Comparative Studies in Jurisprudence, Law, and Politics

Civil Liability Arising from the Design, Production, and Supply of Artificial Intelligence

1. Mohammad Reza Qaboos Shakeri: Department of Law, Ra.C., Islamic Azad University, Rasht, Iran
2. Mahmood Kohani Khoshkbijari*: Department of Law, Ra.C., Islamic Azad University, Rasht, Iran.
Email: mahamood.kohani@iau.ac.ir (Corresponding Author)
3. Vahid Zarei Sharif: Department of Law, La.C., Islamic Azad University, Lahijan, Iran

ABSTRACT

The rapid expansion of artificial intelligence and its growing application in medical, financial, commercial, industrial, and service sectors have created significant legal challenges alongside its technological and economic benefits. The complexity of algorithms, the relative autonomy of intelligent systems, their capacity for learning and behavioral modification, the involvement of multiple actors throughout the technology lifecycle, and the difficulty of identifying the direct cause of harm have exposed certain limitations of traditional civil liability rules. This study examines civil liability arising from the design, production, and supply of artificial intelligence and seeks to determine appropriate legal bases for attributing damage to the various actors involved. The research adopts a descriptive-analytical approach based on general principles of civil liability and the distinctive characteristics of AI systems. The findings indicate that designers may incur liability where harm results from unsafe architecture, unreasonable algorithmic choices, inadequate safeguards, or failure to anticipate foreseeable risks. Producers may be held responsible for defective or biased data, insufficient testing, cybersecurity weaknesses, implementation errors, inadequate quality control, or failure to provide necessary updates. At the supply stage, liability may arise from inadequate disclosure, insufficient warnings, misleading representations, inappropriate deployment, or failure to respond to risks discovered after market entry. The analysis further demonstrates that exclusive reliance on fault-based liability may be insufficient in relation to high-risk AI systems, particularly where injured parties face substantial evidentiary difficulties. Accordingly, a differentiated framework based on the level of risk, facilitated standards of proof, risk-based liability, and compulsory insurance may be necessary in specific fields. Establishing a coherent legal framework that allocates responsibility among designers, producers, suppliers, and users can enhance victim protection while simultaneously promoting the responsible development and deployment of artificial intelligence.

Keywords: *Artificial Intelligence, Civil Liability, Design, Production, Supply, Fault, Risk-Based Liability.*

How to cite: Qaboos Shakeri, M. R., Kohani Khoshkbijari, M., & Zarei Shari, V. (2027). Civil Liability Arising from the Design, Production, and Supply of Artificial Intelligence. *Comparative Studies in Jurisprudence, Law, and Politics*, *9*(3), 1-19.

© 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: 06 April 2026
Revise Date: 08 September 2026
Accept Date: 14 September 2026
Initial Publish Date: 23 September 2026
Final Publish Date: 23 August 2027

پژوهش‌های تطبیقی فقه، حقوق و سیاست

مسئولیت مدنی ناشی از طراحی، تولید و عرضه هوش مصنوعی

۱. محمد رضا کابوس شاکری: گروه حقوق، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران
۲. محمود کهنی خشکبیجاری*: گروه حقوق، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران.
پست الکترونیک: mahamood.kohani@iau.ac.ir (نویسنده مسئول)
۳. وحید زارعی شریف: گروه حقوق، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران

چکیده

گسترش روزافزون هوش مصنوعی و نفوذ آن در حوزه‌های پزشکی، مالی، تجاری، صنعتی و خدماتی، در کنار مزایای گسترده، چالش‌های تازه‌ای در زمینه مسئولیت مدنی ایجاد کرده است. پیچیدگی الگوریتم‌ها، استقلال نسبی سامانه‌های هوشمند، قابلیت یادگیری و تغییر رفتار، تعدد بازیگران دخیل در چرخه حیات فناوری و دشواری شناسایی علت مستقیم زیان، سبب شده است که قواعد سنتی مسئولیت مدنی در برخی موارد با محدودیت مواجه شوند. پژوهش حاضر با هدف تحلیل مسئولیت مدنی ناشی از طراحی، تولید و عرضه هوش مصنوعی و تعیین مبانی مناسب برای انتساب خسارت به بازیگران مختلف این حوزه انجام شده است. روش پژوهش توصیفی ـ تحلیلی و مبتنی بر بررسی قواعد عمومی مسئولیت مدنی و ویژگی‌های خاص سامانه‌های هوش مصنوعی است. یافته‌ها نشان می‌دهد که مسئولیت طراح می‌تواند در مواردی مانند انتخاب معماری نامناسب، استفاده از الگوریتم غیرایمن، فقدان سازوکارهای کنترلی و پیش‌بینی‌نکردن خطرهای قابل انتظار تحقق یابد. مسئولیت تولیدکننده نیز می‌تواند ناشی از استفاده از داده‌های نامعتبر یا جانبدارانه، آزمون ناکافی، ضعف امنیت سایبری، خطا در پیاده‌سازی یا کوتاهی در به‌روزرسانی و کنترل کیفیت باشد. در مرحله عرضه، عدم ارائه اطلاعات کافی، هشدار ناقص، تبلیغات گمراه‌کننده، عرضه سامانه برای کاربرد نامتناسب و قصور در نظارت پس از عرضه می‌تواند مبنای مسئولیت قرار گیرد. نتایج همچنین نشان می‌دهد که در سامانه‌های هوش مصنوعی پرخطر، اتکای صرف بر مسئولیت مبتنی بر تقصیر ممکن است برای حمایت از زیان‌دیدگان کافی نباشد. بنابراین، استفاده از رویکردی مبتنی بر سطح خطر، تسهیل بار اثبات، مسئولیت مبتنی بر خطر و بیمه اجباری در برخی کاربردها ضروری به نظر می‌رسد. تدوین چارچوبی ویژه برای توزیع مسئولیت میان طراح، تولیدکننده، عرضه‌کننده و کاربر می‌تواند ضمن حمایت از زیان‌دیدگان، زمینه توسعه مسئولانه هوش مصنوعی را نیز فراهم سازد.

واژگان کلیدی: *هوش مصنوعی، مسئولیت مدنی، طراحی، تولید، عرضه، تقصیر، مسئولیت مبتنی بر خطر.*

نحوه استناددهی: کابوس شاکری، محمد رضا.، کهنی خشکبیجاری، محمود.، و زارعی شریف، وحید. (۱۴۰۶). مسئولیت مدنی ناشی از طراحی، تولید و عرضه هوش مصنوعی. *پژوهش‌های تطبیقی فقه، حقوق و سیاست*، ۹(۳)، ۱۹-۱.

© 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: ۱۷ فروردین ۱۴۰۵
تاریخ بازنگری: ۱۷ شهریور ۱۴۰۵
تاریخ پذیرش: ۲۳ شهریور ۱۴۰۵
تاریخ چاپ اولیه: ۱ مهر ۱۴۰۵
تاریخ چاپ نهایی: ۱ شهریور ۱۴۰۶

مقدمه

گسترش شتابان هوش مصنوعی در دهه اخیر، ساختار بسیاری از روابط اجتماعی، اقتصادی، حرفه‌ای و حتی شخصی را دگرگون کرده و در کنار فرصت‌های گسترده‌ای که برای افزایش بهره‌وری، کاهش هزینه‌ها، ارتقای دقت تصمیم‌گیری و توسعه خدمات نوین ایجاد کرده است، مجموعه‌ای از مسائل حقوقی جدید را نیز پیش روی نظام‌های حقوقی قرار داده است. هوش مصنوعی دیگر صرفاً یک ابزار فنی محدود به محیط‌های آزمایشگاهی یا صنعتی نیست، بلکه در تصمیم‌گیری‌های بانکی و اعتباری، تشخیص و درمان پزشکی، حمل‌ونقل، بیمه، تجارت الکترونیکی، استخدام، ارزیابی اشخاص، تولید محتوا، تحلیل داده‌های شخصی و ارائه انواع خدمات عمومی و خصوصی حضور یافته است. همین گسترش کاربرد سبب شده است که امکان ورود زیان ناشی از عملکرد سامانه‌های هوشمند نیز به‌طور محسوسی افزایش یابد. در بسیاری از این موارد، زیان نه مستقیماً از رفتار یک شخص انسانی، بلکه از عملکرد سامانه‌ای ناشی می‌شود که طراحی، آموزش، تولید، عرضه و بهره‌برداری از آن حاصل تعامل چندین شخص و نهاد مختلف است. این وضعیت، مدل سنتی مسئولیت مدنی را که معمولاً بر شناسایی رفتار زیان‌بار یک عامل انسانی مشخص، احراز تقصیر، اثبات رابطه سببیت و جبران خسارت استوار است، با دشواری‌های قابل توجهی مواجه می‌کند. پژوهش‌های حقوقی نیز نشان داده‌اند که یکی از چالش‌های اساسی مسئولیت مدنی در حوزه هوش مصنوعی، دشواری در انتساب زیان به یک شخص معین در زنجیره طراحی، توسعه و بهره‌برداری است (Kharitonova et al., 2022).

ویژگی اساسی بسیاری از سامانه‌های هوش مصنوعی در آن است که عملکرد آن‌ها ممکن است در مرحله استفاده دقیقاً مطابق با آنچه طراح اولیه پیش‌بینی کرده است نباشد. سامانه‌های مبتنی بر یادگیری ماشین می‌توانند بر مبنای داده‌های آموزشی، بازخوردهای محیطی و تعامل با کاربران، الگوهایی را استخراج کنند که نتیجه آن‌ها برای طراح یا تولیدکننده به‌صورت کامل قابل پیش‌بینی نباشد. از همین رو، در برخی موارد نمی‌توان به‌سادگی زیان را ناشی از یک «عیب» فنی مشخص دانست. ممکن است سامانه از منظر مهندسی به‌درستی عمل کرده باشد، اما نتیجه‌ای ایجاد کند که از نظر حقوقی تبعیض‌آمیز، زیان‌بار یا ناقض حقوق اشخاص باشد. مسئله مسئولیت مدنی در چنین شرایطی فراتر از تشخیص یک خرابی فنی است و به این پرسش مربوط می‌شود که چه کسی باید خطر ناشی از عملکرد سامانه‌ای پیچیده و تا حدی غیرقابل پیش‌بینی را تحمل کند. برخی مطالعات، ضرورت بازاندیشی در نظریه‌های سنتی تقصیر و انتساب را در مواجهه با این ویژگی‌های هوش مصنوعی مورد تأکید قرار داده‌اند (Biglarbeigi & Solhchi, 2023). در همین راستا، مفهوم الگوریتم متعارف و معقول نیز به‌عنوان معیاری برای ارزیابی رفتار طراحان و توسعه‌دهندگان مطرح شده است تا بتوان بر مبنای استانداردی فنی و حقوقی، حدود مراقبت لازم در طراحی و توسعه سامانه‌های هوشمند را تعیین کرد (Zakerinia & Gholampour, 2024a).

