و ارائه اطلاعات به کاربران) مقررداشته: «۱. سیستم‌های هوش‌مصنوعی با ریسک بالا باید بگونه‌ای طراحی و توسعه یابند که عملکرد آن‌ها به حد کافی شفاف‌باشد به‌گونه‌ای که کاربران قادر باشند خروجی سیستم را درک‌کرده و از آن‌ها به‌شیوه‌ای مناسب بهره‌برداری نمایند. علاوه بر این نوع و درجه مناسبی از شفافیت باید تضمین‌شود... . ۲. سیستم‌های هوش‌مصنوعی با ریسک بالا باید همراه با دستورالعمل‌های استفاده (دفترچه راهنما) در یک قالب دیجیتالی مناسب یا بنحو دیگری ارائه‌شوند که شامل اطلاعاتی مختصر، کامل، صحیح و واضح باشد که برای کاربران مرتبط، قابل دسترسی و فهم باشد.(European Union, 2024, Art. 13)

همچنین ماده ۸۶ در مورد «حق توضیح در تصمیم‌گیری فردی» مقررمی‌کند: «۱. هر شخص متضرری که تصمیمی ... را بعنوان تأثیر منفی بر سلامت، ایمنی یا حقوق اساسی خود تلقی‌کند، حق‌دارد از مجری توضیحات واضح و معناداری در مورد نقش سیستم هوش‌مصنوعی در فرآیند تصمیم‌گیری و عناصر اصلی تصمیم اتخاذشده دریافت‌کند.»(European Union, 2024, Art.86)

در ماده ۵۰ (الزامات شفافیت برای تأمین‌کنندگان و کاربران برخی سیستم‌های هوش‌مصنوعی)، تأمین‌کنندگان ملزم شده‌اند که اطمینان حاصل کنند که سیستم‌های هوش‌مصنوعی که به‌طور مستقیم با افراد تعامل دارند، به‌گونه‌ای طراحی و توسعه یافته‌اند که به افراد اطلاع‌دهند که با سیستم هوش‌مصنوعی در حال تعامل هستند، مگر اینکه این موضوع از منظر فردی که به‌طور معقول آگاه، دقیق و محتاط است، با توجه به شرایط و زمینه استفاده، واضح باشد. ...»(European Union, 2024, Art. 50)

در ایران، تسهیل دسترسی به داده و ساماندهی و شفافیت مدیریت داده و اطلاعات با تکمیل یا ایجاد زیرساختهای قانونی مرتبط با داده امکانپذیر است، ازجمله تکمیل و بهبود «قانون مدیریت داده‌ها و اطلاعات ملی» (۱۴۰۱)، ایجاد ضوابط مرتبط با پردازش داده‌های شخصی افراد و استفاده اشخاص یا سامانه‌های هوش‌مصنوعی مولد از این داده‌ها در قالب تدوین طرح/ لایحه حفاظت و حمایت از داده‌های شخصی و حریم خصوصی، به‌روزرسانی و انطباق قوانین موجود با مصادیق و موضوعات هوش‌مصنوعی مولد ازجمله «قانون جرائم رایانه‌ای» (۱۳۸۸)، «قانون تجارت الکترونیک» (۱۳۸۲) و احکام مرتبط در قوانین مجازات اسلامی و مدنی، قوانین مرتبط با حمایت از مالکیت فکری در موضوعات هوش‌مصنوعی مولد- ازجمله «قانون حمایت از مالکیت صنعتی» (۱۴۰۰)، «قانون حمایت از حقوق پدیدآورندگان نرم‌افزارهای رایانه‌ای»

[1] . Zighra, (Nov.3th, 2023).”Explainable AI: The Key to Unlocking Generative AI’s Potential”, https://www.linkedin.com/pulse/explainable-ai-key-unlocking-generative-ais-potential-zighra-t2lre

(۱۳۷۹)، «قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان» (۱۳۴۸)- و ایجاد زیرساخت‌های قانونی ویژه هوش‌مصنوعی مولد درخصوص کنترل و سطح مسئولیت و خودمختاری. (گزارش مرکز پژوهش‌های مجلس شورای اسلامی، ۱۴۰۳، ص ۹)

ب) نظارت کافی

در راستای تضمین مشروعیت و پاسخگویی در فرایندهای امتیازدهی الگوریتمی، نظارت انسانی مؤثر یکی از ارکان بنیادین حکمرانی مسئولانه بر هوش‌مصنوعی است و شرط لازم برای پیشگیری از تصمیمات خودکار تبعیض‌آمیز یا فاقد شفافیت، و نیز تضمین‌کننده امکان مداخله انسانی در مواقع بروز خطا یا انحراف سیستمی محسوب‌می‌گردد. «قانون هوش‌مصنوعی» هم در ماده ۱۴، با تأکید بر ضرورت طراحی و به‌کارگیری سازوکارهای مؤثر نظارت انسانی، چارچوبی الزام‌آور برای کنترل تصمیمات سیستم‌های پرخطر ایجاد کرده‌است. همچنین فصل ۹ قانون با پیش‌بینی نظام مدیریت ریسک مستمر و به‌روزشونده، پیوندی منطقی میان نظارت انسانی و پایداری ایمنی در چرخه عمر سیستم‌های هوش‌مصنوعی برقرار می‌سازد. نظارت انسانی در سامانه‌های پرخطر نباید صوری و تشریفاتی باشد. شخص یا واحد ناظر باید از صلاحیت، اختیار و اطلاعات کافی برای بررسی تصمیم، توقف اجرای آن، اصلاح نتیجه یا ارجاع موضوع به مرجع مسئول برخوردار باشد. در حوزه‌هایی مانند اعتبارسنجی، استخدام و ارائه خدمات عمومی، صرف تأیید خودکار خروجی الگوریتم توسط انسان، بدون امکان ارزیابی مستقلِ داده‌ها و منطق تصمیم، تضمین مؤثری برای جلوگیری از تبعیض محسوب نمی‌شود. ازاین‌رو، پیش‌بینی ممیزی دوره‌ای الگوریتم، ثبت تصمیمات و امکان رسیدگی به اعتراض اشخاص ضروری است.

