

تحلیل تعهدات بین المللی دولت‌ها در قبال خسارات ناشی از

سامانه‌های هوش مصنوعی خودمختاری فرامرزی

(تاریخ دریافت ۱۴۰۵/۰۱/۲۵ تاریخ تصویب ۱۴۰۵/۰۳/۰۴)

دکتر فرامرز امجدیان^۱

فرگس ملکی

چکیده

تحولات شتابان در حوزه فناوری‌های نوین، به‌ویژه توسعه سامانه‌های هوش مصنوعی خودمختار، چالش‌های بنیادینی را برای نظام حقوق بین‌الملل در زمینه مسئولیت دولت‌ها ایجاد کرده است. این سامانه‌ها با برخورداری از توانایی تصمیم‌گیری مستقل، یادگیری تطبیقی و عملکرد فرامرزی، موجب پیچیده‌تر شدن زنجیره انتساب رفتار و دشواری در تعیین مسئولیت دولت‌ها در قبال خسارات ناشی از آن‌ها شده‌اند. پژوهش حاضر با هدف تحلیل تعهدات بین‌المللی دولت‌ها در قبال خسارات ناشی از هوش مصنوعی خودمختار فرامرزی، به بررسی مبانی نظری مسئولیت بین‌المللی، اصول انتساب، تعهدات پیشگیرانه (Due Diligence) و تعهد جبران خسارت پرداخته است. روش تحقیق در این مطالعه، توصیفی-تحلیلی و مبتنی بر منابع کتابخانه‌ای است و تلاش شده است با بهره‌گیری از اسناد بین‌المللی، دکترین حقوقی و رویه‌های موجود، چارچوبی منسجم برای تحلیل مسئولیت دولت‌ها ارائه شود. یافته‌های پژوهش نشان می‌دهد که قواعد سنتی مسئولیت بین‌المللی، به‌ویژه در حوزه انتساب رفتار و احراز رابطه سببیت، برای پاسخ‌گویی به خسارات ناشی از سامانه‌های خودمختار کافی نبوده و نیازمند تفسیر توسعه‌یافته و بازنگری هنجاری است. همچنین نتایج بیانگر آن است که تعهدات

^۱ نویسنده مسئول

پیشگیرانه دولت‌ها در این حوزه اهمیت فزاینده‌ای یافته و مسئولیت آن‌ها صرفاً محدود به مرحله پس از وقوع خسارت نیست، بلکه شامل نظارت مستمر، کنترل ریسک و تنظیم‌گری فعال در چرخه عمر سامانه‌های هوش مصنوعی نیز می‌شود. در نهایت، پژوهش بر ضرورت توسعه قواعد بین‌المللی جدید، تقویت همکاری میان دولت‌ها و حرکت به سمت نظام مسئولیت مبتنی بر ریسک تأکید می‌نماید.

واژگان کلیدی: مسئولیت بین‌المللی دولت‌ها، هوش مصنوعی خودمختار، تعهد مراقبت مقتضی، انتساب، خسارات فرامرزی، حقوق بین‌الملل.

مقدمه

تحولات شتابان در حوزه فناوری‌های نوین، به‌ویژه توسعه سامانه‌های هوش مصنوعی خودمختار، یکی از بنیادین‌ترین دگرگونی‌ها را در ساختار حقوق بین‌الملل معاصر ایجاد کرده است. این سامانه‌ها که توانایی تصمیم‌گیری، یادگیری و اقدام مستقل از مداخله مستقیم انسان را دارند، به تدریج از سطح ابزارهای صرفاً کمکی فراتر رفته و به بازیگرانی تبدیل شده‌اند که آثار حقوقی و عملی قابل توجهی در عرصه‌های مختلف، از امنیت سایبری و اقتصاد دیجیتال تا حمل‌ونقل، سلامت و حتی حوزه نظامی برجای می‌گذارند. در این میان، ویژگی فرامرزی بودن عملکرد این سامانه‌ها موجب شده است که آثار ناشی از فعالیت آن‌ها محدود به قلمرو یک دولت خاص نباشد و در بسیاری از موارد، چندین نظام حقوقی به‌طور هم‌زمان درگیر پیامدهای ناشی از آن شوند. در چنین بستری، مسئله مسئولیت بین‌المللی دولت‌ها در قبال خسارات ناشی از عملکرد سامانه‌های هوش مصنوعی خودمختار به یکی از چالش‌های نوظهور و پیچیده حقوق بین‌الملل تبدیل شده است. نظام سنتی مسئولیت بین‌المللی دولت‌ها، همان‌گونه که در پیش‌نویس مواد کمیسیون حقوق بین‌الملل سازمان ملل متحد (ILC Articles on State Responsibility) منعکس شده است، مبتنی بر دو عنصر اساسی یعنی انتساب رفتار و نقض

تعهد بین‌المللی است. با این حال، در مواجهه با هوش مصنوعی خودمختار، پرسش‌های بنیادینی مطرح می‌شود؛ از جمله اینکه آیا می‌توان رفتار یک سامانه خودمختار را به دولت منتسب نمود، تا چه میزان کنترل انسانی بر این سامانه‌ها برای تحقق مسئولیت کافی است، و در مواردی که تصمیم‌گیری به صورت غیرقابل پیش‌بینی توسط الگوریتم‌های پیچیده انجام می‌شود، مرز میان خطای فنی و تخلف حقوقی در کجا قرار می‌گیرد. افزون بر این، ماهیت فرامرزی خسارات ناشی از هوش مصنوعی، چالش‌های مضاعفی در حوزه حاکمیت دولت‌ها و اعمال صلاحیت قضایی ایجاد کرده است. در بسیاری از موارد، تولیدکننده الگوریتم، ارائه‌دهنده خدمات، کاربر نهایی و محل وقوع خسارت در حوزه‌های قضایی مختلف قرار دارند؛ امری که تعیین دولت مسئول یا مسئولیت مشترک دولت‌ها را با پیچیدگی‌های جدی مواجه می‌سازد. این وضعیت، ضرورت بازاندیشی در مفاهیم سنتی مانند «کنترل مؤثر»، «مراقبت مقتضی» و «انتساب مستقیم یا غیرمستقیم رفتار» را برجسته می‌سازد. از منظر نظری، این پژوهش در پی آن است که نشان دهد نظام موجود مسئولیت بین‌المللی دولت‌ها، اگرچه ظرفیت‌هایی برای انطباق با فناوری‌های نوین دارد، اما در شکل فعلی خود پاسخ‌گوی کامل چالش‌های ناشی از هوش مصنوعی خودمختار فرامرزی نیست. بنابراین، تحلیل تعهدات بین‌المللی دولت‌ها در این حوزه نه تنها از منظر توصیفی، بلکه از حیث هنجارسازی و توسعه حقوق بین‌الملل نیز اهمیت بنیادین دارد. بر این اساس، مسئله اصلی این پژوهش بر این محور استوار است که دولت‌ها در قبال خسارات ناشی از سامانه‌های هوش مصنوعی خودمختار فرامرزی چه تعهداتی دارند و نظام حقوق بین‌الملل تا چه میزان قادر است این تعهدات را به صورت مؤثر، منسجم و قابل اجرا تنظیم و تضمین کند. این پژوهش با رویکردی توصیفی-تحلیلی و با اتکاء بر منابع کتابخانه‌ای، اسناد بین‌المللی و دکترین حقوقی، درصدد تبیین ابعاد مختلف این مسئولیت و ارائه تصویری روشن از چالش‌ها و الزامات آینده آن است.

