نشریه علمی فقه، حقوق و علوم جزا
مقاله پژوهشی، سال دهم، شماره ٤١، پاییز ١٤٠٥، صفحات ١٣ تا ٤٧
تاریخ دریافت: ١٤٠٥/٠٥/٠٤ – تاریخ پذیرش: ١٤٠٥/٠٧/١١

مسئولیت کیفری دارنده‌ی هوش مصنوعی در ارتکاب جرایم سایبری با تأکید بر چالش هستی‌شناسی در مقررات کیفری ایران، آمریکا و اتحادیه اروپا

| علیرضا رحمانی نعیم آبادی*| استادیار، گروه فقه و مبانی حقوق، دانشکده حقوق، دانشگاه آزاد بندرعباس.
| اسحاق شهدادنژاد| دانشجوی دکتری، گروه فقه و مبانی حقوق، دانشکده حقوق، دانشگاه آزاد بندرعباس.

چکیده
گسترش فناوری‌های هوش مصنوعی و کاربرد گسترده آن در فضای سایبری، زمینه پیدایش جرایم نوینی مانند نفوذ غیرمجاز، بدافزارهای هوشمند، حملات خودکار، فیشینگ پیشرفته، جعل عمیق و سرقت داده‌ها را فراهم کرده است. این تحول، چالش اساسی انتساب رفتار مجرمانه و تعیین مسئولیت کیفری را به دلیل فاصله میان عملکرد الگوریتمی سامانه و تصمیم انسانی ایجاد کرده است. پژوهش حاضر با هدف بررسی مبانی مسئولیت کیفری دارنده هوش مصنوعی در ارتکاب جرایم سایبری و تحلیل تطبیقی رویکردهای حقوق ایران، آمریکا و اتحادیه اروپا انجام شده است. روش پژوهش، توصیفی-تحلیلی و مبتنی بر مطالعه تطبیقی مقررات کیفری، اسناد حقوقی و دیدگاه‌های نظری مرتبط با هوش مصنوعی است. یافته‌ها نشان می‌دهد که هوش مصنوعی به دلیل فقدان اراده مستقل، آگاهی اخلاقی و قابلیت تحمل مجازات، در شرایط کنونی نمی‌تواند به عنوان فاعل کیفری مستقل شناخته شود؛ بنابراین، مسئولیت کیفری باید بر اساس نقش و رفتار اشخاص انسانی مرتبط با سامانه، به‌ویژه دارنده، مالک یا بهره‌بردار آن، ارزیابی شود. در این زمینه، معیارهایی مانند کنترل مؤثر بر سامانه، قابلیت پیش‌بینی خطر، وجود تقصیر یا علم نسبت به احتمال وقوع جرم، رعایت تکلیف مراقبت و نظارت و رابطه سببیت میان فعل یا ترک فعل دارنده و نتیجه مجرمانه، مبنای اصلی انتساب مسئولیت کیفری محسوب می‌شوند. بررسی تطبیقی نشان می‌دهد که حقوق آمریکا بیشتر بر اصول سنتی مسئولیت کیفری و تقصیر انسانی تأکید دارد، در حالی که اتحادیه اروپا با رویکرد پیشگیرانه و نظام مدیریت خطر، الزامات نظارتی گسترده‌تری برای استفاده از هوش مصنوعی ایجاد کرده است. در حقوق ایران نیز با وجود امکان بهره‌گیری از قواعد عمومی کیفری و جرایم رایانه‌ای، خلأ مقررات خاص در زمینه مسئولیت ناشی از کاربرد هوش مصنوعی وجود دارد. بر این اساس، تدوین مقررات ویژه مبتنی بر اصل انسان‌محوری، تعیین حدود وظایف نظارتی و پیش‌بینی معیارهای روشن برای انتساب مسئولیت کیفری، ضروری به نظر می‌رسد.

واژگان کلیدی: هوش مصنوعی، دارنده هوش مصنوعی، جرایم سایبری، مسئولیت کیفری، چالش هستی‌شناختی، کنترل مؤثر، اتحادیه اروپا.

*نویسنده مسئول: hoseinrahimi542@gmail.com

مقدمه

با گسترش شتابان فناوری‌های دیجیتال، هوش مصنوعی از یک ابزار صرفاً کمکی در پردازش داده‌ها و خودکارسازی امور فراتر رفته و به عاملی اثرگذار در تصمیم‌سازی، پیش‌بینی، تولید محتوا و اجرای فرایندهای پیچیده بدل شده است. این تحول، اگرچه فرصت‌های گسترده‌ای در حوزه‌های اقتصادی، آموزشی، مدیریتی و امنیتی فراهم کرده، هم‌زمان ظرفیت‌های نوینی نیز برای ارتکاب رفتارهای مجرمانه پدید آورده است؛ به‌ویژه در قلمرو جرایم سایبری که سرعت، مقیاس و پنهان‌کاری در آن‌ها اهمیت ویژه‌ای دارد. در این میان، قابلیت یادگیری، تحلیل کلان‌داده، تصمیم‌گیری مبتنی بر الگو و اجرای خودکار برخی فرامین سبب شده است که هوش مصنوعی در بستر سایبری، نه‌تنها به‌عنوان ابزار تسهیل‌گر، بلکه گاه به‌مثابه عامل تقویت‌کننده و پیچیده‌کننده رفتار مجرمانه ظاهر شود (عظیمی و اسماعیلی، ۱۴۰۰؛ سهرابی و اکبرنژاد، ۱۴۰۱). از همین رو، بحث از مسئولیت کیفری ناشی از کاربست این فناوری، دیگر موضوعی فرضی یا آینده‌نگرانه نیست، بلکه به یکی از مسائل عینی و رو به گسترش حقوق کیفری معاصر تبدیل شده است.

نقش هوش مصنوعی در ارتکاب جرایم سایبری، تنها به استفاده ساده از نرم‌افزارهای هوشمند محدود نیست، بلکه در سال‌های اخیر اشکال پیشرفته‌تری از سوءاستفاده از این فناوری قابل شناسایی است. تولید و ارسال فیشینگ هوشمند با پیام‌های شخصی‌سازی‌شده، طراحی و انتشار بدافزارهای سازگارشونده، ساخت محتوای جعلی و فریبنده از طریق دیپ‌فیک، شناسایی خودکار آسیب‌پذیری‌های سامانه‌ها و حتی اجرای حملات خودکار در مقیاس وسیع، نمونه‌هایی از دگرگونی الگوی ارتکاب جرم در فضای سایبری‌اند (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲؛ جان‌آبادی، ۱۴۰۳؛ اباذری، ۱۴۰۳). این ویژگی‌ها موجب می‌شود که عنصر انسانی در فرایند وقوع جرم، گاه از حالت مباشرت مستقیم خارج شده و به سطحی از طراحی، آموزش، تنظیم، نظارت یا بهره‌برداری از سامانه‌های هوشمند منتقل شود. در نتیجه، هرچه سامانه هوش مصنوعی خودمختارتر، یادگیرنده‌تر و در تصمیم‌گیری‌های عملیاتی مستقل‌تر شود، پاسخ به این پرسش دشوارتر می‌گردد که در صورت وقوع جرم، مسئولیت کیفری دقیقاً بر عهده چه شخصی است. دشواری یادشده، از آنجا اهمیت بیشتری می‌یابد که در حقوق کیفری، اصل شخصی بودن مسئولیت، ضرورت احراز عنصر روانی و لزوم قابلیت انتساب رفتار مجرمانه به مرتکب، جایگاهی بنیادین دارد. در جرایم سنتی، شناسایی مباشر، شریک یا معاون جرم معمولاً بر پایه رفتار ملموس انسانی صورت می‌گیرد؛ اما در جرایم سایبری مبتنی بر هوش مصنوعی، گاه خروجی مجرمانه محصول سامانه‌ای است که پس از آموزش اولیه، بدون دستور لحظه‌ای انسان، اقدام به تولید محتوا، برقراری ارتباط، تحلیل هدف یا اجرای عملیات می‌کند. همین وضعیت، مسئله محوری این پژوهش را شکل می‌دهد: هرگاه عمل مجرمانه از سامانه‌ای هوشمند و تا حدی خودمختار سر بزند، مسئولیت کیفری دقیقاً متوجه چه کسی است؟ آیا باید مالک سامانه را مسئول دانست، یا بهره‌بردار، یا کنترل‌کننده فنی، یا توسعه‌دهنده، یا شخص حقوقی منتفع از عملکرد سامانه را؟ پاسخ به این مسئله، بدون تحلیل دقیق جایگاه «دارنده» و سنجش رابطه او با سامانه، امکان‌پذیر نیست (اسلامی، ۱۳۹۱؛ اباذری، ۱۴۰۳).

از همین‌جا، چالش هستی‌شناختی هوش مصنوعی نیز وارد بحث می‌شود. تا زمانی که هوش مصنوعی صرفاً ابزار تلقی شود، انتساب مسئولیت به انسان استفاده‌کننده یا هدایت‌کننده آن، کمابیش در چارچوب‌های سنتی حقوق کیفری قابل توجیه است؛ اما هرچه از ابزار بودن صرف فاصله گرفته و به سطحی از خودآیینی عملیاتی نزدیک شود، این پرسش شکل می‌گیرد که آیا می‌توان برای آن نوعی کنشگری مستقل در نظر گرفت یا همچنان باید آن را فاقد اراده حقوقی و قصد مجرمانه دانست. بخش مهمی از ادبیات حقوقی بر این باور است که هوش مصنوعی، علی‌رغم توان پردازشی و خودیادگیرندگی، همچنان فاقد شعور اخلاقی، اراده آزاد، درک هنجاری و قابلیت تحمل مجازات است و از این‌رو، نمی‌تواند فاعل مستقل مسئولیت کیفری به شمار آید (اسلامی، ۱۳۹۱؛ صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲). با وجود این، نفی شخصیت کیفری مستقل هوش مصنوعی، به‌معنای سادگی در انتساب مسئولیت به انسان نیز نیست؛ زیرا میان طراحی، مالکیت، کنترل مؤثر، امکان توقف سامانه، پیش‌بینی‌پذیری خطر و میزان نظارت، تفاوت‌های معناداری وجود دارد که هر یک می‌تواند در احراز مسئولیت کیفری مؤثر باشد.

بر این اساس، پژوهش حاضر در پی پاسخ به یک سؤال کلی است: در پرتو چالش هستی‌شناختی هوش مصنوعی، معیار انتساب مسئولیت کیفری به دارنده سامانه هوش مصنوعی در ارتکاب جرایم سایبری در حقوق ایران، آمریکا و اتحادیه اروپا چیست؟ در پاسخ مقدماتی به این پرسش، فرضیه کلی پژوهش آن است که مسئولیت کیفری دارنده هوش مصنوعی نه بر پایه مالکیت صِرف، بلکه بر اساس نقش واقعی او در ایجاد، هدایت، تداوم یا کنترل خطر مجرمانه، میزان کنترل مؤثر بر سامانه، قابلیت پیش‌بینی رفتار زیان‌بار و نیز نقض تکالیف نظارتی و امنیتی قابل احراز است. به بیان دیگر، هرچه رابطه دارنده با فرایند تصمیم‌سازی و عملکرد سامانه نزدیک‌تر، آگاهانه‌تر و قابل‌کنترل‌تر باشد، امکان انتساب کیفری تقویت می‌شود و هرچه این رابطه صرفاً صوری و فاقد تسلط واقعی باشد، انتساب کیفری با دشواری بیشتری روبه‌رو خواهد بود. مرور مختصر پیشینه پژوهش نشان می‌دهد که در ادبیات داخلی، بیشتر مطالعات بر تبیین کلی مسئولیت ناشی از هوش مصنوعی، آثار این فناوری در ارتکاب جرایم سایبری یا امکان‌سنجی مسئولیت کیفری آن متمرکز بوده‌اند. برای مثال، برخی پژوهش‌ها بر نقش هوش مصنوعی در تحول شیوه‌های ارتکاب جرایم سایبری و اقتصادی تأکید کرده‌اند (جان‌آبادی، ۱۴۰۳)، برخی دیگر بر سوءاستفاده از هوش مصنوعی در ارتکاب جرایم سایبری و پیامدهای کیفری آن تمرکز داشته‌اند (اباذری، ۱۴۰۳)، و گروهی نیز به‌طور خاص از منظر حقوقی به تأثیر هوش مصنوعی در جرم‌زایی سایبری پرداخته‌اند (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲). در آثار بنیادی‌تر نیز، بحث از مسئولیت کیفری هوش مصنوعی و الگوریتم‌های خودآموز، عمدتاً در سطح امکان یا امتناع شناسایی مسئولیت مستقل برای این سامانه‌ها بررسی شده است (اسلامی، ۱۳۹۱). با وجود ارزش این مطالعات، هنوز چارچوبی منسجم که به‌طور هم‌زمان سه مؤلفه «دارنده»، «چالش هستی‌شناختی» و «جرایم سایبری» را در یک مطالعه تطبیقی میان حقوق ایران، آمریکا و اتحادیه اروپا گرد آورد، کمتر دیده می‌شود.

در سطح تطبیقی نیز هر یک از نظام‌های مورد بررسی، رویکرد خاص خود را دارند. در حقوق ایران، به سبب فقدان مقررات اختصاصی درباره مسئولیت کیفری ناشی از هوش مصنوعی، تحلیل‌ها غالباً بر قواعد عامی چون مباشرت، معاونت، تسبیب، ترک فعل و در مواردی مسئولیت اشخاص حقوقی استوار

می‌شود. در حقوق آمریکا، رویکرد غالب پراکنده و جرم‌محور است و مسئولیت بیشتر ذیل عناوین مشخصی چون جرایم دسترسی غیرمجاز، تقلب رایانه‌ای یا مسئولیت شرکت‌ها تحلیل می‌شود. در اتحادیه اروپا نیز گرچه رویکردی پیشرفته‌تر و ریسک‌محور در تنظیم‌گری فناوری مشاهده می‌شود، اما این رویکرد الزاماً به معنای پذیرش یک نظام کیفری مستقل برای هوش مصنوعی نیست، بلکه بیشتر بر تکالیف نظارتی، مدیریت خطر، شفافیت، امنیت و حکمرانی پیشگیرانه تکیه دارد. خلأ اصلی در این میان آن است که هنوز مطالعه‌ای منسجم، این سه نظام را حول یک مسئله واحد، یعنی معیارهای انتساب مسئولیت کیفری به دارنده در بستر چالش هستی‌شناختی هوش مصنوعی، در کنار یکدیگر تحلیل نکرده است. پژوهش حاضر برای پر کردن این خلأ، با روش توصیفی-تحلیلی و رویکرد تطبیقی کارکردی انجام می‌شود؛ بدین معنا که یک مسئله مشخص، یعنی مسئولیت کیفری دارنده هوش مصنوعی در ارتکاب جرایم سایبری، در سه نظام حقوقی ایران، آمریکا و اتحادیه اروپا بر اساس کارکرد قواعد و معیارهای انتساب بررسی می‌شود. نوآوری اصلی این تحقیق در آن است که، افزون بر تحلیل مبانی نظری و تطبیقی، در پایان می‌کوشد الگویی منسجم برای «آزمون چندمرحله‌ای انتساب مسئولیت کیفری دارنده» ارائه دهد؛ الگویی که بتواند با لحاظ عواملی چون کنترل مؤثر، عنصر روانی، رابطه سببیت، میزان خودمختاری سامانه و تکالیف نظارتی و امنیتی، راهنمای تحلیل پرونده‌های مربوط به جرایم سایبری مبتنی بر هوش مصنوعی باشد. از این منظر، پژوهش نه‌تنها در پی توصیف وضعیت موجود، بلکه درصدد ارائه چارچوبی تحلیلی برای مواجهه دقیق‌تر حقوق کیفری با یکی از پیچیده‌ترین مسائل عصر فناوری است.

مبانی مفهومی و نظری

مفهوم هوش مصنوعی و دارنده هوش مصنوعی

هوش مصنوعی در دهه‌های اخیر از یک فناوری صرفاً محاسباتی به یکی از مهم‌ترین فناوری‌های تحول‌آفرین در عرصه‌های اقتصادی، اجتماعی، امنیتی و حقوقی تبدیل شده است. این فناوری با بهره‌گیری از الگوریتم‌های یادگیری ماشین، یادگیری عمیق و شبکه‌های عصبی، قادر است داده‌های گسترده را تحلیل کرده، الگوهای جدید را شناسایی کند و در بسیاری از موارد بدون مداخله مستقیم انسان تصمیم‌گیری نماید. ویژگی اخیر، یعنی برخورداری از نوعی استقلال عملیاتی، موجب شده است که هوش مصنوعی از جایگاه یک ابزار رایانه‌ای متعارف فراتر رفته و به یکی از موضوعات اساسی مطالعات حقوقی، به‌ویژه در حوزه مسئولیت کیفری، تبدیل شود. با وجود این، استقلال عملکردی هوش مصنوعی به معنای برخورداری آن از اراده یا شخصیت حقوقی نیست و اغلب نظام‌های حقوقی همچنان آن را به‌عنوان فناوری‌ای وابسته به انسان تلقی می‌کنند که آثار حقوقی عملکرد آن نهایتاً به اشخاص انسانی یا حقوقی مرتبط با آن بازمی‌گردد (کاوه و بارانی، ۱۴۰۳، ص. ۵۳؛ عبداله‌پور، ۱۴۰۴، صص. ۴-۵). در ادبیات حقوقی، تعریف هوش مصنوعی صرفاً به قابلیت انجام محاسبات یا پردازش داده‌ها محدود نمی‌شود، بلکه بر توانایی سامانه در یادگیری از داده‌های جدید، اصلاح رفتار، تحلیل محیط و اتخاذ تصمیم‌های مبتنی بر الگوریتم نیز تأکید می‌شود. از همین رو، مقررات جدید اتحادیه اروپا، به‌ویژه قانون هوش مصنوعی، سامانه هوش مصنوعی را نظامی معرفی می‌کند که با درجات متفاوتی از خودمختاری،

خروجی‌هایی نظیر پیش‌بینی، توصیه یا تصمیم تولید می‌کند و این خروجی‌ها می‌توانند بر محیط فیزیکی یا دیجیتال اثرگذار باشند. این تعریف، بیش از آنکه بر ساختار فنی سامانه تمرکز داشته باشد، بر میزان استقلال عملکرد و آثار اجتماعی و حقوقی آن تأکید می‌کند؛ امری که در تحلیل مسئولیت کیفری نیز اهمیت بنیادین دارد (کشاورزصفی‌یی، ۱۴۰۴، صص. ۲۷۳-۲۷۵).