ج) کسب رضایت

ریسک‌های مرتبط با سیستم‌های تصمیم‌گیری خودکار که شامل داده‌های زیست‌سنجی هستند (مانند سیستم‌های تشخیص احساسات) عموماً قابل‌توجه است و به‌ویژه شامل تبعیض فرهنگی، تعصبات و انگ‌زنی می‌شود، زیرا الگوریتم‌ها اغلب به تفاوت‌های فرهنگی و اجتماعی توجه‌ندارند. این مسئله نیاز به داده‌های باکیفیت برای آموزش سیستم را برجسته‌می‌کند تا دقت نتایج بهبودیابد. در این راستا تضمین‌های مناسب، مانند مداخله انسانی (کسب رضایت آگاهانه) در زمینه امتیازدهی اهمیت ویژه‌ای دارد. یعنی در تصمیمات خودکار با هدف منفعت‌رساندن به فرد موردنظر (مانند درخواست اعتبار برای دریافت وام)، اخذ رضایت آزادانه[1] می‌تواند مشکل را تاحدی برطرف‌کند. (Comel, 2025, pp. 24-25)

د) مستعارسازی و ناشناس‌سازی

«مستعارسازي[2] به‌طورکلي به فرايندي اطلاق‌مي‌شود كه طي آن، داده‌هاي شخصي كه از طريق آنها مي‌توان به‌طور مستقيم افراد موضوع داده را شناسايي‌كرد، به‌گونه‌اي تغييرمي‌كنند كه اطلاعات هويتي از آنها جدامي‌شود و فقط با كمك يك نام مستعار مشخص و به‌شكل غيرمستقيم، به شناسايي اشخاص مرتبط مي‌انجامد.»(Esmaeili & Narimanpour, 2024, p.127) البته مستعارسازی، به‌تنهایی موجب خروج داده از قلمرو حمایت‌های حقوقی مربوط به داده‌های شخصی نمی‌شود و رعایت اصول حاکم بر پردازش داده‌ها همچنان ضرورت‌دارد. همچنین، میزان کارآمدی مستعارسازی از حیث حمایت از حریم خصوصی، به نحوه اجرای فنی و حقوقی آن و میزان امکان اتصال مجدد داده‌ها به هویت اشخاص وابسته‌است. از این منظر، مستعارسازی را نمی‌توان راهکاری مطلق برای رفع مخاطرات ناشی از پردازش داده‌های شخصی تلقی‌کرد و باید آن را در کنار سایر اصول و الزامات حفاظت از داده مورد ارزیابی قرار داد (Ghorban-Nia & Shakeri, 2025a).

---

[1]. همانطور که در ماده ۷(٤) GDPR آمده: «هنگام ارزیابی اینکه آیا رضایت به‌صورت آزادانه ارائه‌شده‌است یا نه، باید حداکثر توجه به این موضوع معطوف‌شود که آیا اجرای قرارداد یا ارائه خدمات و ...، منوط به دریافت رضایت برای پردازش داده‌های شخصی شده‌است که برای اجرای آن قرارداد ضروری نیست.» پس اگر برای ارائه یک خدمت به شما، شرط‌بگذارند که باید به جمع‌آوری و استفاده از اطلاعات شخصی‌تان که برای همان خدمت لازم نیست هم رضایت‌دهید، این رضایت شما «آزادانه» محسوب‌نمی‌شود.

[2]. Pseudonymization

در کنار مستعارسازی، ناشناس‌سازی[1] نیز یکی از راهکارهای مهم در حوزه حفاظت از داده‌های شخصی است؛ «ناشناس‌سازی فرایندی است که با حذف یا تغییر اطلاعات هویتی، امکان شناسایی مجدد افراد را کاهش می‌دهد و از این طریق، پردازش داده‌ها را با الزامات قانونی و اخلاقی سازگارتر می‌سازد. این فرایند شامل روش‌هایی همچون حذف شناسه‌ها، تعمیم، تصادفی‌سازی و رمزنگاری است. در مقابل، بازشناسایی[2] به فرایندی اطلاق‌می‌شود که طی آن، داده‌های ناشناس‌شده مجدداً به هویت واقعی افراد مرتبط می‌شوند.» (Idem, 2025b, p.67)

بااین‌حال، مطالعات انجام‌شده نشان می‌دهد که ناشناس‌سازی نیز با محدودیت‌های جدی مواجه‌است. وجود داده‌های مکمل درباره اشخاص و امکان ترکیب آن‌ها با داده‌های ناشناس‌شده می‌تواند خطر بازشناسایی افراد را افزایش‌دهد. ازاین‌رو، صرف حذف شناسه‌های مستقیم برای کاهش خطر شناسایی کافی‌نیست و میزان خطر بازشناسایی و شرایطی که داده در آن مورد استفاده قرارمی‌گیرد نیز باید مورد توجه قرارگیرد. بر همین اساس، رویکرد چندلایه در کاهش خطر بازشناسایی، بجای اتکای مطلق بر ناشناس‌سازی، برای سیاست‌گذاری در حوزه داده پیشنهادمی‌شود. این مسئله در مطالعات تطبیقی مربوط به حقوق داده نیز مورد توجه قرارگرفته‌است. در چارچوب «مقررهٔ عمومی حفاظت از داده‌ها» اتحادیه اروپا» هم، تعریف مستعارسازی به‌تنهایی معیار تعیین‌کننده برای تشخیص شخصی‌بودن یا نبودن داده نیست و برای تعیین قابلیت شناسایی شخص باید معیارهای مقرر در چارچوب این مقررات، از جمله وجود وسایلی که به‌طور معقول محتمل است برای شناسایی شخص مورد استفاده قرار گیرد، مورد توجه قرارگیرد. از این منظر، ارزیابی وضعیت داده صرفاً به ویژگی‌های خود داده محدودنمی‌شود و عواملی همچون اطلاعات تکمیلی، نحوه دسترسی به داده و کنترل‌های فنی و سازمانی حاکم بر محیط پردازش نیز می‌توانند در ارزیابی خطر شناسایی مؤثرباشند (Mourby et al., 2018). بنابراین، حفاظت از داده در سامانه‌های هوش مصنوعی را نمی‌توان صرفاً به حذف شناسه‌های مستقیم تقلیل داد؛ بلکه شرایط پردازش داده و امکان واقعی بازشناسایی نیز باید در ارزیابی حقوقی مورد توجه قرارگیردن**) تعمیم مقررات ضد تبعیض موجود به تبعیض‌های الگوریتمی**

