





Criminal Liability of Autonomous Artificial Intelligence with an Emphasis on the Mental Element in the Light of Comparative Studies

Behrooz Javanmard , Department of Criminal Law, Noor Tooba University, Tehran, Iran.
(Corresponding Author) Email: B.javanmard@iau.ir

Noushin Heydarifard , PhD student in Criminal Law and Criminology, Islamic Azad University, Science and Research Branch, Tehran, Iran

Abstract

Introduction: Autonomous artificial intelligence (AI) systems challenge core assumptions of criminal law by performing complex actions with minimal human oversight. This paper asks whether such systems can bear the mental element required for criminal liability.

Methods: Using a comparative method, it first clarifies key concepts such as weak/strong AI, autonomy, and legal personhood, and sets out three theories of AI liability: vicarious liability through a human actor, liability based on reasonably foreseeable consequences, and direct liability where the system itself is treated as an offender. In the first model (the perpetration-via-another liability model), the AI, as an innocent agent such as a child, could be used as a vehicle to perpetrate criminal actions. The second model (the natural-probable-consequence liability model) imposes accountability upon individuals for offenses that arise as a natural and foreseeable consequence of their action, irrespective of their actual awareness of the offense. Last but not least, in the third model, theoretically, if AI is capable of self-determination, it can have will and knowledge of its specific action. In such cases, a third scenario approach is necessitated, allowing the entity itself to be directly liable for its offenses. It then analyzes how elements of crime-physical conduct and mens rea-are defined in selected legal systems (Italy, Slovakia, Germany, and the United States) and reviews recent legislation and judicial practice. The research draws on statutory analysis, doctrinal scholarship, and reported cases.

Results: The findings show that no jurisdiction has yet recognized an AI system as an autonomous perpetrator; criminal liability remains anchored in human intent or negligence. While the European Union's AI Act of 2024 emphasizes human oversight, Italy's 2025 AI Law imposes stricter transparency and harsher penalties for AI-assisted offenses. The article concludes that, although future technological developments may require doctrinal innovation, current law can address AI-related harms through existing concepts of vicarious or negligent liability. Possible sanctions against AI systems are limited to symbolic or functional measures such as deletion of software or suspension of services. The paper advocates continued interdisciplinary collaboration to refine criminal norms as AI evolves.

Keywords: Principle of Autonomy, Mental Element, Criminal Behavior, Generative Artificial Intelligence, Independent Criminal Liability.



پروشکاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مسئولیت کیفری هوش مصنوعی خودمختار با تأکید بر عنصر معنوی در پرتو مطالعات تطبیقی

بهروز جوانمرد^۱، نوشین حیدری فرد^۲

چکیده

این مقاله تلاش دارد به این پرسش پاسخ دهد که آیا سامانه‌های هوش مصنوعی خودمختار می‌توانند دارای عنصر معنوی لازم برای پذیرش مسئولیت کیفری مستقل باشند یا خیر؟ با روش توصیفی تحلیلی با رویکردی تطبیقی، نخست مفاهیم بنیادین هوش مصنوعی ضعیف و قوی و مفهوم خودمختاری و شخصیت حقوقی تشریح می‌شود. سپس سه مدل مسئولیت کیفری شامل «مسئولیت از طریق دیگری»، «نتیجه طبیعی محتمل» و «مسئولیت مستقیم» معرفی و مبنای هر یک بررسی می‌گردد. در ادامه عناصر جرم اعم از رفتار مادی و عنصر معنوی در نظام‌های کیفری منتخب (ایتالیا، اسلواکی، آلمان و ایالات متحده) تحلیل شده و وضعیت حقوقی آن‌ها در مواجهه با رفتارهای مجرمانه مبتنی بر هوش مصنوعی مولد و خودمختار ارزیابی می‌شود. یافته‌های پژوهش نشان می‌دهد که تاکنون هیچ نظام قضایی هوش مصنوعی را به‌عنوان مرتکب مستقل به رسمیت نشناخته و مسئولیت کیفری عمدتاً بر دوش انسان‌ها (طراحان، کاربران یا مالکان) قرار دارد. قانون اتحادیه اروپا موسوم به قانون هوش مصنوعی ۲۰۲۴، بر لزوم نظارت انسانی و شفافیت تأکید دارد و قانون ملی هوش مصنوعی ایتالیا مصوب ۲۰۲۵ با تشدید مجازات و الزامات شفافیت، استفاده از سامانه‌های پرخطر را سامان‌دهی می‌کند. بر اساس نتایج، در شرایط فعلی فناوری، اعمال مجازات مستقیم بر سامانه‌ها معنای عملی ندارد و پیشنهاد می‌شود از تدابیر نمادین مانند حذف نرم‌افزار یا تعلیق خدمات استفاده شود. نوشتار همچنین بر ضرورت همکاری حقوق‌دانان، مهندسان و سیاست‌گذاران برای اصلاح مفاهیم حقوق کیفری همگام با پیشرفت فناوری تأکید می‌کند.

