

مسئولیت مدنی تصمیمات خودکار هوش مصنوعی

مقدمه

گسترش سریع هوش مصنوعی از مرحله ابزارهای محاسباتی و تحلیلی عبور کرده و به سامانه‌هایی رسیده است که می‌توانند در بسیاری از موقعیت‌ها، بدون دخالت مستقیم انسان، تصمیم تولید یا اجرا کنند. اعطای اعتبار بانکی، ارزیابی ریسک، تشخیص تقلب، پیشنهاد یا رد درخواست‌های بیمه‌ای، مدیریت منابع انسانی، کنترل سامانه‌های صنعتی، خودروهای خودران و برخی کاربردهای پزشکی، نمونه‌هایی از حوزه‌هایی هستند که در آن‌ها خروجی الگوریتم می‌تواند آثار مستقیم حقوقی و اقتصادی برای اشخاص ایجاد کند.

این تحول، پرسش مهمی را در حقوق مسئولیت مدنی مطرح می‌کند: اگر تصمیم یک سامانه هوش مصنوعی به شخصی زیان وارد کند، چه کسی باید مسئول جبران خسارت باشد؟

در نظام‌های سنتی مسئولیت مدنی، معمولاً می‌توان میان رفتار زیان‌بار، عامل انسانی، تقصیر یا مبنای مسئولیت، رابطه سببیت و خسارت ارتباط برقرار کرد. اما در سامانه‌های هوش مصنوعی، این زنجیره ممکن است پیچیده‌تر شود. تصمیم نهایی می‌تواند حاصل تعامل داده‌های آموزشی، مدل یادگیری ماشین، تنظیمات نرم‌افزاری، زیرساخت فنی، اطلاعات ورودی و شرایط محیطی باشد. افزون بر این، برخی مدل‌های پیشرفته از ویژگی‌هایی برخوردارند که پیش‌بینی دقیق خروجی آن‌ها حتی برای توسعه‌دهندگان نیز دشوار است.

بنابراین، مسئله مسئولیت مدنی هوش مصنوعی صرفاً به این پرسش محدود نمی‌شود که «چه کسی سیستم را ساخته است؟» بلکه باید مشخص شود چه شخص یا اشخاصی کنترل، نظارت، بهره‌برداری یا نقش مؤثر در ایجاد خطر داشته‌اند و چگونه می‌توان رابطه میان تصمیم خودکار و زیان وارده را اثبات کرد.

۱. مفهوم تصمیم خودکار در سامانه‌های هوش مصنوعی

تصمیم خودکار را می‌توان تصمیمی دانست که فرآیند اصلی ارزیابی، انتخاب یا اقدام در آن توسط یک سامانه محاسباتی و بر اساس قواعد، مدل‌های آماری یا الگوریتم‌های یادگیری ماشین انجام می‌شود و دخالت انسان در تصمیم نهایی محدود یا حذف شده است.

این تصمیم‌ها از نظر میزان استقلال یکسان نیستند. در برخی سامانه‌ها، هوش مصنوعی صرفاً به انسان پیشنهاد می‌دهد و تصمیم نهایی توسط کاربر انسانی اتخاذ می‌شود. در سطح بالاتر، سیستم می‌تواند بدون تأیید انسان تصمیم را اجرا کند.

از منظر حقوق مسئولیت مدنی، این تفاوت اهمیت زیادی دارد. در حالت نخست، می‌توان رفتار انسان استفاده‌کننده را نیز در تحلیل مسئولیت بررسی کرد؛ اما در تصمیم‌گیری کاملاً خودکار، انتساب خسارت به یک رفتار انسانی مشخص دشوارتر می‌شود.