بخش اول: کلیات

گسترش روزافزون سامانه‌های هوش مصنوعی خودمختار در دهه‌های اخیر، ساختار سنتی حقوق بین‌الملل را با چالش‌های بنیادین مواجه ساخته است. این سامانه‌ها که قادر به تحلیل داده‌های پیچیده، یادگیری مستقل و اتخاذ تصمیمات بدون دخالت مستقیم انسان هستند، به‌طور فزاینده‌ای در حوزه‌های حساس و حیاتی مورد استفاده قرار می‌گیرند. در نتیجه، آثار و پیامدهای ناشی از عملکرد آن‌ها، به‌ویژه در مواردی که منجر به خسارات مادی، نقض حقوق افراد یا آسیب به زیرساخت‌های حیاتی می‌شود، دیگر محدود به قلمرو یک دولت خاص نیست و ماهیتی فرامرزی پیدا کرده است. در این چارچوب، مسئله اساسی آن است که نظام حقوق بین‌الملل موجود، به‌ویژه قواعد حاکم بر مسئولیت بین‌المللی دولت‌ها، عمدتاً بر مبنای رفتار انسانی و قابلیت انتساب مستقیم اعمال به دولت‌ها طراحی شده است. در حالی که در مورد سامانه‌های هوش مصنوعی خودمختار، زنجیره تصمیم‌گیری و اقدام به گونه‌ای میان انسان، نرم‌افزار، داده و الگوریتم توزیع شده است که تشخیص عامل مستقیم و انتساب حقوقی رفتار به یک دولت خاص با دشواری‌های جدی مواجه می‌شود. این وضعیت، خلأ مهمی در شناسایی مسئولیت و جبران خسارت در سطح بین‌المللی ایجاد کرده است. در بسیاری از موارد خسارات ناشی از این سامانه‌ها در قلمرو بیش از یک دولت رخ می‌دهد یا منشأ آن در یک کشور و آثار آن در کشور دیگر ظاهر می‌شود. این امر نه تنها مسئله انتساب را پیچیده‌تر می‌سازد، بلکه تعارض میان نظام‌های حقوقی ملی، اختلاف در استانداردهای نظارتی و نبود یک چارچوب الزام‌آور بین‌المللی را نیز برجسته می‌کند. در نتیجه، این پرسش اساسی مطرح می‌شود که در فقدان یک رژیم حقوقی جامع و هماهنگ، چگونه می‌توان مسئولیت بین‌المللی دولت‌ها را در قبال خسارات ناشی از هوش مصنوعی خودمختار فرامرزی به صورت منسجم و قابل اجرا تعریف و اعمال کرد. مسئله اصلی این پژوهش ناظر بر این امر است که با وجود قواعد سنتی مسئولیت بین‌المللی دولت‌ها، تا چه اندازه می‌توان خسارات ناشی از سامانه‌های هوش مصنوعی خودمختار فرامرزی را به دولت‌ها منتسب نمود و چه نوع تعهداتی (پیشگیرانه، نظارتی و

جبرانی) باید در حقوق بین الملل برای پاسخ گویی مؤثر به این پدیده نوظهور شناسایی و توسعه یابد. این مسئله در نقطه تلاقی میان پیشرفت فناوری و نارسایی قواعد سنتی حقوق بین الملل قرار دارد و حل آن مستلزم بازاندیشی در مفاهیم بنیادین مسئولیت دولت ها است.

• سؤالات مطروحه پژوهش

۱. سامانه های هوش مصنوعی خودمختار تا چه اندازه می توانند در چارچوب قواعد موجود حقوق بین الملل موضوع انتساب رفتار به دولت ها قرار گیرند؟
۲. مهم ترین تعهدات بین المللی دولت ها در قبال پیشگیری از خسارات ناشی از سامانه های هوش مصنوعی خودمختار فرامرزی چیست؟
۳. نظام حقوق بین الملل موجود تا چه میزان ظرفیت جبران خسارات ناشی از عملکرد هوش مصنوعی خودمختار را داراست؟
۴. چه اصلاحات یا سازوکارهای حقوقی برای تنظیم مؤثر مسئولیت بین المللی دولت ها در قبال خسارات ناشی از هوش مصنوعی ضروری است؟

بخش دوم: مبانی مفهومی و چارچوب نظری هوش مصنوعی خودمختار

بند اول: مفهوم شناسی هوش مصنوعی خودمختار و ویژگی های سامانه های فرامرزی

هوش مصنوعی خودمختار به دسته ای از سامانه های هوشمند اطلاق می شود که توانایی انجام وظایف پیچیده، تصمیم گیری و اجرای اقدامات را بدون مداخله مستقیم و مستمر انسان دارا هستند. این سامانه ها بر پایه الگوریتم های یادگیری ماشین، شبکه های عصبی و تحلیل کلان داده ها عمل می کنند و قادرند با دریافت داده های محیطی، رفتار خود را به صورت پویا و تطبیقی تنظیم نمایند. وجه تمایز اصلی این نوع از هوش مصنوعی با سیستم های سنتی، سطح