واژگان کلیدی

اصل خودمختاری، عنصر معنوی، رفتار مجرمانه، هوش مصنوعی مولد، مسئولیت کیفری مستقل.

مقدمه

در دهه گذشته، هوش مصنوعی به یکی از برجسته‌ترین نیروهای محرک در تحول فناوری و اقتصاد جهانی تبدیل شده است. سامانه‌های هوشمندی که می‌توانند داده‌ها را تحلیل کرده و تصمیم‌هایی مستقل اتخاذ کنند، نه تنها کارایی صنایع را افزایش داده‌اند بلکه امکان وقوع رفتارهای غیرمنتظره را نیز به وجود آورده‌اند. پیشرفت‌های شتابان در حوزه هوش مصنوعی مولد، به‌ویژه الگوریتم‌های تولید خودکار متن، تصویر و صوت، توجه حقوق‌دانان را به تأثیرات این فناوری بر مفهوم «جرم‌انگاری» جلب کرده است. این سؤال که آیا چنین سامانه‌هایی می‌توانند دارای قصد و آگاهی اخلاقی باشند و در نتیجه به‌عنوان مرتکب جرم شناخته شوند، یکی از چالش‌های بنیادین حقوق کیفری معاصر است.

پیشینه پژوهش‌ها نشان می‌دهد که بسیاری از آثار موجود، مسئولیت حاصل از رفتارهای هوش مصنوعی را در قالب «مسئولیت از طریق دیگری» تبیین کرده‌اند؛ یعنی سامانه هوش مصنوعی صرفاً وسیله‌ای است که رفتار انسان را اجرا می‌کند و عنصر معنوی جرم صرفاً در انسان نهفته است. با این حال، در سال‌های اخیر برخی نویسندگان با اتکا به پیشرفت‌های فناورانه، امکان شکل‌گیری نوعی خودمختاری را در سامانه‌ها مطرح کرده‌اند و خواستار بازنگری در اصول کلاسیک مسئولیت کیفری شده‌اند. از جمله خلیلی و سید هاشمی در مقاله‌ای در سال ۱۴۰۳ با عنوان امکان‌سنجی مسئولیت کیفری هوش مصنوعی با تأکید بر شخصیت مستقل شخص حقوقی در فصلنامه عدلیه تلاش نموده‌اند نظریه‌های مبتنی بر شخصیت مستقل شخص حقوقی را به خود نرم‌افزار هوش مصنوعی ارتباط بدهند. در سطح مطالعات بین‌المللی سید در سال ۲۰۲۴ در مجله Cogent Social Sciences با مقاله‌ای با عنوان «هوش مصنوعی و مسئولیت کیفری در هند: بررسی پیامدها و چالش‌های حقوقی» توضیح می‌دهد که هوش مصنوعی با ایجاد جرائم نوین، مسئولیت کیفری را مبهم کرده است. نویسنده در این مقاله بر ضرورت چارچوب‌های قانونی، اخلاقی و حفاظت داده تأکید دارد؛ در نتیجه‌گیری نیز بر اصلاح قوانین، شفافیت الگوریتم‌ها و مسئولیت توسعه‌دهندگان و کاربران تأکید می‌کند.^۱

از سوی دیگر، مقررات بین‌المللی نوظهور نیز بر نقش نظارت انسانی در استفاده از

1. Hifajatali Sayyed (2024) Artificial intelligence and criminal liability in India: exploring legal implications and challenges, Cogent Social Sciences, 10:1, 2343195, DOI: 10.1080/23311886.2024.2343195