برای مثال، فرض کنیم یک سامانه ارزیابی اعتبار، بر اساس داده‌های موجود، درخواست دریافت وام شخصی را رد کند و شخص مدعی شود این تصمیم موجب ورود خسارت اقتصادی به او شده است. در این وضعیت باید مشخص شود:

- آیا داده‌های مورد استفاده صحیح بوده‌اند؟
- آیا الگوریتم دچار خطا یا سوگیری بوده است؟
- آیا طراحی سیستم مطابق استانداردهای حرفه‌ای انجام شده است؟
- آیا بهره‌بردار وظیفه نظارت بر سیستم را رعایت کرده است؟
- آیا تصمیم سیستم مستقیماً موجب خسارت شده است؟
- و آیا شخص زیان‌دیده می‌تواند رابطه سببیت را اثبات کند؟

این پرسش‌ها نشان می‌دهند که مسئولیت مدنی AI بیشتر از آنکه به «خودمختاری» ماشین وابسته باشد، به ساختار حقوقی و فنی زنجیره تصمیم وابسته است.

۲. آیا خود هوش مصنوعی می‌تواند مسئول مدنی باشد؟

یکی از بحث‌های نظری مهم این است که آیا می‌توان برای سامانه هوش مصنوعی شخصیت حقوقی مستقل در نظر گرفت و مسئولیت مدنی را مستقیماً متوجه آن کرد.

در حقوق سنتی، مسئولیت مدنی معمولاً بر عهده شخص حقیقی یا شخص حقوقی قرار می‌گیرد. هوش مصنوعی، در وضعیت حقوقی کنونی، در اغلب نظام‌های حقوقی به‌عنوان شخص مستقل و هم‌تراز انسان یا شرکت شناخته نمی‌شود.

حتی اگر یک سامانه از درجه بالایی از خودمختاری برخوردار باشد، صرف توانایی آن در تصمیم‌گیری به معنای برخورداری از شخصیت حقوقی نیست. از این‌رو، راهکار عملی‌تر این است که مسئولیت میان بازیگران انسانی و سازمانی دخیل در چرخه عمر سامانه توزیع شود.

این بازیگران می‌توانند شامل موارد زیر باشند:

۱. توسعه‌دهنده مدل یا نرم‌افزار؛
۲. تولیدکننده یا عرضه‌کننده محصول مجهز به AI؛
۳. مالک یا بهره‌بردار سامانه؛

۴. ارائه‌دهنده داده یا زیرساخت؛
۵. ارائه‌دهنده خدمات هوش مصنوعی؛
۶. شخصی که سیستم را برای یک کاربرد خاص تنظیم کرده است؛
۷. اپراتور یا کاربر انسانی؛
۸. و در برخی موارد، ارائه‌دهندگان خدمات ثالث.

بنابراین، مسئله اصلی در حقوق معاصر، نه اعطای فوری شخصیت حقوقی به هوش مصنوعی، بلکه طراحی سازوکار مناسب برای انتساب مسئولیت به انسان‌ها و سازمان‌های پشت سامانه است.

۳. مبانی مسئولیت مدنی در تصمیمات AI

برای تعیین مسئولیت مدنی ناشی از تصمیمات هوش مصنوعی، چند مبنای حقوقی مهم قابل بررسی است.

۳.۱. مسئولیت مبتنی بر تقصیر

یکی از طبیعی‌ترین رویکردها، بررسی وجود تقصیر در طراحی، توسعه، آزمایش، استقرار یا بهره‌برداری از سیستم است.

تقصیر می‌تواند اشکال مختلفی داشته باشد؛ برای مثال:

- استفاده از داده‌های آموزشی نامناسب؛
- عدم آزمایش کافی مدل؛
- بی‌توجهی به خطاهای شناخته‌شده؛
- عدم به‌روزرسانی سیستم؛
- نبود نظارت انسانی در کاربردهای پرخطر؛
- طراحی نادرست فرآیند تصمیم‌گیری؛
- عدم اطلاع‌رسانی درباره محدودیت‌های سیستم؛
- یا استفاده از مدل در حوزه‌ای خارج از محدوده طراحی شده.

در این چارچوب، زیان‌دیده باید تا حد لازم نشان دهد که رفتار یا ترک فعل قابل انتساب به یک شخص یا سازمان، برخلاف استاندارد متعارف مراقبت بوده و به خسارت منجر شده است.

مشکل اصلی این رویکرد در سامانه‌های پیچیده AI، اثبات تقصیر است. ممکن است حتی متخصصان فنی نیز نتوانند به‌سادگی توضیح دهند چرا مدل در یک مورد خاص خروجی مشخصی تولید کرده است.