دولت‌ها موظف‌اند از هرگونه رفتار تبعیض‌آمیز پرهیزکنند و همچنین مانع بروز چنین رفتارهایی در بخش خصوصی شوند. تبعیض‌های الگوریتمی به عنوان نمونه‌ای نوین و پیچیده از رفتارهای تبعیض‌آمیز، ایجاب می‌کند که دولت‌ها مکلف باشند ضمن شناسایی مصادیق و علل آن‌ها، تدابیر قانونی مناسب برای مقابله با این پدیده را پیش‌بینی کنند. تحقق این وظایف نیازمند اتخاذ ابتکارات حقوقی، اداری و قضایی است. بر این اساس، قوانین موجود مقابله با تبعیض تا حد امکان و مطابق با اصول تفسیر حقوقی شامل تبعیض‌های الگوریتمی نیز شوند و در صورت فقدان مقررات خاص، خلأهای قانونی با تصویب مقررات مناسب پر شوند. در حقوق ایران، مقررات متعددی بر ممنوعیت تبعیض تأکیددارند. قانون اساسی در اصول مختلف به این موضوع پرداخته‌است؛ از جمله اصل ۳ (بند ۹) که دولت را مکلف کرده رفع تبعیضات ناروا و ایجاد امکانات عادلانه برای همه در تمامی زمینه‌ها را تضمین نماید، و اصول ۱۹ و ۲۰ که تساوی افراد در برخورداری از حقوق و حمایت قانونی را پیش‌بینی کرده‌اند. همچنین، بند ۷ ماده ۸ قانون رسیدگی به تخلفات اداری (۱۳۷۲) «تبعیض یا اعمال غرض در اجرای قوانین و مقررات نسبت به اشخاص» را از مصادیق تخلفات اداری معرفی کرده‌است. ماده ۵ مصوبه شورای عالی اداری با عنوان «حقوق شهروندی در نظام اداری» (۱۳۹۵) نیز، مصادیق حق مصون ماندن از تبعیض در فرآیندها و تصمیمات اداری را مشخص کرده است، که شامل موارد زیر است:

۱. دستگاه‌های اجرایی موظف‌اند فرآیندها و رویه‌های ارائه خدمات خود را به‌صورت شفاف و یکسان برای همه مراجعان اعلام و اجرا کنند.

۲. مدیران و کارکنان دستگاه‌ها باید تصمیمات خود را مبتنی بر قوانین و مقررات اتخاذ کرده و از هرگونه تبعیض یا اعمال سلیقه شخصی خودداری کنند.

۳. کارکنان دستگاه‌ها در تمام سطوح باید در اعمال صلاحیت‌ها و اختیارات اداری، از جمله احراز صلاحیت‌ها، جذب نیرو و صدور مجوزها، بدون تبعیض عمل نمایند.

---

1. Anonymization
2. Re-identification

با وجود این مقررات و گسترش استفاده از سیستم‌های الگوریتمی، تاکنون هیچ مصوبه یا دستورالعمل خاصی برای مقابله با تبعیض‌های الگوریتمی تدوین نشده‌است.

قانون تجارت الکترونیک که تنها «قانون» مرتبط با حفاظت از داده‌های شخصی در ایران است، از نظر مقابله با تبعیض‌های الگوریتمی کارایی لازم را ندارد (Ansari, 2022, pp. 161-162). هرچند در حقوق ایران مقرره‌ای خاص درباره امتیازدهی اجتماعی مبتنی بر هوش مصنوعی وجود ندارد، برخی اصول حاکم بر روابط دیجیتال، از جمله حمایت از طرف ضعیف‌تر، می‌توانند مبنایی اولیه برای تحلیل این پدیده قرار گیرند(Sadeghi, 2025, p. 232)؛ اصلی که با توجه به نابرابری اطلاعاتی و موقعیت نامتقارن اشخاص در برابر بهره‌برداران سامانه‌های فناورانه، در محیط‌های دیجیتال اهمیتی مضاعف می‌یابد. در میان سایر اسناد تقنینی، «دستورالعمل اجرایی بهبود حفاظت از حریم خصوصی کاربران و شیوه جمع‌آوری، پردازش و نگهداری اطلاعات کاربران در سامانه‌ها و سکوهای فضای مجازی» (مصوب ۱۴۰۲) را باید مهم‌ترین سند تقنینی ایران در حوزه حفاظت از داده‌های شخصی و حریم خصوصی کاربران بشمار آورد. «این دستورالعمل با پیش‌بینی اصولی بنیادین همچون مشروعیت و تعیین هدف مشخص برای پردازش داده‌ها، تناسب و حداقل‌گرایی در جمع‌آوری اطلاعات، اخذ رضایت آگاهانه کاربران، اتخاذ تدابیر امنیتی محدودیت در نگهداری داده‌ها، گامی اساسی در جهت سامان‌دهی پردازش داده‌های شخصی محسوب‌می‌شود.» بااین‌حال، دامنه شمول این دستورالعمل صرفاً به حمایت از داده‌های شخصی و حریم خصوصی محدوداست و هیچ‌گونه مقرره ویژه‌ای درباره شناسایی، پیشگیری یا جبران تبعیض‌های الگوریتمی، بخصوص امتیازدهی اجتماعی، و یا ارزیابی سوگیری سامانه‌های هوش مصنوعی ارائه‌نمی‌دهد. (Ansari, Ibid).

این خلأ تقنینی نشان می‌دهد که سیاست‌گذاری در ایران نباید صرفاً به اعلام کلیِ ممنوعیت تبعیض یا اخذ رضایت کاربر محدود شود؛ بلکه نیازمند اقداماتی عینی و مرحله‌ای است. نخست، باید در لایحه حمایت از داده‌های شخصی یا قانون خاص هوش مصنوعی، امتیازدهی اجتماعیِ سرکوبگرانه و منتهی به طرد یا محرومیت ناموجه اشخاص، به‌عنوان کاربرد ممنوعه تعریف شود. در مرحله دوم، سامانه‌های امتیازدهی در حوزه‌های اعتبارسنجی، استخدام، بیمه، آموزش، اعطای مجوز و خدمات عمومی، به‌عنوان کاربردهای پرخطر شناسایی و مشمول ارزیابی اثرات حقوق بنیادین، ممیزی پیش از بهره‌برداری و ممیزی‌های دوره‌ای قرارگیرند. در گام سوم، حق دریافت توضیحات روشن و اطلاعات مؤثر در تصمیم، حق اعتراض و درخواست بازبینی انسانی برای اشخاص پیش‌بینی شود. بدیهی است اجرای این الزامات مستلزم تعیین مرجع ناظر دارای صلاحیت برای رسیدگی به شکایت‌ها، ارزیابی انطباق سامانه‌ها و اعمال ضمانت اجراهاست؛ زیرا بدون سازوکار نهادی نظارت، الزامات شفافیت و منع تبعیض قابلیت اجرای مؤثر نخواهند داشت؛ لذا این امر باید بعنوان مرحله چهارم مدنظر قرارگیرد.

