پژوهش‌نامه حقوق اسلامی

مقاله پژوهشی | دریافت: ۱۴۰۴/۰۲/۰۲ | پذیرش: ۱۴۰۴/۰۴/۱۹
DOI: 10.30497/law.2025.248106.3746

تحلیل دکترین مسئولیت نیابتی به‌مثابه مبنای مسئولیت ناشی از هوش مصنوعی در حقوق ایران، با الهام از رویه قضایی کامن‌لا

عباس شیری • دانشیار، گروه حقوق جزا و جرم‌شناسی، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.
ashiri@ut.ac.ir
محمدرضا برزگر • دانشجوی دکتری حقوق جزا و جرم‌شناسی، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.
(نویسنده مسئول) mrbarzegar@ut.ac.ir

چکیده

هوش مصنوعی، نظام‌های حقوقی در سراسر جهان و از جمله نظام حقوقی ایران را با مسائل نوینی مواجه ساخته که شاید برجسته‌ترین آن، نحوه انتساب مسئولیت ناشی از خروجی‌های هوش مصنوعی و رفع «خلاء مسئولیت» ناشی از آن است. پژوهش حاضر با هدف ارائه یک چارچوب حقوقی کارآمد، به تحلیل ظرفیت دکترین مسئولیت نیابتی برای پاسخگویی به این مسأله می‌پردازد. هدف اصلی، تبیین مبانی موجِّهه اعمال این دکترین بر اشخاص به‌کارگیرنده سامانه‌های هوشمند، و امکان‌سنجی انطباق عناصر ماهوی آن با ویژگی‌های منحصربه‌فرد هوش مصنوعی خواهد بود. در این راستا، نوشتار حاضر با روش توصیفی تحلیلی، مبانی و عناصر این دکترین را در پرتو ویژگی‌های هوش مصنوعی صورت‌بندی کرده و با اصول مشابه در حقوق ایران تطبیق می‌دهد. تحلیل مبانی مسئولیت نیابتی مؤید آن است که اصولی چون ظرفیت برتر اقتصادی اصیل جهت جبران خسارت بزه‌دیده و مدیریت ریسک، منطق حقوقی اقتصادی ایجاد ریسک توسط شرکت استفاده‌کننده و ضرورت درونی‌سازی هزینه‌ها، و همچنین ماهیت ابزاری و عملکرد ادغام‌شده هوش مصنوعی در فعالیت اصیل و کنترل راهبردی وی بر آن، توجیه‌پذیری اعمال مسئولیت نیابتی در این حوزه را مدلل می‌سازد. بر این اساس، عناصر دکترین مذکور مورد بررسی قرار گرفته و استدلال می‌شود که معیار کامن‌لایی «رابطه شبیه به استخدام» به‌نحو مؤثری قادر به تبیین پیوند حقوقی میان شخص اصیل و سامانه هوشمند است، و معیار «ارتباط نزدیک» اتصال میان فعل مجرمانه ناشی از سیستم و رابطه مذکور را برقرار می‌سازد. نهایتاً، این نتیجه حاصل شده است که دکترین مسئولیت نیابتی، چارچوبی منسجم، قابل دفاع و کارآمد برای انتساب مسئولیت به به‌کارگیرندگان هوش مصنوعی فراهم می‌آورد، و این چارچوب می‌تواند به‌مثابه مبنایی عملی برای قانون‌گذاری هوش مصنوعی در نظام حقوقی ایران و البته دیگر نظام‌های حقوقی که از این «خلاء مسئولیت» رنج می‌برند مورد توجه قرار گیرد؛ بر همین بنیاد، نگارندگان پیشنهادهایی نیز در خاتمه مقاله ارائه کرده‌اند.

واژگان کلیدی: هوش مصنوعی، مسئولیت نیابتی، معیار ارتباط نزدیک، درونی‌سازی هزینه‌ها.

۱– مقدمه

هوش مصنوعی، افق‌هایی را در عرصه‌های گوناگون حیات بشری گشوده و نویدبخش تحول در صنعت، بهداشت، حمل‌ونقل و فراتر از آن است. این تحولات سریع فناوری و تطورات الگوهای کسب و کار نوین با ویژگی‌هایی چون توسعه فناوری‌های تخریب‌گر، غیرملموس بودن، پویایی و تغییر دائمی فضا، تحقق حاکمیت قانون و اصول و قواعد حکمرانی به شکل مرسوم و سنتی آن را ناممکن کرده است. در چنین فضایی، نمی‌توان به قوانین سنتی اتکای چندانی کرد. در این میان، نظام‌های حکمرانی با استفاده از الگوها و سازوکارهای جدید تنظیم‌گری، در تلاش برای حداکثرسازی توسعه این فناوری‌ها و و در عین حال، حفظ و تضمین حقوق شهروندان هستند. (غمامی و هاشمی، ۱۴۰۴، ۱) با این وجود این پیشرفت‌های فناورانه، مجموعه‌ای از مسائل حقوقی بی‌سابقه و بی‌پاسخ را نیز به همراه آورده‌اند[۱] که در کانون آن‌ها، مسئله «خلاء مسئولیت» در قبال آسیب‌های ناشی از عملکرد این سامانه‌ها قرار دارد.[۲] ویژگی‌هایی چون استقلال عملیاتی، ماهیت غیرقابل پیش‌بینی و غیرقابل توضیح، فرآیند سنتی انتساب مسئولیت را با دشواری مواجه ساخته است[۳] و منجر به نگرانی پیرامون تضمین حقوق بزه‌دیدگان و ایجاد بازدارندگی مؤثر شده است؛ چرا

---

[۱] به طوری که حوزه‌های حقوقی متعدد مشخصا در خصوص آن قانون گذاری کرده اند. از جمله می‌توان به قانون هوش مصنوعی اتحادیه اروپا (EU AI Act) (نهایی شده در ۲۰۲۴)، مقررات مربوط به دیپ‌فیک (Deep Synthesis Provisions) (لازم‌الاجرا از ۲۰۲۳) در چین، قانون هوش مصنوعی و داده (AIDA) (پیشنهاد شده در ۲۰۲۲) کانادا اشاره کرد. برای اطلاع جامع از قوانین کشورهای مختلف در خصوص هوش مصنوعی به اینجا رجوع شود:

https://www.aiprm.com/ai-laws-around-the-world/

[۲] اولین بار در سال ۲۰۰۴ بحث خلامسئولیت در یک مقاله مطرح شد و هم اکنون در سال ۲۰۲۵ این بحث آکادمیک ادامه دارد. آندرس ماتیاس –نویسنده– در آن مقاله (که بسیار پیشگامانه و جلوتر از زمانه خود بود) به الگوریتم‌هایی که دارای قابلیت یادگیری و خودمختاری هستند، مانند شبکه‌های عصبی، الگوریتم‌های ژنتیک و معماری‌های عامل، اشاره و استدلال کرد که رفتار این سیستم‌ها می‌تواند برای سازندگان یا اپراتورها غیرقابل پیش‌بینی شود و در نتیجه، تعیین مسئولیت برای اقدامات آن‌ها دشوار می‌گردد. سه سال بعد مقاله دیگری در این خصوص منتشر شد که استدلالات مقاله قبلی را تقویت و تکمیل کرد. به این دو مقاله تاکنون بیش از هزار مرتبه استناد شده است. نویسنده این سطور درک خود از حقوق هوش مصنوعی را مدیون مقاله ماتیاس می‌داند.

-Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and information technology, 6(3), p. 175-183.

-Sparrow, R. (2007). Killer robots. Journal of applied philosophy, 24(1), 62-77.

[۳] برای مطالعه در خصوص چیستی خلا مسئولیت با توجه به پیشرفت های هوش مصنوعی به این مقاله رجوع شود:

که تعیین مسئولیت برای درک علل وقایع نامطلوب، پیشگیری از تکرار آن‌ها و تنظیم انتظارات ما از فناوری ضروری است (Johnson, 2015, p. 713).

در چنین بستری، با توجه به آنکه سیاست‌های مرتبط با مسئولیت دارای تأثیرات اقتصادی و اجتماعی قابل توجهی هستند، انتخاب یک سازوکار (رژیم) مسئولیت مشخص و منسجم برای هوش مصنوعی یک اولویت ضروری خواهد بود. (ذاکری نیا، ۱۴۰۲، ۱۳۶) فقدان چنین سازوکاری می‌تواند منجر به افزایش ریسک‌های حقوقی و اقتصادی برای توسعه‌دهندگان و به‌کارگیرندگان و همچنین ناکارآمدی معیارهای سنتی اعمال مسئولیت (همچون قابلیت پیش بینی و قصد مجرمانه) شوند.[۴] در اینصورت منتفعان هوش مصنوعی، فاقد انگیزه‌های لازم برای به حداقل رساندن آسیب و خسارت، در اعمال دقت لازم قصور خواهند کرد یا حتی ممکن است عمداً باعث آسیب شوند، زیرا می‌توانند این کار را مبتنی بر خلاءهای مسئولیت و بدون مجازات انجام دهند. بنابراین، نگرانی در مورد فقدانِ بازدارنده‌های[۵] رفتار مجرمانه است که یک ساختار انگیزشیِ مفید را از بین می‌برند. (Königs, 2022, p.5) در واکنش به این نیاز مبرم، رویکردهای مختلفی از پذیرش نوعی مسئولیت‌پذیری برای برخی سیستم‌های هوش مصنوعی،[۶] تا تلاش برای یافتن راهکارهای عملی جهت پر کردن خلاء موجود مطرح شده است؛ در میان این راهکارها، رجوع به دکترین انعطاف‌پذیر

Königs, P. (2022). Artificial intelligence and responsibility gaps: What is the problem?. Ethics and Information Technology, 24(3), 36.

۴ . برای مطالعه مبسوط در خصوص اینکه چگونه پدیده هوش مصنوعی منجر به ناکارآمدی معیارهای سنتی مسئولیت(قصد–علیت) می‌شود به اثر ارزشمند و تاثیر گذار زیر رجوع شود:

Bathaee, Y. (2017). The artificial intelligence black box and the failure of intent and causation. Harv. JL & Tech., 31.

5 disincentives

۶ اولین بار بین سال‌های ۲۰۱۰ تا ۲۰۱۵، پروفسور گابریل هالوی نظریه مسئولیت کیفری مستقیم هوش مصنوعی را از طریق یک مقاله علمی و متعاقبا دو کتاب دیگر مطرح کرد. علی‌رغم جذابیت ظاهری و خلاقانه بودن، این نظریه به دلیل فقدان رویکردهای عملیاتی و کاربردی برای حل مسائل موجود مورد نقد قرار گرفت و با وجود بررسی‌های جامع توسط نهادهای قانون‌گذاری، از جمله اتحادیه اروپا در روند تدوین مقررات هوش مصنوعی و سایر حوزه‌های قضایی، دیدگاه هالوی مبنای هیچ قانونگذاری واقع نشد چرا که در حال حاضر امکان پیاده سازی آن وجود ندارد. برای مطالعه این نظر به منابع ذیل مراجعه شود:

- Hallevy, G. (2010). The criminal liability of artificial intelligence entities-from science fiction to legal social control. Akron Intell. Prop. J., 4, 171.
- Hallevy, G. (2013). When robots kill: Artificial intelligence under criminal law. UPNE.
- Hallevy, G. (2015). Liability for crimes involving artificial intelligence systems (Vol. 257). New York, NY, USA: Springer International Publishing.

