

Criminal Liability for Homicide Caused by Artificial Intelligence

1. Hesam Maleki: MA, Department of Criminal Law and Criminology, Faculty of Humanities and Social Sciences, University of Shomal, Amol, Iran
2. Hossein Ranjbar*: Assistant Professor, Department of Law, Faculty of Humanities and Social Sciences, Shomal University, Amol, Iran. Email: h.ranjbar@shomal.ac.ir (Corresponding Author)

ABSTRACT

The expanding use of artificial intelligence in autonomous transportation, healthcare, industry, robotics, and decision-support systems has generated new questions concerning the attribution of criminal liability when harmful consequences, particularly human death, result from the operation of intelligent systems. The central problem arises when a fatal outcome is directly produced by an artificial intelligence system, while multiple natural and legal persons are involved in its design, production, programming, training, deployment, operation, and supervision. This study adopts a descriptive-analytical method to examine the foundations and criteria of criminal attribution in cases of homicide caused by artificial intelligence and analyzes the potential liability of programmers, designers, manufacturers, developing companies, owners, operators, and users. The findings indicate that the technical autonomy of an intelligent system, by itself, cannot justify recognizing artificial intelligence as an independent criminal offender. Under current legal conditions, criminal liability should instead be determined according to the role of human and corporate actors in creating, controlling, or failing to prevent the relevant risk. In this regard, factors such as foreseeability of harm, effective control over the system, access to technical information, the ability to prevent the fatal outcome, breach of a duty of care, and the degree of causal contribution of each actor are of fundamental importance. Where several factors jointly contribute to death, responsibility may need to be distributed among multiple actors rather than attributed exclusively to one person. The study concludes that general principles of criminal law remain applicable to artificial-intelligence-related homicide, but they are not entirely sufficient for autonomous and high-risk systems. Specific legal rules concerning risk assessment, human oversight, algorithmic traceability, safety obligations, post-deployment monitoring, and corporate criminal liability are therefore necessary to prevent responsibility gaps and ensure effective legal accountability.

Keywords: *Artificial Intelligence, Criminal Liability, Homicide, Causation, Autonomous Systems, Criminal Negligence, Corporate Criminal Liability*

How to cite: Maleki, H. & Ranjbar, H. (2027). Criminal Liability for Homicide Caused by Artificial Intelligence. *Comparative Studies in Jurisprudence, Law, and Politics*, 9(5), 1-20.

© 2027 the authors. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Submit Date: 04 May 2026

Revise Date: 23 September 2026

Accept Date: 30 September 2026

Initial Publish Date: 06 October 2026

Final Publish Date: 22 December 2027



مسئولیت کیفری ارتکاب جرم قتل ناشی از هوش مصنوعی

۱. حسام ملکی: کارشناسی ارشد، گروه حقوق کیفری و جرم شناسی، دانشکده علوم انسانی و اجتماعی، دانشگاه شمال، آمل، ایران

۲. حسین رنجبر*: استادیار، گروه حقوق، دانشکده علوم انسانی و اجتماعی، دانشگاه شمال، آمل، ایران.

پست الکترونیک: h.ranjbar@shomal.ac.ir (نویسنده مسئول)

چکیده

گسترش کاربرد هوش مصنوعی در حوزه‌هایی مانند حمل و نقل خودران، پزشکی، صنعت، رباتیک و سامانه‌های تصمیم‌یار، پرسش‌های جدیدی را درباره انتساب مسئولیت کیفری در صورت وقوع نتایج زیان‌بار، به‌ویژه مرگ انسان، ایجاد کرده است. مسئله اصلی زمانی شکل می‌گیرد که نتیجه مرگبار مستقیماً ناشی از عملکرد یک سامانه هوشمند باشد، اما در زنجیره طراحی، تولید، برنامه‌ریزی، آموزش، بهره‌برداری و نظارت بر آن، اشخاص متعدد انسانی و حقوقی نقش داشته باشند. پژوهش حاضر با روش توصیفی - تحلیلی به بررسی مبانی و معیارهای انتساب مسئولیت کیفری در قتل ناشی از هوش مصنوعی می‌پردازد و جایگاه برنامه‌نویس، طراح، تولیدکننده، شرکت توسعه‌دهنده، مالک، بهره‌بردار و کاربر را در فرایند تحقق نتیجه تحلیل می‌کند. یافته‌ها نشان می‌دهد که صرف خودمختاری فنی سامانه هوشمند نمی‌تواند موجب شناسایی آن به‌عنوان مرتکب مستقل جرم شود و در وضعیت کنونی، مسئولیت کیفری باید بر مبنای نقش اشخاص انسانی و حقوقی در ایجاد، کنترل یا عدم پیشگیری از خطر تعیین شود. در این زمینه، معیارهایی مانند قابلیت پیش‌بینی خطر، کنترل مؤثر بر سامانه، دسترسی به اطلاعات فنی، امکان جلوگیری از نتیجه، نقض تکلیف مراقبت و میزان تأثیر رفتار هر شخص در زنجیره سببیت اهمیت اساسی دارند. همچنین، در مواردی که چند عامل به‌طور هم‌زمان در تحقق مرگ نقش دارند، امکان توزیع مسئولیت میان اشخاص متعدد باید مورد پذیرش قرار گیرد. نتیجه پژوهش نشان می‌دهد که قواعد عمومی حقوق کیفری همچنان قابلیت استفاده در این حوزه را دارند، اما برای مواجهه مؤثر با سامانه‌های خودمختار و پرخطر، تدوین قواعد اختصاصی درباره ارزیابی خطر، نظارت انسانی، ثبت تصمیمات الگوریتمی، الزامات ایمنی و مسئولیت اشخاص حقوقی ضروری است.

واژگان کلیدی: کلیدواژه‌ها: هوش مصنوعی، مسئولیت کیفری، قتل، رابطه سببیت، سامانه‌های خودمختار، تقصیر کیفری، مسئولیت اشخاص حقوقی

نحوه استناددهی: ملکی، حسام، و رنجبر، حسین. (۱۴۰۶). مسئولیت کیفری ارتکاب جرم قتل ناشی از هوش مصنوعی. پژوهش‌های تطبیقی فقه، حقوق و سیاست، ۹(۵)، ۲۰-۱.

© ۱۴۰۶ تمامی حقوق انتشار این مقاله متعلق به نویسنده است. انتشار این مقاله به‌صورت دسترسی آزاد مطابق با گواهی (CC BY-NC 4.0) صورت گرفته است.

تاریخ ارسال: ۱۴ اردیبهشت ۱۴۰۵

تاریخ بازنگری: ۱ مهر ۱۴۰۵

تاریخ پذیرش: ۸ مهر ۱۴۰۵

تاریخ چاپ اولیه: ۱۴ مهر ۱۴۰۵

تاریخ چاپ نهایی: ۱ دی ۱۴۰۶

هوش مصنوعی در دهه‌های اخیر از مرحله‌ای که صرفاً به‌عنوان یک فناوری پشتیبان برای پردازش داده‌ها و انجام محاسبات پیچیده مورد استفاده قرار می‌گرفت، عبور کرده و به سامانه‌ای تبدیل شده است که می‌تواند در برخی حوزه‌ها داده‌ها را تحلیل کند، گزینه‌های مختلف را ارزیابی نماید، الگوهای رفتاری را شناسایی کند و حتی بدون مداخله مستقیم و لحظه‌ای انسان تصمیم‌هایی اتخاذ کند که آثار آن در جهان واقعی ظاهر می‌شود. توسعه خودروهای خودران، ربات‌های صنعتی، سامانه‌های تشخیص و درمان پزشکی، سلاح‌های خودکار، نرم‌افزارهای مدیریت زیرساخت، سامانه‌های تصمیم‌یار و ابزارهای مبتنی بر یادگیری ماشین سبب شده است که مسئله مسئولیت حقوقی و به‌ویژه مسئولیت کیفری ناشی از عملکرد این سامانه‌ها به یکی از مسائل پیچیده حقوق کیفری معاصر تبدیل شود. ویژگی مهم این تحول آن است که رفتار سامانه هوشمند همیشه به اجرای مکانیکی یک فرمان صریح از پیش تعیین‌شده محدود نیست، بلکه برخی سامانه‌های پیشرفته می‌توانند براساس داده‌های جدید رفتار خود را اصلاح کنند و در چارچوب معماری فنی خود به نتایجی برسند که جزئیات آن از ابتدا برای طراح یا کاربر قابل پیش‌بینی نبوده است. همین ویژگی موجب شده است که قواعد سنتی مسئولیت کیفری، که اساساً با فرض وجود یک انسان دارای اراده، ادراک و قابلیت انتساب رفتار طراحی شده‌اند، در مواجهه با زیان‌های ناشی از سامانه‌های خودمختار با پرسش‌های جدیدی روبه‌رو شوند (Zandi & Rafiei Alavi, 2024). بحث زمانی اهمیت بیشتری می‌یابد که عملکرد هوش مصنوعی نه صرفاً به خسارت مالی یا اختلال در یک فرایند، بلکه به مرگ یک انسان منتهی شود. در چنین وضعیتی، نظام عدالت کیفری باید تعیین کند که آیا مرگ ایجادشده را می‌توان به شخص یا اشخاص معینی منتسب کرد، عنوان کیفری مناسب چیست، عنصر روانی چگونه احراز می‌شود و در صورت وجود چند عامل انسانی و فناورانه در زنجیره ایجاد نتیجه، مسئولیت چگونه میان آنان توزیع می‌شود.

قتل از مهم‌ترین جرایم علیه تمامیت جسمانی انسان و از شدیدترین اشکال تعرض به حق حیات به شمار می‌رود. در ساختار سنتی حقوق کیفری، انتساب قتل به مرتکب مستلزم آن است که رفتار انسانی مشخصی با نتیجه مرگ ارتباط داشته باشد و شرایط لازم برای مسئولیت کیفری، از جمله رابطه سببیت و حسب مورد عنصر روانی مربوط به نوع قتل، احراز شود (Haji, 2018). ورود هوش مصنوعی این الگوی نسبتاً روشن را پیچیده می‌کند، زیرا ممکن است میان شخصی که سامانه را طراحی کرده، شرکتی که آن را تولید نموده، شخصی که آن را آموزش داده، مالک سامانه، کاربر نهایی و رفتار بالفعل سامانه فاصله قابل توجهی وجود داشته باشد. برای نمونه، در یک خودروی خودران ممکن است نقص در الگوریتم تشخیص عابر، کیفیت نامناسب داده‌های آموزشی، قصور شرکت در به‌روزرسانی نرم‌افزار، بی‌توجهی مالک به هشدارهای فنی یا تصمیم لحظه‌ای سامانه در انتخاب مسیر، هر یک در تحقق حادثه مرگبار نقش داشته باشند. در چنین حالتی تعیین اینکه کدام عامل از منظر حقوق کیفری سبب نتیجه است و مسئولیت کیفری بر عهده چه کسی قرار می‌گیرد، به‌سادگی الگوی کلاسیک فعل مستقیم مرتکب و وقوع نتیجه نیست. مطالعات جدید در زمینه مسئولیت هوش مصنوعی نیز نشان می‌دهد که یکی از مهم‌ترین موانع در این حوزه، پراکندگی نقش‌ها میان طراح، توسعه‌دهنده، تولیدکننده و بهره‌بردار و دشواری شناسایی نقطه‌ای است که در آن مسئولیت کیفری باید مستقر شود (Ghavami Pour Sereshkeh & Mahmoudi, 2025).