۳.۲. مسئولیت ناشی از نقص محصول

در مواردی که هوش مصنوعی بخشی از یک محصول محسوب می‌شود، می‌توان مسئله را از منظر مسئولیت ناشی از محصول معیوب نیز بررسی کرد.

برای مثال، اگر سیستم کنترل یک خودرو یا دستگاه صنعتی مبتنی بر AI به دلیل نقص طراحی تصمیم خطرناکی بگیرد، ممکن است قواعد مربوط به مسئولیت تولیدکننده یا عرضه‌کننده محصول مطرح شود.

در اینجا، تمرکز از رفتار شخص استفاده‌کننده به ایمنی مورد انتظار از محصول منتقل می‌شود.

در چنین رویکردی، عواملی مانند طراحی محصول، فرآیند کنترل کیفیت، هشدارهای ایمنی، قابلیت به‌روزرسانی نرم‌افزار و پیش‌بینی خطرات قابل انتظار اهمیت پیدا می‌کنند.

۳.۳. مسئولیت ناشی از فعالیت‌های خطرناک

برخی کاربردهای هوش مصنوعی می‌توانند در زمره فعالیت‌هایی قرار گیرند که به دلیل ماهیت خود، خطر بالقوه قابل توجهی ایجاد می‌کنند. در چنین شرایطی، نظام حقوقی ممکن است به سمت قواعد مسئولیت شدیدتر از مسئولیت مبتنی بر تقصیر حرکت کند.

این رویکرد به‌ویژه در کاربردهایی اهمیت پیدا می‌کند که یک خطای الگوریتمی می‌تواند خسارت جانی یا مالی گسترده ایجاد کند؛ مانند سامانه‌های حمل‌ونقل خودکار، کنترل زیرساخت‌های حیاتی یا تجهیزات پزشکی.

۴. مسئولیت توسعه‌دهنده هوش مصنوعی

توسعه‌دهنده ممکن است در صورتی مسئول شناخته شود که نقص یا خطای سیستم ناشی از تصمیمات فنی قابل انتساب به او باشد.

برای نمونه، اگر توسعه‌دهنده بدانند مدل در شرایط خاصی عملکرد نامطمئن دارد اما بدون هشدار مناسب آن را برای کاربرد حساس عرضه کند، مسئله تقصیر یا نقص محصول می‌تواند مطرح شود.

با این حال، صرف وقوع خطا لزماً به معنای مسئولیت توسعه‌دهنده نیست. مدل‌های یادگیری ماشین با احتمال و عدم قطعیت کار می‌کنند و نمی‌توان از هر سیستم هوش مصنوعی انتظار خطای صفر داشت.

بنابراین باید میان خطای قابل انتظار در یک سیستم پیچیده و نقصی که با اقدامات متعارف حرفه‌ای قابل پیشگیری بوده است تمایز قائل شد.

استاندارد مراقبت حرفه‌ای می‌تواند عواملی مانند آزمون مدل، ارزیابی ریسک، مستندسازی، کنترل کیفیت داده، بررسی سوگیری و پایش عملکرد پس از استقرار را دربرگیرد.

۵. مسئولیت بهره‌بردار و کاربر سیستم

در بسیاری از موارد، مسئولیت احتمالی تنها متوجه سازنده نیست. سازمانی که سیستم را خریداری، تنظیم و استفاده می‌کند نیز ممکن است نقش تعیین‌کننده‌ای در ایجاد خسارت داشته باشد.

فرض کنیم یک شرکت از سامانه هوش مصنوعی برای انتخاب کارکنان استفاده کند. اگر سیستم دارای محدودیت شناخته‌شده‌ای باشد و شرکت بدون ارزیابی مناسب از آن در تصمیمات استخدامی استفاده کند، نمی‌توان صرفاً توسعه‌دهنده را مسئول دانست.

بهره‌بردار باید متناسب با سطح خطر:

- سیستم را ارزیابی کند؛
- عملکرد آن را پایش کند؛
- خروجی‌های غیرعادی را بررسی کند؛
- امکان مداخله انسانی ایجاد کند؛
- دسترسی‌ها را کنترل کند؛
- و در صورت مشاهده خطا، استفاده از سیستم را متوقف یا اصلاح کند.

