APJ 7(1):61-75, Spring 2026

Arman Process Journal (APJ)

Quarterly Journal of ICT APJ

License Number: 87090

Arman Process Journal (APJ)

Homepage: https://www.armanprocessjournal.ir

The Impact of Artificial Intelligence on Corporate Governance: Focusing on Three Critical Accountability Gaps and Strategies to Strengthen Board Oversight

M. Ghasemi[*,1], **M. Jafarpour Jalali** [2]

[1] Department of Management, Adiban Garmsar Institute of Higher Education, Garmsar, Semnan, Iran

[2] Department of Electrical Engineering, Adiban Garmsar Institute of Higher Education, Garmsar, Semnan, Iran

RESEARCH PAPER

Received: 2025-112-22
Accepted: 2026-3-25

KEYWORDS:
Artificial Intelligence,
Corporate Governance,
Board of Directors,
Accountability Gap,
Algorithmic Decision-Making

[1] Corresponding author:
maziarghasemi@adiban.ac.ir

ABSTRACT

Artificial intelligence has enhanced the accuracy, speed, and efficiency of strategic decision-making processes; however, it has simultaneously posed unprecedented challenges to the traditional accountability framework in corporate governance. Using a critical analytical approach and drawing upon findings from authoritative research on AI and corporate governance, this study argues that existing legal and ethical frameworks are inadequately prepared to address three structural accountability gaps arising from algorithmic decision-making at the board level. Accountability is defined as a relationship among an agent, the board of directors, and the general assembly of shareholders, in which the agent is required to justify their actions, while the assembly is empowered to question, evaluate performance, determine the continuation of cooperation, and take legal action when necessary. The first gap concerns the absence of a clear rule for assigning responsibility when AI systems produce erroneous outcomes, since responsibility is a multidimensional concept and merely attributing responsibility cannot resolve the associated ethical challenges. This gap is particularly evident in augmented intelligence and reinforcement learning systems. The second gap is the lack of a governance mechanism to oversee the board of directors in its role as the supervisor of AI systems, referred to in this study as the *monitor-of-the-monitor* dilemma. The third gap arises from the temporal mismatch between the decision-making cycles of boards of directors and the real-time operation of AI systems, which undermines effective oversight and timely accountability. To address these challenges, the study proposes three policy-oriented solutions: the principle of non-delegable ultimate human responsibility, the establishment of a specialized AI committee within the board of directors with continuous auditing authority, and the redefinition of fiduciary duty so that algorithmic literacy and oversight of the AI model lifecycle become integral components of directors' legal responsibilities. These solutions are examined through the four objectives of accountability-compliance, reporting, oversight, and enforcement- and collectively provide a conceptual framework for strengthening accountability, enhancing board oversight, and aligning corporate governance with the demands of AI-driven decision-making.

Copyright © Author(s).

CC BY NC

نشریه تخصصی آرمان پردازش، دوره ۷، شماره ۱، بهار ۱۴۰۵

فصلنامه تخصصی آرمان پردازش
(APJ)
Homepage: www.armanprocessjournal.ir

فصلنامه تخصصی فناوری اطلاعات و ارتباطات
شماره مجوز: ۸۷۰۹۰

تأثیر هوش مصنوعی بر حاکمیت شرکتی با تمرکز بر سه خلأ حیاتی پاسخگویی و راهکارهای تقویت نظارت هیئت مدیره

مازیار قاسمی*[۱]، محمد جعفرپور جلالی

[۱] گروه مدیریت، موسسه آموزش عالی ادیبان گرمسار ، سمنان، ایران

[۲] گروه مهندسی برق، موسسه آموزش عالی ادیبان گرمسار ، سمنان، ایران

چکیده

مقاله پژوهشی

واژگان کلیدی:
هوش مصنوعی،
حاکمیت شرکتی،
هیئت مدیره،
خلأ پاسخگویی،
تصمیم‌گیری الگوریتمی،