از منظر حقوق کیفری، صرف وقوع یک نتیجه زیان‌بار برای تحمیل مسئولیت کیفری کافی نیست. مسئولیت کیفری اصولاً بر مبنای انتساب شخصی رفتار، وجود عنصر مادی، رابطه سببیت و درجه‌ای از تقصیر یا سوءنیت استوار است. هوش مصنوعی، به‌ویژه در شکل خودمختار، این عناصر را با چالش مواجه می‌کند. ممکن است سامانه‌ای بدون آنکه کاربر به‌صورت مستقیم فرمان‌کننده‌ای صادر کرده باشد، تصمیمی

اتخاذ کند که به مرگ دیگری منتهی شود. از سوی دیگر، طراح الگوریتم ممکن است هیچ‌گونه قصدی نسبت به نتیجه نداشته و حتی وقوع آن را پیش‌بینی نکرده باشد، اما طراحی وی دارای نقصی بوده باشد که در شرایط خاص به تصمیم مرگبار منجر شده است. مسئله دشوارتر زمانی شکل می‌گیرد که سامانه پس از استقرار و برآثر یادگیری از داده‌های محیطی رفتاری از خود نشان دهد که حتی توسعه‌دهنده نیز از قبل قادر به پیش‌بینی دقیق آن نبوده است. در چنین شرایطی، پرسش از امکان احراز عنصر روانی نسبت به انسان‌های دخیل در زنجیره و نیز امکان فرض نوعی مسئولیت مستقیم برای خود سامانه هوشمند مطرح می‌شود (Javanmard & Heydarifard, 2025). برخی دیدگاه‌ها بر آن‌اند که استقلال عملیاتی سامانه‌های هوشمند، الگوهای سنتی مسئولیت را تضعیف کرده و ضرورت بازاندیشی در مبانی انتساب کیفری را مطرح می‌سازد (Younes, 2024). در مقابل، دیدگاهی دیگر بر این باور است که حتی پیشرفته‌ترین سامانه‌های هوش مصنوعی نیز در نهایت در چارچوبی عمل می‌کنند که به‌وسیله انسان طراحی، تأمین مالی، داده‌دهی، استقرار و بهره‌برداری شده و بنابراین نباید استقلال فنی با استقلال حقوقی یکسان دانسته شود (Kaveh & Barani, 2024).

یکی از مهم‌ترین وجوه مسئله آن است که نظام حقوق کیفری در بسیاری از کشورها هنوز شخصیت کیفری مستقلی برای هوش مصنوعی به رسمیت نمی‌شناسد. بنابراین حتی اگر یک سامانه به لحاظ فنی قادر به انتخاب میان گزینه‌های مختلف باشد، این امر لزوماً به معنای برخورداری آن از اراده حقوقی یا قابلیت تحمل مجازات نیست. بررسی امکان مسئولیت کیفری مستقل هوش مصنوعی نیازمند پاسخ به پرسش‌هایی درباره مفهوم اراده، آگاهی، قصد، فهم هنجار و قابلیت سرزنش است. مسئولیت کیفری در معنای کلاسیک خود تنها رابطه‌ای میان رفتار و نتیجه نیست، بلکه واجد بعد هنجاری است؛ یعنی نظام کیفری باید بتواند رفتار مرتکب را از نظر اخلاقی و حقوقی قابل سرزنش بداند. به همین دلیل، برخی پژوهش‌ها امکان مسئولیت کیفری مستقیم هوش مصنوعی را از منظر فلسفی و حقوقی محل تردید دانسته‌اند (Zandi & Rafiei, 2024). در عین حال، افزایش سطح خودمختاری سامانه‌ها موجب شده است که کنار گذاشتن کامل نقش خود سامانه در تحلیل کیفری نیز دشوار باشد، زیرا هرچه فاصله میان تصمیم انسان و نتیجه نهایی بیشتر شود، یافتن عنصر روانی یا رابطه مستقیم سببیت در یکی از اشخاص انسانی دشوارتر خواهد شد (Al-Halahlah, 2025).

در حقوق ایران نیز مسئله مسئولیت ناشی از اعمال هوش مصنوعی هنوز در مرحله شکل‌گیری نظری و تقنینی قرار دارد. قواعد عمومی مسئولیت کیفری، قواعد مربوط به جرایم علیه اشخاص و اصول انتساب نتیجه می‌توانند بخشی از مسائل را پوشش دهند، اما ویژگی‌های فنی سامانه‌های هوشمند موجب شده است که کاربرد این قواعد همیشه بدون دشواری نباشد. پژوهش‌های داخلی در این زمینه نشان داده‌اند که تعیین مسئول در زیان‌های ناشی از هوش مصنوعی نیازمند توجه هم‌زمان به میزان کنترل، سطح خودمختاری سامانه، پیش‌بینی‌پذیری رفتار، نقش هر یک از اشخاص دخیل و قابلیت جلوگیری از نتیجه است (Al-Sadat Makki et al., 2024). همچنین مطرح شده است که خلأ قواعد اختصاصی ممکن است در آینده موجب بروز وضعیت‌هایی شود که در آنها نتیجه مجرمانه آشکار است، اما انتساب آن به یک شخص معین با دشواری جدی مواجه می‌شود (Rostami Zabol, 2025). این مسئله به‌ویژه در قتل اهمیت مضاعف دارد، زیرا نمی‌توان پذیرفت که صرف پیچیدگی فناوری موجب شکل‌گیری منطقه‌ای فاقد مسئولیت کیفری شود و در عین حال نیز نمی‌توان تنها به دلیل شدت نتیجه، شخصی را بدون احراز شرایط قانونی مسئولیت مجازات کرد.

مسئله اصلی پژوهش حاضر آن است که در صورت وقوع مرگ انسان برآثر عملکرد سامانه هوش مصنوعی، مسئولیت کیفری چگونه قابل انتساب است و چه شخص یا اشخاصی را می‌توان مسئول ارتکاب قتل دانست. در این چارچوب، بررسی جایگاه طراح، برنامه‌نویس،

تولیدکننده، شرکت توسعه‌دهنده، مالک، بهره‌بردار و کاربر اهمیت می‌یابد و باید مشخص شود که نقش هر یک از این افراد تحت چه شرایطی می‌تواند مبنای مسئولیت کیفری قرار گیرد. همچنین این پژوهش در پی آن است که حدود کارآمدی مفاهیم سنتی مباشرت، تسبیب، رابطه سببیت، تقصیر، سوءنیت و قابلیت پیش‌بینی را در مواجهه با سامانه‌های خودمختار بررسی کند و مشخص سازد آیا برای این حوزه به معیارهای تکمیلی یا قواعد اختصاصی نیاز است. نوآوری بحث در تمرکز خاص بر نتیجه قتل است؛ زیرا بخش قابل توجهی از مطالعات موجود مسئولیت هوش مصنوعی را به‌طور کلی یا عمدتاً از منظر خسارت و خطای فناوریانه بررسی کرده‌اند، در حالی که قتل از حیث شدت نتیجه، الزامات اثباتی و پیامدهای کیفری از حساسیت ویژه‌ای برخوردار است. روش این پژوهش توصیفی - تحلیلی است و با استفاده از مبانی حقوق کیفری و تحلیل مطالعات مرتبط با مسئولیت هوش مصنوعی، در پی ارائه چارچوبی برای تعیین مسئولیت در مواردی است که عملکرد یک سامانه هوشمند به مرگ انسان منتهی می‌شود.

مبانی مفهومی و حقوقی مسئولیت کیفری در قتل ناشی از هوش مصنوعی

شناخت مسئولیت کیفری ناشی از قتل مرتبط با هوش مصنوعی پیش از هر چیز مستلزم تبیین ماهیت خاص این فناوری در مقایسه با ابزارهای متعارف است. ابزار سنتی معمولاً فاقد ظرفیت تصمیم‌گیری مستقل است و نقش آن در تحقق جرم به رفتار انسانی استفاده‌کننده بازمی‌گردد. برای مثال، در استفاده از یک وسیله مکانیکی، رابطه میان دستور انسان و عملکرد وسیله غالباً مستقیم و قابل پیش‌بینی است. در مقابل، برخی سامانه‌های هوش مصنوعی براساس مدل‌های پیچیده محاسباتی و داده‌های گسترده قادر به پردازش محیط، تشخیص الگو و انتخاب خروجی هستند. در چنین وضعیتی، هرچند خاستگاه سامانه و معماری آن انسانی است، اما فرایند تولید خروجی ممکن است به حدی پیچیده باشد که تشخیص دقیق چرایی یک تصمیم برای کاربر عادی و حتی گاه برای طراحان دشوار باشد. این ویژگی که گاه با مفهوم «جعبه سیاه» توصیف می‌شود، از مهم‌ترین عوامل دشواری انتساب مسئولیت کیفری است (Ghavami Pour Sereshkeh & Mahmoudi, 2025). خودمختاری در این حوزه نیز مفهومی نسبی است و نباید آن را به معنای استقلال مطلق از انسان دانست؛ سامانه ممکن است در مرحله اجرا بدون فرمان لحظه‌ای انسان عمل کند، اما هدف، داده، معماری، محدودیت‌ها و محیط کاربرد آن همچنان محصول تصمیم‌های انسانی است (Amirian Farsani & Hosseini, 2023).

از منظر حقوق کیفری، قتل جرمی علیه حیات انسان است که نتیجه اساسی آن سلب حیات دیگری است و مسئولیت مرتکب بسته به ویژگی‌های رفتار، علم، قصد و کیفیت انتساب نتیجه تعیین می‌شود (Haji, 2018). در تحلیل قتل ناشی از هوش مصنوعی نیز نتیجه مادی تفاوتی با دیگر انواع قتل ندارد؛ انسان فوت کرده است و نظام کیفری باید منشأ مرگ را شناسایی کند. تفاوت اصلی در سازوکار تحقق نتیجه است. ممکن است عمل مادی کشنده مستقیماً توسط ربات، خودرو یا سامانه هوشمند انجام شده باشد، در حالی که شخص انسانی در زمان وقوع حادثه هیچ رفتار فیزیکی مستقیمی انجام نداده باشد. از این رو باید میان «عامل فیزیکی ایجادکننده نتیجه» و «شخص قابل انتساب از منظر کیفری» تمایز قائل شد. این تمایز در حقوق فناوری اهمیت زیادی دارد، زیرا هر موجود یا ابزاری که از نظر فیزیکی نتیجه را ایجاد می‌کند، لزوماً مرتکب جرم در معنای حقوقی نیست. هوش مصنوعی می‌تواند واسطه مادی تحقق نتیجه باشد، اما برای مسئول شناختن آن به‌عنوان مرتکب، باید شرایطی فراتر از رابطه فیزیکی با مرگ احراز شود؛ از جمله اینکه نظام حقوقی آن را واجد اهلیت مسئولیت کیفری بداند، امری که در وضعیت فعلی اغلب نظام‌های حقوقی محل پذیرش نیست (Al-Jamili, 2025).