بالای استقلال عملیاتی و کاهش نقش نظارت لحظه‌ای انسانی در فرآیند تصمیم‌سازی است. از منظر حقوقی، ویژگی مهم این سامانه‌ها «خودمختاری عملیاتی» و «عدم پیش‌بینی‌پذیری کامل نتایج» است. این امر موجب می‌شود که رفتار آن‌ها در برخی موارد از کنترل مستقیم طراح، تولیدکننده یا کاربر خارج شود و نتایجی ایجاد کند که از نظر حقوقی به سختی قابل انتساب به یک فاعل مشخص انسانی باشد. این ویژگی، چالش اساسی برای نظام سنتی مسئولیت بین‌المللی دولت‌ها ایجاد می‌کند که بر فرض وجود اراده انسانی و کنترل قابل انتساب استوار است. ماهیت فرامرزی این سامانه‌ها باعث می‌شود که فرآیند طراحی، آموزش داده، استقرار و استفاده از آن‌ها در حوزه‌های قضایی مختلف انجام گیرد. به عنوان مثال، ممکن است داده‌ها در یک کشور جمع‌آوری، الگوریتم در کشور دیگر طراحی و سامانه در کشور ثالث اجرا شود. این پراکندگی جغرافیایی، مفهوم سنتی «محل وقوع فعل زیان‌بار» را در حقوق بین‌الملل با ابهام مواجه کرده و ضرورت بازتعریف مفاهیم کلاسیک مسئولیت را برجسته می‌سازد.

بند دوم: تمایز میان مسئولیت انسانی، الگوریتمی و خودمختاری سیستم‌های هوشمند

در تحلیل حقوقی هوش مصنوعی خودمختار، تفکیک میان مسئولیت انسانی و عملکرد الگوریتمی اهمیت بنیادین دارد. مسئولیت انسانی ناظر بر اشخاص حقیقی یا حقوقی است که در طراحی، آموزش، استقرار یا بهره‌برداری از سامانه نقش دارند و می‌توان رفتار آنان را تحت قواعد سنتی حقوق مسئولیت مورد ارزیابی قرار داد. در مقابل، عملکرد الگوریتمی به تصمیمات و خروجی‌هایی اشاره دارد که توسط خود سیستم و بر اساس داده‌ها و مدل‌های یادگیری تولید می‌شود. مفهوم «خودمختاری سیستم‌های هوشمند» در این میان به درجه‌ای از استقلال اشاره دارد که در آن، رفتار سیستم صرفاً نتیجه دستورات مستقیم انسانی نیست، بلکه حاصل فرآیندهای یادگیری و تطبیق داخلی سامانه است. این ویژگی سبب می‌شود که زنجیره علیت حقوقی میان انسان و نتیجه نهایی، دچار گسست یا پیچیدگی شود. در چنین شرایطی، تعیین این که آیا خطا یا خسارت ناشی از طراحی ناقص، داده‌های معیوب یا تصمیم مستقل

سیستم بوده است، به یک مسئله پیچیده حقوقی تبدیل می شود. از منظر نظری، این تمایز زمینه ساز شکل گیری بحث هایی درباره امکان یا عدم امکان شناسایی «مسئولیت شبه مستقل» یا «مسئولیت توزیع شده» در حقوق بین الملل شده است. در این رویکردها، تلاش می شود به جای تمرکز صرف بر فاعل انسانی، کل زنجیره طراحی، توسعه و بهره برداری به عنوان یک ساختار مسئولیت پذیر مورد توجه قرار گیرد. با این حال، نظام حقوق بین الملل کلاسیک هنوز به طور کامل این تحول مفهومی را نپذیرفته و همین امر، خلأ نظری مهمی در تحلیل مسئولیت ناشی از هوش مصنوعی ایجاد کرده است.

بخش سوم: مفهوم شناسی هوش مصنوعی خودمختار و ویژگی های سامانه های فرامرزی

هوش مصنوعی خودمختار به مجموعه ای از سیستم ها و برنامه های مبتنی بر الگوریتم های پیشرفته اطلاق می شود که قادرند بدون نیاز به مداخله مستقیم و مستمر انسان، وظایف پیچیده را تحلیل، تصمیم گیری و اجرا کنند. این نوع از هوش مصنوعی بر پایه فناوری هایی نظیر یادگیری ماشین، یادگیری عمیق و تحلیل کلان داده ها عمل می کند و می تواند با دریافت داده های جدید از محیط، رفتار خود را به صورت پویا و تطبیقی تغییر دهد. ویژگی بنیادین این سامانه ها، «توانایی یادگیری از تجربه» و «تولید تصمیمات مستقل از دستور صریح انسانی» است؛ امری که آن ها را از ابزارهای سنتی فناوری اطلاعات متمایز می سازد. مهم ترین خصیصه هوش مصنوعی خودمختار، درجه بالای استقلال عملکردی آن است. این استقلال به معنای آن است که خروجی سیستم همواره قابل پیش بینی کامل توسط طراح یا کاربر نیست و ممکن است در شرایط مشابه، نتایج متفاوتی تولید کند. همین ویژگی، چالش جدی برای نظام های حقوقی ایجاد می کند که بر اصل قابلیت پیش بینی رفتار فاعل انسانی استوار هستند. در نتیجه، رابطه مستقیم میان اراده انسانی و نتیجه نهایی در بسیاری از موارد دچار گسست یا حداقل تضعیف می شود. ویژگی مهم دیگر این سامانه ها، ماهیت فرامرزی و شبکه ای آن ها است. فرآیند توسعه، آموزش، استقرار و بهره برداری از هوش مصنوعی معمولاً در یک قلمرو جغرافیایی واحد انجام

نمی‌شود. داده‌ها ممکن است در یک کشور جمع‌آوری شوند، الگوریتم در کشور دیگر توسعه یابد و سیستم در سطح جهانی مورد استفاده قرار گیرد. این پراکندگی جغرافیایی موجب می‌شود که آثار حقوقی ناشی از عملکرد این سامانه‌ها نیز فرامرزی بوده و در حوزه صلاحیت چندین دولت قرار گیرد. در نتیجه، مفاهیم سنتی حقوق بین‌الملل مانند «قلمرو»، «صلاحیت سرزمینی» و «محل وقوع فعل زیان‌بار» با چالش‌های جدی مواجه می‌شوند.