در اینجا مفهوم وظیفه نظارت انسانی اهمیت ویژه‌ای پیدا می‌کند.

۶. رابطه سببیت و مشکل «جعبه سیاه»

یکی از پیچیده‌ترین مسائل در مسئولیت مدنی AI، اثبات رابطه سببیت است.

در یک پرونده سنتی ممکن است بتوان نشان داد که اقدام شخص X مستقیماً موجب خسارت Y شده است. اما در سامانه هوش مصنوعی، تصمیم ممکن است حاصل زنجیره‌ای از عوامل باشد:

داده ورودی → پیش‌پردازش → مدل → پارامترها → تصمیم الگوریتمی → اقدام سازمان → واکنش شخص ثالث → خسارت.

در چنین شرایطی، مشخص کردن علت حقوقی خسارت دشوارتر می‌شود.

مشکل زمانی بیشتر می‌شود که مدل از قابلیت توضیح‌پذیری محدودی برخوردار باشد. اصطلاح «جعبه سیاه» در این زمینه به سامانه‌هایی اشاره دارد که مسیر دقیق رسیدن از ورودی به خروجی برای انسان به‌سادگی قابل فهم نیست.

البته غیرقابل توضیح بودن مدل لزوماً به معنای نبود امکان اثبات مسئولیت نیست. می‌توان از شواهد دیگری مانند سوابق سیستم، داده‌های ورودی، لاگ‌ها، نتایج آزمون، مستندات طراحی، سیاست‌های سازمان و تحلیل کارشناسی استفاده کرد.

۷. مسئولیت ناشی از داده‌های نادرست یا سوگیرانه

داده در سیستم‌های هوش مصنوعی نقش اساسی دارد. اگر داده‌های آموزشی یا عملیاتی ناقص، نادرست یا نامتوازن باشند، احتمال تصمیم نادرست افزایش می‌یابد.

برای مثال، یک سامانه ارزیابی ریسک بیمه‌ای ممکن است به دلیل داده‌های تاریخی نامناسب، گروه خاصی از افراد را بیش از حد پرخطر ارزیابی کند.

در چنین شرایطی، باید مشخص شود مسئولیت نقص داده بر عهده چه کسی است:

- تأمین‌کننده داده؟
- توسعه‌دهنده مدل؟
- سازمان بهره‌بردار؟
- یا مجموعه‌ای از این عوامل؟

پاسخ به این پرسش به کنترل هر بازیگر بر داده و وظیفه‌ای که در فرآیند پردازش بر عهده داشته است بستگی دارد.

از این‌رو، در قراردادهای مربوط به خدمات AI، تعیین مسئولیت مربوط به کیفیت داده و فرآیند اعتبارسنجی آن اهمیت زیادی دارد.

۸. سوگیری الگوریتمی و تبعیض

یکی از مهم‌ترین زمینه‌های مسئولیت ناشی از تصمیمات خودکار، تصمیم‌های تبعیض‌آمیز است.

اگر یک سیستم استخدام، اعتبارسنجی، بیمه یا ارزیابی ریسک به‌طور نظام‌مند علیه گروه خاصی از اشخاص عمل کند، موضوع می‌تواند علاوه بر مسئولیت مدنی، ابعاد حقوق بشری و ضدتبعیض نیز پیدا کند.

مسئله مهم این است که تبعیض ممکن است بدون درج مستقیم ویژگی‌هایی مانند جنسیت، نژاد یا سایر ویژگی‌های حساس در مدل رخ دهد. متغیرهای جانشین یا الگوهای پنهان موجود در داده می‌توانند باعث تولید نتایج نابرابر شوند.

بنابراین، صرف حذف یک متغیر حساس از داده‌ها الزاماً برای جلوگیری از تبعیض کافی نیست.

ارزیابی fairness یا عدالت الگوریتمی، آزمون‌های آماری، ممیزی مدل و بررسی نتایج واقعی سیستم می‌توانند در کاهش این خطر مؤثر باشند.

۹. تصمیمات خودکار در حوزه‌های حساس

شدت مسئولیت و سطح نظارت مورد نیاز باید با میزان خطر کاربرد متناسب باشد.