ویژگی مهم سامانه‌های نوین هوش مصنوعی، قابلیت یادگیری مستمر و تغییر رفتار بر اساس داده‌های جدید است. برخلاف نرم‌افزارهای سنتی که صرفاً دستورات از پیش تعیین‌شده را اجرا می‌کردند، سامانه‌های مبتنی بر یادگیری ماشین می‌توانند در فرآیند اجرا، راهکارهای تازه‌ای برای حل مسائل پیدا کرده و حتی روش تصمیم‌گیری خود را اصلاح کنند. همین ویژگی سبب شده است که در برخی کاربردها، از جمله امنیت سایبری، سامانه‌های هوشمند قادر باشند به صورت خودکار آسیب‌پذیری‌های شبکه را شناسایی کرده یا در برابر تهدیدهای جدید واکنش نشان دهند. در مقابل، همین قابلیت می‌تواند زمینه سوءاستفاده مجرمانه را نیز فراهم کند و از هوش مصنوعی به‌عنوان ابزاری برای طراحی بدافزارهای هوشمند، حملات سایبری خودکار یا جعل عمیق استفاده شود؛ موضوعی که مسئولیت اشخاص مرتبط با این سامانه‌ها را به یکی از مهم‌ترین مباحث حقوق کیفری معاصر تبدیل کرده است (عبداله‌پور، ۱۴۰۴، صص. ۷-۹). در حقوق کیفری، شناسایی شخص مسئول مستلزم تعیین دقیق مفهوم «دارنده هوش مصنوعی» است. اصطلاح دارنده، مفهومی عام است و صرفاً به مالک قانونی سامانه محدود نمی‌شود، بلکه اشخاص مختلفی را دربرمی‌گیرد که در ایجاد، بهره‌برداری، کنترل یا مدیریت سامانه نقش مؤثر دارند. از این رو، در تحلیل مسئولیت کیفری باید میان مالک، بهره‌بردار، کنترل‌کننده، اپراتور و حتی استفاده‌کننده تجاری تمایز قائل شد؛ زیرا میزان مسئولیت هر یک تابع حدود اختیارات، میزان کنترل مؤثر و نقش واقعی آنان در عملکرد سامانه خواهد بود. در نتیجه، صرف مالکیت یک سامانه هوش مصنوعی، بدون احراز رابطه سببیت یا تقصیر، برای انتساب مسئولیت کیفری کافی نیست (کشاورزصفی‌یی، ۱۴۰۴، صص. ۲۷۸-۲۸۰). مفهوم دارنده در اسناد اتحادیه اروپا نیز عمدتاً با رویکردی کارکردی تفسیر شده است. در این رویکرد، مسئولیت متوجه شخصی است که کنترل عملی بر عملکرد سامانه داشته، آن را در محیط واقعی به کار گرفته یا امکان مدیریت خطرات ناشی از آن را در اختیار دارد. به همین دلیل، مقررات اروپایی میان تولیدکننده، ارائه‌دهنده، بهره‌بردار و استفاده‌کننده نهایی تفکیک قائل شده و برای هر یک متناسب با میزان مداخله در چرخه حیات سامانه، تکالیف متفاوتی در زمینه ارزیابی ریسک، نظارت و رعایت الزامات ایمنی پیش‌بینی کرده است. این نگرش نشان می‌دهد که مفهوم دارنده، مفهومی صرفاً مالکیتی نیست، بلکه مبتنی بر معیار «کنترل بر خطر» و «توانایی پیشگیری از زیان» تحلیل می‌شود (کشاورزصفی‌یی، ۱۴۰۴، صص. ۲۸۲-۲۸۵).

در حقوق ایران، اگرچه تاکنون تعریف قانونی مستقلی از دارنده هوش مصنوعی ارائه نشده است، اما با استناد به قواعد عمومی مسئولیت، می‌توان دارنده را شخصی دانست که اختیار بهره‌برداری، کنترل یا مدیریت سامانه را در اختیار دارد و از منافع آن بهره‌مند می‌شود. از این منظر، مسئولیت کیفری یا مدنی چنین شخصی نه صرفاً بر پایه مالکیت، بلکه بر مبنای میزان دخالت در ایجاد، کنترل یا مدیریت خطرات ناشی از عملکرد سامانه ارزیابی خواهد شد. این برداشت با رویکردهای نوین حقوق تطبیقی نیز همسو است و می‌تواند مبنای مناسبی برای تحلیل مسئولیت دارنده هوش مصنوعی در نظام حقوقی

ایران فراهم آورد (کاوه و بارانی، ۱۴۰۳، صص. ۵۶-۵۸). گسترش کاربرد هوش مصنوعی مولد نیز مفهوم دارنده را با پیچیدگی‌های بیشتری مواجه ساخته است. سامانه‌هایی که قادر به تولید متن، تصویر، صوت یا ویدئو هستند، علاوه بر کاربردهای مشروع، می‌توانند در ارتکاب جرایمی همچون جعل عمیق، جعل هویت، نشر اطلاعات خلاف واقع، اخاذی و کلاهبرداری سایبری مورد استفاده قرار گیرند. پژوهش‌های انجام‌شده نشان می‌دهد که فناوری دیپ‌فیک به دلیل شباهت بسیار زیاد خروجی‌های تولیدشده با نمونه‌های واقعی، تشخیص اصالت محتوا را دشوار ساخته و اعتبار ادله الکترونیکی را نیز با چالش مواجه کرده است. ازاین‌رو، در چنین مواردی، بررسی مسئولیت شخصی که سامانه را آموزش داده، کنترل کرده یا در اختیار دیگران قرار داده است، اهمیت مضاعفی پیدا می‌کند (ماراس و الکساندر[2]، ۲۰۱۹؛ صص ۲۵۷-۲۶۰). علاوه بر این، توسعه فناوری‌های مبتنی بر بلاک‌چین و قراردادهای هوشمند برای مقابله با سوءاستفاده از هوش مصنوعی، نشان می‌دهد که مدیریت مخاطرات این فناوری صرفاً از طریق ابزارهای فنی امکان‌پذیر نیست و باید با سازوکارهای حقوقی و کیفری همراه شود. پیشنهاد استفاده از بلاک‌چین برای احراز اصالت محتوای دیجیتال و مقابله با دیپ‌فیک، هرچند می‌تواند احتمال وقوع برخی جرایم را کاهش دهد، اما مسئولیت اشخاصی را که آگاهانه از سامانه‌های هوش مصنوعی برای ارتکاب جرایم سایبری استفاده می‌کنند، منتفی نمی‌سازد. بنابراین، حتی با توسعه ابزارهای پیشگیرانه، نقش دارنده یا بهره‌بردار سامانه در مدیریت خطر و جلوگیری از سوءاستفاده‌های مجرمانه همچنان محور اصلی تحلیل مسئولیت حقوقی و کیفری باقی خواهد ماند (حسن و صلاح[3]، ۲۰۱۹؛ صص ۶۰۰-۶۰۳).

مفهوم مسئولیت کیفری و جرایم سایبری

مسئولیت کیفری یکی از بنیادی‌ترین نهادهای حقوق جزا است که بر اساس آن، شخصی که با تحقق شرایط قانونی، رفتار مجرمانه‌ای را مرتکب شده باشد، در برابر جامعه و نظام عدالت کیفری پاسخگو شناخته می‌شود. تحقق این مسئولیت مستلزم وجود عناصر قانونی، مادی و روانی جرم است و تنها زمانی امکان‌پذیر خواهد بود که رفتار مجرمانه به شخصی دارای اراده، اختیار و قابلیت سرزنش منتسب شود. اصل شخصی بودن مسئولیت کیفری نیز بر همین مبنا استوار است و اقتضا می‌کند که هیچ شخصی جز مرتکب رفتار مجرمانه مورد تعقیب یا مجازات قرار نگیرد. با ظهور فناوری‌های نوین، به‌ویژه سامانه‌های مبتنی بر هوش مصنوعی، این مبانی سنتی با پرسش‌های جدیدی روبه‌رو شده‌اند؛ زیرا در بسیاری از موارد، نتیجه مجرمانه مستقیماً توسط انسان ایجاد نمی‌شود، بلکه سامانه‌ای هوشمند در فرآیند ارتکاب جرم نقش مؤثر ایفا می‌کند. ازاین‌رو، مهم‌ترین مسئله در حقوق کیفری معاصر، نه شناسایی مسئولیت برای خود هوش مصنوعی، بلکه تعیین مبنای انتساب رفتار مجرمانه به اشخاصی است که در طراحی، آموزش، کنترل یا بهره‌برداری از آن نقش دارند (کاوه و بارانی، ۱۴۰۳، صص. ۵۴-۵۸). در حقوق ایران، مسئولیت کیفری همچنان بر پایه اصول کلاسیک حقوق جزا استوار است و قانون‌گذار، تحقق آن را منوط به وجود رفتار مجرمانه منتسب به شخص دارای اراده و سوءنیت یا تقصیر

---

[2]. Maras & Alexandrou
[3]. Hasan & Salah

دانسته است. از آنجا که هوش مصنوعی فاقد شخصیت حقوقی، اراده مستقل و عنصر روانی به مفهوم حقوق کیفری است، نمی‌توان آن را به‌عنوان مرتکب جرم شناخت. در نتیجه، مسئولیت کیفری ناشی از عملکرد سامانه‌های هوشمند باید از طریق قواعد سنتی مباشرت، معاونت، سببیت و تقصیر نسبت به اشخاص انسانی یا حقوقی مرتبط با سامانه تحلیل شود. این رویکرد اگرچه با اصول بنیادین حقوق کیفری سازگار است، اما در مواجهه با سامانه‌های خودآموز و دارای استقلال عملکردی نسبی، پاسخگوی همه پیچیدگی‌های موجود نیست و نیازمند بازنگری در معیارهای انتساب مسئولیت است (کاوه و بارانی، ۱۴۰۳، صص. ۵۸-۶۰). در مقابل، نظام‌های حقوقی پیشرو، به‌ویژه اتحادیه اروپا، مفهوم مسئولیت را بیش از گذشته با مدیریت خطر و تکلیف به مراقبت پیوند زده‌اند. در اسناد اتحادیه اروپا، به جای آنکه بر فاعلیت کیفری هوش مصنوعی تأکید شود، مسئولیت اشخاصی که سامانه را طراحی، عرضه یا بهره‌برداری می‌کنند، بر پایه میزان کنترل مؤثر، قابلیت پیش‌بینی خطر و رعایت الزامات قانونی ارزیابی می‌شود. به همین دلیل، مسئولیت در این نظام حقوقی بیش از آنکه نتیجه صرف وقوع جرم باشد، به نقض تعهدات پیشگیرانه، ارزیابی ریسک و کوتاهی در نظارت وابسته است. چنین رویکردی نشان می‌دهد که تحول فناوری، مفهوم مسئولیت را از یک نهاد صرفاً واکنشی به نهادی پیشگیرانه و مبتنی بر مدیریت مخاطره سوق داده است (کشاورزصفی‌یی، ۱۴۰۴، صص. ۲۸۱-۲۸۶).

از منظر حقوق تطبیقی، توسعه هوش مصنوعی موجب تحول در مفهوم عنصر روانی نیز شده است. در بسیاری از جرایم سنتی، احراز قصد مجرمانه یا بی‌احتیاطی مرتکب نسبتاً آسان است؛ اما در جرایمی که با بهره‌گیری از سامانه‌های خودآموز ارتکاب می‌یابند، ممکن است دارنده سامانه قصد مستقیمی برای تحقق نتیجه مجرمانه نداشته باشد، هرچند از خطرات بالقوه آن آگاه بوده یا باید از آن آگاه می‌بود. به همین دلیل، برخی نظام‌های حقوقی به‌ویژه در اروپا، مفهوم «نقض وظیفه مراقبت» و «قابلیت پیش‌بینی خطر» را به‌عنوان معیارهای مکمل عنصر روانی مورد توجه قرار داده‌اند تا بتوانند مسئولیت اشخاص را متناسب با میزان مداخله آنان در مدیریت سامانه تحلیل کنند (عبداله‌پور، ۱۴۰۴، صص. ۱۰—۱۲). جرایم سایبری نیز از مهم‌ترین حوزه‌هایی هستند که توسعه هوش مصنوعی، ماهیت و شیوه ارتکاب آن‌ها را دگرگون ساخته است. جرایم سایبری به رفتارهای مجرمانه‌ای اطلاق می‌شود که موضوع، وسیله یا بستر ارتکاب آن‌ها سامانه‌های رایانه‌ای، شبکه‌های ارتباطی یا داده‌های دیجیتال است. اگرچه در گذشته بیشتر این جرایم توسط کاربران انسانی و از طریق ابزارهای متعارف رایانه‌ای انجام می‌شد، اما امروزه سامانه‌های هوش مصنوعی قادرند بخش قابل توجهی از فرآیند طراحی، اجرا و توسعه حملات سایبری را به صورت خودکار انجام دهند. همین تحول، مرز میان ابزار و فاعل را تا حدی مبهم ساخته و تعیین شخص مسئول را با دشواری‌های جدیدی مواجه کرده است (عبداله‌پور، ۱۴۰۴، صص. ۵-۸).

یکی از مهم‌ترین مصادیق جرایم سایبری مبتنی بر هوش مصنوعی، جعل عمیق است. فناوری دیپ‌فیک با استفاده از الگوریتم‌های یادگیری عمیق، تصاویر، صداها و ویدئوهایی تولید می‌کند که از نظر ظاهری تفاوت اندکی با نمونه‌های واقعی دارند. این فناوری می‌تواند در ارتکاب جرایمی همچون کلاهبرداری، جعل هویت، اخاذی، انتشار اخبار جعلی، تخریب شخصیت و حتی مداخله در فرآیندهای انتخاباتی مورد استفاده قرار گیرد. پژوهش‌های انجام‌شده نشان می‌دهد که گسترش دیپ‌فیک، اعتبار ادله الکترونیکی

را نیز با چالش جدی روبه‌رو ساخته و تشخیص اصالت محتوا را برای مراجع قضایی دشوار کرده است؛ امری که ضرورت بازنگری در قواعد اثبات جرم و مسئولیت کیفری را آشکار می‌سازد (ماراس و الکساندر، ۲۰۱۹؛ صص ۲۵۶-۲۶۱). علاوه بر جعل عمیق، هوش مصنوعی نقش مهمی در توسعه بدافزارهای هوشمند، حملات خودکار، فیشینگ هدفمند و نفوذ به سامانه‌های اطلاعاتی ایفا می‌کند. برخلاف حملات سنتی که به هدایت مستمر انسان نیاز داشتند، سامانه‌های مبتنی بر یادگیری ماشین می‌توانند آسیب‌پذیری‌های امنیتی را به صورت خودکار شناسایی کرده، روش حمله را متناسب با شرایط تغییر دهند و حتی برای عبور از سامانه‌های دفاعی، رفتار خود را اصلاح کنند. این ویژگی‌ها موجب افزایش سرعت، دقت و گستره حملات سایبری شده و مقابله با آن‌ها را برای نهادهای امنیتی و قضایی دشوارتر ساخته است (عبداله‌پور، ۱۴۰۴، صص. ۹-۱۲). در برابر این تهدیدها، پژوهشگران راهکارهای فناورانه مختلفی را نیز پیشنهاد کرده‌اند. یکی از مهم‌ترین این راهکارها، استفاده از فناوری بلاک‌چین و قراردادهای هوشمند برای ثبت، رهگیری و احراز اصالت داده‌های دیجیتال است. استفاده از این فناوری‌ها می‌تواند امکان شناسایی محتوای جعلی، رهگیری منشأ تولید اطلاعات و کاهش احتمال سوءاستفاده از سامانه‌های هوش مصنوعی را فراهم آورد. با این حال، حتی پیشرفته‌ترین سازوکارهای فنی نیز جایگزین قواعد حقوق کیفری نخواهند شد؛ زیرا مسئولیت ناشی از بهره‌برداری نادرست یا مجرمانه از هوش مصنوعی همچنان متوجه اشخاص انسانی باقی می‌ماند و فناوری صرفاً نقش ابزار کمکی در پیشگیری و کشف جرم را ایفا می‌کند (حسن و صلاح، ۲۰۱۹؛ صص ۲۹۸-۶۰۴). برآیند مباحث فوق نشان می‌دهد که مفهوم مسئولیت کیفری و جرایم سایبری در عصر هوش مصنوعی دچار تحول اساسی شده است. هرچند اصول بنیادین حقوق کیفری، از جمله اصل شخصی بودن مسئولیت، همچنان اعتبار خود را حفظ کرده‌اند، اما ویژگی‌هایی مانند خودمختاری نسبی سامانه‌های هوشمند، یادگیری مستمر و پیچیدگی الگوریتمی، ضرورت بازتفسیر معیارهای سنتی انتساب، تقصیر، رابطه سببیت و وظیفه مراقبت را آشکار ساخته‌اند. ازاین‌رو، تحلیل مسئولیت کیفری دارنده هوش مصنوعی دیگر صرفاً بر مبنای مالکیت یا استفاده از سامانه امکان‌پذیر نیست، بلکه باید میزان کنترل، قابلیت پیش‌بینی خطر و نقش واقعی وی در مدیریت سامانه نیز مورد توجه قرار گیرد؛ مباحثی که زمینه ورود به بحث «چالش هستی‌شناسی در حقوق کیفری» را فراهم می‌سازند.

چالش هستی‌شناسی در حقوق کیفری

چالش هستی‌شناسی یکی از بنیادی‌ترین مباحث نظری در حقوق کیفری هوش مصنوعی است که به ماهیت وجودی سامانه‌های هوشمند و جایگاه آن‌ها در نظام مسئولیت کیفری می‌پردازد. پرسش اصلی این است که آیا هوش مصنوعی صرفاً ابزاری پیشرفته در اختیار انسان است یا آنکه به دلیل برخورداری از قابلیت یادگیری، تصمیم‌گیری و عملکرد نسبتاً مستقل، باید جایگاهی متفاوت از ابزارهای سنتی برای آن در نظر گرفت. پاسخ به این پرسش، تأثیر مستقیمی بر تعیین فاعل جرم، انتساب رفتار مجرمانه و شناسایی شخص مسئول دارد؛ زیرا حقوق کیفری کلاسیک بر این فرض استوار است که جرم تنها از سوی شخصی دارای اراده، شعور و قابلیت سرزنش ارتکاب می‌یابد. ازاین‌رو، توسعه سامانه‌های هوشمند، مبانی فلسفی و حقوقی مسئولیت کیفری را با چالشی جدی مواجه ساخته است (کاوه و بارانی، ۱۴۰۳،

صص. ۵۳-۵۶). در حقوق کیفری سنتی، رابطه میان انسان و ابزار، رابطه‌ای کاملاً تبعی تلقی می‌شود؛ به این معنا که ابزار فاقد هرگونه استقلال در تصمیم‌گیری است و تمامی آثار حقوقی عملکرد آن به استفاده‌کننده انسانی منتسب می‌شود. با این حال، سامانه‌های مبتنی بر یادگیری ماشین و یادگیری عمیق، برخلاف ابزارهای متعارف، قادرند داده‌های جدید را تحلیل کرده، الگوهای رفتاری خود را اصلاح کنند و در شرایط پیش‌بینی‌نشده، تصمیم‌هایی اتخاذ نمایند که لزوماً از پیش توسط برنامه‌نویس تعیین نشده است. همین استقلال عملکردی نسبی موجب شده است که برخی صاحب‌نظران از «شکاف انتساب[4]» سخن بگویند؛ وضعیتی که در آن تعیین شخص مسئول میان طراح، تولیدکننده، دارنده، بهره‌بردار و حتی خود سامانه با ابهام مواجه می‌شود. با وجود این، اکثر نظام‌های حقوقی هنوز این استقلال عملکردی را برای شناسایی شخصیت کیفری مستقل کافی نمی‌دانند (عبداله‌پور، ۱۴۰۴، صص. ۹-۱۱).