مسئولیت نیابتی[۷]، که سابقه‌ای طولانی در نظام‌های حقوقی کامن‌لا و رومی–ژرمنی دارد، گزینه‌ای قابل تأمل به نظر می‌رسد. مسئولیت نیابتی، به عنوان سازوکاری که تحت شرایط خاص، بار مسئولیت ناشی از فعل مجرمانه یک شخص (عامل) را بر شخص دیگری (اصیل یا مسئول نیابتی) که با او رابطه ویژه‌ای دارد (مانند کارفرما–کارمند والد–فرزند)، تحمیل می‌کند – حتی بدون اثبات تقصیر یا عمد مستقیم اصیل – ظرفیت بالقوه‌ای برای فائق آمدن بر برخی از موانع مسئولیت در عصر هوش‌مصنوعی دارد (Diamantis, 2023, p.319).

البته، اعمال مفاهیم مسئولیت بر هوش‌مصنوعی خالی از پیچیدگی نیست و مواردی همچون دشواری در احراز عنصر روانی (اعم از قصد یا تقصیر) و معضل توزیع مسئولیت در ساختارهای سازمانی پیچیده یا زنجیره‌های تأمین فناوری که از آن با عنوان «مشکل دست‌های زیاد»[۸] یاد می‌شود، در خصوص هوش‌مصنوعی تشدید می‌گردند (Kerrigan, 2022, p. 149). اما این پیچیدگی‌ها، به جای آنکه مبنایی برای رد کامل امکان اعمال مسئولیت نیابتی باشند، ضرورت یک تحلیل دقیق و تطبیقی برای بررسی چگونگی انطباق این دکترین با ویژگی‌های منحصر به فرد هوش‌مصنوعی را آشکار می‌سازند؛ رویکردی که برخی اندیشمندان نیز با تأکید بر لزوم بسط و به‌روزرسانی چارچوب‌های حقوقی موجود برای مواجهه با فناوری‌های نوین، بر آن صحه می‌گذارد (Johnson, 2015, p. 709).

از طرف دیگر، در نظام حقوقی ایران، چارچوب کلی مسئولیت نیابتی در ماده ۱۴۲ قانون مجازات اسلامی تبیین شده است که شرایط آن را منوط به وجود مسئولیت قانونی یا ارتکاب تقصیر در قبال نتیجه می‌داند. در عین حال صرف تلاش برای انطباق پدیده هوش مصنوعی با مقررات موجود که برای وضعیت‌های حقوقی متفاوتی تدوین شده‌اند، اساساً قادر به ارائه راهکاری جامع و پایدار برای خلاء مسئولیت ناشی از هوش مصنوعی نبوده و حتی می‌تواند غیرسازنده[۹]باشد و منابع پژوهشی را از وظیفه اولی، یعنی بحث و تحلیل در خصوص ایجاد قواعد و مقررات اختصاصی برای هوش مصنوعی–همچون کشورهای پیشرو اروپایی– منحرف سازد. بدین جهت، پژوهش حاضر درصدد تحلیل انطباقی هوش مصنوعی با ماده ۱۴۲ نیست.[۱۰]هدف اصلی این مقاله، ارائه و ارزیابی

---

[7] Vicarious Liability
[8] Many hands problem
[9] counterproductive
[۱۰] علاوه بر این باید توجه داشت که مسئولیت نیابتی در حقوق کیفری ایران، در حد تصویب ماده ۱۴۲ باقی مانده است و هیچ گونه پشتوانه‌ی اجرایی در رویه قضایی ندارد؛ بنابراین زمانی که اعمال آن بر جرایم و مجرم انسانی با ابهامات فراوانی مواجه است

یک چارچوب مفهومی و تحلیلی با الهام از دکترین مسئولیت نیابتی (با تمرکز بر نظام کامن‌لا) همچنین، تمهید مبانی نظری و ارائه ایده‌های عملی برای تدوین مقررات اختصاصی و کارآمد برای هوش مصنوعی در حقوق ایران است. این امر از آن جهت ضروری و کمک‌کننده می‌نماید که نظام تنظیم‌گری فناوری‌های جدید در ایران، ضعیف، غیرساختارمند، جزیره‌ای، نامنسجم و ناهماهنگ توصیف شده است و نظام حقوقی ایران در حوزه هوش مصنوعی تاکنون رویکرد فعال و تنظیمی روشنی در پیش نگرفته و سیاست‌گذاری و حکمرانی در این حوزه معطل مانده است (غمامی و هاشمی، ۱۴۰۴، ۱).

با توجه به مراتب مذکور، مقاله حاضر، به تحلیل مبانی موجِّهه اعمال مسئولیت نیابتی بر اصیل در قبال عملکرد هوش مصنوعی می‌پردازد و چارچوب مفهومی و تحلیلیِ عناصر دوگانه این مسئولیت – یعنی احراز رابطه حقوقی خاص (با تمرکز بر معیار «شبیه به استخدام») و احراز ارتباط فعل مجرمانه با آن رابطه (با تمرکز بر معیار «ارتباط نزدیک») – را صورتبندی و تشریح می‌نماید. بدین منظور، ابتدا مبانی اقتصادی و مبتنی بر ریسک (۲) و سپس مبانی عملیاتی و کنترلی توجیه‌گر (۳) این مسئولیت که در رویه قضایی کامن‌لا مطرح شده‌اند، تحلیل می‌شوند. در ادامه، عناصر مسئولیت نیابتی در پرتو هوش مصنوعی (۴) مورد بررسی قرار می‌گیرد.

۲– مبانی اقتصادی و مبتنی بر ریسک در توجیه مسئولیت نیابتی هوش‌مصنوعی

در نظام‌های حقوقی مختلف، رویکردهای تنظیم‌گری فناوری‌های نوپدید به سمت ارزیابی و مدیریت ریسک حرکت می‌کنند؛ این رویکردها که گاه تحت عنوان «نظریه تنظیمی مبتنی بر ریسک» شناخته می‌شوند، به جای محدود کردن فناوری‌ها، بر شناسایی، ارزیابی و مدیریت ریسک‌های مرتبط با آن‌ها از طریق تدوین سیاست‌ها و مقررات مناسب تمرکز دارند. (غمامی و هاشمی، ۱۴۰۴، ۱۰). در همین راستا، تبیین قابلیت اعمال مسئولیت نیابتی بر جرائم ناشی از هوش مصنوعی نیز نیازمند تحلیل دقیق مبانی موجِّهه آن است. بخش نخست این تحلیل، بر ابعاد اقتصادی و ملاحظات ناظر بر تخصیص بهینه ریسک متمرکز است؛ ملاحظاتی که در رویه قضایی کامن‌لا نقش اساسی در توجیه این نوع مسئولیت ایفا می‌کنند و در این چارچوب، دو محور اساسی مورد مداقه قرار می‌گیرد. نخست، ظرفیت برتر اصیل در قیاس با عامل هوشمند جهت جبران خسارت و مدیریت تبعات مالی ریسک (۲–۱) و دوم، مبنای حقوقی–اقتصادیِ مسئولیت ناشی از ایجاد ریسک توسط شرکت و اصل درونی‌سازی

و این ابهامات منجر به بی استفاده ماندن ماده ۱۴۲ در نظام حقوقی ایران شده است هر تلاشی برای تطبیق جرایم به کلی متفاوت هوش مصنوعی با آن محکوم به بی‌نتیجگی کارکردی است.

هزینه‌های جرم (۲-۲). بررسی این دو محور، مبنایی برای درک چراییِ انتساب مسئولیت به اصیل فراهم می‌آورد.

۱-۲- ظرفیت برتر اصیل در جبران خسارت و مدیریت ریسک‌های نوظهور

ملاحظات عمل‌گرایانه و اقتصادی به مثابه مبنایی برای توجیه مسئولیت نیابتی در رویه قضایی کامن‌لا به‌طور مکرر مورد تأکید قرار گرفته است. این رویکرد، مسئولیت را به نقطه‌ای در زنجیره فعالیت اقتصادی تخصیص می‌دهد که بیشترین ظرفیت را برای مدیریت و جبران خسارت‌های احتمالی دارا باشد. بر این اساس، دو استدلال مطرح می‌شود؛ نخست، توانایی مالی و ساختاری برتر «اصیل» (به‌کارگیرنده) (۲-۱-۱) و دوم، چالش‌های ذاتی و موانع حقوقی در تخصیص شخصیت حقوقی به سامانه هوش مصنوعی که ضرورت تمرکز بر مسئولیت اصیل را تقویت می‌کند (۲-۱-۲).

۱-۱-۲- توانایی اصیل به مثابه بهترین اجتناب کننده هزینه

این واقعیت انکارناپذیر که اصیل[11]، که غالباً یک کارفرما، شرکت تجاری یا سازمان است، معمولاً در موقعیت به مراتب بهتری برای جبران آسیب وارده به اشخاص ثالث بزه‌دیده قرار دارد (ژوردن، ۱۳۹۴، ۱۴۶) یکی از عمل‌گرایانه‌ترین مبانی سیاستی توجیه‌گر دکترین مسئولیت نیابتی، که در رویه قضایی کامن‌لا مورد شناسایی و تأکید قرار گرفته است. دادگاه‌های کامن‌لا به مراتب تصریح نموده‌اند که هدف سیاستی زیربنایی مسئولیت نیابتی، حصول اطمینان از این موضوع است که مسئولیت نهایی تا حد امکان «منصفانه، عادلانه و معقول»، بر عهده شخصی قرار گیرد که واجد وسیله‌ی جبران خسارت قربانی باشد. (Various Claimants v. Catholic Child Welfare Society, 2012, para 34)[12] این توانمندی برتر اصیل، ریشه در عوامل متعددی دارد؛ سازمان‌ها و شرکت‌های تجاری دارایی‌های ثابت، سرمایه در گردش، جریان‌های درآمدی مستمر و دسترسی به بازارهای سرمایه برخوردارند و این بنیه مالی که گاه در ادبیات حقوقی از آن با عنوان «جیب عمیق»[13] کارفرما یاد می‌شود (Petrin & Witting, 2023, p. 291) ، به آن‌ها امکان می‌دهد تا آسیب‌های مالی ناشی از فعالیت‌های خود یا کارکنانشان را، حتی در مقیاس وسیع، جذب و جبران نمایند.

نهادها و شرکت‌هایی که در خط مقدم تجاری‌سازی و بهره‌برداری از سامانه‌های هوشمند قرار دارند، اعم از غول‌های فناوری یا کسب‌وکارهای بهره‌بردار، همان بازیگرانی هستند که ضمن انتفاع از مزایای اقتصادی هوش مصنوعی، ریسک‌های مرتبط با آن را نیز به جامعه می‌آورند

---

11 Principal

۱۲ از این پس: CCWS

13 Deep pocket

(Taplin, 2022, p. 6). این نهادها، به واسطه ساختار سازمانی و منابع مالی خود، در بهترین موقعیت برای پیش‌بینی، مدیریت و نهایتاً پوشش مالی این ریسک‌ها قرار دارند. تحمیل مسئولیت نهایی بر دوش آن‌ها تضمین می‌نماید که بزه‌دیدگان ناشی از هوش مصنوعی، از یک مسیر عملی و مؤثر برای احقاق حق برخوردار خواهند بود (Dremliuga et al., 2019, p. 111) و با یک عامل فاقد توانایی مالی یا زنجیره پیچیده علیت مواجه نخواهند شد. به بیان دیگر، این رویکرد موجب درونی‌سازی[۱۴] هزینه‌های ناشی از حوادث و خطاهای عملیاتی به عنوان جزئی از هزینه‌های متعارف آن فعالیت خاص در محاسبات اقتصادی شرکت می‌گردد.