چالش مسئولیت مدنی زمانی پیچیده‌تر می‌شود که چرخه حیات یک سامانه هوش مصنوعی مورد توجه قرار گیرد. یک محصول یا خدمت مبتنی بر هوش مصنوعی ممکن است توسط یک شرکت طراحی شود، داده‌های آموزشی آن از منابع دیگری گردآوری شده باشد، مدل پایه آن توسط یک توسعه‌دهنده مستقل ایجاد شده باشد، سامانه توسط شرکت دیگری تجاری‌سازی شود، فروشنده یا پلتفرم دیگری آن را در اختیار مصرف‌کننده قرار دهد و در نهایت کاربر نیز آن را در محیط خاصی به کار گیرد. در چنین زنجیره‌ای، وقوع خسارت می‌تواند نتیجه رفتار یا قصور یک یا چند عامل باشد. برای مثال، نقص در معماری نرم‌افزار، کیفیت نامناسب داده‌های آموزشی، عدم شناسایی سوگیری‌ها، نبود آزمون‌های کافی پیش از عرضه، ضعف در هشداردهی به مصرف‌کننده، عدم ارائه به‌روزرسانی‌های ضروری یا استفاده نادرست کاربر، هر یک می‌تواند در ایجاد نتیجه زیان‌بار دخالت داشته باشد. بررسی مسئولیت بازیگران مرتبط با بهره‌برداری از سامانه‌های هوش مصنوعی نشان می‌دهد

که تعیین مسئولیت صرفاً با تمرکز بر کاربر نهایی یا تولیدکننده اصلی کافی نیست و باید شبکه‌ای از بازیگران و میزان کنترل هر یک بر خطر مورد تحلیل قرار گیرد (Keshavarz Safiei, 2025).

در حقوق ایران، قواعد عمومی مسئولیت مدنی بر مفاهیمی همچون ورود ضرر، فعل زیان‌بار، رابطه سببیت و در بسیاری موارد تقصیر استوار است. با این حال، اعمال این مفاهیم در زمینه هوش مصنوعی با دشواری‌های اثباتی و نظری روبه‌رو است. زیان‌دیده ممکن است هیچ دسترسی‌ای به کد، ساختار الگوریتم، داده‌های آموزشی یا سوابق تصمیم‌گیری سامانه نداشته باشد و بنابراین اثبات اینکه خطای سامانه ناشی از قصور طراح یا تولیدکننده بوده است برای او دشوار باشد. افزون بر این، برخی سامانه‌های هوش مصنوعی ماهیت «جعبه سیاه» دارند؛ بدین معنا که حتی متخصصان نیز ممکن است نتوانند به‌طور کامل توضیح دهند که چرا الگوریتم به یک نتیجه مشخص رسیده است. این عدم شفافیت می‌تواند رابطه میان رفتار عامل مسئول و خسارت را مبهم کند و بار اثبات را به‌گونه‌ای بر زیان‌دیده تحمیل نماید که عملاً امکان دریافت خسارت از بین برود. چالش‌های مشابه در مطالعات مربوط به نظام حقوقی ایران نیز مورد توجه قرار گرفته و ضرورت استفاده از تجربه مقررات‌گذاری اتحادیه اروپا برای بازنگری در قواعد داخلی مطرح شده است (Haji-Esmaeili, 2024).

مسئله دیگری که در تحلیل مسئولیت مدنی هوش مصنوعی اهمیت دارد، تمایز میان طراحی، تولید و عرضه است. طراحی به مرحله‌ای مربوط می‌شود که ساختار فنی، اهداف، معماری الگوریتم، داده‌های مورد استفاده و روش تصمیم‌گیری سامانه تعیین می‌شود. تولید می‌تواند شامل توسعه عملی نرم‌افزار، آموزش مدل، ادغام اجزای سخت‌افزاری و نرم‌افزاری و آماده‌سازی سامانه برای استفاده باشد. عرضه نیز مرحله‌ای است که سامانه به بازار یا در اختیار کاربر قرار می‌گیرد و در آن تعهداتی همچون ارائه اطلاعات صحیح، هشدار نسبت به خطرات، آموزش کاربر، تضمین امنیت، به‌روزرسانی و نظارت پس از عرضه اهمیت پیدا می‌کند. هر یک از این مراحل می‌تواند منشأ نوع متفاوتی از مسئولیت باشد. برای مثال، طراح ممکن است به دلیل انتخاب معماری غیرایمن مسئول شناخته شود، تولیدکننده به دلیل استفاده از داده‌های ناقص یا کنترل کیفیت ناکافی مسئول باشد و عرضه‌کننده به سبب عدم هشدار درباره محدودیت‌های سامانه یا ارائه محصولی که برای کاربرد خاصی مناسب نیست، مسئولیت پیدا کند. تحلیل الزامات مسئولیت مدنی در توسعه و استقرار هوش مصنوعی نشان می‌دهد که ارزیابی وظایف هر عامل باید متناسب با نقش او در چرخه حیات سامانه انجام گیرد (Fadaei & Taherkhani, 2023).

از سوی دیگر، ماهیت خسارات ناشی از هوش مصنوعی نیز متنوع است. بخشی از خسارات ممکن است مالی باشد، مانند تصمیم اشتباه یک الگوریتم اعتبارسنجی که منجر به محرومیت شخص از تسهیلات بانکی یا از دست رفتن فرصت اقتصادی شود. بخشی دیگر ممکن است جسمانی باشد، مانند اشتباه یک سامانه تشخیص پزشکی یا کنترل یک وسیله خودکار. خسارات معنوی نیز می‌توانند در نتیجه نقض حریم خصوصی، افشای اطلاعات، تبعیض الگوریتمی یا آسیب به اعتبار شخص ایجاد شوند. مطالعات مربوط به حریم خصوصی نشان داده‌اند که استفاده گسترده از هوش مصنوعی در پردازش داده‌های شخصی می‌تواند زمینه نقض حریم خصوصی و در نتیجه مسئولیت مدنی را فراهم آورد (Mortazavi & Khodaeifam, 2026). همین امر نشان می‌دهد که مسئولیت مدنی ناشی از هوش مصنوعی را نباید صرفاً در قالب مسئولیت ناشی از عیب محصول تحلیل کرد، بلکه باید آن را در پیوند با حقوق شخصیت، حریم خصوصی، امنیت داده، منع تبعیض و حق برخورداری از تصمیم‌گیری منصفانه نیز مورد بررسی قرار داد.

در حوزه پزشکی نیز کاربرد هوش مصنوعی به‌خوبی نشان می‌دهد که چگونه مرز میان مسئولیت طراح، تولیدکننده، عرضه‌کننده و استفاده‌کننده می‌تواند مبهم شود. اگر یک سامانه تشخیصی بر مبنای داده‌های آموزشی نامتناسب، بیماری را تشخیص ندهد، ممکن است مسئولیت متوجه

تولیدکننده یا طراح باشد، اما اگر پزشک بدون توجه به محدودیت‌های اعلام‌شده سامانه صرفاً بر نتیجه الگوریتم اتکا کند، مسئولیت او نیز قابل بررسی خواهد بود. تحلیل‌های تطبیقی در زمینه مسئولیت ناشی از کاربرد پزشکی هوش مصنوعی، بر ضرورت تعیین دقیق نقش پزشک، تولیدکننده و توسعه‌دهنده تأکید کرده‌اند (Badini, 2024). همچنین امکان تحقق مسئولیت در کاربردهای پزشکی هوش مصنوعی زمانی اهمیت بیشتری می‌یابد که سامانه مستقیماً بر سلامت و حیات افراد اثر می‌گذارد و اشتباه آن می‌تواند زیان‌های جبران‌ناپذیری ایجاد کند (Nikbakht Nasrabadi & Abbasi, 2023). بنابراین، هر الگویی که برای مسئولیت مدنی هوش مصنوعی طراحی می‌شود باید بتواند هم خسارات صرفاً اقتصادی و هم خسارات جسمانی، معنوی و مرتبط با حریم خصوصی را پوشش دهد.

پژوهش حاضر با هدف تحلیل مسئولیت مدنی ناشی از طراحی، تولید و عرضه هوش مصنوعی و تعیین مبانی مناسب برای انتساب خسارت در هر یک از این مراحل انجام می‌شود. مسئله اصلی آن است که قواعد موجود مسئولیت مدنی تا چه اندازه قادرند خسارات ناشی از سامانه‌های هوش مصنوعی را پوشش دهند و در چه مواردی نیاز به تفسیر جدید یا تدوین قواعد ویژه وجود دارد. همچنین باید روشن شود که معیار تشخیص تقصیر در طراحی و تولید سامانه‌های هوشمند چیست، رابطه سببیت در شرایط تعدد عوامل چگونه احراز می‌شود، چه تعهداتی بر عهده عرضه‌کننده قرار دارد و در چه شرایطی می‌توان از مسئولیت مبتنی بر تقصیر فاصله گرفت و به الگوهای مسئولیت مبتنی بر خطر، مسئولیت محض یا سازوکارهای بیمه‌ای نزدیک شد. رویکرد پژوهش، تحلیلی و حقوقی است و با تمرکز بر مفاهیم مسئولیت مدنی و ویژگی‌های فنی هوش مصنوعی، تلاش می‌کند الگویی منسجم برای توزیع مسئولیت میان طراح، تولیدکننده، عرضه‌کننده و در موارد لازم کاربر ارائه کند.

مبانی مفهومی و حقوقی مسئولیت مدنی در حوزه هوش مصنوعی

هوش مصنوعی از منظر حقوقی مفهومی یکدست و کاملاً ثابت نیست. در معنای گسترده، این اصطلاح به سامانه‌هایی اطلاق می‌شود که با استفاده از روش‌های محاسباتی، داده‌ها و الگوریتم‌ها قادرند وظایفی را انجام دهند که معمولاً نیازمند نوعی ادراک، استدلال، پیش‌بینی، طبقه‌بندی، تصمیم‌گیری یا یادگیری است. اهمیت این تعریف از آن جهت است که میزان استقلال و پیچیدگی سامانه می‌تواند بر نوع تحلیل مسئولیت مدنی اثر بگذارد. یک نرم‌افزار ساده که بر مبنای قواعد ثابت عمل می‌کند از نظر انتساب مسئولیت به طراح تفاوت قابل توجهی با سامانه‌ای دارد که بر مبنای یادگیری ماشین رفتار خود را با داده‌های جدید تنظیم می‌کند. در مورد نوع دوم، فاصله میان رفتار اولیه طراح و نتیجه نهایی سامانه بیشتر است و همین امر مسئله انتساب را دشوار می‌سازد. برخی مطالعات مربوط به شخصیت حقوقی هوش مصنوعی نیز از همین ویژگی استقلال عملکردی آغاز کرده‌اند و پرسیده‌اند که آیا در آینده می‌توان نوعی شخصیت مستقل برای سامانه‌های پیشرفته در نظر گرفت (Bagheri, 2025). با این حال، اعطای شخصیت حقوقی به سامانه هوشمند لزوماً مشکل جبران خسارت را حل نمی‌کند، زیرا شخصیت حقوقی بدون دارایی، بیمه یا سازوکار مالی مستقل، ممکن است صرفاً مسئولیت اشخاص واقعی یا حقوقی پشت سامانه را تضعیف کند.