برای بازنگری در سند ملی هوش‌مصنوعی، اتخاذ رویکرد بینابینی توسعه و تنظیم و ایجاد یک تعادل راهبردی مهم است (Abd Khodaee et al., 2021) لذا توسعه مقررات برای حاکمیت هوش‌مصنوعی مسئولیت‌پذیر، با تأکید بر هم‌افزایی و انسجام سیاستی این حوزه در ذیل دستگاه‌های اجرایی، از طریق ایجاد دستورالعمل‌ها و کدهای اخلاقی در بهبود بکارگیری اخلاقی داده‌ها، افزایش شفافیت و پاسخگویی، اعتبارسنجی هوش‌مصنوعی با توجه به ماهیت هر نوع بهره‌برداری (هوشمند یا غیر هوشمند، مخاطره‌آمیز یا رایج) و ارزیابی، ضروری است. (گزارش مرکز پژوهش‌های مجلس شورای اسلامی، ۱۴۰۳، ص۲۶)

نتیجه‌گیری

پژوهش حاضر با تمرکز بر مفهوم‌شناسی هوش‌مصنوعی و امتیازدهی اجتماعی، جایگاه پدیده اخیر در قانون هوش‌مصنوعی اتحادیه اروپا و پیامدهای آن برای نظام حقوقی ایران را بررسی کرد. یافته‌ها نشان‌می‌دهد که امتیازدهی اجتماعی، به‌رغم ظاهر کارآمد خود، واجد مخاطرات جدی در حوزه حقوق بنیادین است و می‌تواند به بازتولید تبعیض‌های ساختاری و نقض آزادی‌های اساسی منجرشود. براین اساس نتایج ذیل حاصل گردید:

۱.هوش‌مصنوعی از یک‌سو ظرفیت ارائه تصمیمات بی‌طرفانه و مبتنی بر داده‌های عینی را دارد، اما از سوی دیگر، در صورت آموزش بر داده‌های ناقص یا سوگیرانه، به تولید نتایج تبعیض‌آمیز منجرمی‌شود؛ لذا بی‌طرفی الگوریتمی یک فرض ساده‌انگارانه است و بدون نظارت حقوقی و اخلاقی، به بازتولید

تبعیض‌های اجتماعی منجرخواهدشد. الگوریتم‌ها بر اساس تمایزگذاری میان افراد عمل‌می‌کنند که درصورت فقدان داده‌های متنوع و باکیفیت، به تبعیض سیستماتیک بدل‌می‌شود.

۲. امتیازدهی اجتماعی فرآیندی است که در آن رفتارها و ویژگی‌های انسانی به داده‌های کمی تبدیل و بر اساس آن، رتبه یا امتیاز فرد تعیین می‌گردد. امتیازدهی و ارزیابی الگوریتمی ممکن است در حوزه‌های مختلفی مانند اعتبارسنجی مالی، استخدام، بیمه، خدمات اجتماعی و ارزیابی برخی خطرات مورد استفاده قرار گیرد؛ بااین‌حال، صرف استفاده از سازوکار امتیازدهی در این حوزه‌ها به معنای تحقق مفهوم «امتیازدهی اجتماعی» ممنوع در قانون هوش‌مصنوعی اتحادیه اروپا نیست و وضعیت حقوقی هر کاربرد باید با توجه به نوع، هدف، زمینه استفاده و آثار آن ارزیابی شود. این سازوکارها در صورت استفاده نامتناسب یا سوگیرانه می‌توانند به بازتولید تبعیض و محدود شدن حقوق افراد منجر شوند؛ به‌ویژه زمانی که شهروندان بدون اطلاع کافی از فرایند امتیازدهی، در معرض تصمیمات خودکار قرار می‌گیرند، امکان اعمال حقوقی مانند دسترسی به داده‌ها و اعتراض نیز می‌تواند با چالش مواجه شود.

۳. »قانون هوش‌مصنوعی» اتحادیه اروپا به دلیل نگرانی‌های جدی نسبت به نقض اصل برابری، حریم خصوصی و آزادی انتخاب، با اتخاذ سیاستی پیشگیرانه، امتیازدهی اجتماعی را - البته تحت شرایطی- در زمرهٔ کاربردهای ممنوعه قرارداده‌است و سوگیری الگوریتمی را به‌عنوان یکی از ریسک‌های سیستماتیک به رسمیت شناخته‌است. البته طبق رأی صادره در پرونده شوفا، حتی در غیاب تصریح قانونی، رویه‌های امتیازدهی اجتماعی می‌توانند ذیل ماده ۲۲ مقرره عمومی حفاظت از داده‌ها قرار گیرند و مشمول محدودیت‌های سختگیرانه شوند.

۴. در مواجهه با سامانه‌های پرخطر، از جمله سازوکارهایی مانند شفافیت، توضیح‌پذیری، نظارت انسانی مؤثر، ممیزی الگوریتمی، ثبت رویدادها، ارزیابی اثرات حقوق بنیادین و حق اعتراض اشخاص، کسب رضایت آگاهانه، مستعارسازی و ناشناس‌سازی، نقش تضمینی دارند و می‌توانند از تبدیل امتیازدهی به ابزار تبعیض و محروم‌سازی جلوگیری کنند. صرف اتکای صوری به تصمیم انسان در کنار خروجی الگوریتم، بدون اختیار واقعی برای بازبینی و اصلاح، کفایت ندارد.

۵. در حقوق داخلی ایران، با وجود برخی اسناد ملی در حوزه هوش‌مصنوعی، مقررات مشخصی برای امتیازدهی اجتماعی وجودندارد. این خلأ می‌تواند زمینه‌ساز گسترش تبعیض‌های الگوریتمی و نقض حقوق بنیادین شهروندان شود. ازاین‌رو، لازم است با تکمیل مقررات موجود بخصوص لایحه حمایت از داده‌های شخصی، میان کاربردهای ممنوع و پرخطر امتیازدهی تفکیک شود و الزامات ناظر بر ارزیابی سوگیری، شفافیت، حق اعتراض، بازبینی انسانی و نظارت مؤثر بر این سامانه‌ها پیش‌بینی گردد.