عنصر مادی قتل ناشی از هوش مصنوعی می‌تواند اشکال متنوعی داشته باشد. در یک خودرو خودران، تصمیم سامانه درباره سرعت، ترمز یا تغییر مسیر ممکن است موجب برخورد مرگبار شود. در حوزه پزشکی، سامانه هوشمند ممکن است تشخیص یا توصیه درمانی نادرستی ارائه کند که در صورت اعتماد حرفه‌ای بدون کنترل لازم به مرگ بیمار منتهی شود. در محیط صنعتی، ربات ممکن است در اثر خطای تشخیص یا نقص برنامه‌ریزی، حرکتی انجام دهد که جان کارگر را بگیرد. در هر یک از این موارد، رابطه فیزیکی میان سامانه و نتیجه قابل شناسایی است، اما مسئله حقوق کیفری فراتر از این مرحله قرار دارد و باید بررسی کند کدام رفتار انسانی در زنجیره رخداد از اهمیت حقوقی برخوردار است. خطای طراحی، استفاده از داده‌های نامناسب، بی‌توجهی به استانداردهای ایمنی، عدم نظارت، استفاده برخلاف دستورالعمل، عدم نصب به‌روزرسانی ضروری یا واگذاری تصمیمی که عرفاً نیازمند نظارت انسانی بوده است، می‌تواند به‌عنوان رفتار قابل ارزیابی مطرح شود (Al-Hasanat, 2024). از این‌رو عنصر مادی در این حوزه ممکن است ترکیبی از فعل سامانه و فعل یا ترک فعل چند شخص انسانی باشد.

رابطه سببیت یکی از دشوارترین مباحث در قتل ناشی از هوش مصنوعی است. در مدل کلاسیک، دادگاه بررسی می‌کند که آیا رفتار متهم در وقوع مرگ نقش مؤثر داشته و آیا می‌توان نتیجه را از نظر حقوقی به او منتسب کرد. در سامانه‌های هوشمند، میان رفتار اولیه طراح و مرگ ممکن است مجموعه‌ای از عوامل واسطه‌ای قرار گیرد. برای مثال، خطای برنامه‌نویس تنها در صورتی فعال شود که داده خاصی وارد سامانه گردد؛ کاربر ممکن است برخلاف هشدار فنی از سامانه استفاده کرده باشد؛ سامانه نیز بر اثر یادگیری از داده‌های جدید تصمیمی گرفته باشد که در نسخه اولیه پیش‌بینی نشده بوده است. این زنجیره طولانی سبب می‌شود که رابطه سببیت نتواند تنها براساس توالی زمانی تحلیل شود. لازم است مشخص گردد کدام عامل خطر حقوقاً مرتبط را ایجاد یا افزایش داده و آیا نتیجه واقع شده در محدوده همان خطر قرار داشته است. در ادبیات مسئولیت هوش مصنوعی نیز قابلیت پیش‌بینی و کنترل به‌عنوان دو معیار مهم برای تحلیل انتساب مطرح شده‌اند (Meghdadi & Pourcheriki, 2024).

عنصر روانی در قتل ناشی از هوش مصنوعی از پیچیدگی بیشتری برخوردار است. اگر شخصی عمداً سامانه هوشمند را برای کشتن فردی طراحی یا به کار گیرد، مسئله از حیث عنصر روانی تفاوت اساسی با استفاده از سایر ابزارها ندارد؛ هوش مصنوعی در این وضعیت ابزار تحقق قصد انسانی است. اما مسئله اصلی در مواردی پدیدار می‌شود که هیچ شخصی قصد مرگ نداشته است و نتیجه بر اثر خطا یا رفتار غیرقابل انتظار سامانه ایجاد می‌شود. در چنین وضعیتی، باید بررسی شود که آیا یکی از اشخاص مرتبط خطر مرگ را پیش‌بینی کرده و با وجود آن به رفتار خود ادامه داده است، یا آنکه با رعایت استاندارد متعارف باید چنین خطری را پیش‌بینی می‌کرد. پژوهش درباره مسئولیت کیفری هوش مصنوعی خودمختار نشان می‌دهد که عنصر روانی در این حوزه نمی‌تواند صرفاً از تصمیم نهایی سامانه استنباط شود، بلکه باید به وضعیت ذهنی و تکالیف انسان‌هایی توجه شود که امکان کنترل، نظارت یا مداخله داشته‌اند (Javanmard & Heydarifard, 2025). بنابراین اگر الگوریتم نتیجه‌ای غیرقابل پیش‌بینی ایجاد کرده باشد، نمی‌توان به‌صورت خودکار علم یا قصد را به طراح نسبت داد؛ همان‌گونه که صرف عدم قصد نیز همیشه موجب رفع مسئولیت نیست و ممکن است بی‌احتیاطی یا بی‌مبالاتی وجود داشته باشد.

در این زمینه، تمایز میان عمد، تقصیر آگاهانه و تقصیر ناآگاهانه اهمیت پیدا می‌کند. ممکن است تولیدکننده از نقص مهم سامانه و احتمال ایجاد آسیب شدید آگاه باشد، اما به دلیل ملاحظات اقتصادی عرضه محصول را ادامه دهد. در چنین وضعیتی درجه سرزنش متفاوت از حالتی است که نقصی کاملاً نامتعارف و براساس دانش موجود غیرقابل کشف باشد. همچنین ممکن است بهره‌بردار هشدارهای سامانه را نادیده بگیرد یا آن را در محیطی استفاده کند که برای آن طراحی نشده است. مسئولیت کیفری باید بر پایه سهم واقعی هر شخص در ایجاد خطر و

میزان آگاهی یا قابلیت پیش‌بینی وی تعیین شود (Younes, 2024). استفاده از معیار واحد برای همه اشخاص دخیل می‌تواند به نتایج ناعادلانه منتهی شود، زیرا برنامه‌نویس، تولیدکننده و کاربر از نظر دانش فنی، قدرت کنترل و دسترسی به اطلاعات در موقعیت یکسانی قرار ندارند.

یکی دیگر از مبانی مهم، تمایز میان مباشرت و تسبیب است. در قتل‌های سنتی، مباشر معمولاً کسی است که عمل مادی را مستقیماً انجام می‌دهد و سبب کسی است که شرایط مؤثر در تحقق نتیجه را فراهم می‌سازد. حضور یک سامانه هوشمند ممکن است این تمایز را پیچیده کند، زیرا انجام فیزیکی رفتار کشنده به سامانه نسبت داده می‌شود، در حالی که سامانه در اغلب نظام‌ها شخص مسئول کیفری محسوب نمی‌شود. بنابراین تحلیل باید به سمت شناسایی انسانی حرکت کند که از نظر حقوقی رفتار سامانه قابل انتساب به اوست یا با فعل و ترک فعل خود سبب مؤثر نتیجه شده است. در برخی پژوهش‌های داخلی نیز مسئله تعدد اسباب و فاصله میان رفتار انسانی و نتیجه ناشی از سامانه هوشمند به عنوان یکی از مسائل اصلی مطرح شده است (Makki et al., 2024). قواعد مربوط به اجتماع اسباب می‌تواند نقطه شروع مناسبی باشد، اما پیچیدگی شبکه تولید و بهره‌برداری از هوش مصنوعی ممکن است نیازمند شاخص‌های دقیق‌تر برای تعیین نقش غالب یا مؤثر باشد.

مسئله مسئولیت خود هوش مصنوعی نیز از منظر نظری قابل طرح است. طرفداران چنین رویکردی ممکن است استدلال کنند که افزایش استقلال و توانایی انتخاب سامانه‌ها سبب شده است که نگاه صرفاً ابزاری دیگر پاسخ‌گوی همه موارد نباشد. با این حال مسئولیت کیفری صرفاً پاسخ به توانایی انتخاب فنی نیست. مجازات مفاهیمی چون سرزنش، بازدارندگی، اصلاح و در مواردی ناتوان‌سازی را دنبال می‌کند که اعمال آنها نسبت به یک نرم‌افزار یا سامانه فاقد شخصیت حقوقی و تجربه انسانی با دشواری نظری مواجه است (Zandi & Rafiei Alavi, 2024). همچنین پذیرش مسئولیت مستقیم هوش مصنوعی می‌تواند به‌طور ناخواسته موجب انتقال مسئولیت از انسان‌ها و شرکت‌هایی شود که از فناوری سود می‌برند و امکان کنترل خطر را دارند. از این منظر، شناسایی هوش مصنوعی به عنوان «برده مسئولیت» می‌تواند عدالت کیفری را تضعیف کند. بنابراین تا زمانی که تحولی بنیادین در مفهوم شخصیت و مسئولیت کیفری رخ نداده است، رویکرد مناسب‌تر آن است که هوش مصنوعی به عنوان عامل فناورانه در زنجیره علی تحلیل شود و مسئولیت بر پایه نقش اشخاص انسانی و حقوقی مرتبط تعیین گردد (Al-M'ebid, 2024).

در نهایت، مبانی حقوقی مسئولیت کیفری در قتل ناشی از هوش مصنوعی باید بر ترکیبی از اصول سنتی و معیارهای متناسب با فناوری استوار باشد. اصولی مانند شخصی بودن مسئولیت، ضرورت احراز رابطه سببیت، منع مسئولیت صرفاً عینی و لزوم اثبات عنصر روانی همچنان باید حفظ شوند. در عین حال، معیارهایی نظیر سطح کنترل بر سامانه، دانش فنی قابل انتظار، قابلیت پیش‌بینی خطر، وظیفه نظارت، امکان مداخله، سطح خودمختاری سامانه و میزان انحراف رفتار واقعی از کارکرد مورد انتظار باید در تحلیل وارد شوند. مطالعات تطبیقی در زمینه مسئولیت کیفری هوش مصنوعی نیز نشان می‌دهد که رویکردهای جدید عمدتاً به جای کنار گذاشتن اصول عمومی، در پی انطباق آنها با ساختارهای جدید فناوری هستند (Kaveh & Barani, 2024). این مبانی زمینه لازم را برای بررسی دقیق‌تر چالش‌های انتساب قتل ناشی از هوش مصنوعی فراهم می‌کند.

چالش‌های انتساب قتل ناشی از عملکرد هوش مصنوعی

مهم‌ترین مسئله در مسئولیت کیفری ناشی از عملکرد هوش مصنوعی، چگونگی انتساب نتیجه به یک شخص معین است. انتساب در حقوق کیفری مفهومی فراتر از شناسایی علت فیزیکی نتیجه است. ممکن است یک سامانه هوشمند از نظر فنی عامل مستقیم مرگ باشد، اما حقوق کیفری باید مشخص کند که آیا رفتار و نتیجه را می‌توان به انسانی منتسب کرد که شرایط لازم برای تحمل مسئولیت کیفری را داراست. در سیستم‌های سنتی، مسیر انتساب غالباً کوتاه‌تر است؛ مرتکب رفتار را انجام می‌دهد و نتیجه نیز در ادامه همان رفتار شکل می‌گیرد. در هوش مصنوعی، زنجیره ممکن است شامل طراحی، گردآوری داده، آموزش مدل، آزمون، تولید سخت‌افزار، عرضه، نصب، تنظیمات کاربر، به‌روزرسانی و تصمیم خودکار سامانه باشد. هر حلقه می‌تواند در نتیجه نقش داشته باشد و در عین حال هیچ‌یک به تنهایی علت کامل آن نباشد. این پراکندگی مسئولیت به یکی از دشوارترین ویژگی‌های حقوق کیفری هوش مصنوعی تبدیل شده است (Ghavamipour, Sarshkeh & Mahmoudi, 2024).

