IJCITAI
International Journal of Computer, Information Technology and Artificial Intelligence (IJCITAI)

مجله بین‌المللی کامپیوتر، فناوری اطلاعات و هوش مصنوعی

https://ijcitai.ir

دوره ۱ – شماره ۱ (پاییز ۱۴۰۴) – صفحه ۲۹ تا ۳۸

چالش‌های حقوقی و مسئولیت‌پذیری سامانه‌های هوش مصنوعی در تصمیم‌گیری‌های خودکار و اثر آن بر نظام مسئولیت مدنی و کیفری

علیرضا براتی

کارشناس ارشد حقوق خصوصی، گروه حقوق خصوصی ، واحد نورآباد ممسنی ،دانشگاه ازاد نورآباد ممسنی ، ایران

alireza.barati۵۳۹۶@iau.ir

چکیده

پیشرفت‌های سریع در فناوری هوش مصنوعی و به‌ویژه سامانه‌های تصمیم‌گیر خودکار، چالش‌های عمیقی را برای نظام‌های حقوقی موجود به وجود آورده است. این مقاله به بررسی مبانی حقوقی مسئولیت‌پذیری در قبال تصمیم‌گیری‌های خودکار هوش مصنوعی و تأثیر آن بر دکترین‌های سنتی مسئولیت مدنی و کیفری می‌پردازد. با تحلیل مفاهیم عاملیت، قصد و تقصیر در بستر فناوری‌های غیرانسانی، نشان می‌دهیم که چارچوب‌های فعلی پاسخگوی پیچیدگی‌های ناشی از خودمختاری الگوریتمی نیستند. این پژوهش با رویکردی توصیفی-تحلیلی و با مرور منابع کتابخانه‌ای معتبر، به ارزیابی راهکارهای حقوقی پیشنهادی از جمله شخصیت حقوقی برای هوش مصنوعی، مسئولیت بدون تقصیر، و مدل‌های تقسیم مسئولیت می‌پردازد. یافته‌ها حاکی از نیاز به بازنگری اساسی در مفاهیم بنیادین مسئولیت در عصر هوش مصنوعی است و پیشنهاد می‌کند که قانون‌گذاری‌های نوین می‌بایست توأمان اهداف جبران خسارت، بازدارندگی و اخلاق‌محوری را دنبال کنند.

کلمات کلیدی: هوش مصنوعی، مسئولیت مدنی، مسئولیت کیفری، تصمیم‌گیری خودکار، عاملیت الگوریتمی، شخصیت حقوقی، تقصیر، جبران خسارت.

۱. مقدمه

ظهور سامانه‌های هوش مصنوعی (AI) که قادر به تصمیم‌گیری مستقل و عمل بدون دستور مستقیم انسانی هستند، پارادایم سنتی مسئولیت حقوقی را به چالش کشیده است. از خودروهای خودران و ربات‌های جراح تا سیستم‌های اعطای وام و تشخیص چهره، هوش مصنوعی به شکلی فزاینده در حال اتخاذ تصمیماتی است که پیامدهای مستقیم و گاه جبران‌ناپذیری بر حقوق، اموال و تمامیت جسمانی افراد دارد. سؤال محوری این است: در صورت وقوع زیان یا ارتکاب جرم ناشی از تصمیم یا عمل یک سامانه هوش مصنوعی خودمختار، مسئولیت بر عهده کیست؟ آیا می‌توان موجودیتی غیرانسانی را مسئول شناخت، یا باید دایره مسئولیت را به سازندگان، برنامه‌نویسان، کاربران یا مالکان آن محدود کرد؟

نظام‌های حقوقی معاصر، اعم از مدنی و کیفری، بر پایه مفاهیمی انسانی چون تقصیر، قصد، سوءنیت، علم و اراده بنا شده‌اند (Ashrafian,۲۰۱۵). این مفاهیم در مواجهه با موجودیتی که فاقد هوشیاری، عواطف و ادراک انسانی است، دچار ابهام و ناکارآمدی می‌شوند. از سوی دیگر، پیچیدگی ذاتی برخی الگوریتم‌ها، به‌ویژه در یادگیری عمیق، که گاه حتی برای طراحانشان نیز به یک "جعبه سیاه" تبدیل می‌شوند، شناسایی رابطه سببیت و انتساب عمل را دشوار می‌سازد (Burrell,۲۰۱٦).