سلامت و پزشکی

اگر سامانه هوش مصنوعی در تشخیص یا پیشنهاد درمان نقش داشته باشد و خروجی نادرست آن موجب آسیب بیمار شود، مسئولیت می‌تواند میان پزشک، بیمارستان، تولیدکننده نرم‌افزار، ارائه‌دهنده داده و سایر عوامل تقسیم شود.

در این حوزه، AI معمولاً نباید جایگزین کامل سازوکارهای تخصصی و نظارت حرفه‌ای شود؛ زیرا پیامد خطا می‌تواند جانی باشد.

خدمات مالی

در اعتبارسنجی یا کشف تقلب، تصمیم الگوریتمی ممکن است موجب رد خدمات مالی یا ایجاد خسارت اقتصادی شود. شفافیت معیارها، صحت داده و امکان اعتراض به تصمیم اهمیت ویژه‌ای پیدا می‌کند.

استخدام

اگر تصمیم خودکار موجب رد یا پذیرش ناموجه متقاضی شود، مسئله تبعیض الگوریتمی و مسئولیت سازمان استفاده‌کننده مطرح می‌شود.

خودروهای خودران

در خودروهای خودران، زنجیره مسئولیت پیچیده‌تر است؛ زیرا نرم‌افزار، سخت‌افزار، سازنده خودرو، مالک، اپراتور، ارائه‌دهنده نقشه و سایر اجزای فنی ممکن است در نتیجه نهایی نقش داشته باشند.

این مثال‌ها نشان می‌دهند که یک رژیم واحد برای تمام کاربردهای AI لزوماً مناسب نیست و باید میان کاربردهای کم‌خطر و پرخطر تمایز ایجاد شود.

۱۰. مسئولیت قراردادی در خدمات هوش مصنوعی

بخش مهمی از کاربردهای تجاری AI در قالب قراردادهای ارائه خدمات نرم‌افزاری، ابری یا API انجام می‌شود. در این قراردادهای، مسئله مسئولیت قراردادی اهمیت ویژه‌ای دارد.

قرارداد می‌تواند درباره موضوعاتی مانند:

- سطح خدمت؛
- کیفیت و دسترسی‌پذیری سیستم؛
- مسئولیت داده؛
- امنیت اطلاعات؛
- مالکیت داده و خروجی؛
- نحوه اعلام خطا؛
- محدودیت مسئولیت؛
- جبران خسارت؛
- بیمه؛
- و شرایط توقف استفاده از مدل

تعیین تکلیف کند.

با این حال، قرارداد نمی‌تواند در همه نظام‌های حقوقی و در همه شرایط، مسئولیت قانونی را به‌طور کامل حذف کند. به‌ویژه در مواردی که حقوق اشخاص ثالث یا قواعد آمره مطرح است، توافق میان توسعه‌دهنده و بهره‌بردار لزوماً مانع مطالبه خسارت از سوی زیان‌دیده نخواهد بود.

۱۱. نقش ثبت سوابق و قابلیت حسابرسی

یکی از راهکارهای مهم برای کاهش اختلافات حقوقی، طراحی سیستم‌های AI به‌گونه‌ای است که فرآیند تصمیم‌گیری قابل ردیابی باشد.

ثبت موارد زیر می‌تواند اهمیت زیادی داشته باشد:

- نسخه مدل؛
- داده‌های ورودی؛

- زمان تصمیم؛
- تنظیمات سیستم؛
- خروجی مدل؛
- مداخلات انسانی؛
- هشدارهای سیستم؛
- تغییرات نرم‌افزاری؛
- و اقدامات انجام‌شده پس از تصمیم.

این اطلاعات می‌تواند در صورت وقوع حادثه برای بازسازی زنجیره علت و تعیین مسئولیت مورد استفاده قرار گیرد.

به همین دلیل، **auditability** یا قابلیت حسابرسی باید بخشی از طراحی فنی سامانه‌های پرخطر باشد، نه اقدامی که تنها پس از وقوع خسارت مورد توجه قرار گیرد.