سامانه‌های پرخطر تأکید می‌کنند؛ برای نمونه، «قانون هوش مصنوعی اتحادیه اروپا»^۱ مصوب ۲۰۲۴ بر ایجاد چارچوبی هماهنگ برای ایمنی و مسئولیت تأکید دارد. علاوه بر آن، ایتالیا در سپتامبر ۲۰۲۵ نخستین قانون ملی هوش مصنوعی^۲ در اروپا را تصویب کرد که علاوه بر اصول کلی، مجازاتی را برای سوءاستفاده‌های کیفری پیش‌بینی می‌کند. نوآوری مقاله حاضر نسبت به نوشتارهای پیشین داخلی و خارجی این است که نویسندگان از زاویه امکان‌سنجی وجود عنصر معنوی برای هوش مصنوعی مولد و خودمختار به موضوع نگریسته‌اند.

هدف این مقاله پاسخ‌گویی به این پرسش است که آیا سامانه‌های هوش مصنوعی خودمختار، می‌توانند به‌طور مستقل دارای عنصر معنوی، یعنی سوءنیت مجرمانه، باشند یا خیر. برای نیل به این هدف، نخست مفاهیم پایه‌ای هوش مصنوعی و سطوح خودمختاری آن تعریف می‌شود و سپس سه نظریه مسئولیت کیفری سامانه‌های هوشمند شامل مسئولیت از طریق دیگری، مسئولیت مبتنی بر نتیجه طبیعی و مسئولیت مستقیم تبیین می‌گردد. به‌منظور بررسی دقیق‌تر عنصر مادی و معنوی، نظام‌های کیفری چهار کشور (ایتالیا، اسلواکی، آلمان و ایالات متحده) مطالعه و عناصر جرم در این نظام‌ها بررسی می‌شود. روش تحقیق این نوشتار بر مبنای تحلیل اسناد قانونی، مطالعات نظری و بررسی رویه‌های قضایی گزارش شده است و تلاش می‌شود ضمن نقد آثار قبلی، تصویر روشن‌تری از جایگاه مسئولیت کیفری هوش مصنوعی ارائه شود. نویسندگان علاوه بر این، به‌نقد دیدگاه‌های مطرح‌شده در ادبیات حقوقی پرداخته و تلاش می‌کنند با تلفیق یافته‌های تطبیقی، ساختاری پیشنهادی برای مواجهه قانون‌گذاران با پدیده هوش مصنوعی ترسیم نمایند.

۱. چارچوب مفهومی

۱-۱. پیشینه، تعریف و گونه‌شناسی هوش مصنوعی

از منظر پیشینه، موضوع امکان‌سنجی مسئولیت کیفری برای هوش مصنوعی نخست در آمریکا توسط پژوهشگرانی چون چوپرا و وایت^۳ (۲۰۱۱) استادان دانشگاه میشیگان

1. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)

۲. قانون ملی ایتالیا با عنوان «Provisions and delegations to the Government on arti- (ficial intelligence)» در تاریخ ۲۳ سپتامبر ۲۰۲۵ تصویب شده است.

3. Chopra & White

مطرح شد. سپس در اروپا در دانشگاه ماستریخت هلند با رساله دکتری آلیس جیانینی^۱ (۲۰۲۳) و در ادامه در آفریقا در دانشگاه تنتا در استان غربیه مصر با مطالعات یاسر اللامی^۲ و عمرو الواکد^۳ (۲۰۲۳) این بحث را توسعه دادند. از جنبه تعریف شناسی، هوش مصنوعی اصطلاحی فراگیر است که به مجموعه‌ای از روش‌ها و فناوری‌ها برای شبیه‌سازی توانایی‌های شناختی انسان در ماشین‌ها اطلاق می‌شود. دودسته عمده هوش مصنوعی قابل تفکیک است: «هوش مصنوعی ضعیف» که سامانه‌هایی با الگوریتم‌های از پیش تعریف‌شده و فاقد شعور را شامل می‌شود و «هوش مصنوعی قوی» که هدف آن ایجاد سامانه‌ای است که بتواند آگاهانه بیندیشد و به گونه‌ای مشابه انسان یاد بگیرد و تصمیم بگیرد؛ چنین سامانه‌ای هنوز محقق نشده است. سیری^۴ و الکسا^۵ نمونه‌هایی از هوش مصنوعی ضعیف هستند که بر اساس داده‌های گذشته پاسخ‌های محدودی ارائه می‌دهند و هیچ اراده مستقلی ندارند. برخلاف آن، تصور هوش مصنوعی قوی مستلزم ایجاد نظامی است که واجد خودآگاهی و نیت باشد، امری که فعلاً دور از دسترس فناوری است.

