

The Encyclopedia of Comparative Jurisprudence and Law

An Examination of the Challenges of Determining Criminal Liability in Cases of Harm Caused by Artificial Intelligence Systems and Robots with an Approach Based on Islamic Jurisprudence

Reyhaneh Zare Dehabadi^{1*}, Hosein Shokreian Amiri²

1. Assistant Professor, Department of Fiqh and Principles of Islamic Law, Faculty of Humanities and Social Sciences, Ardakan University, Ardakan, Iran

2. PhD in Jurisprudence and Fundamentals of Islamic Law, University of Mazandaran, Babolsar, Iran

* Corresponding Author's Email: r.zaredehabadi@ardakan.ac.ir

ABSTRACT

The integration of artificial intelligence and autonomous systems into decision-making domains has posed a serious challenge to the traditional concepts of duty and liability in Imamiyya jurisprudence. A jurisprudential system grounded in human intention and volition must now confront an agent that lacks human consciousness and intent, yet is capable in practice of generating legal and criminal consequences. Drawing upon the concept of "meta-obligation," this study attempts to present a novel model of jurisprudential order in the post-artificial intelligence era; an order founded upon distributive liability between human beings and technology. Within this framework, the concepts of causation, justice, and legal personality are redefined on the basis of networked relations. The jurisprudential foundations of criminal liability regarding damages caused by artificial intelligence are grounded in the preservation of human social order and justice; through the principle of public interest, the consequences of technology are assessed and harm is prevented. Reason serves as the criterion for determining the extent of damage and proportional punishment, while the principles of no-harm and destruction guarantee compensation for losses. The findings indicate that Imamiyya jurisprudence possesses the capacity to transition from an individual-centered model toward a technological and network-oriented system; an order in which "meta-obligation" establishes a connection between human rationality and technological intelligence, and justice is conceptualized through balance among humans, data, and automated decision-making.

Keywords: *Meta-obligation; post-artificial intelligence jurisprudential order; distributive liability; network justice; technological ijthad.*

How to cite: Zare Dehabadi, R. & Shokreian Amiri, H. (2027). An Examination of the Challenges of Determining Criminal Liability in Cases of Harm Caused by Artificial Intelligence Systems and Robots with an Approach Based on Islamic Jurisprudence. *The Encyclopedia of Comparative Jurisprudence and Law*, 5(3), 1-17.



تاریخ ارسال: ۱ دی ۱۴۰۴

تاریخ بازنگری: ۸ اردیبهشت ۱۴۰۵

تاریخ پذیرش: ۱۶ اردیبهشت ۱۴۰۵

تاریخ چاپ اولیه: ۱ مرداد ۱۴۰۵

تاریخ چاپ نهایی: ۱ شهریور ۱۴۰۶

دانشنامه فقه و حقوق تطبیقی

واکاوی چالش های تعیین مسئولیت کیفری در صورت ورود صدمه به افراد توسط سیستم های هوش مصنوعی و ربات ها با رویکردی بر فقه

ریحانه زارع ده آبادی^{۱*}، حسین شکران امیری^۲

۱. استادیار، گروه فقه و مبانی حقوق اسلامی، دانشکده علوم انسانی و اجتماعی، دانشگاه اردکان، اردکان، ایران

۲. دکتری فقه و مبانی حقوق اسلامی، دانشگاه مازندران، بابلسر، ایران

* پست الکترونیک نویسنده مسئول: r.zaredehabadi@ardakan.ac.ir

چکیده

ورود هوش مصنوعی و سامانه های خودکار به حوزه های تصمیم گیری، مفاهیم سنتی تکلیف و مسئولیت در فقه امامیه را با چالشی جدی روبه رو کرده است. فقهی که بر اراده و قصد انسان استوار است، اکنون باید با عاملی مواجه شود که فاقد شعور و نیت انسانی است اما در واقعیت می تواند منشأ آثار حقوقی و کیفری باشد. این پژوهش، با تکیه بر مفهوم «فرا تکلیف»، کوشیده است الگویی تازه از نظم فقهی در دوران پسا هوش مصنوعی ارائه دهد؛ نظمی مبتنی بر مسئولیت توزیعی میان انسان و فناوری. در این چارچوب، مفاهیم سببیت، عدالت، و شخصیت حقوقی، بر اساس ارتباط شبکه ای باز تعریف می شوند. مبانی فقهی مسئولیت کیفری در برابر خسارات هوش مصنوعی بر حفظ نظام بشری و عدالت استوار است؛ با قاعده مصلحت، آثار فناوری سنجیده و از ضرر جلوگیری می شود. عقل میزان خسارت و مجازات متناسب را تعیین می کند و قاعده لاضرر و اتلاف نیز جبران خسارت را تضمین می نماید. نتیجه آن است که فقه امامیه قابلیت گذار از الگوی فردمحور به نظامی فناورانه و شبکه ای را دارد؛ نظمی که در آن «فرا تکلیف» پیوند میان عقل انسانی و خرد فناورانه را برقرار می سازد و عدالت را در توازن میان انسان، داده و تصمیم خودکار معنا می کند.

کلیدواژگان: فرا تکلیف؛ نظم فقهی پسا هوش مصنوعی؛ مسئولیت توزیعی؛ عدالت شبکه ای؛ اجتهاد فناورانه.

نحوه استناددهی: زارع ده آبادی، ریحانه، و شکران امیری، حسین. (۱۴۰۶). واکاوی چالش های تعیین مسئولیت کیفری در صورت ورود صدمه به افراد توسط سیستم های هوش مصنوعی و ربات ها با رویکردی بر فقه. *دانشنامه فقه و حقوق تطبیقی*، ۵(۳)، ۱۷-۱.

می‌کند. بر این اساس، مسئولیت حقوقی نه صرفاً در نقطه‌ی اراده‌ی انسانی، بلکه در زنجیره‌ای از تصمیم‌های انسانی و الگوریتمی معنا می‌یابد. این بازتعریف، زمینه‌ی پیدایش نظم فقهی پساہوش مصنوعی را فراهم می‌کند؛ نظمی که عدالت را نه به صورت انتزاعی یا فردی، بلکه در قالب توازن میان انسان، داده و تصمیم خودکار بازمی‌سازد.

پژوهش حاضر تلاش دارد این تغییر بنیادین را در چارچوب مبانی فقه امامیه بررسی کند و نشان دهد که اجتهاد اسلامی ظرفیت درک شرایط جدید و تولید پاسخ‌های نو را داراست. تلفیق عقلانیت فقهی با منطق شبکه‌ای فناوری، گامی در جهت حفظ روح شریعت در جهان هوشمند است. در این جهان، مفاهیمی مانند سببیت، عمد، تقصیر و عدالت، بازتعریف می‌شوند تا هم با اصول شرعی سازگار بمانند و هم نزدیکی بیشتری با واقعیت‌های عصر داده و الگوریتم پیدا کنند.

در نهایت، نوشتار حاضر در پی آن است که از دل این تأمل، مبانی نظری «نظام فراتکلیف» را ترسیم کند؛ نظامی که می‌تواند پیوند میان انسان و فناوری را در قالبی فقهی، مسئولانه و اخلاق‌محور سامان دهد. بدین ترتیب، فقه امامیه نه تنها از مواجهه با هوش مصنوعی دچار گسست نمی‌شود، بلکه از این مواجهه فرصتی تازه می‌سازد تا مفاهیم خود را در افقی گسترده‌تر و متناسب با جهان پساہوش مصنوعی بازسازی کند.

مفاهیم

مسئولیت کیفری

مسئولیت کیفری در قلمرو حقوق جزا به وضعیتی اطلاق می‌شود که در آن، شخص به لحاظ انجام رفتار مجرمانه، از سوی قانون‌گذار شایسته‌ی انتساب و تحمل نتایج کیفری دانسته می‌شود. به تعبیر دیگر، این مسئولیت بیانگر اهلیت فرد برای پاسخ‌گویی جزایی و پذیرش مجازات در برابر افعالی است که با اراده و اختیار خویش مرتکب شده است (Mirsaeedi, 2007). بنابراین، اساس و

تحولات پرشتاب هوش مصنوعی و فناوری‌های خودکار، در دهه‌ی اخیر یکی از عمیق‌ترین دگرگونی‌های معرفتی بشر را رقم زده‌اند. هوش مصنوعی دیگر صرفاً ابزار پردازش داده نیست، بلکه به عامل تصمیم‌گیر با قدرت اثرگذاری در عرصه‌های اجتماعی، اقتصادی و حتی اخلاقی بدل شده است. تصمیمات خودکار، تعاملات در شبکه‌های داده‌ای و کنش‌های هوشمند در نظام‌های حقوقی و کیفری، به تدریج جایگاه انسان را از «فاعل مستقل» به «عنصر همکار در یک شبکه‌ی هوشمند تصمیم‌گیری» تغییر داده‌اند. این دگرگونی، به‌ویژه در حوزه‌ی فقه و حقوق اسلامی، پرسش‌هایی بنیادین پدید آورده است: مسئولیت در جهان هوشمند چگونه تعریف می‌شود؟ فاعل بی‌اراده چگونه ممکن است سبب جرم تلقی شود؟ عدالت در نظامی که تصمیمات به صورت الگوریتمی اتخاذ می‌شوند، چه معنا و مبنایی دارد؟

فقه امامیه، که ریشه‌ای عمیق در عقلانیت و اجتهاد دارد، همواره توانایی انطباق با شرایط نو را از خود نشان داده است. این فقه بر سه رکن اراده، قصد و عقل انسانی استوار است و بر پایه‌ی همین ویژگی‌ها، نظام مسئولیت کیفری و مدنی خود را شکل داده است. با این حال، ورود به عصر پساہوش مصنوعی، این بنیان‌ها را در معرض پرسش قرار می‌دهد. هرچند عامل فناورانه فاقد قصد و نیت انسانی است، اما در عمل می‌تواند منشأ کنش‌هایی شود که پیامدهای حقیقی دارند؛ از آسیب به انسان گرفته تا تخلف در جریان داده‌های اخلاقی و مالی. این وضعیت، مرز میان «سبب انسانی» و «تسبیب فناورانه» را مبهم ساخته و نظریه‌های کلاسیک مسئولیت را به بازنگری فرا می‌خواند.