۱۲. توضیح‌پذیری و حق اعتراض

در تصمیماتی که آثار جدی بر حقوق و منافع افراد دارند، امکان فهم و اعتراض به تصمیم الگوریتمی اهمیت زیادی پیدا می‌کند.

توضیح‌پذیری به این معنا نیست که همیشه باید تمام جزئیات ریاضی مدل برای کاربر ارائه شود. در بسیاری از موارد، ارائه اطلاعات قابل فهم درباره عوامل مؤثر در تصمیم، محدودیت‌های مدل و مسیر اعتراض می‌تواند کافی‌تر و کاربردی‌تر باشد.

از منظر مسئولیت مدنی، وجود سازوکار اعتراض می‌تواند امکان شناسایی خطا را پیش از تبدیل آن به خسارت گسترده افزایش دهد.

بنابراین، طراحی یک سیستم مسئولانه باید علاوه بر «دقت پیش‌بینی»، به قابلیت اعتراض، اصلاح و مداخله انسانی نیز توجه کند.

۱۳. هوش مصنوعی مولد و چالش مسئولیت

ظهور مدل‌های زبانی بزرگ و هوش مصنوعی مولد، دامنه بحث مسئولیت را گسترده‌تر کرده است. این مدل‌ها می‌توانند متن، کد، تحلیل، توصیه و محتوای تصمیم‌ساز تولید کنند.

یکی از مشکلات اصلی، تولید خروجی نادرست یا به اصطلاح hallucination است. اگر شخص یا سازمانی بر اساس چنین خروجی‌ای اقدام کند و زیان ایجاد شود، تعیین مسئولیت به عوامل متعددی بستگی خواهد داشت.

برای مثال باید بررسی شود:

- آیا ارائه‌دهنده درباره محدودیت‌های مدل هشدار داده بود؟
- آیا کاربر موظف به بررسی انسانی خروجی بود؟
- آیا سیستم برای کاربرد موردنظر مناسب بوده است؟
- آیا سازمان فرآیند اعتبارسنجی خروجی داشته است؟
- آیا خسارت نتیجه مستقیم خروجی مدل بوده یا نتیجه تصمیم مستقل کاربر؟

در این حوزه، اصل مهم آن است که تولید خودکار یک خروجی نباید به تنهایی جایگزین تحلیل حقوقی رابطه سببیت و نقش واقعی انسان‌ها شود.

۱۴. رویکرد مبتنی بر ریسک

یکی از مناسب‌ترین چارچوب‌ها برای تنظیم مسئولیت هوش مصنوعی، رویکرد مبتنی بر ریسک است.

بر اساس این رویکرد، هرچه احتمال و شدت خسارت بالقوه بیشتر باشد، الزامات مربوط به ارزیابی، نظارت، مستندسازی و کنترل سیستم نیز باید افزایش یابد.

برای نمونه، استفاده از AI برای پیشنهاد فیلم با استفاده از آن در کنترل تجهیزات پزشکی یکسان نیست. در کاربرد نخست، خطای سیستم معمولاً خسارت محدودی ایجاد می‌کند؛ در کاربرد دوم، خطا می‌تواند پیامد جانی داشته باشد.

بنابراین می‌توان الزامات را بر اساس سطح ریسک طبقه‌بندی کرد:

ریسک پایین: نظارت عمومی و شفافیت مناسب.

ریسک متوسط: آزمون، مستندسازی، پایش و کنترل انسانی بیشتر.

ریسک بالا: ارزیابی دقیق ریسک، ممیزی، قابلیت ردیابی، نظارت انسانی مؤثر و سازوکارهای مشخص جبران خسارت.

این رویکرد با تحولات جدید مقررات‌گذاری هوش مصنوعی در جهان نیز هم‌راستا است.

۱۵. چالش اثبات دعوا برای زیان دیده

یکی از نگرانی‌های اصلی در مسئولیت مدنی AI این است که بار اثبات برای شخص زیان دیده بیش از حد سنگین شود.

ممکن است فرد به سیستم، کد، داده‌های آموزشی یا سوابق تصمیم‌گیری دسترسی نداشته باشد. در مقابل، شرکت توسعه‌دهنده یا بهره‌بردار اطلاعات فنی گسترده‌ای در اختیار دارد.