این مقاله در پی آن است تا با واکاوی چالش‌های مذکور، به تحلیل تأثیر تصمیم‌گیری خودکار هوش مصنوعی بر مبانی مسئولیت مدنی و کیفری بپردازد. در این راستا، ابتدا مفهوم خودمختاری در هوش مصنوعی و طیف تصمیم‌گیری از انسان‌محور تا ماشین‌محور تبیین می‌شود. سپس، چالش‌های اعمال دکترین‌های سنتی مسئولیت مدنی (از جمله تقصیر، خطر و نفی عیب) و مسئولیت کیفری مورد بررسی قرار می‌گیرد. در ادامه، راهکارهای پیشنهادی در ادبیات حقوقی و رویه‌های نوظهور تحلیل شده و در نهایت، با جمع‌بندی یافته‌ها، چشماندازها و پیشنهادهایی برای نظام‌های حقوقی ارائه می‌گردد.

۲. خودمختاری هوش مصنوعی و طیف تصمیم‌گیری

پیش از پرداخت به مباحث حقوقی، ضروری است که درکی روشن از سطوح مختلف خودمختاری در سامانه‌های هوش مصنوعی حاصل شود. خودمختاری در اینجا به معنای توانایی سیستم برای انجام وظایف و اتخاذ تصمیمات در محیط‌های پیچیده و غیرقابل پیش‌بینی، بدون نظارت یا کنترل مستقیم انسانی است (European Commission,۲۰۱۹).

همان‌گونه که مشاهده می‌شود، چالش‌های حقوقی عمدتاً از سطح ۳ به بالا آغاز می‌گردد، جایی که سیستم قادر به تفسیر محیط، انتخاب از میان گزینه‌ها و اقدام بدون دستور صریح لحظه‌ای است. ویژگی "یادگیری (Learning)" عمق این چالش را بیشتر می‌کند، چراکه رفتار سیستم پس از طراحی و حتی پس از خروج از کارخانه می‌تواند متحول شود و تحت تأثیر داده‌های جدید قرار گیرد (Matthias,۲۰۰٤). این پویایی، انتساب یک عمل خاص به دستور اولیه برنامه‌نویس یا حتی الگوی داده‌ای اولیه را با مشکل مواجه می‌سازد.

جدول ۱: سطح‌بندی خودمختاری در سامانه‌های هوش مصنوعی (بر اساس اصلاحاتی از SAE J۳۰۱۶ و EU Guidelines)

سطح	عنوان	توضیح	مثال	سطح دخالت انسانی
۰	بدون خودمختاری	تمامی وظایف توسط انسان انجام می‌شود.	ماشین حساب ساده	کامل
۱	کمک به تصمیم‌گیرنده	سیستم پیشنهادهایی ارائه می‌دهد، اما تصمیم‌گیری نهایی با انسان است.	سیستم پیشنهاد تشخیص پزشکی	بالا
۲	خودمختاری جزئی	سیستم می‌تواند برخی وظایف را در شرایط خاص به صورت خودکار انجام دهد، اما نظارت مستمر انسان ضروری است.	کروز کنترل تطبیقی در خودرو	مداوم
۳	خودمختاری مشروط	سیستم در شرایط تعریف شده می‌تواند تمامی وظایف را انجام دهد، اما درخواست می‌کند انسان در صورت نیاز مداخله کند.	خودروی خودران در شاهراه	آماده به مداخله
۴	خودمختاری بالا	سیستم در اغلب شرایط بدون نیاز به مداخله انسان عمل می‌کند. محدوده عملیاتی مشخص است.	ربات جراح در مرحله خاصی از عمل	محدود
۵	خودمختاری کامل	سیستم در تمام شرایط، با عملکردی برابر یا فراتر از انسان، بدون نیاز به هیچ مداخله‌ای عمل می‌کند.	عامل هوشمند کاملاً مستقل (فرضی)	هیچ

۳. چالش‌های اعمال مسئولیت مدنی سنتی

مسئولیت مدنی به دنبال جبران خسارت شخص زیان‌دیده است و معمولاً بر پایه دو مبنا استوار است: **مسئولیت مبتنی بر تقصیر** و **مسئولیت مبتنی بر خطر یا بدون تقصیر**. هر دو رهیافت در مواجهه با هوش مصنوعی با موانعی روبه‌رو هستند.