۲-۱. خودمختاری و شخصیت حقوقی

مفهوم خودمختاری در سامانه‌های هوشمند به توانایی آن‌ها برای تعیین اهداف و اتخاذ تصمیم بدون دخالت مستقیم انسان اشاره دارد (Danaher, 2021). خودروهای خودران و ربات‌های پیشرفته نمونه‌هایی از سامانه‌هایی هستند که در شرایط مشخصی رفتار مستقل از انسان نشان می‌دهند. باین‌حال، آنچه به نظر «تصمیم مستقل» می‌رسد، در واقع اجرای مجموعه‌ای از دستورالعمل‌ها و الگوریتم‌های یادگیری ماشین است و سامانه بدون آگاهی یا درک اخلاقی واقعی است (Bryson, Diamantis, & Grant, 2017).

در حوزه شخصیت حقوقی هوش مصنوعی، برخی پژوهشگران پیشنهاد داده‌اند که با اعطای نوعی شخصیت حقوقی محدود بتوان سازوکار قضایی مستقلی برای سامانه‌های هوشمند ایجاد کرد (European Parliament, 2017). باین‌حال، بیش از ۱۵۰ پژوهشگر در نامه‌ای سرگشاده با این ایده مخالفت کرده‌اند و آن را موجب بی‌مسئولیتی و گریز از پاسخ‌گویی دانسته‌اند (Nemitz, 2018). در وضعیت کنونی،

1. Alice Gianini
2. Yasser Allami
3. Amr Elwakad

۴. Siri، دستیار صوتی شرکت Apple است که به‌عنوان نمونه‌ای از هوش مصنوعی ضعیف به کار می‌رود.

۵. Alexa، دستیار صوتی شرکت Amazon و نمونه‌ای دیگر از هوش مصنوعی ضعیف است.

اعطای شخصیت حقوقی به ماشین‌ها هنوز موضوعی نظری است و به همین دلیل، مسئله «خودمختاری» در تعیین مسئولیت و پاسخ‌گویی حقوقی نقشی اساسی دارد (Bryson et al. 2017; Danaher, 2021).

۲. الگوهای مسئولیت کیفری هوش مصنوعی

۱-۲. الگوی مسئولیت از طریق دیگری

در این الگو، سامانه هوش مصنوعی صرفاً وسیله‌ای بی‌اراده در نظر گرفته می‌شود و تمام مسئولیت کیفری بر عهده شخص انسانی است که از آن استفاده کرده یا آن را برنامه‌ریزی نموده است (Hallevy, 2010). برای مثال، اگر برنامه‌نویسی رباتی را به گونه‌ای طراحی کند که موجب آتش‌سوزی شود یا کاربری به ربات دستور دهد که حیوانی را از بین ببرد، عنصر معنوی جرم در عامل انسانی وجود دارد و هوش مصنوعی صرفاً ابزار مادی عمل محسوب می‌شود. نظام کیفری اسلواکی نیز چنین نگرشی را پذیرفته و تنها افراد حقیقی و حقوقی را مرتکب جرم می‌داند و برای نرم‌افزار و ماشین‌ها مسئولیت مستقلاً در نظر نمی‌گیرد (European Parliament, 2017).

۲-۲. الگوی نتیجه‌طبیعی محتمل

این رویکرد بر این نکته تأکید دارد که حتی در نبود قصد مستقیم، اگر بتوان به‌طور منطقی پیش‌بینی کرد که استفاده از سامانه هوش مصنوعی باعث بروز زیان می‌شود، انسان مسئول خواهد بود (Sartor, 2020). برای نمونه، اگر سازندگان سامانه هدایت خودکار هواپیما نقصی را در سیستم ایمنی نادیده بگیرند و این امر موجب حادثه‌ای مرگبار شود، آن‌ها به علت سهل‌انگاری مسئول شناخته می‌شوند. در حقوق ایتالیا نیز بی‌احتیاطی و نادیده گرفتن عواقب قابل پیش‌بینی می‌تواند به مسئولیت شبه عمد منجر شود. در این الگو، کلیدواژه اصلی «تعهد مراقبتی» است؛ هرچند سامانه هوش مصنوعی عمل را اجرا کرده، ولی عدم پیش‌بینی و پیشگیری از خطر توسط انسان موجب تحقق مسئولیت می‌شود (Sartor, 2020).