بند اول: تمایز میان مسئولیت انسانی، الگوریتمی و خودمختاری سیستم‌های هوشمند

در تحلیل حقوقی هوش مصنوعی، تفکیک میان سه سطح مسئولیت یعنی مسئولیت انسانی، مسئولیت الگوریتمی و سطح خودمختاری سیستم از اهمیت بنیادین برخوردار است. مسئولیت انسانی ناظر بر اشخاص حقیقی یا حقوقی است که در طراحی، توسعه، آموزش داده‌ها، استقرار یا بهره‌برداری از سامانه نقش دارند. در این سطح، رفتار قابل انتساب به انسان‌هاست و قواعد سنتی حقوق بین‌الملل، به‌ویژه اصول مسئولیت دولت‌ها، قابلیت اعمال دارند. مسئولیت الگوریتمی به نقش ساختارهای فنی و نرم‌افزاری در ایجاد نتیجه زیان‌بار اشاره دارد. در این سطح، خطا ممکن است ناشی از داده‌های معیوب، طراحی ناقص مدل، سوگیری الگوریتمی یا عدم پیش‌بینی رفتار سیستم در شرایط خاص باشد. با این حال، از منظر حقوقی، الگوریتم به‌عنوان یک موجود مستقل فاقد شخصیت حقوقی است و نمی‌توان آن را به‌طور مستقیم مسئول دانست. بنابراین، مسئولیت الگوریتمی در نهایت به اشخاص انسانی یا نهادهای حقوقی بازمی‌گردد که کنترل یا نظارت بر آن را بر عهده داشته‌اند. سطح سوم، یعنی خودمختاری سیستم‌های هوشمند، به حالتی اشاره دارد که در آن سیستم پس از استقرار، قادر است بدون دخالت انسانی تصمیم‌گیری کرده و حتی رفتارهایی بروز دهد که پیش‌بینی آن‌ها در زمان طراحی ممکن نبوده است. این ویژگی موجب ایجاد «شکاف مسئولیت» می‌شود؛ زیرا رابطه علی میان رفتار انسانی اولیه و نتیجه نهایی تضعیف می‌گردد. در چنین وضعیتی، نظام سنتی مسئولیت که مبتنی بر انتساب مستقیم و کنترل مؤثر است، با محدودیت‌های جدی مواجه

می‌شود. تمایز میان این سه سطح نشان می‌دهد که حقوق بین‌الملل با یک چالش ساختاری مواجه است؛ زیرا هنوز چارچوبی منسجم برای توزیع مسئولیت در میان انسان، فناوری و سیستم‌های خودمختار ارائه نکرده است. این وضعیت ضرورت بازتعریف مفاهیم بنیادین مسئولیت، انتساب و کنترل در عصر هوش مصنوعی را آشکار می‌سازد.

بخش چهارم: مبانی حقوقی تعهدات بین‌المللی دولت‌ها

تعهدات بین‌المللی دولت‌ها در نظام حقوق بین‌الملل معاصر، هسته اصلی تنظیم روابط میان کشورها و تضمین نظم حقوقی جهانی را تشکیل می‌دهد. این تعهدات مجموعه‌ای از الزامات الزام‌آور هستند که دولت‌ها بر اساس منابع معتبر حقوق بین‌الملل، از جمله معاهدات، عرف بین‌المللی، اصول کلی حقوق، رویه قضایی و دکترین حقوقی ملزم به رعایت آن‌ها می‌باشند. ماهیت این تعهدات به گونه‌ای است که نقض آن‌ها موجب تحقق مسئولیت بین‌المللی دولت و ایجاد آثار حقوقی همچون جبران خسارت، اعاده وضع سابق، توقف رفتار متخلفانه و تضمین عدم تکرار می‌شود. از منظر ساختاری، تعهدات بین‌المللی دولت‌ها بر اصل حاکمیت قانون در عرصه بین‌المللی استوار است؛ به این معنا که حاکمیت دولت‌ها مطلق و بی‌قید و شرط نیست، بلکه در چارچوب قواعد پذیرفته‌شده حقوق بین‌الملل محدود می‌شود. این محدودیت نه تنها از طریق معاهدات الزام‌آور، بلکه از طریق شکل‌گیری تدریجی عرف‌های بین‌المللی و رویه مستمر دولت‌ها نیز تحقق می‌یابد. در نتیجه، دولت‌ها در کنار حقوق خود، دارای تکالیف مشخصی در قبال جامعه بین‌المللی هستند که رعایت آن‌ها شرط حفظ نظم و ثبات بین‌المللی محسوب می‌شود. منابع تعهدات بین‌المللی دولت‌ها در چارچوب ماده ۳۸ اساسنامه دیوان بین‌المللی دادگستری شامل معاهدات، عرف بین‌المللی و اصول کلی حقوق است که هر یک نقش مکمل در شکل‌دهی به نظام تعهدات دارند. در کنار این منابع، رویه قضایی و دکترین حقوقی نیز در تفسیر و توسعه این تعهدات نقش مهمی ایفا می‌کنند. در حوزه‌های نوظهور مانند فناوری‌های دیجیتال و هوش مصنوعی، اگرچه معاهدات جامع و یکپارچه هنوز محدود هستند، اما اصولی همچون منع آسیب‌فرامرزی، اصل همکاری بین‌المللی و تعهد مراقبت مقتضی به

تدریج در حال تثبیت به عنوان قواعد عرفی و هنجارهای الزام آور هستند. این روند نشان می‌دهد که تعهدات بین‌المللی دولت‌ها همواره در حال تحول و انطباق با واقعیت‌های جدید جهانی هستند.