۳.۱. مسئولیت مبتنی بر تقصیر

در این مدل، زیان‌دیده باید ثابت کند که خوانده با فعل یا ترک فعل عمدی یا غیرعمدی خود، مرتکب تقصیر شده و این تقصیر سبب ورود زیان شده است. تقصیر معمولاً به معنای عدول از رفتار یک انسان متعارف (الگوی رفتار معقول) در شرایط و موقعیت مشابه است.

- **چالش نخست: انتساب تقصیر به چه کسی؟** در یک زنجیره طولانی، چندین actor ممکن است دخیل باشند: **سازنده سخت‌افزار، توسعه‌دهنده الگوریتم/نرم‌افزار، تأمین‌کننده داده‌های آموزشی، مالک یا بهره‌بردار** سیستم، و حتی **کاربر نهایی**. تعیین اینکه تقصیر کدام یک از این اشخاص منجر به نتیجه زیان‌بار شده، به دلیل پیچیدگی فنی و اشتراک عوامل، بسیار دشوار است (Chesterman, ۲۰۲۰). آیا تقصیر از برنامه‌نویسی ناقص است، یا از داده‌های مغرضانه، یا از عدم نگهداری مناسب توسط مالک، یا از ترکیبی از این عوامل؟
- **چالش دوم: معیار "انسان متعارف" چه معنایی دارد؟** معیار رفتار معقول انسانی برای ارزیابی عملکرد یک سیستم غیرانسانی نامناسب به نظر می‌رسد. آیا باید سیستم را با یک "متخصص متعارف" در آن حوزه مقایسه کرد؟ یا باید

استانداردهای صنعتی و فنی را ملاک قرار داد؟ این امر به ویژه زمانی مشکل‌ساز است که سیستم از توانایی‌های انسانی فراتر رود یا در حوزه‌ای کاملاً جدید عمل کند (Pagallo, ۲۰۱۳).

- **چالش سوم: فقدان علم و قصد.** تقصیر غالباً مستلزم عنصری از آگاهی یا بی‌احتیاطی است. یک الگوریتم فاقد آگاهی است. بنابراین، نمی‌توان به معنای سنتی، تقصیر را به خود سیستم نسبت داد. متمرکز شدن بر تقصیر اشخاص انسانی مرتبط نیز، به دلیل مشکل "جعبه سیاه" (عدم شفافیت در تصمیم‌گیری داخلی سیستم)، با مانع اثباتی مواجه است (Zerilli et al., ۲۰۱۹).

۳.۲. مسئولیت مبتنی بر خطر یا بدون تقصیر

برای رفع برخی مشکلات مذکور، در بسیاری از حوزه‌ها (مانند مسئولیت تولیدکننده) از نظریه مسئولیت مطلق یا مبتنی بر خطر استفاده می‌شود. در این نظریه، برای مسئول شناختن شخص (مانند تولیدکننده)، نیازی به اثبات تقصیر او نیست؛ کافی است ثابت شود محصول او معیوب بوده و این عیب سبب ورود زیان شده است.

- **مزیت نسبی:** این رویکرد بار اثبات را از دوش زیان‌دیده برمی‌دارد و با توزیع خطر فعالیت‌های پرخطر اما سودمند بر عهده سازنده، انگیزه بیشتری برای ایمن‌سازی ایجاد می‌کند (European Parliament, ۲۰۲۰). این مدل برای برخی مصادیق هوش مصنوعی مانند خودروهای خودران قابل تطبیق‌تر به نظر می‌رسد.
- **چالش اصلی: تعریف "عیب".** در حقوق سنتی محصول، عیب می‌تواند در طراحی، ساخت یا هشدار ندادن باشد. اما عیب در یک سیستم هوش مصنوعی پیچیده چیست؟ آیا سیستم‌ای که مطابق با آخرین استانداردهای فنی روز طراحی شده، اما به دلیل یک تعامل غیرمنتظره با محیط باعث حادثه شده، معیوب محسوب می‌شود؟ آیا "یادگیری" سیستم پس از عرضه به بازار می‌تواند ایجادکننده عیب تلقی شود؟ تعیین ملاکی برای "عملکرد ایمن مورد انتظار" یک سیستم خودمختار کار ساده‌ای نیست (Bertolini & Episcopo, ۲۰۲۲).
- **چالش تخصیص مسئولیت:** حتی با پذیرش مسئولیت مطلق، این سؤال باقی است که کدام نهاد در زنجیره ارزش باید مسئول شناخته شود. آیا تولیدکننده نهایی است، یا توسعه‌دهنده الگوریتم کلیدی؟ این امر می‌تواند به دادگاه‌ها اجازه دهد تا با توجه به میزان کنترل و سود هر یک از طرفین، مسئولیت را توزیع کنند.