فهرست منابع

الف- منابع فارسی

۱. قانون اساسی (۱۳۵۸).
۲. قانون رسیدگی به تخلفات اداری (۱۳۷۲).
۳. سند ملی هوش‌مصنوعی (۱۴۰۳).
۴. مصوبه شورای‌عالی اداری درباره حقوق شهروندی در نظام اداری (۱۳۹۵).
۵. اسماعیلی، محسن، و مهدی نریمان‌پور. (۱۴۰۳). مبانی حقوقی تبادل داده‌های شخصی (مطالعه تطبیقی در مقررات عمومی حفاظت از داده اتحادیه اروپا و حقوق ایران). حقوق اسلامی، ۲۱(۸۲).
۶. انصاری، باقر. (۱۴۰۱). مطالعه حقوقی تبعیض الگوریتمی. دانش حقوق عمومی، ۱۱(۳۸)، ۱۴۱-۱۷۰.
۷. خدایی، زهره، توحیدی، نسیم، کریمی واقف، نرگس و براتی جوآبادی، محمد. (۱۴۰۰). هوش سفید: از اخلاق ماشین تا ماشین اخلاق‌مند. مؤسسه مطالعات فرهنگی و اجتماعی.
۸. خردمندنیا، سهیلا، و تقی‌زاده، مسلم (۱۴۰۳). هوش‌مصنوعی مولد: چالش‌ها و الزامات برای توسعه و پیاده‌سازی. مرکز پژوهش‌های مجلس، دفتر مطالعات انرژی، صنعت و معدن.
۹. سلحشور، عباس، توکلی، احمدرضا، و حیدری، محمدعلی. (۱۴۰۴). حوزه‌های بهره‌مندی هوش مصنوعی در حقوق بنیادین افراد در پرتو بررسی جهات مشروعیت آن. دانشنامه فقه و حقوق تطبیقی، ۳(۱).
۱۰. صادقی، محسن. (۱۴۰۴). نقدی بر آراء دادگاه‌های حقوقی ایران از منظر اصول کلی حقوق تجارت الکترونیک. نقد و تحلیل آراء قضایی، ۴(۸)، ۲۱۰-۲۴۰. Doi: 10.22034/jpl.2025.2070457.1237

۱۱. صادقی، محسن، و ملک‌زاده قلعه، امیر. (۱۴۰۵). بازاندیشی مبانی مالکیت خصوصی در بستر دارایی‌های دیجیتال. تحقیق و توسعه در حقوق خصوصی، ۳(۵)، ۲۷۸-۳۲۳. Doi: 10.22034/analysis.2025.2076273.1136

۱۲. قربان‌نیا، امیرمحمد، و شاکری، زهرا. (۱۴۰۴). اهمیت و چالش‌های حقوقی مستعارسازی داده‌ها؛ با نگاهی به اصول پردازش داده‌های شخصی. پژوهش‌های حقوق اقتصادی و تجاری، ۳(۲)، ۷۱-۱۰۰.

۱۳. قربان‌نیا، امیرمحمد، و شاکری، زهرا. (۱۴۰۴). چالش‌های حقوقی ناشناس‌سازی و بازشناسایی داده‌های شخصی. بررسی‌های بازرگانی، ۲۳(۱۳۲)، ۶۵-۸۶.

۱۴. مشرفیان، محمدرضا. (۱۴۰۴). کنوانسیون چارچوب شورای اروپا در مورد هوش مصنوعی و حقوق بشر، دموکراسی و حاکمیت قانون. مجله حقوقی بین‌المللی، پیاپی ۷۷، ۱۳۷-۱۵۱.

۱۵. مصطفوی اردبیلی، سیدمحمدمهدی، تقی‌زاده انصاری، مصطفی، و رحمت‌فر، سمانه (۱۴۰۱). کارکردها و بایسته‌های هوش‌مصنوعی ازمنظر دادرسی منصفانه. حقوق فناوری‌های نوین، ۳(۶)، ۴۷-۶۰.

۱۶. مصطفوی اردبیلی، سیدمحمدمهدی، تقی‌زاده انصاری، مصطفی، و رحمت‌فر، سمانه. (۱۴۰۲). تأثیر هوش‌مصنوعی بر نظام حقوق‌بشر بین‌الملل. حقوق فناوری‌های نوین، ۴(۸)، ۸۵-۱۰۰.

ب- منابع انگلیسی

1. Abid, A., Farooqi, M., & Zou, J. (2021). Persistent anti-Muslim bias in large language models. arXiv. https://doi.org/10.48550/arXiv.2101.05783
2. Access Now. (2023). Extreme vetting and the future of immigration control. https://www.accessnow.org
3. Article 29 Data Protection Working Party. (2018). Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (wp251rev.01). https://ec.europa.eu/newsroom/article29/items/
4. Backer, L. C. (2017, June). Measurement, assessment and reward: The challenges of building institutionalized social credit and rating systems in China and in the West. Paper presented at The Chinese Social Credit System 2017 Conference, Shanghai, China.
5. Batte, B. (2025). Algorithmic borders: The role of artificial intelligence in immigration control and refugee protection. University of the Cumberlands. https://www.researchgate.net/publication/301600648
6. Betzalel, E., Penso, C. Navon, A. & Fetaya, E. (June.22th, 2022)."A study on the Evaluation of Generative Models", Retrieved from https://arxiv.org/abs/2206.10935
7. Bhardwaj, G., Singh, S. V., & Kumar, V. (2020). An empirical study of artificial intelligence and its impact on human resource functions. In 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 47-51). IEEE.
8. Brennan, D., Fry, W., & O'Flynn, C (2018). Artificial intelligence in the workplace (part 3): ai and gender equality. https://www.williamfry.com/newsandinsights/news-article/2018/08/01/artificial-intelligence-in-theworkplace-(part-3)-ai-and-gender-equality
9. C., Stanton, N. A., & Salmon, P. M. (2021). The risks associated with artificial general intelligence: A systematic review. Journal of Experimental & Theoretical Artificial Intelligence, 35(5), 649-663. https://doi.org/10.1080/0952813X.2021.1964003
10. Case C-634/21 SCHUFA Holding (Scoring) [2023] ECLI:EU:C:2023:957
11. Cobbe, J. (2021). Administrative law and the machines of government: Judicial review of automated public-sector decision-making. Legal Studies, 39(4), 636-655.
12. Cofone, I. N. (2019). Algorithmic discrimination is an information problem. Hastings Law Journal, 70(6), 1389-1406. https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/1
13. Comel, L. (2025). Scoring systems in the General Data Protection Regulation and Artificial Intelligence Act: A critical appraisal (MCEL Master's Thesis Series No. 2025/03). Maastricht University, Faculty of Law, Maastricht Centre for European Law.
14. Committee of experts on internet intermediaries (2018), Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-onthe-human-rights-dimensions-of-automated-data-processing-techniques-andpossible-regulatory-implications.html.
15. Convention concerning Discrimination in respect of Employment and Occupation. https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C111