این عدم تقارن اطلاعاتی می‌تواند عدالت فرآیند رسیدگی را تحت تأثیر قرار دهد.

راهکارهایی مانند:

- الزامات نگهداری سوابق؛
- افشای کنترل‌شده اطلاعات فنی؛
- استانداردهای مستندسازی؛
- استفاده از کارشناسان مستقل؛
- و در برخی نظام‌ها، قواعد ویژه درباره بار اثبات

می‌توانند به کاهش این مشکل کمک کنند.

۱۶. بیمه و انتقال ریسک

بیمه می‌تواند یکی از ابزارهای مکمل برای مدیریت خسارت ناشی از سامانه‌های هوش مصنوعی باشد.

سازمان‌ها ممکن است برای خطرات ناشی از خطای الگوریتمی، نقص نرم‌افزار، نقض امنیت یا مسئولیت حرفه‌ای به پوشش بیمه‌ای نیاز داشته باشند.

اما توسعه بازار بیمه AI با چالش‌هایی مانند دشواری برآورد احتمال خسارت، کمبود داده تاریخی و سرعت تغییر فناوری روبه‌رو است.

در بلندمدت، ترکیب استانداردهای فنی، مدیریت ریسک، قرارداد، بیمه و قواعد مسئولیت مدنی می‌تواند چارچوب جامع‌تری برای جبران خسارت ایجاد کند.

۱۷. چارچوب پیشنهادی برای تعیین مسئولیت

برای تحلیل حقوقی یک حادثه ناشی از تصمیم AI می‌توان یک فرآیند چندمرحله‌ای در نظر گرفت:

مرحله اول: شناسایی تصمیم

مشخص شود چه تصمیمی توسط سامانه اتخاذ شده و چه اثر حقوقی یا اقتصادی ایجاد کرده است.

مرحله دوم: شناسایی بازیگران

تمام اشخاص دخیل در طراحی، توسعه، داده، استقرار، تنظیم و بهره‌برداری شناسایی شوند.

مرحله سوم: بررسی استاندارد مراقبت

بررسی شود آیا هر بازیگر اقدامات متعارف و مورد انتظار را انجام داده است یا خیر.

مرحله چهارم: تحلیل نقص یا خطا

مشخص شود منشأ خطا در داده، مدل، نرم‌افزار، سخت‌افزار، تنظیمات یا فرآیند انسانی بوده است.

مرحله پنجم: اثبات رابطه سببیت

بررسی شود آیا تصمیم AI علت مؤثر و قابل انتساب خسارت بوده است یا عوامل دیگری نیز دخالت داشته‌اند.

مرحله ششم: تعیین سهم مسئولیت

در صورت دخالت چند عامل، سهم هر یک بر اساس نقش، کنترل، تقصیر و قواعد حقوقی تعیین شود.

مرحله هفتم: ارزیابی خسارت

نوع و میزان خسارت مالی، بدنی، معنوی یا سایر زیان‌های قابل جبران بررسی شود.

این چارچوب می‌تواند مبنایی برای تحلیل پرونده‌های پیچیده‌ای باشد که در آن‌ها یک تصمیم الگوریتمی به خسارت منجر شده است.

۱۸. چالش قانون‌گذاری

یکی از مسائل اصلی این است که فناوری AI سریع‌تر از بسیاری از نظام‌های حقوقی تحول پیدا می‌کند. قوانین سنتی مسئولیت مدنی در بسیاری از موارد قابلیت انطباق با فناوری جدید را دارند، اما در پرونده‌های پیچیده ممکن است خلأهایی در زمینه اثبات، شفافیت و توزیع مسئولیت ایجاد شود.

قانون‌گذاری جدید باید میان دو هدف تعادل ایجاد کند:

۱. حمایت مؤثر از حقوق زیان‌دیدگان؛

۲. جلوگیری از تحمیل مسئولیت غیرمنطقی که نوآوری را متوقف کند.