هوش مصنوعی با ورود به فرآیندهای تصمیم‌گیری راهبردی، دقت، سرعت و کارایی را افزایش داده، اما همزمان نظام سنتی پاسخگویی در حاکمیت شرکتی را با چالش‌های بی‌سابقه‌ای مواجه ساخته است. این مقاله با روش تحلیل انتقادی و بهره‌گیری از یافته‌های پژوهش‌های معتبر در حوزه هوش مصنوعی و حاکمیت شرکتی، استدلال می‌کند که نظام حقوقی و اخلاقی موجود برای پاسخگویی به سه خلأ ساختاری ناشی از تصمیم‌گیری الگوریتمی در سطح هیئت‌مدیره آماده نیست. پاسخگویی به‌عنوان رابطه‌ای میان عامل، هیئت‌مدیره و مجمع عمومی سهامداران تعریف می‌شود که در آن عامل موظف به توجیه عملکرد خود است و مجمع اختیار پرسش، ارزیابی عملکرد، تصمیم‌گیری درباره تداوم همکاری و اقدام قانونی در صورت لزوم را دارد. نخستین خلأ، نبود قاعده روشن برای تعیین مسئول در هنگام بروز خطای سامانه‌های هوش مصنوعی است، زیرا مفهوم مسئولیت دارای مؤلفه‌های متعددی است و صرفاً نسبت دادن آن نمی‌تواند مسائل اخلاقی را حل کند. این خلأ به‌ویژه در سطوح هوش افزایشی و تقویتی آشکار است. دومین خلأ، فقدان سازوکاری برای نظارت بر خود هیئت‌مدیره در مقام ناظر بر سامانه‌های هوش مصنوعی یا معمای ناظر بر ناظر است. سومین خلأ، ناهمگونی بازه‌های زمانی تصمیم‌گیری میان هیئت‌مدیره و سامانه‌های هوش مصنوعی است که امکان نظارت مؤثر و پاسخگویی به‌موقع را تضعیف می‌کند. در این پژوهش سه راهکار سیاستی شامل قاعده تفویض‌ناپذیری مسئولیت نهایی انسانی، ایجاد کمیته تخصصی هوش مصنوعی در هیئت‌مدیره با اختیارات حسابرسی مستمر و بازتعریف وظیفه امانتداری به‌گونه‌ای پیشنهاد می شود که سواد الگوریتمی و نظارت بر چرخه حیات مدل‌های هوش مصنوعی جزئی از مسئولیت قانونی مدیران باشد. این راهکارها در چارچوب چهار هدف پاسخگویی شامل انطباق، گزارش، نظارت و اجرا تحلیل شده و چارچوبی مفهومی برای تقویت پاسخگویی، ارتقای نظارت هیئت‌مدیره و انطباق حاکمیت شرکتی با الزامات تصمیم‌گیری مبتنی بر هوش مصنوعی ارائه می‌کنند.