16. Cossette-Lefebvre, H., & Maclure, J. (2022). AI's fairness problem: understanding wrongful discrimination in the context of automated decision-making. AI and Ethics, 3(4), 1255-1269.
17. Court of Justice of the European Union. (2023). Case C-634/21 SCHUFA Holding (Scoring). ECLI:EU:C:2023:957. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=ECLI:EU:C:2023:957
18. Dencik, L., Redden, J., Hintz, A., & Warne, H. (2019). The "golden view": Data-driven governance in the scoring society. Internet Policy Review, 8(2), 1-24. Retrieved July 20, 2024, from https://policyreview.info/node/1413
19. Duncan Minty, The EU's Social Scoring Ban - What Does It Mean?, Oct 23, 2024, https://www.ethicsandinsurance.info/the-eus-social-scoring-ban-what-does-it-mean/
20. European Commission, 2019, High-Level Expert Group on Artificial Intelligence, https://www.aepd.es/sites/default/files/2019-12/aidefinition.pdf.
21. European Commission, Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts,Brussels, 21.4.2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.
22. Ferrer Aran, X., van Nuenen, T., Such, J., Coté, M., & Criado Pacheco, N. (2021). Bias and discrimination in AI: A cross-disciplinary perspective. IEEE Technology and Society Magazine, 40(2), 72-80. https://doi.org/10.1109/MTS.2021.3056293
23. Genicot, N. (2025). Scoring the European citizen in the AI era. Computer Law & Security Review, 57, 106130. , https://doi.org/10.1016/j.clsr.2025.106130. https://www.sciencedirect.com/science/article/abs/pii/S2212473X25000033
24. Gerards, J., & Xenidis, R. (2021). Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. European Commission. https://doi.org/10.2838/544956
25. Heinrichs, B. (2021). Discrimination in the Age of Artificial Intelligence. Ethics and Information Technology, 37, pp. 143 - 154.
26. High-Level Expert Group on Artificial Intelligence. (2019). Ethics Guidelines for Trustworthy AI. European Commission.
27. Hinton, G.E, Osindero, S.& Whye The, Y. (2006). "A Fast-Learning Algorithm for Deep Belief Nets", Retrieved from https://www.cs.toronto.edu/~hinton/absps/ncfast.pdf .
28. HiPeople. (2023). Candidate scoring. HiPeople. Retrieved July 20, 2024, from https://www.hipeople.io/glossary/candidate-scoring
29. https://ai-act-law.eu/recital/31/
30. https://www.holisticai.com/blog/prohibitions-under-eu-ai-act
31. Javier Sánchez-Monedero and Lina Dencik, 'The Datafication of the Workplace' (Data Justice Lab, Cardiff University 2019) Working paper; Wolfie Christl, 'Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions' (Cracked Labs 2017) <https://crackedlabs.org/en/corporate-surveillance> accessed 14 January 2025.
32. Kaye, David, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment, 29 August 2018, https://www.undocs.org/A/73/348.https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportGA73.aspx.
33. Law Library of Congress (2019), Regulation of Artificial Intelligence in Selected Jurisdictions, January 2019, pp. 16-132. https://www.loc.gov/law/help/artificial-intelligence/regulation-artificialintelligence.pdf. Ko¨chling, Alina, Marius Claus Wehner, Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HRdevelopment,https://link.springer.com/article/10.1007/s40685--020-00134-w.
34. Lina Dencik and others, 'Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services. Project Report' [2018] Data Justice Lab Cardiff University 10.
35. Louise Amoore, The Politics of Possibility: Risk and Security beyond Probability (Duke University Press 2013).
36. Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. Big Data & Society, 6(1), 205395171989580. https://doi.org/10.1177/205395171989580
37. Mikella Hurley and Julius Adebayo, 'Credit Scoring in the Era of Big Data' [2016] Yale Journal of Law and Technology 148.
38. Moore, J. (Aprill 27th, 2023). "7 Generative AI Challenges that Businesses should consider", Retrieved from https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-challenges-that-businesses-should-consider.
39. Mourby, M., Mackey, E., Elliot, M., Gowans, H., Wallace, S. E., Bell, J., Smith, H., Aidinlis, S., & Kaye, J. (2018). Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. Computer Law & Security Review, 34(2), 222-233.

40. R (JCWI) v. Home Office. (2020). [UK visa algorithm case].Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) [2024] OJ L.
41. Rovatsos, Michael, Brent Mittelstadt and Ansgar Koene (2019), Landscape Summary: Bias in Algorithmic Decision-Making, The Centre for Data Ethics and Innovation.
42. Wang, J., Li, H., Xu, W., & Xu, W. (2023). Envisioning a credit society: Social credit systems and the institutionalization of moral standards in China. Media, Culture & Society, 45(3), 451-470. https://doi.org/10.1177/01634437221127364
43. Wiltrud Weidner, Fabian WG Transchel and Robert Weidner, 'Telematic Driving Profile Classification in Car Insurance Pricing' (2017) 11 Annals of Actuarial Science.
44. Xenidis, Raphaël e and Linda Senden (2020), EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination in Ulf Bernitz et al (eds), General Principles of EU law and the EU Digital Order (Kluwer Law International), https://cadmus.eui.eu/handle/1814/65845.
45. Zuiderveen Borgesius, Frederik (2018), Discrimination, Artificial Intelligence, and Algorithmic Decision-making. Council of Europe, https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmicdecision-making/1680925d73

ج- منابع فرانسوی

1. Dubois, V. (2021). Contrô ler les assisté s : Genè ses et usages d'un mot d'ordre. Paris. Raisons d'agir.
2. Institut Montaigne (2020), Algorithmes: contrôle des biais S.V.P., Rapport Mars 2020, pp.17-18. https://www.institutmontaigne.org/ressources/pdfs/publications/algorithmescontrole-des-biais-svp.pdf.
3. Lucchesi, Laure, Simon Chignard (2019), Rapport-Ethique et responsabilité des algorithmes publics,Juin2019.https://www.etalab.gouv.fr/wpcontent/uploads/2020/01/RapportENAEthique-et-responsabilité -des-algorithmes-publics.pdf .