یکی از مهم‌ترین ابعاد چالش هستی‌شناسی، مسئله اراده و عنصر روانی جرم است. مسئولیت کیفری در اغلب نظام‌های حقوقی بر پایه وجود قصد مجرمانه یا دست‌کم تقصیر استوار است و این عناصر مستلزم برخورداری مرتکب از ادراک، اختیار و توانایی تشخیص رفتار خویش هستند. اگرچه سامانه‌های هوش مصنوعی قادر به پردازش اطلاعات، انتخاب میان گزینه‌های مختلف و حتی اصلاح تصمیم‌های خود هستند، اما این قابلیت‌ها از منظر حقوق کیفری معادل اراده انسانی محسوب نمی‌شوند. تصمیم‌های الگوریتمی نتیجه پردازش داده‌ها و اجرای مدل‌های ریاضی هستند و فاقد آگاهی، قصد اخلاقی و مسئولیت‌پذیری به معنای حقوقی‌اند. از همین رو، انتساب مستقیم جرم به هوش مصنوعی با اصول بنیادین حقوق کیفری، از جمله اصل شخصی بودن مسئولیت و اصل قابلیت سرزنش، ناسازگار تلقی می‌شود (کاوه و بارانی، ۱۴۰۳، صص. ۵۷-۶۰). در نظام حقوقی ایران، قوانین کیفری بر مبنای مسئولیت اشخاص حقیقی و در موارد محدود اشخاص حقوقی تنظیم شده‌اند و هیچ‌گونه شناسایی مستقلی برای سامانه‌های هوش مصنوعی به‌عنوان فاعل جرم وجود ندارد. قانون مجازات اسلامی تحقق مسئولیت کیفری را به وجود اراده، ادراک و قابلیت انتساب رفتار مجرمانه به شخص مسئول منوط کرده است؛ شرایطی که در مورد سامانه‌های هوشمند قابل احراز نیست. بنابراین، هرچند ممکن است هوش مصنوعی در فرآیند ارتکاب جرم نقش مؤثری ایفا کند، اما این نقش از منظر حقوق ایران همچنان در قالب ابزار یا وسیله ارتکاب جرم تحلیل می‌شود و مسئولیت متوجه اشخاصی خواهد بود که در طراحی، کنترل، آموزش یا بهره‌برداری از آن نقش داشته‌اند (کاوه و بارانی، ۱۴۰۳، صص. ۵۸-۶۱). در ایالات متحده آمریکا نیز اگرچه مقررات جامعی درباره مسئولیت کیفری هوش مصنوعی وجود ندارد، اما رویه‌های حقوقی و اسناد سیاست‌گذاری، مسئولیت را عمدتاً بر مبنای رفتار انسان تحلیل می‌کنند. در این رویکرد، سامانه هوش مصنوعی فاقد شخصیت حقوقی مستقل است و مسئولیت بر عهده توسعه‌دهنده، عرضه‌کننده یا بهره‌برداری قرار می‌گیرد که بر عملکرد سامانه کنترل مؤثر داشته یا امکان پیش‌بینی خطرات ناشی از آن را دارا بوده است. از این منظر، خودمختاری فنی سامانه موجب قطع رابطه انتساب نمی‌شود، بلکه صرفاً ارزیابی رابطه سببیت و میزان تقصیر اشخاص مرتبط را پیچیده‌تر می‌سازد. بنابراین، حقوق آمریکا نیز علی‌رغم پذیرش واقعیت استقلال عملکردی هوش مصنوعی،

---

4. Attribution Gap

همچنان بر محور مسئولیت انسانی استوار باقی مانده است (عبداله‌پور، ۱۴۰۴، صص. ۱۰-۱۲). در مقابل، اتحادیه اروپا رویکردی آینده‌نگرتر اتخاذ کرده و ضمن پرهیز از اعطای شخصیت کیفری به هوش مصنوعی، تلاش کرده است آثار ناشی از استقلال عملکردی این سامانه‌ها را از طریق مقررات ویژه مدیریت کند. اسناد اروپایی، به‌ویژه قانون هوش مصنوعی[5]، بر این مبنا تنظیم شده‌اند که هرچه میزان خودمختاری و سطح خطر سامانه بیشتر باشد، تکالیف قانونی اشخاص مرتبط نیز افزایش می‌یابد. به همین دلیل، مسئولیت در این نظام بیش از آنکه بر شخصیت سامانه مبتنی باشد، بر مفهوم «کنترل مؤثر»، «ارزیابی ریسک» و «قابلیت پیش‌بینی خطر» استوار است. این رویکرد، ضمن حفظ اصول سنتی حقوق کیفری، می‌کوشد خلأ ناشی از استقلال عملکردی سامانه‌های هوشمند را از طریق توسعه تعهدات پیشگیرانه جبران کند (کشاورزصفی‌یی، ۱۴۰۴، صص. ۲۸۲-۲۸۷).

چالش هستی‌شناسی در جرایم سایبری نمود آشکارتری پیدا می‌کند؛ زیرا بسیاری از حملات سایبری مبتنی بر هوش مصنوعی بدون مداخله مستقیم و لحظه‌ای انسان انجام می‌شوند. سامانه‌های هوشمند می‌توانند آسیب‌پذیری‌های شبکه را شناسایی، بدافزارهای جدید تولید، حملات فیشینگ را شخصی‌سازی و حتی روش‌های نفوذ خود را در حین اجرا اصلاح کنند. در چنین شرایطی، تشخیص اینکه نتیجه مجرمانه تا چه اندازه ناشی از تصمیم انسان و تا چه میزان حاصل فرآیند یادگیری سامانه است، دشوار می‌شود. این وضعیت نه تنها انتساب رفتار مجرمانه را با ابهام روبه‌رو می‌کند، بلکه ارزیابی رابطه سببیت، تقصیر و قابلیت پیش‌بینی را نیز پیچیده‌تر می‌سازد و ضرورت بازاندیشی در معیارهای سنتی مسئولیت کیفری را آشکار می‌کند (عبداله‌پور، ۱۴۰۴، صص. ۸-۱۲). چالش هستی‌شناسی در حوزه ادله کیفری نیز آثار قابل توجهی بر جای گذاشته است. فناوری‌هایی مانند دیپ‌فیک با بهره‌گیری از هوش مصنوعی قادرند تصاویر، صداها و ویدئوهایی تولید کنند که تشخیص آن‌ها از نمونه‌های واقعی بسیار دشوار است. این وضعیت، علاوه بر افزایش امکان ارتکاب جرایمی مانند جعل، اخاذی، کلاهبرداری و انتشار اطلاعات خلاف واقع، اعتبار ادله دیجیتال را نیز با تردید مواجه ساخته است. پژوهش‌های انجام‌شده نشان می‌دهد که گسترش این فناوری‌ها، مراجع قضایی را ناگزیر از استفاده هم‌زمان از ابزارهای فنی و معیارهای حقوقی برای احراز اصالت ادله ساخته و در نتیجه، نقش دارنده و بهره‌بردار سامانه در مدیریت مخاطرات ناشی از آن بیش از گذشته اهمیت یافته است. در مجموع، چالش هستی‌شناسی در حقوق کیفری را نمی‌توان صرفاً مسئله‌ای فلسفی یا فناورانه دانست، بلکه این چالش مستقیماً بر ارکان مسئولیت کیفری، معیارهای انتساب جرم، رابطه سببیت، عنصر روانی و حدود پاسخگویی اشخاص اثر می‌گذارد. بررسی تطبیقی حقوق ایران، آمریکا و اتحادیه اروپا نشان می‌دهد که اگرچه هیچ‌یک از این نظام‌ها تاکنون برای هوش مصنوعی شخصیت کیفری مستقل قائل نشده‌اند، اما همگی پذیرفته‌اند که ویژگی‌های نوین این فناوری، به‌ویژه استقلال عملکردی و یادگیری مستمر، تفسیر سنتی قواعد مسئولیت کیفری را با چالش مواجه ساخته است. ازاین‌رو، تحلیل مسئولیت کیفری دارنده هوش مصنوعی باید بر پایه معیارهایی همچون کنترل مؤثر، قابلیت پیش‌بینی خطر، میزان مداخله در چرخه عملکرد سامانه و رعایت تکالیف قانونی استوار شود؛ معیارهایی که در بخش‌های

---

[5]. AI Act

بعدی مقاله به صورت تطبیقی و با تمرکز بر نظام‌های حقوقی ایران، آمریکا و اتحادیه اروپا بررسی خواهند شد.

تحلیل تطبیقی چالش هستی‌شناسی در ایران، آمریکا و اتحادیه اروپا (حدود ۲۰۰۰ کلمه)

هوش مصنوعی؛ ابزار یا فاعل؟

یکی از بنیادی‌ترین پرسش‌های حقوق کیفری معاصر آن است که هوش مصنوعی در فرآیند ارتکاب جرم چه جایگاهی دارد؛ آیا باید آن را همانند ابزارهای سنتی وسیله‌ای در اختیار انسان دانست یا آنکه به دلیل برخورداری از قابلیت یادگیری، تصمیم‌گیری و اجرای مستقل برخی فرایندها، می‌توان آن را نوعی فاعل نوظهور تلقی کرد؟ پاسخ به این پرسش صرفاً جنبه نظری ندارد، بلکه آثار مستقیمی بر تعیین مسئولیت کیفری دارنده، تولیدکننده، برنامه‌نویس و بهره‌بردار سامانه‌های هوشمند برجای می‌گذارد. در حقوق کیفری کلاسیک، ابزار فاقد هرگونه اراده و استقلال است و تمامی نتایج ناشی از عملکرد آن به شخص استفاده‌کننده منتسب می‌شود؛ اما ظهور سامانه‌های مبتنی بر یادگیری ماشین و یادگیری عمیق، این برداشت سنتی را با چالش مواجه ساخته است. امروزه بسیاری از سامانه‌های هوش مصنوعی قادرند بدون صدور دستور مستقیم انسان، داده‌های جدید را تحلیل کرده، راهبردهای تصمیم‌گیری خود را اصلاح کنند و حتی در محیط‌های ناشناخته واکنش‌های متفاوتی از خود نشان دهند. همین ویژگی سبب شده است که برخی اندیشمندان از «فاعل الگوریتمی» سخن بگویند، هرچند هنوز اکثریت نظام‌های حقوقی چنین جایگاهی را برای هوش مصنوعی به رسمیت نشناخته‌اند (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۴-۶؛ کوورت[6]، ۲۰۱۷؛ صص ۱۸۳-۱۸۶). از منظر فلسفه حقوق، فاعل جرم موجودی است که رفتار مجرمانه از اراده آگاهانه او ناشی شود و بتوان وی را به لحاظ اخلاقی و حقوقی مورد سرزنش قرار داد. بر این اساس، شخصیت کیفری همواره با مفاهیمی همچون آگاهی، اختیار، ادراک و قابلیت انتساب رفتار پیوند خورده است. اگرچه سامانه‌های هوش مصنوعی در بسیاری از موارد قادر به انتخاب میان گزینه‌های مختلف و حتی اصلاح رفتار خود هستند، اما این تصمیم‌ها محصول پردازش آماری داده‌ها و اجرای مدل‌های ریاضی است و نه حاصل اراده یا آگاهی اخلاقی. بنابراین، استقلال عملیاتی سامانه نباید با استقلال حقوقی یا شخصیت کیفری آن خلط شود؛ زیرا میان «تصمیم الگوریتمی» و «اراده انسانی» تفاوتی ماهوی وجود دارد. ازاین‌رو، اکثر نظریه‌های حقوق کیفری همچنان انسان را تنها فاعل واقعی جرم می‌دانند و هوش مصنوعی را صرفاً در جایگاه ابزاری پیچیده و هوشمند تحلیل می‌کنند (عظیمی و اسماعیلی، ۱۴۰۰، صص. ۹۸-۱۰۱).

در حقوق ایران نیز مبانی قانون مجازات اسلامی بر مسئولیت اشخاص حقیقی و در موارد استثنایی اشخاص حقوقی استوار است و هیچ جایگاهی برای شناسایی شخصیت کیفری سامانه‌های هوش مصنوعی پیش‌بینی نشده است. اصل شخصی بودن مسئولیت کیفری، اصل قانونی بودن جرایم و مجازات‌ها و لزوم احراز عنصر روانی، همگی بیانگر آن‌اند که مرتکب جرم باید دارای ادراک و اراده باشد؛ شرایطی که درباره سامانه‌های هوش مصنوعی قابل تصور نیست. در نتیجه، هرگاه یک سامانه هوشمند

---

6. Kowert

در ارتکاب جرم سایبری نقش داشته باشد، مسئولیت کیفری همچنان متوجه اشخاصی خواهد بود که سامانه را طراحی، آموزش، کنترل یا مورد بهره‌برداری قرار داده‌اند. به بیان دیگر، حقوق ایران هنوز از نظریه «ابزار بودن هوش مصنوعی» فاصله نگرفته است، هرچند افزایش سطح استقلال الگوریتمی موجب شده است که تشخیص شخص مسئول در برخی موارد با دشواری‌های عملی روبه‌رو شود (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۷-۱۰؛ فیروزی‌نیا، ۱۴۰۱، صص. ۶-۹). در ایالات متحده آمریکا نیز با وجود توسعه گسترده سامانه‌های هوشمند، هیچ مبنای قانونی برای شناسایی هوش مصنوعی به‌عنوان فاعل جرم وجود ندارد. بااین‌حال، ادبیات حقوقی آمریکا بیش از سایر نظام‌های حقوقی بر مسئله قابلیت پیش‌بینی رفتار سامانه‌های هوشمند تمرکز کرده است. به اعتقاد کوورت (۲۰۱۷)، هرچه رفتار سامانه برای توسعه‌دهنده یا بهره‌بردار قابل پیش‌بینی‌تر باشد، امکان انتساب نتایج زیان‌بار به وی نیز بیشتر خواهد بود. در مقابل، هرچه میزان خودآموزی و پیچیدگی الگوریتم افزایش یابد، اثبات رابطه میان رفتار انسان و نتیجه مجرمانه دشوارتر می‌شود؛ اما این دشواری به معنای انتقال شخصیت کیفری به هوش مصنوعی نیست، بلکه صرفاً موجب پیچیده‌تر شدن ارزیابی تقصیر و رابطه سببیت می‌شود (کوورت، ۲۰۱۷؛ صص ۱۸۸-۱۹۳).

رویکرد اتحادیه اروپا تا حدودی متفاوت است. اسناد جدید اروپایی، به‌ویژه قانون هوش مصنوعی، اگرچه برای سامانه‌های هوشمند شخصیت حقوقی یا کیفری مستقلی قائل نیستند، اما برخلاف نگرش سنتی، آن‌ها را صرفاً ابزارهای خنثی نیز تلقی نمی‌کنند. قانون‌گذار اروپایی پذیرفته است که برخی سامانه‌های هوشمند، به‌ویژه سامانه‌های پرخطر، دارای سطحی از استقلال عملکردی هستند که می‌تواند مخاطرات جدیدی برای امنیت عمومی ایجاد کند. ازاین‌رو، به جای تغییر مفهوم فاعل جرم، مسئولیت اشخاص انسانی را بر مبنای میزان کنترل مؤثر، مدیریت ریسک و رعایت الزامات قانونی توسعه داده است. بدین ترتیب، اتحادیه اروپا تلاش کرده است بدون عدول از اصول سنتی حقوق کیفری، آثار ناشی از استقلال عملکردی هوش مصنوعی را از طریق ایجاد تعهدات جدید برای اشخاص مرتبط مدیریت کند (کاظمی، ۱۴۰۴، صص. ۴-۶). یکی از مهم‌ترین دلایل استمرار رویکرد «ابزار بودن هوش مصنوعی»، فقدان مسئولیت اخلاقی در سامانه‌های هوشمند است. مسئولیت کیفری صرفاً ناظر بر وقوع نتیجه زیان‌بار نیست، بلکه مستلزم آن است که مرتکب بتواند ارزش اخلاقی رفتار خود را درک کرده و میان رفتار مجاز و غیرمجاز تمایز قائل شود. سامانه‌های هوش مصنوعی، حتی پیشرفته‌ترین انواع آن‌ها، فاقد تجربه آگاهانه، احساس مسئولیت، وجدان اخلاقی و توانایی درک هنجارهای اجتماعی هستند. بنابراین، هرچند ممکن است خروجی آن‌ها آثار مجرمانه‌ای ایجاد کند، اما این آثار از منظر حقوق کیفری قابل انتساب به خود سامانه نیست و مسئولیت همچنان متوجه اشخاصی خواهد بود که زمینه بهره‌برداری از آن را فراهم کرده‌اند (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۱۰-۱۲). با این حال، افزایش نقش هوش مصنوعی در امنیت سایبری، مرز میان ابزار و فاعل را تا حدی مبهم ساخته است. امروزه سامانه‌های مبتنی بر یادگیری ماشین می‌توانند آسیب‌پذیری‌های امنیتی را شناسایی، بدافزارهای جدید تولید، حملات فیشینگ را شخصی‌سازی و حتی در برابر سامانه‌های دفاعی راهبردهای خود را تغییر دهند. پژوهش‌های انجام‌شده در حوزه امنیت سایبری نشان می‌دهد که این قابلیت‌ها موجب شده است فرآیند ارتکاب جرم بیش از گذشته از کنترل لحظه‌ای انسان خارج شود و نقش سامانه در تحقق نتیجه

مجرمانه افزایش یابد. بااین‌حال، همین پژوهش‌ها تأکید می‌کنند که این استقلال عملکردی نباید با استقلال حقوقی اشتباه گرفته شود؛ زیرا تمامی الگوریتم‌ها در نهایت در چارچوب اهداف، داده‌ها و محدودیت‌هایی عمل می‌کنند که توسط انسان تعیین شده است. نمونه بارز این وضعیت در حملات سایبری مبتنی بر تزریق داده‌های جعلی[7] و حملات اخلالگر مشاهده می‌شود. در این حملات، سامانه‌های هوشمند قادرند به صورت خودکار اطلاعات شبکه را تحلیل کرده، بهترین زمان حمله را انتخاب و حتی راهبردهای نفوذ خود را متناسب با واکنش سامانه دفاعی اصلاح کنند. اگرچه اجرای عملیات به صورت مستقل انجام می‌شود، اما طراحی الگوریتم، آموزش مدل و تعیین اهداف حمله توسط انسان صورت گرفته است؛ ازاین‌رو، مسئولیت کیفری نیز باید متناسب با نقش اشخاص در ایجاد این فرآیند ارزیابی شود، نه بر مبنای فرض وجود شخصیت مستقل برای سامانه (گوان و جی، ۲۰۱۷؛ صص ۵۰-۵۵).