در چنین زمینه‌ای، مفهوم «فراتکلیف» به عنوان طرحی نظری برای گذار از انسان‌محوری مطلق به نظامی شبکه‌ای از مسئولیت مطرح می‌شود. فراتکلیف، به جای محدود ساختن تکلیف به فرد دارنده‌ی اراده و اختیار، آن را در گستره‌ی تعامل میان انسان و فناوری توزیع

مبنای تمام اقسام مسئولیت کیفری، وجود اختیار و آزادی در رفتار است (Jafari, 2016).

در فقه امامیه، مسئولیت کیفری بر مبنای اختیار و قابلیت انتساب فعل به فاعل استوار است؛ بدین معنا که تحقق کیفر در حق شخص منوط به آن است که رفتار مجرمانه از او به صورت ارادی و آگاهانه سر زده باشد (Khomeini, 2000). در فقه امامیه، بنیان مسئولیت کیفری با قواعد عام ضمان پیوند دارد؛ از جمله قاعده‌ی اتلاف، تسبیب و غرور. هرچند این قواعد غالباً در حوزه‌ی ضمان مالی کاربرد دارند، اما منطق مشترک انتساب فعل زیان‌بار به فاعل مختار، همان مبنای عقلی مسئولیت در قلمرو کیفری است (Mohaghegh Damad, 1999). در قلمرو حقوق جزای اسلامی نیز اراده و آگاهی، رکن تحقق مسئولیت کیفری شناخته می‌شوند. چنان‌که هرگاه رفتار بدون اراده‌ی معتبر، مانند اجبار، اکراه، جنون یا جهل به حرمت عمل، واقع گردد، انتساب کیفری از مرتکب منتفی است.

هوش مصنوعی

هوش مصنوعی، شاخه‌ای میان‌رشته‌ای از علوم رایانه است که هدف اصلی آن طراحی و تولید ماشین‌هایی با توانایی فکری و شناختی مشابه انسان می‌باشد؛ ماشین‌هایی که قادرند کارکردهایی چون سخن گفتن، یادگیری، استدلال و حل مسئله را به صورت خودکار انجام دهند.

در تعریف دقیق‌تر، هوش مصنوعی به سامانه‌هایی اطلاق می‌شود که می‌توانند در شرایط پیچیده، رفتاری مشابه انسان از خود بروز دهند؛ از جمله: درک موقعیت‌های متغیر، شبیه‌سازی فرایند تفکر و روش‌های استدلال انسانی، یادگیری از تجربه و بهره‌گیری از دانش برای حل مسائل جدید. به بیان دیگر، هوش مصنوعی نشان‌دهنده‌ی سطحی از هوش است که ماشین در مواجهه با موقعیت‌های مختلف از خود بروز می‌دهد.

اغلب پژوهش‌ها و نوشته‌های علمی در این حوزه، هوش مصنوعی را دانش مطالعه، شناخت و طراحی «عامل‌های هوشمند» می‌دانند؛ عامل‌هایی که قادرند بر اساس داده‌ها تصمیم بگیرند و رفتار خود را متناسب با محیط اصلاح کنند. به همین دلیل، متخصصان این حوزه می‌کوشند با بهره‌گیری از علوم مختلف، از جمله ریاضیات، فیزیک، زیست‌شناسی و رایانه، توانایی‌های ذهنی انسان را بازسازی و شبیه‌سازی نمایند تا از دستاوردهای آن در جهت پیشرفت انسان و خدمت به جوامع بشری بهره‌برداری کنند (Vaseghi, 2020).

ربات

واژه‌ی «ربات» به معنای «کار اجباری یا کارگرانه» ریشه گرفته است. در معنای علمی امروز، ربات موجودی مصنوع است که با ترکیب سخت‌افزار مکانیکی، حسگر، سامانه‌ی پردازش داده و عملگرها می‌تواند بخشی از افعال انسانی را به صورت خودکار یا هوشمند انجام دهد (Ghafouri, 2019).

ارکان مسئولیت کیفری

فقه‌ای امامیه برای تحقق مسئولیت کیفری سه رکن اساسی قائل‌اند که مبنای قانون مجازات اسلامی نیز به شمار می‌رود:

الف) عنصر قانونی «نص شرعی»

در فقه جزایی، بنابر اصول مسلمی همچون قبح عقاب بلا بیان و اصل برائت، مجازات هیچ فعل یا ترک فعلی جایز نیست مگر آنکه دلیل یا نص شرعی معتبر بر حرمت و مجازات آن وجود داشته باشد. محقق حلی در شرایع الإسلام تصریح می‌کند: «الْحُدُودُ تَقْتَرُ إِلَى الشَّرْعِ، وَلَا يُعَاقَبُ الْعَبْدُ إِلَّا بِمَا وَرَدَ فِيهِ النَّصُّ» (Muhaqqiq Hilli, 1988).

بر پایه‌ی این مبنا، تا زمانی که حکم صریح شرعی یا قانونی جدیدی در خصوص اعمال ربات یا سامانه‌های هوش مصنوعی صادر نشده باشد، رفتارهای آنان در قلمرو جرم‌انگاری فقهی قرار نمی‌گیرد و مشمول کیفر شرعی نخواهد بود.

ب) عنصر مادی «کردار مجرمانه»

عنصر مادی به فعل یا ترک فعل خارجی گفته می‌شود که موجب تحقق جرم گردد. در مباحث فقهی، غالباً از دو عنوان مباشرت و تسبیب برای توصیف چگونگی تحقق این عنصر استفاده می‌شود. مباشر، کسی است که فعل زیان‌بار را رأساً انجام دهد؛ اما سبب، کسی است که مقدمات فعل زیان‌بار را فراهم می‌سازد، بدون اینکه خود آن را مستقیماً انجام دهد.

فقه‌ها در «جواهرالکلام» و «تحریرالوسیله» قاعده کرده‌اند که اگر سبب اقوی از مباشر باشد، ضمان به سبب منتقل می‌شود (Najafi, 1984).

این قاعده در مورد ربات‌ها بسیار مهم است؛ زیرا ربات نه مباشر ذی‌شعور است و نه منشأ قصد، پس فعل او از قبیل سببیت غیرعقل تلقی می‌شود.

ج) عنصر معنوی «قصد و اختیار»

از نظر فقهی، عنصر معنوی رکن تعیین‌کننده در انتساب جرم است. فقه‌ها، «قصد» «نیت فعل حرام» و «اختیار» «اراده در انجام آن» را شرط صحت عقاب دانسته‌اند. امام خمینی در مبحث قتل عمد می‌گوید: «وَيَتَحَقَّقُ الْقَتْلُ الْعَمْدُ إِذَا وَقَعَ عَنْ قَصْدٍ وَاخْتِيَارٍ مِنَ الْعَاقِلِ الْبَالِغِ»؛ یعنی فعل جزایی تنها در صورتی عمدی محسوب می‌شود که از شخصی عاقل، بالغ و مختار صادر گردد (Khomeini, 2000). لذا، در فقه امامیه، هر موجود فاقد عقل یا اراده، اهلیت انتساب کیفری ندارد. از این قاعده، نتیجه‌ی فقهی روشن به دست می‌آید: ربات یا هوش مصنوعی، حتی اگر رفتار ظاهراً عمدی انجام دهد، فاقد قصد واقعی است؛ چون شعور تکلیفی و اراده‌ی مستقل ندارد.