مقدمه

طی یک دهه اخیر، هوش مصنوعی از جایگاه یک ابزار فنی صرف خارج شده و به پدیده‌ای ساختارشکن در نظام حاکمیت شرکتی تبدیل شده است. شواهد تجربی معتبر حاکی از آن است که هوش مصنوعی با شتاب فزاینده‌ای در حال تبدیل شدن به یک دستور کار اصلی و همیشگی در هیئت‌های مدیره است. بر اساس نظرسنجی انجام شده در سال ۲۰۲۵ که از دویست و یک عضو هیئت مدیره شرکت‌های سهامی عام انجام شده، ۶۲ درصد از مدیران گزارش داده‌اند که در حال حاضر زمان مشخصی را در دستور کار جلسات هیئت مدیره به بحث و بررسی هوش مصنوعی اختصاص می‌دهند، در حالی که این رقم در سال ۲۰۲۳ تنها ۲۸ درصد بوده است [۱]. این افزایش چشمگیر نشان می‌دهد که موضوع هوش مصنوعی با چه سرعتی در حال نفوذ به بالاترین سطوح تصمیم‌گیری شرکتی است. تحلیل مؤسسه مشاوره بین‌المللی نیز این یافته را تأیید می‌کند و هوش مصنوعی را به عنوان یک موضوع همیشگی در دستور کار جلسات هیئت مدیره و مدیران ارشد معرفی می‌نماید که اکنون در قلب مباحث مربوط به استراتژی، رشد و رهبری قرار گرفته است [۲] . با این وجود، شواهد نشان می‌دهد که این حضور گسترده در دستور کار به معنای آمادگی کامل و بلوغ دانش اعضا نیست. این شکاف دانشی در حالی رخ می‌دهد که کارکردهای هوش مصنوعی در بهینه‌سازی فرایندها و افزایش سرعت و کارایی، صرفاً به حوزه‌های فنی محدود نبوده و به‌تدریج در حال نفوذ به عرصه‌های راهبردی و مدیریتی نظیر حاکمیت شرکتی نیز می‌باشد [۳] . نظرسنجی مؤسسه فناوری ماساچوست از هزاران مدیر ارشد حاکی از آن است که ۷۹ درصد از پاسخ‌دهندگان معتقدند هیئت مدیره آنها دانش لازم برای درک هوش مصنوعی را ندارد [۴]. همچنین ۳۱ درصد از مدیران اعتراف کرده‌اند که هوش مصنوعی در دستور کار هیئت مدیره آنها قرار ندارد و ۶۶ درصد اذعان داشته‌اند که دانش محدود درباره هوش مصنوعی دارند. این شکاف دانشی در حالی رخ می‌دهد که سازمان همکاری و توسعه اقتصادی تأکید کرده است که نادیده گرفتن تحول هوش مصنوعی یا انتظار برای حل همه ناشناخته‌ها، دولت‌ها و شرکت‌ها را به پذیرنده منفعل فناوری تنزل می‌دهد و آنان را از نقش شکل‌دهنده فعال گزینه‌ها بازمی‌دارد و هزینه‌های سنگین و عقب‌ماندگی جبران‌ناپذیری در پی خواهد داشت [۵]. هیئت‌های مدیره امروزه با چالش دوگانه حضور هوش مصنوعی در دستور کار و فقدان دانش تخصصی مواجه هستند و به دنبال راه‌هایی برای پر کردن این شکاف از طریق آموزش، استخدام مشاور و حتی تغییر ترکیب خود می‌باشند [۴][۱] . بنابراین، علی‌رغم افزایش توجه هیئت‌های مدیره به موضوع هوش مصنوعی، فاصله عمیقی بین آگاهی از ضرورت این فناوری و درک عملی و تخصصی از ابعاد و ریسک‌های آن وجود دارد و این فاصله، نظام سنتی پاسخگویی در حاکمیت شرکتی را با چالش‌های بی‌سابقه‌ای مواجه ساخته است. پاسخگویی به عنوان یکی از ارکان اصلی حاکمیت شرکتی در عصر هوش مصنوعی، نیازمند تعریف دقیق است. رکن مذکور به عنوان یک رابطه پاسخ‌گویی میان عامل و مجمع تعریف می‌شود که در آن عامل موظف به توجیه و تشریح رفتار خود است و مجمع اختیار پرسش، قضاوت و اعمال پیامدها را دارد [۶]. با این حال، برخی پژوهشگران به این نکته اشاره کرده‌اند که اصطلاح مسئولیت خود یک اصطلاح بسته‌ای با مؤلفه‌های متعدد شامل پاسخگویی به عنوان توانایی ارائه دلیل، قابل سرزنش بودن به عنوان مستحق ملامت بودن و مسئولیت قانونی به عنوان قابلیت جبران خسارت است و صرفاً نسبت دادن مسئولیت نمی‌تواند مسائل اخلاقی ناشی از تصمیمات الگوریتمی را حل کند [۷]. از سوی دیگر، برخی پژوهشگران نشان داده‌اند که مشکل شکاف مسئولیت در هوش مصنوعی نه یک مسئله واحد، بلکه مجموعه‌ای از چهار شکاف به هم پیوسته شامل شکاف قابلیت سرزنش، شکاف پاسخگویی اخلاقی، شکاف پاسخگویی عمومی و شکاف مسئولیت فعال است [۸]. در بطور مثال در خصوص نظارت مستمر، از منظر حقوق استرالیا نیز تأکید شده است که مدیران دارای وظیفه امانی مستمر برای نظارت بر استفاده از سیستم‌های مبتنی بر هوش مصنوعی هستند و این نظارت نمی‌تواند به صورت مقطعی و دوره‌ای باشد [۹]. شواهد تجربی حاکی از آن است که کاربست هوش مصنوعی به طور معناداری انحراف راهبردی شرکت‌ها را از هنجارهای صنعت افزایش می‌دهد و این موضوع نشان می‌دهد که هوش مصنوعی به مدیران جرأت می‌بخشد تا استراتژی‌های شخصی‌شده و غیرمتعارف را دنبال کنند [۱۰]. با این حال، همین پژوهش نشان می‌دهد که در ابعاد ریسک مالی نظیر اهرم مالی و شدت سرمایه، شرکت‌ها به طرز شگفت‌آوری به میانگین صنعت نزدیک می‌شوند و رفتاری از خود بروز می‌دهند که از آن با عنوان «انقباض عقلانی» یاد شده است. این الگوی دوگانه حاکی از آن است که مدیران از هوش مصنوعی به گونه‌ای استفاده می‌کنند که هم می‌تواند نوآوری فعالانه را تسهیل کند و هم محافظه‌کاری عقلانی را توجیه نماید [۱۰]. این یافته اگرچه به روشنی تأثیر هوش مصنوعی بر محتوای تصمیمات راهبردی را مستند می‌کند، اما به پرسش بنیادین پاسخ نمی‌دهد که در فرآیند اتخاذ این تصمیمات، مسئولیت نهایی بر عهده کیست و نظام حاکمیت شرکتی چگونه باید با خطاهای احتمالی الگوریتمی مواجه شود.

پژوهش‌های مروری در حوزه حاکمیت هوش مصنوعی نشان می‌دهد که با وجود تعهدات گسترده به اصولی چون انصاف، شفافیت، پاسخگویی و نظارت انسانی، بسیاری از سازمان‌ها شواهد محدودی از ترجمه این اصول به کنترل‌های عملیاتی برای مقابله با آسیب‌های عینی هوش مصنوعی ارائه می‌دهند و این شکاف میان چهارچوب‌های مفهومی و پیاده‌سازی عملی، به‌ویژه در حوزه سوگیری الگوریتمی، شفافیت، ایمنی و ریسک‌های اجتماعی، چالشی جدی برای حاکمیت شرکتی محسوب می‌شود [۱۱].