از سوی دیگر، توسعه کاربردهای هوش مصنوعی در بانکداری الکترونیکی، دستگاه‌های خودپرداز و سامانه‌های مالی نیز ابعاد جدیدی از این چالش را آشکار ساخته است. پژوهش آدکونله[8] (۲۰۱۹) نشان می‌دهد که هرچند سامانه‌های هوشمند قادرند بسیاری از فرآیندهای تصمیم‌گیری مالی را به صورت خودکار انجام دهند، اما تمامی این تصمیم‌ها در چارچوب سیاست‌ها، محدودیت‌ها و الگوریتم‌هایی اتخاذ می‌شود که توسط انسان طراحی شده‌اند. بنابراین، حتی در پیچیده‌ترین سامانه‌های خودکار نیز نمی‌توان نقش انسان را به‌طور کامل حذف کرد و این امر، استمرار مسئولیت کیفری اشخاص انسانی را توجیه می‌کند (آدکونله، ۲۰۱۹؛ ص ۱۹). در نتیجه، تحلیل تطبیقی حقوق ایران، آمریکا و اتحادیه اروپا نشان می‌دهد که اگرچه میزان توجه این نظام‌ها به استقلال عملکردی هوش مصنوعی متفاوت است، اما هیچ‌یک تاکنون آن را به‌عنوان فاعل مستقل جرم به رسمیت نشناخته‌اند. تفاوت اصلی میان این نظام‌ها در نحوه پاسخ به آثار ناشی از استقلال الگوریتمی است؛ به‌گونه‌ای که حقوق ایران همچنان بر قواعد سنتی انتساب و مسئولیت مبتنی است، حقوق آمریکا بر معیار قابلیت پیش‌بینی رفتار سامانه تأکید می‌کند و اتحادیه اروپا با توسعه مفهوم کنترل مؤثر، مدیریت ریسک و تعهدات پیشگیرانه، می‌کوشد شکاف موجود میان پیشرفت فناوری و قواعد کلاسیک حقوق کیفری را کاهش دهد. این تفاوت‌ها نشان می‌دهد که مسئله اصلی دیگر «فاعل بودن یا نبودن هوش مصنوعی» نیست، بلکه یافتن مبنایی منسجم برای انتساب نتایج رفتار سامانه‌های هوشمند به اشخاصی است که بیشترین نقش را در ایجاد، کنترل و بهره‌برداری از آن‌ها دارند؛ موضوعی که در بخش بعدی با عنوان «استقلال الگوریتمی و انتساب رفتار مجرمانه» به‌طور تفصیلی بررسی خواهد شد.

استقلال الگوریتمی و انتساب رفتار مجرمانه

یکی از مهم‌ترین پیامدهای گسترش هوش مصنوعی در حقوق کیفری، ظهور مفهوم «استقلال الگوریتمی» و تأثیر آن بر انتساب رفتار مجرمانه است. در حقوق کیفری سنتی، انتساب جرم مبتنی بر وجود رابطه‌ای مستقیم میان رفتار مرتکب و نتیجه مجرمانه است؛ به این معنا که شخص با انجام فعل

---

7. False Data Injection
8. Adekunle

یا ترک فعل، نتیجه زیان‌بار را ایجاد کرده و ازاین‌رو مسئول شناخته می‌شود. اما در سامانه‌های هوش مصنوعی، به‌ویژه سامانه‌های مبتنی بر یادگیری ماشین و یادگیری عمیق، فرآیند تصمیم‌گیری تنها به اجرای دستورات از پیش تعیین‌شده محدود نیست، بلکه سامانه با تحلیل داده‌های جدید، اصلاح پارامترهای داخلی و انطباق مستمر با محیط، رفتارهایی از خود بروز می‌دهد که در بسیاری از موارد از پیش توسط برنامه‌نویس قابل پیش‌بینی دقیق نبوده است. همین ویژگی، مسئله انتساب رفتار مجرمانه را به یکی از پیچیده‌ترین مباحث حقوق کیفری فناوری تبدیل کرده است؛ زیرا مشخص نیست در صورت وقوع جرم، مسئولیت باید متوجه طراح الگوریتم، تولیدکننده، دارنده، بهره‌بردار یا استفاده‌کننده نهایی باشد (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۸-۱۰). استقلال الگوریتمی به معنای استقلال حقوقی نیست، بلکه بیانگر توانایی سامانه در اتخاذ تصمیم یا تولید خروجی بر پایه تحلیل داده‌ها، بدون دخالت مستقیم انسان در هر مرحله از فرآیند اجرا است. این استقلال عمدتاً نتیجه استفاده از الگوریتم‌های یادگیری ماشین است که امکان اصلاح مدل تصمیم‌گیری را در طول زمان فراهم می‌کنند. با وجود این، استقلال عملکردی سامانه همواره در چارچوب داده‌های آموزشی، اهداف طراحی‌شده و محدودیت‌هایی تحقق می‌یابد که توسط انسان تعیین شده‌اند. ازاین‌رو، بسیاری از پژوهشگران معتقدند که هرچند استقلال الگوریتمی موجب افزایش فاصله میان رفتار انسان و نتیجه نهایی می‌شود، اما این فاصله به‌تنهایی برای قطع رابطه انتساب کفایت نمی‌کند؛ زیرا منشأ اصلی شکل‌گیری رفتار سامانه همچنان تصمیم‌های انسانی در مرحله طراحی، آموزش و بهره‌برداری است (کشاورز و حسینی[9]، ۲۰۲۳؛ صص ۵-۸).

در حقوق ایران، قواعد حاکم بر انتساب رفتار مجرمانه همچنان بر مبنای رابطه سببیت و نقش مؤثر مرتکب در تحقق نتیجه تنظیم شده است. هرچند قانون مجازات اسلامی مقررات ویژه‌ای درباره سامانه‌های هوش مصنوعی ندارد، اما قواعد عمومی مسئولیت کیفری ایجاب می‌کند که میان رفتار شخص و نتیجه مجرمانه رابطه‌ای قابل احراز وجود داشته باشد. بنابراین، اگر دارنده یا بهره‌بردار سامانه هوش مصنوعی با آگاهی از مخاطرات آن، زمینه ارتکاب جرم سایبری را فراهم کند یا از انجام اقدامات متعارف برای کنترل سامانه خودداری نماید، استقلال الگوریتمی مانع انتساب نتیجه به وی نخواهد بود. در مقابل، هرگاه نتیجه صرفاً ناشی از نقص غیرقابل پیش‌بینی سامانه باشد و امکان پیشگیری متعارف نیز وجود نداشته باشد، احراز مسئولیت کیفری با دشواری بیشتری همراه خواهد شد. این امر نشان می‌دهد که قواعد سنتی حقوق ایران، اگرچه قابلیت اعمال بر سامانه‌های هوشمند را دارند، اما در تبیین مرزهای مسئولیت در برابر الگوریتم‌های خودآموز با محدودیت‌هایی مواجه هستند (فیروزی‌نیا، ۱۴۰۱، صص. ۸-۱۱). در نظام حقوقی آمریکا، تحلیل استقلال الگوریتمی بیش از هر چیز بر معیار «قابلیت پیش‌بینی» استوار است. بر اساس این رویکرد، پرسش اصلی آن نیست که سامانه تا چه اندازه مستقل عمل کرده، بلکه این است که آیا رفتار زیان‌بار آن برای شخص مسئول قابل پیش‌بینی بوده است یا خیر. کوورت (۲۰۱۷) بر این باور است که هرچه توسعه‌دهنده یا دارنده سامانه بتواند احتمال وقوع رفتارهای مخاطره‌آمیز را پیش‌بینی کند، مسئولیت وی نیز افزایش می‌یابد؛ زیرا استقلال سامانه، تعهد اشخاص به ارزیابی ریسک و مدیریت مخاطرات را از بین نمی‌برد، بلکه اهمیت آن را دوچندان می‌کند.

---

[9]. Keshavarz & Hosseini

ازاین‌رو، حقوق آمریکا استقلال الگوریتمی را عاملی برای تغییر معیارهای ارزیابی تقصیر می‌داند، نه دلیلی برای نفی مسئولیت انسان (کوورت، ۲۰۱۷؛ صص ۱۸۹-۱۹۳).

در اتحادیه اروپا نیز استقلال الگوریتمی به‌عنوان واقعیتی فناورانه پذیرفته شده است، اما قانون‌گذار اروپایی تلاش کرده است آثار حقوقی آن را از طریق نظام مدیریت ریسک کنترل کند. در این رویکرد، هرچه سامانه از استقلال بیشتری برخوردار باشد و احتمال ورود خسارت ناشی از عملکرد آن افزایش یابد، تکالیف اشخاص مرتبط نیز گسترده‌تر خواهد شد. الزام به ارزیابی مستمر ریسک، مستندسازی فرآیند تصمیم‌گیری، نظارت انسانی مؤثر و ثبت عملکرد سامانه، همگی با هدف کاهش فاصله میان رفتار الگوریتم و مسئولیت انسان پیش‌بینی شده‌اند. بنابراین، اتحادیه اروپا نیز استقلال الگوریتمی را موجب انتقال شخصیت کیفری به هوش مصنوعی نمی‌داند، بلکه آن را مبنایی برای تشدید تعهدات اشخاص مسئول در چرخه طراحی و بهره‌برداری تلقی می‌کند (کاظمی، ۱۴۰۴، صص. ۵–۸). پیچیدگی مسئله زمانی افزایش می‌یابد که سامانه هوش مصنوعی در بستر جرایم سایبری به صورت خودکار عمل کند. امروزه الگوریتم‌های هوشمند قادرند بدون صدور دستور جدید از سوی کاربر، آسیب‌پذیری‌های شبکه را شناسایی، اهداف مناسب را انتخاب، بدافزارها را متناسب با شرایط اصلاح و حتی راهبرد حمله را در جریان عملیات تغییر دهند. در چنین وضعیتی، فاصله زمانی و عملی میان رفتار اولیه انسان و نتیجه مجرمانه افزایش می‌یابد و این پرسش مطرح می‌شود که آیا همچنان می‌توان نتیجه را به شخص انسانی منتسب کرد یا خیر. پاسخ غالب در حقوق تطبیقی آن است که تا زمانی که استقلال سامانه در چارچوب اهداف و محدودیت‌های تعیین‌شده توسط انسان شکل گرفته باشد، رابطه انتساب از بین نمی‌رود؛ بلکه صرفاً اثبات آن نیازمند تحلیل دقیق‌تر رابطه سببیت و میزان کنترل اشخاص خواهد بود.

نمونه روشن این مسئله در حملات مبتنی بر تزریق داده‌های جعلی مشاهده می‌شود. پژوهش گوان و جی (۲۰۱۷) نشان می‌دهد که سامانه‌های هوشمند می‌توانند با تحلیل مداوم داده‌های شبکه، زمان مناسب برای حمله را انتخاب کرده و در برابر اقدامات دفاعی، راهبرد خود را تغییر دهند. اگرچه اجرای این فرآیند بدون دخالت مستقیم انسان انجام می‌شود، اما طراحی معماری الگوریتم، تعیین اهداف، آموزش مدل و فعال‌سازی سامانه همگی توسط انسان صورت گرفته است. بنابراین، استقلال الگوریتمی موجب حذف نقش انسان نمی‌شود، بلکه زنجیره مسئولیت را میان اشخاص مختلف توزیع می‌کند و دادگاه را ناگزیر می‌سازد سهم هر یک از آنان را در ایجاد نتیجه مجرمانه ارزیابی کند (گوان و جی، ۲۰۱۷؛ صص ۵۱-۵۶). در حوزه بانکداری الکترونیکی نیز توسعه سامانه‌های هوشمند، چالش مشابهی ایجاد کرده است. پژوهش آدکونله (۲۰۱۹) نشان می‌دهد که الگوریتم‌های مورد استفاده در دستگاه‌های خودپرداز و سامانه‌های بانکداری اینترنتی قادرند بسیاری از تصمیم‌های امنیتی را به صورت خودکار اتخاذ کنند؛ بااین‌حال، تمامی این تصمیم‌ها در چارچوب سیاست‌های تعیین‌شده توسط انسان اجرا می‌شوند. ازاین‌رو، در صورت وقوع جرایم سایبری، استقلال الگوریتمی نمی‌تواند مبنایی برای سلب مسئولیت از مؤسسات مالی یا مدیران سامانه باشد، بلکه آنان مکلف‌اند با ارزیابی مستمر ریسک، به‌روزرسانی الگوریتم‌ها و نظارت مؤثر، از وقوع رفتارهای زیان‌بار جلوگیری کنند (آدکونله، ۲۰۱۹؛ ص ۱۹). از منظر حقوق کیفری، مهم‌ترین پیامد استقلال الگوریتمی، تغییر در مفهوم رابطه سببیت است. در جرایم سنتی، زنجیره سببیت غالباً کوتاه و روشن است؛ اما در جرایم مبتنی بر هوش مصنوعی، میان

رفتار اولیه انسان و نتیجه مجرمانه ممکن است فرآیندهای پیچیده‌ای از یادگیری، پردازش داده و تصمیم‌گیری خودکار قرار گیرد. همین امر موجب شده است که برخی نویسندگان از «سببیت چندلایه» سخن بگویند؛ بدین معنا که تحقق نتیجه مجرمانه حاصل مجموعه‌ای از عوامل انسانی، نرم‌افزاری و محیطی است و نمی‌توان آن را صرفاً به عملکرد سامانه یا صرفاً به رفتار انسان نسبت داد. در چنین شرایطی، معیارهایی مانند کنترل مؤثر، قابلیت پیش‌بینی، میزان مداخله در چرخه عملکرد سامانه و رعایت استانداردهای فنی و قانونی اهمیت بیشتری نسبت به معیارهای سنتی پیدا می‌کنند (کاظمی، ۱۴۰۴، صص. ۸-۱۰؛ صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۱۳-۱۵).

اراده، سوءنیت و قابلیت سرزنش

یکی از اساسی‌ترین مبانی مسئولیت کیفری، وجود اراده، سوءنیت و قابلیت سرزنش مرتکب است. حقوق کیفری کلاسیک بر این اصل استوار است که تنها شخصی قابل مجازات است که رفتار مجرمانه را با آگاهی، اختیار و قابلیت درک نتایج عمل خود انجام داده باشد. به همین دلیل، عنصر روانی جرم همواره در کنار عنصر مادی، رکن اساسی تحقق مسئولیت کیفری محسوب می‌شود. با ظهور سامانه‌های هوش مصنوعی، این مبنا با چالشی بنیادین روبه‌رو شده است؛ زیرا هرچند سامانه‌های هوشمند قادرند تصمیم‌هایی مستقل از دخالت لحظه‌ای انسان اتخاذ کنند، اما این تصمیم‌ها فاقد آگاهی، قصد اخلاقی و اراده به مفهوم حقوقی هستند. ازاین‌رو، پرسش اصلی آن است که آیا می‌توان تصمیم الگوریتمی را معادل اراده کیفری دانست یا همچنان باید عنصر روانی را صرفاً به اشخاص انسانی منتسب کرد؟ پاسخ به این پرسش، نقش تعیین‌کننده‌ای در تبیین مسئولیت کیفری دارنده هوش مصنوعی در جرایم سایبری دارد (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۱۲-۱۴). از منظر حقوق کیفری، اراده مفهومی فراتر از انتخاب میان گزینه‌های مختلف است. اراده کیفری مستلزم آگاهی نسبت به ماهیت رفتار، امکان انتخاب آزادانه و توانایی تشخیص مشروع یا نامشروع بودن عمل است. سامانه‌های هوش مصنوعی اگرچه قادر به تحلیل میلیون‌ها داده و انتخاب مناسب‌ترین پاسخ بر اساس الگوریتم‌های آماری هستند، اما فاقد شعور، وجدان اخلاقی و ادراک هنجارهای اجتماعی‌اند. بنابراین، انتخاب الگوریتمی را نمی‌توان با اراده انسانی یکسان دانست. در واقع، تصمیم‌های سامانه نتیجه محاسبات ریاضی و مدل‌های یادگیری است و نه حاصل قصد مجرمانه یا آگاهی نسبت به ارزش‌های حقوقی. همین تفاوت بنیادین سبب شده است که تقریباً تمامی نظام‌های حقوقی، عنصر روانی جرم را همچنان به انسان اختصاص دهند و برای هوش مصنوعی شخصیت کیفری مستقلی قائل نشوند (عظیمی و اسماعیلی، ۱۴۰۰، صص. ۹۹–۱۰۲). در حقوق ایران، قانون مجازات اسلامی تحقق مسئولیت کیفری را منوط به وجود عمد، قصد یا در مواردی تقصیر مرتکب می‌داند. این عناصر مستلزم آن است که شخص از ماهیت رفتار خود آگاه باشد و بتواند نتایج آن را پیش‌بینی کند. از آنجا که سامانه‌های هوش مصنوعی فاقد چنین ویژگی‌هایی هستند، امکان انتساب سوءنیت به آن‌ها وجود ندارد. بنابراین، حتی اگر یک سامانه هوشمند به صورت مستقل حمله سایبری را اجرا یا محتوای مجرمانه تولید کند، سوءنیت قابل انتساب به اشخاصی خواهد بود که با طراحی، آموزش، راه‌اندازی یا بهره‌برداری از سامانه، زمینه تحقق رفتار مجرمانه را فراهم کرده‌اند. این تحلیل با اصل شخصی بودن مسئولیت کیفری نیز هماهنگ است؛ زیرا

مسئولیت بر پایه رفتار و وضعیت روانی انسان استوار می‌شود، نه عملکرد مستقل فناوری (فیروزی‌نیا، ۱۴۰۱، صص. ۹–۱۲). در نظام حقوقی آمریکا نیز عنصر روانی همچنان محور اصلی مسئولیت کیفری محسوب می‌شود، اما در مواجهه با سامانه‌های هوشمند، مفهوم سوءنیت تا حدی با معیار «قابلیت پیش‌بینی» تکمیل شده است. به بیان دیگر، در بسیاری از موارد، دادگاه‌ها علاوه بر بررسی قصد مستقیم، ارزیابی می‌کنند که آیا توسعه‌دهنده یا دارنده سامانه می‌توانسته احتمال سوءاستفاده یا بروز رفتار مجرمانه را پیش‌بینی کند یا خیر. کوورت (۲۰۱۷) معتقد است که هرچه رفتار سامانه برای اشخاص مسئول قابل پیش‌بینی‌تر باشد، عدم اتخاذ تدابیر پیشگیرانه می‌تواند مبنای سرزنش کیفری آنان قرار گیرد. بنابراین، حقوق آمریکا اگرچه سوءنیت را همچنان ویژگی انسان می‌داند، اما دامنه ارزیابی عنصر روانی را به حوزه مدیریت خطر و پیش‌بینی رفتار سامانه نیز گسترش داده است (کوورت، ۲۰۱۷؛ صص ۱۹۰-۱۹۴).