مبنای کلاسیک مسئولیت مدنی بر این ایده استوار است که شخصی که به دیگری زیان وارد می‌کند، در شرایط مقرر قانونی باید آن زیان را جبران کند. در بسیاری از نظام‌های حقوقی، مسئولیت بر تقصیر بنا شده است؛ به این معنا که زیان‌دیده باید نشان دهد عامل زیان از رفتار شخص متعارف یا استاندارد حرفه‌ای مورد انتظار انحراف داشته است. این الگو در حوزه هوش مصنوعی با دو مانع اساسی مواجه می‌شود. نخست آنکه رفتار زیان‌بار ممکن است مستقیماً توسط انسان انجام نشده باشد، بلکه حاصل خروجی الگوریتم باشد. دوم آنکه اثبات قصور

انسانی در پشت آن خروجی می‌تواند دشوار باشد. بررسی مبانی مسئولیت مدنی مرتبط با هوش مصنوعی نشان داده است که نظریه تقصیر همچنان ظرفیت کاربرد دارد، اما باید با معیارهای جدیدی مانند وظیفه مراقبت در طراحی الگوریتم، کیفیت داده، قابلیت توضیح، آزمون‌پذیری و نظارت مستمر تکمیل شود (Hosseini & Yazdani, 2024). بنابراین تقصیر در این حوزه لزوماً به معنای خطای آشکار برنامه‌نویسی نیست، بلکه می‌تواند شامل کوتاهی در مدیریت خطرات قابل پیش‌بینی، عدم استفاده از داده مناسب، فقدان ارزیابی اثرات یا عدم ایجاد سازوکار توقف و مداخله انسانی باشد.

مفهوم «الگوریتم معقول و متعارف» در همین زمینه اهمیت پیدا می‌کند. همان‌گونه که در مسئولیت حرفه‌ای رفتار پزشک، مهندس یا وکیل با رفتار شخص حرفه‌ای متعارف مقایسه می‌شود، در مسئولیت ناشی از طراحی هوش مصنوعی نیز می‌توان پرسید آیا الگوریتم از نظر فنی و حقوقی مطابق سطح قابل انتظار ایمنی و دقت طراحی شده است یا خیر. این معیار امکان آن را فراهم می‌کند که بدون الزام به اثبات قصد یا علم طراح، عملکرد او بر اساس استانداردهای حرفه‌ای سنجیده شود. در ادبیات حقوقی ایران، نظریه انتساب مسئولیت به کمک مفهوم الگوریتم متعارف تقویت شده و بر این نکته تأکید شده است که اگر طراح از استانداردهای متعارف فنی و ایمنی عدول کند، می‌توان نتیجه زیان‌بار را با سهولت بیشتری به او منتسب ساخت (Zakerinia & Gholampour, 2024b). این رویکرد به‌ویژه در شرایطی مفید است که الگوریتم از نظر داخلی پیچیده است اما می‌توان روش طراحی، مجموعه داده، آزمون‌ها و تدابیر پیشگیرانه آن را با استانداردهای حرفه‌ای مقایسه کرد.

در برابر مسئولیت مبتنی بر تقصیر، نظریه خطر بر این اندیشه استوار است که شخصی که فعالیتی خطرناک را ایجاد یا از آن منتفع می‌شود، باید بخشی از زیان‌های ناشی از آن را نیز تحمل کند. کاربرد این نظریه در حوزه هوش مصنوعی به‌ویژه برای سامانه‌های پرخطر قابل دفاع است. برای مثال، اگر یک شرکت از استقرار سامانه خودکار در حوزه‌ای مانند حمل‌ونقل، پزشکی، زیرساخت‌های حیاتی یا تصمیم‌گیری مالی گسترده سود اقتصادی به دست می‌آورد، می‌توان استدلال کرد که تحمیل تمام بار اثبات تقصیر به زیان‌دیده عادلانه نیست. در برخی تحلیل‌های تطبیقی، حرکت از مسئولیت صرفاً مبتنی بر تقصیر به سمت الگوهای مبتنی بر خطر برای برخی کاربردهای هوش مصنوعی پیشنهاد شده است (Alipanahi et al., 2024). مزیت این رویکرد آن است که احتمال جبران خسارت افزایش می‌یابد، اما در مقابل باید از گسترش بیش از حد مسئولیت که ممکن است نوآوری را سرکوب کند جلوگیری شود. بنابراین تمایز میان سامانه‌های کم‌خطر و پرخطر اهمیت بنیادین پیدا می‌کند.

مسئولیت محض نیز شکل شدیدتری از مسئولیت بدون تقصیر است که در آن صرف تحقق رابطه میان فعالیت و زیان می‌تواند برای ایجاد مسئولیت کافی باشد. استفاده گسترده از این نظریه در تمام حوزه‌های هوش مصنوعی مناسب نیست، زیرا بسیاری از سامانه‌ها خطر محدودی دارند و تحمیل مسئولیت محض به تمامی توسعه‌دهندگان می‌تواند هزینه نوآوری را به‌طور نامتناسب افزایش دهد. با این حال، در مواردی که سامانه به‌طور بالقوه قادر به ایجاد خسارات شدید و گسترده است، مسئولیت محض می‌تواند از منظر حمایت از زیان‌دیده توجیه‌پذیر باشد. در آثار مربوط به مسئولیت ناشی از خسارات هوش مصنوعی نیز از ضرورت انتخاب مبنای متفاوت بر حسب ماهیت خطر و نوع کاربرد سخن گفته شده است (Ali, 2024). چنین رویکردی با اصل تناسب نیز سازگارتر است؛ زیرا یک سامانه توصیه‌گر محتوایی و یک سامانه کنترل وسیله خودران نباید الزاماً تابع یک رژیم واحد مسئولیت باشند.

یکی دیگر از عناصر بنیادین مسئولیت مدنی، تحقق ضرر است. در زمینه هوش مصنوعی، مفهوم ضرر باید به‌گونه‌ای تفسیر شود که علاوه بر خسارت مالی و جسمانی، خسارت ناشی از نقض حقوق اطلاعاتی و شخصیتی را نیز شامل شود. پردازش کلان‌داده، شناسایی الگوهای رفتاری

و پیش‌بینی ویژگی‌های افراد می‌تواند بدون ایجاد خسارت مالی فوری، حریم خصوصی یا استقلال تصمیم‌گیری آن‌ها را نقض کند. بررسی آثار هوش مصنوعی بر حریم خصوصی نشان می‌دهد که جمع‌آوری، ترکیب و تحلیل داده‌های شخصی می‌تواند در شرایطی به مسئولیت مدنی منجر شود، به‌ویژه هنگامی که استفاده از داده فراتر از رضایت یا انتظار معقول شخص انجام گیرد (Mortazavi & Khodaeifam, 2026). در محیط‌های پزشکی، این مسئله حساس‌تر است، زیرا داده‌های سلامت از خصوصی‌ترین اطلاعات افراد محسوب می‌شوند و افشا یا استفاده ناموجه از آن‌ها می‌تواند هم خسارت معنوی و هم زیان اقتصادی به همراه داشته باشد (Falahati Katilateh, 2025).

رابطه سببیت شاید پیچیده‌ترین رکن مسئولیت مدنی در حوزه هوش مصنوعی باشد. در الگوی سنتی، دادرس تلاش می‌کند نشان دهد اگر رفتار خوانده نبود، خسارت نیز رخ نمی‌داد یا اینکه رفتار او علت متعارف و قابل انتساب خسارت بوده است. اما در سامانه‌های هوشمند، زنجیره علی ممکن است از عوامل متعددی تشکیل شود. طراح معماری، تولیدکننده مدل، فراهم‌کننده داده، ارائه‌دهنده زیرساخت ابری، عرضه‌کننده نهایی و کاربر ممکن است هر یک در نتیجه دخیل باشند. علاوه بر این، سامانه ممکن است در طول زمان به‌روزرسانی شود و خروجی زیان‌بار حاصل تعامل نسخه‌های متعدد نرم‌افزار و داده‌های جدید باشد. مطالعات مربوط به مسئولیت ناشی از تصمیم‌گیری خودکار بر همین دشواری تأکید دارند و نشان می‌دهند که تحلیل رابطه سببیت باید از الگوی خطی فاصله گرفته و به سمت تحلیل چندعاملی حرکت کند (Farajpour, 2025). در این شرایط، ممکن است قواعد مربوط به تعدد اسباب، مسئولیت مشترک یا تقسیم مسئولیت بر اساس میزان مشارکت در خطر اهمیت بیشتری پیدا کند.

انتساب حقوقی نیز باید از سببیت فنی تفکیک شود. ممکن است از منظر فنی بتوان نشان داد که یک بخش از کد یا یک داده خاص در ایجاد خروجی نقش داشته است، اما این امر به‌تنهایی برای مسئول شناختن شخص مربوط کافی نیست. مسئولیت حقوقی معمولاً مستلزم این است که زیان در حوزه خطر قابل انتساب به آن شخص قرار گیرد. برای نمونه، اگر تولیدکننده سامانه را برای استفاده محدود در محیط تحقیقاتی عرضه کرده اما شخص دیگری بدون مجوز آن را در محیط درمانی به کار برده باشد، صرف وجود رابطه فنی میان سامانه و زیان لزوماً برای انتساب کامل مسئولیت به تولیدکننده کافی نخواهد بود. برعکس، اگر تولیدکننده از امکان استفاده پزشکی آگاه بوده و محصول را با ادعاهای تبلیغاتی مشخص برای چنین کاربردی عرضه کرده باشد، دامنه خطر قابل انتساب به او گسترده‌تر می‌شود. این نکته در تحلیل مسئولیت بازیگران مرتبط با بهره‌برداری از هوش مصنوعی اهمیت ویژه دارد (Keshavarz Safiei, 2025).