در همین راستا و در بستری متفاوت، مطالعه‌ای بر روی شرکت‌های بورس فلسطین نشان داده است که هوش مصنوعی به طور انتخابی و گزینشی رابطه میان مکانیسم‌های سنتی حاکمیت شرکتی و عملکرد مالی را تعدیل می‌کند. یافته‌های این پژوهش حاکی از آن است که هوش مصنوعی تأثیر تنوع جنسیتی و سطح تحصیلات اعضای هیئت مدیره بر بازده دارایی‌ها را تقویت می‌نماید، اما در عین حال تأثیر مثبت تعداد جلسات هیئت مدیره بر عملکرد مالی را تضعیف کرده و حتی آن را معکوس ساخته است [۱۲]. این یافته از دو جهت حائز اهمیت است. نخست آنکه نشان می‌دهد هوش مصنوعی صرفاً یک متغیر تعدیلگر خنثی نیست، بلکه می‌تواند برخی از مکانیسم‌های نظارتی سنتی هیئت مدیره را کنار بزند یا بی‌اثر سازد. دوم آنکه این نتیجه ضمنی را به همراه دارد که اگر هوش مصنوعی بتواند نیاز به جلسات مکرر هیئت مدیره را کاهش دهد، در آن صورت پرسش بنیادین ناظر بر پاسخگویی خود فرآیندهای الگوریتمی به طور فزاینده‌ای برجسته می‌شود. به عبارت دیگر، همان فناوری که وعده افزایش کارایی و کاهش هزینه‌های نظارت را می‌دهد، می‌تواند ساختارهایی را که برای پاسخگویی طراحی شده‌اند، تضعیف کند [۱۲]. از سوی دیگر، برخی پژوهش‌ها به آینده دورتر نظر داشته و سناریوهایی برای مدیریت شرکتی در عصری ترسیم کرده‌اند که هوش مصنوعی به طور کامل جایگزین مدیران انسانی می‌شود. بر این اساس، سه سناریو برای آینده مسئولیت مدیران قابل تصور است: نخست آنکه خود سامانه‌های هوش مصنوعی به عنوان موجودیت‌هایی با شخصیت حقوقی محدود شناخته شوند و بتوان از آنها در دعاوی سهامداری شکایت کرد. دوم، حذف کامل مسئولیت شخصی و جایگزینی آن با مکانیسم‌های جبران خسارت از طریق بیمه یا صندوق‌های تضمین است. سوم، که محتمل‌ترین گزینه به شمار می‌رود، نظامی شبیه به مسئولیت محصول است که در آن تأمین‌کنندگان و توسعه‌دهندگان نرم‌افزارهای مدیریتی هوش مصنوعی به جای مدیران انسانی در قبال خطاهای الگوریتمی پاسخگو خواهند بود [۱۳]. اما آنچه در این تحلیل کم‌رنگ باقی می‌ماند، وضعیت میانی و کنونی است؛ وضعیتی که در آن انسان و هوش مصنوعی با یکدیگر هم‌زیستی دارند، تصمیمات به صورت مشترک اتخاذ می‌شوند، و مرز میان توصیه الگوریتمی و تصمیم انسانی چنان در هم تنیده است که نمی‌توان به سادگی یکی را از دیگری تفکیک کرد.