اتحادیه اروپا نیز رویکرد مشابهی اتخاذ کرده، اما بیش از حقوق آمریکا بر مفهوم «حکمرانی مسئولانه» و «وظیفه مراقبت» تأکید دارد. در مقررات اروپایی، هرچه سطح خطر سامانه هوش مصنوعی افزایش یابد، اشخاص مسئول مکلف‌اند ارزیابی ریسک، نظارت مستمر و کنترل انسانی مؤثر را در تمام مراحل چرخه حیات سامانه اعمال کنند. در این چارچوب، مسئولیت کیفری یا شبه‌کیفری ناشی از عملکرد سامانه، بیش از آنکه بر قصد مجرمانه مستقیم مبتنی باشد، بر نقض تعهدات قانونی و قصور در مدیریت مخاطرات استوار می‌شود. این رویکرد نشان می‌دهد که اتحادیه اروپا تلاش کرده است بدون تغییر مفهوم کلاسیک سوءنیت، معیارهای جدیدی برای ارزیابی قابلیت سرزنش اشخاص در عصر هوش مصنوعی ایجاد کند (کاظمی، ۱۴۰۴، صص. ۷–۱۰). موضوع قابلیت سرزنش[10] نیز در تحلیل مسئولیت کیفری دارنده هوش مصنوعی اهمیت ویژه‌ای دارد. قابلیت سرزنش زمانی محقق می‌شود که شخص امکان جلوگیری از وقوع نتیجه زیان‌بار را داشته، اما با وجود این، از انجام اقدامات لازم خودداری کرده باشد. در جرایم سایبری مبتنی بر هوش مصنوعی، ممکن است دارنده سامانه هیچ قصد مستقیمی برای ارتکاب جرم نداشته باشد، اما با بی‌توجهی به به‌روزرسانی سامانه، استفاده از داده‌های آموزشی نامناسب یا فقدان نظارت کافی، زمینه سوءاستفاده مجرمانه را فراهم کرده باشد. در چنین مواردی، سرزنش کیفری نه بر پایه قصد ارتکاب جرم، بلکه بر مبنای نقض وظیفه مراقبت و بی‌احتیاطی در مدیریت سامانه شکل می‌گیرد (صوفی و صالح‌نژاد بهرستاقی، ۱۴۰۲، صص. ۱۴-۱۶).

اهمیت این موضوع در حوزه امنیت سایبری بیش از سایر حوزه‌هاست. پژوهش‌های انجام‌شده نشان می‌دهد که سامانه‌های مبتنی بر هوش مصنوعی قادرند حملات سایبری را به صورت خودکار طراحی و اجرا کنند و حتی در برابر اقدامات دفاعی، رفتار خود را اصلاح نمایند. اگر دارنده سامانه از چنین قابلیت‌هایی آگاه باشد و با وجود احتمال معقول سوءاستفاده، هیچ‌گونه محدودیت یا نظارتی بر عملکرد سامانه اعمال نکند، قابلیت سرزنش وی افزایش خواهد یافت. در مقابل، اگر تمامی استانداردهای فنی و امنیتی رعایت شده باشد و رفتار مجرمانه ناشی از شرایطی کاملاً غیرقابل پیش‌بینی باشد، انتساب تقصیر با دشواری بیشتری همراه خواهد بود .نمونه عملی این وضعیت در حملات مبتنی بر هوش مصنوعی علیه زیرساخت‌های حیاتی مشاهده می‌شود. پژوهش گوان و جی (۲۰۱۷) نشان می‌دهد که

---

[10]. Blameworthiness

الگوریتم‌های هوشمند می‌توانند بدون مداخله مستمر انسان، مسیر حمله را تغییر داده و متناسب با واکنش سامانه دفاعی تصمیم‌های جدید اتخاذ کنند. با وجود این، چنین تصمیم‌هایی در چارچوب اهداف و پارامترهایی انجام می‌شود که توسط انسان طراحی شده‌اند. بنابراین، قابلیت سرزنش همچنان متوجه اشخاصی است که سامانه را طراحی یا فعال کرده‌اند و نه خود الگوریتم؛ زیرا سامانه فاقد درک ارزش‌های حقوقی و اخلاقی است و تنها بر اساس داده‌های دریافتی عمل می‌کند (گوان. جی، ۲۰۱۷؛ صص ۵۰-۵۴). از منظر حقوق تطبیقی، تفاوت اصلی میان ایران، آمریکا و اتحادیه اروپا در اصل ضرورت وجود عنصر روانی نیست، بلکه در شیوه اثبات و ارزیابی آن است. حقوق ایران همچنان بر مفاهیم سنتی عمد و تقصیر تأکید دارد و قواعد خاصی برای ارزیابی رفتار سامانه‌های هوشمند پیش‌بینی نکرده است. در مقابل، حقوق آمریکا مفهوم قابلیت پیش‌بینی را به‌عنوان معیاری برای توسعه قلمرو سرزنش کیفری مورد استفاده قرار می‌دهد و اتحادیه اروپا نیز با تأکید بر مدیریت ریسک، نظارت انسانی و رعایت الزامات قانونی، دامنه مسئولیت اشخاص مرتبط با سامانه‌های هوشمند را توسعه داده است. وجه مشترک هر سه نظام آن است که اراده، سوءنیت و قابلیت سرزنش همچنان به انسان تعلق دارد و هوش مصنوعی، علی‌رغم برخورداری از استقلال عملکردی، فاقد شرایط لازم برای تحمل مسئولیت کیفری مستقل است. بر این اساس، می‌توان نتیجه گرفت که چالش اصلی حقوق کیفری در عصر هوش مصنوعی، انتقال عنصر روانی به سامانه‌های هوشمند نیست، بلکه بازتعریف معیارهای انتساب سوءنیت و قابلیت سرزنش نسبت به اشخاصی است که در چرخه طراحی، آموزش، استقرار و بهره‌برداری از سامانه نقش دارند .به همین دلیل، در تحلیل مسئولیت کیفری دارنده هوش مصنوعی، معیارهایی مانند کنترل مؤثر، قابلیت پیش‌بینی خطر، میزان نظارت انسانی، رعایت استانداردهای امنیت سایبری و انجام تکالیف قانونی اهمیت بیشتری از تمرکز صرف بر عملکرد مستقل الگوریتم پیدا می‌کنند. این مبنا، چارچوب مشترک نظام‌های حقوقی ایران، آمریکا و اتحادیه اروپا را تشکیل می‌دهد و زمینه را برای ارائه یک جمع‌بندی تطبیقی و استخراج الگوی مناسب برای حقوق ایران فراهم می‌سازد.

جمع‌بندی تطبیقی

بررسی تطبیقی حقوق ایران، آمریکا و اتحادیه اروپا نشان می‌دهد که اگرچه این سه نظام حقوقی در مواجهه با هوش مصنوعی از رویکردهای تقنینی متفاوتی بهره می‌گیرند، اما در یک اصل بنیادین اشتراک دارند و آن اینکه هوش مصنوعی هنوز به‌عنوان فاعل مستقل جرم یا دارنده شخصیت کیفری شناخته نمی‌شود .در هر سه نظام، مسئولیت کیفری همچنان بر مبنای رفتار و وضعیت روانی اشخاص انسانی استوار است و سامانه هوش مصنوعی، علی‌رغم برخورداری از استقلال عملکردی، فاقد اراده، شعور اخلاقی و قابلیت تحمل مسئولیت کیفری تلقی می‌شود. اختلاف میان این نظام‌ها نه در پذیرش یا رد شخصیت کیفری هوش مصنوعی، بلکه در نحوه تحلیل آثار ناشی از استقلال الگوریتمی، معیارهای انتساب رفتار مجرمانه و حدود مسئولیت اشخاص مرتبط با سامانه‌های هوشمند است. در حقوق ایران، قواعد حاکم بر مسئولیت کیفری همچنان بر اصول سنتی حقوق جزا، از جمله اصل شخصی بودن مسئولیت، وجود عنصر روانی، رابطه سببیت و قابلیت انتساب رفتار استوار است. این قواعد اگرچه امکان انتساب رفتار مجرمانه ناشی از عملکرد هوش مصنوعی را به اشخاصی مانند دارنده، بهره‌بردار یا

برنامه‌نویس فراهم می‌کنند، اما برای مواجهه با پیچیدگی‌های ناشی از سامانه‌های خودآموز و الگوریتم‌های مستقل، پاسخ‌های صریح و اختصاصی ارائه نمی‌دهند. در نتیجه، حقوق ایران هنوز فاقد چارچوب تقنینی ویژه‌ای است که بتواند حدود مسئولیت کیفری اشخاص مختلف در چرخه طراحی، آموزش، استقرار و بهره‌برداری از سامانه‌های هوش مصنوعی را به‌طور دقیق مشخص کند. در مقابل، حقوق آمریکا اگرچه شخصیت کیفری مستقلی برای هوش مصنوعی قائل نیست، اما از طریق توسعه معیار قابلیت پیش‌بینی رفتار سامانه[11] تلاش کرده است خلأهای ناشی از استقلال الگوریتمی را جبران کند. در این رویکرد، هرچه رفتار مخاطره‌آمیز سامانه برای توسعه‌دهنده، دارنده یا بهره‌بردار قابل پیش‌بینی‌تر باشد، مسئولیت کیفری آنان نیز تقویت می‌شود. این نگرش نشان می‌دهد که استقلال الگوریتمی، رابطه انتساب را از بین نمی‌برد، بلکه ارزیابی تقصیر و حدود مسئولیت اشخاص را پیچیده‌تر می‌کند و دادگاه را به بررسی دقیق میزان آگاهی، کنترل و امکان پیشگیری از وقوع جرم ملزم می‌سازد. در اتحادیه اروپا نیز قانون‌گذار رویکردی پیشگیرانه و مبتنی بر مدیریت ریسک اتخاذ کرده است. مقررات اروپایی، به‌ویژه قانون هوش مصنوعی، به جای تمرکز بر شخصیت حقوقی سامانه، مسئولیت را بر پایه کنترل مؤثر، نظارت انسانی، ارزیابی مستمر ریسک و رعایت الزامات قانونی سامان‌دهی کرده‌اند. در این نظام، هرچه سامانه از استقلال بیشتری برخوردار باشد یا خطرات بیشتری برای امنیت سایبری ایجاد کند، تکالیف قانونی اشخاص مرتبط نیز افزایش می‌یابد. بدین ترتیب، اتحادیه اروپا تلاش کرده است بدون عدول از اصول کلاسیک حقوق کیفری، از طریق توسعه تعهدات پیشگیرانه و الزامات فنی، فاصله میان پیشرفت فناوری و قواعد سنتی مسئولیت را کاهش دهد (کاظمی، ۱۴۰۴، صص. ۶-۱۰).

مقایسه این سه رویکرد نشان می‌دهد که هرچند مبانی فلسفی مسئولیت کیفری همچنان بر محور انسان قرار دارد، اما معیارهای تحلیل مسئولیت در حال تحول است. در حقوق سنتی، تمرکز اصلی بر رفتار مستقیم مرتکب و سوءنیت وی بود؛ اما در عصر هوش مصنوعی، مفاهیمی مانند کنترل مؤثر، قابلیت پیش‌بینی، مدیریت ریسک، نظارت مستمر، کیفیت داده‌های آموزشی و رعایت استانداردهای امنیت سایبری به عناصر مکمل تحلیل مسئولیت تبدیل شده‌اند. این تحول، بیانگر گذار تدریجی حقوق کیفری از رویکرد صرفاً واکنشی به رویکردی پیشگیرانه و فناورانه است که در آن، مسئولیت نه تنها بر مبنای وقوع جرم، بلکه بر پایه میزان ایفای تعهدات قانونی در پیشگیری از وقوع آن نیز ارزیابی می‌شود. از منظر موضوع این پژوهش، مهم‌ترین نتیجه تطبیقی آن است که دارنده هوش مصنوعی در هیچ‌یک از نظام‌های حقوقی مورد مطالعه صرفاً به دلیل مالکیت سامانه مسئول شناخته نمی‌شود، بلکه مسئولیت وی تابع عواملی همچون میزان کنترل بر سامانه، قابلیت پیش‌بینی مخاطرات، نقش در آموزش و استقرار الگوریتم، رعایت استانداردهای امنیتی و انجام تکالیف نظارتی است. بنابراین، مبنای مسئولیت کیفری دارنده هوش مصنوعی باید از مفهوم سنتی مالکیت فاصله گرفته و بر معیارهای کارکردی و مبتنی بر مدیریت خطر استوار شود. این نتیجه می‌تواند مبنای مناسبی برای اصلاح مقررات کیفری ایران نیز باشد؛ زیرا چارچوب کنونی، پاسخگوی تمامی چالش‌های ناشی از استقلال الگوریتمی و جرایم سایبری مبتنی بر هوش مصنوعی نیست .نتیجه نهایی این بخش آن است که اختلاف اصلی میان ایران، آمریکا و اتحادیه اروپا بر سر فاعل بودن هوش مصنوعی نیست؛ زیرا هر سه نظام این فرض را رد

---

[11]. Foreseeability

می‌کنند. اختلاف واقعی در مبنای انتساب مسئولیت کیفری به اشخاص انسانی است. حقوق ایران بر قواعد سنتی سببیت و تقصیر تکیه دارد، حقوق آمریکا معیار قابلیت پیش‌بینی را توسعه داده و اتحادیه اروپا نظامی مبتنی بر مدیریت ریسک و کنترل مؤثر ایجاد کرده است. این تفاوت‌ها نشان می‌دهد که برای مواجهه با جرایم سایبری مبتنی بر هوش مصنوعی، اصلاح قوانین کیفری ایران باید بیش از آنکه به اعطای شخصیت کیفری به هوش مصنوعی بیندیشد، بر تعیین دقیق جایگاه دارنده هوش مصنوعی، تبیین تکالیف قانونی وی، توسعه معیارهای انتساب رفتار مجرمانه و پذیرش الگوی مسئولیت مبتنی بر کنترل و پیشگیری متمرکز شود. این نتیجه، مبنای ورود به بخش بعدی مقاله، یعنی «تحلیل مبانی مسئولیت کیفری دارنده هوش مصنوعی در ایران، آمریکا و اتحادیه اروپا» خواهد بود.

مبانی مسئولیت کیفری دارنده هوش مصنوعی در نظام‌های حقوقی

با گسترش استفاده از سامانه‌های هوش مصنوعی در حوزه‌های مختلف، یکی از مهم‌ترین چالش‌های حقوق کیفری معاصر، تعیین مبنای مسئولیت اشخاصی است که در چرخه ایجاد، کنترل، بهره‌برداری یا مدیریت این سامانه‌ها نقش دارند. برخلاف ابزارهای سنتی ارتکاب جرم که معمولاً رابطه میان رفتار انسانی و نتیجه مجرمانه روشن و مستقیم است، سامانه‌های هوش مصنوعی به دلیل برخورداری از قابلیت یادگیری، پردازش داده‌های گسترده و اتخاذ تصمیم‌های خودکار، فرآیند انتساب رفتار مجرمانه را پیچیده ساخته‌اند. در این میان، مسئله اصلی آن نیست که آیا خود هوش مصنوعی باید به‌عنوان مرتکب جرم شناخته شود یا خیر، بلکه پرسش بنیادین این است که در صورت تحقق یک رفتار مجرمانه ناشی از عملکرد سامانه هوشمند، مسئولیت کیفری متوجه کدام شخص یا اشخاص خواهد بود. به همین دلیل، مفهوم «دارنده هوش مصنوعی» در حقوق کیفری اهمیت ویژه‌ای یافته است؛ زیرا دارنده ممکن است مالک، بهره‌بردار، اپراتور، ارائه‌دهنده خدمات، توسعه‌دهنده یا شخص حقوقی باشد که کنترل یا منفعت اصلی از عملکرد سامانه را در اختیار دارد. بر این اساس، مبانی مسئولیت کیفری دارنده هوش مصنوعی را نمی‌توان صرفاً بر پایه مالکیت ظاهری سامانه تحلیل کرد، بلکه باید مجموعه‌ای از معیارهای عملی همچون کنترل مؤثر، قابلیت پیش‌بینی خطر، رعایت وظیفه مراقبت و رابطه سببیت میان رفتار دارنده و نتیجه مجرمانه مورد بررسی قرار گیرد (حکمت‌نیا و همکاران، ۱۳۹۸، صص. ۲۳۱-۲۵۵؛ هالوری، ۲۰۱۵؛ صص ۴۵-۵۲).

کنترل مؤثر

یکی از مهم‌ترین معیارهای تعیین مسئولیت کیفری دارنده هوش مصنوعی، مفهوم «کنترل مؤثر» است. منظور از کنترل مؤثر، میزان توانایی شخص در هدایت، نظارت، محدودسازی یا توقف عملکرد سامانه هوشمند است. در نظام مسئولیت کیفری سنتی، شخص زمانی مسئول شناخته می‌شود که بر رفتار خود یا ابزار مورد استفاده برای ارتکاب جرم تسلط داشته باشد؛ اما در محیط هوش مصنوعی، این کنترل دیگر صرفاً به معنای صدور دستور مستقیم نیست، بلکه شامل امکان طراحی محدودیت‌های امنیتی، تنظیم سطح دسترسی، نظارت بر عملکرد الگوریتم و اتخاذ اقدامات پیشگیرانه نیز می‌شود. بنابراین، دارنده‌ای که کنترل واقعی بر نحوه فعالیت سامانه دارد، حتی اگر در لحظه وقوع جرم دستور

مستقیم صادر نکرده باشد، ممکن است به دلیل نقش تعیین‌کننده خود در ایجاد یا تداوم وضعیت خطرناک، مسئول شناخته شود (ابوت[۱۲]، ۲۰۲۰؛ صص ۸۹-۹۵). در جرایم سایبری مبتنی بر هوش مصنوعی، معیار کنترل مؤثر اهمیت بیشتری پیدا می‌کند؛ زیرا بسیاری از سامانه‌های هوشمند می‌توانند پس از آموزش اولیه، به صورت خودکار اقدام به شناسایی آسیب‌پذیری‌ها، تولید محتوای جعلی یا اجرای حملات سایبری کنند. در چنین شرایطی، صرف استناد به استقلال الگوریتمی سامانه نمی‌تواند موجب رفع مسئولیت دارنده شود؛ زیرا استقلال عملیاتی سامانه در بسیاری از موارد نتیجه تصمیمات انسانی در مراحل طراحی، آموزش، استقرار و تنظیم آن است. برای مثال، شخصی که سامانه‌ای را برای انجام عملیات خودکار در فضای سایبری به کار گرفته و از قابلیت‌های خطرناک آن آگاه بوده است، نمی‌تواند صرفاً به دلیل عدم دخالت لحظه‌ای خود، از مسئولیت کیفری رهایی یابد. در مقابل، اگر دارنده سامانه هیچ امکان واقعی برای کنترل عملکرد آن نداشته و رفتار مجرمانه ناشی از نقصی کاملاً غیرقابل پیش‌بینی باشد، انتساب کیفری با دشواری بیشتری مواجه خواهد شد (آلن و همکاران[۱۳]، ۲۰۲۵؛ صص ۳۸-۴۲).

در حقوق ایران، مفهوم کنترل مؤثر را می‌توان از طریق قواعد عمومی مسئولیت کیفری مانند مباشرت، معاونت، تسبیب و ترک فعل تحلیل کرد. هرچند قانون‌گذار ایرانی هنوز مقررات خاصی درباره مسئولیت کیفری ناشی از عملکرد هوش مصنوعی وضع نکرده است، اما در صورتی که دارنده سامانه امکان جلوگیری از وقوع نتیجه مجرمانه را داشته و با وجود آگاهی از خطر، از اعمال کنترل لازم خودداری کند، می‌توان بر اساس مبانی سنتی مسئولیت، امکان انتساب رفتار را بررسی کرد. به بیان دیگر، در حقوق ایران نیز مانند بسیاری از نظام‌های حقوقی دیگر، معیار اصلی نه مالکیت، بلکه میزان تسلط واقعی شخص بر فرآیند خطرآفرین است (رجبی، ۱۳۹۸، صص. ۴۴۹-۴۶۶).