۲-۳. الگوی مسئولیت مستقیم

مدل سوم فرض می‌کند که اگر روزی سامانه‌های هوشمند به سطحی از خودآگاهی و اختیار برسند که بتوان نیت و علم را در آن‌ها تشخیص داد، باید به‌عنوان مرتکب

۱. «Autopilot» به سیستم هدایت خودکار هواپیما گفته می‌شود.

مستقل مجازات شوند (Hallevy, 2015). طرفداران این دیدگاه معتقدند که نمی‌توان میان عامل انسانی و هوش مصنوعی تفاوتی قائل شد، اگر هر دو دارای قصد و آگاهی باشند. با این حال، در وضعیت فعلی چنین سطحی از خودآگاهی در هیچ سامانه‌ای وجود ندارد و این نظریه عمدتاً جنبه‌ای فرضی و آینده‌پژوهانه دارد. در بسیاری از نظام‌های حقوقی، حتی مسئولیت کیفری شرکت‌ها نیز محدود است؛ بنابراین، پذیرفتن مسئولیت مستقل برای ماشین‌هایی که فاقد وجدان و انگیزه اخلاقی هستند، هنوز امکان‌پذیر نشده است (Hallevy, 2015).

۳. عناصر جرم و رویکرد نظام‌های حقوقی منتخب

۳-۱. نظام کیفری ایتالیا و اسلواکی

در حقوق کیفری ایتالیا، تحقق جرم منوط به وجود رفتار مادی^۱ و عنصر روانی عمد یا بی‌احتیاطی^۲ است. قانون مجازات ایتالیا میان عمد مستقیم، عمد احتمالی و بی‌احتیاطی تمایز قائل می‌شود و هرگونه انتساب مسئولیت به ماشین‌ها را رد می‌کند (Sartor, 2020). نظام کیفری اسلواکی نیز بر دو سطح عمد و غیر عمد استوار است و مرتکب را صرفاً انسان یا نهادهای حقوقی می‌داند (European Parliamentary Research Service, 2020). در هر دو کشور، در صورت بروز جرم مرتبط با هوش مصنوعی، مجازات متوجه انسان طراح، کاربر یا توسعه‌دهنده نرم‌افزار است که نتایج قابل پیش‌بینی را نادیده گرفته است (Sartor, 2020; European Parliamentary Research Service, 2020).

۳-۲. نظام کیفری آلمان

حقوق کیفری آلمان برای وقوع جرم تحقق «تاتبستند»^۳ (عناصر مادی و روانی) و «شولد»^۴ (قابل انتساب بودن عمل به مرتکب) را الزامی می‌داند. عنصر روانی به عمد در رفتار یا سهل‌انگاری^۵ تقسیم می‌شود. در وضعیت سهل‌انگاری، مرتکب به علت نادیده گرفتن معیار مراقبت فرد عاقل مسئول است. در این کشور نیز هیچ پیش‌بینی قانونی برای مجازات هوش مصنوعی وجود ندارد و تنها افراد انسانی مسئول شناخته

1. Actus Reus

2. Mens Rea

۳. «Tatbestand»، واژه آلمانی به معنای عناصر مادی و روانی تشکیل جرم است.

۴. «Schuld»، اصطلاح حقوق کیفری آلمان به معنای قابلیت انتساب عمل مجرمانه به مرتکب است.

۵. «Vorsatz»، واژه آلمانی به معنای قصد یا عمد در ارتکاب جرم است.

۶. «Fahrlässigkeit»، واژه آلمانی به معنای سهل‌انگاری یا بی‌احتیاطی است.

مسئولیت کیفری هوش مصنوعی خودمختار [...] |

می‌شوند. باین‌حال، قوانین ایمنی محصول و نظارت فنی به‌گونه‌ای است که اگر تولیدکننده‌ای نقص سیستم را نادیده بگیرد و زیانی رخ دهد، می‌تواند به جرم قتل خطایی یا جرح خطایی محکوم شود.