بار اثبات نیز یکی از موضوعات محوری است. اگر زیان‌دیده موظف باشد ساختار درونی الگوریتم، داده‌های آموزشی، خطاهای آزمون و تصمیم‌های توسعه‌دهندگان را اثبات کند، عملاً در برابر شرکت‌های بزرگ فناوری در موقعیت بسیار نابرابری قرار خواهد گرفت. بنابراین یکی از راهکارهای حقوقی می‌تواند تسهیل ادله اثبات، الزام به نگهداری سوابق فنی، ایجاد حق دسترسی به اطلاعات لازم و در برخی موارد جابه‌جایی یا تعدیل بار اثبات باشد. تجربه اتحادیه اروپا در این زمینه بر شفافیت، مستندسازی و امکان ردیابی تأکید دارد و تحلیل‌های داخلی نیز از ظرفیت این رویکرد برای اصلاح حقوق ایران سخن گفته‌اند (Haji-Esmaeili, 2024). چنین سازوکارهایی نه تنها به زیان‌دیده کمک می‌کند، بلکه تولیدکننده را نیز به ایجاد نظام مستندسازی منظم و مدیریت ریسک ترغیب می‌نماید.

در نهایت، مبانی حقوقی مسئولیت هوش مصنوعی باید میان دو هدف تعادل برقرار کند: حمایت مؤثر از زیان‌دیدگان و حفظ امکان نوآوری. مسئولیتی که بیش از حد سنگین و غیرقابل پیش‌بینی باشد می‌تواند مانع توسعه فناوری شود، اما رژیمی که اثبات مسئولیت را تقریباً ناممکن کند نیز هزینه‌های اجتماعی فناوری را به اشخاص بی‌دفاع منتقل می‌کند. راه‌حل مناسب، طراحی نظامی لایه‌ای است که در آن میزان مسئولیت

با نوع سامانه، سطح خطر، قدرت کنترل عامل، میزان انتفاع اقتصادی، امکان پیش‌بینی و توان پیشگیری از زیان ارتباط داشته باشد. چنین رویکردی می‌تواند مسئولیت مبتنی بر تقصیر را برای سامانه‌های عادی حفظ کند، در حوزه‌های پرخطر از فرض تقصیر یا تعدیل بار اثبات استفاده نماید و در موارد بسیار خطرناک به مسئولیت بدون تقصیر یا بیمه اجباری متوسل شود.

مسئولیت مدنی ناشی از طراحی و تولید هوش مصنوعی

طراحی نخستین مرحله‌ای است که بسیاری از خطرهای مرتبط با یک سامانه هوش مصنوعی در آن شکل می‌گیرند. تصمیم درباره هدف سامانه، نوع داده‌های مورد استفاده، متغیرهای ورودی، معماری مدل، معیارهای بهینه‌سازی، سطح خودکاربودن و امکان مداخله انسانی همگی در این مرحله اتخاذ می‌شوند. از منظر مسئولیت مدنی، طراح زمانی می‌تواند مسئول شناخته شود که در انتخاب یا پیاده‌سازی این عناصر از سطح مراقبت متعارف عدول کرده و این انحراف در ایجاد خسارت نقش داشته باشد. برای مثال، طراحی سامانه‌ای که بدون در نظر گرفتن داده‌های متنوع جمعیتی برای تصمیم‌گیری درباره افراد استفاده می‌شود می‌تواند به نتایج تبعیض‌آمیز منجر شود. اگر چنین خطری در زمان طراحی قابل پیش‌بینی بوده باشد، طراح نمی‌تواند صرفاً به این استدلال متوسل شود که خروجی نهایی توسط الگوریتم تولید شده است. در تحلیل‌های مربوط به الگوریتم متعارف نیز بر این نکته تأکید شده که رعایت استانداردهای حرفه‌ای در انتخاب داده و طراحی فرایند تصمیم‌گیری باید یکی از معیارهای اصلی انتساب باشد (Zakerinia & Gholampour, 2024a).

نقص طراحی زمانی رخ می‌دهد که محصول مطابق نقشه و ساختار پیش‌بینی‌شده تولید شده اما خود آن ساختار به‌گونه‌ای است که خطر نامعقول ایجاد می‌کند. در هوش مصنوعی، این نقص می‌تواند در قالب انتخاب نامناسب مدل، تعریف اشتباه تابع هدف، عدم پیش‌بینی مداخله انسانی، ضعف سازوکار امنیتی یا نادیده گرفتن شرایط خاص استفاده ظاهر شود. تفاوت اساسی با محصولات سنتی آن است که عیب طراحی در هوش مصنوعی همیشه به شکل یک نقص ثابت و قابل مشاهده نیست. ممکن است سامانه در اغلب موارد عملکرد مناسبی داشته باشد اما در شرایط خاص به‌صورت سیستماتیک خطا کند. برای تشخیص مسئولیت در این موارد باید پرسید آیا طراح می‌توانسته با هزینه و فناوری معقول، طراحی ایمن‌تری انتخاب کند. بررسی مبانی مسئولیت ناشی از توسعه هوش مصنوعی نیز نشان می‌دهد که آزمون گزینه جایگزین ایمن‌تر می‌تواند در ارزیابی تقصیر تولیدکننده یا طراح نقش مهمی ایفا کند (Fadaei & Taherkhani, 2023).

کیفیت داده‌های آموزشی یکی از مهم‌ترین ابعاد مسئولیت در مرحله طراحی و تولید است. الگوریتم‌های یادگیری ماشین رفتار خود را بر مبنای الگوهای موجود در داده می‌آموزند و اگر داده ناقص، منحرف، قدیمی یا تبعیض‌آمیز باشد، خروجی نیز ممکن است همان کاستی‌ها را بازتولید کند. بنابراین، وظیفه مراقبت طراح فقط به کدنویسی محدود نیست، بلکه شامل بررسی منشأ، اعتبار، تنوع و تناسب داده‌ها نیز می‌شود. اگر تولیدکننده سامانه‌ای را برای تشخیص پزشکی طراحی کند اما داده‌های آموزشی آن عمدتاً مربوط به یک گروه خاص باشد، استفاده از سامانه در جمعیتی متفاوت ممکن است دقت آن را کاهش دهد. مطالعات حوزه پزشکی هوش مصنوعی این مشکل را به‌عنوان یکی از منابع بالقوه مسئولیت مورد توجه قرار داده‌اند (Badini, 2024). در چنین شرایطی، اگر عدم نمایندگی مناسب داده برای متخصصان قابل پیش‌بینی باشد، کوتاهی در اصلاح آن می‌تواند مصداق تقصیر حرفه‌ای تلقی شود.

نقص داده ممکن است به‌طور مستقیم به تولیدکننده منتسب نباشد، زیرا داده‌ها گاه از تأمین‌کنندگان مستقل خریداری می‌شوند. با این حال، استفاده از داده شخص ثالث تولیدکننده را به‌طور کامل از مسئولیت معاف نمی‌کند. همان‌طور که تولیدکننده یک کالای فیزیکی موظف است کیفیت اجزای مورد استفاده را تا حد متعارف کنترل کند، توسعه‌دهنده هوش مصنوعی نیز باید درباره کیفیت داده‌های حیاتی ارزیابی انجام

دهد. اگر داده از منبعی ناشناخته یا فاقد استاندارد مناسب دریافت شود و هیچ کنترل معقولی بر آن صورت نگیرد، اتکا به شخص ثالث نمی‌تواند دفاع کامل باشد. در مقابل، اگر تأمین‌کننده داده تضمین‌های نادرست ارائه کرده باشد، امکان رجوع تولیدکننده به او در روابط داخلی قابل تصور است. این رویکرد با تحلیل چندعاملی مسئولیت سازگار است و مانع از آن می‌شود که پیچیدگی زنجیره تأمین به ابزاری برای فرار از مسئولیت تبدیل شود.

آزمون و ارزیابی پیش از عرضه بخش دیگری از مسئولیت تولیدکننده است. سامانه هوش مصنوعی باید پیش از ورود به بازار در شرایط مختلف مورد آزمایش قرار گیرد و نقاط ضعف آن شناسایی شود. هرچه کاربرد سامانه پرخطرتر باشد، سطح آزمون مورد انتظار نیز باید بالاتر باشد. برای مثال، استاندارد قابل قبول برای یک سامانه توصیه فیلم نمی‌تواند با استاندارد سامانه‌ای که در تشخیص بیماری یا کنترل تجهیزات صنعتی استفاده می‌شود یکسان باشد. مطالعات تطبیقی در زمینه مسئولیت سامانه‌های هوشمند بر اهمیت ارزیابی خطر پیش از استقرار تأکید کرده‌اند (Kharitonova et al., 2022). از این منظر، تقصیر تولیدکننده می‌تواند نه فقط در طراحی معیوب، بلکه در فقدان آزمون‌های کافی، نادیده گرفتن نتایج هشداردهنده یا عرضه شتاب‌زده سامانه تحقق یابد.

موضوع قابلیت توضیح نیز با مسئولیت طراحی ارتباط مستقیم دارد. هرچند برخی مدل‌های پیشرفته از نظر فنی پیچیده‌اند، تولیدکننده باید تا حد معقول سازوکاری برای فهم منطق تصمیم یا دست‌کم ردیابی عوامل مؤثر بر آن فراهم کند، به‌ویژه زمانی که تصمیم الگوریتمی بر حقوق و منافع اساسی اشخاص اثر می‌گذارد. اگر سامانه به‌گونه‌ای طراحی شود که هیچ امکان مؤثری برای توضیح یا بررسی بعدی تصمیم وجود نداشته باشد، اثبات خطا و اصلاح آن بسیار دشوار خواهد شد. در حوزه تصمیم‌گیری خودکار، فقدان شفافیت یکی از چالش‌های اصلی مسئولیت شناخته شده است (Farajpour, 2025). بنابراین، قابلیت توضیح صرفاً ویژگی فنی مطلوب نیست، بلکه می‌تواند بخشی از تکلیف حقوقی مراقبت تلقی شود.

امنیت سایبری نیز یکی از ابعاد مهم مسئولیت تولیدکننده است. سامانه هوش مصنوعی ممکن است در برابر دستکاری داده، حملات خصمانه، تزریق داده‌های مسموم یا سوءاستفاده از رابط‌های برنامه‌نویسی آسیب‌پذیر باشد. اگر تولیدکننده تدابیر متعارف امنیتی را رعایت نکرده باشد و مهاجم با بهره‌گیری از همین ضعف موجب خسارت شود، مداخله شخص ثالث لزوماً رابطه سببیت را قطع نمی‌کند. در حقوق مسئولیت، اگر رفتار مجرمانه یا مداخله شخص ثالث قابل پیش‌بینی باشد، عامل اولیه ممکن است همچنان مسئول باقی بماند. این اصل در مورد هوش مصنوعی اهمیت بیشتری دارد، زیرا حملات سایبری جزء خطرات شناخته‌شده فناوری محسوب می‌شوند. بنابراین طراحی ایمن، آزمون نفوذ، مدیریت دسترسی، به‌روزرسانی امنیتی و پیش‌بینی واکنش به رخداد باید بخشی از معیار رفتار متعارف تولیدکننده باشد.