علاوه بر درونی‌سازی هزینه، رویکرد تحمیل مسئولیت بر اصیل، با دیگر معیارهای اقتصادی شناخت مسئولیت نیز هم‌خوانی دارد. به عنوان مثال، مبتنی بر «اصل ارزان‌ترین اجتناب‌کننده هزینه»[۱۵] شخصی باید مسئول شناخته شود که می‌تواند از خسارت با کمترین هزینه جلوگیری کند یا آن را کاهش دهد، به شرطی که هزینه اجتناب یا کاهش خسارت کمتر از سود باشد. در مورد سیستم‌های هوش مصنوعی، غالباً این تولیدکننده یا به‌کارگیرنده اصلی است که در موقعیتی مناسب‌تر برای تعیین و بهبود ویژگی‌های ایمنی سیستم قرار دارد و به لحاظ اقتصادی می‌تواند ارزان‌ترین اجتناب‌کننده باشد. (حکمت نیا و همکاران، ۱۳۹۸، ۲۵۰) از منظر مبانی حقوق ایران نیز می‌توان مسئولیت شرکت در قبال زیان‌های ناشی از به‌کارگیری هوش مصنوعی را در راستای «اصل احترام» ارزیابی نمود. این اصل، به عنوان یک مبنای فقهی با پشتوانه استدلالی قوی اقتضا دارد که حقوق اشخاص و جامعه محترم شمرده شود و هرگونه ضرر و زیان وارده، جبران گردد. (حسینی و همکاران، ۱۴۰۳، ۵۰) بر این اساس، تحمیل مسئولیت بر «اصیل» (به‌کارگیرنده هوش مصنوعی) که از فناوری منتفع شده، ریسک‌های آن را ایجاد کرده و توانایی مدیریت و پوشش این ریسک‌ها را دارد، علاوه بر عدالت و انصاف با جوهره «اصل احترام» در فقه نیز هم‌سویی دارد و در راستای محترم شمردن حقوق بزه‌دیدگان است.

مآلاً، قرار دادن مسئولیت بر عهده نهادِ به‌کارگیرنده یا توسعه‌دهنده هوش مصنوعی، به دلیل توانایی برتر آن در جبران خسارت و مدیریت ریسک (دیلمی و دیگران، ۱۳۹۹، ۷۱)، در انطباق کامل با معیار «منصفانه، عادلانه و معقول» بودن، که توسط دیوان عالی در پرونده CCWS تبیین شد، قرار می‌گیرد. این رهیافت، منصفانه است زیرا مسئولیت را بر دوش نهادی می‌نهد که از فناوری

[14] Internalization
[15] Best Cost Avoider

سود می‌برد و توانایی مدیریت ریسک آن را داراست، ضمن آنکه امکان جبران خسارت را برای بزه‌دیده فراهم می‌کند؛ عادلانه است زیرا به تحقق عدالت جبرانی[۱۶] و عدالت توزیعی[۱۷و۱۸] یاری رسانده و معقول و کارآمد است زیرا راه حلی عملی برای مسئله خلاء مسئولیت ارائه داده و همزمان، انگیزه‌های اقتصادی لازم برای سرمایه‌گذاری در ایمنی و کاهش ریسک را ایجاد می‌کند (Diamantis, 2023, p. 320) و بدین ترتیب به هدف مهم بازدارندگی نیز به طور غیرمستقیم کمک می‌نماید.

۲-۱-۲- شخصیت حقوقی هوش مصنوعی و ضرورت تمرکز بر مسئولیت اصیل

در حالی که تحلیل پیشین، ظرفیت برتر «اصیل» (به‌کارگیرنده) در جبران خسارت را از منظر اقتصادی مدلل ساخت، رویکرد علمی اقتضای این را دارد که وضعیت حقوقی بازیگر دیگر این معادله، یعنی سامانه هوشمند، نیز بررسی شود. این بررسی، به طرح این پرسش می‌انجامد که آیا اعطای شخصیت حقوقی به هوش مصنوعی می‌تواند راهکاری جایگزین و مطلوب برای تخصیص مسئولیت باشد؟ این ایده که در پی رخدادهایی نظیر اعطای شهروندی به ربات سوفیا و پیشنهادهای اولیه پارلمان اروپا برای ایجاد «شخصیت الکترونیک» مورد توجه قرار گرفت، (Lanni & Monterossi, 2017, p. 579; Szentgáli-Tóth, 2021, p. 102) علی‌رغم جذابیت اولیه، به دلیل ناکارآمدی عملی و موانع مفهومی، در نهایت مورد پذیرش نظام‌های حقوقی قرار نگرفت (Dremliuga et al., 2019, p. 108)، حتی اقداماتی مانند اعطای شهروندی به سوفیا، بیش از آنکه مبتنی بر یک تحلیل حقوقی عمیق باشد، برآمده از ملاحظات سیاسی و راهبردهای بازاریابی بود(Szentgáli-Tóth, 2021, p. 106).

مهم‌ترین مانع در این مسیر، بی‌اثر بودن تحمیل مسئولیت مستقیم بر یک عامل غیرانسانی است که فاقد دارایی و توانایی واقعی برای پاسخگویی است و در نقطه مقابل، «اصیل» که غالباً یک شرکت تجاری یا سازمان است، توانایی جبران خسارت بزه‌دیده را دارد؛ وضعیتی که حتی در مقایسه

---

16 Corrective Justice

17 Distributive Justice

۱۸ فعالیت‌های اقتصادی ذاتاً با ریسک همراه‌اند و عدالت توزیعی به این می‌پردازد که ریسک‌های ناشی از این فعالیت‌ها چگونه باید در جامعه توزیع شوند و از طرف دیگر، مسئولیت نیابتی، ریسک مالی ناشی از خطای عامل در حین کار را به اصیل منتقل می‌کند و توجیه این انتقال ریسک، بر پایه ملاحظات اصل منفعت (Benefit Principle) انصاف (Fairness) استوار است که در قلب عدالت توزیعی قرار دارد. برای مطالعه بیشتر به مهمترین اثر در این حوزه رجوع شود:

رالز، جان. (۱۴۰۳). نظریه‌ای در باب عدالت (مرتضی نوری، مترجم). نشر مرکز. (اثر اصلی در سال ۱۹۷۱ منتشر شده است).

با کارکنان انسانی نیز صادق است، زیرا آن‌ها نیز به ندرت از دارایی شخصی کافی برای جبران خسارات سنگین برخوردارند (Oliphant & Wagner, 2012, pp. 26-27) و این ناتوانی در مورد سامانه‌های هوش مصنوعی که ذاتاً فاقد هرگونه دارایی مستقل هستند بارزتر است. افزون بر مانع عملی، اعطای شخصیت حقوقی به هوش مصنوعی با این پرسش اساسی روبروست که چه منفعتی برای جامعه به همراه خواهد داشت؟ برای مثال، شناسایی شخصیت حقوقی برای شرکت‌ها، منافع اقتصادی واضحی مانند افزایش کارایی، شفافیت مالی و تسهیل پاسخگویی را به دنبال داشته است (Dremliuga et al., 2019, p. 111)؛ در حالی که چنین مزایای ملموسی برای اعطای شخصیت به هوش مصنوعی متصور نیست و شاید به همین دلیل هم، تلاش‌های بین‌المللی نیز در این زمینه با مقاومت جدی مواجه شده است، چنانچه هم اتحادیه اروپا و هم ایالات متحده، با این ایده مخالفتی صریح دارند (Dremliuga et al., 2019, p. 108) و نهادهای تخصصی مانند سازمان جهانی مالکیت فکری[۱۹] نیز اقداماتی را در جهت ممانعت از شناسایی حقوق مالکیت فکری برای هوش مصنوعی اتخاذ کرده‌اند[۲۰] (Chakrabarty & Baral, 2023, p. 3).

این مسئله در نظام حقوقی ایران و مبانی فقهی آن، به نحو بهتری درک می‌شود چرا که هوش مصنوعی نه در قالب «شخص حقیقی» می‌گنجد، زیرا فاقد عناصر اساسی انسان بودن نظیر حیات، اراده و اهلیت است (ولی پور و اسماعیلی، ۱۴۰۰، صص. ۴-۵) و نه می‌توان آن را «شخص حقوقی» به شمار آورد. اشخاص حقوقی، هویتی اعتباری هستند که به اراده قانون‌گذار و افراد انسانی ایجاد می‌شوند، در حالی که هوش مصنوعی فاقد چنین مبنای تأسیسی و ارادی است (همان منبع). در نتیجه، عدم امکان شناسایی هوش مصنوعی به عنوان یک شخص، هرگونه تلاش برای مسئول دانستن مستقیم آن را در حقوق ایران با مانع مواجه می‌سازد (حسینی و همکاران، ۱۴۰۳، ص. ۵۰).

---

[19] WIPO

[۲۰] در پرونده‌های قضایی بین‌المللی معروف به «دابوس» (DABUS)، تلاش شد تا یک سیستم هوش مصنوعی به عنوان مخترع دو اختراع ثبت شود ولی نظام‌های حقوقی از جمله در بریتانیا، ایالات متحده، و دفتر ثبت اختراع اروپا (EPO)، با این استدلال که یک ماشین فاقد شخصیت حقوقی لازم برای این عنوان است، این درخواست را رد کردند. برای مطالعه به دو لینک زیر مراجعه شود، لینک دوم مربوط به دادگاه بریتانیا است:

https://www.reuters.com/legal/litigation/us-appeals-court-says-artificial-intelligence-cant-be-patent-inventor-2022-08-05/
https://www.supremecourt.uk/cases/uksc-2021-0201

بنابراین می‌توان گفت ایده اعطای شخصیت حقوقی به هوش مصنوعی، راهکاری کارآمد و مطلوب برای رفع خلاء مسئولیت نیست و شاید تنها در آینده‌ای دور، پیشرفته‌ترین ربات‌های کاملاً خودمختار بتوانند کاندیدای کسب نوعی شخصیت حقوقی باشند، اما چنین چشم‌اندازی در حال حاضر موضوعیت ندارد. (Szentgáli-Tóth, 2021, p. 113) بر این اساس، منطقی‌ترین و عملی‌ترین راهکار، تمرکز بر مسئولیت‌پذیر ساختن اشخاص و نهادهای انسانی کنترل‌کننده این فناوری، یعنی کاربران، مالکان و تولیدکنندگان آن است (Chakrabarty & Baral, 2023, p. 3). بر این اساس، به جای خلق یک مفهوم حقوقی ساختگی و ناکارآمد، باید مسئولیت را به همان کانونی منتسب کرد که از این فناوری منتفع شده، ریسک‌های آن را به جامعه تحمیل کرده و نهایتاً کنترل راهبردی بر آن را در اختیار دارد.