قابلیت پیش‌بینی خطر

یکی دیگر از مبانی مهم مسئولیت کیفری دارنده هوش مصنوعی، قابلیت پیش‌بینی خطر ناشی از عملکرد سامانه است. در حقوق کیفری، مسئولیت معمولاً زمانی شکل می‌گیرد که شخص امکان آگاهی از پیامدهای زیان‌بار رفتار خود را داشته باشد. در فضای هوش مصنوعی نیز پرسش اساسی این است که آیا دارنده یا بهره‌بردار می‌توانسته احتمال استفاده مجرمانه از سامانه یا وقوع نتیجه زیان‌بار را پیش‌بینی کند یا خیر. هرچه سطح آگاهی و امکان پیش‌بینی خطر بیشتر باشد، مبنای سرزنش کیفری نیز تقویت خواهد شد (ابوت، ۲۰۲۰؛ صص ۱۱۲-۱۱۸). قابلیت پیش‌بینی در سامانه‌های هوش مصنوعی به دلیل ماهیت پیچیده الگوریتم‌ها اهمیت مضاعف دارد. برخی سامانه‌ها ممکن است رفتارهایی از خود نشان دهند که حتی برای توسعه‌دهندگان اولیه نیز کاملاً قابل پیش‌بینی نباشد. با وجود این، پیچیدگی فناوری به‌تنهایی نمی‌تواند موجب حذف مسئولیت شود؛ زیرا اشخاصی که سامانه‌های پرخطر را طراحی یا به کار می‌گیرند، موظف‌اند مخاطرات احتمالی آن را ارزیابی کرده و تدابیر لازم برای کاهش خطر را اتخاذ کنند. در واقع، در عصر هوش مصنوعی، معیار پیش‌بینی از «دانستن قطعی نتیجه» به «امکان

12. Abbott
13. Allen et al

منطقی تشخیص خطر» تغییر یافته است (خان[14]، ۲۰۲۴؛ صص ۵۲-۵۶). در نظام حقوقی آمریکا، مفهوم قابلیت پیش‌بینی جایگاه برجسته‌ای در تحلیل مسئولیت ناشی از فناوری‌های نوین دارد. حقوق آمریکا معمولاً به جای تمرکز بر شخصیت مستقل فناوری، رفتار اشخاص انسانی مرتبط با فناوری را بررسی می‌کند. بنابراین، اگر دارنده یا بهره‌بردار سامانه می‌دانسته یا باید می‌دانسته که استفاده از سامانه می‌تواند زمینه ارتکاب جرم سایبری را فراهم کند، عدم اتخاذ اقدامات پیشگیرانه می‌تواند مبنای مسئولیت وی قرار گیرد. این رویکرد در حقوق جرایم رایانه‌ای آمریکا نیز مشاهده می‌شود که بر عنصر آگاهی، قصد و امکان کنترل رفتار مجرمانه تأکید دارد (کر، ۲۰۲۲؛ صص ۲۰۱-۲۱۰). در اتحادیه اروپا نیز قابلیت پیش‌بینی خطر در قالب نظام مدیریت ریسک مورد توجه قرار گرفته است. قانون هوش مصنوعی اتحادیه اروپا با طبقه‌بندی سامانه‌ها بر اساس میزان خطر، تکالیف متفاوتی برای توسعه‌دهندگان و بهره‌برداران ایجاد کرده است. در این نظام، اشخاص مرتبط با سامانه‌های پرخطر مکلف‌اند پیش از استفاده، خطرات احتمالی را ارزیابی کرده، سازوکارهای نظارت انسانی ایجاد کنند و عملکرد سامانه را به صورت مستمر بررسی نمایند. بنابراین، عدم پیش‌بینی یا نادیده گرفتن خطرات شناخته‌شده می‌تواند مبنای مسئولیت حقوقی و در برخی موارد کیفری قرار گیرد.

تقصیر و وظیفه مراقبت

تقصیر یکی از مهم‌ترین مبانی سنتی مسئولیت کیفری است که در عصر هوش مصنوعی نیز اهمیت خود را حفظ کرده است. با این تفاوت که در مورد سامانه‌های هوشمند، تقصیر دیگر صرفاً به قصد مستقیم ارتکاب جرم محدود نمی‌شود، بلکه می‌تواند در قالب کوتاهی در نظارت، عدم به‌روزرسانی امنیتی، استفاده از داده‌های آموزشی نامناسب یا فقدان کنترل انسانی مناسب تحقق یابد. به همین دلیل، مفهوم «وظیفه مراقبت» در حقوق فناوری اهمیت فزاینده‌ای یافته است (ذاکری‌نیا، ۱۴۰۲، صص. ۱۵۲-۱۳۵). دارنده هوش مصنوعی به‌عنوان شخصی که از منافع سامانه بهره‌مند می‌شود، در برابر خطرات ناشی از عملکرد آن نیز دارای مسئولیت مراقبتی است. این وظیفه شامل انتخاب صحیح سامانه، بررسی قابلیت‌های آن، اعمال محدودیت‌های امنیتی، نظارت بر عملکرد و جلوگیری از استفاده مجرمانه است. برای مثال، اگر یک شرکت از سامانه هوش مصنوعی برای مدیریت امنیت شبکه استفاده کند اما بدون آزمایش‌های لازم آن را در معرض حملات سایبری قرار دهد، وقوع جرم ناشی از ضعف نظارتی می‌تواند زمینه مسئولیت مدیران یا بهره‌برداران را فراهم کند (حامدزاده و خوشدل واجاری، ۱۳۹۶، صص. ۱۴۴-۱۶۵). در اتحادیه اروپا، وظیفه مراقبت یکی از پایه‌های اصلی تنظیم مسئولیت هوش مصنوعی است. مقرراتی مانند AI Act و دستورالعمل NIS2 بر ضرورت امنیت سایبری، مدیریت خطر و نظارت مستمر تأکید دارند. این رویکرد نشان می‌دهد که مسئولیت در نظام اروپایی صرفاً پس از وقوع آسیب ایجاد نمی‌شود، بلکه از مرحله طراحی و بهره‌برداری سامانه آغاز می‌گردد.

14. Khan

رابطه سببیت و انتساب نتیجه

رابطه سببیت یکی از دشوارترین مسائل در مسئولیت کیفری ناشی از هوش مصنوعی است؛ زیرا میان رفتار انسانی و نتیجه مجرمانه ممکن است چندین مرحله پردازش الگوریتمی وجود داشته باشد. در جرایم سنتی، اثبات اینکه رفتار مرتکب موجب نتیجه شده است معمولاً ساده‌تر است، اما در جرایم مبتنی بر هوش مصنوعی، نتیجه ممکن است حاصل تعامل پیچیده میان داده‌ها، الگوریتم، محیط اجرایی و تصمیم‌های انسانی باشد. ازاین‌رو، برخی پژوهشگران از مفهوم «سببیت چندلایه» در تحلیل جرایم هوش مصنوعی استفاده کرده‌اند (هالوی، ۲۰۱۵؛ صص ۱۱۶-۱۲۱). در حقوق ایران، انتساب نتیجه مجرمانه همچنان نیازمند اثبات رابطه میان رفتار شخص و وقوع جرم است. بنابراین، دارنده سامانه تنها زمانی مسئول خواهد بود که بتوان نشان داد رفتار یا ترک فعل او در ایجاد نتیجه نقش مؤثر داشته است. برای نمونه، اگر شخصی سامانه هوشمندی را با قابلیت حمله سایبری فعال کند و نتیجه مجرمانه ناشی از همین طراحی یا بهره‌برداری باشد، رابطه سببیت قابل بررسی خواهد بود؛ اما اگر نتیجه ناشی از نقص کاملاً غیرقابل پیش‌بینی باشد، انتساب کیفری دشوارتر می‌شود (محمدی کردخیلی و همکاران، ۱۴۰۰، صص. ۶۳۰-۶۴۷). در حقوق آمریکا، رابطه سببیت با توجه به معیارهای سنتی حقوق کیفری و مسئولیت فناوری بررسی می‌شود. دادگاه‌ها معمولاً بررسی می‌کنند که آیا رفتار شخص، عامل قابل توجهی در وقوع جرم بوده و آیا نتیجه قابل پیش‌بینی بوده است یا خیر. در مقابل، اتحادیه اروپا با تمرکز بر مدیریت خطر، تلاش کرده است پیش از رسیدن به مرحله وقوع جرم، زنجیره سببیت را از طریق الزامات نظارتی و فنی کنترل کند.

تحلیل تطبیقی ایران، آمریکا و اتحادیه اروپا

تحلیل تطبیقی سه نظام حقوقی ایران، آمریکا و اتحادیه اروپا نشان می‌دهد که هیچ‌یک از این نظام‌ها تاکنون هوش مصنوعی را به‌عنوان فاعل مستقل جرم به رسمیت نشناخته‌اند، بلکه مسئولیت را همچنان متوجه اشخاص انسانی مرتبط با سامانه می‌دانند. با این حال، تفاوت اساسی میان آن‌ها در معیار تعیین شخص مسئول است. در حقوق ایران، مسئولیت دارنده هوش مصنوعی عمدتاً بر مبنای قواعد سنتی مسئولیت کیفری مانند تقصیر، سببیت و معاونت قابل تحلیل است. ضعف اصلی این نظام، فقدان مقررات اختصاصی درباره چرخه حیات سامانه‌های هوشمند و تقسیم مسئولیت میان مالک، توسعه‌دهنده و بهره‌بردار است. در حقوق آمریکا، معیارهایی مانند قابلیت پیش‌بینی خطر، آگاهی از ریسک و امکان کنترل رفتار سامانه نقش برجسته‌تری دارند. این نظام تلاش می‌کند بدون ایجاد شخصیت کیفری برای هوش مصنوعی، مسئولیت اشخاص انسانی را بر اساس میزان ارتباط آنان با نتیجه مجرمانه تعیین کند (خان، ۲۰۲۴؛ کر، ۲۰۲۲). در اتحادیه اروپا، رویکرد غالب مبتنی بر مدیریت خطر، کنترل مؤثر و وظیفه مراقبت است. قانون هوش مصنوعی اتحادیه اروپا نشان می‌دهد که این نظام به جای تمرکز بر مجازات پسینی، بر پیشگیری، ارزیابی ریسک و ایجاد تعهدات قانونی برای اشخاص فعال در چرخه حیات هوش مصنوعی تأکید دارد (پارلمان و شورای اروپا[15]، ۲۰۲۴). در مجموع، می‌توان گفت مبنای مشترک مسئولیت کیفری دارنده هوش مصنوعی در هر سه نظام، نه مالکیت

[15] .European Parliament and Council

صرف، بلکه ترکیبی از کنترل مؤثر، قابلیت پیش‌بینی خطر، رعایت وظیفه مراقبت و وجود رابطه سببیت میان رفتار دارنده و نتیجه مجرمانه است. بنابراین، پاسخ مناسب حقوق کیفری به چالش هوش مصنوعی نه در پذیرش شخصیت کیفری برای ماشین، بلکه در بازتعریف معیارهای انتساب مسئولیت به انسان‌هایی است که امکان کنترل، مدیریت و پیشگیری از خطرات ناشی از این فناوری را دارند.

مسئولیت کیفری دارنده در جرایم سایبری مبتنی بر هوش مصنوعی

گسترش استفاده از سامانه‌های هوش مصنوعی در فضای سایبری، الگوی سنتی ارتکاب جرم را با تغییرات اساسی مواجه ساخته است. در گذشته، بسیاری از جرایم سایبری نیازمند حضور مستقیم مرتکب، دانش فنی بالا و مداخله مستمر انسانی بود؛ اما امروزه سامانه‌های هوشمند قادرند بخش قابل توجهی از فرآیند ارتکاب جرم، از شناسایی هدف و تحلیل آسیب‌پذیری‌ها تا طراحی حمله و اجرای عملیات را به صورت خودکار انجام دهند. این تحول موجب شده است که نقش انسان از «عامل مستقیم ارتکاب جرم» به «طراح، مالک، کنترل‌کننده یا بهره‌بردار سامانه» تغییر یابد. بنابراین، در جرایم سایبری مبتنی بر هوش مصنوعی، مسئله اصلی تعیین جایگاه دارنده سامانه و بررسی این موضوع است که در چه شرایطی می‌توان نتیجه مجرمانه ناشی از عملکرد هوش مصنوعی را به وی منتسب کرد. در این زمینه، مسئولیت کیفری دارنده نه بر مبنای صرف مالکیت فناوری، بلکه بر اساس میزان کنترل، آگاهی از قابلیت‌های خطرناک سامانه، امکان پیشگیری از سوءاستفاده و نقش وی در ایجاد وضعیت مجرمانه ارزیابی می‌شود (معبودیان، ۱۴۰۴، صص. ۵-۱۰؛ هرکه و توت[16]، ۲۰۲۴؛ صص ۱۸۰-۱۸۵). ویژگی اصلی جرایم سایبری مبتنی بر هوش مصنوعی، کاهش فاصله میان تصمیم انسانی و وقوع نتیجه مجرمانه است. سامانه‌های هوشمند با استفاده از الگوریتم‌های یادگیری ماشین و تحلیل داده‌های گسترده می‌توانند رفتارهای خود را متناسب با شرایط تغییر داده و روش‌های جدیدی برای دستیابی به هدف پیدا کنند. این ویژگی اگرچه در کاربردهای مشروع موجب افزایش کارایی و سرعت می‌شود، اما در حوزه مجرمانه امکان طراحی حملات پیچیده‌تر، پنهان‌تر و گسترده‌تر را فراهم می‌سازد. از همین رو، حقوق کیفری با مسئله‌ای جدید مواجه است که در آن، نتیجه مجرمانه ممکن است بدون صدور دستور لحظه‌ای از سوی انسان رخ دهد، اما همچنان در چارچوب اهداف و محدودیت‌هایی شکل گرفته باشد که توسط انسان ایجاد شده است. بنابراین، تحلیل مسئولیت دارنده هوش مصنوعی در جرایم سایبری نیازمند بررسی دقیق نوع جرم، میزان استقلال سامانه و نقش واقعی دارنده در فرآیند وقوع جرم است (راسل و نورویگ[17]، ۲۰۲۱؛ صص ۱۰۵۰-۱۰۵۵).

نفوذ غیرمجاز و حملات خودکار

یکی از مهم‌ترین حوزه‌های سوءاستفاده از هوش مصنوعی در فضای سایبری، ارتکاب نفوذ غیرمجاز و حملات خودکار علیه سامانه‌های رایانه‌ای است. در حملات سنتی، مهاجم معمولاً به صورت مستقیم اقدام به شناسایی آسیب‌پذیری‌ها، انتخاب هدف و اجرای عملیات نفوذ می‌کرد؛ اما سامانه‌های هوشمند

16. Herke & Tóth
17. Russell & Norvig

این امکان را فراهم کرده‌اند که بخش زیادی از این فرآیند به صورت خودکار انجام شود. الگوریتم‌های هوش مصنوعی قادرند شبکه‌ها را تحلیل کرده، نقاط ضعف امنیتی را شناسایی کنند، الگوهای دفاعی سیستم مقابل را بررسی نمایند و حتی راهبرد حمله را در جریان عملیات تغییر دهند. این قابلیت موجب شده است که حملات سایبری از نظر سرعت، مقیاس و پیچیدگی وارد مرحله جدیدی شوند (سعود[18]، ۲۰۲۴؛ صص ۵-۷). در چنین شرایطی، تعیین مسئولیت کیفری دارنده سامانه با دشواری‌های جدیدی روبه‌رو می‌شود. برای مثال، اگر شخصی سامانه هوش مصنوعی را با هدف انجام عملیات امنیتی توسعه دهد، اما آن را به گونه‌ای تنظیم کند که قابلیت نفوذ خودکار به سامانه‌های دیگر را داشته باشد، در صورت تحقق جرم، نمی‌توان صرفاً به عملکرد مستقل الگوریتم استناد کرد. زیرا ایجاد قابلیت خطرناک، انتخاب معماری سامانه و تعیین اهداف عملیاتی، همگی در حوزه تصمیمات انسانی قرار دارند. بنابراین، دارنده یا کنترل‌کننده سامانه در صورتی که از قابلیت‌های مجرمانه آن آگاه باشد یا امکان جلوگیری از سوءاستفاده را داشته باشد، می‌تواند بر اساس معیارهای تقصیر، کنترل مؤثر و قابلیت پیش‌بینی خطر مسئول شناخته شود (ترنر، ۲۰۱۹؛ صص ۱۲۵-۱۳۰).

در حقوق کیفری ایران، نفوذ غیرمجاز در سامانه‌های رایانه‌ای بر اساس مقررات جرایم رایانه‌ای و قواعد عمومی مسئولیت کیفری قابل تحلیل است. هرچند قانون‌گذار ایرانی مقررات خاصی درباره حملات خودکار مبتنی بر هوش مصنوعی پیش‌بینی نکرده است، اما در صورت احراز نقش مؤثر دارنده در ایجاد یا تسهیل جرم، امکان انتساب رفتار از طریق قواعد معاونت، تسبیب یا ترک فعل قابل بررسی خواهد بود. برای نمونه، شخصی که سامانه‌ای را با قابلیت نفوذ خودکار در اختیار دیگران قرار می‌دهد و هیچ‌گونه محدودیت یا نظارتی بر استفاده از آن اعمال نمی‌کند، ممکن است به دلیل ایجاد زمینه وقوع جرم، مسئولیت کیفری پیدا کند (میرمحمدصادقی، ۱۴۰۲، صص. ۲۴۵-۲۵۰). در نظام‌های حقوقی پیشرفته نیز رویکرد غالب، پذیرش مسئولیت انسانی به جای مسئولیت مستقل هوش مصنوعی است. اتحادیه اروپا و ایالات متحده آمریکا، هرچند استقلال عملیاتی سامانه‌های هوشمند را پذیرفته‌اند، اما آن را موجب انتقال مسئولیت به ماشین نمی‌دانند. در این نظام‌ها، تمرکز اصلی بر نقش توسعه‌دهنده، مالک، اپراتور و بهره‌بردار در کنترل خطرات ناشی از سامانه است. بنابراین، خودکار بودن حمله سایبری، مانع از مسئولیت شخصی که امکان مدیریت یا جلوگیری از آن را داشته است، نخواهد بود (هرکه و توت، ۲۰۲۴؛ صص ۱۹۰-۱۹۵).

بدافزارهای هوشمند و فیشینگ مبتنی بر AI

یکی دیگر از مهم‌ترین کاربردهای مجرمانه هوش مصنوعی در فضای سایبری، تولید و توسعه بدافزارهای هوشمند و حملات فیشینگ پیشرفته است. بدافزارهای سنتی معمولاً بر اساس کدهای از پیش طراحی‌شده عمل می‌کردند، اما بدافزارهای مبتنی بر هوش مصنوعی می‌توانند رفتار خود را تغییر دهند، از روش‌های دفاعی عبور کنند و حتی متناسب با محیط آلوده، ساختار خود را اصلاح نمایند. این قابلیت، شناسایی و مقابله با آن‌ها را برای نهادهای امنیتی دشوارتر کرده است (سعود، ۲۰۲۴؛ صص ۷-۱۲). در حوزه فیشینگ نیز هوش مصنوعی موجب تحول جدی در شیوه ارتکاب جرم شده است.