۴. چالش‌های اعمال مسئولیت کیفری سنتی

مسئولیت کیفری، که متضمن مجازات است، شرایط دشوارتری دارد و بر عناصری مانند **قصد مجرمانه** (mens rea)، **عمل مجرمانه** (actus reus)، **تطابق قصد و عمل** و **سببیت** استوار است. تطبیق این عناصر با هوش مصنوعی به مراتب چالش‌برانگیزتر است.

- **فقدان قصد و علم:** یک سیستم هوش مصنوعی فاقد ذهنیت، نیت، سوءنیت، عمد یا حتی بی‌احتیاطی کیفری است. بنابراین، انتساب عنصر روانی جرم به آن غیرممکن به نظر می‌رسد (Hallevy, ۲۰۱۰). حتی اگر سیستم طوری برنامه‌ریزی شده باشد که عملی ممنوع را انجام دهد (مثلاً یک ربات جنگی برای هدف گرفتن غیرنظامیان)، این "قصد" از آنِ برنامه‌نویس یا فرمانده است، نه خود سیستم.

- **عمل مجرمانه:** اگر چه یک سیستم هوش مصنوعی می‌تواند عملی فیزیکی انجام دهد (مانند تصادف کردن، تزریق دارو) یا حتی عملی مجازی (مانند انتقال غیرقانونیfunds)، اما این عمل فاقد اراده اخلاقی و کیفری است. با این حال، برخی نظریه‌پردازان استدلال می‌کنند که برای اهداف کاربردی حقوق کیفری، می‌توان صرف وقوع عمل فیزیکی خطرناک توسط یک موجود غیرانسانی را کافی دانست، مشابه مسئولیت کیفری برای اشخاص حقوقی (نهادهای شرکتی) که امروزه در بسیاری از نظام‌ها پذیرفته شده است (Gabriel,۲۰۱۸).
- **سببیت:** همانند مسئولیت مدنی، اثبات رابطه سببیت مستقیم بین دستور انسان و نتیجه نهایی در سیستم‌های پیچیده و خودآموز می‌تواند بسیار مشکل باشد.
- **مجازات:** هدف مجازات‌های کیفری معمولاً سزادهی، اصلاح مجرم و بازدارندگی است. مجازات یک موجود غیرانسانی (مانند از کار انداختن، پاک‌کردن یا "اصلاح" آن) معنای سزادهی ندارد، اما ممکن است جنبه بازدارندگی برای سازندگان و کاربرانش داشته باشد.

به دلیل این موانع جدی، اکثر صاحب‌نظران معتقدند اعمال مستقیم مسئولیت کیفری به خود سیستم هوش مصنوعی در چارچوب فعلی ناممکن است. تمرکز اصلی باید بر مسئولیت **اشخاص حقیقی و حقوقی مرتبط** (سازندگان، بهره‌برداران، فرماندهان) باشد، آن هم عمدتاً بر مبنای تقصیر (مانند بی‌احتیاطی، عدم نظارت کافی) یا از طریق توسعه نظریه‌های معاونیت و مشارکت در جرم (Hallevy, ۲۰۱۰).

۵. راهکارهای پیشنهادی و رهیافت‌های نوین

برای فائق آمدن بر این چالش‌ها، راهکارهای متعددی در محافل آکادمیک، نهادهای قانون‌گذاری و سازمان‌های بین‌المللی مطرح شده است.