د- منبع ایتالیایی

1. Garante per la Protezione dei Dati Personali. (2022, June 8). "Cittadinanza a punti": Garante privacy ha avviato tre istruttorie. Preoccupanti i meccanismi di scoring che premiano i cittadini "virtuosi".

Refrences

1. Abd Khodaee, Z., Tohidi, N., Karimi Vagef, N., & Barati Jowabadi, M. (2021). White Intelligence: From Machine Ethics to Ethical Machine. Institute for Cultural and Social Studies.(in Persian)
2. Abid, A., Farooqi, M., & Zou, J. (2021). Persistent anti-Muslim bias in large language models. arXiv. https://doi.org/10.48550/arXiv.2101.05783
3. Access Now. (2023). Extreme vetting and the future of immigration control. https://www.accessnow.org
4. Ansari, B. (2023). Algorithmic discrimination: Legal analysis. Journal of Public Law Knowledge, 11(38), 147-178.(in Persian)
5. Article 29 Data Protection Working Party. (2018). Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (wp251rev.01). https://ec.europa.eu/newsroom/article29/items/
6. Backer, L. C. (2017, June). Measurement, assessment and reward: The challenges of building institutionalized social credit and rating systems in China and in the West. Paper presented at The Chinese Social Credit System 2017 Conference, Shanghai, China.

7. Batte, B. (2025). Algorithmic borders: The role of artificial intelligence in immigration control and refugee protection. University of the Cumberlands. https://www.researchgate.net/publication/301600648
8. Betzalel, E., Penso, C. Navon, A. & Fetaya, E. (June.22th, 2022)."A study on the Evaluation of Generative Models", Retrieved from https://arxiv.org/abs/2206.10935
9. Bhardwaj, G., Singh, S. V., & Kumar, V. (2020). An empirical study of artificial intelligence and its impact on human resource functions. In 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 47-51). IEEE.
10. Brennan, D., Fry, W., & O'Flynn, C (2018). Artificial intelligence in the workplace (part 3): ai and gender equality. https://www.williamfry.com/newsandinsights/news-article/2018/08/01/artificial-intelligence-in-theworkplace-(part-3)-ai-and-gender-equality
11. C., Stanton, N. A., & Salmon, P. M. (2021). The risks associated with artificial general intelligence: A systematic review. Journal of Experimental & Theoretical Artificial Intelligence, 35(5), 649-663. https://doi.org/10.1080/0952813X.2021.1964003
12. Case C-634/21 SCHUFA Holding (Scoring) [2023] ECLI:EU:C:2023:957
13. Cobbe, J. (2021). Administrative law and the machines of government: Judicial review of automated public-sector decision-making. Legal Studies, 39(4), 636-655.
14. Cofone, I. N. (2019). Algorithmic discrimination is an information problem. Hastings Law Journal, 70(6), 1389-1406. https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/1
15. Comel, L. (2025). Scoring systems in the General Data Protection Regulation and Artificial Intelligence Act: A critical appraisal (MCEL Master's Thesis Series No. 2025/03). Maastricht University, Faculty of Law, Maastricht Centre for European Law.
16. Committee of experts on internet intermediaries (2018), Study on the Human Rights Dimensions of Automated Data Processing Techniques (in particular algorithms) and possible regulatory implications, https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-onthe-human-rights-dimensions-of-automated-data-processing-techniques-andpossible-regulatory-implications.html.
17. Constitution of the Islamic Republic of Iran. (1979). Available at: https://www.refworld.org/docid/3ae6b56710.htm
18. Convention concerning Discrimination in respect of Employment and Occupation. https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_ILO_CODE:C111
19. Cossette-Lefebvre, H., & Maclure, J. (2022). AI's fairness problem: understanding wrongful discrimination in the context of automated decision-making. AI and Ethics, 3(4), 1255-1269.
20. Court of Justice of the European Union. (2023). Case C-634/21 SCHUFA Holding (Scoring). ECLI:EU:C:2023:957. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=ECLI:EU:C:2023:957
21. Dencik, L., Redden, J., Hintz, A., & Warne, H. (2019). The "golden view": Data-driven governance in the scoring society. Internet Policy Review, 8(2), 1-24. Retrieved July 20, 2024, from https://policyreview.info/node/1413
22. Dubois, V. (2021). Contrô ler les assisté s : Genè ses et usages d'un mot d'ordre. Paris. Raisons d'agir.(in French)
23. Duncan Minty, The EU's Social Scoring Ban - What Does It Mean?, Oct 23, 2024, https://www.ethicsandinsurance.info/the-eus-social-scoring-ban-what-does-it-mean/
24. Esmaeili, M., & Narimanpour, M. (2024). Legal Basis of Personal Data Exchange (Comparative Study of EU General Data Protection Regulations and Iranian Law). Islamic Law, 21(82), 131-176.(in Persian)
25. European Commission, 2019, High-Level Expert Group on Artificial Intelligence, https://www.aepd.es/sites/default/files/2019-12/aidefinition.pdf.
26. European Commission, Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts,Brussels, 21.4.2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206.
27. Ferrer Aran, X., van Nuenen, T., Such, J., Coté, M., & Criado Pacheco, N. (2021). Bias and discrimination in AI: A cross-disciplinary perspective. IEEE Technology and Society Magazine, 40(2), 72-80. https://doi.org/10.1109/MTS.2021.3056293
28. Garante per la Protezione dei Dati Personali. (2022, June 8). "Cittadinanza a punti": Garante privacy ha avviato tre istruttorie. Preoccupanti i meccanismi di scoring che premiano i cittadini "virtuosi".(in Italian)