18. Saud

سامانه‌های مولد هوش مصنوعی می‌توانند پیام‌هایی با ساختار بسیار طبیعی، متناسب با ویژگی‌های فردی قربانی و بدون نشانه‌های معمول فیشینگ تولید کنند. در نتیجه، حملات فیشینگ از پیام‌های عمومی و ساده به حملاتی شخصی‌سازی‌شده و مبتنی بر تحلیل رفتار کاربران تبدیل شده‌اند. این امر احتمال موفقیت کلاهبرداری‌های اینترنتی و سرقت اطلاعات را افزایش داده است (راسل و نورویگ، ۲۰۲۱؛ صص ۱۰۵۸-۱۰۶۲). در این نوع جرایم، مسئولیت کیفری دارنده هوش مصنوعی بیش از هر چیز به میزان آگاهی و نقش وی در بهره‌برداری از سامانه وابسته است. اگر شخصی آگاهانه از قابلیت‌های تولید محتوای فریبنده یا ایجاد بدافزار برای ارتکاب جرم استفاده کند، عملکرد خودکار سامانه نمی‌تواند مانع انتساب جرم به وی شود. همچنین، اگر شرکت یا سازمانی سامانه‌ای را برای خدمات مشروع توسعه دهد، اما بدون اتخاذ تدابیر امنیتی لازم آن را در معرض سوءاستفاده قرار دهد، موضوع مسئولیت ناشی از قصور در مراقبت و مدیریت خطر مطرح خواهد شد (معبودیان، ۱۴۰۴، صص. ۱۵-۱۲). در حقوق تطبیقی، این موضوع با مفهوم «مسئولیت مبتنی بر کنترل خطر» تحلیل می‌شود. بدین معنا که هر شخصی که بیشترین امکان کنترل خطر ناشی از سامانه را دارد، باید بیشترین مسئولیت را در جلوگیری از سوءاستفاده بر عهده داشته باشد. این رویکرد در اتحادیه اروپا به‌ویژه در مقررات مربوط به سامانه‌های پرخطر هوش مصنوعی مورد توجه قرار گرفته و بر ارزیابی ریسک، نظارت انسانی و امنیت داده‌ها تأکید شده است.

جعل عمیق و سرقت داده‌ها

فناوری جعل عمیق[19] یکی از مهم‌ترین مصادیق جرایم سایبری مبتنی بر هوش مصنوعی است که چالش‌های جدیدی برای حقوق کیفری ایجاد کرده است. این فناوری با استفاده از الگوریتم‌های یادگیری عمیق قادر است تصاویر، صداها و ویدئوهایی تولید کند که شباهت بسیار زیادی به محتوای واقعی دارند. استفاده مجرمانه از دیپ‌فیک می‌تواند در قالب جعل هویت، اخاذی، کلاهبرداری، انتشار اطلاعات نادرست و حتی تهدید امنیت عمومی ظهور پیدا کند (کرامت میرشکارلو و خالقی، ۱۴۰۳، صص. ۵-۸). مسئولیت کیفری دارنده سامانه در جرایم مرتبط با جعل عمیق، وابسته به بررسی نقش وی در تولید، انتشار یا بهره‌برداری از محتوای جعلی است. شخصی که سامانه هوش مصنوعی را با هدف ساخت محتوای جعلی و فریب دیگران به کار می‌گیرد، نمی‌تواند مسئولیت خود را به دلیل نقش الگوریتم در تولید محتوا نفی کند. زیرا الگوریتم صرفاً وسیله‌ای برای تحقق هدف انسانی بوده است. در مقابل، اگر یک سامانه عمومی تولید محتوا بدون قصد مجرمانه مورد سوءاستفاده قرار گیرد، بررسی مسئولیت توسعه‌دهنده یا ارائه‌دهنده خدمت نیازمند احراز تقصیر، عدم رعایت الزامات امنیتی یا فقدان تدابیر پیشگیرانه خواهد بود (ترنر، ۲۰۱۹؛ صص ۱۴۵-۱۵۰). سرقت داده‌ها نیز از دیگر حوزه‌هایی است که هوش مصنوعی می‌تواند نقش مهمی در آن ایفا کند. سامانه‌های هوشمند قادرند حجم عظیمی از اطلاعات را تحلیل کرده، داده‌های حساس را شناسایی و مسیرهای دسترسی غیرمجاز را پیدا کنند. در چنین مواردی، دارنده سامانه‌ای که از قابلیت‌های آن برای جمع‌آوری غیرقانونی داده‌ها استفاده می‌کند، بر اساس اصول عمومی حقوق کیفری مسئول خواهد بود. همچنین، سازمان‌هایی که مسئول نگهداری

---

[19]. Deepfake

داده‌های کاربران هستند، در صورت فقدان اقدامات امنیتی لازم ممکن است با مسئولیت ناشی از بی‌احتیاطی یا نقض تکلیف مراقبت مواجه شوند (میرمحمدصادقی، ۱۴۰۲، صص. ۳۱۰-۳۱۵).

تحلیل تطبیقی پاسخ نظام‌های حقوقی

بررسی تطبیقی ایران، آمریکا و اتحادیه اروپا نشان می‌دهد که هر سه نظام حقوقی، علی‌رغم تفاوت در شیوه تنظیم مقررات، بر اصل مسئولیت انسانی در جرایم سایبری مبتنی بر هوش مصنوعی تأکید دارند. هیچ‌یک از این نظام‌ها تاکنون هوش مصنوعی را به‌عنوان فاعل مستقل جرم نپذیرفته‌اند؛ بلکه مسئولیت را متوجه اشخاصی می‌دانند که در طراحی، کنترل، بهره‌برداری یا مدیریت سامانه نقش دارند. در حقوق ایران، مسئولیت دارنده هوش مصنوعی عمدتاً از طریق قواعد سنتی حقوق کیفری مانند مباشرت، معاونت، تسبیب و تقصیر تحلیل می‌شود. مزیت این رویکرد، هماهنگی با اصول بنیادین حقوق جزا است، اما ضعف آن در نبود مقررات اختصاصی برای تعیین حدود مسئولیت اشخاص مختلف در چرخه حیات هوش مصنوعی است. در حقوق آمریکا، رویکرد غالب بر قابلیت پیش‌بینی خطر، میزان کنترل و نقش واقعی شخص در ایجاد نتیجه مجرمانه استوار است. این نظام تلاش می‌کند بدون ایجاد شخصیت کیفری برای هوش مصنوعی، مسئولیت اشخاص انسانی را بر اساس میزان آگاهی و امکان جلوگیری از جرم تعیین کند (سعود، ۲۰۲۴؛ کر، ۲۰۲۲). در اتحادیه اروپا، پاسخ حقوقی بیشتر ماهیت پیشگیرانه دارد. قانون هوش مصنوعی اتحادیه اروپا با ایجاد الزامات ارزیابی خطر، شفافیت، نظارت انسانی و امنیت سامانه‌ها تلاش کرده است پیش از وقوع جرم، زمینه‌های مسئولیت را کنترل کند. بنابراین، این نظام بیش از آنکه بر مجازات پسینی تمرکز کند، بر مدیریت خطر و حکمرانی مسئولانه فناوری تأکید دارد (پارلمان و شورای اروپا، ۲۰۲۴). در نهایت، می‌توان گفت که مسئولیت کیفری دارنده هوش مصنوعی در جرایم سایبری مبتنی بر سه محور اصلی استوار است: نخست، میزان کنترل مؤثر دارنده بر سامانه؛ دوم، قابلیت پیش‌بینی و مدیریت خطرات ناشی از عملکرد آن؛ و سوم، وجود رابطه سببیت میان رفتار دارنده و نتیجه مجرمانه. بر این اساس، چالش اصلی حقوق کیفری آینده، شناسایی شخصیت کیفری برای هوش مصنوعی نیست، بلکه ایجاد معیارهای دقیق‌تر برای تعیین مسئولیت انسان‌هایی است که در ایجاد، توسعه و بهره‌برداری از این فناوری نقش دارند.

چالش‌های تقنینی و الگوی پیشنهادی برای حقوق ایران

خلأهای حقوق ایران

گسترش روزافزون کاربرد سامانه‌های هوش مصنوعی در حوزه‌های مختلف اجتماعی، اقتصادی و به‌ویژه فضای سایبری، نظام‌های حقوقی را با چالش‌های جدیدی در زمینه تعیین مسئولیت مواجه ساخته است. در حقوق ایران، اگرچه مجموعه‌ای از مقررات عمومی مانند قانون مجازات اسلامی مصوب ۱۳۹۲، قانون جرایم رایانه‌ای مصوب ۱۳۸۸ و قانون مسئولیت مدنی مصوب ۱۳۳۹ ظرفیت‌هایی برای پاسخ‌گویی به برخی رفتارهای زیان‌بار ناشی از فناوری‌های نوین دارند، اما این قوانین عمدتاً بر مبنای فرض سنتیِ وجود یک انسان به عنوان فاعل مستقیم جرم یا عامل ورود زیان طراحی شده‌اند. در نتیجه، در مواردی که یک سامانه هوش مصنوعی با درجه‌ای از استقلال عملیاتی اقدام به ارتکاب رفتار

مجرمانه در فضای سایبری می‌کند، تعیین شخص مسئول و مبنای انتساب رفتار با دشواری‌های جدی مواجه می‌شود (اردبیلی، ۱۴۰۱، صص. ۱۱۰-۱۰۸). یکی از مهم‌ترین خلأهای حقوق ایران، فقدان مقررات اختصاصی درباره جایگاه حقوقی هوش مصنوعی و تعیین حدود مسئولیت اشخاص مرتبط با آن است. در نظام کیفری ایران، مسئولیت کیفری بر پایه اصولی مانند اراده، اختیار، علم، عمد یا تقصیر انسانی استوار شده است؛ در حالی که سامانه‌های هوش مصنوعی ممکن است در فرآیند تصمیم‌گیری خود از الگوهای یادگیری ماشینی استفاده کنند و عملکرد آن‌ها همیشه به شکل مستقیم قابل پیش‌بینی توسط طراح، تولیدکننده یا دارنده نباشد. بنابراین، پرسش اساسی این است که در صورت وقوع جرم سایبری توسط یک سامانه هوشمند، آیا باید مسئولیت را متوجه دارنده، بهره‌بردار، تولیدکننده، برنامه‌نویس یا ارائه‌دهنده خدمات دانست؟ پاسخ روشن به این پرسش در قوانین کیفری ایران وجود ندارد (شامبیاتی، ۱۴۰۲).

از منظر حقوق کیفری، یکی از ارکان اساسی مسئولیت، وجود رابطه سببیت میان رفتار مرتکب و نتیجه مجرمانه است. در جرایم سنتی، این رابطه معمولاً میان رفتار مستقیم انسان و نتیجه زیان‌بار برقرار می‌شود؛ اما در جرایم مبتنی بر هوش مصنوعی، فاصله میان تصمیم انسانی و نتیجه نهایی افزایش یافته و ممکن است اقدامات سامانه هوشمند، الگوریتم‌ها و فرآیندهای خودآموز در تحقق نتیجه نقش داشته باشند. این وضعیت موجب می‌شود که قواعد موجود درباره معاونت، مباشرت، سببیت و تقصیر برای پاسخ‌گویی کامل به جرایم نوظهور کافی نباشند (نظرپور، ایزدی‌فرد و محسنی دهکلانی، ۱۴۰۰، صص. ۱۸۰-۱۸۵). علاوه بر این، در حقوق ایران هنوز معیار مشخصی برای ارزیابی «وظیفه مراقبت» دارندگان و بهره‌برداران هوش مصنوعی وجود ندارد. در حالی که استفاده‌کننده از یک فناوری پیشرفته باید نسبت به خطرات احتمالی آن آگاهی داشته و اقدامات پیشگیرانه لازم را انجام دهد، فقدان معیارهای قانونی درباره سطح مراقبت مورد انتظار، امکان اثبات تقصیر کیفری یا مدنی دارنده را دشوار می‌سازد. به همین دلیل، برخی پژوهشگران بر ضرورت توسعه مفهوم تقصیر فناورانه و پذیرش استانداردهای جدید مسئولیت متناسب با ویژگی‌های هوش مصنوعی تأکید کرده‌اند (ولی‌پور و اسماعیلی، ۱۴۰۰، صص. ۶-۴).

دستاوردهای حقوق آمریکا و اتحادیه اروپا

در نظام حقوقی آمریکا، مواجهه با مسئولیت ناشی از فناوری‌های هوشمند بیشتر بر مبنای توسعه اصول موجود در حقوق کیفری و مدنی، به‌ویژه مفاهیمی مانند تقصیر، سهل‌انگاری، کنترل مؤثر و پیش‌بینی‌پذیری خطر صورت گرفته است. حقوق آمریکا به جای پذیرش شخصیت کیفری مستقل برای هوش مصنوعی، تلاش می‌کند مسئولیت را به اشخاص انسانی مرتبط با سامانه، از جمله توسعه‌دهندگان، مالکان و بهره‌برداران منتسب کند. در این رویکرد، هرگاه شخصی با علم به خطرات احتمالی یک سامانه هوشمند، بدون اتخاذ تدابیر لازم از آن استفاده کند و سامانه مزبور در ارتکاب جرم نقش داشته باشد، امکان انتساب مسئولیت به وی تقویت می‌شود (ترنر، ۲۰۱۹). در اتحادیه اروپا، رویکرد قانون‌گذار به سمت ایجاد یک نظام تنظیم‌گری پیشگیرانه حرکت کرده است. تصویب قانون

هوش مصنوعی اتحادیه اروپا[20] نشان‌دهنده تلاش برای ایجاد چارچوبی مبتنی بر سطح خطر سامانه‌های هوش مصنوعی است. این مقررات با دسته‌بندی سامانه‌ها بر اساس میزان ریسک، الزامات مشخصی برای شفافیت، نظارت انسانی، ارزیابی خطر و کنترل عملکرد وضع کرده است. چنین رویکردی می‌تواند مبنایی برای تعیین مسئولیت دارندگان و بهره‌برداران سامانه‌های هوشمند در آینده باشد؛ زیرا عدم رعایت الزامات مراقبتی می‌تواند به عنوان نشانه‌ای از تقصیر یا بی‌احتیاطی تلقی شود (پارلمان و شورای اروپا، ۲۰۲۴). همچنین در حوزه مسئولیت ناشی از هوش مصنوعی، اتحادیه اروپا با طرح دستورالعمل مسئولیت هوش مصنوعی تلاش کرده است قواعد سنتی مسئولیت را با ویژگی‌های فناوری‌های نوین هماهنگ کند. این رویکرد بر کاهش دشواری اثبات رابطه سببیت و ایجاد امکان دسترسی زیان‌دیدگان به اطلاعات فنی تأکید دارد؛ زیرا یکی از مشکلات اصلی در دعاوی مرتبط با هوش مصنوعی، عدم امکان شناخت دقیق فرآیند تصمیم‌گیری الگوریتم‌ها است.

الگوی پیشنهادی برای اصلاح مقررات کیفری ایران

با توجه به چالش‌های موجود، اصلاح نظام کیفری ایران در حوزه مسئولیت ناشی از هوش مصنوعی باید بر مبنای پذیرش یک الگوی ترکیبی انجام شود؛ الگویی که ضمن حفظ اصل شخصی بودن مسئولیت کیفری، امکان انتساب رفتارهای مجرمانه ناشی از استفاده نادرست از هوش مصنوعی را فراهم کند. نخستین گام در این مسیر، تعریف جایگاه حقوقی هوش مصنوعی و تعیین نقش اشخاص مرتبط با آن است. به نظر می‌رسد اعطای شخصیت کیفری مستقل به هوش مصنوعی در شرایط فعلی با مبانی حقوق کیفری ایران سازگار نباشد؛ زیرا مسئولیت کیفری نیازمند وجود اراده و قابلیت سرزنش اخلاقی است. بنابراین، مسئولیت باید متوجه اشخاص انسانی مرتبط با سامانه باشد (نظرپور، ایزدی‌فرد و محسنی دهکلانی، ۱۴۰۰). در این چارچوب، پیشنهاد می‌شود قانون‌گذار ایرانی مفهوم «دارنده یا بهره‌بردار هوش مصنوعی» را به عنوان یک عنوان حقوقی مستقل شناسایی کرده و برای وی تکالیفی مانند نظارت مستمر، کنترل عملکرد، به‌روزرسانی امنیتی، جلوگیری از سوءاستفاده و ارزیابی خطرات احتمالی مقرر کند. نقض این تکالیف، در صورتی که منجر به وقوع جرم سایبری شود، می‌تواند مبنای تحقق تقصیر کیفری قرار گیرد. همچنین ضروری است مقررات قانون جرایم رایانه‌ای با توجه به ظهور جرایمی مانند حملات خودکار، بدافزارهای هوشمند، جعل عمیق و سوءاستفاده الگوریتمی بازنگری شود. فناوری‌هایی مانند دیپ‌فیک نشان داده‌اند که هوش مصنوعی می‌تواند امکان تولید محتوای جعلی با قابلیت فریب گسترده را فراهم کند و تهدیدهای جدیدی علیه امنیت فردی و اجتماعی ایجاد نماید (چاولا[21]، ۲۰۱۹؛ فلچر[22]، ۲۰۱۸). بنابراین، جرم‌انگاری برخی رفتارهای خاص مانند استفاده عامدانه از هوش مصنوعی برای تولید محتوای جعلی زیان‌بار یا طراحی سامانه‌های خودکار با هدف ارتکاب جرم، می‌تواند خلأهای موجود را کاهش دهد (دی[23]، ۲۰۱۹). در نهایت، اصلاح قوانین ایران باید با بهره‌گیری از تجربه آمریکا و اتحادیه اروپا، بر ایجاد نظامی مبتنی بر «مسئولیت تدریجی و مبتنی بر میزان

---

20. Artificial Intelligence Act
21. Chawla
22. Fletcher
23. Day

کنترل» استوار شود. به این معنا که هرچه میزان کنترل، دسترسی و امکان پیشگیری دارنده نسبت به عملکرد سامانه بیشتر باشد، مسئولیت وی نیز افزایش یابد. چنین الگویی ضمن هماهنگی با اصول بنیادین حقوق کیفری ایران، می‌تواند پاسخی مناسب به چالش‌های ناشی از توسعه هوش مصنوعی در جرایم سایبری ارائه دهد و زمینه شکل‌گیری یک نظام حقوقی آینده‌نگر در مواجهه با فناوری‌های نوین را فراهم سازد.