۲-۲- فعالیت هوش مصنوعی به مثابه بخش جدایی‌ناپذیر از فعالیت اصیل و لزوم درونی‌سازی هزینه‌های جرم

یکی از مبانی دوگانه‌ای که توجیه نظریِ تحمیل مسئولیت نیابتی بر اصیل را فراهم می‌آورد، حول دو محور شکل گرفته است. نخست، ادغام ماهوی و کارکردی فعالیت عامل در ساختار فعالیت اقتصادی اصیل؛ و دوم، ایجاد یا تشدید ریسک توسط اصیل به واسطه همین فعالیت‌ها. بدین معنا که انتساب عمل زیانبار عامل به اصیل از منظر مسئولیت نیابتی، منوط به احراز این شرط است که عمل مذکور در جریان فعالیتی ارتکاب یافته باشد که بتوان آن را به عنوان جزئی لاینفک و ذاتی[۲۱] از مجموعه فعالیت‌های تجاری اصیل تلقی نمود. این اصل در چارچوب روابط استخدامی سنتی، مصداق بارزی می‌یابد؛ زیرا وظایف محوله به مستخدم (کارگر یا کارمند)، ماهیتاً در راستای تحقق اهداف کارفرما تعریف شده و عمل وی، از منظر کارکردی، امتداد و تجلی فعالیت کارفرما محسوب می‌گردد (دیلمی و همکاران، ۱۳۹۹، ص ۷۱؛ ژوردن، ۱۳۹۴، ص ۱۴۵).

استقرار و به‌کارگیری سامانه‌های هوش مصنوعی مصداق بارز منطق «ادغام در فعالیت تجاری»[۲۲] است. چرا که هوش مصنوعی در پارادایم کنونی، به یک مؤلفه محوری در هسته عملیات تجاری مدرن مبدل شده است. سازمان‌ها، سامانه‌های هوشمند را نه برای انجام امور حاشیه‌ای، بلکه دقیقاً به منظور اجرای وظایف مهم یا ارتقای کیفی فرآیندهای ارزش‌آفرین خود به خدمت

[21] integral
[22] Integration into Business Activity

می‌گمارند.[۲۳] تحلیل کلان‌داده‌ها جهت شخصی‌سازی پویا در استراتژی‌های بازاریابی شخصی سازی شده (Ebers & Gamito, 2021, p. 2)، اتوماسیون فرآیندهای تولید به منظور افزایش بهره‌وری، ارائه خدمات مالی خودکار[۲۴] و استفاده از هوش مصنوعی در فرآیندهای حساس مدیریتی و منابع انسانی، همگی مؤید این ادغام عمیق هستند. این سطح از یکپارچگی عملکردی، علی‌الخصوص با در نظر گرفتن ریسک‌های ذاتی مترتب بر پردازش کلان‌داده‌ها نظیر نقض حریم خصوصی یا بروز تبعیضات الگوریتمی (Ebers & Gamito, 2021, p. 3)، نشان می‌دهد که فعل مجرمانه سامانه هوشمند در اثنای اجرای این وظایف، امری منفک از فعالیت کلی اصیل نبوده، بلکه به مثابه امتداد فناورانه آن قابل ارزیابی است.

این پیوند ذاتی میان فعالیت هوشمند و عملیات تجاری، مستقیماً به سنگ بنای دوم مسئولیت یعنی نظریه «خلق ریسک»[۲۵] پیوند می‌خورد. منطق حقوقی حکم می‌کند شخصی که برای کسب سود، یک فعالیت اقتصادی پرخطر را وارد جامعه می‌کند، باید مسئولیت تحقق آن خطرات را نیز بر عهده بگیرد (Brodie, 2010, p. 61). رویه قضایی نیز بر این منطق صحه گذارده است؛ چنانکه در پرونده اسکاندیناویسکا[۲۶و27] دادگاه تصریح نمود که تحمیل مسئولیت بر شخصی که «یک شرکت اقتصادی پرخطر را وارد جامعه می‌کند» در زمان تجلی آن خطرات، امری عادلانه محسوب شده و علاوه بر تأمین هدف جبران خسارت مؤثر، کارکرد بازدارندگی و ترغیب به اتخاذ تدابیر پیشگیرانه را نیز ایفا می‌نماید (Giliker, 2022, p. 125). این استدلال به نحو انکارناپذیری بر به‌کارگیری سامانه‌های هوش مصنوعی قابل انطباق است. تصمیم یک بنگاه برای به‌کارگیری هوش مصنوعی به منزله ایجاد آگاهانه عدم قطعیت و ریسک است که از ویژگی‌های ماهوی این فناوری نشأت می‌گیرد (Nikolinakos, 2023, p. 699). پدیده جعبه سیاه به مثابه یکی از ویژگی‌های ذاتی هوش مصنوعی با ایجاد گسست در زنجیره علیتِ قابل پیش‌بینی، فهم چرایی یک تصمیم مخرب را حتی برای خالقان

---

[۲۳] به عنوان مثال می‌توان به استفاده پلیس از هوش مصنوعی یا استفاده دادگاه‌های آمریکایی از هوش مصنوعی اشاره کرد. برای مطالعه در خصوص چالش‌های استفاده پلیس از هوش مصنوعی به این منبع رجوع شود:

-Berk, R. A. (2021). Artificial intelligence, predictive policing, and risk assessment for law enforcement. Annual Review of Criminology, 4(1), 209-237.

[24] Robo-advisors
[25] Risk Creation
[26] Skandinaviska
[27] Skandinaviska Enskilda Banken AB v Asia Pacific Breweries (S) Pte Ltd, 2011

سیستم نیز ناممکن ساخته و مفهوم پیش‌بینی‌پذیری را به چالش می‌کشد (تخشید، ۱۴۰۰، ص ۲۳۸) و منجر به ریسک‌هایی می‌شود که اعم از رفتارهای غیرقابل پیش‌بینی در شرایط نامتعارف، سوگیری‌های نظام‌مند ناشی از داده‌ها در استخدام هوشمند بوسیله هوش مصنوعی (Szapiro & Kacprzyk, 2022, p. 160)، یا آسیب‌پذیری‌های امنیتی هستند. در نتیجه، اصیل با انتخاب استقرار این فناوری، این مجموعه از ریسک‌ها را آگاهانه به حوزه فعالیت خود و متعاقباً به جامعه تحمیل می‌نماید.

بر این اساس، تلاقی این دو مبنا یک استدلال حقوقی متقن را صورت‌بندی می‌کند و آن اینکه هنگامی که خسارتی از دل ریسک‌های ذاتی یک سیستم هوشمند سر بر می‌آورد، مسئولیت نیابتی اصیل، به دلیل ادغام ماهوی آن سیستم در فعالیت تجاری‌اش و هم به دلیل نقش او به عنوان خالق آن ریسک، امری موجه است و تضمین می‌کند که هزینه‌های اجتماعی ناشی از این فناوری، بر عهده همان بازیگرانی قرار گیرد که از مزایای آن منتفع شده و به صورت همزمان، توانایی و مسئولیت مدیریت ریسک‌های آن را به عنوان جزئی جدایی‌ناپذیر از فعالیت اقتصادی خود پذیرفته‌اند.

۳- مبانی عملیاتی و کنترلی در توجیه مسئولیت نیابتی هوش‌مصنوعی

فراتر از منطق اقتصادی تخصیص ریسک، معماری حقوقی مسئولیت نیابتی برای هوش مصنوعی بر دو بُعد مکمل و درهم‌تنیده تمرکز دارد که اقتدار حاکمیتی اصیل را آشکار می‌سازد. نخست، جایگاه کارکردی هوش مصنوعی به مثابه ابزاری که در خدمت اهداف و منافع غایی اصیل قرار دارد (۱-۳)؛ و دوم، اثبات کنترل مؤثر اصیل بر چارچوب کلان فعالیت سامانه، حتی در صورت وجود خودمختاری عملیاتی (۲-۳). تبیین این دو مؤلفه، پیوند ناگسستنی میان اراده راهبردی اصیل و عملکرد عامل هوشمند را آشکار ساخته و مبنای حقوقی انتساب مسئولیت را مستحکم می‌نماید.

۱-۳- تبعیت ابزاری هوش مصنوعی از اهداف اصیل

یکی از مبانی دکترین مسئولیت نیابتی، که در رویه قضایی نیز مورد توجه قرار گرفته،[۲۸] بر این گزاره استوار است که فعل زیانبار باید در جریان فعالیتی ارتکاب یابد که عامل «از طرف»[۲۹] و در جهت «منافع» اصیل[۳۰] انجام می‌دهد. این اصل، مسئولیت را به ریسک‌های ذاتی و ملازم یک فعالیت اقتصادی سازمان‌یافته پیوند می‌زند. بدین معنا که اصیل باید پاسخگوی آسیب‌هایی باشد که در مسیر

---

[۲۸] از جمله به‌طور ضمنی در تحلیل‌های پرونده CCWS

[29] On Behalf Of

[30] Principal

تحقق اهداف سازمانش حادث می‌شوند، همچنانکه این منطق به سهولت در روابط استخدامی سنتی قابل مشاهده است؛ فعالیت‌های کارمند، تجلی کارکردی بنگاه کارفرما در جهان خارج است و چنانچه در اثنای انجام وظایف خود مرتکب خطایی شود، تبعات آن منطقاً بر عهده همان نهادی قرار می‌گیرد که از آن فعالیت منتفع شده و ریسک‌های آن را ایجاد نموده است (دیلمی و دیگران، ۱۳۹۹، ۷۱).

این پیوند غایی میان فعل زیانبار و منافع اصیل، به سادگی در خصوص هوش مصنوعی تحقق می‌یابد، چرا که یک سامانه هوشمند، فارغ از درجه خودمختاری عملیاتی‌اش،[۳۱] ذاتاً ابزاری است که برای تحقق اهداف معین یک نهاد پشتیبان (اصیل) طراحی، آموزش و به کار گرفته می‌شود. بنابراین، کنش‌های آن، حتی پیچیده‌ترین و غیرشفاف‌ترین آن‌ها، همواره در چارچوب حاکمیتی و راهبردی همان اهداف قابل تفسیر است. هوش مصنوعی به نیابت از اصیل عمل می‌کند و خطاهای آن صرفاً یک «نقص فنی»[۳۲]یا رویداد پیش‌بینی‌ناپذیر نیست، بلکه اغلب، پیامد منطقی و در عین حال ناخواسته بهینه‌سازی برای همان اهدافی هست که برایش تعریف شده است. به عنوان مثال، یک الگوریتم توصیه‌گر که برای بهینه‌سازی تعامل کاربر و درآمدزایی طراحی شده، ممکن است با ترویج افراط‌گرایی یا توزیع اطلاعات نادرست، به این هدف دست یابد (Soni, Sharma, & Sinha, 2024, p. 99). یا یک خودروی خودران[۳۳] که هدف آن انتقال ایمن و کارآمد است (Krömker, 2022, p. 420)، ممکن است در یک سناریوی بغرنج، با اتخاذ تصمیمی که پارامترهای ایمنی و کارایی را به شکلی نامطلوب موازنه می‌کند، موجب بروز صدمه جانی گردد (برزگر و الهام، ۱۳۹۹، ۲۰۸). در هر دو مورد، فعل زیانبار، فعلیت‌یافتن یکی از ریسک‌های ذاتی است که اصیل با تصمیم به استفاده از این ابزار فناورانه برای آن هدف خاص، آگاهانه آن را پذیرفته و در چرخه فعالیت خود ادغام کرده است.