مکمل این یافته‌ها، چهارچوبی است که پنج سطح مختلف برای حکمرانی هوش مصنوعی ترسیم می‌کند و هوش کمکی، هوش افزایشی، هوش تقویتی، هوش خودگردان و هوش خودپویا را از یکدیگر متمایز می‌سازد. این چهارچوب ابزار تحلیلی سودمندی برای فهم وضعیت میانی مورد اشاره فراهم می‌آورد [۱۴]. بر اساس این چهارچوب، در سطح هوش کمکی که انسان هنوز تصمیم‌گیرنده نهایی است و هوش مصنوعی صرفاً پشتیبانی گزینشی ارائه می‌دهد، مسئله پاسخگویی چندان حاد نیست. در این سطح، مدیر انسانی همچنان می‌تواند توصیه الگوریتمی را بپذیرد یا رد کند و در هر دو صورت، مسئولیت تصمیم نهایی بر عهده او باقی می‌ماند. اما در سطح هوش افزایشی که سامانه‌های هوش مصنوعی توصیه‌های پیچیده‌تری ارائه می‌دهند و مدیر انسانی برای اتخاذ تصمیم نهایی به شدت به این توصیه‌ها وابسته است، و به ویژه در سطح هوش تقویتی که تصمیم‌گیری به شکلی متوازن و بعضاً غیرقابل تفکیک میان انسان و ماشین توزیع می‌شود، انتساب مسئولیت به طور چشمگیری دشوار می‌گردد. در چنین وضعیتی نمی‌توان به سادگی تعیین کرد که آیا خطا ناشی از قصور مدیر انسانی در ارزیابی توصیه الگوریتمی بوده، یا ناشی از نقص در خود الگوریتم [۱۴]. وی خود اذعان دارد که در این دو سطح، نظام‌های حقوقی و اخلاقی موجود با یک معمای حل‌نشده مواجه هستند و تا زمانی که شفافیت الگوریتم‌ها به سطحی نرسد که مدیر انسانی بتواند به درستی مبنای توصیه‌ها را ارزیابی کند، نمی‌توان از مدیران انسانی انتظار داشت در قبال تصمیماتی که بر مبنای خروجی الگوریتم‌های جعبه سیاه گرفته‌اند، به طور کامل پاسخگو باشند [۱۴]. در سطح کلان و تطبیقی نیز پژوهشی نشان داده است که نظام‌های حقوقی با الگوی ذی‌نفع‌گرا مانند کشورهای عضو اتحادیه اروپا، در مقایسه با نظام‌های سهامدارمحور مانند ایالات متحده، ظرفیت بیشتری برای یکپارچه‌سازی سازوکارهای پاسخگویی الگوریتمی دارند. این مزیت عمدتاً ناشی از وجود نهادهای نظارتی قدرتمند و الزامات شفافیت پیشینی در این نظام‌هاست [۱۵]. قانون هوش مصنوعی اتحادیه اروپا که در سال ۲۰۲۴ به تصویب رسید، سامانه‌های هوش مصنوعی پرخطر را موظف به رعایت الزامات سختگیرانه‌ای از جمله ایجاد نظام مدیریت ریسک، حاکمیت بر داده‌ها، مستندسازی فنی، نظارت انسانی، و دقت و امنیت سایبری کرده است [۱۵]. با این حال، این تحلیل تطبیقی عمدتاً بر سطح کلان حقوقی و تفاوت‌های نهادی متمرکز است و به پویایی‌های درون هیئت مدیره، روابط قدرت میان اعضا، و فرآیندهای روزمره تصمیم‌گیری که در آنها خطاهای الگوریتمی رخ می‌دهد، نمی‌پردازد. به عبارت دیگر، تحلیل یادشده هرچند در سطح قانونگذاری و تنظیم‌گری راهگشاست، اما در سطح خرد و اجرایی یعنی همان جایی که هیئت مدیره با خروجی یک سامانه هوش مصنوعی مواجه می‌شود و باید در لحظه تصمیم بگیرد، اطلاعات چندانی ارائه نمی‌دهد. [۱۶] در یک مطالعه فراترکیب بر روی ۳۶ پژوهش پیشین، چهارچوبی چهاربعدی برای تصمیم‌گیری استراتژیک داده محور ارائه داده‌اند که شامل شرایط، ویژگی‌ها، ابعاد و پیامدها می‌شود و به طور خاص بر تلفیق شهود و تحلیل و نیز تعادل میان نقش انسان و ماشین تأکید دارد. در این چهارچوب، یکی از ابعاد کلیدی تصمیم‌گیری تلفیق شهود و تحلیل معرفی شده است. بر این اساس، هرچه شرایط محیطی به سمت تلاطم و پیچیدگی بیشتر پیش رود، سهم شهود انسانی در تصمیم‌گیری افزایش می‌یابد و نقش هوش مصنوعی به عنوان ابزار تحلیل‌گر داده‌ها، هرچند ضروری، اما ناکافی است [۱۶]. این یافته از

آن جهت اهمیت دارد که نشان می‌دهد تصمیم‌گیری در شرایط ابهام و عدم قطعیت، ذاتاً مستعد خطاهایی است که نه به داده‌ها و الگوریتم‌ها، بلکه به قضاوت انسانی بازمی‌گردد. در چنین شرایطی، تفکیک سهم هر یک از این دو عامل در بروز خطا نه تنها دشوار، بلکه گاه ناممکن می‌شود. بر اساس این پیشینه، مقاله حاضر استدلال می‌کند که ادبیات موجود با وجود غنای توصیفی و تنوع روش‌شناختی، به طور سیستماتیک از کنار سه خلأ ساختاری پاسخگویی عبور کرده است. نخستین خلأ، نبود قاعده روشن برای تعیین مسئول در خطای سامانه‌های هوش مصنوعی است. دومین خلأ، فقدان سازوکاری برای نظارت بر خود هیئت مدیره در مقام ناظر بر سامانه‌های هوش مصنوعی می‌باشد. سومین خلأ، ناهمگونی بازه‌های زمانی تصمیم‌گیری میان هیئت مدیره و سامانه‌های هوش مصنوعی است. در بخش‌های بعدی، هر یک از این خلأها به تفکیک تشریح و تحلیل می‌شوند و در پایان، راهکارهایی برای پر کردن آنها پیشنهاد می‌گردد.