یکی از چالش‌ها، خودمختاری نسبی سامانه است. هرچه رفتار سامانه به اجرای مستقیم فرمان‌های انسانی نزدیک‌تر باشد، انتساب نیز ساده‌تر است. اگر کاربر دستور مشخصی برای انجام عمل‌کننده صادر کند، سامانه صرفاً وسیله اجرای اراده است. اما در سامانه‌های یادگیرنده، تصمیم نهایی ممکن است از تعامل میان مدل، داده‌ها و شرایط محیطی شکل گیرد. در چنین حالتی، طراح ممکن است منطق کلی را ایجاد کرده باشد، اما خروجی خاصی که به مرگ منجر شده است مستقیماً توسط او تعیین نشده باشد. پژوهش‌ها در زمینه مسئولیت هوش مصنوعی خودمختار نشان می‌دهند که فاصله گرفتن خروجی از تصمیم مستقیم انسان یکی از مبانی طرح مجدد مفهوم عنصر روانی و انتساب کیفری است (Javanmard & Heydarifard, 2025). باین‌حال، خودمختاری فنی نباید به‌تنهایی موجب قطع انتساب شود. سؤال اصلی آن است که آیا انسانی در زنجیره وجود داشته که خطر را ایجاد کرده، آن را مدیریت نکرده یا امکان جلوگیری از نتیجه را داشته است. چالش دیگر، قابلیت پیش‌بینی است. مسئولیت ناشی از تقصیر تا حد زیادی بر این ایده استوار است که شخص باید خطری را که فرد متعارف در شرایط مشابه می‌توانست یا باید پیش‌بینی می‌کرد، مورد توجه قرار دهد. در هوش مصنوعی، تعیین سطح پیش‌بینی مورد انتظار دشوار است. توسعه‌دهنده متخصص ممکن است موظف باشد رفتارهای نامتعارفی را پیش‌بینی کند که برای کاربر عادی کاملاً ناشناخته است. از سوی دیگر، نمی‌توان از برنامه‌نویس انتظار داشت تمام وضعیت‌های ممکن یک سامانه پیچیده را پیش‌بینی کند. بنابراین پیش‌بینی‌پذیری باید متناسب با نقش و تخصص هر شخص سنجیده شود (Al-Falasi, 2023). اگر خطر مرگ ناشی از یک نقص شناخته‌شده بوده و شرکت با وجود گزارش‌های مکرر از اصلاح آن خودداری کرده باشد، مسئولیت به‌مراتب قابل توجیه‌تر از حالتی است که رفتار کاملاً جدیدی بر اثر تعامل پیش‌بینی‌ناپذیر عوامل متعدد ایجاد شده باشد.

«مسئله جعبه سیاه» نیز انتساب را دشوار می‌کند. در برخی مدل‌های هوش مصنوعی، می‌توان خروجی سامانه را مشاهده کرد، اما توضیح دقیق مسیر داخلی شکل‌گیری آن دشوار است. این ویژگی در دادرسی کیفری اهمیت اساسی دارد، زیرا دادگاه باید رابطه میان نقص ادعایی و نتیجه مرگ را احراز کند. اگر مشخص نباشد که چرا سامانه تصمیم خاصی گرفته است، اثبات اینکه کدام خطای طراحی یا داده‌ای علت تصمیم بوده دشوار می‌شود. این مشکل در حوزه پزشکی برجسته‌تر است، زیرا ممکن است الگوریتم تشخیصی نتیجه‌ای ارائه کند که پزشک بر آن تکیه کرده و بعداً مشخص شود اشتباه بوده است. در چنین وضعیتی باید معلوم شود که آیا مسئولیت بر عهده طراح سامانه، بیمارستان، پزشک یا

مجموعه‌ای از آنان است. بررسی مسئولیت هوش مصنوعی در حوزه سلامت نیز نشان داده است که توزیع نقش میان سیستم تصمیم‌یار و متخصص انسانی می‌تواند تعیین مسئول نهایی را بسیار پیچیده کند (Goel, 2023).

داده‌های آموزشی نیز می‌توانند منشأ مستقل خطر باشند. هوش مصنوعی معمولاً از داده برای ساخت الگو استفاده می‌کند و کیفیت داده مستقیماً بر عملکرد آن اثر دارد. اگر داده‌ها ناقص، نامتوازن، آلوده یا غیرمتناسب با محیط واقعی باشند، خروجی سامانه ممکن است خطرناک شود. در یک سامانه پزشکی، داده نامناسب می‌تواند تشخیص اشتباه ایجاد کند و در یک خودروی خودران، ضعف داده‌های مربوط به شرایط جوی یا اشیای خاص می‌تواند موجب عدم تشخیص خطر شود. پرسش حقوقی آن است که مسئولیت چنین نقضی بر عهده چه کسی است: گردآورنده داده، تیم توسعه، شرکت تولیدکننده یا نهادی که سامانه را بدون آزمون کافی به کار گرفته است. پاسخ باید براساس تقسیم وظایف و میزان امکان تشخیص نقص تعیین شود. مطالعات حقوقی جدید نیز بر ضرورت حرکت از تصور مسئولیت فردی ساده به سمت تحلیل زنجیره مسئولیت در سامانه‌های هوشمند تأکید دارند (Al-Halahlah, 2025).

پیچیدگی رابطه سببیت هنگامی افزایش می‌یابد که خطای انسانی و تصمیم سامانه هم‌زمان وجود داشته باشند. فرض شود خودروی خودران دارای نقص جزئی در سامانه تشخیص مانع است، اما مالک نیز به‌رغم هشدارهای مکرر فنی خودرو را تعمیر نکرده است. در زمان حادثه نیز عابر رفتار غیرمنتظره‌ای انجام می‌دهد. در چنین حالتی چند عامل در تحقق مرگ نقش دارند. حقوق کیفری باید تعیین کند کدام عوامل از نظر هنجاری به اندازه کافی مهم‌اند که بتوان نتیجه را به آنها منتسب کرد. صرف اینکه «بدون» یک عامل نتیجه رخ نمی‌داد، همیشه برای مسئولیت کیفری کافی نیست؛ زیرا این معیار ممکن است تعداد بسیار زیادی از عوامل دور را وارد زنجیره کند. لازم است میان علت واقعی و علت قابل انتساب حقوقی تمایز گذاشته شود. در این زمینه، قواعد مربوط به تسبیب در قتل و ارزیابی نقش اسباب می‌تواند مبنای اولیه تحلیل باشد (Nozari Ferdowsieh et al., 2022)، اما ویژگی‌های چندلایه فناوری احتمالاً نیازمند تکمیل این قواعد با معیار خطر و کنترل است.

مسئله دیگری که می‌تواند انتساب را مخدوش کند، اعتماد بیش از حد انسان به سامانه هوشمند است. در بسیاری از حوزه‌ها هوش مصنوعی نقش تصمیم‌یار دارد نه تصمیم‌گیر نهایی؛ با این حال کاربران ممکن است در عمل پیشنهاد سامانه را بدون بررسی بپذیرند. اگر پزشک توصیه الگوریتم را بدون ارزیابی مستقل اجرا کند و بیمار فوت کند، نمی‌توان صرف وجود هوش مصنوعی را موجب رفع مسئولیت حرفه‌ای دانست. از سوی دیگر، اگر سامانه با سطح اعتبار بسیار بالا توسط یک مؤسسه رسمی عرضه شده و خطایی ایجاد کرده که برای کاربر متخصص نیز قابل تشخیص نبوده است، تحمیل تمام مسئولیت بر کاربر عادلانه نیست. بنابراین باید میان «اتکای معقول» و «اتکای غیرموجه» به هوش مصنوعی تمایز قائل شد (Goel, 2023). این تمایز در آینده اهمیت بیشتری خواهد یافت، زیرا هرچه سامانه‌ها دقیق‌تر شوند، اعتماد انسان نیز افزایش می‌یابد و مرز وظیفه نظارت انسانی می‌تواند مبهم شود.

یکی از چالش‌های اساسی دیگر، احراز عنصر روانی در شرکت‌ها و اشخاص حقوقی است. بسیاری از سامانه‌های هوش مصنوعی توسط شرکت‌ها طراحی و عرضه می‌شوند و تصمیم‌های مربوط به ایمنی، بودجه آزمون، سرعت ورود محصول به بازار و مدیریت هشدارها در سطح سازمانی اتخاذ می‌شود. ممکن است هیچ فرد واحدی به تنهایی همه اطلاعات لازم درباره خطر را در اختیار نداشته باشد، اما مجموعه سازمان از طریق واحدهای مختلف نسبت به خطر آگاهی داشته باشد. در چنین حالتی مسئولیت فردی مدیر، مهندس یا کارمند ممکن است به دلیل پراکندگی دانش دشوار باشد. این مسئله ضرورت توجه به مسئولیت کیفری اشخاص حقوقی و فرهنگ سازمانی را مطرح می‌سازد (Al-

(Sadat Makki et al., 2024). اگر شرکت آگاهانه سامانه‌ای را بدون آزمون کافی عرضه کرده یا گزارش‌های خطر را به‌صورت ساختاری نادیده گرفته باشد، محدود کردن تحلیل به تقصیر یک برنامه‌نویس خاص ممکن است تصویر واقعی مسئولیت را منعکس نکند. در کنار این مسائل، نظریه مسئولیت مستقیم هوش مصنوعی گاه به‌عنوان راه‌حلی برای خلأ انتساب مطرح می‌شود. استدلال می‌شود که اگر سامانه قادر است بدون فرمان لحظه‌ای انسان رفتار کند، شاید بتوان آن را مسئول نتیجه دانست. این رویکرد با موانع بنیادین روبه‌روست. مسئولیت کیفری مبتنی بر قابلیت سرزنش است و سرزنش مستلزم نوعی آگاهی هنجاری، فهم ممنوعیت و توانایی انتخاب در معنایی است که فراتر از پردازش محاسباتی باشد. پژوهش‌های فلسفی درباره امکان مسئولیت هوش مصنوعی نیز نشان داده‌اند که تشبیه تصمیم محاسباتی به اراده انسانی هنوز از منظر مبانی کیفری محل مناقشه است (Zandi & Rafiei Alavi, 2024). افزون بر این، پذیرش مسئولیت هوش مصنوعی می‌تواند سبب شود تولیدکنندگان و بهره‌برداران به سادگی رفتار خطرناک را به «تصمیم مستقل ماشین» نسبت دهند و از پاسخ‌گویی بگریزند. از این رو رویکرد محتاطانه اقتضا می‌کند که استقلال عملیاتی سامانه به‌عنوان عامل مؤثر در تعیین دامنه مسئولیت انسان‌ها لحاظ شود، نه آنکه بهانه‌ای برای حذف آنان از زنجیره مسئولیت باشد.