در این تحلیل، خودمختاری عملیاتی هوش مصنوعی، نافی این پیوند بنیادین نیست. تصمیمات یک سیستم هوشمند، گرچه بدون مداخله آنی انسان اتخاذ می‌شوند، اما همچنان محصول مستقیم محدودیت‌ها، داده‌های آموزشی و اهداف کلانی هستند که توسط اصیل و یا به سفارش اصیل، تعریف و مهندسی شده‌اند (Szentgáli-Tóth, 2021, p. 104). در واقع، خودمختاری صرفاً به

---

[31] Operational Autonomy

[32] Technical Glitch

[۳۳] بخش هشتم قانون ایالت نوادا خودروی خودران را چنین تعریف میکند: «وسیلهٔ نقلیّهٔ موتوری‌ای است که از هوش مصنوعی، سنسورها و سیستم موقعیّت یاب جهانی برای راندن خودش بدون مداخلهٔ یک اپراتور انسانی استفاده می‌کند» (Bill AB511 Nevada Legislature, section 8)

«چگونگی» اجرای وظیفه مربوط است، حال آنکه «چرایی» و هدف نهایی آن همچنان در ید قدرت اصیل باقی می‌ماند. از این رو، با عنایت به اینکه فعل مجرمانه هوش مصنوعی در مسیر تحقق منافع اصیل و به نیابت از او رخ می‌دهد، اعمال مسئولیت نیابتی بر وی، انطباق کامل با جوهره این دکترین حقوقی دارد.

۲-۳- کنترل راهبردی و عملیاتی اصیل بر فعالیت سامانه هوشمند

معیار سنتی توجیه‌کننده مسئولیت نیابتی، مفهوم «کنترل»[۳۴] اصیل بر عامل در زمان ارتکاب فعل مجرمانه است. این معیار در مواجهه با سیستم‌های هوشمند با ویژگی‌های خاص آن، از جمله پیچیدگی‌های فنی، دینامیسم تکاملی و سرعت تحولات، (غمامی و هاشمی، ۱۴۰۴، ۵-۶) با یک چالش روبرو می‌شود. اما این چالش با ارجاع به تحولات خود دکترین حقوقی، پاسخی درخور می‌یابد. درک معیار کنترل در حقوق مدرن، با توجه به تحولات بازار کار و پیچیدگی وظایف، دستخوش تحول شده است و رأی دیوان عالی در پرونده CCWS به مثابه یک نقطه عطف نظری، این گذار را صورت‌بندی کرده است (CCWS, 2012, para 36). در این پارادایم نوین، «کنترل» دیگر به معنای توانایی دیکته کردن جزئیات اجرایی وظایف، آن‌گونه که در دوران «ارباب و خدمتکار» متصور بود، نیست. بلکه، هسته کنترل در قدرت و اختیار راهبردی اصیل برای تعیین ماهیت و غایت فعالیت عامل، یعنی تعریف وظایف، اهداف و قلمرو کلی عملکرد در راستای منافع بنگاه، نهفته است و این‌گونه شرط مسئولیت از «مدیریت خُرد روش» جدا شده و به «حاکمیت کلان بر فعالیت» پیوند می‌خورد.

این بازتعریف مفهومی، دارای یک منطق سیاستی قابل توجه نیز هست و آن اینکه هنگامی که اصیل، عاملی را در ساختار سازمانی خود ادغام نموده و فعالیت او را در جهت اهداف خود هدایت می‌کند، آن فعالیت را به بخشی از قلمرو مسئولیت خویش بدل می‌سازد. در این شرایط، حتی اگر کنترل کامل بر کنش‌های عامل در عمل ناممکن باشد، از منظر سیاست‌گذاری تحمیل مسئولیت نیابتی تضمین می‌کند که بنگاه‌ها با هزینه‌های کامل مورد انتظار ناشی از صدمات مواجه شده و در نتیجه، از پذیرش ریسک‌های بیش از حد بازداشته شوند (Petrin & Witting, 2023, p. 291).

حال پرسش اساسی این است که آیا مفهوم مدرن «کنترل بر فعالیت»، آن‌گونه که در رویه قضایی مسئولیت نیابتی و مشخصاً رأی CCWS تبیین شد، می‌تواند بر رابطه میان اصیل (به‌کارگیرنده یا توسعه‌دهنده) و سامانه‌های هوشمند نیز اعمال گردد؟ پاسخ، با در نظر گرفتن ابعاد چندگانه این

[34] Control

رابطه، مثبت به نظر می‌رسد. این کنترل، اگرچه ممکن است متفاوت از نظارت مستقیم بر یک کارمند انسانی باشد، اما در لایه‌های مختلفی قابل شناسایی و از منظر حقوقی معنادار است. نخستین و شاید مهم‌ترین جنبه، کنترل بر تعریف هدف و وظیفه سامانه هوشمند است و در این راستا، این اصیل است که تصمیم می‌گیرد هوش مصنوعی برای چه مقصودی به کار گرفته شود، چه مشکلی را حل کند و در چه چارچوبی فعالیت نماید؛ این تعیین استراتژیک کارکرد و هدف، هسته اصلی کنترل اصیل بر فعالیت هوش مصنوعی را تشکیل می‌دهد و مسیر کلی عملکرد آن را مشخص می‌سازد. در کنار این، کنترل از طریق طراحی و پیکربندی اولیه نیز وجود دارد؛ اصیل، چه مستقیماً و چه با تعیین الزامات برای تیم توسعه، بر انتخاب الگوریتم‌ها، تعریف پارامترهای عملیاتی، تعیین محدودیت‌ها و قواعد رفتاری و پیاده‌سازی سازوکارهای ایمنی، نظارت و تأثیرگذاری قابل توجهی دارد و تصمیمات مراحل اولیه طراحی، شالوده و چارچوب رفتار آتی سامانه را بنیان می‌نهند (Gero & Sudweeks, 2012, p. 136). سرانجام و شاید در مؤثرترین شکل کنترل، این سیطره به حوزه کنترل معرفت‌شناختی بر داده‌های آموزشی[۳۵] گسترش می‌یابد که در آن اصیل با انتخاب و مهندسی داده‌هایی که سیستم با آن‌ها جهان را می‌آموزد، به طور مستقیم جهان‌بینی و الگوهای تصمیم‌گیری آن را شکل می‌دهد (Azad & Nazari-Heris, 2024, p. 45). این لایه‌های کنترل، شبکه‌ای درهم‌تنیده از اقتدار را می‌سازند که از منظر حقوقی، بسیار معنادارتر از نظارت آنی و مستقیم است و مبنایی مستحکم برای انتساب مسئولیت نیابتی به اصیل فراهم می‌آورد.

این حاکمیت راهبردی اصیل، از چارچوب ایستا و اولیه طراحی فراتر رفته و در تمام چرخه حیات عملیاتی سامانه هوشمند در قالب نظارت مستمر، قابلیت به‌روزرسانی و بازآموزی[۳۶] سیستم بر اساس بازخوردهای دریافتی به شکلی پویا استمرار می‌یابد (Charles Sr., 2025, pp. 214-215). مهم‌تر اینکه، این حاکمیت با توجه به اختیار نهایی اصیل برای مداخله، تحدید یا توقف کامل سامانه به اوج خود می‌رسد. این قدرت برای قطع کامل جریان و از کار انداختن سیستم (Cristiano et al., 2023, p. 173)، نافی هرگونه توهم استقلال مطلق و نبود کنترل است.[۳۷]

---

[35] Control via Data

[36] Control via Monitoring, Updating & Intervention

[۳۷] تحلیلی که در اینجا ارائه شده است، دارای درون مایه های عمیق منطقی است و دقیقا منطبق بر رویکردی عملگرایانه است که با قانونگذاری کشورهای مختلف هم‌راستا است. به عنوان مثال ماده ۴ قانون هوش مصنوعی چین مصوب سال ۲۰۲۳ (Interim Measures for the Management of Generative AI Services) به صراحت ارائه‌دهنده سرویس را به عنوان نهاد مسئول محتوای تولید شده توسط سیستم هوش مصنوعی‌اش معرفی می‌کند، همانطور که یک ناشر یا وب‌سایت مسئول محتوایی

علاوه بر این، وابستگی وجودی سامانه به منابع محاسباتی، انرژی و مالی که تماماً توسط اصیل تأمین و مدیریت می‌شود، لایه دیگری از این کنترل ساختاری و انکارناپذیر را به نمایش می‌گذارد.

ممکن است در برابر این تحلیل، به خودمختاری عملیاتی[38و39] برخی سیستم‌های هوشمند استناد شود که به آن‌ها اجازه تصمیم‌گیری بدون دخالت آنی انسان را می‌دهد (Gatti, 2016, p. 142). این استدلال اما خودمختاری عملیاتی در «روش» را با استقلال غایی در «هدف» یکسان می‌پندارد و از این منظر مغالطه می‌کند. همان‌گونه که در تفسیر مدرن مسئولیت نیابتی و رأی CCWS تبیین شد، کنترل حقوقی معنادار، ناظر بر هدایت کلان فعالیت است و نه لزوماً مدیریت خُرد اجرای آن. بنابراین، خودمختاری هوش مصنوعی در نحوه رسیدن به هدف، ناقض کنترل اصیل بر تعیین خودِ هدف، چارچوب عملکرد، محدودیت‌های عملیاتی و داده‌های ورودی نیست. این خودمختاری، یک خودمختاری محصور و هدایت‌شده است که در چارچوب حاکمیتی عمل می‌کند که ابعاد آن از تعریف غایت تا مدیریت داده و نظارت پویا، تماماً توسط اصیل مهندسی و مدیریت می‌شود.

این چارچوب تحلیلی، در انطباق کامل با «معیار قدرت و کنترل» که در ادبیات حقوقی مطرح است نیز قرار دارد، چرا که مطابق این معیار، مسئولیت باید بر عهده شخصی قرار گیرد که بیشترین قدرت و کنترل را بر سیستم اعمال می‌کند (حکمت نیا و همکاران، ۱۳۹۸، صص ۲۵۰-۲۵۱) و از آنجا که اصیل، همان‌طور که تشریح شد، از طریق مجموعه‌ای از کنترل‌های غایی، معماری‌شناختی، معرفت‌شناختی و نظارتی، حاکمیت مؤثری بر سامانه هوشمند دارد، انتساب مسئولیت به وی بر مبنای این معیار نیز کاملاً موجه است و در این راستا می‌توان اذعان داشت که قدرت تصمیم‌گیری در خصوص چرایی و چگونگی به‌کارگیری هوش مصنوعی و توانایی شکل‌دهی به چارچوب عملکرد آن، مسئولیت متناظر را نیز به همراه دارد.

۴- تحلیل عناصر مسئولیت نیابتی در پرتو هوش‌مصنوعی

بررسی‌های تطبیقی در نظام‌های حقوقی مختلف، از جمله کشورهای عضو اتحادیه اروپا، نشان می‌دهد که در جستجوی مبنایی برای مسئولیت ناشی از هوش مصنوعی، اغلب چهارچوب‌های مسئولیت موجود، از جمله مسئولیت ناشی از اشیاء، فعالیت‌های خطرناک، حیوانات و همچنین مسئولیت نیابتی، مورد توجه و تحلیل قرار می‌گیرند. (ذاکری نیا، ۱۴۰۲، ۱۴۳) نکته مهم آن که فراتر

است که منتشر می‌کند. سایر مواد مانند ماده ۱۴ و ۱۵ نیز به اقدامات لازم در صورت کشف محتوای غیرقانونی و رسیدگی به شکایات کاربران می‌پردازند که مکمل این مسئولیت‌پذیری کلی است.