خلأ پاسخگویی در انتساب مسئولیت خطای سامانه‌های هوش مصنوعی

نخستین و بنیادی‌ترین خلأ پاسخگویی که ادبیات موجود با وجود گستردگی، به طور نظام‌مند از کنار آن عبور کرده است، به نبود قاعده روشن برای تعیین مسئول در هنگام بروز خطا از سوی سامانه‌های هوش مصنوعی بازمی‌گردد. هنگامی که یک سامانه هوش مصنوعی تصمیمی را پیشنهاد می‌دهد یا به صورت خودگردان اجرا می‌کند و هیئت مدیره آن تصمیم را تأیید یا نادیده می‌گیرد، در نهایت خسارتی متوجه شرکت یا ذی‌نفعان آن می‌شود، نظام حقوقی و اخلاقی موجود قادر نیست به روشنی تعیین کند که مسئولیت نهایی بر عهده چه نهادی یا چه شخصی است و هر یک از بازیگران درگیر در فرآیند تصمیم‌گیری می‌توانند مسئولیت را به دیگری نسبت دهند و عملاً از پاسخگویی شانه خالی کنند. این ابهام نه در قوانین تجاری کشورهای مختلف و نه در رویه‌های رایج حاکمیت شرکتی پاسخ روشنی ندارد و ریشه در سرشت دوگانه هوش مصنوعی دارد که از یک سو یک ابزار محاسباتی و از سوی دیگر یک عامل تصمیم‌گیر با درجات مختلفی از خودگردانی است. در واقع نظام حقوقی سنتی بر پایه این تمایز بنیادین طراحی شده که ابزارها فاقد مسئولیت هستند و مسئولیت بر عهده انسان بهره‌بردار است، اما هوش مصنوعی این تمایز را بر هم می‌زند و موجودیتی را وارد عرصه تصمیم‌گیری می‌کند که نه کاملاً ابزار است و نه کاملاً عامل مستقل [۱۷][۱۸] . این مسئله از آن جهت پیچیده‌تر می‌شود که «مسئولیت» خود مجموعه ای از مؤلفه‌های مختلف است. برخی پژوهشگران استدلال می‌کنند که استفاده از اصطلاح «مسئولیت» به عنوان ابزاری تحلیلی برای مسائل اخلاقی هوش مصنوعی گمراه‌کننده است، زیرا این اصطلاح مؤلفه‌های متعددی از جمله پاسخگویی به عنوان توانایی ارائه دلیل، قابل سرزنش بودن به عنوان مستحق ملامت بودن و مسئولیت قانونی به عنوان قابلیت جبران خسارت را در خود جمع کرده است که هر یک الزامات متفاوتی برای عامل ایجاد می‌کنند [۷]. به عبارت دیگر، صرف بیان اینکه چه کسی مسئول است بدون تفکیک این مؤلفه‌ها، تحلیل اخلاقی را سطحی و ناقص می‌سازد. تحقیقاتی که تأثیر هوش مصنوعی بر حاکمیت شرکتی بررسی نموده اند اکثرا به شماری از مزایای کاربست هوش مصنوعی در هیئت مدیره از جمله افزایش کارایی عملیاتی، کاهش ریسک‌های ناشی از خطای انسانی و بهبود چشمگیر قابلیت‌های پیش‌بینی اشاره داشته اند [۱۹] اما در جای خود به این پرسش بنیادین نمی‌پردازند که اگر هوش مصنوعی مرتکب خطایی شود که یک مدیر انسانی با همان داده‌ها و در همان شرایط مرتکب نمی‌شد، چه کسی باید پاسخگو باشد و نظام حاکمیت شرکتی چگونه باید با خطاهایی مواجه شود که نه از قصور مدیر انسانی، بلکه از نقص در طراحی الگوریتم یا کیفیت داده‌های آموزشی ناشی می‌شوند. [۲۰] در مقاله خود با عنوان تصمیم‌گیری مبتنی بر هوش مصنوعی، سه چالش اصلی را برای حاکمیت شرکتی در عصر الگوریتم‌ها شناسایی کرده است: سوگیری الگوریتمی، فقدان شفافیت در فرآیند تصمیم‌گیری و نبود پاسخگویی روشن در قبال تصمیمات الگوریتمی. او به درستی تأکید می‌کند که در نظام سنتی حاکمیت شرکتی، پاسخگویی بر عهده مدیران انسانی است و سهامداران از طریق دعاوی نیابتی می‌توانند مدیران خاطی را تحت پیگرد قرار دهند، اما هنگامی که سامانه‌های هوش مصنوعی در فرآیند تصمیم‌گیری دخالت می‌کنند، این الگوی روشن پاسخگویی با ابهام روبرو می‌شود [۲۰]. راهکاری که [۲۰] پیشنهاد می‌کند، ایجاد چهارچوب‌های روشن برای پاسخگویی هوش مصنوعی و تعیین افرادی است که مسئولیت نظارت بر سامانه‌های الگوریتمی را بر عهده دارند، با این حال او به این پرسش اساسی پاسخ نمی‌دهد که اگر خود آن افراد ناظر، به دلیل پیچیدگی فنی الگوریتم‌ها یا ماهیت جعبه سیاه مدل‌های یادگیری عمیق، نتوانند خطای الگوریتمی را تشخیص دهند یا درک کنند، در آن صورت سازوکار جایگزین پاسخگویی چیست [۲۰][۱۴]. [۱۳] در تحلیل خود از سه سناریوی آینده مسئولیت، با دقت بیشتری به وضعیت میانی و کنونی پرداخته است که در آن انسان و هوش مصنوعی با یکدیگر هم‌زیستی دارند و مدیران انسانی با پرسش‌های دشواری درباره میزان نظارت بر الگوریتم‌ها و میزان واگذاری وظایف به ماشین‌ها بدون آنکه خود را در معرض مسئولیت شخصی قرار دهند، مواجه خواهند شد. او استدلال می‌کند که قواعد سنتی که به مدیران اجازه می‌دهد به گزارش‌های کارشناسان، حسابرسان، و مشاوران حقوقی تکیه کنند و در صورت حسن نیت از مسئولیت معاف شوند، به طور کامل برای تکیه بر توصیه‌های سامانه‌های هوش مصنوعی قابل تعمیم نیست، زیرا این سامانه‌ها بر خلاف کارشناسان انسانی، فاقد مسئولیت حرفه‌ای در قبال توصیه‌های خود هستند و در صورت خطا نمی‌توان آنها را به همان سادگی که یک حسابرس یا