در مرحله تولید، مسئله «نقص تولید» نیز مطرح است. ممکن است طراحی اولیه صحیح باشد اما نسخه‌ای از سامانه به دلیل خطای برنامه‌نویسی، پیکربندی نادرست، نصب اشتباه یا خرابی یک جزء به شکلی متفاوت از طراحی اصلی عمل کند. در این حالت، مسئولیت می‌تواند به واحد تولید، یکپارچه‌ساز یا پیمانکاری که سامانه را نصب کرده است نسبت داده شود. برای تشخیص مسئولیت باید زنجیره کنترل کیفیت بررسی شود. اگر شرکت سازنده فرایند مؤثری برای کنترل نسخه‌ها، ثبت تغییرات و آزمون نهایی نداشته باشد، همین قصور می‌تواند مبنای مسئولیت قرار گیرد. تحلیل مسئولیت در حقوق ایران نیز نشان می‌دهد که قواعد عمومی مسئولیت قابلیت آن را دارند که نقص‌های ناشی از تولید یا استقرار نادرست را پوشش دهند، هرچند برای سامانه‌های پیچیده ممکن است قواعد تخصصی‌تر لازم باشد (Shafizadeh, 2023).

به‌روزرسانی مدل‌ها مرز میان تولید و دوره پس از عرضه را نیز تغییر داده است. برخلاف محصول سنتی که پس از فروش تقریباً ثابت می‌ماند، سامانه هوش مصنوعی می‌تواند از راه دور تغییر کند. تولیدکننده ممکن است الگوریتم، مدل یا داده‌های سامانه را پس از عرضه به‌روزرسانی کند و همین به‌روزرسانی منشأ خطر جدید شود. بنابراین مسئولیت تولیدکننده پایان‌یافته تلقی نمی‌شود. اگر نسخه اولیه ایمن باشد ولی به‌روزرسانی بعدی باعث کاهش دقت یا افزایش خطر شود، مسئولیت باید با توجه به وضعیت جدید بررسی گردد. همچنین اگر آسیب‌پذیری شناخته‌شده‌ای کشف شود و تولیدکننده با وجود امکان اصلاح، از ارائه به‌روزرسانی خودداری کند، این ترک فعل می‌تواند مبنای مسئولیت باشد. در اینجا رابطه زمانی میان علم تولیدکننده به خطر و زمان وقوع خسارت اهمیت پیدا می‌کند.

یکی از دشوارترین مسائل، استقلال یادگیری سامانه است. ممکن است تولیدکننده ادعا کند که رفتار زیان‌بار پس از ماه‌ها استفاده و یادگیری در محیط کاربر شکل گرفته و در زمان عرضه قابل پیش‌بینی نبوده است. این دفاع تنها در صورتی قابل پذیرش است که تولیدکننده واقعاً تدابیر معقول برای محدودکردن رفتار نامطلوب، نظارت و تشخیص انحراف طراحی کرده باشد. اگر سامانه به‌گونه‌ای ساخته شده باشد که امکان تغییر قابل توجه رفتار داشته باشد، همین قابلیت بخشی از خطر محصول است و تولیدکننده باید برای آن سازوکار کنترلی مناسب ایجاد کند. نظریه مسئولیت بازیگران مرتبط با بهره‌برداری نیز بر میزان کنترل و امکان پیشگیری هر عامل تأکید می‌کند (Keshavarz Safiei, 2025).

در نتیجه، هرچه تولیدکننده کنترل بیشتری بر مدل، به‌روزرسانی و سازوکار یادگیری داشته باشد، امکان انتساب زیان به او بیشتر است.

تعدد طراحان و توسعه‌دهندگان نیز توزیع مسئولیت را پیچیده می‌کند. در پروژه‌های بزرگ، یک شرکت ممکن است مدل پایه را توسعه دهد، شرکت دیگر آن را برای کاربرد خاص تنظیم کند و مجموعه‌ای دیگر آن را با داده محلی آموزش دهد. در چنین وضعی نباید به دنبال یک مسئول واحد بود، بلکه باید میزان مشارکت هر عامل در ایجاد خطر بررسی شود. اگر چند عامل به‌طور مشترک شرایط وقوع زیان را فراهم کرده باشند و تفکیک دقیق سهم آن‌ها دشوار باشد، استفاده از قواعد مسئولیت مشترک یا تضامنی می‌تواند حمایت مؤثرتری از زیان‌دیده ایجاد کند. سپس مسئولان می‌توانند در روابط داخلی بر اساس سهم تقصیر یا میزان دخالت به یکدیگر رجوع کنند. این روش از تحمیل دشواری فنی شناسایی سهم دقیق هر عامل بر زیان‌دیده جلوگیری می‌کند.

در تحلیل مسئولیت طراح و تولیدکننده باید میان خطای قابل اجتناب و محدودیت ذاتی فناوری تفاوت گذاشت. هیچ سامانه‌ای از خطا مصون نیست و مسئولیت نباید به معنای تضمین مطلق نتیجه تلقی شود. اگر تولیدکننده تمام استانداردهای شناخته‌شده را رعایت کرده، خطرها را ارزیابی کرده، محدودیت‌ها را اعلام نموده و سامانه در شرایطی نادر و غیرقابل پیش‌بینی خطا کرده باشد، مسئولیت مبتنی بر تقصیر ممکن است قابل احراز نباشد. با این حال، در حوزه‌های بسیار پرخطر می‌توان همچنان از مسئولیت مبتنی بر خطر یا بیمه اجباری استفاده کرد تا زیان‌دیده بدون نیاز به اثبات تقصیر جبران شود. چنین تفکیکی میان مسئولیت تقصیری و مسئولیت ناشی از خطر می‌تواند تعادل میان نوآوری و حمایت از زیان‌دیدگان را حفظ کند.

مسئولیت مدنی ناشی از عرضه و بهره‌برداری از هوش مصنوعی

عرضه هوش مصنوعی صرفاً به معنای انتقال یک کالا یا نرم‌افزار از تولیدکننده به مصرف‌کننده نیست، بلکه مرحله‌ای است که در آن سامانه وارد محیط واقعی می‌شود و خطرهای بالقوه آن می‌توانند تحقق خارجی پیدا کنند. عرضه‌کننده، توزیع‌کننده، پلتفرم ارائه‌دهنده و شرکت خدماتی ممکن است هر یک نقش مستقلی در این فرایند داشته باشند. از منظر مسئولیت مدنی، شخصی که محصول را به بازار ارائه می‌کند وظیفه دارد اطلاعات کافی درباره ماهیت، حدود کارکرد، خطرها و شیوه صحیح استفاده در اختیار مصرف‌کننده قرار دهد. اگر سامانه فقط برای

کاربرد خاصی طراحی شده باشد اما عرضه‌کننده آن را برای استفاده گسترده‌تر تبلیغ کند، مسئولیت او می‌تواند مستقل از مسئولیت تولیدکننده شکل گیرد. در مطالعات مربوط به حقوق اتحادیه اروپا نیز مسئولیت ناشی از استفاده و عرضه سامانه‌های هوش مصنوعی در پیوند با تعهدات اطلاع‌رسانی و مدیریت خطر بررسی شده است (Alipanahi et al., 2024).

تعهد به هشدار یکی از مهم‌ترین تعهدات عرضه‌کننده است. در محصولاتی که خطر آن‌ها برای مصرف‌کننده آشکار نیست، صرف نبود عیب فنی برای رفع مسئولیت کافی نیست. اگر عرضه‌کننده بداند یا باید بداند که سامانه در شرایط خاص عملکرد ضعیفی دارد، باید این محدودیت را به‌روشنی اعلام کند. برای نمونه، اگر یک سامانه تشخیص پزشکی برای گروه سنی مشخصی دقت پایین‌تری دارد، این مسئله باید به کاربر حرفه‌ای اطلاع داده شود. در غیر این صورت، حتی اگر طراحی سامانه مطابق استاندارد باشد، فقدان هشدار کافی می‌تواند خود منشأ مسئولیت باشد. در حوزه پزشکی، اهمیت اطلاع‌رسانی درباره محدودیت‌های سامانه به دلیل اثر مستقیم آن بر تصمیم پزشک و ایمنی بیمار بیشتر است (Heydari, 2023). همچنین مطالعات مربوط به امکان مسئولیت در کاربرد پزشکی تأکید کرده‌اند که استفاده‌کننده حرفه‌ای باید بتواند حدود اعتماد مجاز به خروجی سامانه را تشخیص دهد (Nikbakht Nasrabadi & Abbasi, 2023).

شیوه تبلیغ و معرفی محصول نیز می‌تواند در ارزیابی مسئولیت اثرگذار باشد. اگر عرضه‌کننده سامانه را به‌عنوان «کاملاً دقیق»، «بدون خطا» یا «جایگزین کامل قضاوت انسانی» معرفی کند، انتظار مصرف‌کننده را به سطحی افزایش می‌دهد که ممکن است از ظرفیت واقعی محصول فراتر باشد. در چنین شرایطی، زیان ناشی از اعتماد متعارف مصرف‌کننده به این ادعاها می‌تواند به عرضه‌کننده منتسب شود. از این رو، مسئولیت در حوزه هوش مصنوعی فقط از ویژگی‌های فنی سامانه ناشی نمی‌شود، بلکه محتوای قرارداد، تبلیغات، دستورالعمل‌ها و اظهارات بازاریابی نیز در تعیین دامنه تعهدات مؤثر است. اگر عرضه‌کننده خطرهای شناخته‌شده را پنهان کند یا درباره دقت و قابلیت محصول اغراق نماید، مسئولیت او می‌تواند بر مبنای تقصیر در اطلاع‌رسانی یا حتی نقض تعهد قراردادی توجیه شود.

تعهد به شفافیت نیز در این مرحله اهمیت دارد. مصرف‌کننده یا کاربر حرفه‌ای باید بداند که با یک سامانه هوش مصنوعی تعامل دارد، تصمیم چگونه تولید می‌شود و تا چه حد امکان اعتراض یا بررسی انسانی وجود دارد. هرچند ارائه جزئیات کامل فنی همیشه لازم یا ممکن نیست، اطلاعاتی که برای استفاده ایمن و ارزیابی نتیجه ضروری است باید قابل دسترس باشد. در حوزه تصمیم‌گیری خودکار، شفافیت اهمیت مضاعفی دارد زیرا فرد ممکن است بر اثر تصمیمی الگوریتمی از اعتبار، استخدام، بیمه یا فرصت دیگر محروم شود بدون اینکه علت آن را بداند. تحلیل مسئولیت در تصمیم‌گیری خودکار نشان داده است که فقدان شفافیت می‌تواند هم حقوق اشخاص را تضعیف کند و هم احراز مسئولیت را دشوار سازد (Farajpour, 2025).