بند اول: اصول حاکم بر مسئولیت بین‌المللی دولت‌ها در حقوق بین‌الملل

مسئولیت بین‌المللی دولت‌ها یکی از بنیان‌های اساسی نظام حقوق بین‌الملل معاصر است که بر این اصل استوار است که هر فعل متخلفانه بین‌المللی دولت، موجب مسئولیت آن دولت در برابر جامعه بین‌المللی یا دولت متضرر می‌گردد. این نظام عمدتاً بر مبنای پیش‌نویس مواد کمیسیون حقوق بین‌الملل (ARSIWA) شکل گرفته و شامل قواعدی است که شرایط تحقق مسئولیت، آثار آن و نحوه جبران خسارت را تبیین می‌کند. در این چارچوب، مسئولیت بین‌المللی نه یک امر اختیاری، بلکه یک پیامد حقوقی اجتناب‌ناپذیر ناشی از نقض تعهدات بین‌المللی محسوب می‌شود. یکی از اصول بنیادین حاکم بر مسئولیت بین‌المللی دولت‌ها، اصل انتساب رفتار (Attribution) است. بر اساس این اصل، تنها رفتارهایی می‌توانند موجب مسئولیت دولت شوند که به دولت قابل انتساب باشند؛ اعم از اینکه توسط ارکان رسمی دولت، اشخاص مأمور، یا در برخی شرایط خاص توسط اشخاص خصوصی تحت کنترل مؤثر دولت انجام شده باشند. این اصل نقش تعیین‌کننده‌ای در تفکیک میان اعمال دولت و رفتار اشخاص ثالث دارد و مانع از توسعه بی‌ضابطه مسئولیت دولت‌ها می‌شود. در واقع، بدون احراز انتساب، حتی در صورت وقوع خسارت، مسئولیت بین‌المللی دولت قابل تحقق نخواهد بود. اصل دوم، اصل نقض تعهد بین‌المللی (Breach of an International Obligation) است که بیان می‌دارد مسئولیت زمانی محقق می‌شود که دولت از انجام تعهدی که بر عهده دارد، اعم از تعهدات معاهده‌ای یا عرفی، تخلف کند. این نقض می‌تواند به صورت فعل یا ترک فعل باشد و نیازی به اثبات سوءنیت دولت ندارد، بلکه صرف عدم انطباق رفتار با تعهد بین‌المللی کافی است. این ویژگی، مسئولیت بین‌المللی را به یک مسئولیت عمدتاً عینی و مبتنی بر نتیجه تبدیل می‌کند.

اصل سوم، اصل قابلیت انتساب خسارت و رابطه سببیت (Causation) است که بر اساس آن باید میان رفتار متخلفانه دولت و خسارت وارده رابطه مستقیم یا قابل قبول حقوقی وجود داشته باشد. این اصل از گسترش نامحدود مسئولیت جلوگیری کرده و تضمین می کند که تنها خساراتی قابل جبران هستند که به طور منطقی و حقوقی ناشی از فعل متخلفانه باشند. اصل جبران کامل خسارت (Full Reparation) از اصول بنیادین مسئولیت بین المللی دولت ها محسوب می شود که بر اساس آن دولت مسئول موظف است وضعیت را تا حد امکان به حالت قبل از وقوع فعل متخلفانه بازگرداند. جبران خسارت می تواند به صورت اعاده وضع سابق، پرداخت غرامت یا رضایت (Satisfaction) باشد. این اصل در کنار سایر اصول، چارچوبی منسجم برای تضمین کارآمدی نظام مسئولیت بین المللی و حفظ نظم حقوقی بین المللی ایجاد می کند.

بند دوم: اصل انتساب رفتار سامانه های هوش مصنوعی به دولت ها

اصل انتساب (Attribution) در حقوق بین الملل یکی از ارکان بنیادین مسئولیت بین المللی دولت ها محسوب می شود و تعیین می کند که چه رفتارهایی می توانند به دولت نسبت داده شوند. در چارچوب سنتی، این اصل بر رفتار ارکان دولتی، اشخاص مأمور دولت، یا اشخاص خصوصی که تحت کنترل مؤثر دولت عمل می کنند، مبتنی است. با این حال، ظهور سامانه های هوش مصنوعی خودمختار، به ویژه در سطح فرامرزی، این چارچوب کلاسیک را با چالش های جدی مواجه ساخته است؛ زیرا رفتار این سامانه ها نه به طور مستقیم انسانی است و نه همواره قابل پیش بینی یا کنترل کامل توسط دولت ها. در تحلیل نوین، انتساب رفتار سامانه های هوش مصنوعی به دولت ها عمدتاً از طریق زنجیره انسانی دخالت و کنترل قابل تصور است. به این معنا که اگر دولت در طراحی، توسعه، آموزش داده ها، استقرار یا بهره برداری از سامانه نقش اساسی داشته باشد، رفتار نهایی سیستم می تواند به صورت غیرمستقیم به دولت منتسب شود. این رویکرد با تکیه بر معیارهایی همچون «کنترل مؤثر»، «جهت دهی عملیاتی» و «آگاهی از ریسک های قابل پیش بینی» تلاش می کند خلأ ناشی از فقدان فاعلیت انسانی مستقیم در

سامانه‌های خودمختار را جبران نماید. یکی از چالش‌های اساسی در این زمینه، استقلال تصمیم‌گیری سامانه‌های هوش مصنوعی پس از استقرار است. در بسیاری از موارد، سیستم‌های یادگیرنده ممکن است بر اساس داده‌های جدید، رفتارهایی تولید کنند که پیش‌بینی آن در زمان طراحی امکان‌پذیر نبوده است. در چنین شرایطی، تردید جدی در خصوص امکان انتساب مستقیم رفتار به دولت ایجاد می‌شود؛ زیرا عنصر کنترل مؤثر و پیش‌بینی‌پذیری که از شروط سنتی انتساب است، تضعیف می‌گردد. برخی رویکردهای نوین در دکتترین حقوق بین‌الملل تلاش کرده‌اند مفهوم «انتساب توسعه‌یافته» یا «انتساب ساختاری» را مطرح کنند. در این دیدگاه، به جای تمرکز صرف بر کنترل لحظه‌ای، کل چرخه عمر سامانه هوش مصنوعی — از طراحی تا بهره‌برداری — به عنوان یک زنجیره واحد مورد توجه قرار می‌گیرد و مسئولیت دولت بر اساس میزان نقش و میزان ریسک‌پذیری آن تعیین می‌شود. با این وجود، این رویکرد هنوز در حقوق بین‌الملل به صورت قطعی پذیرفته نشده و در مرحله شکل‌گیری هنجاری قرار دارد.