اگر هر خطای پیش‌بینی‌ناپذیر AI به‌طور خودکار موجب مسئولیت مطلق توسعه‌دهنده شود، انگیزه توسعه فناوری‌های جدید کاهش می‌یابد. در مقابل، اگر اثبات مسئولیت برای زیان‌دیده تقریباً غیرممکن باشد، ریسک فناوری بدون جبران مناسب بر دوش اشخاص قرار خواهد گرفت.

بنابراین، سیاست مطلوب احتمالاً ترکیبی از مسئولیت متناسب با ریسک، استانداردهای مراقبت، الزامات مستندسازی و سازوکارهای مناسب جبران خسارت خواهد بود.

نتیجه‌گیری

تصمیمات خودکار هوش مصنوعی یکی از مهم‌ترین چالش‌های حقوق مسئولیت مدنی در عصر دیجیتال است. ویژگی اصلی این چالش آن است که تصمیم زیان‌بار ممکن است مستقیماً از رفتار یک انسان منفرد ناشی نشود، بلکه نتیجه تعامل پیچیده میان داده، الگوریتم، نرم‌افزار، زیرساخت و تصمیمات سازمانی باشد.

در شرایط فعلی، اعطای شخصیت حقوقی مستقل به هوش مصنوعی برای حل مسئله مسئولیت مدنی، راهکار اصلی و ضروری به نظر نمی‌رسد. مسئله مهم‌تر، ایجاد سازوکار مناسب برای انتساب مسئولیت به توسعه‌دهندگان، تولیدکنندگان، ارائه‌دهندگان خدمات و بهره‌برداران بر اساس میزان کنترل، نقش، خطر ایجادشده و استاندارد مراقبت مورد انتظار است.

مسئولیت مبتنی بر تقصیر، مسئولیت ناشی از نقص محصول، قواعد مربوط به فعالیت‌های پرخطر، مسئولیت قراردادی و سازوکارهای بیمه‌ای می‌توانند در کنار یکدیگر چارچوبی چندلایه ایجاد کنند.

از سوی دیگر، افزایش قابلیت حسابرسی، ثبت سوابق تصمیم، کیفیت داده، ارزیابی سوگیری، توضیح‌پذیری، نظارت انسانی و مدیریت ریسک می‌تواند علاوه بر کاهش احتمال خسارت، فرآیند اثبات مسئولیت را نیز ساده‌تر کند.

در نهایت، آینده مسئولیت مدنی هوش مصنوعی احتمالاً نه در پذیرش یک قاعده واحد، بلکه در ایجاد نظامی مبتنی بر ریسک و متناسب با سطح خودمختاری و پیامد تصمیمات AI قرار خواهد داشت. هرچه تصمیم یک سامانه تأثیر بیشتری بر جان، آزادی، حقوق یا دارایی افراد داشته باشد، انتظار حقوقی از شفافیت، کنترل، نظارت و پاسخگویی نیز باید بیشتر شود.

European Union. (٢٠٢٤). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*.

European Commission. (٢٠٢٤). *AI Act: Regulatory framework for artificial intelligence*.

OECD. (٢٠١٩). *Recommendation of the Council on Artificial Intelligence*.

NIST. (٢٠٢٣). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology.

UNESCO. (٢٠٢١). *Recommendation on the Ethics of Artificial Intelligence*.

European Commission. (٢٠٢٢). *Liability for defective products: adapting liability rules to the digital age, AI and the circular economy*.

European Parliament. (٢٠٢٠). *Artificial Intelligence and Civil Liability*.

European Commission. (٢٠٢٢). *Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)*.

Russell, S., & Norvig, P. (٢٠٢١). *Artificial Intelligence: A Modern Approach* (٤th ed.). Pearson.

Goodfellow, I., Bengio, Y., & Courville, A. (٢٠١٦). *Deep Learning*. MIT Press.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (٢٠١٧). Accountable algorithms. *University of Pennsylvania Law Review*, 165, ٦٣٣–٧٠٥.

Wachter, S., Mittelstadt, B., & Floridi, L. (٢٠١٧). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(٢), ٧٦–٩٩.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (٢٠١٦). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(٢).

Bryson, J. J. (٢٠١٨). Patency is not a virtue: The design of intelligent systems and systems of ethics. *Ethics and Information Technology*, 20, ١٥–٢٦.