چالش مهم دیگر تعیین نوع قتل است. در برخی موارد، سامانه صرفاً ابزار ارتکاب قتل عمدی است. شخصی ممکن است ربات، پهپاد یا سامانه دیگری را عمداً برای کشتن فرد برنامه‌ریزی کند و در این حالت انتساب قتل به او دشواری خاصی ندارد. اما بیشتر مسائل نوظهور مربوط به مرگ‌هایی است که بدون قصد مستقیم رخ می‌دهند. باید بررسی کرد آیا شخص انسانی از احتمال نتیجه آگاه بوده، آیا رفتار وی نقض فاحش تکلیف مراقبت محسوب می‌شود، آیا خطر مرگ عرفاً قابل پیش‌بینی بوده و آیا می‌توان نوعی بی‌احتیاطی یا بی‌مبالاتی کیفری را احراز کرد. ادبیات حقوقی مرتبط با ویژگی‌های خاص مسئولیت هوش مصنوعی تأکید دارد که سطح خودمختاری نباید باعث کنار گذاشتن بررسی دقیق حالات روانی انسان‌های درگیر شود (Younes, 2024). بنابراین طبقه‌بندی قتل باید بر رفتار و وضعیت ذهنی اشخاص مسئول استوار باشد، نه صرفاً بر ماهیت پیشرفته فناوری.

در نهایت، حل مسئله انتساب مستلزم پذیرش یک رویکرد چندعاملی است. هیچ معیار واحدی مانند مالکیت، طراحی یا استفاده نمی‌تواند در همه پرونده‌ها پاسخ مناسب ارائه دهد. ممکن است در یک پرونده تولیدکننده بیشترین کنترل و آگاهی را داشته باشد، در پرونده‌ای دیگر بهره‌بردار با استفاده خلاف دستورالعمل عامل اصلی خطر باشد و در پرونده‌ای دیگر ترکیبی از قصور سازمانی و نقص طراحی به نتیجه منجر شود. از این رو باید عواملی مانند قابلیت پیش‌بینی، میزان کنترل، دسترسی به اطلاعات، توانایی پیشگیری، نقض استاندارد فنی، نقش در ایجاد خطر و فاصله میان رفتار و نتیجه به‌صورت هم‌زمان بررسی شود. این رویکرد با گرایش جدید پژوهش‌ها به ایجاد چارچوب‌های چندلایه برای مسئولیت سامانه‌های هوش مصنوعی هماهنگ است (Rostami Zabol, 2025).

تعیین شخص مسئول در قتل ناشی از هوش مصنوعی

تعیین شخص مسئول در قتل ناشی از هوش مصنوعی باید با شناسایی همه اشخاصی آغاز شود که در چرخه حیات سامانه نقش دارند. برخلاف ابزارهای ساده، هوش مصنوعی معمولاً حاصل فعالیت یک شخص واحد نیست. شرکت یا مؤسسه‌ای منابع مالی را تأمین می‌کند، مهندسان معماری سامانه را طراحی می‌کنند، تیم‌های دیگر داده جمع‌آوری و پالایش می‌کنند، برنامه‌نویسان مدل را توسعه می‌دهند، واحدهای کنترل کیفیت آن را آزمایش می‌کنند، شرکت دیگری ممکن است سخت‌افزار را تولید کند و نهایتاً کاربر یا سازمان بهره‌بردار سامانه را در محیط واقعی به کار گیرد. مسئولیت کیفری نمی‌تواند صرفاً بر اساس نزدیکی فیزیکی به حادثه تعیین شود. شخصی که در زمان مرگ در محل حضور ندارد

ممکن است تصمیمی گرفته باشد که نقش تعیین‌کننده در ایجاد خطر داشته است. به همین دلیل، تحلیل مسئولیت باید از «لحظه حادثه» به «چرخه ایجاد و مدیریت خطر» گسترش یابد (Ghavami Pour Sereshkeh & Mahmoudi, 2025).

برنامه‌نویس یا طراح الگوریتم نخستین فردی است که ممکن است به‌عنوان مسئول احتمالی مطرح شود، اما مسئول شناختن او صرفاً به دلیل مشارکت در ساخت سامانه قابل قبول نیست. باید مشخص شود که خطای قابل انتسابی در طراحی وجود داشته و این خطا با نتیجه مرگ رابطه مؤثر داشته است. اگر برنامه‌نویس برخلاف استانداردهای فنی شناخته‌شده، آزمون‌های ضروری را نادیده گرفته یا از نقصی جدی آگاه بوده و آن را گزارش نکرده باشد، امکان احراز تقصیر کیفری افزایش می‌یابد. در مقابل، اگر وی مطابق دانش و استانداردهای متعارف عمل کرده و رفتار مرگبار ناشی از پدیده‌ای کاملاً جدید و غیرقابل پیش‌بینی بوده باشد، تحمیل مسئولیت کیفری دشوار خواهد بود. ادبیات مربوط به موانع مسئولیت کیفری ربات‌های هوشمند نیز نشان داده است که فاصله میان طراحی اولیه و رفتار نهایی سامانه یکی از عوامل محدودکننده مسئولیت طراح است (Amirian Farsani & Hosseini, 2023). بنابراین مسئولیت برنامه‌نویس باید مبتنی بر تقصیر مشخص و اثبات‌شده باشد.

تولیدکننده و شرکت توسعه‌دهنده در بسیاری از موارد نقش گسترده‌تری از یک برنامه‌نویس منفرد دارند. شرکت مسئول انتخاب معماری، سیاست‌های ایمنی، سطح آزمون، زمان عرضه، نحوه اطلاع‌رسانی خطرها و نظام پایش پس از عرضه است. اگر محصول دارای خطر شناخته‌شده‌ای باشد و شرکت با وجود امکان رفع نقص آن را نادیده بگیرد، این رفتار می‌تواند مبنای قوی‌تری برای مسئولیت باشد. همچنین اگر شرکت آگاهانه سامانه را برای کاربردی با سطح خطر بالا عرضه کند، در حالی که آزمون‌های کافی انجام نشده‌اند، نتیجه مرگبار می‌تواند در محدوده خطر ایجادشده قرار گیرد (Al-Jamili, 2025). این مسئله نشان می‌دهد که در برخی پرونده‌ها مسئولیت اصلی ممکن است نه در سطح مهندس فنی، بلکه در سطح تصمیم‌گیری سازمانی باشد.

شخص حقوقی نیز باید در تحلیل جایگاه مستقل داشته باشد. سامانه‌های هوش مصنوعی غالباً توسط شرکت‌های بزرگ و ساختارهای پیچیده سازمانی تولید می‌شوند و ممکن است نقص ایمنی ناشی از تصمیم جمعی یا سیاست سازمان باشد. برای مثال، کاهش هزینه آزمون، محدود کردن زمان ارزیابی محصول یا بی‌اعتنایی به گزارش کارکنان ممکن است تصمیم‌هایی باشند که به‌صورت سازمانی اتخاذ شده‌اند. در چنین وضعیتی، جست‌وجوی یک فرد خاص که به تنهایی واجد تمام عناصر تقصیر باشد ممکن است منجر به نادیده گرفتن مسئولیت واقعی شود. مطالعات حقوقی ایران درباره مسئولیت ناشی از اعمال هوش مصنوعی نیز بر اهمیت بررسی نقش اشخاص حقوقی در کنار اشخاص حقیقی تأکید کرده‌اند (Makki et al., 2024). بدین ترتیب، هرچند نظام حقوقی شرایط مسئولیت کیفری شخص حقوقی را فراهم بداند، این ظرفیت می‌تواند در پرونده‌های هوش مصنوعی اهمیت ویژه‌ای پیدا کند.

مالک سامانه هوشمند نیز همیشه مسئول نیست، اما مالکیت می‌تواند در کنار کنترل و تکلیف مراقبت اهمیت یابد. اگر فردی مالک یک دستگاه هوشمند باشد اما هیچ دانش یا کنترل واقعی بر نقص فنی آن نداشته باشد، صرف مالکیت نباید سبب مسئولیت کیفری شود. در مقابل، اگر مالک هشدارهای مکرر ایمنی را دریافت کرده باشد، از نقص خطرناک مطلع باشد و با وجود آن استفاده را ادامه دهد، مسئولیت می‌تواند مطرح شود. در خودروهای خودران، برای مثال، بی‌توجهی مالک به الزام به‌روزرسانی نرم‌افزار یا تعمیر حسگرهای حیاتی ممکن است در وقوع حادثه مرگبار نقش اساسی داشته باشد. مسئولیت کیفری در چنین حالتی از مالکیت ناشی نمی‌شود، بلکه از ترک تکلیف مراقبتی ناشی می‌شود که با قابلیت پیش‌بینی خطر همراه بوده است (Meghdadi & Pourcheriki, 2024).

کاربر یا بهره‌بردار یکی دیگر از مهم‌ترین اشخاص در زنجیره مسئولیت است. سطح مسئولیت کاربر به نوع سامانه بستگی دارد. در برخی سامانه‌ها کاربر تنها دستور کلی را تعیین می‌کند و کنترل عملیاتی عمدتاً با ماشین است، اما در برخی دیگر کاربر باید به‌طور مستمر نظارت کند و در شرایط خطرناک مداخله نماید. اگر کاربر از سامانه در شرایط ممنوع استفاده کند، محدودیت‌های فنی آن را نادیده بگیرد یا عمداً کنترل ایمنی را غیرفعال نماید، ارتباط رفتار او با نتیجه قوی خواهد بود. در مقابل، اگر طراحی سامانه به‌گونه‌ای باشد که کاربر معقولانه تصور کند نیازی به مداخله نیست، نمی‌توان پس از حادثه تمام مسئولیت را به او منتقل کرد. قوانین و دستورالعمل‌های آینده باید به‌طور روشن سطح نظارت مورد انتظار از کاربران را تعیین کنند تا مسئولیت پس از حادثه بر مبنای استانداردهای مبهم شکل نگیرد (Kaveh & Barani, 2024).

در خودروهای خودران، مسئله توزیع مسئولیت نمونه روشنی از پیچیدگی موضوع است. اگر سامانه در تشخیص عابر دچار خطا شود، باید بررسی شود آیا نقص از حسگر، الگوریتم، داده، تنظیمات، نگهداری یا رفتار کاربر ناشی بوده است. ممکن است تولیدکننده از محدودیت سامانه در شرایط خاص جوی اطلاع داشته ولی هشدار کافی ارائه نکرده باشد. ممکن است راننده نیز برخلاف الزام قانونی توجهی به مسیر نداشته باشد. در چنین حالتی مسئولیت می‌تواند مشترک باشد و تحلیل باید بر سهم هر عامل در ایجاد خطر متمرکز شود. رویکردی که به‌طور خودکار «راننده» یا «سازنده» را مسئول بداند، با پیچیدگی واقعی فناوری سازگار نیست. چارچوب‌های جدید مسئولیت هوش مصنوعی نیز بر ارزیابی سطح کنترل و قابلیت پیشگیری تأکید می‌کنند (Ghavamipour Sarshkeh & Mahmoudi, 2024).

در حوزه پزشکی، موضوع حساس‌تر است، زیرا تصمیم هوش مصنوعی با داور پزشکی درهم می‌آمیزد. سامانه ممکن است تشخیص، دوز دارو یا روش درمان را پیشنهاد کند، اما پرسش آن است که آیا پزشک مکلف به بررسی مستقل است یا می‌تواند به سامانه اعتماد کند. اگر الگوریتم صرفاً ابزار مشورتی باشد، معمولاً انتظار می‌رود متخصص انسانی تصمیم نهایی را کنترل کند. در عین حال، در مواردی که سامانه از نظر فنی بسیار پیچیده‌تر از توان ارزیابی پزشک است، نمی‌توان مسئولیت نقص پنهان را تماماً بر متخصص نهایی تحمیل کرد. مطالعات حوزه سلامت نشان داده‌اند که تعیین مسئولیت در این وضعیت نیازمند تمایز میان خطای الگوریتم، خطای استفاده و اعتماد حرفه‌ای متعارف است (Goel, 2023). بنابراین در پرونده قتل ناشی از خطای پزشکی مبتنی بر هوش مصنوعی، باید استاندارد حرفه‌ای، اطلاعات ارائه‌شده توسط تولیدکننده، قابلیت کشف خطا و میزان استقلال واقعی پزشک بررسی شود.