[38] Functional Autonomy
[39] Operational Autonomy

از مبانی توجیهی، استقرار مسئولیت نیابتی منوط به احراز عناصر ماهوی دوگانه آن است؛ نخست وجود رابطه حقوقی خاص میان اصیل و عامل مستقیم فعل مجرمانه؛ و دوم، ارتکاب آن فعل در جریان یا به تعبیر فنی‌تر، در محدوده همان رابطه. در این قسمت برای بررسی این موضوع، ابتدا ظرفیت معیار انعطاف‌پذیر «شبیه به استخدام»[40] برای تبیین رابطه میان اصیل و عامل هوشمند ارزیابی شده و سپس، معیار «ارتباط نزدیک»[41] به عنوان ضابطه سنجش پیوند میان فعل مجرمانه و آن رابطه، معرفی و تشریح می‌گردد.

۱-۴-عنصر رابطه؛ از استخدام سنتی تا روابط «شبیه به استخدام» در عصر هوش‌مصنوعی

در طول عمر مسئولیت نیابتی در نظام کامن‌لا، پارادایم مسلط و تاریخی برای استقرار این نوع مسئولیت، همواره مبتنی بر رابطه کارفرمایی-کارگری[42] بوده است. معیارهای نخستین جهت تشخیص و احراز این رابطه، در چارچوب نظریه سنتی «ارباب و خدمتکار»[43] صورت‌بندی گردید که مبتنی بر عنصر کنترل مستقیم، جزئی و آمرانه ارباب بر شیوه اجرای تکالیف خدمتکار بود. معهذا، با دگرگونی ساختاری در نظامات اقتصادی و اجتماعی از اواخر قرن هفدهم به بعد و در پی انقلاب صنعتی، دکترین سنتی کنترل مستقیم کارآمدی عملیاتی خود را از دست داد و عدم انطباق ماهوی خود را با واقعیات حاکم بر بازار کار، که متضمن ظهور نیروهای کار متخصص و برخوردار از استقلال عملکردی قابل توجه بود، آشکار ساخت (Giliker, 2022, p. 4). این تحولات منجر به پیدایش درک گسترده‌تر و انعطاف‌پذیرتری از رابطه استخدامی گردید.

دکترین معاصر مسئولیت نیابتی، با عبور از الزامات شکلی و صوری، دیگر بر احراز یک قرارداد استخدامی مدون یا قابلیت اِعمال کنترل جزئی‌نگر بر روش‌های انجام کار استوار نیست (ژوردن، ۱۳۹۴، ۱۴۸-۱۴۹). رویه قضایی، در پرتو رأی مهم دیوان عالی بریتانیا در پرونده CCWS، به سمت یک رویکرد ماهیت‌گرا و انعطاف‌پذیر حرکت کرده است. این رویکرد، قابلیت انتساب مسئولیت را در چارچوب روابطی به رسمیت می‌شناسد که هرچند فاقد وصف حقوقی استخدام رسمی هستند، اما در جوهره خود، قرابت و شباهت قابل ملاحظه‌ای با آن دارند؛ روابطی که در ادبیات حقوقی تحت عنوان «شبیه به استخدام»[44]مفهوم‌پردازی شده‌اند (Giliker, 2022, p. 128). دادگاه در رأی مذکور، با تحلیل پرونده فرعی GE (ناظر بر مسئولیت یک نهاد کلیسایی در قبال آزار

---

40 Akin to Employment
41 Close Connection
42Employer-Employee
43Master-Servant
44 Akin to Employment

جنسی ارتکابی توسط کشیش فاقد قرارداد استخدامی)[45]، تأکید نمود که برای احراز رابطه موجد مسئولیت، باید از عناوین ظاهری فراتر رفته و بر «ویژگی‌های حیاتی»[46]آن رابطه تمرکز کرد (CCWS, 2012, para 48) که مهم‌ترین آن، این است که آیا عامل از طرف اصیل منصوب شده تا به نیابت از او[47] عمل کند یا خیر. این تحول مفهومی، پژواک خود را در جدیدترین پیشنهادهای حقوقی برای راهبری هوش مصنوعی نیز یافته است؛ تأکید پارلمان اروپا بر مسئولیت «اپراتور» سیستم‌های پرخطر و عدم امکان فرار از مسئولیت با استناد به استقلال عملکرد سیستم (ذاکری نیا، ۱۴۰۲، ۱۴۹)، نشان از همگرایی جهانی به سمت شناسایی مسئول در جایگاه کنترل‌کننده و منتفع اصلی دارد.

این جابجایی کانون مسئولیت از کاربر مستقیم به یک نهاد کنترل‌کننده، در مواردی مانند خودروهای خودران، اهمیت بیشتری می‌یابد، چرا که در این سیستم‌ها، کاربر انسانی حتی در صورت حضور فیزیکی، غالباً در حکم مسافر تلقی شده و به دلیل عدم مشارکت فعال در کنترل وسیله، بحث تقصیر او منتفی می‌گردد (مشهدی زاده و قلی نیا، ۱۴۰۱، ۳۰۹–۳۱۰). این بازتعریف نقش‌ها، که در آن سیستم خودران به «عامل» یا «ابزار» تحت کنترل راهبردی یک «اصیل» (سازنده یا ارائه‌دهنده خدمات مثلا شرکت دارای تاکسی خودران) تبدیل می‌شود، زمینه را برای انطباق این رابطه با الگوی «شبیه به استخدام» به بهترین شکل فراهم می‌سازد. در تلاش برای توصیف این رابطه، ممکن است در نگاه نخست نظریه کلاسیک نمایندگی به ذهن متبادر شود، اما این رویکرد، با در نظر گرفتن استقلال عملکردی هوش مصنوعی، قابل پذیرش نیست چرا که لازمه نمایندگی سنتی، انطباق کامل اراده نماینده با اراده و نفع اصیل است، حال آنکه یک هوش مصنوعی پیشرفته ممکن است اهداف کلی (مانند بهینه‌سازی خود) را دنبال کند که لزوماً منطبق بر نفع مستقیم و آنی کاربر نیست (ولی پور و اسماعیلی، ۱۴۰۰، ۵) و این استقلال نسبی در عملکرد، انطباق کامل با مفهوم سنتی نمایندگی را دشوار می‌سازد.

چارچوب «رابطه شبیه به استخدام» که بر ماهیت و کارکرد تمرکز دارد، برای تحلیل رابطه اصیل و هوش مصنوعی از نظر حقوقی و منطقی مناسب و مدلل است چرا که سیستم‌های هوش مصنوعی، فراتر از صرف ابزار بودن، به گونه‌ای در عملیات نهادها ادغام می‌شوند که وظایف محوله را به نیابت و در جهت تحقق اهداف سازمانی اصیل به انجام می‌رسانند؛ خواه این وظیفه جایگزینی یک راننده انسانی باشد، یا ایفای نقش یک مدیر گزینش یا حتی یک مشاور. این انتصاب کارکردی

[45] GE v The Trustees of the Portsmouth Roman Catholic Diocesan Trust [2012].
[46] Crucial Features
[47] on its behalf

که از طریق طراحی، برنامه‌ریزی و استقرار سیستم صورت می‌گیرد، جوهره اصلی معیار «منصوب و مجاز شدن برای عمل از طرف اصیل» را، که در رویه قضایی برای روابط شبیه به استخدام تبیین شده، به طور کامل منعکس می‌کند و اولین شرط لازم برای استقرار مسئولیت نیابتی را فراهم می‌آورد.

در پرتو این تحلیل، امتناع از شناسایی وضعیت «شبیه به استخدام» برای هوش مصنوعی صرفاً به دلیل ماهیت غیرانسانی آن، یک عقب‌گرد نظری محسوب شده و می‌تواند منجر به ایجاد خلاء مسئولیت و تضعیف اهداف حقوق که از جمله آن جبران خسارت و بازدارندگی است (شیری، ۱۳۹۶، ۷۴-۷۳) شود. افزون بر این، انطباق رابطه اصیل و هوش مصنوعی با الگوی «شبیه به استخدام» از طریق تحلیل هم‌زمان دو مؤلفه «کنترل» و «انتفاع» که در کانون منطق حقوقی مسئولیت نیابتی قرار دارند نیز قابل توجیه است؛ همان‌طور که پیشتر تشریح شد(در قسمت ۳-۲ و ۲-۲)، اصیل از طریق تعیین اهداف، پیکربندی معماری سیستم، مدیریت داده‌های آموزشی و تعریف محدودیت‌های عملیاتی، یک کنترل هنجاری و راهبردی بر رفتار سیستم اعمال می‌کند و از سوی دیگر، اصیل به طور مستقیم از کارایی، بهره‌وری و قابلیت‌های ناشی از به‌کارگیری همان سیستم منتفع می‌گردد. این دوگانگی کنترل و انتفاع، حتی در تعاریف فنی نیز بازتاب یافته است. به عنوان مثال، سازمان بین‌المللی استانداردسازی[۴۸] ربات را یک «مکانیسم فعال‌شده»[۴۹] و «قابل برنامه‌ریزی» توصیف می‌کند که برای انجام «وظایف مورد نظر» به کار گرفته می‌شود (Simmler & Markwalder, 2019, p. 5). قید «فعال‌شده» به قدرت اصیل در ایجاد و استمرار فعالیت اشاره دارد و عبارت «قابل برنامه‌ریزی برای وظایف مورد نظر»، بیانگر کنترل او بر غایت و کارکرد سیستم است و این هم‌زمانی کنترل بر منشأ ریسک و انتفاع از فعالیت، جوهره اصلی مسئولیت نیابتی را به طور کامل محقق می‌سازد. بنابراین، با اتکا به ویژگی‌های پیش‌گفته، می‌توان اولین شرط ماهوی برای استقرار مسئولیت نیابتی را احراز شده تلقی نمود.

۲-۴- عنصر ارتباط فعل مجرمانه هوش‌مصنوعی بر اساس معیار «ارتباط نزدیک»

پس از احراز وجود یک رابطه موجد مسئولیت، گام دوم در تحلیل مسئولیت نیابتی، ارزیابی پیوند میان فعل زیانبار عامل و آن رابطه است. این عنصر که به طور سنتی تحت عنوان ارتکاب فعل «در جریان»[۵۰]رابطه شناخته می‌شود، قلمرو مسئولیت اصیل را مشخص می‌کند و به دلیل پیچیدگی‌های تعاملات مدرن، همواره محل چالش‌های تفسیری بوده است (ژوردن، ۱۳۹۴، ۱۵۱) چنانکه

[48] ISO
[49] actuated mechanism
[50] In the Course of

تبارشناسی این عنصر، یک جستجوی طولانی برای گذار از معیارهای صلب و ناکارآمد به سمت ضابطه‌ای منعطف و کارآمد را نشان می‌دهد. در رویه قضایی رویکردهای اولیه که مسئولیت را به دستور مستقیم اصیل برای ارتکاب فعل محدود می‌کردند، به سرعت جای خود را به معیار کلاسیک «سالموند»[۵۱]دادند. معیار سالموند[۵۲]، مسئولیت را به افعال مجاز یا شیوه‌های غیرمجاز برای انجام یک فعل مجاز محدود می‌کرد، اما در مواجهه با تخلفات عمدی کارکنان که ظاهراً خارج از چارچوب وظایف شغلی بودند، با دشواری‌های جدی روبرو بود و به نتایج غیرمنصفانه می‌انجامید(Giliker, 2022, pp. 9-11).