مباشرت و تسبیب

در فقه جزایی اسلام، انتساب هر رفتار مجرمانه تنها زمانی صحیح و معتبر است که میان آن رفتار و فاعل، رابطه‌ای علی و قابل استناد از منظر شرع یا عرف برقرار باشد. این پیوند سببی، اساس مسئولیت

و ضمان را تشکیل می‌دهد و در نظام فقه امامیه به دو شیوه‌ی متمایز تحقق می‌یابد: ۱. مباشرت ۲. تسبیب. تشخیص اینکه فعل در زمره‌ی مباشرت است یا تسبیب، در تعیین مسئولیت کیفری و ضمان تفاوت اساسی ایجاد می‌کند. این تمایز مبنای انتقال مسئولیت است. عناصر تحقق مباشرت عبارت است از: ارتباط مستقیم بین فعل و نتیجه؛ وجود اراده و اختیار در فاعل؛ فقدان واسطه‌ی مؤثر مستقل (Sadeghi, 1981). با تطبیق این شرایط بر رفتار ربات یا سامانه‌ی هوش مصنوعی، روشن می‌شود که: ربات فاقد ادراک و اراده‌ی انسانی «قصد» است، فاقد اختیار آگاهانه نسبت به نتیجه‌ی عمل خویش است و عمل او همواره متکی به علل و داده‌های پیش‌فرض ساخته‌شده توسط انسان است. بنابراین، ربات هرچند ابزار انجام عمل است، اما فاعل مباشر کیفری محسوب نمی‌شود. در فقه امامیه، همانند مورد دیوانه، حیوان یا ابزار غیرعقل، فعل ربات به خود او منسوب نمی‌گردد، بلکه در نهایت به انسان علت‌ساز بازمی‌گردد. زمانی که فرد انسانی از طریق طراحی الگوریتم، آموزش سیستم یا صدور فرمان باعث می‌شود ربات مرتکب فعل زیان‌بار شود، عمل او در فقه از مصادیق تسبیب است؛ یعنی ایجاد علت واسطه «سیستم یا ربات» برای تحقق اثر. به حکم قاعده‌ی «السبب اقوی من المباشرة»، اگر نتیجه‌ی زیان‌بار تنها به واسطه‌ی برنامه‌ریزی یا هدایت انسان تحقق یافته باشد، سبب انسانی ضامن است، نه ربات. رابطه‌ی انسان و ربات دقیقاً به رابطه‌ی «سبب و مباشر» شباهت دارد، با این تفاوت که مباشر در اینجا غیرعقل و غیرمختار است.

مبانی فقهی مسئولیت کیفری در صورت ورود صدمه به افراد توسط سیستم‌های هوش مصنوعی و ربات‌ها

حفظ نظام بشری

بدیهی است که وقتی «نظام» به عنوان عام‌ترین مفهوم سامان‌بخش حیات اجتماعی تعریف می‌شود، مسئولیت حقوقی نیز باید بر مبنای همان منطق آراستگی درونی جامعه و نهادهای آن، پیش‌بینی

وصفی» و «عدالت عملی» راه را برای تحلیل مسئولیت کیفری در حوزه‌ی سامانه‌های هوش مصنوعی و ربات‌ها روشن می‌کند (Mohaghegh Damad, 1999).

در سطح وصفی، عدالت ایجاب می‌کند که تصمیم‌گیری و عملکرد سامانه‌های هوشمند از نظر طراحی، داده‌های ورودی، معیارهای ارزیابی و سازوکارهای کنترل، واجد معیارهای سلامت و مشروعیت باشند تا از «فسق»، به معنای خروج از چارچوب درست، منحرف نشوند. بنابراین، اگر الگوریتم یا منطق حاکم بر ربات به شکل غیرمنصفانه یا غیرمنطبق با معیارهای لازم عمل کند و زمینه‌ی صدمه به افراد فراهم شود، مسئله صرفاً یک حادثه نیست، بلکه تخلف از عدالت وصفی است و می‌تواند مبنای مداخله‌ی کیفری قرار گیرد.

در سطح عملی نیز، عدالت در برابر جور قرار می‌گیرد؛ یعنی وقتی صدمه‌ای رخ می‌دهد، نظام کیفری باید پاسخ حقوقی متناسبی ارائه دهد که از بی‌سامانی و نابرابری در پیامدها جلوگیری کند و «حق آسبیده» را در عمل بازسازی کند (Mohaghegh Damad, 1999). از این منظر، مسئولیت کیفری نه‌تنها برای سرزنش صرف، بلکه برای تحقق عدالت عملی است: تعیین اینکه چه کسی در زنجیره‌ی طراحی، تولید، بهره‌برداری یا نظارت، به‌علت قصور یا بی‌احتیاطی، شرایط جور را ایجاد کرده است، تا تکرار صدمات ناشی از هوش مصنوعی و ربات مهار شود.

قاعده‌ی التعزیر لکل عمل محرم

قاعده‌ی «التعزیر لکل عمل محرم» این اصل را بیان می‌کند که هر رفتاری که ممنوع و زیان‌بار باشد، اگر مشمول حدود نشود، می‌تواند در قلمرو تعزیر قرار گیرد. با در نظر گرفتن این قاعده و با توجه به آنچه در معنای لغوی تعزیر آمده است، یعنی نکوهش، توانمند ساختن، نوعی پیروزی دادن از طریق زبان یا شمشیر، ضرب و تنبیهی کمتر از حد، تأدیب، شدیدترین گونه‌های مجازات، سیاست و عقوبت (Dehkhoda, 1946)، می‌توان مسئولیت

ارتباط معقول میان ساختارها و فراهم‌سازی بستر اجرای احکام الهی تحلیل گردد (Tabatabaei Yazdi, 1991). در این چارچوب، ورود صدمه به افراد از سوی سیستم‌های هوش مصنوعی و ربات‌ها تنها یک خطای فنی یا رخداد پراکنده تلقی نمی‌شود، بلکه می‌تواند برهم‌زننده‌ی نظم عمومی، ایجاد اختلال در آرامش اجتماعی و به‌هم‌ریختن «این‌همانی جای هر چیز» در روابط حقوقی و نهادی جامعه باشد؛ ازاین‌رو، واکنش کیفری در جهت بازگرداندن تعادل و جلوگیری از هرج‌ومرج ناشی از فناوری، واجد معنا و ضرورت است.

از سوی دیگر، چون در نظریه‌ی فوق، «کمال و سعادت انسان‌ها»، که همان قرب و منزلت نزد خداوند است، هدف برتر نظام دانسته شده و تحقق آن مشروط به آرامش و سامان زندگی مردم است، مسئولیت کیفری باید تضمین کند که رفتارهای مضر ربات‌ها و سامانه‌های هوشمند به‌گونه‌ای کنترل و پاسخ داده شود که حقوق بنیادین افراد تضعیف نگردد و امکان اجرای قواعد و تکالیف اجتماعی از مسیر صحیح خود منحرف نشود (Tabatabaei Yazdi, 1991). بنابراین، اگر صدمه‌ای واقع شود، نظام کیفری باید با تعیین دقیق مسئولیت، اعم از سطح بهره‌بردار، سازنده، برنامه‌ریز یا اداره‌کننده‌ی سامانه، از تداوم یا تکرار زیان جلوگیری کند و میان «آزادی عمل فناوری» و «الزام به نظم و عدالت» توازن برقرار سازد. در نهایت، پیگیری کیفری، ابزاری برای آماده‌سازی محیط جامعه جهت اجرای احکام و تأمین امنیت روانی و اجتماعی است تا انسان‌ها با آسودگی به مسیر سعادت نزدیک شوند.

اصل عدالت

عدالت، در نسبت با دو مفهوم «فسق» و «جور» معنا می‌یابد؛ به‌این‌ترتیب، هنگامی که رفتار یا سازوکاری به سمت فسق میل می‌کند، از عدالت فاصله می‌گیرد، و در مقابل جور نیز، عدالت خود را به شکل تحقق عینی نشان می‌دهد. از همین رو، هرچند فسق نیز می‌تواند جلوه‌ای از «جور» باشد، اما تفکیک میان «عدالت

نگرش، ساختار حقوق جزای اسلامی بر سنجش مستمر مصالح و مفاسد استوار است؛ یعنی جرم‌انگاری و نهی شارع زمانی تحقق می‌یابد که فعل یا ترک فعل، منجر به زیان، آشفتگی در نظم عمومی، لطمه به باورها، تعرض به جان، مال، حیثیت، احساس امنیت یا سایر ارکان حیاتی حیات اجتماعی شود؛ اموری که بقای جامعه و پرهیز از تفریط و تجاوز در آن‌ها، ضرورت دارد (Awdeh, 1993).

در حوزه‌ی فناوری‌های نوین، هرگاه سامانه‌های هوش مصنوعی یا ربات‌ها سبب ورود صدمه به افراد شوند، مصلحت ایجاب می‌کند که این وضعیت صرفاً به‌عنوان خطای فنی خنثی تلقی نشود؛ بلکه در پرتو آثار واقعی آن بر امنیت، مال، جان و کرامت افراد ارزیابی شود. اگر استفاده، طراحی یا بهره‌برداری از این فناوری‌ها به‌گونه‌ای است که احتمال ضرر و اختلال در جامعه را بالا می‌برد، مقتضای مصلحت آن است که نظام کیفری با تعیین مسئولیت برای طراحان، تولیدکنندگان، بهره‌برداران یا نهادهای ناظر، از این مفاسد پیشگیری کند و ضمانت‌اجرای متناسب پیش‌بینی نماید.

بدین ترتیب، جرم‌انگاری یا تشدید و تخفیف مجازات در زمینه‌ی صدمات ناشی از هوش مصنوعی، باید در راستای تأمین مصالح اجتماعی، استقرار نظم و حمایت از جامعه‌ای نیرومند، متعهد و متخلق به فضائل اخلاقی تفسیر گردد، تا فناوری به ابزاری در خدمت مصلحت جمعی بدل شود، نه منشأ مفسده و ناامنی.