۵.۱. ایجاد "شخصیت حقوقی الکترونیکی" یا "شخصیت برای ربات"

برخی، به پیروی از الگوی شخصیت حقوقی برای شرکت‌ها، پیشنهاد ایجاد یک شخصیت حقوقی مجزا برای سامانه‌های خودمختار پیشرفته را داده‌اند (European Parliament,۲۰۱۷). تحت این مدل، سیستم هوش مصنوعی می‌تواند دارای اموال (مانند یک صندوق جبران خسارت اجباری)، حقوق و تکالیف محدود شود. در صورت ورود زیان، ابتدا از دارایی این "شخص" خسارت جبران می‌شود. این مدل می‌تواند دسترسی به جبران خسارت را برای زیان‌دیده تسهیل کند.

- **منتقدان** استدلال می‌کنند که این کار خطر "خفگی مسئولیت "(responsibility gap) را افزایش می‌دهد، چرا که به سازندگان و بهره‌برداران اجازه می‌دهد پشت این شخصیت مصنوعی پنهان شوند و از مسئولیت شخصی شانه خالی کنند. همچنین، اعطای شخصیت حقوقی می‌تواند بهانه‌ای برای مطالبه حقوق برای ربات‌ها شود که خود بحثی اخلاقی و فلسفی پیچیده است (Bryson,Kime-Dyche & Danks, ۲۰۱۷).

۵.۲. توسعه یک رژیم ویژه مسئولیت برای هوش مصنوعی

رهیافت عملی‌تر، ایجاد یک چارچوب قانونی خاص برای هوش مصنوعی‌های پرخطر است. این رژیم می‌تواند ترکیبی از عناصر زیر باشد:

- **مسئولیت مطلق یا مبتنی بر خطر برای توسعه‌دهندگان و تولیدکنندگان:** به ویژه برای حوزه‌های پرخطر مانند کاربردهای پزشکی، حمل و نقل، و انرژی.
- **الزام بیمه اجباری:** تولیدکنندگان یا بهره‌برداران ملزم به داشتن بیمه مسئولیت برای پوشش خسارات بالقوه شوند.
- **ایجاد صندوق‌های جبران خسارت:** برای مواردی که منبع زیان قابل شناسایی نیست یا خسارت از حد پوشش بیمه فراتر می‌رود.
- **معکوس کردن بار اثبات:** در موارد خاص، بار اثبات بی‌تقصیری یا عدم عیب محصول بر عهده تولیدکننده گذاشته شود.
- **تقسیم مسئولیت متناسب:** دادگاه بتواند مسئولیت را بر اساس میزان کنترل، سود و تقصیر هر یک از بازیگران زنجیره ارزش تقسیم کند.

اتحادیه اروپا در پیش‌نویس "قانون هوش مصنوعی (AI Act)" خود و نیز در اسناد مرتبط با مسئولیت، به سمت چنین رهیافتی حرکت کرده است (European Commission, ۲۰۲۲). در این مدل، تمرکز بر "انسان محوری" و اطمینان از جبران مؤثر خسارت وارده به افراد است.

۵.۳. الزام به شفافیت، تبیین‌پذیری و ثبت وقایع

یکی از ریشه‌های مشکل مسئولیت، عدم شفافیت است. مقررات می‌توانند توسعه و استقرار سیستم‌های "پر-خطر" را منوط به رعایت اصولی کنند مانند:

- **قابلیت تبیین:** توانایی توضیح علت تصمیم گرفته شده به زبانی قابل درک برای انسان.
- **ثبت وقایع (Logging & Data Recording):** الزام به نصب "جعبه سیاه" (مانند Event Data Recorder در خودروهای خودران) که داده‌های ورودی، پارامترهای تصمیم‌گیری و اقدامات سیستم را در بازه زمانی قبل از حادثه ذخیره کند. این امر اثبات سببیت را ممکن‌تر می‌سازد (Umbrello, ۲۰۲۱).
- **ارزیابی انطباق و تأیید پیش از بازار:** سیستم‌های پرخطر باید قبل از عرضه، از نظر ایمنی، امنیت و انطباق با مقررات ارزیابی شوند.