در مواجهه با این ناکارآمدی ضوابط سنتی، نظام‌های حقوقی کامن‌لا به تدریج به سوی پذیرش یک معیار انعطاف‌پذیر و سیاست‌محور برای احراز مسئولیت نیابتی حرکت کرده‌اند که تحت عنوان «آزمون ارتباط نزدیک»[۵۳] شناخته شد. دیوان عالی بریتانیا در رأی CCWS، این معیار را چنین تبیین می‌کند که آیا فعل زیانبار عامل «آنقدر با اعمالی که او مجاز به انجامشان بوده است ارتباط نزدیک دارد که ممکن است آن فعل، به طور منصفانه و مناسب[۵۴]، در جریان عادی آن رابطه تلقی شود؟» (Giliker, 2022, pp. 8-9, 126; Campbell & Lindsay, 2024, p. 191). بر این اساس، تحلیل معاصر فراتر از جستجوی مجوز صریح از سوی اصیل است و مبتنی بر شناسایی یک پیوند ماهوی و کیفی میان فعل زیانبار و ماهیت رابطه است؛ پیوندی که انتساب مسئولیت به اصیل را از منظر عدالت توزیعی و سیاست‌گذاری حقوقی موجه سازد و محاکم برای احراز آن، به بررسی عواملی چون فرصت[۵۵] ایجاد شده توسط رابطه، ارتباط فعل با وظایف[۵۶]محوله، و تحقق ریسک‌های ذاتی[۵۷] آن فعالیت می‌پردازند.

لازمه انطباق این ضابطه با حوزه عملکرد سامانه‌های هوشمند، اتخاذ یک رویکرد کارکردی است. بدین معنا که اصیل، با تعریف اهداف غایی برای سیستم خواه بیشینه‌سازی سود و کارایی،

---

51 Salmond Test

۵۲ "معیار سالموند" (Salmond test) برای اولین بار توسط حقوقدانی به نام جان ویلیام سالِموند (John William Salmond) در کتاب زیر مطرح شد.

- Salmond, J. W. (1907). The law of torts. Stevens and Haynes.

53 Close Connection Test

54 fairly and properly

55 Opportunity

56 Connection to Duties

57 Inherent Risk

خواه بهینه‌سازی تعامل با کاربر، در واقع یک «محدوده ریسک»[۵۸] مشخص را ایجاد و مدیریت می‌کند ((Barfield & Pagallo, 2018, p.399 و در این بستر، موجودیت و کارکرد عملیاتی سامانه هوشمند، در خدمت نیل به همین اهداف از پیش مقرر شده قرار دارد. در همین راستا، پرونده کمبریج آنالیتیکا[۵۹]نمونه‌ای بارز از این پدیده است که نشان می‌دهد چگونه تعقیب یک هدف سازمانیِ ظاهراً مشروع (هدف‌گذاری تبلیغاتی) می‌تواند به طور مستقیم به نقض گسترده حقوق و آسیب‌های اجتماعی منجر شود. در این مثال، فعل زیانبار سیستم، یک انحراف غیرمرتبط از فعالیت آن نیست، بلکه تحقق یکی از ریسک‌های ذاتی است که اصیل با تعریف آن هدف خاص، آن را ایجاد کرده است. چنین فعلی از «فرصت» ایجاد شده توسط پلتفرم بهره برده و با «وظیفه» اصلی آن یعنی بهینه‌سازی برای یک هدف معین، «ارتباطی تنگاتنگ» دارد. بنابراین، پیوند میان عملکرد زیانبار هوش مصنوعی و اهداف تعیین‌شده توسط اصیل، آنچنان عمیق و ماهوی است که انتساب مسئولیت به وی بر اساس معیار «ارتباط نزدیک»، هم ممکن و هم «منصفانه و مناسب» به نظر می‌رسد.

این انطباق معیار «ارتباط نزدیک» بر کنش‌های زیانبار هوش مصنوعی را می‌توان اینگونه مستدل کرد که نفسِ تصمیم اصیل مبنی بر استقرار یک سامانه هوشمند برای انجام یک وظیفه، به خودی خود «فرصت» ایراد خسارت توسط آن سامانه را خلق می‌کند؛ فرصتی که بدون اقدام اصیل هرگز وجود نداشت. مهم‌تر آنکه، بسیاری از افعال زیانبار هوش مصنوعی، تجلی مستقیم «ریسک‌های ذاتی» و نوعاً قابل پیش‌بینی هستند که با سپردن آن وظیفه خاص به یک عامل الگوریتمی ملازمه دارند. به عنوان مثال، در مورد الگوریتم استخدام شرکت آمازون که به حذف سیستماتیک متقاضیان زن انجامید (Albert et al., 2022, p. 279)، فرصت تبعیض توسط تصمیم به خودکارسازی فرآیند گزینش ایجاد شد و فعل زیانبار، تحقق ریسک ذاتی سوگیری در داده‌های

---

[58] field of risk

[۵۹] رسوایی شرکت کمبریج آنالیتیکا (Cambridge Analytica scandal) مرتبط با حریم خصوصی داده‌ها، سوء استفاده از اطلاعات شخصی در شبکه‌های اجتماعی و تأثیرگذاری بر روندهای سیاسی است که در سال ۲۰۱۸ علنی شد. برای مطالعه تفضیلی به کتاب زیر که توسط افشاگر اصلی ماجرای کمبریج آنالیتیکا نوشته شده است و روایتی دست اول است رجوع شود. نویسنده به عنوان مدیر تحقیقات سابق شرکت، در این کتاب به تفصیل توضیح می‌دهد که چگونه این شرکت از داده‌های فیس‌بوک برای ساخت سلاح‌های روانی و تأثیرگذاری بر انتخابات در سراسر جهان، از جمله انتخابات ریاست جمهوری ۲۰۱۶ آمریکا و رفراندوم برگزیت، استفاده کرد.

- Wylie, C. (2020). Mindf* ck: Cambridge Analytica. La trama para desestabilizar el mundo. Roca Editorial.

آموزشی بود. انتساب این نتیجه به آمازون، حتی در غیاب قصد تبعیض، از آن رو «منصفانه و مناسب» است که آسیب، مستقیماً از ریسک فعالیت انتفاعی و تصمیمات عملیاتی همان شرکت ناشی شده است.

این ارتباط تنگاتنگ، در عملکرد الگوریتم‌های توصیه‌گر محتوا که در پلتفرم‌های بزرگ به کار گرفته می‌شوند (Chakraborty, 2025, pp. 378-379) نیز نمایان می‌شود. با این توضیح که این الگوریتم‌ها وظیفه دارند تعامل کاربر را به نفع اهداف تجاری پلتفرم افزایش دهند و اگر در راستای انجام این وظیفه، محتوای افراطی یا کذب را ترویج کنند (Shin, 2024, p. 57)، این کنش یک انحراف از وظیفه نیست، بلکه محصول جانبی و منطقیِ بهینه‌سازی افراطی برای همان وظیفه است. این فعل زیانبار، اگرچه نامطلوب، اما دارای پیوندی عمیق با وظیفه محوله و اهداف اصیل است و نمی‌توان آن را عملی کاملاً منفک از جریان عادی فعالیت پلتفرم دانست. بنابراین، با در نظر گرفتن این پیوند میان خلق فرصت توسط اصیل، تحقق ریسک‌های ذاتی ناشی از انتخاب فناوری توسط او و ارتباط تنگاتنگ فعل زیانبار با وظایف محوله، معیار «ارتباط نزدیک» قابلیت انطباق مؤثری با وضعیت هوش مصنوعی دارد. لذا، اعمال دکترین مسئولیت نیابتی و تلقی این افعال به عنوان امری که «به طور منصفانه و مناسب» در جریان عادی رابطه با اصیل رخ داده، با منطق حقوقی این معیار سازگار بوده و دومین و آخرین عنصر لازم برای استقرار این مسئولیت را محقق می‌سازد.

نتیجه‌گیری

مقاله حاضر در پی ارائه و توجیه چارچوبی منسجم برای انتساب مسئولیت ناشی از عملکرد سامانه‌های هوش مصنوعی در حقوق ایران بود. با در نظر گرفتن واقعیات موضوعی و تحلیل حقوقی امر این نتیجه حاصل شد که ظرفیت‌های نظری و عملی دکترین مسئولیت نیابتی، آن را به راهکاری مناسب برای مسئول شناختن اشخاص به‌کارگیرنده (اصیل) هوش مصنوعی در نظام حقوقی ایران تبدیل می‌کند. تحلیل مبانی این رویکرد نشان داد که اصول بنیادین سیاست‌گذاری از جمله منطق اقتصادیِ «بهره‌مندی از منافع و در عین حال خلق ریسک»، ضرورت «درونی‌سازی هزینه‌ها» و توانایی برتر «اصیل» در مدیریت خسارات توجیهی متقن برای مسئول دانستن نهادِ به‌کارگیرنده فراهم می‌آورد و قابل توجه است که این مفاهیم قرابت مفهومی عمیقی با قواعدی چون «من له الغنم فعلیه الغرم» و «اصل احترام» در حقوق ایران نیز دارند. همچنین، صورت‌بندی تحلیلی و نظام‌مند از ارکان این نوع مسئولیت نشان داد که معیار «رابطه شبه‌استخدامی» با در نظر گرفتن انتصاب کارکردی سامانه هوشمند توسط اصیل، کنترل وی بر اهداف و پارامترهای عملیاتی آن و انتفاع حاصله از به‌کارگیری آن، قابلیت انطباق با وضعیت هوش مصنوعی را دارد و احراز پیوند ذاتی میان فرصت‌ها و ریسک‌های مرتبط با هوش مصنوعی و وظایفی که اصیل به سامانه محول می‌کند، با در نظر گرفتن معیار «ارتباط نزدیک» میان خروجی مجرمانه سیستم و فعالیت‌های اصیل، امکان‌پذیر است. این چارچوب با پر کردن خلاء مسئولیت، نظام حقوقی را در مواجهه با هوش مصنوعی کارآمد می‌سازد و با قرار دادن مسئولیت بر عهده گروهی که از ریسک منتفع شده و آن را کنترل می‌کند، بازدارندگی را تقویت کرده و به ذینفعان، انگیزه لازم برای توسعه و استفاده مسئولانه از این فناوری را می‌دهد. در نهایت، این پژوهش با توجه به نیاز نظام حقوقی ایران به ایجاد چارچوب قانونی برای مسئولیت‌پذیری در حوزه فناوری‌های نوظهور، مبنایی تحلیلی و کاربردی در اختیار قانون‌گذار قرار می‌دهد. بر این اساس، پیشنهاد می‌شود سنگ بنای نظام مقرراتی ایران بر اصل مسئولیت اولیه منتفع یا «به‌کارگیرنده» استوار گردد.

پیشنهاد برای قانون‌گذاری در ایران

۱. مسئولیت اولیه برای «به‌کارگیرنده»

پیشنهاد می‌شود قانون‌گذار، مسئولیت ناشی از سامانه‌های هوشمند را بر عهده شخص «به‌کارگیرنده» قرار دهد. این ایده با قاعده فقهی–حقوقی «من له الغنم فعلیه الغرم» و جوهره «اصل احترام» در فقه اسلامی هم‌سو است و علاوه بر منصفانه بودن، بر این واقعیت مبتنی است که اصیل (شرکت یا سازمان) به دلیل برخورداری از منابع مالی بیشتر (جیب عمیق)، در بهترین موقعیت برای جبران خسارت زیان‌دیده قرار دارد. مسئولیت به‌کارگیرنده همچنین کاملاً با منطق اقتصادی «ارزان‌ترین پیشگیری‌کننده از هزینه» مطابقت دارد، زیرا به‌کارگیرنده در بهترین و مقرون‌به‌صرفه‌ترین جایگاه برای مدیریت، کنترل و کاهش ریسک‌های سیستم قرار دارد.