ربات‌های صنعتی نیز نمونه دیگری هستند. ربات ممکن است به‌دلیل نقص نرم‌افزاری، تغییر محیط یا داده حسگر حرکتی انجام دهد که به مرگ کارگر منتهی شود. در این حوزه، مسئولیت ممکن است میان تولیدکننده، کارفرما، تیم تعمیر و نگهداری و اپراتور توزیع شود. اگر کارفرما حفاظت ایمنی را برای افزایش سرعت تولید غیرفعال کرده باشد، مسئولیت او برجسته خواهد بود. اگر نقص طراحی شناخته‌شده‌ای وجود داشته باشد، تولیدکننده در معرض مسئولیت قرار می‌گیرد. اگر تعمیرکار تغییر غیرمجاز ایجاد کرده باشد، تحلیل متفاوت است. این نمونه‌ها نشان می‌دهند که «نزدیک‌ترین شخص به ربات» لزوماً مسئول اصلی نیست و باید زنجیره خطر از طراحی تا استفاده بررسی شود (Al-Hasanat, 2024).

در حوزه سلاح‌های خودکار، مسئله دارای حساسیت مضاعف است، زیرا هدف فناوری از ابتدا اعمال نیروی بالقوه کشنده است. اگر یک سامانه براساس داده یا الگوریتم معیوب هدفی را اشتباه شناسایی کند، مسئله مسئولیت می‌تواند به طراح، فرمانده، اپراتور یا سازمان استفاده‌کننده مرتبط شود. سطح نظارت انسانی در این حوزه اهمیت بنیادین دارد. واگذاری کامل تصمیم مرگ و زندگی به سامانه‌ای که امکان خطای

شناخته شده دارد، می‌تواند خود به‌عنوان رفتار پرخطر ارزیابی شود. هرچه پیامد بالقوه یک سامانه شدیدتر باشد، استاندارد مراقبت مورد انتظار از توسعه‌دهندگان و بهره‌برداران نیز باید بالاتر باشد. این منطق با رویکردهای مبتنی بر تنظیم فناوری‌های پرخطر هماهنگ است (Al-M'ebid, 2024).

یکی از معیارهای مناسب برای تعیین مسئول، «کنترل مؤثر» است. شخصی که بیشترین توان واقعی برای هدایت، متوقف کردن یا اصلاح سامانه داشته است، از نظر حقوقی موقعیت متفاوتی نسبت به شخصی دارد که هیچ امکان مداخله‌ای نداشته است. باین‌حال کنترل باید در زمان مناسب ارزیابی شود. ممکن است طراح پس از عرضه هیچ کنترلی نداشته باشد، اما در مرحله طراحی فرصت کامل برای حذف خطر داشته باشد. تولیدکننده ممکن است با به‌روزرسانی از راه دور امکان اصلاح نقص را داشته باشد. کاربر نیز ممکن است در لحظه حادثه امکان توقف سامانه را داشته باشد. بنابراین کنترل مفهومی مرحله‌ای است و باید در کل چرخه فناوری بررسی شود (Rostami Zabol, 2025).

معیار دیگر «قابلیت پیش‌بینی خطر» است. مسئولیت کیفری باید بیش از همه بر اشخاصی متمرکز شود که وقوع نتیجه برای آنان، با توجه به دانش تخصصی و اطلاعات موجود، قابل پیش‌بینی بوده است. شرکت تولیدکننده نمی‌تواند به همان استاندارد کاربر عادی سنجیده شود، زیرا به داده‌های آزمایش، گزارش‌های فنی و ساختار داخلی سامانه دسترسی بیشتری دارد. از سوی دیگر، نمی‌توان انتظار داشت هر برنامه‌نویس جزء از تصمیم‌های مدیریتی یا کارکرد کل سامانه اطلاع داشته باشد. فردی‌سازی استاندارد پیش‌بینی از ضروریات عدالت کیفری در این حوزه است (Younes, 2024).

«ایجاد خطر نامتعارف» نیز می‌تواند معیار مکمل باشد. اگر شخصی سامانه‌ای را در شرایطی استفاده کند که خطر بسیار بیشتر از کارکرد معمول ایجاد می‌کند، حتی بدون نقص طراحی ممکن است مسئول باشد. برای مثال، استفاده از سامانه آزمایشی در محیط عمومی بدون تدابیر حفاظتی می‌تواند ایجاد خطر غیرموجه محسوب شود. همین منطق درباره شرکت تولیدکننده‌ای نیز صادق است که محصول را پیش از تکمیل ارزیابی ایمنی وارد بازار می‌کند. تمرکز بر خطر به دادگاه اجازه می‌دهد به جای جست‌وجوی صرف علت فیزیکی، بررسی کند چه کسی وضعیت خطرناک را ایجاد یا تحمل کرده است.

معیار مهم دیگر «نقض تکلیف مراقبت» است. مسئولیت کیفری ناشی از بی‌احتیاطی یا بی‌مبالاتی تنها زمانی موجه است که مشخص باشد شخص تکلیفی داشته و آن را نقض کرده است. تعیین این تکلیف نیازمند استانداردهای فنی و حقوقی روشن است. بدون مقررات مشخص، دادگاه ممکن است پس از حادثه استنادی را به‌صورت پسینی ایجاد کند که برای فعالان فناوری از قبل قابل پیش‌بینی نبوده است. توسعه قواعد تخصصی برای آزمون، مستندسازی، نظارت انسانی، ثبت تصمیمات الگوریتمی و گزارش نقض می‌تواند هم ایمنی را افزایش دهد و هم معیار مسئولیت را شفاف سازد (Al-Halahlah, 2025).

در کنار مسئولیت فردی، امکان تعدد مسئولان نیز باید پذیرفته شود. در یک حادثه ممکن است نقص طراحی، ضعف نظارت سازمانی و استفاده نادرست کاربر هم‌زمان وجود داشته باشند. در چنین حالتی لزومی ندارد حقوق کیفری به دنبال یک مسئول منحصر باشد. اگر رفتار چند شخص هر یک به‌طور مؤثر در نتیجه نقش داشته و شرایط تقصیر در مورد آنان احراز شود، مسئولیت می‌تواند براساس قواعد مربوط به اجتماع اسباب و مشارکت در ایجاد نتیجه تحلیل شود. مباحث فقهی و حقوقی مربوط به انتساب قتل در وضعیت دخالت اشخاص غیرمباشر نیز نشان می‌دهد که صرف غیرمستقیم بودن نقش مانع مسئولیت نیست، مشروط بر اینکه ارتباط حقوقی کافی با نتیجه وجود داشته باشد (Nozari, 2022).

از نظر سیاست جنایی، یکی از خطرهای مهم آن است که پیچیدگی فنی سبب «خلاء مسئولیت» شود؛ یعنی همه بازیگران بتوانند نتیجه را به عامل دیگری نسبت دهند. برنامه‌نویس می‌تواند به تصمیم شرکت اشاره کند، شرکت به رفتار خودمختار سامانه، سامانه به داده‌ها و بهره‌بردار به اعتماد خود به فناوری. اگر نظام حقوقی معیار روشنی نداشته باشد، این زنجیره ممکن است نهایتاً به عدم مسئولیت هیچ‌کس منجر شود. از سوی دیگر، راه‌حل نباید مسئولیت مطلق باشد، زیرا چنین رویکردی اصول بنیادین حقوق کیفری را نقض و نوآوری مشروع را تهدید می‌کند. راهکار متوازن، تعریف دقیق تکالیف متناسب با نقش هر شخص و سپس ارزیابی نقض این تکالیف در پرتو نتیجه واقع شده است (Ghavami, Pour Sereshkeh & Mahmoudi, 2025).

در حقوق ایران، توسعه چارچوب مشخص برای سامانه‌های پرخطر می‌تواند به تعیین مسئول کمک کند. می‌توان سطحی از تعهدات ویژه برای تولیدکنندگان، بهره‌برداران حرفه‌ای و مالکان سامانه‌های خاص تعریف کرد؛ از جمله ارزیابی خطر پیش از عرضه، ثبت تغییرات الگوریتم، نگهداری گزارش تصمیمات، الزام به به‌روزرسانی ایمنی و تعیین سطح نظارت انسانی. چنین قواعدی نه تنها نقش پیشگیرانه دارد، بلکه پس از وقوع حادثه نیز امکان شناسایی نقطه شکست را فراهم می‌کند. پژوهش‌های داخلی نیز بر نیاز به تقنین متناسب با ویژگی‌های سامانه‌های خودمختار و پرهیز از اتکای کامل به قواعد عمومی تأکید کرده‌اند (Rostami Zabol, 2025). بنابراین مسئولیت قتل ناشی از هوش مصنوعی باید در چارچوبی تحلیل شود که هم اصول کلاسیک مسئولیت را حفظ کند و هم ساختار پیچیده فناوری را به رسمیت بشناسد.

نتیجه‌گیری

گسترش سامانه‌های هوش مصنوعی به حوزه‌هایی که مستقیماً با جان و سلامت انسان ارتباط دارند، مفهوم مسئولیت کیفری را در برابر یکی از مهم‌ترین آزمون‌های خود قرار داده است. مسئله اصلی در قتل ناشی از هوش مصنوعی آن نیست که آیا نتیجه مرگبار می‌تواند از نظر فیزیکی به یک سامانه هوشمند مرتبط باشد، بلکه این است که چگونه می‌توان این نتیجه را در چارچوب اصول حقوق کیفری به شخص یا اشخاصی که واجد قابلیت مسئولیت هستند منتسب کرد. تفاوت اصلی هوش مصنوعی با ابزارهای سنتی در آن است که سامانه می‌تواند میان فرمان اولیه انسان و نتیجه نهایی نقش تصمیم‌ساز و پردازشگر فعالی ایفا کند و در برخی موارد رفتار آن به‌سادگی قابل پیش‌بینی یا توضیح نباشد. همین امر سبب می‌شود رابطه خطی میان تصمیم انسانی و نتیجه جای خود را به زنجیره‌ای از تصمیم‌ها، داده‌ها، طراحی‌ها و مداخلات بدهد. تحلیل انجام‌شده نشان می‌دهد که هوش مصنوعی در وضعیت فعلی نباید صرفاً به دلیل سطح بالای خودمختاری به‌عنوان مرتکب مستقل قتل در معنای حقوق کیفری تلقی شود. مسئولیت کیفری مستلزم بیش از توانایی تولید خروجی یا انتخاب میان گزینه‌هاست و مفاهیمی همچون قابلیت سرزنش، فهم هنجاری و تحمل مجازات در ساختار آن نقش بنیادی دارند. پذیرش مسئولیت مستقیم هوش مصنوعی، بدون حل این مسائل نظری، نه تنها دشوار است، بلکه می‌تواند به انتقال ناموجه مسئولیت از اشخاصی منجر شود که فناوری را طراحی، تولید، عرضه و بهره‌برداری کرده‌اند. از این رو، رویکرد مناسب در شرایط کنونی آن است که هوش مصنوعی به‌عنوان عامل فناورانه مؤثر در زنجیره تحقق نتیجه بررسی شود و تمرکز اصلی بر شناسایی رفتار انسانی یا سازمانی قابل انتساب قرار گیرد.