۵.۴. مسئولیت کیفری برای اشخاص مرتبط

در حوزه کیفری، راهکار اصلی توسعه دکترین‌های موجود برای پوشش رفتارهای انسان‌ها در قبال استفاده از هوش مصنوعی است. این شامل:

- **مسئولیت ناشی از بی‌احتیاطی:** اگر شخصی با بی‌احتیاطی از یک سیستم هوش مصنوعی استفاده کند یا آن را به کار گیرد (مثلاً بدون نظارت کافی)، و این بی‌احتیاطی منجر به نتیجه مجرمانه شود، می‌توان او را مسئول دانست.
- **معاونت و مشارکت:** اگر فردی با علم و قصد، از یک سیستم هوش مصنوعی به عنوان ابزاری برای ارتکاب جرم استفاده کند (مثلاً با برنامه‌ریزی سیستم برای انجام کلاهبرداری)، می‌توان او را به عنوان مرتکب اصلی یا معاون جرم محکوم کرد.
- **مسئولیت فرماندهی و ریاست:** در موارد نظامی یا سازمانی، فرمانده یا رئیس می‌تواند به دلیل عدم اعمال نظارت و کنترل مؤثر بر سیستم‌های تحت امرش مسئول شناخته شود.

جدول ۲: مقایسه رهیافت‌های مختلف به مسئولیت در قبال هوش مصنوعی

نمونه احتمالی کاربرد	معایب	مزایا	رهیافت
سیستم‌های سطح ۱ و ۲ هوش مصنوعی (کمک‌کننده).	ناکارآمد در موارد خودمختاری بالا، مشکل در اثبات.	حفظ ثبات حقوقی، عدم نیاز به قانون‌گذاری جدید.	تطبیق قوانین موجود (مسئولیت تقصیر)
محصولات فیزیکی مجهز به هوش مصنوعی (مثل اسباب‌بازی، دستگاه‌های پزشکی).	مشکل در تعریف "عیب"، ممکن است نوآوری را خفه کند.	تسهیل جبران خسارت برای زیان‌دیده، انگیزه برای ایمن‌سازی.	تطبیق قوانین موجود (مسئولیت مطلق محصول)
ربات‌های بسیار خودمختار با دارایی مستقل (آینده‌نگرانه).	خطر خفگی مسئولیت انسان‌ها، پیچیدگی فلسفی و حقوقی.	ایجاد وضوح، تسهیل جبران خسارت از طریق دارایی اختصاصی.	شخصیت حقوقی برای ربات
سیستم‌های هوش مصنوعی پر-خطر (خودروهای خودران سطح ۴ و ۵، سیستم‌های تشخیص پزشکی).	نیاز به قانون‌گذاری جدید و پیچیده، احتمال تداخل مقررات.	انعطاف‌پذیری، هدف‌گیری دقیق‌تر ریسک‌ها، ترکیب ابزارهای مختلف.	رژیم ویژه مسئولیت
مکمل ضروری برای هر رژیم مسئولیت، به ویژه در حمل و نقل خودران.	هزینه اضافه برای صنعت، ممکن است رفتار بی‌احتیاطی را تشویق کند.	تضمین جبران خسارت برای قربانیان، توزیع ریسک در جامعه.	تأکید بر بیمه و صندوق جبران

۶. جمع‌بندی و نتیجه‌گیری

سامانه‌های هوش مصنوعی خودمختار نه تنها فناوری را متحول می‌کنند، بلکه بنیان‌های فلسفی و حقوقی جوامع انسانی را به پرسش می‌کشند. چالش اصلی در تقاطع دو جهان قرار دارد: جهان سنتی مسئولیت که بر محوریت انسان، قصد و تقصیر می‌چرخد، و جهان جدید عامل‌های غیرانسانی که با الگوهای پیچیده آماری و بدون درک انسانی تصمیم می‌گیرند.

نظام‌های حقوقی نمی‌توانند در برابر این تحول واکنشی منفعلانه داشته باشند. یافته‌های این مقاله نشان می‌دهد که قواعد سنتی مسئولیت مدنی و کیفری، به ویژه در مواجهه با سیستم‌های با خودمختاری بالا و قابلیت یادگیری، ناکارآمد و ناکافی هستند. خطر ایجاد "شکاف مسئولیت (Responsibility Gap)" که در آن هیچ شخص انسانی قابل سرزنش نباشد و زیان نیز جبران نشود، واقعی است.