بخش پنجم: انواع تعهدات دولت‌ها در قبال خسارات ناشی از هوش مصنوعی

تعهدات دولت‌ها در قبال خسارات ناشی از هوش مصنوعی در حقوق بین‌الملل معاصر را می‌توان در قالب یک نظام چندلایه از تعهدات پیشگیرانه، مراقبتی، حمایتی و جبرانی تحلیل کرد. این تعهدات، برخلاف ساختار سنتی مسئولیت بین‌المللی که عمدتاً پسینی و مبتنی بر وقوع خسارت است، ماهیتی ترکیبی داشته و هم پیش از وقوع خسارت و هم پس از تحقق آن، دولت‌ها را ملزم به اتخاذ رفتارهای مشخص می‌نمایند. این تحول نشان‌دهنده گذار حقوق بین‌الملل از الگوی صرفاً واکنشی به الگوی پیش‌دستانه در مواجهه با فناوری‌های نوین است. نخستین دسته، تعهدات پیشگیرانه (Preventive Obligations) است که دولت‌ها را مکلف می‌سازد پیش از وقوع هرگونه خسارت احتمالی ناشی از سامانه‌های هوش مصنوعی، اقدامات لازم برای کاهش ریسک را اتخاذ نمایند. این تعهدات شامل وضع مقررات داخلی مناسب، ایجاد نظام‌های نظارتی مؤثر، ارزیابی اثرات حقوقی و امنیتی سامانه‌های خودمختار و

کنترل فعالیت نهادهای توسعه دهنده و بهره بردار است. در واقع، اصل «مراقبت مقتضی (Due Diligence)» در این حوزه نقش محوری دارد و دولت ها را ملزم می کند تا از قلمرو خود برای جلوگیری از ورود ضرر به سایر دولت ها استفاده کنند، حتی اگر رفتار زیان بار مستقیماً توسط دولت انجام نشود. دومین دسته، تعهدات نظارتی و کنترلی است که ناظر بر استمرار نظارت دولت بر چرخه حیات سامانه های هوش مصنوعی می باشد. این تعهدات ایجاب می کند دولت ها صرفاً به مرحله طراحی یا صدور مجوز اکتفا نکنند، بلکه در طول بهره برداری نیز سازوکارهای پایش، ارزیابی و اصلاح را برقرار نمایند. در سامانه های خودمختار، به دلیل امکان یادگیری و تغییر رفتار سیستم، این نوع نظارت اهمیت مضاعف دارد؛ زیرا خطرات ممکن است پس از استقرار سامانه و در نتیجه تعامل با محیط های جدید پدیدار شوند. از این رو، تعهد نظارتی دولت ها به عنوان یک تعهد مستمر و پویا تلقی می شود. سومین دسته، تعهدات حمایتی و جبرانی (Protective and Reparation Obligations) است که پس از وقوع خسارت فعال می شود. در این مرحله، دولت مسئول مکلف است اقداماتی برای توقف وضعیت متخلفانه، جلوگیری از تداوم خسارت و جبران آثار زیان بار انجام دهد. این جبران می تواند شامل اعاده وضع سابق، پرداخت غرامت مالی یا ارائه اشکال غیرمالی جبران مانند عذرخواهی رسمی یا تضمین عدم تکرار باشد. در حوزه هوش مصنوعی، چالش اصلی در این مرحله، تعیین دقیق منشأ خسارت و تفکیک میان نقش دولت، توسعه دهندگان خصوصی و خود سامانه است. تعهدات همکاری بین المللی نیز بخش مهمی از نظام تعهدات دولت ها را تشکیل می دهد. با توجه به ماهیت فرامرزی فناوری هوش مصنوعی، هیچ دولتی به تنهایی قادر به کنترل کامل پیامدهای آن نیست. بنابراین، دولت ها موظف اند در چارچوب همکاری های دو جانبه، منطقه ای و جهانی، اطلاعات، استانداردها و سازوکارهای نظارتی را به اشتراک بگذارند. این تعهد همکاری، به ویژه در زمینه مدیریت بحران های ناشی از سامانه های خودمختار، نقش کلیدی در کاهش آثار مخرب و ارتقای کارآمدی نظام حقوق بین الملل ایفا می کند.

بند اول: تعهد پیشگیری (Due Diligence) و کنترل خطرات فناوری‌های هوشمند

تعهد پیشگیری یا «مراقبت مقتضی» (Due Diligence) «یکی از اصول بنیادین حقوق بین‌الملل در حوزه مسئولیت دولت‌ها است که بر اساس آن، دولت‌ها موظف‌اند تمامی اقدامات معقول و متناسب را برای جلوگیری از ورود خسارت به سایر دولت‌ها یا حقوق بین‌المللی انجام دهند. در زمینه فناوری‌های هوشمند و به‌ویژه سامانه‌های هوش مصنوعی خودمختار، این تعهد ماهیتی پیش‌دستانه و مستمر پیدا می‌کند و صرفاً به مرحله پس از وقوع خسارت محدود نمی‌شود، بلکه از مرحله طراحی، توسعه و استقرار این فناوری‌ها آغاز می‌گردد.

در چارچوب فناوری‌های هوشمند، تعهد پیشگیری دولت‌ها شامل ایجاد نظام‌های حقوقی و نهادی مناسب برای تنظیم فعالیت توسعه‌دهندگان، اعمال استانداردهای ایمنی و امنیتی، ارزیابی ریسک‌های احتمالی و نظارت بر استفاده از سامانه‌های پرخطر است. دولت‌ها مکلف‌اند اطمینان حاصل کنند که فعالیت اشخاص حقیقی و حقوقی تحت صلاحیت آن‌ها، منجر به ورود خسارت فرامرزی از طریق سامانه‌های هوش مصنوعی نمی‌شود. این امر به‌ویژه در مورد سامانه‌هایی که قابلیت یادگیری مستقل و تصمیم‌گیری خودکار دارند، اهمیت مضاعف دارد. از منظر تحلیلی، استاندارد «مراقبت مقتضی» در حوزه هوش مصنوعی یک استاندارد پویا و متناسب با سطح خطر است. به این معنا که هرچه توانایی سامانه در ایجاد خسارت گسترده‌تر و پیش‌بینی ناپذیرتر باشد، سطح مراقبت و نظارت مورد انتظار از دولت نیز افزایش می‌یابد. بنابراین، دولت‌ها نمی‌توانند صرفاً به وضع مقررات کلی اکتفا کنند، بلکه باید سازوکارهای فنی، نظارتی و حتی همکاری بین‌المللی برای مدیریت ریسک‌های ناشی از این فناوری‌ها را به کار گیرند.