حفظ حریم خصوصی نیز یکی از تعهدات اصلی عرضه‌کنندگان خدمات مبتنی بر هوش مصنوعی است. بسیاری از سامانه‌ها پس از عرضه همچنان داده‌های کاربران را جمع‌آوری و پردازش می‌کنند. بنابراین رابطه مسئولیت به مرحله فروش ختم نمی‌شود. اگر عرضه‌کننده داده‌های بیشتری از آنچه برای ارائه خدمت ضروری است جمع‌آوری کند، آن‌ها را برای اهداف دیگری به کار گیرد یا حفاظت کافی از اطلاعات نداشته باشد، ممکن است مسئولیت مدنی ایجاد شود. آثار مربوط به هوش مصنوعی و حریم خصوصی نشان می‌دهد که خسارت ناشی از پردازش ناموجه داده می‌تواند حتی در نبود زیان مالی مستقیم، به دلیل نقض حق شخصیت و آرامش روانی فرد قابل بررسی باشد (Mortazavi & Khodaeifam, 2026). در محیط‌های پزشکی نیز نقض محرمانگی داده بیمار از طریق سامانه هوشمند می‌تواند مبنای مستقل مسئولیت تلقی شود (Falahati Katilateh, 2025).

مسئولیت پس از عرضه یکی از مهم‌ترین تفاوت‌های محصولات هوشمند با محصولات سنتی است. عرضه‌کننده و تولیدکننده ممکن است پس از ورود محصول به بازار از وجود نقص یا آسیب‌پذیری جدید آگاه شوند. در چنین وضعی، سکوت یا عدم اقدام می‌تواند تقصیر محسوب شود. وظیفه مراقبت می‌تواند شامل هشدار فوری، ارائه اصلاحیه نرم‌افزاری، تعلیق خدمت یا حتی فراخوان سامانه باشد. هرچه خطر شدیدتر باشد، سرعت و گستره اقدام مورد انتظار نیز بیشتر است. اگر شرکت پس از دریافت گزارش‌های متعدد درباره خطای خطرناک، همچنان سامانه را بدون تغییر در اختیار کاربران قرار دهد، مسئولیت او به‌سادگی قابل توجیه خواهد بود. این امر نشان می‌دهد که ارزیابی مسئولیت باید کل چرخه عمر محصول را پوشش دهد.

از سوی دیگر، کاربر نیز ممکن است در وقوع خسارت نقش داشته باشد. اگر دستورالعمل‌های روشن سامانه نادیده گرفته شود، محصول برای کاربردی خارج از محدوده مجاز استفاده گردد یا هشدارهای ارائه‌شده عمداً بی‌اعتنا گذاشته شود، سهم کاربر در ایجاد خسارت باید مورد توجه قرار گیرد. در حوزه پزشکی، پزشک نمی‌تواند صرفاً به دلیل استفاده از سامانه هوش مصنوعی مسئولیت حرفه‌ای خود را به‌طور کامل منتقل کند. اگر نتیجه الگوریتم با علائم آشکار بیمار ناسازگار باشد، اتکای کورکورانه به آن می‌تواند تقصیر کاربر حرفه‌ای محسوب شود. مطالعات تطبیقی در حقوق ایران و عراق نیز نقش استفاده‌کننده حرفه‌ای را در مسئولیت ناشی از هوش مصنوعی پزشکی مورد توجه قرار داده‌اند (Heydari, 2023). بنابراین، مسئولیت باید با توجه به میزان استقلال سامانه و سطح اختیاری که برای کاربر باقی مانده است توزیع شود.

با این حال، نمی‌توان به بهانه وجود کاربر انسانی، تمام مسئولیت را به او منتقل کرد. در بسیاری از سامانه‌ها، کاربر عملاً امکان ارزیابی فنی نتیجه را ندارد و به دلیل پیچیدگی فناوری ناگزیر به اعتماد به تولیدکننده است. برای مثال، بیمار یا مصرف‌کننده عادی نمی‌تواند صحت مدل یادگیری ماشین را بررسی کند. حتی پزشک نیز ممکن است به جزئیات داده‌های آموزشی یا معماری مدل دسترسی نداشته باشد. در چنین شرایطی، اصل تخصیص مسئولیت باید بر مبنای «قدرت کنترل بر خطر» شکل گیرد. شخصی که امکان بیشتری برای طراحی، اصلاح، آزمون یا توقف سامانه دارد باید سهم بیشتری از مسئولیت را تحمل کند. این معیار در تحلیل مسئولیت بازیگران هوش مصنوعی نیز مورد توجه قرار گرفته است (Keshavarz Safiei, 2025).

موقعیت پلتفرم‌های واسط نیز نیازمند توجه است. برخی شرکت‌ها خود تولیدکننده مدل نیستند اما آن را از طریق خدمات ابری، رابط‌های برنامه‌نویسی یا فروشگاه‌های نرم‌افزاری در اختیار دیگران قرار می‌دهند. اگر نقش آن‌ها صرفاً انتقال فنی باشد، مسئولیتشان ممکن است محدودتر باشد، اما اگر در تنظیم، سفارشی‌سازی، تبلیغ یا کنترل محصول نقش فعال داشته باشند، نمی‌توان آن‌ها را صرفاً واسطه تلقی کرد. دامنه مسئولیت باید با سطح مداخله و آگاهی آن‌ها از خطر متناسب باشد. اگر پلتفرم از گزارش‌های متعدد درباره زیان آگاه شده اما همچنان بدون محدودیت خدمت را ارائه کند، ترک اقدام می‌تواند مبنای مسئولیت قرار گیرد.

از منظر قراردادی نیز عرضه هوش مصنوعی پرسش‌های خاصی ایجاد می‌کند. شرکت‌ها ممکن است تلاش کنند با درج شروط محدودکننده مسئولیت، تعهد خود را کاهش دهند. اعتبار چنین شروطی باید با قواعد آمره، حمایت از مصرف‌کننده و ماهیت خسارت سنجیده شود. در مواردی که خسارت ناشی از تقصیر سنگین، نقض حقوق اساسی یا آسیب جسمانی باشد، پذیرش معافیت کامل از مسئولیت می‌تواند با نظم عمومی ناسازگار باشد. به‌ویژه در قراردادهای الحاقی که کاربر قدرت مذاکره واقعی ندارد، نباید اجازه داد عرضه‌کننده تمام خطر فناوری را به مصرف‌کننده منتقل کند. بنابراین، شروط قراردادی می‌توانند مسئولیت را تنظیم کنند اما نباید جایگزین استانداردهای حداقلی ایمنی و مراقبت شوند.

در نظام‌های مسئولیت مدنی مربوط به هوش مصنوعی، بیمه می‌تواند نقش مکمل مهمی داشته باشد. برای سامانه‌های پرخطر، بیمه اجباری می‌تواند اطمینان دهد که حتی در شرایطی که شناسایی مسئول دشوار است، زیان‌دیده امکان جبران دارد. این راهکار به‌ویژه در حوزه‌هایی مناسب است که خسارت بالقوه بالا و پیچیدگی فنی زیاد است. در چنین مدلی، شرکت بیمه پس از جبران خسارت می‌تواند بر اساس سهم مسئولیت به تولیدکننده، عرضه‌کننده یا کاربر رجوع کند. ترکیب مسئولیت مدنی و بیمه همچنین می‌تواند تولیدکنندگان را به رعایت استانداردهای ایمنی ترغیب کند، زیرا حق بیمه می‌تواند بر اساس سطح خطر و کیفیت مدیریت ریسک تعیین شود.

مسئله دیگری که باید در حوزه عرضه مورد توجه قرار گیرد، تفاوت میان عرضه عمومی و عرضه تخصصی است. وقتی سامانه برای مصرف‌کننده عادی ارائه می‌شود، تعهد اطلاع‌رسانی باید ساده‌تر و روشن‌تر باشد و نمی‌توان از کاربر انتظار دانش فنی داشت. در مقابل، در عرضه به شرکت یا متخصص، سطح دانش مورد انتظار بالاتر است، اما همچنان عرضه‌کننده باید اطلاعات حیاتی را ارائه کند. بنابراین معیار تقصیر در هشداردهی نسبی است و با نوع مخاطب تغییر می‌کند. همین منطق درباره آموزش نیز صدق می‌کند. اگر استفاده ایمن از سامانه مستلزم آموزش تخصصی باشد، عرضه بدون فراهم‌کردن آموزش کافی می‌تواند رفتار زیان‌بار تلقی شود.

در حقوق ایران، بهره‌گیری از قواعد عمومی مسئولیت مدنی می‌تواند بخش قابل توجهی از این مسائل را پوشش دهد، اما پیچیدگی فناوری ضرورت تدوین استانداردهای خاص را آشکار می‌کند. مطالعات داخلی بر چالش‌های موجود در انتساب مسئولیت، اثبات تقصیر و رابطه سببیت تأکید کرده‌اند (Haji-Esmaeili, 2024). همچنین پژوهش‌های مربوط به مسئولیت حقوقی هوش مصنوعی نشان می‌دهند که عدم وجود مقررات تخصصی ممکن است باعث تشتت رویه و عدم پیش‌بینی‌پذیری شود (Biglarbeigi & Solhchi, 2023). بر همین اساس، تدوین مقرراتی درباره سطح خطر، الزامات ثبت و مستندسازی، شفافیت، حفظ داده، به‌روزرسانی، بیمه و تقسیم مسئولیت میان بازیگران مختلف می‌تواند امنیت حقوقی را برای توسعه‌دهندگان و زیان‌دیدگان افزایش دهد.

الگوی مناسب مسئولیت در مرحله عرضه باید انعطاف‌پذیر باشد. برای سامانه‌های کم‌خطر، مسئولیت مبتنی بر تقصیر و قواعد قراردادی می‌تواند کفایت کند. در سامانه‌های متوسط‌خطر، می‌توان از فرض تقصیر، الزام به مستندسازی و انتقال بخشی از بار اثبات استفاده کرد. در حوزه‌های پرخطر، مسئولیت بدون تقصیر، بیمه اجباری و الزامات نظارتی شدیدتر قابل توجیه است. این تفکیک بر اساس خطر، امکان می‌دهد مقررات نه بیش از حد سخت‌گیرانه و نه بیش از حد مسامحه‌آمیز باشد. هدف اصلی باید آن باشد که شخصی که از فعالیت اقتصادی مبتنی بر هوش مصنوعی منفعت می‌برد و بیشترین قدرت را برای کنترل خطر دارد، نتواند هزینه زیان را به اشخاصی منتقل کند که هیچ نقشی در ایجاد یا مدیریت آن نداشته‌اند.