در مواردی که هوش مصنوعی صرفاً ابزار تحقق قصد مجرمانه انسان است، قواعد موجود مسئولیت کیفری اصولاً پاسخ‌گو هستند. اگر فردی آگاهانه سامانه‌ای را برای کشتن دیگری برنامه‌ریزی یا به کار گیرد، استقلال فنی ابزار تأثیری در ماهیت عمدی رفتار ندارد. دشواری واقعی در مواردی پدیدار می‌شود که هیچ‌کس مستقیماً قصد مرگ نداشته و نتیجه بر اثر خطای طراحی، نقص داده، ضعف نظارت، تصمیم الگوریتمی یا

مجموعه‌ای از این عوامل ایجاد شده باشد. در این موارد، تعیین نوع قتل و شخص مسئول باید بر پایه وضعیت ذهنی و رفتاری اشخاص مرتبط انجام شود و شدت نتیجه نمی‌تواند جایگزین اثبات تقصیر شود.

بررسی زنجیره ایجاد فناوری نشان داد که مسئولیت ممکن است حسب مورد بر عهده برنامه‌نویس، طراح، تولیدکننده، شرکت توسعه‌دهنده، شخص حقوقی، مالک، بهره‌بردار یا کاربر قرار گیرد. هیچ‌یک از این عناوین به‌تنهایی و به‌طور خودکار موجب مسئولیت نیستند. برنامه‌نویس تنها زمانی باید مسئول شناخته شود که خطای مشخص و قابل انتساب او در ایجاد نتیجه مؤثر بوده باشد. تولیدکننده زمانی در معرض مسئولیت جدی‌تر قرار می‌گیرد که سامانه را با وجود نقص شناخته‌شده یا بدون رعایت الزامات ایمنی عرضه کرده باشد. کاربر نیز هنگامی می‌تواند مسئول شناخته شود که برخلاف تکلیف مراقبتی عمل کند، هشدارهای ایمنی را نادیده بگیرد یا سامانه را خارج از حدود متعارف به کار ببرد. شخص حقوقی نیز در مواردی که خطر محصول نتیجه سیاست سازمانی یا تصمیم‌های مدیریتی باشد، باید در کانون تحلیل قرار گیرد.

مناسب‌ترین رویکرد برای تعیین مسئول، استفاده هم‌زمان از چند معیار است. نخست باید میزان کنترل مؤثر هر شخص بر سامانه در مرحله‌ای که خطر قابل پیشگیری بوده است سنجیده شود. سپس قابلیت پیش‌بینی خطر با توجه به سطح تخصص و اطلاعات در دسترس او بررسی گردد. در ادامه باید مشخص شود آیا شخص با رفتار خود خطر نامتعارفی ایجاد کرده یا تکلیف مراقبتی را نقض کرده است. همچنین امکان جلوگیری از نتیجه، سطح استقلال سامانه و فاصله میان رفتار شخص و مرگ باید در تحلیل وارد شود. چنین چارچوبی از دو افراط جلوگیری می‌کند: از یک سو مانع آن می‌شود که پیچیدگی فناوری موجب عدم مسئولیت همه افراد شود و از سوی دیگر اجازه نمی‌دهد صرف وقوع مرگ به تحمیل مسئولیت کیفری بر نزدیک‌ترین شخص به حادثه منجر گردد.

رابطه سببیت در این حوزه باید با حساسیت بیشتری تحلیل شود. وجود چند عامل در زنجیره تحقق مرگ به معنای قطع خودکار مسئولیت نیست. ممکن است چند سبب به‌طور هم‌زمان یا متوالی در نتیجه نقش داشته باشند. در چنین وضعیتی باید مشخص شود کدام عامل خطر حقوقاً مرتبط را ایجاد کرده و آیا نتیجه واقع‌شده در محدوده همان خطر قابل پیش‌بینی قرار داشته است. معیار صرف «شرط ضروری بودن» کافی نیست، زیرا می‌تواند دامنه مسئولیت را بیش از حد گسترش دهد. در مقابل، تمرکز صرف بر آخرین عامل نیز ممکن است نقش تصمیم‌های مهم پیشین را نادیده بگیرد. تحلیل هنجاری سببیت باید میان این دو رویکرد تعادل ایجاد کند.

یکی از نتایج اساسی پژوهش حاضر آن است که قواعد عمومی حقوق کیفری همچنان قابلیت استفاده دارند، اما برای پوشش کامل مسائل هوش مصنوعی کافی نیستند. مفاهیمی مانند تقصیر، رابطه سببیت، مباشرت، تسبیب و مسئولیت شخص حقوقی می‌توانند پایه تحلیل را تشکیل دهند، اما نیازمند تفسیر متناسب با ویژگی‌های سامانه‌های خودمختار هستند. فقدان مقررات مشخص درباره سطح نظارت انسانی، الزامات ثبت تصمیمات الگوریتمی، استانداردهای آزمون ایمنی و مسئولیت زنجیره تولید می‌تواند اثبات پرونده‌های کیفری آینده را دشوار کند. بنابراین ضرورت دارد که قانون‌گذار به‌جای انتظار برای شکل‌گیری پرونده‌های متعدد، به‌صورت پیشینی چارچوبی ویژه برای فناوری‌های پرخطر ایجاد کند.

در این چارچوب، پیشنهاد می‌شود سامانه‌های هوش مصنوعی براساس سطح خطر طبقه‌بندی شوند و برای سامانه‌هایی که می‌توانند به‌طور مستقیم یا غیرمستقیم جان انسان را به خطر اندازند، الزامات سخت‌گیرانه‌تری مقرر شود. تولیدکنندگان چنین سامانه‌هایی باید ملزم به ارزیابی مستمر خطر، مستندسازی فرایند طراحی و آزمون، ثبت تغییرات مهم الگوریتمی و اعلام سریع نقص‌های ایمنی شوند. بهره‌برداران حرفه‌ای نیز باید تکالیف مشخصی در زمینه نظارت، آموزش کاربران، نگهداری فنی و گزارش حوادث داشته باشند. وجود این تکالیف به دادگاه اجازه

خواهد داد پس از وقوع حادثه به جای اتکا به برداشت‌های کلی، بررسی کند کدام وظیفه قانونی نقض شده و این نقض چه نقشی در مرگ داشته است.

همچنین لازم است شفافیت فنی و قابلیت ثبت وقایع در سامانه‌های پرخطر به یک الزام حقوقی تبدیل شود. اگر پس از حادثه امکان بازسازی تصمیم سامانه وجود نداشته باشد، اثبات علت مرگ و شناسایی مسئول با مانع جدی روبه‌رو خواهد شد. طراحی سامانه باید به گونه‌ای باشد که ورودی‌های اساسی، تصمیم‌های مهم، هشدارهای صادرشده و مداخلات انسانی قابل بازبینی باشند. این امر تنها یک الزام فنی نیست، بلکه بخشی از تضمین عدالت کیفری است؛ زیرا بدون دسترسی به اطلاعات قابل اعتماد، احتمال محکومیت نادرست یا عدم شناسایی مسئول واقعی افزایش می‌یابد.

در نهایت، مسئولیت کیفری قتل ناشی از هوش مصنوعی باید بر اصل پاسخ‌گویی انسانی و سازمانی استوار باقی بماند، اما این پاسخ‌گویی باید متناسب با ساختار واقعی فناوری بازتعریف شود. نباید سامانه هوشمند به پوششی برای فرار از مسئولیت تبدیل شود و در عین حال نباید اشخاص صرفاً به دلیل ارتباط فنی یا مالکیتی با سامانه مسئول شناخته شوند. معیار نهایی باید میزان نقش در ایجاد خطر، قابلیت پیش‌بینی، قدرت کنترل، امکان پیشگیری و درجه نقض تکلیف باشد. با چنین رویکردی می‌توان میان حمایت از حق حیات، تضمین اصول بنیادین حقوق کیفری و ضرورت توسعه مسئولانه فناوری تعادل برقرار کرد و از شکل‌گیری خلأ مسئولیت در برابر مرگ‌های ناشی از سامانه‌های هوشمند جلوگیری نمود.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافع وجود ندارد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

حامی مالی

این پژوهش حامی مالی نداشته است.

EXTENDED SUMMARY

The rapid expansion of artificial intelligence into fields directly connected with human life, bodily integrity, mobility, healthcare, industrial activity, and automated decision-making has created a new generation of criminal-law problems that cannot always be resolved through conventional models of attribution. Artificial intelligence systems are no longer confined to passive computational functions; increasingly, they are capable of processing large volumes of data, recognizing patterns, adapting to changing environments, and producing outputs that may not be fully predetermined by a human operator at the moment of action. This transformation becomes especially significant when the operation of an autonomous or semi-autonomous system contributes to the death of a person. In such cases, criminal law must determine whether the fatal outcome can be attributed to a programmer, designer, manufacturer, corporate entity, owner, operator, user, or another participant in the technological chain, and whether the conduct in question satisfies the material and mental requirements of homicide. The challenge is intensified by the fact that the system may function with a degree of operational independence and may generate decisions that are difficult to predict or explain in advance (Zandi & Rafiei Alavi, 2024). Recent scholarship has therefore emphasized that artificial intelligence creates a potential gap between the physical production of harm and the legal attribution of that harm to a responsible human actor (Ghavami Pour Sereshkeh & Mahmoudi, 2025).

At the same time, the existence of sophisticated technology cannot itself justify abandoning the fundamental principles of criminal law, including personal responsibility, culpability, causation, and the requirement that punishment be based on legally attributable conduct. The objective of the present study is to examine the criminal liability arising from homicide caused by artificial intelligence, to identify the persons who may bear responsibility within the design-to-use chain, and to evaluate whether traditional concepts such as direct perpetration, causation, negligence, foreseeability, and mens rea remain sufficient in the context of autonomous systems. The study adopts a descriptive-analytical legal method and relies on doctrinal analysis of criminal liability together with contemporary scholarship concerning the legal status and responsibility of artificial intelligence.

From a conceptual perspective, homicide caused by artificial intelligence must be distinguished from cases in which a human offender merely uses an intelligent system as an instrument for intentionally killing another person. Where a person deliberately programs, directs, or deploys an artificial intelligence system for the purpose of causing death, the technology does not substantially alter the ordinary analysis of intentional homicide, because the system functions as a means through which the perpetrator's criminal intent is implemented. The more difficult cases arise when no human actor specifically intends death and the fatal result emerges from a combination of algorithmic design, machine learning, defective data, inadequate testing, insufficient supervision, system malfunction, or unexpected autonomous behavior. Homicide, as a crime against human life, traditionally requires identification of a legally relevant act or omission, a causal relationship between that conduct and death, and the appropriate mental element corresponding to the specific category of homicide (Haji, 2018). Artificial intelligence complicates each of these requirements because the physical act producing death may be performed by a machine rather than a human being, while the human contribution may be temporally and causally remote from the final event. This difficulty is particularly evident in systems that continue learning after deployment, because the ultimate output may differ from the behavior originally anticipated by designers and developers (Javanmard & Heydarifard, 2025). Nevertheless, technical autonomy should not automatically be equated with legal autonomy. Existing research on artificial intelligence in criminal law indicates that most legal systems do not recognize an autonomous artificial intelligence system as an independent bearer of criminal responsibility, largely because criminal liability presupposes normative blameworthiness, legal capacity, and a meaningful mental relationship to the prohibited conduct (Younes, 2024). Consequently, the more appropriate approach is to treat artificial intelligence as an active technological factor within the causal chain while continuing to examine whether human or corporate actors created, controlled, tolerated, or failed to prevent the relevant risk (Kaveh & Barani, 2024).