بند دوم: تعهد جبران خسارت و پاسخ گویی در قبال زیان های فرامرزی

تعهد جبران خسارت یکی از آثار کلاسیک مسئولیت بین المللی دولت ها است که در صورت وقوع فعل متخلفانه، دولت مسئول را مکلف به جبران کامل خسارات وارده می سازد. این تعهد در حقوق بین الملل معاصر بر اصل «اعاده وضع سابق» (Restitution) «پرداخت غرامت (Compensation) و ارائه رضایت (Satisfaction) استوار است. در زمینه خسارات ناشی از فناوری های هوش مصنوعی، اعمال این تعهد با چالش های پیچیده ای مواجه است، به ویژه در مواردی که خسارت ناشی از سامانه های خودمختار فرامرزی باشد. در خسارات فرامرزی ناشی از هوش مصنوعی، نخستین چالش، تعیین دقیق رابطه سببیت میان رفتار دولت یا اشخاص تحت صلاحیت آن و خسارت وارده است. از آنجا که این سامانه ها ممکن است در چندین کشور طراحی، آموزش و اجرا شوند، تفکیک نقش هر دولت در ایجاد خسارت دشوار می شود. با این حال، در صورت احراز انتساب یا قصور در انجام تعهد پیشگیری، دولت مسئول مکلف به جبران کامل خسارت خواهد بود، حتی اگر خسارت به صورت غیرمستقیم یا غیرقابل پیش بینی کامل رخ داده باشد. تعهد پاسخ گویی دولت ها صرفاً به جبران مالی محدود نمی شود، بلکه شامل اتخاذ تدابیر اصلاحی و تضمینی نیز هست. این اقدامات می تواند شامل توقف فعالیت سامانه های پرخطر، اصلاح چارچوب های نظارتی داخلی، ارائه تضمین های عدم تکرار و همکاری با دولت های متضرر برای کاهش آثار خسارت باشد. در نتیجه، نظام جبران خسارت در حوزه هوش مصنوعی ماهیتی چندبعدی دارد که ترکیبی از ابزارهای مالی، حقوقی و نهادی را در بر می گیرد. با توجه به ماهیت فرامرزی خسارات ناشی از هوش مصنوعی، تحقق مؤثر تعهد جبران خسارت نیازمند توسعه سازوکارهای همکاری بین المللی و ایجاد استانداردهای مشترک است. بدون چنین هماهنگی ای، اجرای مؤثر مسئولیت بین المللی دولت ها در این حوزه با محدودیت های جدی مواجه خواهد شد و خطر گسترش بی پاسخ گویی در نظام حقوق بین الملل افزایش می یابد.

نتیجه گیری

بررسی تعهدات بین المللی دولت ها در قبال خسارات ناشی از سامانه های هوش مصنوعی خودمختار فرامرزی نشان می دهد که نظام حقوق بین الملل در وضعیت گذار میان قواعد سنتی مسئولیت و الزامات نوظهور فناوری محور قرار دارد. ساختار کلاسیک مسئولیت بین المللی که بر انتساب مستقیم، کنترل مؤثر و رابطه علیت نسبتاً خطی استوار است، در مواجهه با سامانه های هوشمند که دارای سطح بالایی از خودمختاری، یادگیری تطبیقی و رفتارهای غیرقابل پیش بینی هستند، با محدودیت های جدی مواجه شده است. در نتیجه، شکاف قابل توجهی میان واقعیت های فناوری و ظرفیت های حقوقی موجود ایجاد گردیده است.

یافته های این پژوهش نشان می دهد که اگرچه اصول بنیادین حقوق بین الملل مانند اصل انتساب، اصل منع آسیب فرامرزی، تعهد مراقبت مقتضی (Due Diligence) و اصل جبران کامل خسارت همچنان چارچوب اصلی تحلیل مسئولیت دولت ها را تشکیل می دهند، اما تفسیر و کاربرست آن ها در حوزه هوش مصنوعی نیازمند بازنگری و توسعه تفسیری است. به ویژه در حوزه انتساب رفتار سامانه های خودمختار، اتکای صرف به معیارهای سنتی مانند «کنترل مؤثر» کافی نبوده و باید رویکردهای جدیدی همچون انتساب ساختاری و مسئولیت مبتنی بر ریسک مورد توجه قرار گیرد. همچنین، تعهدات پیشگیرانه دولت ها در این حوزه اهمیت فزاینده ای یافته اند؛ به گونه ای که مسئولیت دولت ها صرفاً پسینی نبوده، بلکه به صورت پیش دستانه و مستمر در تمام چرخه عمر سامانه های هوش مصنوعی جریان دارد. در کنار آن، تعهد جبران خسارت نیز با چالش های عملی مهمی همچون تعیین رابطه سببیت، شناسایی فاعل مسئول و تفکیک نقش دولت ها از توسعه دهندگان خصوصی مواجه است. این امر نشان می دهد که نظام مسئولیت بین المللی در حوزه هوش مصنوعی نیازمند بازتعریف نهادی و هنجاری است. می توان نتیجه گرفت که حقوق بین الملل در حوزه هوش مصنوعی در مرحله «تکوین هنجاری» قرار دارد و هنوز به یک رژیم جامع و یکپارچه نرسیده است. با این حال، روندهای موجود در دکترین حقوقی، رویه های پراکنده و اسناد نرم حقوقی نشان دهنده حرکت تدریجی به سمت

شکل گیری چارچوبی نوین برای تنظیم مسئولیت دولت ها در قبال فناوری های خودمختار است.