حکم عقل

عقل، به‌عنوان هبه‌ای بزرگ از جانب خداوند، اصلی‌ترین منبع درک و فهم حقایق و همچنین راهنمایی انسان در مسیری است که به سعادت می‌انجامد (Mohaghegh Damad, 2006; Mousavi Bojnourdi, 1996). از این رو، در نظام حقوقی، خصوصاً حوزه‌ی مسئولیت کیفری، نقش عقل در ارزیابی و تشخیص صدمات وارده توسط سیستم‌های هوش مصنوعی و ربات‌ها اهمیت ویژه‌ای دارد؛ زیرا عقل انسان است که می‌تواند

کیفری در صدمات ناشی از ربات‌ها و سامانه‌های هوش مصنوعی را تبیین کرد.

وقتی سامانه‌ای هوشمند مرتکب رفتاری می‌شود که به زیان اشخاص می‌انجامد، چه در قالب خطای پردازشی، نقص در برنامه‌ریزی یا عملکرد نایمن، این رفتار در اصل به منبع انسانی آن بازمی‌گردد؛ زیرا تصمیم‌گیری، طراحی، آموزش و بهره‌برداری از چنین سیستم‌هایی زیر نظر اشخاص حقیقی یا حقوقی صورت می‌گیرد. بر همین اساس، چنانچه عملی که منجر به صدمه شده، در شمار افعالی باشد که قانون آن را نامشروع می‌داند، تعزیر می‌تواند ابزار واکنشی برای نكوهش و تأدیب اشخاص مسئول باشد. این واکنش، بنا بر ماهیت تعزیر، انعطاف‌پذیر است و اجازه می‌دهد مجازات متناسب با نوع قصور، شدت آسیب و نقش هر یک از عوامل انسانی تعیین شود.

از آنجا که تعزیر علاوه بر جنبه‌ی تنبیهی، متضمن «سیاست» و «توانا ساختن» نیز هست، نظام کیفری می‌تواند از این ظرفیت بهره گیرد تا علاوه بر سرزنش و عقوبت افراد خاطی، سازوکارهای ایمنی و نظارتی در حوزه‌ی هوش مصنوعی تقویت شود. در واقع، اعمال تعزیر پاسخی است برای جلوگیری از رواج رفتارهای ممنوع و بی‌احتیاطی‌های فناورانه و در عین حال، ابزاری برای سامان‌دهی و قدرت‌بخشی به ساختارهای نظارتی، تا از تکرار صدمات مشابه جلوگیری شود.

قاعده‌ی مصلحت

قاعده‌ی مصلحت، با گستره‌ی مفهومی وسیعی که دارد و با عنوان‌هایی چون نبود مفسده، ضرورت، منفعت، حقیقت، حسن و قبح عقلی، استحسان، سدّ و فتح ذریعه، عرف، مقاصد شریعت، علت و حکمت حکم درهم‌تنیده است (Dehkhoda, 1946; Ibn Manzur, 1988; Shartouni, 1983)، چارچوبی کارآمد برای تحلیل مسئولیت کیفری در قبال صدمات ناشی از سیستم‌های هوش مصنوعی و ربات‌ها فراهم می‌کند. بر پایه‌ی این

تکلیفی یا وضعی که عاقبت آن ضرر باشد، صادر نکرده است (Ansari).

در مواجهه با فناوری‌های نوین مانند سیستم‌های هوش مصنوعی و ربات‌ها، این قاعده حکم می‌کند که هیچ‌گونه مسئولیت کیفری برای اشخاص یا نهادهایی که سبب ورود خسارت و آسیب به افراد می‌شوند، در صورتی که این خسارت بر اثر عملکرد نادرست، نقص استانداردها یا سوءاستفاده باشد، نمی‌تواند بر پایه‌ی احکام شرعی یا قانونی مشروع تلقی گردد. بنابراین، هرگونه قصور، بی‌احتیاطی یا خطای فنی در استفاده یا توسعه‌ی این فناوری‌ها که منجر به صدمه به افراد شود، باید به‌گونه‌ای بررسی و اقدامات متناسب در راستای عدم ورود ضرر و تضمین حقوق زیان‌دیدگان صورت گیرد، تا از انحراف مبانی «لا ضرر» و رعایت اصل عدم آسیب در نظام کیفری پیروی شود.

قاعده‌ی اتلاف

قاعده‌ی اتلاف در فقه، به معنای از بین بردن مال دیگری است و برای تحقق این فعل، باید شخص مرتکب از قبل آگاه و هدفمند در عمل انجام‌شده باشد (Jafari Langroudi, 2006). بر اساس آموزه‌های قرآنی، هرگاه شخصی مال دیگری را تلف کند، مسئول جبران آن است و در این باره، از اصطلاح «مثل» یا قیمتی برای پرداخت دیه یا غرامت استفاده می‌شود (Hurr Amili, 1989). علاوه بر این، مشهور است که «مَنْ اتْلَفَ مَالَ الْغَيْرِ فَهُوَ لِهَ ضَامِنٌ»، و این نشان می‌دهد که تلف مال دیگران، تابع ضمان و مسئولیت است (Najafi, 1984). این قاعده هم شامل اتلاف واقعی است؛ یعنی نابودی فیزیکی مال، مانند سوزاندن کالا، و هم اتلاف حکمی؛ یعنی نابودی ارزش مال در نتیجه‌ی اقداماتی مثل ممانعت از فروش مال فاسد و در نتیجه، کاهش ارزش آن است. در موارد اتلاف، ضمان ممکن است حقیقی باشد؛ یعنی تلف کالا به‌صورت مستقیم، یا حکمی؛ یعنی کاهش ارزش آن بدون تلف فیزیکی. در زمینه‌ی فناوری‌های نوین، مانند سیستم‌های هوش

میزان و نوع آسیب‌ها، وظایف و تخلفات مرتبط با طراحی، بهره‌برداری و کنترل این فناوری‌ها را تشخیص دهد. در مواردی که فعل یا ترک فعل سیستم‌ها سبب صدمه به اشخاص شود، عقل مصداقی از منابع کشف احکام شریعت است و می‌تواند راهنمای تعیین مسئول و میزان مسئولیت باشد.

عقل در کنار کتاب و سنت، ابزاری است برای تحلیل صحیح و منطقی پیامدهای استفاده از فناوری‌های نوین و کمک می‌کند تا در قوانین کیفری، میزان و نوع مجازات در راستای تأمین مصالح اجتماعی و حفظ نظم قرار گیرد (Mousavi Bojnourdi, 1996). در اینجا، عقل نقشی کلیدی در تشخیص تقصیر، تخطی و تعیین میزان ضرر دارد و می‌تواند پایه و اساس وضع مجازات‌های تعزیری یا حتی حدی باشد. بنابراین، در مواجهه با خسارات ناشی از هوش مصنوعی و ربات‌ها، بهره‌گیری از حکمت و قوه‌ی عاقله در تعیین مسئولیت، چراغی است که هم نقش راهنما دارد و هم ابزاری است برای اصلاح و بازدارندگی، تا فناوری به خدمت مصالح و عدالت درآید و از آسیب‌ها و نواقص جلوگیری شود.

قاعده‌ی لا ضرر

قاعده‌ی «لا ضرر» یکی از اصول بنیادین و قطعی در فقه اسلام است که بر اصل عدم وارد آوردن ضرر و ضرار به دیگران تأکید می‌کند و بر پایه‌ی روایت پیامبر «ص»: «لا ضرر و لا ضرار فی الدّین أو فی الإسلام» است (Rashti Najafi). بر اساس این قاعده، در نظام حقوقی اسلام، احکام و دستورات شرعی به‌گونه‌ای طراحی شده است که منجر به ضرر و زیان فرد یا جامعه نشود، و هر حکم یا عمل شرعی که موجب آسیب، ضرر یا اختلال در حقوق و مصالح عمومی باشد، مردود است (Jafari Tabrizi). از این رو، هر عملی که لزوماً منجر به خسارت یا نقصان در حقوق دیگران گردد، در اسلام قابل تشریع نمی‌باشد و شارع هیچ حکم

که مرتکب را به سوی ارتکاب رفتار ممنوع سوق می‌دهد (Mansourabadi, 2016).

در نظام‌های حقوقی، یکی از مباحث اساسی در خصوص هوش مصنوعی، امکان انتساب نیت و قصد مجرمانه به سامانه‌های هوشمند است. در واقع، مسئله‌ای که مطرح می‌شود آن است که اگر یک ربات یا الگوریتم یادگیرنده موجب ایراد صدمه، قتل یا خسارت گردد، آیا می‌توان گفت که وی واجد سوءنیت بوده و رفتار او واجد رکن معنوی جرم است. بررسی این موضوع نشان می‌دهد که پاسخ منفی است؛ زیرا مفهوم رکن معنوی جرم، بر سه عنصر بنیادین استوار است که در ساختار مصنوعی هوش مصنوعی تحقق نمی‌یابد.

الف. عنصر آگاهی

در حقوق کیفری، آگاهی به منزله‌ی ادراک نتایج رفتار و علم به حرمت یا غیرقانونی بودن آن تلقی می‌شود (Zeraat, 2017). سامانه‌های هوشمند اگرچه قادر به تحلیل حجم وسیعی از داده‌ها و ارائه‌ی پاسخ‌های منطقی هستند، از توانایی درک ارزشی و اخلاقی بی‌بهره‌اند. تصمیم‌گیری آن‌ها بر پایه‌ی محاسبات الگوریتمی صورت می‌گیرد. از این رو، مفاهیمی چون خوب و بد، مشروع و نامشروع، برای آن‌ها قابل تصور نیست.