۲. مقابله با خلاء مسئولیت با به‌کارگیری معیار «رابطه شبیه به استخدام»

جهت جلوگیری از ایجاد خلأهای مسئولیت، پیشنهاد می‌شود قانون‌گذار از چارچوب‌های سنتی فراتر رفته و با استفاده از چارچوب انعطاف‌پذیر «رابطه شبه‌استخدامی»، مسئولیت را بر اساس ماهیت عملکردیِ رابطه تعریف کند. این معیار زمانی محقق می‌شود که یک سامانه هوشمند در فعالیت اصلی بنگاه ادغام شده و به نیابت از آن برای تحقق اهداف سازمانی عمل می‌کند؛ در این حالت، هوش مصنوعی به یک امتداد فناورانه از فعالیت اصیل تبدیل می‌شود، مانند استفاده از آن به جای راننده، مسئول گزینش در استخدام یا حتی مشاور روانشناسی. پذیرفتن این «انتصاب کارکردی»، که جوهره اصلی رویه قضایی مدرن است مانع از فرار شرکت‌ها از مسئولیت از طریق ساختارهای حقوقی سنتی می‌گردد.

۳. به‌کارگیری معیار «ارتباط نزدیک» برای احراز رابطه استناد

به دلیل ماهیت «جعبه سیاه» هوش مصنوعی، یک خلأ مسئولیت از باب سخت بودن احراز رابطه ستناد به وجود می‌آید. برای رفع این مشکل، پیشنهاد می‌شود قانون‌گذار معیار مدرن «ارتباط نزدیک» را برای احراز پیوند میان عمل زیان‌بار هوش مصنوعی و فعالیت اصیل به کار گیرد. این معیار با بررسی این موضوع که آیا عمل زیان‌بار، تحقق یکی از ریسک‌های ذاتی و به نوعی ملازم با فعالیتی است که اصیل با سپردن آن وظیفه به هوش مصنوعی ایجاد کرده، بر مشکل «جعبه سیاه» بودن الگوریتم‌ها فائق می‌آید. همان‌طور که در نمونه‌های عملی (مانند الگوریتم استخدامی آمازون یا الگوریتم‌های توصیه‌گر محتوا) مشاهده شد، این معیار به خوبی نشان می‌دهد که چگونه عملکرد زیانبار سیستم، ارتباطی تنگاتنگ با وظایف محوله و اهداف تجاری اصیل دارد.

۴. اتخاذ رویکرد نظارتی ریسک‌محور و تدوین استانداردهای فنی–حقوقی متناسب با ریسک‌ها

به منظور همگامی با روندهای حقوقی نوین جهانی، پیشنهاد می‌شود که نظام نظارتی ایران مدل مؤثر «نظارت ریسک‌محور» را به عنوان مبنای سیاست‌گذاری خود اتخاذ کند. این روش امکان آن را فراهم می‌آورد که الزامات قانونی و سطح مسئولیت برای هر سیستم، بر اساس سطح ریسکِ به‌کارگیری آن تعیین شود. برای عملیاتی کردن این مدل، قانون‌گذار باید نهادهای نظارتی را موظف به تدوین استانداردهای فنی–حقوقی پویا کند. این استانداردها باید جنبه‌های فنی، ایمنی و اخلاقی (مانند کاهش سوگیری) را در بر گرفته و همچنین به حل «مشکل دست‌های زیاد» کمک کنند.

منابع

فارسی

۱. برزگر، محمد رضا و الهام، غلامحسین. (۱۳۹۹). مسئولیت کیفری کاربر خودروی خودران در قبال صدمات وارده توسط آن. فصلنامه پژوهش حقوق کیفری، ۸(۳۰)، ۲۰۱-۲۲۹. doi: 10.22054/jclr.2020.39009.1838

۲. تخشید، زهرا. (۱۴۰۰). مقدمه ای بر چالشهای هوش مصنوعی در حوزه مسئولیت. حقوق خصوصی، ۱۸(۱)، ۲۲۷-۲۵۰. https://doi.org/10.22059/jolt.2021.319529.1006965

۳. حسینی، الهام السادات.، محمدیان امیری، مهدی.، و خیراللهی، محمدعلی. (۱۴۰۳). بررسی فقهی مسئولیت در فناوری هوش مصنوعی. پژوهشهای تطبیقی فقه، حقوق و سیاست، ۶(۴)، ۴۹-۶۶.

۴. حکمت نیا، محمود.، محمدی، مرتضی.، و واثقی، محسن. (۱۳۹۸). مسئولیت ناشی از تولید رباتهای مبتنی بر هوش مصنوعی خودمختار. حقوق اسلامی، ۱۵(۶۰)، ۲۳۱-۲۵۵.

۵. دیلمی، احمد، زند لشنی، ناهید. و جوادیه، علی. (۱۳۹۹). مبانی مسئولیت ناشی از فعل غیر در حقوق انگلستان و ایران. دوفصلنامه علمی دانش حقوق مدنی، ۹(۲)، ۶۷-۷۹. doi: 10.30473/clk.2021.49776.2660

۶. ذاکری نیا، حانیه. (۱۴۰۲). ماهیت و مبنای مسئولیت ناشی از هوش مصنوعی در حقوق ایران و کشورهای اتحادیه اروپا. حقوق خصوصی، ۲۰(۱)، ۱۳۵-۱۵۲. https://doi.org/10.22059/jolt.2023.356703.1007186

۷. ژوردن، پاتریس. (۱۳۹۴). اصول مسئولیت (چاپ چهارم). (مجید ادیب، مترجم). نشر میزان.

۸. شیری، عباس. (۱۳۹۶). عدالت ترمیمی (چاپ اول). نشر میزان.

۹. غمامی، سید محمد مهدی، و هاشمی، سید حسین. (۲۰۲۵). چالشهای حقوقی نظام تنظیم گری فناوریهای نوظهور مجازی در ایران. پژوهش نامه حقوق اسلامی. انتشار آنلاین پیش از چاپ. https://doi.org/10.30497/law.2025.247604.3690

۱۰. مشهدی زاده، علیرضا.، و قلی نیا، رضا. (۱۴۰۱). مسؤولیت مدنی کاربر در به کارگیری سیستم هوش مصنوعی در خودرو. مجله پژوهشهای حقوقی، ۲۱(۵۰)، ۳۰۵-۳۳۱. https://doi.org/10.48300/JLR.2021.285755.1653

۱۱. ولی پور، علی.، و اسماعیلی، محسن. (۱۴۰۰). امکان سنجی مسؤولیت مدنی هوش مصنوعی عمومی ناشی از ایجاد ضرر در حقوق مدنی. فصلنامه اندیشه حقوقی معاصر، ۲(ششم)، ۱-۸.

انگلیسی

12. Albert, M. V., Lin, L., Spector, M. J., & Dunn, L. S. (Eds.). (2022). Bridging human intelligence and artificial intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-030-84729-6
13. Azad, S., & Nazari-Heris, M. (Eds.). (2024). Artificial Intelligence in the Operation and Control of Digitalized Power Systems. Springer Nature.

14. Barfield, W., & Pagallo, U. (Eds.). (2018). Research handbook on the law of artificial intelligence. Edward Elgar Publishing
15. Brodie, D. (2010). Enterprise Liability and the Common Law. Cambridge University Press.
16. Campbell, M., & Lindsay, B. (2024). Refining Vicarious Liability. Edinburgh Law Review, 28(2), 174-206.
17. Chakrabarty, I., & Baral, A. (2023). Artificial intelligence and personhood in the 21st century. Indian Journal of Integrated Research in Law, 3(1), 1–10.
18. Chakraborty, S. (Ed.). (2025). Ethical AI solutions for addressing social media influence and hate speech. IGI Global.pp. https://doi.org/10.4018/979-8-3693-9904-0
19. Charles Sr., C. (2025). Security automation with Python: Practical Python solutions for automating and scaling security operations. Packt Publishing.
20. Cristiano, F., Broeders, D., Delerue, F., Douzet, F., & Géry, A. (2023). Artificial intelligence and international conflict in cyberspace. Taylor & Francis.
21. Dremliuga, R., Kuznetcov, P., & Mamychev, A. (2019). Criteria for recognition of ai as legal person. Journal of Politics and Law, 12(3), 105–112.
22. Ebers, M., & Gamito, M. C. (2021). Algorithmic Governance and Governance of Algorithms. New York: Springer.
23. Gatti, M. (2016). European external action Service: Promoting coherence through autonomy and coordination (Vol. 11). Brill.
24. GE v The Trustees of the Portsmouth Roman Catholic Diocesan Trust 2012.
25. Gero, J. S., & Sudweeks, F. (Eds.). (2012). Artificial Intelligence in Design'96. Springer Science & Business Media.
26. Johnson, D. G. (2015). Technology with no human responsibility?. Journal of Business Ethics, 127(4), 707-715.
27. Kerrigan, C. (Ed.). (2022). Artificial intelligence: Law and regulation. Edward Elgar Publishing.
28. Königs, P. (2022). Artificial intelligence and responsibility gaps: What is the problem?. Ethics and Information Technology, 24(3), 36.
29. Krömker, H. (2022). HCI in Mobility, Transport, and Automotive Systems. Berlin, Germany: Springer.
30. Nikolinakos, N. T. (2023). EU policy and legal framework for Artificial intelligence, Robotics and related Technologies-the AI Act. Springer International Publishing AG.
31. Oliphant, K., & Wagner, G. (Eds.). (2012). Employers' liability and workers' compensation (Vol. 31). Walter de Gruyter.
32. Petrin, M., & Witting, C. (Eds.). (2023). Research Handbook on Corporate Liability. Edward Elgar Publishing.
33. Shin, D. (2024). Artificial Misinformation: Exploring Human-Algorithm Interaction Online. Springer Nature.
34. Simmler, M., & Markwalder, N. (2019). Guilty robots?–rethinking the nature of culpability and legal personhood in an age of artificial intelligence. In Criminal law forum (Vol. 30, pp. 1-31). Springer Netherlands.
35. Skandinaviska Enskilda Banken AB v Asia Pacific Breweries (S) Pte Ltd, 2011.
36. Soni, H. K., Sharma, S., & Sinha, G. R. (Eds.). (2024). Text and Social Media Analytics for Fake News and Hate Speech Detection. CRC Press.

37. Szapiro, T., & Kacprzyk, J. (Eds.). (2022). Collective decisions: Theory, algorithms and decision support systems. Springer.
38. Szentgali-Toth, B. (2021). The source of unexplored opportunities or an unpredictable risk factor? could artificial intelligences be subject to the same laws as human beings?. Public Governance, Administration and Finances Law Review (PGAF LR), 6(2), 101–120.
39. Taplin, R. (2022). Artificial intelligence, intellectual property, cyber risk and robotics: a new digital age. Routledge.
40. Various Claimants v. Catholic Child Welfare Society, 2012.
41. Wagner, A. (2025). ChatGPT Can Make Mistakes: The Liability of Programmers for Defects in Artificial Intelligence. BoD – Books on Demand.