A central issue in this field is causation. In traditional homicide cases, the causal sequence between human conduct and death is often relatively direct, whereas artificial intelligence systems introduce multiple intermediate stages, including data collection, model training, software architecture, hardware integration, testing, deployment, maintenance, updates, and real-time automated decision-making. Each stage may contribute to the fatal outcome, and in some cases no single actor can be identified as the exclusive cause of death. A defective autonomous vehicle, for example, may fail to recognize a pedestrian because of inadequate training data; the manufacturer may have released the system despite known limitations; the owner may have ignored software updates; and the user may have failed to exercise the degree of supervision required under the operating instructions. Determining criminal liability in such a situation requires more than a simple "but-for"

test, because an excessively broad causal inquiry could extend responsibility to actors whose connection to the death is too remote. The relevant question should therefore focus on whether a person created or materially increased a legally significant risk and whether the death occurred within the scope of that risk. The jurisprudential and doctrinal discussion of indirect responsibility in homicide also supports the proposition that an actor need not physically perform the final harmful act in order to bear responsibility where the conduct constitutes an effective cause of death (Nozari Ferdowsieh et al., 2022). Artificial intelligence further complicates causation through the so-called black-box problem, under which even experts may be unable to explain precisely how a machine-learning system arrived at a particular output. This problem may impede the identification of the specific defect responsible for death and thereby create evidentiary difficulties in criminal proceedings (Ghavami Pour Sereshkeh & Mahmoudi, 2025). These concerns become especially acute in medical artificial intelligence, where an incorrect diagnosis or treatment recommendation may interact with professional judgment and institutional procedures before producing a fatal outcome (Goel, 2023). Accordingly, causation in artificial-intelligence-related homicide should be assessed through a multidimensional framework that considers technical contribution, normative relevance, foreseeability, control, and the possibility of prevention.

The mental element presents an equally significant challenge. Criminal law generally distinguishes between intentional conduct, conscious risk-taking, negligence, and situations in which a harmful outcome was genuinely unforeseeable. Artificial intelligence complicates this distinction because the person involved in designing or using the system may not know the precise manner in which the system will behave in every situation. A programmer may create the architecture of a learning system without anticipating a specific fatal output, while a manufacturer may become aware of recurring safety warnings and nevertheless continue commercial deployment. Likewise, an operator may fail to intervene despite visible indications that the system is malfunctioning. These actors occupy different epistemic positions and should not be judged by an identical standard of foresight. Scholarship on autonomous artificial intelligence has stressed that the mental element must be assessed with reference to the knowledge, professional expertise, available information, and actual capacity of each human participant to foresee and prevent the harmful outcome (Javanmard & Heydarifard, 2025). Foreseeability is therefore not a uniform concept; what is unforeseeable to an ordinary consumer may be readily foreseeable to a specialist developer or manufacturer with access to testing data and internal risk reports. Comparative studies also suggest that the degree of control exercised by a person over the system is a crucial factor in determining criminal attribution (Meghdadi & Pourcheriki, 2024). Where a company is aware of a defect that may cause serious bodily harm but fails to correct it for economic reasons, the degree of culpability is far greater than in a case involving a genuinely unknown technical anomaly. Similarly, where a user knowingly disables a safety feature or operates a system outside the conditions for which it was designed, the user's conduct may constitute a legally relevant contribution to the death. Thus, the classification of homicide in cases involving artificial intelligence should not be based merely on the presence of an autonomous system, but on a careful examination of the mental state and duty of care of the persons who participated in the creation, deployment, or operation of the technology (Al-Halahlah, 2025).

Identifying the responsible person requires examination of the entire lifecycle of the artificial intelligence system rather than focusing only on the individual who was physically closest to the fatal event. The programmer or algorithm designer may be liable where a serious and foreseeable coding or design defect results from a culpable departure from recognized professional standards, but mere participation in the development process is not sufficient to establish criminal responsibility

(Amirian Farsani & Hosseini, 2023). The manufacturer or corporate developer may bear greater responsibility where the organization controls testing, release decisions, safety procedures, warnings, updates, and post-market monitoring, particularly if it knowingly places a dangerous system into operation (Al-Jamili, 2025). Corporate criminal responsibility also becomes relevant where the risk results from organizational policies rather than the isolated conduct of a single employee; decisions to reduce testing, conceal defects, disregard internal warnings, or accelerate deployment may represent institutional failures that cannot be adequately explained through individual liability alone (Al-Sadat Makki et al., 2024). The owner or operator may also bear responsibility when death results from ignoring known warnings, failing to maintain the system, or using it outside its approved operational limits, while liability should generally be excluded where the defect was concealed and could not reasonably have been detected. In autonomous vehicles, medical systems, industrial robots, and other high-risk applications, liability may therefore be distributed among several persons where multiple acts and omissions materially contribute to the fatal result (Al-Hasanat, 2024). A more convincing legal framework should rely on several interconnected criteria: effective control over the system, foreseeability of the relevant danger, access to risk information, ability to prevent the outcome, breach of a legally recognized duty of care, and the extent to which the person's conduct created or increased an abnormal risk. Such an approach reduces the danger of a responsibility gap in which every participant attempts to attribute the death to the machine or to another actor in the chain (Rostami Zabol, 2025). At the same time, it avoids imposing strict criminal liability merely because an individual or organization had some relationship with the artificial intelligence system.

The findings of this study indicate that existing principles of criminal law remain indispensable for addressing homicide caused by artificial intelligence, but they require adaptation to the technological structure within which fatal outcomes occur. Artificial intelligence should not presently be treated as an independent criminal offender merely because it is technically capable of making autonomous decisions. Rather, the legal analysis should identify the human and corporate actors who created, managed, controlled, or failed to mitigate the relevant risk and should evaluate their responsibility through the combined criteria of causation, foreseeability, control, and culpable breach of duty. No single actor, such as the programmer, manufacturer, owner, or user, can be presumed responsible in all cases, and criminal liability must instead be individualized according to the actual role of each participant. The study further concludes that conventional rules concerning causation, complicity, indirect contribution, negligence, and corporate responsibility can provide the basic legal foundation for these cases, but they should be supplemented by specific regulatory duties for high-risk artificial intelligence systems. Such duties should include mandatory risk assessment, technical documentation, traceability of significant algorithmic decisions, preservation of event logs, safety testing, post-deployment monitoring, clear human-supervision requirements, and prompt correction of known defects. These measures would serve both preventive and evidentiary functions by reducing the likelihood of fatal incidents and enabling courts to reconstruct the causal chain when death occurs. Ultimately, the legal response to artificial-intelligence-related homicide should preserve the core principle that serious criminal punishment must be grounded in proven personal or organizational culpability while preventing technological complexity from becoming a mechanism through which legally responsible actors evade accountability.

References

Al-Falasi, A. A. M. (2023). *Criminal Liability Arising from Artificial Intelligence Errors*.

- Al-Halahlah, A. K. F. (2025). *Artificial intelligence and crime: Legal frameworks of criminal liability in the digital era*.
- Al-Hasanat, H. K. H. (2024). Criminal Liability Arising from Artificial Intelligence Errors. *Al-Zaytoonah University of Jordan Journal for Legal Studies*, 43(4), 1849-1952.
- Al-Jamili, O. H. J. (2025). *Criminal Liability Arising from Artificial Intelligence Errors*. Dar Misr for Printing and Publishing.
- Al-M'ebid, O. I. (2024). Criminal Liability Arising from the Use of Artificial Intelligence Technologies in the Saudi Arabian Legal System: An Analytical-Comparative Study. *Arab Journal of Security Studies*, 40(2), 143-157.
- Al-Sadat Makki, A., Al-Sadat Makki, Z., & Kashkoulia, E. (2024). An Examination of Liability Arising from Artificial Intelligence Acts in the Iranian Legal System. *Scientific Journal of Jurisprudence, Law and Criminal Sciences*, 8(32), 71-79.
- Amirian Farsani, A., & Hosseini, S. M. (2023). Challenges and Barriers to Criminal Liability in Robots with Artificial Intelligence Capabilities. *Legal Civilization*(18).
- Ghavami Pour Sereshkeh, M., & Mahmoudi, A. (2025). An Introduction to Criminal Liability Frameworks for Artificial Intelligence Systems. *Biannual Scientific Journal of New Technologies Law*, 6(11), 209-232.
- Ghavamipour Sarshkeh, M., & Mahmoudi, M. (2024). An Introduction to Legal Frameworks for the Criminal Liability of Artificial Intelligence Systems. *Quarterly Journal of Legal Studies*.
- Goel, A. T. V. (2023). Determining the 'Responsibility' Paradox- The Criminal Liability of Artificial Intelligence in the Healthcare Sector. *Russian Law Journal*, 11(2s). <https://doi.org/10.52783/rlj.v11i2s.581>
- Haji, D. (2018). *Crimes Against Persons: Homicide* (2nd ed.). Mizan.
- Javanmard, M., & Heydarifard, S. M. (2025). Criminal Liability of Autonomous Artificial Intelligence with Emphasis on the Mental Element in Light of Comparative Studies. *Quarterly Journal of Legal Research*.
- Kaveh, M. H., & Barani, M. (2024). Criminal Liability of Artificial Intelligence in Iranian Criminal Law with a View to European Union Legislation. *Comparative Criminal Jurisprudence*(3).
- Makki, A. a.-S., Makki, Z. a.-S., & Kashkoulia, E. (2024). Examining Liability Arising from Artificial Intelligence Actions in the Iranian Legal System. *Jurisprudence, Law and Criminal Sciences*(32).
- Meghdadi, M., & Pourcheriki, F. (2024). A Comparative Study of the Criminal Liability of Artificial Intelligence in the Legal Systems of Iran and Saudi Arabia. *Journal of Comparative Legal Studies*.
- Nozari Ferdowsieh, M., Rostami Ghafasabadi, A., & Jahani, M. H. (2022). A jurisprudential re-examination of the criminal liability of an unseen supervisor in a deliberate homicide crime. *Research on Islamic Jurisprudence and Law*, 18.
- Rostami Zabol, B. (2025). Criminal liability arising from the actions of autonomous Artificial Intelligence in the Iranian criminal legal system: Challenges and legislative necessities. 14th International and National Conference on Management, Accounting, and Law Studies, Tehran.
- Younes, Y. Z. K. (2024). Special features of criminal liability concerning artificial intelligence. *Iraqi Journal*, 68. <https://iasj.rdd.edu.iq/journals/uploads/2025/05/03/c31dd10d02a6678fed56bccb1e823948.pdf>
- Zandi, M., & Rafiei Alavi, S. E. (2024). Feasibility of criminal liability in Artificial Intelligence based on its philosophical foundations. *Philosophy of Law*, 3(1).