ب. عنصر اراده و اختیار

مسئولیت جزایی یا تکلیف شرعی، هنگامی تحقق می‌یابد که فاعل از قدرت تشخیص و تصمیم‌گیری ارادی برخوردار باشد (Sadeghi, 1981). در مقابل، رفتار سامانه‌های هوشمند همواره در محدوده‌ی الگوریتم‌ها و دستورالعمل‌های از پیش تعیین شده جریان دارد و حتی در فرآیند یادگیری عمیق نیز چارچوب عمل آن‌ها توسط طراح یا برنامه‌نویس مشخص می‌شود. بدین ترتیب، تصمیمات صادره از سوی هوش مصنوعی را نمی‌توان جلوه‌ای از اراده‌ی آزاد دانست؛ بلکه بازتابی از واکنش‌های داده‌محور است.

مصنوعی و ربات‌ها، اگر آسیب یا خسارتی وارد آید، باید بر اساس اصل ضمان، مسئولیت تعیین شود. در صورت ایجاد ضرر، هرگونه کوتاهی در طراحی، بهره‌برداری یا نگهداری سیستم‌های هوشمند می‌تواند منجر به مسئولیت کیفری گردد؛ چه در صورت تلف فیزیکی و چه در صورت کاهش ارزش مال، و باید جبران خسارت به گونه‌ای باشد که حق زیان‌دیده حفظ شود (Sami, 2021).

مهم‌ترین چالش‌ها

فقدان شخصیت کیفری مستقل برای هوش مصنوعی

در نظام‌های حقوقی فعلی، مسئولیت کیفری صرفاً به اشخاص حقیقی و حقوقی محدود است و سامانه‌های هوشمند و ربات‌های خودمختار، به دلیل فقدان شخصیت کیفری و صلاحیت قانونی، نمی‌توانند فاعل جرم شناخته شوند. قوانین موجود، تنها انسان‌ها و شرکت‌ها را مسئول می‌دانند و نظریه‌ی شخصیت الکترونیکی محدود نیز با مخالفت مواجه شده است. در فقه امامیه، مسئولیت بر «عقل» و «اختیار» استوار است و سامانه‌های هوشمند غیرعقل تلقی می‌شوند، بنابراین مسئولیت به عهده‌ی اصحاب آن‌ها، یعنی طراح، برنامه‌نویس و بهره‌بردار، محول است. قواعد فقهی مانند اتلاف و تسبیب نشان می‌دهد حتی اگر ماشین مباشر زیان باشد، مسئولیت نهایی بر عاقل و دارای اختیار است. در نتیجه، نمی‌توان برای سامانه‌های هوشمند شخصیت حقوقی قائل شد، اما می‌توان نظام بیمه و جبران خسارت فناوری مخصوص برای مدیریت آثار و خسارت‌های اقتصادی و عملی آنان ایجاد کرد که با اصول حقوق و فقه سازگار است (Mir Mohammad Sadeghi, 2004).

احراز سوءنیت

در حقوق کیفری، تحقق رفتار مجرمانه منوط به وجود دو رکن بنیادین است: رکن مادی، ناظر به انجام دادن یا خودداری از انجام فعلی است که قانون‌گذار آن را ممنوع کرده است؛ رکن معنوی، که بیانگر نیت، قصد یا انگیزه‌ی مجرمانه است؛ یعنی آگاهی و اراده‌ی

ج. عنصر قصد و نیت

قصد مجرمانه در حقوق کیفری به معنای جهت‌گیری عامدانه و آگاهانه‌ی فاعل به سوی تحقق نتیجه‌ای ممنوع است؛ به گونه‌ای که انسان در زمان ارتکاب جرم از پیامد رفتار خود آگاه بوده و وقوع آن را می‌پذیرد یا بر آن تعمد دارد (Mir Mohammad Sadeghi, 2024). سامانه‌های هوشمند فاقد نیت درونی و آگاهی هدفمند هستند و حتی اگر رفتار آن‌ها به نتیجه‌ای مجرمانه منجر شود، نمی‌توان وجود قصد ارتکاب جرم را برایشان مفروض دانست؛ زیرا هیچ‌گونه پیش‌بینی، اختیار یا انتخاب آگاهانه در تصمیماتشان وجود ندارد.

بنابراین، عدم احراز عناصر آگاهی، اراده و قصد آگاهانه در ساختار هوش مصنوعی سبب می‌شود که رکن معنوی جرم در رفتار این سامانه‌ها قابل تحقق نبوده و انتساب سوءنیت به آن‌ها از دیدگاه حقوقی و فقهی منتفی گردد.

پیچیدگی رابطه‌ی سببیت در سیستم‌های هوشمند

رابطه‌ی سببیت از ارکان بنیادین مسئولیت کیفری و از مؤلفه‌های اساسی تحقق جرم به شمار می‌آید؛ زیرا تا زمانی که پیوند میان رفتار و نتیجه‌ی زیان‌بار به‌صورت قانونی و منطقی احراز نگردد، انتساب فعل به فاعل امکان‌پذیر نخواهد بود. در نظام کلاسیک حقوق کیفری، این ارتباط غالباً میان فعل مستقیم انسانی و پیامد زیان‌آور آن برقرار است (Aghaei Nia). لیکن در عصر فناوری‌های هوشمند و سامانه‌های خودکار، مفهوم سببیت با پیچیدگی قابل‌توجهی مواجه شده است؛ به‌ویژه هنگامی که پیامد زیان‌بار، نتیجه‌ی تعامل و پیوستگی تصمیمات و عملکرد طراح، برنامه‌نویس، بهره‌بردار، کاربر و خود سیستم یادگیرنده باشد. در چنین ساختارهای چندعاملی، احراز اینکه کدام یک از این عناصر «سبب حقیقی» زیان تلقی می‌شود دشوار است؛ زیرا نتیجه‌ی نهایی حاصل از شبکه‌ای متشکل از کنش‌ها و تصمیمات انسانی و غیرانسانی است که در جریان اجرای فرایندها، یادگیری مستمر و

خوداصلاحی الگوریتم‌ها با یکدیگر ترکیب شده‌اند. چالش رابطه‌ی سببیت در هوش مصنوعی پیامدهای مستقیمی در سطح اثباتی، تقنینی و فقهی دارد:

بعد اثباتی؛ دشواری تشخیص علت در پرونده‌های مربوط به هوش مصنوعی

در حقوق کیفری سنتی، تشخیص عامل جرم و ارتباط مستقیم عمل با نتیجه معمولاً آسان است؛ اما در هوش مصنوعی، تصمیمات نتیجه‌ی تعامل چند عامل‌اند: الگوریتم خودآموز، داده‌های ورودی انسان، طراحی نرم‌افزار و سخت‌افزار و شرایط محیطی. در نتیجه، تعیین مقصر واقعی و رابطه‌ی سببی دشوار می‌شود و رسیدگی طولانی، فنی و پرهزینه خواهد شد. دادگاه‌ها برای اثبات باید به گزارش‌های الگوریتم، سوابق عملکرد و داده‌های دیجیتال استناد کنند. از این رو، تشکیل واحدهای تخصصی یا دادگاه‌های فناورانه ضروری به نظر می‌رسد تا با کمک کارشناسان، رابطه‌ی علی دقیق‌تر تحلیل و از خطای قضایی جلوگیری شود.

بعد تقنینی رابطه‌ی سببیت در هوش مصنوعی

در نظام کیفری سنتی، قانون‌گذار فرض می‌کند که اگر جرم یا صدمه‌ای رخ دهد، یک انسان دارای عقل و اراده آن را انجام داده است. اما وقتی یک سیستم هوشمند یا ربات مرتکب عملی زیان‌بار می‌شود، این فرض دیگر درست نیست. هوش مصنوعی نه اراده دارد، نه قصد و نه احساس مسئولیت. در نتیجه، ساختار قوانین کیفری موجود برای چنین وضعیتی کافی نیست و باید تغییر کند. قانون‌گذار باید پاسخ دهد که در صورت آسیب‌دیدن کسی توسط هوش مصنوعی، چه کسی مسئول شناخته می‌شود؟ آیا مسئولیت بر عهده‌ی کسی است که هوش مصنوعی را طراحی کرده؟ بر عهده‌ی شرکت تولیدکننده است؟ یا بر عهده‌ی شخصی که از آن استفاده کرده و دستور داده است؟

در عمل، ممکن است همه‌ی این افراد نقش‌هایی در ایجاد حادثه داشته باشند. بنابراین، تعیین مسئولیت در قانون باید چندلایه و

اخلاقی بر «جبران ضرر و پیشگیری از تکرار» استوار است، نه صرفاً مجازات فرد خطاکار.

صندوق یا نظام جبران جمعی خسارت فناورانه

در مواردی که علت دقیق زیان مشخص نیست، مثلاً چند الگوریتم مختلف با هم عمل کرده‌اند، می‌توان قانوناً صندوقی از شرکت‌های تولیدکننده، سرمایه‌گذاران و طراحان تشکیل داد تا هزینه‌ی جبران خسارت از آن پرداخت شود. این صندوق می‌تواند زیر نظر دولت یا اتحادیه‌ی صنفی فناوری اداره شود و به قربانی کمک کند بدون نیاز به دادرسی طولانی، غرامت خود را دریافت کند. از نظر عدالت اجتماعی، چنین صندوقی مانع آن می‌شود که قربانی در فضای پیچیده‌ی فناوری بی‌دفاع بماند یا به دلیل ناتوانی در اثبات رابطه‌ی سببیت، از حق خود محروم شود.