۳-۳. نظام کیفری ایالات متحده

در ایالات متحده، تحقق جرم نیازمند وجود actus reus و mens rea است. مدل قانون مجازات آمریکا چهار سطح از سوءنیت را تعریف می‌کند: عمد، علم، بی‌ملاحظه‌بودن و سهل‌انگاری (American Law Institute, 1985). در پرونده‌های مربوط به خودروهای خودران، معمولاً به‌جای تعقیب سامانه هوش مصنوعی، کاربر یا سازنده به جرم قتل شبه‌عمد یا جرح ناشی از سهل‌انگاری محکوم می‌شود (Calo, 2016). از این‌رو، مسئولیت کیفری در حقوق آمریکا نیز بر عنصر روانی انسان مبتنی است و تاکنون هیچ پرونده‌ای هوش مصنوعی را به‌عنوان مرتکب مستقل نشناخته است (American Law Institute, 1985; Calo, 2016).

جمع‌بندی تطبیقی مبانی تحقق جرم، عنصر روانی و وضعیت مسئولیت کیفری هوش مصنوعی در نظام‌های حقوقی منتخب در جدول شماره ۱ ارائه شده است. همان‌گونه که جدول ذیل نشان می‌دهد، مسئولیت کیفری همواره متوجه انسان باقی می‌ماند.

جدول ۱. مسئولیت کیفری هوش مصنوعی در نظام‌های منتخب

نظام کیفری	مبانی تحقق جرم و عنصر روانی	وضعیت مسئولیت هوش مصنوعی	مسئول نهایی و نوع اتهام
ایتالیا و اسلواکی	<ul style="list-style-type: none"> رفتار مادی (actus reus) + عنصر روانی (mens rea) تفکیک میان عمد (مستقیم/احتمالی) و بی‌احتیاطی 	<ul style="list-style-type: none"> رد کامل مسئولیت ماشین‌ها مرتکب صرفاً انسان یا نهاد حقوقی (در اسلواکی) است. 	طراح، کاربر یا توسعه‌دهنده به دلیل نادیده‌گرفتن نتایج قابل پیش‌بینی مجازات می‌شوند.
آلمان	<ul style="list-style-type: none"> تابستند (عناصر مادی/روانی) + شولد (قابلیت انتساب) تقسیم به عمد (Vorsatz) و سهل‌انگاری (Fahrlässigkeit) 	<ul style="list-style-type: none"> عدم پیش‌بینی قانونی برای مجازات هوش مصنوعی تنها انسان مسئول شناخته می‌شود. 	تولیدکننده در صورت نادیده‌گرفتن نقص سیستم یا ایمنی، به جرم قتل یا جرح خطایی محکوم می‌شود.

۱. Model Penal Code، مدلی از قانون مجازات در ایالات متحده است که سطوح مختلف عنصر معنوی مرتکب را تعریف می‌کند.

نظام کیفری	مبانی تحقق جرم و عنصر روانی	وضعیت مسئولیت هوش مصنوعی	مسئول نهایی و نوع اتهام
ایالات متحده آمریکا	<ul style="list-style-type: none"> actus reus + mens rea ۴ سطح سوءنیت: عمد، علم، بی‌ملاحظگی و سهل‌انگاری (طبق MPC) 	<ul style="list-style-type: none"> عدم شناسایی هوش مصنوعی به‌عنوان مرتکب مستقل مسئولیت بر پایه عنصر روانی انسان است. 	<p>کاربر یا سازنده معمولاً به جرم قتل شبه‌عمد یا جرح ناشی از سهل‌انگاری محکوم می‌شوند.</p>

۴. رویه قضایی در پرتو مقررات موجود

تاکنون هیچ مرجع قضایی داخلی یا بین‌المللی هوش مصنوعی را به‌عنوان مرتکب مستقل با عنصر معنوی مجرمانه محکوم نکرده است. بیشتر پرونده‌ها به مسئولیت انسان‌های پشت سامانه‌ها پرداخته‌اند و در صورت وقوع حادثه، سازنده یا کاربر به اتهام بی‌احتیاطی یا عمد محاکمه شده‌اند. از منظر قوانین، اتحادیه اروپا در ژوئن ۲۰۲۴ مقرر ۱۶۸۹/۲۰۲۴ موسوم به قانون هوش مصنوعی را تصویب کرد که قواعد هماهنگ برای استفاده از سامانه‌های هوش مصنوعی را تعیین می‌کند. این مقرر بر شفافیت، تناسب، امنیت، حفاظت از داده‌ها و مسئولیت‌پذیری تأکید کرده و تصریح می‌کند که استفاده از سامانه‌های AI نباید جایگزین قدرت تصمیم‌گیری انسان شود (European Union, 2024).