برای رویارویی با این چالش، نمی‌توان یک راه‌حل واحد و جهانی ارائه داد. بلکه، نیاز به یک **چارچوب چندلایه و منعطف** است که بتواند با توجه به سطح خودمختاری، حوزه کاربرد و میزان خطر سیستم، پاسخ مناسب حقوقی را فراهم کند. به نظر می‌رسد مؤثرترین راهکار، توسعه **یک رژیم ویژه مسئولیت برای کاربردهای پر-خطر هوش مصنوعی** باشد که ترکیبی هوشمندانه از مسئولیت مطلق برای تولیدکنندگان، الزامات شدید بیمه‌ای، ثبت اجباری داده‌ها و امکان تقسیم مسئولیت را در خود جای دهد. این رژیم باید در کنار توسعه دکترین‌های مسئولیت کیفری برای اشخاص مرتبط (بر مبنای بی‌احتیاطی و معاونت) عمل کند.

پیشنهادهای سیاستی:

۱. **سطح‌بندی ریسک:** قانون‌گذاری باید سیستم‌های هوش مصنوعی را بر اساس میزان خطر (ریسک) دسته‌بندی کند و قواعد سخت‌گیرانه‌تر مسئولیت را برای دسته‌های پر-خطر (مانند سیستم‌های مرتبط با سلامت، امنیت و حقوق اساسی) اعمال نماید.
۲. **تأکید بر جبران خسارت:** هدف اولیه قوانین مدنی جدید باید تضمین جبران سریع و مؤثر خسارت برای زیان‌دیدگان باشد، حتی اگر تعیین مقصر دقیق مشکل باشد. مکانیسم‌هایی مانند بیمه اجباری و صندوق‌های جبران جمعی باید مورد توجه قرار گیرند.
۳. **شفافیت و پاسخگویی:** الزام به قابلیت تبیین و ثبت داده‌ها باید به عنوان یک اصل اساسی در طراحی و استقرار سیستم‌های خودمختار گنجانده شود تا امکان ارزیابی و پیگیری حقوقی پس از وقوع حادثه فراهم شود.
٤. **همکاری بین‌المللی:** ماهیت فرامرزی فناوری هوش مصنوعی مستلزم هماهنگی بین نظام‌های حقوقی مختلف است. تلاش‌های سازمان‌هایی مانند یونسیترال، OECDو شورای اروپا برای تدوین اصول مشترک باید تقویت گردد.
٥. **نظارت مستمر و به‌روزرسانی:** قوانین مصوب باید به گونه‌ای طراحی شوند که بتوانند با شتاب پیشرفت فناوری به‌روز شوند. ایجاد نهادهای نظارتی تخصصی در حوزه هوش مصنوعی می‌تواند در این زمینه مفید باشد.

در نهایت، تنظیم‌مقررات هوش مصنوعی و مسئولیت ناشی از آن، تنها یک چالش فنی نیست؛ بلکه آزمونی برای تعهد جامعه به اصول عدالت، اخلاق و محوریت انسان در عصر فناوری‌های نوظهور است. رویکردی متوازن که هم از نوآوری حمایت کند و هم از حقوق و ایمنی افراد محافظت نماید، کلید ساختن آینده‌ای است که در آن انسان و ماشین بتوانند به شکلی مسئولانه و سودمند همزیستی داشته باشند.

منابع

۱. Ashrafian, H. (۲۰۱۵). Artificial intelligence and robot responsibilities: Innovating beyond rights. *Science and Engineering Ethics*, ۲۱(۲), ۳۱۷-۳۲۶.

۲. Bertolini, A., & Episcopo, F. (۲۰۲۲). The European approach to AI and liability: A critical analysis. *European Journal of Risk Regulation*, ۱۳(۱), ۱-۲۱.

۳. Bryson, J. J., Kime-Dyche, J., & Danks, D. (۲۰۱۷). Of robots and responsibilities: Liability in the age of AI. *The Journal of Sociotechnical Critique*, ۱(۱), ۱-۲۰.

۴. Burrell, J. (۲۰۱۶). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, ۳(۱).

۵. Chesterman, S. (۲۰۲۰). Artificial intelligence and the limits of legal personality. *International & Comparative Law Quarterly*, ٦۹(۴), ۸۱۹-۸٤٤.

۶. European Commission. (۲۰۱۹). *Ethics guidelines for trustworthy AI*. High-Level Expert Group on Artificial Intelligence.

۷. European Commission. (۲۰۲۲). *Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*. COM(۲۰۲۲) ٤۹٦ final.

۸. European Parliament. (۲۰۱۷). *Resolution of ۱٦ February ۲۰۱۷ with recommendations to the Commission on Civil Law Rules on Robotics (۲۰۱۵/۲۱۰۳(INL)*.