پیامد کلی برای قانون‌گذار

در چارچوب فقهی و نظام کیفری ایران، اصل بر آن است که مسئولیت تنها بر موجود عاقل و صاحب اختیار تعلق می‌گیرد (Ardabili, 2013). چون هوش مصنوعی فاقد این ویژگی‌هاست، قانون‌گذار باید قواعد جدید مسئولیت را حول انسان صاحب کنترل طراحی کند؛ کسی که بیشترین قدرت تصمیم‌گیری، آگاهی و اختیار را در زنجیره‌ی علت دارد؛ همان «سبب اقوی» در فقه امامیه. بنابراین، مناسب‌ترین اصلاح حقوقی در ایران، ترکیبی از این سه راهکار است:

شناسایی مسئولیت اشتراکی میان طراح، تولیدکننده و کاربر

ایجاد بیمه‌ی اجباری برای سیستم‌های هوش مصنوعی

تأسیس صندوق مشترک جبران خسارت‌های ناشی از فناوری‌های هوشمند.

این ساختار باعث می‌شود قانون‌گذاری از الگوی مجازات فردی سنتی فراتر برود و به سمت عدالت جبرانی و پیشگیرانه‌ی هوشمند حرکت کند. عدالتی که هم با تفکر فقهی امامیه سازگار است و هم پاسخگوی پیچیدگی‌های عصر فناوری است.

شبکه‌ای باشد، نه صرفاً نسبت‌دادن آن به یک فرد واحد. برای روشن‌تر شدن، سه نگاه یا الگوی حقوقی اصلی در این زمینه وجود دارد:

مسئولیت اشتراکی یا شبکه‌ای

در این مدل، قانون می‌گوید هر شخصی که در زنجیره‌ی عملکرد هوش مصنوعی دخالت دارد، باید به‌اندازه‌ی نقش خود پاسخگو باشد. مثلاً: طراح نرم‌افزار، اگر الگوریتم را به‌گونه‌ای نوشته که پیش‌بینی خطر در آن سخت یا غیرممکن باشد، مسئول «تقصیر در طراحی» است. شرکت سازنده، اگر سیستم را بدون آزمایش کافی روانه‌ی بازار کرده، مسئول «سهل‌انگاری در تولید» است. کاربر، اگر از هوش مصنوعی در موقعیتی خطرناک و برخلاف دستورالعمل استفاده کرده، مسئول «استفاده‌ی نابجا یا غیرمجاز» خواهد بود. به این ترتیب، مسئولیت میان همه‌ی سبب‌های انسانی تقسیم می‌شود و دیگر نیازی نیست به ربات یا الگوریتم، شخصیت کیفری مستقل داده شود. این رویکرد باعث کاهش فرار از مسئولیت و افزایش شفافیت می‌شود (Ghafari, 2019).

بیمه‌ی هوشمند برای خسارات فناوری

در بسیاری از کشورها پیشنهاد شده که دولت یا شرکت‌های بیمه، بیمه‌های تخصصی برای هوش مصنوعی ایجاد کنند. یعنی اگر یک ربات یا نرم‌افزار باعث آسیب مالی یا جانی شود، خسارت از محل بیمه پرداخت شود. هدف این طرح آن است که قربانی بدون نیاز به تشخیص دقیق مقصر، بتواند جبران خسارت دریافت کند (Molla Karami & Naeimi, 2023). این شیوه، فشار بر دادگاه‌ها را کم می‌کند و عدالت عملی را سریع‌تر برقرار می‌سازد. همچنین، شرکت‌ها و کاربران را تشویق می‌کند تا از بیمه‌های پیشگیرانه استفاده کنند و ایمنی سیستم‌هایشان را بالا ببرند. در واقع، بیمه‌ی هوشمند تبدیل می‌شود به یک جایگزین مدرن برای مسئولیت کیفری مستقیم؛ زیرا اصل عدالت کیفری اسلامی و

بعد فقهی رابطه‌ی سببیت در هوش مصنوعی

در فقه اسلامی، چه در مکتب امامیه و چه در مذاهب اهل سنت، مسئولیت کیفری فقط بر موجودی بار می‌شود که عاقل و دارای اراده و اختیار باشد (Hurr Amili, 1989; Tusi, 1981). هوش مصنوعی و ربات، گرچه عملی را انجام می‌دهند، اما فاقد عقل انسانی و اراده‌ی مستقل‌اند. از این رو، در فقه در حکم ابزار یا مباشر غیرعاقل هستند و نمی‌توان آن‌ها را شخصاً مسئول شناخت (Khoei, 2001). طبق قاعده‌ی تسبیب، اگر مباشر فاقد عقل یا اختیار باشد، مسئولیت بر عهده‌ی سبب انسانی می‌آید؛ یعنی طراح، ناظر یا استفاده‌کننده‌ای که بیشترین کنترل و آگاهی از عملکرد سیستم داشته است. فقه امامیه این فرد را سبب اقوی می‌نامد و ضمان را بر او تحمیل می‌کند. در اهل سنت نیز همین منطق پذیرفته شده است. به این ترتیب، حتی با پیچیدگی‌های فنی و شبکه‌ای در عمل هوش مصنوعی، اصل فقهی روشن است: مسئولیت همیشه بر انسان صاحب عقل و اختیار باقی می‌ماند، نه بر ماشین.

فرا تکلیف؛ تولد نظم فقهی پسا هوش مصنوعی

عبور از انسان مکلف به سامانه‌ی مسئول

فقه اسلامی از آغاز بر «انسان مکلف» مبتنی بوده است؛ اما با ورود هوش مصنوعی، پرسش جدیدی پدید آمده: اگر تصمیم و نتیجه از سامانه‌ای بدون قصد و شعور انسانی صادر شود، آیا همچنان باید تمام بار مسئولیت بر انسان مکلف باشد؟ با توجه به پویایی اجتهاد و امکان بازخوانی قواعد (Janati, 1993)، مفهوم «فرا تکلیف» مطرح می‌شود؛ یعنی فقه از انحصار تکلیف فردی فراتر رود و نظامی فراگیر برای همه‌ی کنشگران، انسان، الگوریتم، نهاد یا شبکه، بسازد. در جهان تصمیم‌های شبکه‌ای، توضیح مسئولیت با مباشرت و تسبیب خطی کافی نیست و «فقه شبکه‌ای» برای توزیع مسئولیت میان اجزا ضروری است.

مبانی نظری فراتکلیف در فقه امامیه

در فقه امامیه، اصل بر اختیار و عقل مکلف است، اما ظرفیت‌هایی برای گسترش آن وجود دارد؛ از جمله مفهوم «شخصیت اعتباری». همان‌طور که برای نهادهایی مانند وقف، شرکت و بیت‌المال شخصیت مستقل فرض شده (Hilli, 1999)، می‌توان این منطق را به سامانه‌های هوشمند نیز تعمیم داد. چون شخصیت حقوقی اعتباری و قراردادی است، می‌توان مسئولیت را با قواعدی مانند تسلیط و الید به سامانه‌ها نسبت داد، بی‌نیاز از اثبات عقل یا نیت اخلاقی انسانی. در نگاه «فرا تکلیف»، مسئولیت طیفی است: از انسان تا فاعل فناورانه و سپس نهاد ناظر حقوقی؛ شبکه‌ای از «تکلف جمعی» که هیچ فعلی بی‌پاسخ نمی‌ماند.

بازتعریف رکن تسبیب در نظام فراتکلیفی

قاعده‌ی «سبب اقوی مقدم است» می‌تواند تعارض‌های علی در مسئولیت هوش مصنوعی را حل کند، اما در منظومه‌ی «فرا تکلیف» معنایی گسترده‌تر می‌یابد. در نظم شبکه‌ای کنش و داده، «سبب اقوی» ممکن است در قالب گره مسلط الگوریتمی یا نهادی باشد که بیشترین سهم در پیش‌بینی/کنترل خطر دارد. در نتیجه، فقه از تسبیب خطی به تسبیب شبکه‌ای می‌رسد و نقش هر عامل متناسب با مشارکتش در زنجیره‌ی تصمیم و داده سنجیده می‌شود. همچنین، تدارک ضرر جمعی‌تر می‌گردد و زمینه‌ساز نهادهایی مانند بیمه‌ی هوشمند فقهی و ضمان جمعی فناورانه می‌شود.

شخصیت شبه‌حقوقی هوش مصنوعی؛ پلی میان فقه و فناوری

اندیشه‌ی «فرا تکلیف» راهی برای عبور از دوگانه‌ی انسان دارای اراده و ابزار فاقد شعور است؛ بی‌آنکه هوش مصنوعی را دارای عقل مستقل بدانند، آن را در قلمرو مسئولیت وارد می‌کند. این نگاه بر «انتساب آثار حقوقی» به سامانه‌ها تکیه دارد، نه بر نیت انسانی. از ظرفیت فقه امامیه برای ضمان در نهادهای شبه‌ارادی (Eskini, 2001) و نیز قواعدی چون «لا ضرر»، «من أتلف...» و «السبب اقوی...» در راستای همین منطق می‌توان نظام ضمان فناورانه

پیوند می‌زند، و اصل استمرار شریعت توان حفظ ارزش‌های شرعی در فضای دیجیتال را دارد (Isfahani, 1995). این رویکرد، فقه را از سنتی به سمت سیاست جنایی مدرن برای مدیریت عدالت در عصر هوشمند سوق می‌دهد.