مشاور را تحت پیگرد قرار داد، پاسخگو ساخت (همان منبع قبلی). این نکته از آن جهت حائز اهمیت است که نشان می‌دهد حتی در وضعیت کنونی که هوش مصنوعی هنوز جایگزین کامل مدیران نشده است، خلأ پاسخگویی به صورت بالقوه وجود دارد و نظام حقوقی برای پر کردن آن یا باید مسئولیت حرفه‌ای جدیدی برای توسعه‌دهندگان الگوریتم‌ها تعریف کند، یا قواعد سنتی تکیه بر نظرات کارشناسان را به گونه‌ای اصلاح نماید که استفاده از سامانه‌های هوش مصنوعی را نیز در بر گیرد. [۸] در تحلیل چهارگانه خود، این وضعیت را شکاف قابلیت سرزنش می‌نامند و نشان می‌دهند که این شکاف نه تنها در سیستم‌های یادگیرنده، بلکه در سیستم‌های غیریادگیرنده با زنجیره پیچیده عوامل انسانی نیز رخ می‌دهد. به باور آنان، خطر این شکاف در آن است که هرچه عوامل درگیر در طراحی، نظارت و بهره‌برداری از سیستم بتوانند به طور مشروع از سرزنش اجتناب کنند، انگیزه کمتری برای پیشگیری از رفتارهای نادرست خواهند داشت و قربانیان نیز کمتر احتمال دارد که جبران خسارت دریافت کنند (همان منبع قبلی). از سوی دیگر، این شکاف همچنین می‌تواند توانایی افراد را برای درک و ابراز قضاوت اخلاقی در مورد خطاها و حوادث تضعیف کند و احساس درماندگی و شکاکیت اخلاقی را نسبت به امکان فهم و اصلاح خطاها دامن بزند. [۲۱] با بسط نظریه نمایندگی به عصر هوش مصنوعی، مفهوم عامل مصنوعی را معرفی کرده و نشان می‌دهند که یک سامانه هوش مصنوعی با درجه بالایی از خودگردانی و خودتعیینی می‌تواند به عنوان عاملی در نظر گرفته شود که برای همسوسازی منافع آن با شرکت، نیازمند بازنگری اساسی در مکانیسم‌های نظارت و انگیزش سنتی است. آنان استدلال می‌کنند که نظارت بر عامل مصنوعی باید از مراحل اولیه تکامل هوش مصنوعی آغاز شود و در مراحل بالاتر، مکانیسم‌های انگیزشی مبتنی بر منابع محاسباتی و ذخیره‌سازی جایگزین انگیزش‌های مالی رایج برای مدیران انسانی گردد (همان منبع قبلی). در جمع‌بندی این بخش می‌توان گفت که نخستین خلأ پاسخگویی ریشه در سرشت دوگانه هوش مصنوعی دارد و پر کردن آن نیازمند بازنگری در مبانی فلسفی و حقوقی مسئولیت در عصر الگوریتم‌ها است. آنچه از این تحلیل به دست می‌آید این است که مسئله اصلی نه صرفاً انتساب حقوقی مسئولیت به یک عامل خاص، بلکه پرسش بنیادین آن است که کدام عامل از توانایی واقعی برای پاسخگویی و ایفای کامل مؤلفه‌های مسئولیت، مانند کنترل، آگاهی، اختیار و قابلیت جبران خسارت، برخوردار است چراکه صرف انتساب مسئولیت بدون چنین توانایی، پاسخگویی را به امری شکلی و ناکارآمد تبدیل می‌کند.