29. Genicot, N. (2025). Scoring the European citizen in the AI era. Computer Law & Security Review, 57, 106130. , https://doi.org/10.1016/j.clsr.2025.106130. https://www.sciencedirect.com/science/article/abs/pii/S2212473X25000033
30. Gerards, J., & Xenidis, R. (2021). Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. European Commission. https://doi.org/10.2838/544956
31. Ghorban Nia, A., & Shakeri, Z. (2025). The importance and legal challenges of data pseudonymization: Considering the principles of personal data processing Economic and Commercial Law Researches, 3(2), 71-100.(in Persian)
32. Ghorbannia, A. M., & Shakeri, Z. (2026). Legal challenges of anonymization and re-identification of personal data. Commercial Surveys, 23(132), 65-86.(in Persian)
33. Heinrichs, B. (2021). Discrimination in the Age of Artificial Intelligence. Ethics and Information Technology, 37, pp. 143 - 154.
34. High-Level Expert Group on Artificial Intelligence. (2019). Ethics Guidelines for Trustworthy AI. European Commission.
35. Hinton, G.E, Osindero, S.& Whye The, Y. (2006). "A Fast-Learning Algorithm for Deep Belief Nets", Retrieved from https://www.cs.toronto.edu/~hinton/absps/ncfast.pdf .
36. HiPeople. (2023). Candidate scoring. HiPeople. Retrieved July 20, 2024, from https://www.hipeople.io/glossary/candidate-scoring
37. https://ai-act-law.eu/recital/31/
38. https://www.holisticai.com/blog/prohibitions-under-eu-ai-act
39. Institut Montaigne (2020), Algorithmes: contrôle des biais S.V.P., Rapport Mars 2020, pp.17-18. https://www.institutmontaigne.org/ressources/pdfs/publications/algorithmescontrole-des-biais-svp.pdf.(in French)
40. Javier Sánchez-Monedero and Lina Dencik, 'The Datafication of the Workplace' (Data Justice Lab, Cardiff University 2019) Working paper; Wolfie Christl, 'Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions' (Cracked Labs 2017) <https://crackedlabs.org/en/corporate-surveillance> accessed 14 January 2025.
41. Kaye, David, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (2018), Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment, 29 August 2018, https://www.undocs.org/A/73/348.https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportGA73.aspx.
42. Khordmandnia, S., & Taghizadeh, M. (2024). Generative artificial intelligence: Challenges and requirements for development and implementation. Research Center of the Islamic Consultative Assembly, Office for Energy, Industry, and Mining Studies.(in Persian)
43. Law Library of Congress (2019), Regulation of Artificial Intelligence in Selected Jurisdictions, January 2019, pp. 16-132. https://www.loc.gov/law/help/artificial-intelligence/regulation-artificialintelligence.pdf. Ko¨chling, Alina, Marius Claus Wehner, Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HRdevelopment,https://link.springer.com/article/10.1007/s40685--020-00134-w.
44. Law on Administrative Violations. (1993). Islamic Republic of Iran.(in Persian)
45. Lina Dencik and others, 'Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services. Project Report' [2018] Data Justice Lab Cardiff University 10.
46. Louise Amoore, The Politics of Possibility: Risk and Security beyond Probability (Duke University Press 2013).
47. Lucchesi, Laure, Simon Chignard (2019), Rapport-Ethique et responsabilité des algorithmes publics,Juin2019.https://www.etalab.gouv.fr/wpcontent/uploads/2020/01/RapportENAEthique-et-responsabilité-des-algorithmes-publics.pdf (in French).
48. Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. Big Data & Society, 6(1), 205395171989580. https://doi.org/10.1177/205395171989580
49. Mikella Hurley and Julius Adebayo, 'Credit Scoring in the Era of Big Data' [2016] Yale Journal of Law and Technology 148.
50. Moore, J. (Aprill 27th, 2023). "7 Generative AI Challenges that Businesses should consider", Retrieved from https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-challenges-that-businesses-should-consider.
51. Moshrefian, M. R. (2025). Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. International Legal Journal, (77), 137-151.(in Persian)
52. Mostafavi Ardebili, M., Taghizadeh Ansari, M., & Rahmatifar, S. (2023). The impact of artificial intelligence on the international human rights system. Modern Technologies Law, 4(8), 85-100.(in Persian)

53. Mostafavi Ardebili, S. M. M., Taghizadeh Ansari, M., & Rahmatifar, S. (2022). Functions and Requirements of Artificial Intelligence in the Light of Fair Trial. Modern Technologies Law, 3(6), 47-60.(in Persian)
54. Mourby, M., Mackey, E., Elliot, M., Gowans, H., Wallace, S. E., Bell, J., Smith, H., Aidinlis, S., & Kaye, J. (2018). Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK. Computer Law & Security Review, 34(2), 222-233.
55. National Council for Cultural Revolution & Islamic Republic of Iran. (2024). National Artificial Intelligence Document of the Islamic Republic of Iran [National AI Document of the Islamic Republic of Iran] (Approved in session No. 901, June 18, 2024). Retrieved from https://rc.majlis.ir/fa/law/show/1811432 (in Persian)
56. R (JCWI) v. Home Office. (2020). [UK visa algorithm case].Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) [2024] OJ L.
57. Rovatsos, Michael, Brent Mittelstadt and Ansgar Koene (2019), Landscape Summary: Bias in Algorithmic Decision-Making, The Centre for Data Ethics and Innovation.
58. Sadeghi, M. (2025). A critique of judgments rendered by the Iranian civil courts from the perspective of general principles governing e-commerce law. Journal of Critical Analysis of Judicial Decisions, 4(8), 210-240. Doi: 10.22034/analysis.2025.2076273.1136 (in Persian)
59. Sadeghi, M., & Malekzadeh Qhaleh, A. (2026). Rethinking the foundations of private ownership in the context of digital assets. Journal of Research and Development in Private Law, 3(5), 278-323. Doi: 10.22034/jpl.2025.2070457.1237 (in Persian)
60. 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.(in Persian)
61. Supreme Administrative Council. (2016). Resolution on Citizenship Rights in the Administrative System. Islamic Republic of Iran.(in Persian)
62. Wang, J., Li, H., Xu, W., & Xu, W. (2023). Envisioning a credit society: Social credit systems and the institutionalization of moral standards in China. Media, Culture & Society, 45(3), 451-470. https://doi.org/10.1177/01634437221127364
63. Wiltrud Weidner, Fabian WG Transchel and Robert Weidner, 'Telematic Driving Profile Classification in Car Insurance Pricing' (2017) 11 Annals of Actuarial Science.
64. Xenidis, Raphaël e and Linda Senden (2020), EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination in Ulf Bernitz et al (eds), General Principles of EU law and the EU Digital Order (Kluwer Law International), https://cadmus.eui.eu/handle/1814/65845.
65. Zuiderveen Borgesius, Frederik (2018), Discrimination, Artificial Intelligence, and Algorithmic Decision-making. Council of Europe, https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmicdecision-making/1680925d73