در سپتامبر ۲۰۲۵، پارلمان ایتالیا قانون شماره ۱۳۲/۲۰۲۵ را تصویب کرد که به عنوان نخستین قانون ملی هوش مصنوعی در اروپا شناخته می‌شود. این قانون که طبق خطوط راهنمای مقرر اتحادیه تصویب شده، اصول استفاده از AI را تبیین و اجرا می‌کند و بر حفظ نظارت انسانی، شفافیت و امنیت تأکید دارد. همچنین، استفاده از AI در بخش‌های مختلف مانند سلامت، اشتغال، خدمات حرفه‌ای و دادگاه‌ها را تنظیم کرده و برای کاربردهای مخرب مجازات شدیدتری از جمله حبس پیش‌بینی می‌کند (Linklaters, 2025).

۵. راهکارهای تقنینی

با فرض پذیرش احتمالی مسئولیت مستقل برای سامانه‌های هوش مصنوعی در آینده، این پرسش مطرح می‌شود که چه نوع مجازاتی برای ماشین‌ها قابل تصور است. هدف اصلی مجازات کیفری انسانی بازدارندگی و اصلاح فرد مجرم است؛ اما اعمال این

اهداف بر سامانه‌های بی‌احساس دشوار است. جریمه نقدی برای ربان معنا ندارد و نابودی فیزیکی سامانه نیز اثری نمادین خواهد داشت. از این رو، پیشنهاد نویسندگان این است مجازات به تدابیر نمادین یا عملکردی معطوف گردد. برای نمونه، حذف نرم‌افزار یا قطع موقت دسترسی سیستم به شبکه‌ها می‌تواند معادل مجازات سخت تلقی شود. همچنین می‌توان از تولیدکننده یا مالک سامانه خواست الگوریتم‌ها را بازطراحی و گزارش‌های شفاف‌تری ارائه کند. در کنار آن، مجازات اجتماعی و اقتصادی مانند تعلیق فروش نسخه معیوب یا الزام به انجام خدمات عام‌المنفعه می‌تواند در نظر گرفته شود. این تدابیر بیش از آنکه بر «خود دستگاه» اثر بگذارند، در واقع رفتار انسان‌ها و نهادهای پشت سامانه را کنترل می‌کنند.

انواع مجازات و تدابیر قابل‌تصور در فرض پذیرش مسئولیت کیفری برای سامانه‌های هوش مصنوعی، همراه باهدف و اثر هر یک، در جدول شماره ۲ طبقه‌بندی شده است. مطابق جدول شماره ۲، این مجازات عمدتاً ماهیتی نمادین یا عملکردی دارند و بیش از آنکه متوجه خود سامانه باشند، رفتار مالکان و تولیدکنندگان را تنظیم می‌کنند.

جدول ۲. انواع مجازات قابل‌اعمال بر هوش مصنوعی

نوع مجازات	توضیح	هدف یا اثر اصلی
جریمه نقدی	برای ماشین بی‌معناست.	بی‌اثر بر سامانه (فقط بر مالک اثر دارد).
نابودی فیزیکی سامانه	حذف سخت‌افزار یا دستگاه فیزیکی.	اثر نمادین، بدون اصلاح واقعی مشکل.
حذف نرم‌افزار / قطع دسترسی شبکه	توقف عملکرد یا محدودسازی ارتباطات سیستم.	معادل «مجازات سخت» (اعدام/حبس) برای ماشین؛ بازدارندگی عملکردی.
بازطراحی الگوریتم‌ها و گزارش شفاف	الزام تولیدکننده یا مالک به اصلاح کدهای معیوب.	افزایش مسئولیت‌پذیری، شفافیت و کاهش خطای آینده.
تعلیق فروش نسخه معیوب	جلوگیری قانونی از عرضه محصول مشکل‌دار به بازار.	اعمال فشار اقتصادی و اجتماعی بر شرکت‌های سازنده.
خدمات عام‌المنفعه	الزام شرکت یا مالک به انجام فعالیتهای اجتماعی رایگان.	جبران خسارت وارده به جامعه و تقویت مسئولیت اجتماعی.