۹. European Parliament. (۲۰۲۰). *The European framework on ethical aspects of artificial intelligence, robotics and related technologies*. Study for the JURI Committee.

۱۰. Gabriel, I. (۲۰۱۸). Artificial intelligence, values, and alignment. *Minds and Machines*, ۳۰(۳), ٤۱۱-٤۳۷.

۱۱. Hallevy, G. (۲۰۱۰). The criminal liability of artificial intelligence entities. *Social Science Research Network*.

۱۲. Matthias, A. (۲۰۰۴). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, ٦(۳), ۱۷۵-۱۸۳.

۱۳. Pagallo, U. (۲۰۱۳). *The laws of robots: Crimes, contracts, and torts*. Springer.

۱۴. Umbrello, S. (۲۰۲۱). Coupling levels of abstraction in understanding meaningful human control of autonomous weapons: A two-tiered approach. *Ethics and Information Technology*, ۲۳(۳), ٤٥٥-٤٦٤.

۱۵. Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (۲۰۱۹). Transparency in algorithmic and human decision-making: Is there a double standard? *Philosophy & Technology*, ۳۲(۴), ٦٦۱-٦۸۳.

۱۶. Beck, S. (۲۰۱۶). The problem of ascribing legal responsibility in the case of robotics. *AI & Society*, ۳۱(۴), ٤٧٣-٤٨١.
۱۷. Calo, R. (۲۰۱۵). Robotics and the lessons of cyberlaw. *California Law Review*, ۱۰۳(۳), ٥١٣-٥٦٣.
۱۸. Čerka, P., Grigienė, J., & Sirbikytė, G. (۲۰۱۵). Liability for damages caused by artificial intelligence. *Computer Law & Security Review*, ۳۱(۳), ٣٧٦-٣٨٩.
۱۹. Elish, M. C., & boyd, d. (۲۰۱۸). Situating methods in the magic of Big Data and AI. *Communication Monographs*, ۸٥(۱), ٥٧-٨٠.
۲۰. Gless, S., Silverman, E., & Weigend, T. (۲۰۱۶). If robots cause harm, who is to blame? Self-driving cars and criminal liability. *New Criminal Law Review*, ۱۹(۳), ٤١٢-٤٣٦.
۲۱. Goodall, N. J. (۲۰۱۶). Away from trolley problems and toward risk management. *Applied Artificial Intelligence*, ۳۰(۸), ٨١٠-٨٢١.
۲۲. Kerr, I. (۲۰۱۹). The legibility of algorithms: The challenge of governing artificial intelligence. *University of Toronto Law Journal*, ٦٩(S٢), ١٢١-١٣٩.
۲۳. Koops, B. J., et al. (۲۰۱۰). Bridging the accountability gap: Rights for new entities in the information society? *Minnesota Journal of Law, Science & Technology*, ۱۱(۲), ٤٩٧-٥٦١.
۲۴. Marchant, G. E., & Lindor, R. A. (۲۰۱۲). The coming collision between autonomous vehicles and the liability system. *Santa Clara Law Review*, ٥٢(۴), ١٣٢١-١٣٤٠.
۲۵. Nissenbaum, H. (۱۹۹۶). Accountability in a computerized society. *Science and Engineering Ethics*, ٢(۱), ٢٥-٤٢.
۲۶. Rahwan, I., et al. (۲۰۱۹). Machine behaviour. *Nature*, ٥٦٨(٧٧٥٣), ٤٧٧-٤٨٦.
۲۷. Scherer, M. U. (۲۰۱۶). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, ٢٩(۲), ٣٥٣-٤٠٠.
۲۸. Selbst, A. D., & Barocas, S. (۲۰۱۸). The intuitive appeal of explainable machines. *Fordham Law Review*, ٨٧(۳), ١٠٨٥-١١٣٩.
۲۹. Sullivan, H. R., & Schweikart, S. J. (۲۰۱۹). Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA Journal of Ethics*, ٢١(۲), E١٦٠-١٦٦.
۳۰. Teubner, G. (۲۰۱۸). Digital personhood? The status of autonomous software agents in private law. *Ancilla Iuris*, ١-٢٠.