نتیجه‌گیری

با گسترش روزافزون هوش مصنوعی و شکل‌گیری سامانه‌هایی که می‌توانند به‌صورت مستقل تصمیم بگیرند، مبانی سنتی مسئولیت در فقه اسلامی نیازمند بازنگری عمیق است. در این پژوهش، مفهوم فرائد تکلیف به‌عنوان چارچوب تازه‌ای برای فهم رفتار و مسئولیت در عصر پسا هوش مصنوعی معرفی شد. این مفهوم نشان می‌دهد که تکلیف دیگر محدود به انسانِ مباشر نیست، بلکه در میان شبکه‌ای از کنشگران انسانی و فناورانه توزیع می‌شود؛ شبکه‌ای که در آن، علت و تأثیر به‌صورت پویا در گردش است و مسئولیت باید در سطحی توزیعی و چندلایه بررسی شود.

چنین برداشتی از مسئولیت، زمینه‌ساز شکل‌گیری نظم فقهی جدید است؛ نظمی که در آن، فاعلیت الگوریتمی همچون «سبب اقوی» مورد توجه قرار می‌گیرد و ارزیابی مسئولیت صرفاً بر اساس قصد و مباشرت انسانی نیست، بلکه بر پایه‌ی تعامل میان انسان، فناوری و ساختارهای تصمیم‌گیری پیچیده سنجیده می‌شود. مفاهیمی مانند عدالت شبکه‌ای، مسئولیت توزیعی و اجتهاد فناورانه در این نظام فقهی جدید نقش بنیادین دارند و می‌توانند جهت حرکت فقه اسلامی را در جهان پسا هوش مصنوعی تعیین کنند. مبانی فقهی مسئولیت کیفری در مواجهه با خسارات ناشی از سیستم‌های هوش مصنوعی و ربات‌ها بر حفظ نظام بشری و اصول عدالت استوار است و باید به‌گونه‌ای باشد که نظم عمومی و حقوق بنیادین رعایت شده و عدالت عملی تأمین گردد تا جامعه در مسیر سعادت و امنیت روانی قرار گیرد. بر این اساس، قاعده‌ی مصلحت بر اهمیت سنجش آثار فناوری‌های نوین بر امنیت و نظم عمومی تأکید دارد و مسئولیت کیفری باید در جهت جلوگیری از تضرر

ساخت. در این نظام، الگوریتم به‌عنوان واحد دارای حساب ضمان مستقل تلقی می‌شود و جبران خسارت از صندوق‌ها یا بیمه‌های هوشمند انجام می‌گیرد، نه از قصد اخلاقی.

پیامدهای نظری نظم فقهی پسا هوش مصنوعی

نظم پسا هوش مصنوعی، یعنی مرحله‌ای که در آن فناوری به سطح مشارکت در ساختار عدالت و جرم رسیده است.

در این مرحله، سه تحول نظری رخ می‌دهد:

الف) تغییر ماهیت مکلف

مکلف، دیگر تنها انسانِ مکلف شرعی نیست؛ بلکه شبکه‌ای از عوامل است که به‌طور مشترک در تکلیف و تخلف سهیم‌اند. این دیدگاه، مفهوم «اهلیت کیفری» را از انحصار انسان رها و آن را به الگویی تقسیم‌پذیر تبدیل می‌کند. به تعبیری دیگر، مفهوم «اهلیت کیفری» از حالت فردمحور خارج و به‌صورت اهلیت شبکه‌ای یا توزیعی تعریف می‌شود؛ یعنی میزان مسئولیت هر عامل «انسانی یا غیرانسانی» بر پایه‌ی نقش و تأثیرش در زنجیره‌ی تصمیم و خطا مشخص می‌گردد. به این ترتیب، نظام مسئولیت از تک‌فاعل به چندفاعل تغییر پیدا می‌کند.

ب) عدالت شبکه‌ای

در نگاه «فرائد تکلیف»، عدالت فردی نیست، بلکه توازن سیستمی میان انسان و فناوری است؛ وقتی میان داده‌ها، تصمیم‌های هوشمند و اثرات اجتماعی هماهنگی باشد، عدالت محقق می‌شود. در نتیجه، «مصلحت عامه» نیز از تعیین صرف حاکم/شارع به تعامل پویا میان انسان و الگوریتم‌ها وابسته می‌گردد.

ج) ضرورت اجتهاد فناورانه

فقه برای پاسخ‌گویی به چالش‌های پسا هوش مصنوعی، نیازمند توسعه‌ی ابزارهای اجتهادی خود است. «اصول فقه فناوری» با تکیه بر قواعدی مانند تسبیب، ضمان و عدالت توزیعی تدوین می‌شود و مسئولیت شبکه‌ای و عدالت فناورانه را تبیین می‌کند. در این نسل نوین، مفهوم «فرائد تکلیف» مسئولیت میان انسان و فناوری را

یادشده، این پژوهش نشان می‌دهد که فقه اسلامی با تکیه بر ظرفیت اجتهاد پویا و نگاه مقاصدی، توان انطباق با تحولات فناورانه را دارد. فراتکلیف به‌جای آنکه نشانه‌ای از پایان مسئولیت باشد، افقی تازه برای بازتعریف مسئولیت انسانی و اخلاقی در جهان هوشمند و فناورانه‌ی آینده است؛ جهانی که در آن انسان، فناوری و شریعت در نظامی هم‌پیوند و پویا با هم معنا می‌یابند.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

EXTENDED ABSTRACT

The rapid development of artificial intelligence, autonomous technologies, and robotic systems has created one of the most significant conceptual challenges for criminal law and Islamic jurisprudence: how can responsibility be determined when harm is caused by a system that acts, decides, learns, or responds without possessing human intention, moral consciousness, or legally recognizable will? Classical theories of criminal liability generally presuppose a human actor who can be addressed by legal norms, understand prohibition, exercise choice, and be blamed for violating a protected interest. In Imamiyya jurisprudence, this structure is even more closely tied to the concepts of reason, volition, intention, and imputability, because criminal responsibility is traditionally grounded in the capacity of the mukallaf to understand the legal-moral meaning of conduct and to act with awareness and choice (Khoei, 2001; Khomeini, 2000). The emergence of artificial intelligence unsettles this structure because autonomous systems may produce legally harmful consequences while lacking the internal qualities that make punishment meaningful in Islamic criminal thought. A robot may physically injure a person, an algorithm may cause financial or

جامعه و افراد و هدایت فناوری در مسیر خدمت به مصالح جمعی و پرهیز از مفسده باشد. همچنین، عقل، به‌عنوان منبع اصلی درک حقایق، نقش کلیدی در تشخیص میزان و نوع خسارت و تعیین مجازات‌های متناسب دارد و قاعده‌ی لاضرر تضمین می‌کند که هیچ عملی نباید منجر به وارد کردن ضرر به دیگران شود. در کنار این، قاعده‌ی اتلاف نیز مسئولیت جبران خسارت‌های فیزیکی یا ارزش مالی ناشی از آسیب‌های فناوری‌های نوین را بر عهده می‌گیرد و به‌طور کلی، نظام حقوقی باید از این مبانی ولذت بهره‌مند شده و در جهت حفظ امنیت، نظم و حقوق افراد، مسئولیت‌های کیفری دقیق و منصفانه‌ای تعیین کند. نظر به استدلال

reputational damage, and an intelligent system may make decisions that affect life, health, property, or public order; yet these systems cannot be treated as morally blameworthy in the same way as human offenders. Accordingly, the central problem is not merely technical but jurisprudential: whether the traditional foundations of causation, intent, fault, and liability can be reinterpreted in a way that preserves the normative integrity of Islamic law while responding to the realities of algorithmic decision-making and distributed technological agency (Mirsaeedi, 2007; Zeraat, 2017).

The study is based on a jurisprudential-legal analysis of criminal liability for harms caused by artificial intelligence systems and robots, with particular attention to the Imamiyya legal tradition. It examines the doctrinal foundations of liability, including the legal element, material element, and mental element of crime, and then evaluates whether these elements can be attributed to artificial intelligence. The legal element requires a valid textual or statutory basis for punishment, consistent with the principle that punishment without prior declaration is impermissible; this principle prevents the direct criminalization of artificial intelligence systems unless a recognized norm clearly

establishes the relevant obligation or prohibition (Ardabili, 2013; Muhaqqiq Hilli, 1988). The material element may appear more readily applicable because robots and intelligent systems can physically perform harmful acts, but material performance alone is insufficient for criminal responsibility when the acting entity lacks consciousness and volition. The mental element creates the deepest challenge, because intention, awareness, and choice constitute the core of criminal imputability in both modern criminal law and Islamic jurisprudence (Jafari, 2016; Mansourabadi, 2016). Since artificial intelligence operates through data processing, algorithmic optimization, and programmed or learned responses, its conduct cannot be equated with human intentionality. Therefore, even when the external form of conduct resembles intentional action, the absence of genuine awareness and legal-moral comprehension prevents attribution of mens rea to the machine itself.