نتیجه‌گیری

یافته‌های این پژوهش نشان می‌دهد که در وضعیت کنونی فناوری و حقوق، سامانه‌های هوش مصنوعی فاقد آگاهی و اراده لازم برای تحقق عنصر معنوی جرم

هستند و بنابراین نمی‌توان آن‌ها را مرتکب مستقل تلقی کرد. نظام‌های حقوقی مختلف با وجود تفاوت‌های ساختاری، در این نکته اشتراک دارند که مسئولیت کیفری را به انسان‌ها نسبت می‌دهند و سامانه هوش مصنوعی را ابزار مادی قلمداد می‌کنند. با وجود این، قوانین جدید همچون مقررده ۱۶۸۹/۲۰۲۴ اتحادیه اروپا و قانون ملی ایتالیا نشان می‌دهد که کشورها در پی ایجاد چارچوب‌هایی برای تضمین شفافیت، نظارت انسانی و پیشگیری از سوءاستفاده از AI هستند.

از منظر سیاست‌گذاری، توصیه می‌شود در حال حاضر مسئولیت کیفری همچنان بر عهده انسان‌ها باقی بماند و تمرکز بر تقویت مقررات ایمنی، استانداردهای طراحی و الزامات شفافیت برای سامانه‌های هوش مصنوعی باشد. در عین حال، لازم است که حقوق‌دانان، مهندسان و سیاست‌گذاران همکاری نزدیکی داشته باشند تا در صورت دستیابی فناوری به سطحی از خودمختاری واقعی، مبانی نظری و حقوقی لازم برای تعیین مسئولیت فراهم شود. بازنگری در مفاهیم سنتی مانند قصد، علم و بی‌احتیاطی و تطبیق آن‌ها با شرایط جدید فناوری می‌تواند به ایجاد نظامی منصفانه‌تر و کارآمدتر در برخورد با چالش‌های آینده کمک کند.

تعارض منافع

تعارض منافع ندارم.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

منابع و مأخذ

«Autopilot» به سیستم هدایت خودکار هواپیما گفته می شود.

- American Law Institute (ALI). (1985). *Model Penal Code and Commentaries* (Official Draft and Revised Comments). Philadelphia: The American Law Institute.
- Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2017). *Of, for, and by the people: The legal lacuna of synthetic persons*. *Artificial Intelligence and Law*, 25(3), 273-291.
- Calo, R. (2016). *Robotics and the lessons of cyberlaw*. *California Law Review*, 103(3), 513-563.
- European Parliament. (2017). *Report with recommendations to the Commission on Civil Law Rules on Robotics* (2015/2103(INL)).
- European Union. (2024). *Artificial Intelligence Act* (Regulation (EU) 2024/1689). Publications Office of the European Union.
- Gless, S., Silverman, E., & Weigend, T. (2016). *If robots cause harm, who is to blame? Self-driving cars and criminal liability*. *New Criminal Law Review*, 19(3), 412-436. <https://doi.org/10.1525/nclr.2016.19.3.412>
- Hallevy, G. (2010). *The criminal liability of artificial intelligence entities - from science fiction to legal social control*. *Akron Intellectual Property Journal*, 4(2), 171-201.
- Hallevy, G. (2015). *When robots kill: Artificial intelligence under criminal law*. University Press of New England.
- Linklaters. (2025). *Italy - A pioneering national framework to complement the EU AI Act*. <https://www.linklaters.com>
- Nemitz, P. (2018). *Constitutional democracy and technology in the age of artificial intelligence*. *Philosophical Transactions of the Royal Society A*, 376(2133), 20180089.
- Piparo, C., & Blažek, R. (2024). *Criminal algorithms and their punishment in modern constitutionalism*. *Bratislava Law Review*, 8(2), 199-222. <https://doi.org/10.46282/blr.2024.8.2.832>
- Sartor, G. (2020). *AI and responsibility: A legal perspective*. *Philosophical Transactions of the Royal Society A*, 378(2166), 20190117.
- Wang, H., Blok, V., & van Hilten, M. (2025). *ELSA Labs for responsible AI: A novel approach for addressing ethical, legal, social issues*. *Journal of Responsible Innovation*, 12(1), 2563944. <https://doi.org/10.1080/23299460.2025.2563944>