نتیجه‌گیری

تحولات شتابان فناوری هوش مصنوعی و ورود گسترده آن به عرصه‌های مختلف زندگی اجتماعی، اقتصادی و ارتباطی، ساختارهای سنتی حقوق کیفری را با پرسش‌های بنیادینی درباره مفهوم فاعل جرم، معیارهای انتساب رفتار مجرمانه و حدود مسئولیت اشخاص مرتبط با این فناوری مواجه ساخته است. در گذشته، ارتکاب جرم سایبری عمدتاً مستلزم دخالت مستقیم و ارادی انسان بود؛ اما امروزه سامانه‌های هوش مصنوعی قادرند با بهره‌گیری از الگوریتم‌های یادگیری ماشین، داده‌های گسترده و فرآیندهای تصمیم‌گیری خودکار، اقداماتی انجام دهند که ممکن است به نقض امنیت اطلاعات، سرقت داده‌ها، نفوذ غیرمجاز، تولید محتوای جعلی و سایر اشکال رفتار مجرمانه منجر شود. ازاین‌رو، مسئله اصلی دیگر صرفاً شناسایی مرتکب فیزیکی جرم نیست، بلکه تعیین این موضوع است که در زنجیره پیچیده طراحی، توسعه، مالکیت و بهره‌برداری از هوش مصنوعی، مسئولیت کیفری باید متوجه چه شخص یا اشخاصی باشد. بررسی مبانی مسئولیت کیفری دارنده هوش مصنوعی نشان داد که مهم‌ترین معیارهای انتساب مسئولیت در نظام‌های حقوقی مختلف بر مفاهیمی همچون کنترل مؤثر، قابلیت پیش‌بینی خطر، تقصیر، وظیفه مراقبت و رابطه سببیت استوار شده است. دارنده یا بهره‌بردار هوش مصنوعی، اگرچه ممکن است مستقیماً رفتار مجرمانه را انجام نداده باشد، اما در صورتی که اختیار کنترل سامانه را داشته، از خطرات احتمالی آن آگاه بوده یا امکان پیشگیری از وقوع جرم را داشته باشد، می‌تواند مسئول شناخته شود. بنابراین، معیار اصلی در تعیین مسئولیت، نه صرف مالکیت یا استفاده از فناوری، بلکه میزان کنترل، آگاهی و امکان مداخله مؤثر در عملکرد سامانه هوشمند است.

تحلیل مفهوم «هوش مصنوعی؛ ابزار یا فاعل؟» نیز نشان داد که در وضعیت فعلی، پذیرش شخصیت کیفری مستقل برای هوش مصنوعی با اصول بنیادین حقوق کیفری، از جمله اصل شخصی بودن مسئولیت کیفری و ضرورت وجود اراده و قابلیت سرزنش اخلاقی، سازگاری کامل ندارد. هوش مصنوعی هرچند ممکن است در تصمیم‌گیری و اجرای عملیات پیچیده استقلال فنی داشته باشد، اما فاقد آگاهی، قصد مجرمانه و مسئولیت اخلاقی انسانی است. بنابراین، رویکرد منطقی‌تر آن است که هوش مصنوعی به عنوان ابزار یا عامل واسطه‌ای در ارتکاب جرم تلقی شود و مسئولیت کیفری بر اساس میزان نقش و تقصیر اشخاص انسانی مرتبط با آن، مانند دارنده، بهره‌بردار، تولیدکننده یا ارائه‌دهنده خدمات، تعیین گردد. بررسی جرایم سایبری مبتنی بر هوش مصنوعی نیز نشان داد که این فناوری می‌تواند دامنه و پیچیدگی جرایم رایانه‌ای را افزایش دهد. حملات خودکار، بدافزارهای هوشمند، فیشینگ مبتنی بر الگوریتم‌های یادگیری ماشین، جعل عمیق و سرقت داده‌ها نمونه‌هایی از جرایمی هستند که به دلیل سرعت، مقیاس و قابلیت پنهان‌سازی بالا، چالش‌های جدیدی را برای نظام عدالت کیفری ایجاد

کرده‌اند. در چنین شرایطی، قواعد سنتی حقوق کیفری که عمدتاً بر رفتار مستقیم انسان تمرکز دارند، برای پاسخ‌گویی کامل کافی نیستند و نیازمند تفسیر توسعه‌ای یا اصلاح تقنینی هستند. مطالعه تطبیقی نظام‌های حقوقی آمریکا و اتحادیه اروپا نشان داد که این نظام‌ها بدون پذیرش شخصیت کیفری مستقل برای هوش مصنوعی، تلاش کرده‌اند مسئولیت را از طریق توسعه مفاهیمی مانند مراقبت متعارف، مدیریت خطر و مسئولیت ناشی از عدم نظارت کافی تنظیم کنند. در ایالات متحده، تمرکز اصلی بر انتساب تقصیر به اشخاص انسانی مرتبط با سامانه هوشمند است؛ به گونه‌ای که هرگاه شخصی با وجود امکان پیش‌بینی خطر، اقدامات لازم برای جلوگیری از سوءاستفاده را انجام ندهد، امکان مسئولیت وی افزایش می‌یابد. در اتحادیه اروپا نیز با تصویب مقررات مرتبط با هوش مصنوعی، رویکردی مبتنی بر مدیریت ریسک، شفافیت، نظارت انسانی و طبقه‌بندی سامانه‌ها بر اساس میزان خطر اتخاذ شده است. این تجربه‌ها می‌تواند الگویی مناسب برای توسعه حقوق ایران باشد.

در حقوق ایران، اگرچه قانون جرایم رایانه‌ای، قانون مجازات اسلامی و قواعد عمومی مسئولیت مدنی ظرفیت‌هایی برای برخورد با برخی نتایج زیان‌بار ناشی از استفاده از هوش مصنوعی دارند، اما هنوز مقررات مشخصی درباره مسئولیت کیفری دارنده یا بهره‌بردار هوش مصنوعی وجود ندارد. مهم‌ترین خلأ موجود، نبود معیار قانونی برای تعیین حدود وظیفه مراقبت، میزان کنترل مورد انتظار و نحوه اثبات رابطه سببیت میان عملکرد سامانه هوشمند و نتیجه مجرمانه است. این خلأ می‌تواند موجب ایجاد اختلاف در رویه قضایی و دشواری در تعقیب کیفری مرتکبان واقعی شود. بر این اساس، پیشنهاد می‌شود قانون‌گذار ایرانی ضمن حفظ مبانی سنتی حقوق کیفری، مقررات ویژه‌ای درباره مسئولیت ناشی از استفاده از هوش مصنوعی تدوین کند. در این چارچوب، به جای اعطای شخصیت کیفری مستقل به هوش مصنوعی، باید مسئولیت اشخاص انسانی مرتبط با آن بر اساس معیارهایی مانند میزان کنترل، سطح آگاهی، امکان پیشگیری، رعایت الزامات امنیتی و درجه تقصیر تعیین شود. همچنین ضروری است قانون جرایم رایانه‌ای با توجه به ظهور جرایم جدید مانند جعل عمیق، حملات خودکار و سوءاستفاده از الگوریتم‌های هوشمند بازنگری شود و برای دارندگان و بهره‌برداران سامانه‌های پرخطر، تکالیف مشخصی در زمینه نظارت و پیشگیری ایجاد گردد. در نهایت، می‌توان نتیجه گرفت که چالش اصلی حقوق کیفری در عصر هوش مصنوعی، نه جایگزینی انسان با ماشین، بلکه یافتن شیوه‌ای مناسب برای تنظیم رابطه میان فناوری، کنترل انسانی و مسئولیت حقوقی است. هوش مصنوعی باید به عنوان یک فناوری قدرتمند اما فاقد شخصیت اخلاقی و کیفری مستقل تلقی شود و مسئولیت ناشی از کاربرد زیان‌بار آن بر اساس اصول انتساب، تقصیر و کنترل انسانی تعیین گردد. اتخاذ چنین رویکردی می‌تواند ضمن جلوگیری از خلأ مسئولیت در جرایم سایبری نوین، تعادلی میان حمایت از امنیت عمومی، توسعه فناوری و رعایت اصول بنیادین حقوق کیفری ایران ایجاد کند.

منابع و مراجع

۱) اباذری، کسری. (۱۴۰۳). بررسی مسئولیت کیفری ناشی از سوءاستفاده از هوش مصنوعی در جرایم سایبری. در مجموعه مقالات پانزدهمین کنفرانس بین‌المللی دستاوردهای نوین پژوهشی در علوم تربیتی، روانشناسی و علوم اجتماعی.

۲) اردبیلی، محمدعلی. (۱۴۰۱). *حقوق جزای عمومی* (جلد اول، چاپ سی‌وهشتم). تهران: میزان.

۳) اسلامی، محمدرضا. (۱۳۹۱). مسئولیت کیفری هوش مصنوعی و الگوریتم‌های خودآموز. تهران: نشر عدالت.

۴) جان‌آبادی، نرگس. (۱۴۰۳). بررسی نقش هوش مصنوعی در ارتکاب جرایم سایبری و اقتصادی. در مجموعه مقالات دومین کنفرانس بین‌المللی حقوق، مدیریت، علوم تربیتی، روانشناسی و مدیریت برنامه‌ریزی آموزشی.

۵) الجیزانی، حیدر کریم عبود؛ حیدری، مسعود؛ علی، حیدر حسین؛ صعیدی، یاسین (۱۴۰۴). مسئولیت کیفری ناشی از جرائم ارتکابی به وسیله ربات‌ها، دوره ۳، شماره ۳ (۱۱)، صص ۱-۱۷.

۶) حامدزاده، حامد و علی خوشدل واجاری، بررسی مسئولیت مدنی ناشی از ارتباطات الکترونیکی، فصلنامه مطالعات حقوق، شماره ۱۶، ۱۳۹۶، صص ۱۶۵-۱۴۴.

۷) حکمت‌نیا، محمود، مرتضی محمدی و محسن واثقی، مسئولیت مدنی ناشی از تولید ربات‌های حاوی هوش‌مصنوعی خودمختار، فصلنامه حقوقی اسلامی، دوره ۱۶، شماره ۶۰، ۱۳۹۸، صص ۲۵۵-۲۳۱.

۸) ذاکری‌نیا، هانیه، ماهیت و مبنای مسئولیت مدنی ناشی از هوش‌مصنوعی در حقوق ایران و کشورهای اتحادیه اروپا، مجله حقوق خصوصی، دوره ۲۰، شماره ۱، ۱۴۰۲، صص ۱۵۲-۱۳۵.

۹) رجبی، عبدالله، ضمان در هوش‌مصنوعی، مطالعات حقوق تطبیقی، دوره ۱۰، شماره ۲، ۱۳۹۸، صص ۴۶۶-۴۴۹.

۱۰) سهرابی، ایمان و اکبرنژاد، حسن (۱۴۰۱). نقش هوش مصنوعی در فرهنگ نوین تمدن اسلامی، جستارنامه فرهنگ و هنر اسلامی، دوره اول، شماره ۲.

۱۱) شامبیاتی، هوشنگ (۱۴۰۲). حقوق جزای عمومی، انتشارات ژوبین، تهران.

۱۲) صادقی، حسین، مسئولیت مدنی واسطه‌ها و تأمین‌کنندگان خدمات ارتباطی الکترونیک، فصلنامه حقوق خصوصی، دوره ۴۰، شماره ۲، ۱۳۸۹، صص ۲۱۸-۱۹۹.

۱۳) صوفی، سارا؛ صالح‌نژاد بهرستاقی، صابر؛ (۱۴۰۲). تأثیر هوش مصنوعی در ارتکاب جرایم سایبری، مجله مطالعات حقوق، شماره ۵۱، دوره ۱۱، سال پنجم، ۱-۱۸.

۱۴) عبداله‌پور، وحید (۱۴۰۴). تحلیل حقوقی و چالش‌های قانونی هوش مصنوعی در پیشگیری و مقابله با جرائم رایانه‌ای و تقویت امنیت سایبری، *فصلنامه پژوهش‌های تطبیقی فقه، حقوق و سیاست*، دوره ۷، شماره ۴، صص. ۱—۱۶.

۱۵) عظیمی، محمدحسن و اسماعیلی، سمیرا (۱۴۰۰). شناسایی مؤلفه‌های هوش مصنوعی در پایگاه‌های اطلاعاتی ایرانی، مجله دانش‌شناسی، دوره ۱۴، شماره ۵۴.

۱۶) فیروزی‌نیا، اکبر. (۱۴۰۱). مسئولیت کیفری اشخاص حقوقی ناشی از ارتکاب جرایم سایبری. در مجموعه مقالات دومین کنفرانس ملی حقوق، فقه و فرهنگ.

۱۷) قایدرحمتی، الهام. (۱۳۹۵). نقش شبکه‌های اجتماعی در جرم‌زدایی، اولین همایش ملی آینده-پژوهی، علوم انسانی و امنیت اجتماعی.

۱۸) کاظمی، محمد. (۱۴۰۴). تأثیر هوش مصنوعی بر ارتکاب جرائم سایبری. در مجموعه مقالات بیست و یکمین کنفرانس بین‌المللی مطالعات حقوقی و قضایی.
۱۹) کاوه، محمدهادی؛ بارانی، محمد (۱۴۰۳). مسؤولیت کیفری هوش مصنوعی در حقوق کیفری ایران با نگاهی به قوانین اتحادیه اروپا، *فصلنامه فقه جزای تطبیقی*، دوره ۴، شماره ۳، صص. ۵۱–۶۱.
۲۰) کرامت میرشکارلو، یاسین، خالقی، ابوالفتح. (۱۴۰۳). مسئولیت کیفری ارتکاب ترور سایبری با فناوری هوش مصنوعی در حقوق کیفری بین‌الملل. در مجموعه مقالات اولین کنفرانس ملی پژوهش‌های نوظهور در مدیریت و حقوق با رویکرد شعار سال.
۲۱) کشاورزصفی‌یی، اسماعیل (۱۴۰۴). مبنای مسئولیت مدنی عوامل مرتبط با عملکرد سیستم‌های هوش مصنوعی در اسناد اتحادیه اروپا و حقوق ایران، *فصلنامه تحقیقات حقوقی*، دوره ۲۸، شماره ۴، صص. ۲۶۷–۲۸۸.
۲۲) محمدی کردخیلی، حمید، مجید صفوی، محمد غلامعلی‌زاده و حمیدرضا اسمعیلی، تبیین ارکان مسئولیت مدنی و شیوه‌های جبران خسارت ناشی از ارتباطات اینترنتی در حقوق ایران و اتحادیه اروپا، ماهنامه جامعه‌شناسی سیاسی ایران، دوره ۴، شماره ۴، ۱۴۰۰، صص ۶۴۷-۶۳۰.
۲۳) معبودیان، فاطمه؛ (۱۴۰۴). مسئولیت کیفری هوش مصنوعی از دیدگاه حقوق جزا و جرم‌شناسی، فصلنامه علمی فقه و حقوق نوین، سال ششم، شماره ۲۳، ۱-۲۰.
۲۴) میرمحمدصادقی، حسین. (۱۴۰۲). جرایم علیه امنیت و آسایش عمومی، انتشارات میزان، تهران.
۲۵) نجاتی، نریمان؛ کلانتری، سعیده و بمانیان، محمدرضا (۱۴۰۰). آموزش طراحی معماری مبتنی بر هوش مصنوعی، مجله معماری نوین، دوره اول، شماره ۱.
۲۶) نظرپور، حمزه، علی‌اکبر ایزدی‌فرد و محمد محسنی دهکلانی، «بررسی فقهی حقوقی مسئولیت مدنی ربات»، فصلنامه فقه و اصول، دوره ۵۳، شماره ۴، ۱۴۰۰، صص ۱۹۰-۱۷۱.
۲۷) نظرپور، محمد؛ موسوی، سیدحسن و حسینی، میرسعید (۱۳۹۹). کاربرد هوش مصنوعی در حسابرسی مالیاتی، دانش حسابرسی، سال ۲۰، شماره ۸۱.
۲۸) ورویی، اکبر. سعادتی، رضا. و هاشمی، حمید (۱۳۹۳). تأثیر اخلاق در جرم‌انگاری و جرم‌زدایی در نظام حقوق کیفری اسلامی ایران، پژوهش حقوق کیفری، شماره ۸.
۲۹) ولی‌پور، علی و محسن اسماعیلی، امکان‌سنجی مسئولیت مدنی هوش‌مصنوعی عمومی ناشی از ایجاد ضرر در حقوق مدنی، فصلنامه اندیشه حقوقی معاصر، دوره ۲، شماره ۳، ۱۴۰۰، صص ۸-۱.
۳۰) قانون اساسی جمهوری اسلامی ایران.
۳۱) قانون جرایم رایانه‌ای مصوب ۱۳۸۸.
۳۲) قانون مجازات اسلامی مصوب ۱۳۹۲ با اصلاحات بعدی.
۳۳) قانون مسئولیت مدنی مصوب ۱۳۳۹.
34) Abbott, R. (2020). *The Reasonable Robot: Artificial Intelligence and the Law*. Cambridge University Press.
35) Adekunle, Y. (2019). Holistic exploration of gaps vis-à-vis artificial intelligence in automated teller machine and internet banking. International Journal of Applied Information Systems,2(25), 12.
36) Allen, H., Surono, A., & Pamungkas, R. (2025). Criminal Liability of Artificial Intelligence Operators in Cybercrime. *Journal of Information Law*, 18(1), 33–58.
37) Chawla, R (2019). “Deepfakes How a pervert shook the world”. International Journal of Advance Research and Development, 4(6): 4–8.

38) Day, C (2019). “The Future of Misinformation”. Journal of Computing in Science & Engineering, 21(1): 108–110.
39) Duran, M., & García, P. (2003). Internet Addiction and Behavioral Consequences. *CyberPsychology & Behavior*, 6(2), 143–151.
40) European Commission. (2022). Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive). Brussels.
41) European Parliament and Council. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation – GDPR).
42) European Parliament and Council. (2022). Directive (EU) 2022/2555 (NIS2 Directive).
43) Fletcher, J (2018). “Deepfakes-Artificial Intelligence and Some Kind of Dystopia The New Faces of Online Post-Fact Performance”. Theatre Journal, 70(4): 455–471.
44) Ghorbanpour Vishkasouqi, S., & Mirazimi Tabalvandani, S. A. (2023). A comprehensive guide to the applications of artificial intelligence in cybersecurity. Proceedings of the 11th National Conference on Novel Ideas in Humanities and Engineering, Rasht, Iran.
45) Guan, Y., & Ge, X. (2017). Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks. IEEE Transactions on Signal and Information Processing over Networks,4(1), 48-59.
46) Hallevy, G. (2015). *Liability for Crimes Involving Artificial Intelligence Systems*. Springer.
47) Hasan, H & Salah, K (2019). “Combating Deepfake Videos Using Blockchain and Smart Contracts”. IEEE Access, 4(7): 41596–41606.
48) Herke, J., & Tóth, B. (2024). Criminal Responsibility and Artificial Intelligence in Autonomous Systems. *European Journal of Crime, Criminal Law and Criminal Justice*, 32(2), 177–201.
49) Kerr, O. S. (2022). *Computer Crime Law* (5th ed.). West Academic Publishing.
50) Keshavarz, Z., & Hosseini, H. R. (2023). Artificial intelligence in cybersecurity: Applications, challenges, and opportunities. Proceedings of the 6th National Conference on New Technologies in Electrical, Computer, and Mechanical Engineering in Iran, Tehran, Iran.
51) Khan, S. (2024). Criminal Liability for Artificial Intelligence Systems in Cybercrime. *International Journal of Law and Information Technology*, 32(1), 45–69.
52) Kowert, W. (2017). The foreseeability of human-artificial intelligence interactions. Texas Law Review,96(1), 181.
53) Maras, M & Alexandrou, A (2019). “Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos”. International Journal of Evidence & Proof, 23(3): 255–262.
54) Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
55) Saud, A. (2024). Artificial Intelligence and Cybercrime: Criminal Law Challenges. *Computer Law & Security Review*, 53, 105912.

56) Turner, J. (2019). *Robot Rules: Regulating Artificial Intelligence*. Palgrave Macmillan.
57) European Parliament and Council. (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act). Official Journal of the European Union.