The findings indicate that the most coherent jurisprudential approach is to analyze artificial intelligence and robots through the doctrines of direct commission and causation rather than through direct criminal personality. In Imamiyya jurisprudence, responsibility may arise through mubasharat when a person directly causes harm, or through tasbib when a person creates the conditions or means through which harm occurs. In cases involving artificial intelligence, the machine may function as the immediate physical medium of harm, but it is not a conscious direct perpetrator. Therefore, the inquiry must shift toward the human or institutional causes behind the harmful outcome: the designer, programmer, producer, operator, owner, user, supervisor, or regulatory authority that enabled, controlled, deployed, trained, or failed to monitor the system. The jurisprudential maxim that the stronger cause may prevail over the immediate actor is

especially important in this context, because the robot or algorithm may be the apparent direct actor while the legally relevant cause remains the human agent who had knowledge, control, or capacity to prevent harm (Mohaghegh Damad, 1999; Najafi, 1984). This framework allows criminal responsibility to remain attached to rational and legally accountable persons without ignoring the complex causal role of autonomous systems. It also enables the transition from a linear model of causation to a networked model in which liability is distributed according to the degree of control, foreseeability, negligence, and contribution to risk.

The jurisprudential foundations of liability in this field are further supported by broader principles such as preservation of social order, justice, public interest, reason, no-harm, and destruction. The preservation of human social order requires that technological harm not be treated as a neutral technical accident when it threatens life, property, dignity, security, or public trust. Islamic law recognizes the necessity of maintaining social order and preventing disruption in the conditions required for human welfare and moral life (Tabatabaei Yazdi, 1991). The principle of justice also demands that the legal system respond proportionately to harm, not merely by assigning blame but by restoring the rights of victims and preventing unequal or arbitrary distribution of technological risk (Mohaghegh Damad, 2006). The principle of public interest provides a flexible basis for regulating artificial intelligence where technological benefits must be balanced against risks to persons and society (Awdeh, 1993; Shartouni, 1983). Likewise, the rule of no-harm prevents the legitimacy of conduct, systems, or regulatory omissions that produce unjustified injury to others (Ansari; Rashti Najafi). The rule of destruction also reinforces compensation when property, value, or legally protected interests are damaged through technological

conduct (Jafari Langroudi, 2006; Sami, 2021). These principles collectively support a model of criminal and compensatory responsibility that is preventive, restorative, and proportionate.

One of the major theoretical contributions of the study is the proposal of “meta-obligation” as a jurisprudential framework for the post-artificial intelligence order. Meta-obligation does not mean abandoning the classical requirement that responsibility ultimately belongs to rational and accountable subjects; rather, it means expanding the field in which obligation and responsibility are analyzed. In a technological environment, harmful outcomes are often produced by networks of human decisions, institutional policies, algorithmic structures, datasets, machine learning processes, and operational contexts. A purely individual-centered theory of responsibility may fail to identify the real sources of risk, while direct criminal personality for artificial intelligence may conflict with the jurisprudential foundations of reason and volition. Meta-obligation therefore offers an intermediate framework: responsibility remains normatively anchored in human and institutional accountability, but it is distributed across the network that produces the harmful decision. This model is compatible with the idea of legal personality as a constructed and conventional concept, especially where legal systems already recognize certain non-natural entities for functional purposes (Eskini, 2001; Rahmanpour, 2013). It may also justify mechanisms such as mandatory technological insurance, collective compensation funds, regulatory duties, and specialized standards of care for designers and operators of intelligent systems (Molla Karami & Naeimi, 2023; Vaseghi, 2020). In this sense, meta-obligation transforms liability from a single-point inquiry into a layered assessment of control, risk creation, supervision, and preventability.

In conclusion, the study shows that Imamiyya jurisprudence possesses significant interpretive capacity to respond to the challenges of artificial intelligence and robotics without abandoning its fundamental principles. Artificial intelligence cannot presently be treated as a genuine criminal subject because it lacks reason, volition, moral understanding, and intentionality; nevertheless, its harmful effects cannot remain outside the legal and jurisprudential order. The appropriate response is a networked model of liability that identifies the human and institutional causes behind technological harm and distributes responsibility according to control, foreseeability, fault, and contribution to risk. The proposed concept of meta-obligation provides a jurisprudential bridge between classical doctrines and the post-artificial intelligence world by preserving the centrality of human accountability while recognizing that harm increasingly arises within complex systems of data, design, automation, and decision-making. This framework supports justice, compensation, prevention, and public order, while also encouraging the development of technological *ijtihad*, specialized regulation, insurance mechanisms, and collective compensation structures. Ultimately, the encounter between Islamic jurisprudence and artificial intelligence need not produce rupture; rather, it can become an opportunity to reconstruct the concepts of causation, responsibility, and justice in a broader and more realistic horizon suited to the age of intelligent technologies.

References

- Aghaei Nia, H. *Crimes against the Physical Integrity of Persons* (20th ed.). Mizan.
- Ansari, M. i. M. A. *The Rules of No Harm, Possession, Validity, and Drawing Lots (Fara'id al-Usul)*. Islamic Publications Office Affiliated with the Society of Seminary Teachers of Qom.
- Ardabili, M. A. (2013). *General Criminal Law, Volume 1* (8th ed.). Mizan.

- Awdeh, A. a.-Q. (1993). *Islamic Criminal Legislation Compared with Positive Law* (N. Ghorban, Trans.). Shahid Beheshti University Academic Jihad.
- Dehkhoda, A. A. (1946). *Dehkhoda Dictionary*. University of Tehran Press and Publishing Institute.
- Eskini, R. (2001). *Commercial Law, Volume 1* (5th ed.). SAMT.
- Ghafouri, S. A. A. (2019). *Rights of Robots and Artificial Intelligence in Legal Systems*. SAMT.
- Hilli, H. i. Y. (1999). *Qawa'id al-Ahkam, Volume 2*. Islamic Publishing Institute.
- Hurr Amili, M. i. H. (1989). *Wasa'il al-Shi'a, Volume 28*. Al al-Bayt Institute.
- Ibn Manzur, M. i. M. (1988). *Lisan al-Arab*. Dar Ihya' al-Turath al-Arabi.
- Isfahani, M. H. (1995). *Nihayat al-Dirayah fi Sharh al-Kifayah, Volume 1*. Seyed al-Shohada Institute.
- Jafari Langroudi, J. (2006). *Legal Terminology*. Ganj Danesh.
- Jafari, M. (2016). A Reflection on the Nature and Scope of Will in the Realization of Criminal Responsibility. *Criminal Law Research*(15).
- Jafari Tabrizi, M. T. *Sources of Jurisprudence*.
- Janati, M. E. (1993). *Periods of Ijtihad from the Perspective of Islamic Schools* (1st ed.). Kayhan.
- Khoei, A. (2001). *Mabani Takmilat al-Minhaj, Volume 2*. Revival of the Works of Imam al-Khoei.
- Khomeini, R. (2000). *Tahrir al-Wasilah, Volume 2* (2nd ed.). Institute for Compilation and Publication of Imam Khomeini's Works.
- Mansourabadi, A. (2016). *General Criminal Law 1* (1st ed.). Mizan.
- Mir Mohammad Sadeghi, H. (2004). *The International Criminal Court*. Dadgostar.
- Mir Mohammad Sadeghi, H. (2024). *Specific Criminal Law: Crimes against Persons* (23rd ed.). Mizan.
- Mirsaeedi, S. M. (2007). *Criminal Responsibility: Scope and Elements, Volume 1* (2nd ed.). Mizan.
- Mohaghegh Damad, M. (1999). *Jurisprudential Rules: Criminal Section* (3rd ed.). University Publishing Center.
- Mohaghegh Damad, M. (2006). *Jurisprudential Rules: Civil Section*. Islamic Sciences Publishing Center.
- Molla Karami, S., & Naeimi, O. (2023). *Occupational and Professional Civil Liability Insurance* (2nd ed.). Majd.
- Mousavi Bojnourdi, S. M. H. (1996). The Rule of Assumption of Risk. *Legal Perspectives*(2).
- Muhaqqiq Hilli, N. a.-D. J. i. H. (1988). *Shara'i al-Islam: Kitab al-Hudud, Volume 4*. Society of Seminary Teachers of Qom.
- Najafi, M. H. i. B. (1984). *Jawahir al-Kalam, Volume 43*. Dar Ihya' al-Turath al-Arabi.
- Rahmanpour, A. (2013). *Criminal Responsibility of Legal Persons*. Imam Hossein University.
- Rashti Najafi, M. H. *Kitab al-Ghasb*.
- Sadeghi, M. H. (1981). *Crimes against Persons*. Mizan.
- Sami, M. J. (2021). Applying the Rule of Ghurur to the Rules of Destruction, Causation, and No Harm in Iranian Criminal Law. *Quarterly Journal of Modern Interdisciplinary Legal Research*, 1(2), 1-12.
- Shartouni, S. (1983). *Aqrab al-Mawarid*. Library of Grand Ayatollah Marashi Najafi.
- Tabatabaei Yazdi, S. M. K. (1991). *Hashiyat al-Makasib*. Ismailian Institute.
- Tusi, M. i. H. (1981). *Tahdhib al-Ahkam, Volume 10*. Dar al-Saab.
- Vaseghi, M. (2020). Feasibility of Granting Legal Personality to Intelligent Robots Based on the European Union Resolution. *Majlis and Rahbord*(103).
- Zeraat, A. (2017). *General Criminal Law, Volume 1* (6th ed.). Jangal.