• پیشنهادها

۱. تدوین معاهده بین المللی جامع در حوزه هوش مصنوعی پیشنهاد می شود جامعه بین المللی نسبت به تدوین یک سند الزام آور جهانی با هدف تنظیم مسئولیت دولت ها در قبال خسارات ناشی از سامانه های هوش مصنوعی خودمختار اقدام نماید تا خلأهای موجود در قواعد سنتی مسئولیت برطرف گردد .

۲. توسعه مفهوم انتساب توسعه یافته (Expanded Attribution) لازم است در دکترین حقوق بین الملل، مفهوم انتساب از سطح کنترل مستقیم فراتر رفته و شامل زنجیره طراحی، توسعه، آموزش و بهره برداری از سامانه های هوشمند نیز گردد .

۳. تقویت تعهد مراقبت مقتضی در حوزه فناوری های پرخطر دولت ها باید ملزم به ایجاد نظام های ارزیابی ریسک، نظارت مستمر و کنترل فنی بر سامانه های هوش مصنوعی پرخطر شوند و استانداردهای حداقلی بین المللی در این زمینه تدوین گردد .

۴. ایجاد نظام مسئولیت مبتنی بر ریسک در کنار مسئولیت مبتنی بر تقصیر با توجه به پیچیدگی فناوری های خودمختار، پیشنهاد می شود در کنار مسئولیت سنتی، مدل مسئولیت مبتنی بر ریسک برای جبران خسارات فرامرزی نیز مورد پذیرش قرار گیرد .

۵. توسعه سازوکارهای همکاری بین المللی و تبادل داده ایجاد نهادهای مشترک نظارتی و سامانه های تبادل اطلاعات میان دولت ها می تواند

نقش مهمی در پیشگیری از خسارات و افزایش شفافیت فعالیت سامانه‌های هوش مصنوعی ایفا کند .

۶. ایجاد صندوق بین‌المللی جبران خسارات ناشی از هوش مصنوعی برای حل مشکلات مربوط به تعیین مسئول مستقیم، تشکیل یک صندوق بین‌المللی می‌تواند به جبران سریع خسارات قربانیان کمک کرده و از اطلاع دادرسی جلوگیری نماید .

۷. تدوین استانداردهای فنی-حقوقی مشترک برای توسعه‌دهندگان هوش مصنوعی پیشنهاد می‌شود اصول حقوقی در قالب استانداردهای فنی قابل اجرا برای شرکت‌ها و توسعه‌دهندگان تبدیل شود تا از مرحله طراحی، انطباق با الزامات حقوق بین‌الملل تضمین گردد.

منابع و مآخذ

الف) منابع فارسی

۱. عباسعلی کدخدایی، حقوق بین الملل عمومی، تهران: نشر میزان، ۱۴۰۲.
۲. محمدعلی موحد، حقوق بین الملل عمومی، تهران: نشر خوارزمی، ۱۴۰۱.
۳. سید حسین صفایی، مسئولیت مدنی و بین المللی دولت ها، تهران: سمت، ۱۴۰۰.
۴. ناصر کاتوزیان، مسئولیت مدنی، تهران: نشر میزان، ۱۳۹۹.
۵. محسن محبی، حقوق بین الملل عمومی (جلد دوم)، تهران: جنگل، ۱۴۰۱.
۶. فریدون عظیمی، مبانی حقوق بین الملل معاصر، تهران: نشر سمت، ۱۴۰۲.
۷. عباس کریمی، حقوق فناوری های نوین و اینترنت، تهران: دانشگاه تهران، ۱۴۰۱.
۸. محمدهادی محمدی، حقوق سایبری و امنیت اطلاعات، تهران: نشر میزان، ۱۴۰۲.
۹. علی خالقی، حقوق کیفری بین المللی، تهران: شهر دانش، ۱۴۰۰.
۱۰. ابراهیم گلستانی، مبانی مسئولیت بین المللی دولت ها، تهران: مجد، ۱۳۹۹.
۱۱. رضا اسلامی، حقوق بین الملل و فناوری های نوین، تهران: سمت، ۱۴۰۲.
۱۲. سعید عابدی، حقوق اینترنت و فضای مجازی، تهران: جنگل، ۱۴۰۱.
۱۳. محمد رفیعی، مسئولیت دولت ها در حقوق بین الملل، تهران: میزان، ۱۴۰۰.
۱۴. علی نجفی ابرندآبادی، حقوق بین الملل عمومی کاربردی، تهران: سمت، ۱۴۰۱.

۱۵. حسین مهرپور، حقوق بشر و فناوری‌های نوین، تهران: نشر دادگستر، ۱۴۰۲.

۱۶. مهدی ذاکریان، حقوق بین‌الملل معاصر، تهران: میزان، ۱۴۰۳.

۱۷. کاظم غریب‌آبادی، حقوق بین‌الملل عمومی، تهران: مجد، ۱۴۰۱.

ب) منابع خارجی

۱. Crawford, James, *State Responsibility: The General Principles of International Law*, Cambridge University Press, ۲۰۱۳.
۲. Crawford, James; Pellet, Alain; Olleson, Simon, *The Law of International Responsibility*, Oxford University Press, ۲۰۱۰.
۳. Schmitt, Michael N. (ed.), *Tallinn Manual ۲.۰ on the International Law Applicable to Cyber Operations*, Cambridge University Press, ۲۰۱۷.
۴. Scherer, Matthew, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies", *Harvard Journal of Law & Technology*, ۲۰۱۶.
۵. Crootof, Rebecca, "The Killer Robots Are Here", *Northwestern University Law Review*, ۲۰۱۹.
۶. Floridi, Luciano, *The Ethics of Artificial Intelligence*, Oxford University Press, ۲۰۲۱.
۷. Russell, Stuart & Norvig, Peter, *Artificial Intelligence: A Modern Approach*, Pearson, ۲۰۲۰.

۸. Brownsword, Roger, "Regulating AI and Robotics", *Law, Innovation and Technology Journal*, ۲۰۲۰.
۹. United Nations, *Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA)*, ۲۰۰۱.
۱۰. ICJ Reports, *Case Law on State Responsibility and Due Diligence*, International Court of Justice Publications, various years.