نتیجه‌گیری

هوش مصنوعی ساختار کلاسیک مسئولیت مدنی را با چالش‌هایی مواجه کرده است که از ماهیت فنی، شبکه‌ای و تا حدی خودمختار این فناوری ناشی می‌شوند. در الگوهای سنتی مسئولیت، معمولاً امکان شناسایی رفتار زیان‌بار یک شخص، بررسی تقصیر او و احراز رابطه مستقیم میان رفتار و خسارت وجود دارد، اما در سامانه‌های هوشمند این زنجیره می‌تواند میان طراح، توسعه‌دهنده، تأمین‌کننده داده، تولیدکننده، عرضه‌کننده، اپراتور، کاربر و خود فرایند یادگیری الگوریتم توزیع شود. بنابراین، تلاش برای تحمیل تمام مسئولیت به یک بازیگر واحد نه با واقعیت فنی هوش مصنوعی سازگار است و نه می‌تواند همواره به نتیجه عادلانه منجر شود. راه‌حل مناسب مستلزم پذیرش این واقعیت است

که مسئولیت مدنی هوش مصنوعی ماهیتی چندلایه دارد و باید بر اساس نقش، میزان کنترل، قابلیت پیش‌بینی و قدرت پیشگیری هر یک از عوامل تعیین شود.

در مرحله طراحی، معیار اصلی باید رعایت استانداردهای فنی، ایمنی و حقوقی متعارف باشد. طراح موظف است اهداف سامانه، معماری الگوریتم، داده‌های مورد استفاده و سازوکارهای کنترلی را به‌گونه‌ای انتخاب کند که خطرهای قابل پیش‌بینی کاهش یابد. طراحی هوش مصنوعی نباید صرفاً مسئله‌ای فنی تلقی شود؛ زیرا انتخاب داده، تابع هدف، معیار تصمیم‌گیری و میزان خودکارسازی می‌تواند مستقیماً بر حقوق اشخاص اثر بگذارد. به همین دلیل، مسئولیت طراح زمانی قابل توجیه است که خطر ناشی از یک تصمیم طراحی در زمان ایجاد سامانه قابل پیش‌بینی بوده و با استفاده از روش معقول و در دسترس می‌توانسته کاهش یابد.

در مرحله تولید، کنترل کیفیت، آزمون، ارزیابی ایمنی، بررسی سوگیری داده، امنیت سایبری و مستندسازی باید از عناصر اصلی وظیفه مراقبت محسوب شوند. تولیدکننده نمی‌تواند صرفاً با استناد به پیچیدگی الگوریتم از مسئولیت رهایی یابد. پیچیدگی فناوری بخشی از فعالیتی است که او انتخاب کرده و از آن منفعت اقتصادی به دست می‌آورد. در نتیجه، هرچه سامانه پیچیده‌تر و بالقوه خطرناک‌تر باشد، سطح مراقبت مورد انتظار نیز باید بالاتر باشد. در عین حال، مسئولیت نباید به تضمین مطلق عملکرد تبدیل شود. وقوع هر خطا به‌خودی خود اثبات‌کننده تقصیر نیست و باید میان خطای اجتناب‌پذیر و محدودیت ذاتی فناوری تفاوت گذاشت.

در مرحله عرضه، تعهدات اطلاع‌رسانی و هشدار باید جایگاهی اساسی داشته باشند. عرضه‌کننده باید کاربر را از محدودیت‌های سامانه، خطرهای شناخته‌شده، شرایط استفاده صحیح و سطح قابل اعتماد بودن نتیجه آگاه سازد. اغراق در تبلیغ، پنهان‌کردن محدودیت‌ها یا معرفی سامانه برای کاربردی فراتر از ظرفیت واقعی آن می‌تواند منشأ مستقل مسئولیت باشد. همچنین عرضه‌کننده و تولیدکننده پس از ورود سامانه به بازار نیز باید نسبت به خطرهای جدید بی‌تفاوت نباشند. ظهور آسیب‌پذیری یا خطای تازه می‌تواند وظیفه هشدار، اصلاح، به‌روزرسانی یا تعلیق خدمت را ایجاد کند. بنابراین، مسئولیت هوش مصنوعی باید کل چرخه عمر سامانه را در بر گیرد.

در حوزه رابطه سببیت نیز باید از نگاه کاملاً خطی فاصله گرفت. در بسیاری از خسارات ناشی از هوش مصنوعی چند عامل هم‌زمان نقش دارند و ممکن است تعیین سهم دقیق هر یک برای زیان‌دیده امکان‌پذیر نباشد. در چنین شرایطی، نظام حقوقی باید از سازوکارهایی استفاده کند که بار دشواری فنی را به‌طور کامل بر زیان‌دیده تحمیل نکند. مسئولیت مشترک، فرض سببیت در شرایط خاص، تعدیل بار اثبات و امکان رجوع مسئولان به یکدیگر می‌تواند از راهکارهای مناسب باشد. معیار اصلی باید آن باشد که شخصی که بیشترین کنترل را بر منشأ خطر دارد، متناسب با همان کنترل مسئولیت بیشتری نیز بر عهده گیرد.

در خصوص مبنای مسئولیت، نمی‌توان برای تمام کاربردهای هوش مصنوعی یک الگوی واحد در نظر گرفت. سامانه‌هایی که خطر محدود دارند می‌توانند همچنان تابع مسئولیت مبتنی بر تقصیر باشند. در سامانه‌هایی که بر تصمیمات مهم اقتصادی، حرفه‌ای یا شخصی اثر می‌گذارند، می‌توان با فرض تقصیر یا تسهیل بار اثبات، حمایت بیشتری از زیان‌دیده ایجاد کرد. در حوزه‌هایی که احتمال خسارت شدید جسمانی، مالی یا اجتماعی وجود دارد، مسئولیت بدون تقصیر یا مبتنی بر خطر قابل دفاع‌تر است. بدین ترتیب، رژیم مسئولیت باید با سطح خطر متناسب شود.

در حقوق ایران، قواعد عمومی مسئولیت مدنی ظرفیت اولیه لازم برای رسیدگی به بسیاری از خسارات ناشی از هوش مصنوعی را دارند، اما این قواعد برای مواجهه با همه پیچیدگی‌های فناوری کافی نیستند. مفاهیم سنتی تقصیر، سببیت و انتساب باید در پرتو ویژگی‌های خاص

سامانه‌های هوشمند تفسیر شوند و در کنار آن‌ها مقررات ویژه‌ای درباره شفافیت، مستندسازی، حفاظت از داده، آزمون ایمنی، به‌روزرسانی، هشدار و تقسیم مسئولیت تدوین شود. نبود چنین مقرراتی می‌تواند به تفاوت در تصمیم‌های قضایی و کاهش قابلیت پیش‌بینی حقوقی منجر شود.

یکی از پیشنهادهای اساسی، طبقه‌بندی سامانه‌های هوش مصنوعی بر اساس سطح خطر و تعیین الزامات متفاوت برای هر طبقه است. سامانه‌های پرخطر باید پیش از عرضه تحت ارزیابی تخصصی قرار گیرند و تولیدکننده آن‌ها ملزم به نگهداری سوابق طراحی، داده، آزمون و به‌روزرسانی باشد. در صورت وقوع خسارت، این سوابق باید در اختیار مرجع قضایی قرار گیرد تا امکان بررسی دقیق‌تر رابطه سببیت فراهم شود. در مواردی که تولیدکننده از ارائه اطلاعات خودداری می‌کند یا اسناد لازم را نگهداری نکرده است، امکان ایجاد اماره به زیان او می‌تواند مورد توجه قرار گیرد.

پیشنهاد دیگر، توسعه بیمه مسئولیت ویژه برای هوش مصنوعی است. بیمه می‌تواند در حوزه‌های پرخطر، شکاف میان وقوع خسارت و تعیین قطعی مسئول را کاهش دهد و از این طریق تضمین بیشتری برای جبران زیان فراهم سازد. به‌ویژه در مواردی که چند عامل در ایجاد خسارت نقش دارند یا احراز دقیق سهم تقصیر زمان‌بر است، جبران اولیه توسط بیمه و رجوع بعدی به عاملان مسئول می‌تواند کارآمدتر باشد.

در نهایت، نظام مطلوب مسئولیت مدنی هوش مصنوعی باید بر سه اصل اساسی استوار شود: جبران مؤثر خسارت، تخصیص مسئولیت به شخصی که بیشترین کنترل و توان پیشگیری را دارد، و حفظ تعادل میان حمایت حقوقی و امکان نوآوری. هدف نباید جلوگیری از توسعه هوش مصنوعی باشد، بلکه باید هزینه‌های فناوری به‌گونه‌ای منصفانه میان بازیگران آن توزیع شود. اگر منافع اقتصادی و اجتماعی هوش مصنوعی عمدتاً نصیب طراحان، تولیدکنندگان و عرضه‌کنندگان می‌شود، منطقی است که خطرهای غیرمتعارف ناشی از آن نیز صرفاً بر دوش مصرف‌کنندگان و زیان‌دیدگان قرار نگیرد. بر این مبنا، تدوین یک چارچوب حقوقی منسجم که مسئولیت طراح، تولیدکننده، عرضه‌کننده و کاربر را به‌صورت تفکیک‌شده ولی هماهنگ تنظیم کند، می‌تواند مهم‌ترین گام در ایجاد نظام عادلانه مسئولیت مدنی در عصر هوش مصنوعی باشد.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافعی 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 has transformed the structure of economic, professional, medical, commercial, and social relations, while simultaneously creating new legal questions regarding the allocation of civil liability for harms caused by intelligent systems. Unlike conventional products, AI systems may operate through complex algorithmic processes, learn from data, modify their behavior over time, and generate outputs that are not always fully predictable by their designers or users. These characteristics challenge traditional models of civil liability that rely on identifying a specific human act, proving fault, establishing causation, and attributing damage to a clearly identifiable actor. The legal difficulty becomes more acute because the lifecycle of an AI

system usually involves multiple actors, including designers, developers, data providers, producers, distributors, service providers, platform operators, and end users. Existing scholarship has emphasized that the diffusion of control among multiple participants complicates the attribution of harmful consequences and requires a reconsideration of conventional liability doctrines (Kharitonova et al., 2022). Similar concerns have been raised regarding the adequacy of Iranian civil liability rules in dealing with algorithmic opacity, autonomous decision-making, and the evidentiary difficulties faced by injured parties (Haji-Esmaeili, 2024). The objective of this study is therefore to analyze the civil liability arising from the design, production, and supply of artificial intelligence, to identify the legal basis for assigning responsibility at each stage of the AI lifecycle, and to determine whether fault-based liability remains sufficient or should be complemented by risk-based, strict, or insurance-supported liability mechanisms.

From a conceptual and doctrinal perspective, civil liability in the field of artificial intelligence must be examined in light of the specific technological characteristics of AI systems. Traditional fault-based liability generally requires proof that the defendant failed to comply with a legally expected standard of care; however, in AI-related disputes, fault may not take the form of an obvious programming error. It may instead consist of negligent model selection, inadequate testing, failure to detect data bias, insufficient cybersecurity, lack of transparency, omission of appropriate human oversight, or disregard of foreseeable risks. Jurisprudential and legal analyses have therefore suggested that conventional fault should be reinterpreted through standards adapted to algorithmic systems and professional technological conduct (Hosseini & Yazdani, 2024). In this context, the notion of a reasonable and customary algorithm provides a useful benchmark for evaluating whether designers and developers have complied with expected technical and legal standards (Zakerinia & Gholampour, 2024b). Rather than requiring courts to understand every internal computational process, this approach enables them to assess whether the developer adopted reasonable design methods, appropriate datasets, sufficient testing procedures, and adequate safeguards. At the same time, sole reliance on fault may be inadequate for high-risk AI applications because injured persons often lack access to source code, training data, validation records, and other technical evidence necessary to demonstrate negligence. Comparative research has therefore supported differentiated liability regimes that vary according to the level of technological risk and the extent of control exercised by the responsible actor (Alipanahi et al., 2024). Such a model can preserve fault-based liability for low-risk systems while allowing evidentiary presumptions, risk-based liability, or stricter forms of responsibility for systems capable of causing serious personal, financial, or social harm.

Civil liability arising from the design and production of AI begins with the recognition that many risks are created long before a system reaches the market. Decisions concerning the architecture of the model, the definition of its objectives, the selection of input variables, the scope of automation, the quality of training data, and the availability of human intervention can substantially determine the probability and severity of future harm. A design defect may exist where the system operates precisely as intended but the underlying design itself generates an unreasonable level of risk. In AI systems, such defects may appear as systematic bias, unsafe optimization goals, inadequate safeguards, non-representative datasets, or insufficient mechanisms for detecting abnormal outputs. Research on the development and deployment of artificial intelligence has identified the duty to test, validate, monitor, and document system performance as an important component of the producer's standard of care (Fadaei & Taherkhani, 2023). Data quality is particularly significant because machine-learning systems derive patterns from training datasets, meaning that incomplete, outdated, or discriminatory data may produce harmful results even in the absence of conventional

software malfunction. This issue is especially important in medical applications, where unrepresentative data may lead to inaccurate diagnoses or treatment recommendations for particular populations (Badini, 2024). The producer's responsibility should therefore extend beyond programming and include the reasonable verification of data provenance, reliability, representativeness, and suitability. Moreover, the producer cannot automatically escape responsibility by relying on third-party data suppliers, because the use of external components does not eliminate the duty to conduct reasonable quality control. Where several developers or entities jointly contribute to the system, liability may need to be distributed according to their contribution to the risk, degree of control, and capacity to prevent the harm.

The supply and post-market operation of artificial intelligence create a distinct set of legal duties that cannot be reduced to design and manufacturing obligations. A supplier may be liable where a technically sound product is marketed inaccurately, provided without adequate warnings, or used in contexts for which it was not reasonably suitable. Information duties are therefore central to AI liability, particularly where the product contains limitations that are not evident to ordinary users. Suppliers should communicate the intended use of the system, known limitations, foreseeable risks, necessary human supervision, and circumstances in which the output should not be relied upon. This obligation is especially important in medical AI, where the relationship between algorithmic recommendations and professional judgment must remain clear (Heydari, 2023). Excessive or misleading claims concerning accuracy, autonomy, or reliability may increase the supplier's responsibility because users may reasonably rely on such representations when deciding how much trust to place in the system. Automated decision-making also raises transparency concerns, since individuals affected by algorithmic decisions may be unable to understand or challenge outcomes concerning employment, credit, insurance, or access to services. Legal analysis of automated decisions has highlighted the importance of explainability and procedural safeguards in establishing accountability (Farajpour, 2025). In addition, AI suppliers may have continuing obligations after the initial release of the product because intelligent systems can be remotely updated, modified, or exposed to newly discovered vulnerabilities. Failure to warn users, issue corrective updates, suspend unsafe functionality, or respond to known defects may constitute an independent basis of liability. Thus, the relevant duty of care extends throughout the lifecycle of the system rather than ending at the moment of sale.

The allocation of liability becomes even more complex when multiple actors jointly contribute to the occurrence of damage. In AI ecosystems, the designer may determine the model architecture, another company may provide the training data, a third party may customize the system, a platform may distribute it, and a professional user may rely on the final output. In such cases, a linear model of causation is often inadequate because the harmful outcome may result from the interaction of several independent or cumulative factors. Contemporary research on AI liability has therefore stressed the need to analyze responsibility in terms of control, foreseeability, contribution to risk, and capacity to prevent harm (Keshavarz Safiei, 2025). The user's role should also be considered, particularly where clear instructions are ignored or the system is used beyond its intended scope. Nevertheless, the mere presence of a human user should not automatically shift responsibility away from producers and suppliers. Ordinary users, and even many professional users, may have no realistic ability to inspect source code, evaluate training data, or identify hidden algorithmic defects. This imbalance of information and technical capacity justifies mechanisms that facilitate proof for injured parties, including disclosure duties, mandatory documentation, retention of technical logs, rebuttable presumptions, and in appropriate cases a partial shift in the burden of proof. Privacy-

related damage further demonstrates why AI liability cannot be limited to bodily or property harm. Intelligent systems frequently collect, infer, combine, and analyze personal information in ways that may violate privacy, dignity, or informational autonomy (Mortazavi & Khodaeifam, 2026). In sensitive settings such as healthcare, unauthorized processing or disclosure of patient data may itself constitute an independent source of civil liability (Falahati Katilateh, 2025). Accordingly, an effective liability regime must encompass economic, bodily, moral, informational, and privacy-related harms. The analysis ultimately demonstrates that civil liability arising from the design, production, and supply of artificial intelligence should be structured as a differentiated, lifecycle-based regime rather than through a single uniform rule. Designers should be responsible for foreseeable risks created by defective architecture, inappropriate objectives, inadequate safeguards, or unreasonable algorithmic choices; producers should bear responsibility for failures in implementation, testing, data quality, cybersecurity, quality control, and post-production modification; and suppliers should be accountable for inaccurate marketing, insufficient warnings, inadequate disclosure, unsafe deployment, and failure to respond to newly discovered risks. The degree of liability should correspond to the level of risk created by the system, the actor's control over that risk, the foreseeability of harm, and the practical ability to prevent or mitigate it. Low-risk systems may remain primarily subject to fault-based liability, whereas medium- and high-risk systems may justify evidentiary presumptions, partial reversal of the burden of proof, risk-based responsibility, compulsory insurance, or strict liability in narrowly defined circumstances. Iranian civil liability law contains general principles capable of addressing some AI-related harms, but these principles require further clarification and supplementation through specific rules on documentation, transparency, algorithmic auditing, data governance, cybersecurity, post-market monitoring, and allocation of responsibility among multiple actors. A coherent legal framework should neither inhibit technological innovation nor allow technological complexity to operate as a shield against accountability. Its central purpose should be to ensure effective compensation for injured persons while assigning the economic consequences of AI-related risks to those who create, control, and benefit from such technologies.

References

Ali, A. H. (2024). *Civil Liability for Artificial Intelligence Damages*. New University Publishing House.

Alipanahi, M., Nasiran Najafabadi, D., & Shirani, M. (2024). Civil liability arising from artificial intelligence use in the European Union. *Scientific Quarterly Journal of Economic Jurisprudence Studies*, *6*(5), 1-18.

Badini, N. (2024). An Analytical Perspective on Civil Liability Arising from the Use of Artificial Intelligence in Medicine: A Comparative Study. *Medical Law Scientific Research Journal*, *18*(59), 99-112LA - Persian.

Bagheri, P. (2025). Legal Personality and Civil Liability of Artificial Intelligence: Legal Challenges and Solutions. *Scientific-Research Quarterly of Private Law*, *13*(51), 47-82.

Biglarbeigi, K., & Solhchi, S. (2023). Artificial Intelligence and Legal Liability. *Legal Civilization Quarterly*, *6*(18).

Fadaei, H., & Taherkhani, M. (2023). Essentials of civil liability arising from the development and deployment of artificial intelligence. Seventh International Conference on Modern Studies in Computer Science and Electrical Engineering,

Falahati Katilateh, N. (2025). Civil Liability Arising from Violation of Patients' Privacy Rights through the Use of Artificial Intelligence Systems in Medical Environments.

Farajpour, R. (2025). Legal and Comparative Analysis of Civil Liability of Artificial Intelligence in Automated Decision-Making. *AI and Tech in Behavioral and Social Sciences*, *3*(3), 1-9. https://doi.org/10.61838/kman.aitech.3.3.4

Haji-Esmaeili, M. (2024). Challenges of Civil Liability for Artificial Intelligence in Iran's Legal System: A Look at EU Regulations. *State and Law*, *15*(5), 81-98.

Heydari, C. (2023). *Civil Liability Arising from the Application of Artificial Intelligence in Medicine: A Comparative Study in Iraqi and Iranian Law* Shiraz University].

Hosseini, A. S., & Yazdani, S. (2024). Challenges of artificial intelligence and its civil liability foundations in jurisprudence. *Comparative Research in Jurisprudence, Law, and Politics*, *6*(5), 2-20.

Keshavarz Safiei, E. (2025). The Basis of Civil Liability of Actors Related to the Operation of Artificial Intelligence Systems in European Union Documents and Iranian Law. *Legal Research*, *28*(4).

Kharitonova, Y. S., Savina, V., & Pagnini, F. (2022). Civil Liability in the Development and Application of Artificial Intelligence and Robotic Systems: Basic Approaches. *Вестник Пермского Университета Юридические Науки*(4(58)), 683-708. https://doi.org/10.17072/1995-4190-2022-58-683-708

Mortazavi, S. M., & Khodaeifam, H. (2026). Investigating the effects of using artificial intelligence on personal privacy and the resulting civil liability. *Modern Fiqh and Law Quarterly*, *6*(22), 1-14.

Nikbakht Nasrabadi, M., & Abbasi, M. (2023). Feasibility of Civil Liability Arising from the Use of Artificial Intelligence in Medicine. *Health Law Journal*, *1*(1).

Shafizadeh, S. (2023). A study and analysis of civil liability arising from artificial intelligence performance in Iran's legal system. First International Conference on Advocacy, Law, and Humanities,

Zakerinia, H., & Gholampour, Z. (2024a). Reasonable and Customary Algorithms and Reinforcing the Theory of Attributability of Civil Liability for Artificial Intelligence. *Journal of New Technologies Law*, *5*(9).

Zakerinia, H., & Gholampour, Z. (2024b). Reasonable and customary algorithms and strengthening the theory of imputability of civil liability for artificial intelligence. *Journal of Contract Law and Emerging Technologies*, *5*(1), 155-168.