خلأ پاسخگویی در معمای ناظر بر ناظر در نظام راهبری شرکتی

دومین خلأ پاسخگویی که ادبیات موجود به گونه‌ای نظام‌مند از کنار آن عبور کرده است، به فقدان سازوکاری برای نظارت بر خود هیئت مدیره در مقام ناظر بر سامانه‌های هوش مصنوعی بازمی‌گردد که می‌توان آن را معمای ناظر بر ناظر نامید، زیرا در نظام سنتی حاکمیت شرکتی، هیئت مدیره به عنوان عالی‌ترین نهاد نظارتی تعریف می‌شود که بر عملکرد مدیران اجرایی و یکپارچگی فرآیندهای سازمانی اشراف دارد، اما هنگامی که همان هیئت مدیره به سامانه‌های هوش مصنوعی برای نظارت بر ریسک‌ها یا ارزیابی صحت گزارش‌های عملکردی متکی می‌شود، دیگر نهادی بیرون از این چرخه وجود ندارد که بتواند بر خود هیئت مدیره در مقام کاربر و ناظر همزمان هوش مصنوعی نظارت کند. [۲۲] در مطالعه‌ای بر روی ۴۴۴ شرکت بورسی تایلند نشان داده است که نه تنها اثربخشی هیئت مدیره، بلکه اثربخشی کمیته حسابرسی نیز به طور معناداری با کیفیت نظارت بر ریسک مرتبط است و ویژگی‌هایی مانند اندازه، استقلال، تخصص و تعداد جلسات در این اثربخشی نقش حیاتی ایفا می‌کنند. این یافته نشان می‌دهد که اثربخشی نظارت بر ریسک در شرکت‌ها به شدت به ویژگی‌های ساختاری و کارکردی خود هیئت مدیره و کمیته‌های آن وابسته است و هرگونه جایگزینی یا تضعیف این ویژگی‌ها توسط سامانه‌های هوش مصنوعی می‌تواند خلأ نظارتی ایجاد کند. [۲۳] در مطالعه‌ای دیگر نشان داده‌اند که استقلال و تخصص اعضای هیئت مدیره به طور غیرمستقیم و از طریق کیفیت مدیریت ریسک راهبردی بر عملکرد شرکت تأثیر می‌گذارند، اما اندازه هیئت مدیره نه تنها تأثیر مستقیمی بر عملکرد ندارد، بلکه حتی با مدیریت ریسک راهبردی نیز رابطه معناداری نشان نمی‌دهد. این یافته از آن جهت قابل توجه است که نشان می‌دهد اثربخشی هیئت مدیره در مدیریت ریسک و نظارت بر آن، بیشتر به قابلیت و تخصص وابسته است تا ساختار و اندازه آن. به عبارت دیگر، آنچه هیئت مدیره را در نظارت بر ریسک موفق می‌کند، صرفاً تعداد اعضا یا وجود کمیته‌های متعدد نیست، بلکه دانش، تخصص، و استقلال قضاوت اعضا است. این نکته در عصر هوش مصنوعی از دو جهت اهمیت پیدا می‌کند، از یک سو، اگر هیئت مدیره نتواند تخصص فنی لازم برای ارزیابی خروجی سامانه‌های هوش مصنوعی را در خود ایجاد کند، در آن صورت نظارت آن بر این سامانه‌ها صرفاً صوری و نمادین خواهد بود و عملاً نمی‌تواند از بروز خطاهای الگوریتمی جلوگیری کند. از سوی دیگر، اگر هوش مصنوعی بتواند به گونه‌ای طراحی شود که کاستی‌های ناشی از کمبود تخصص یا استقلال اعضای هیئت مدیره را جبران کند، در آن صورت ممکن است این سؤال پیش آید که اساساً به هیئت مدیره انسانی با ساختارهای سنتی خود نیاز است یا خیر و اگر پاسخ مثبت است، نقش نظارتی هیئت مدیره در قبال سامانه‌ای که از خود او توانمندتر است، چیست. [۸] در تحلیل چهارگانه خود، دو نوع شکاف پاسخگویی

را شناسایی کرده‌اند که مستقیماً با معمای ناظر بر ناظر مرتبط هستند. شکاف اول، شکاف پاسخگویی اخلاقی است که عبارت است از دشواری افراد در درک، تبیین و بازتاب منطق حاکم بر رفتار خود و دیگران به دلیل پیچیدگی و ابهام فرآیندهای تصمیم‌گیری الگوریتمی است. شکاف دوم، شکاف پاسخگویی عمومی است که ناظر بر دشواری شهروندان در دریافت توضیح درباره تصمیمات اتخاذشده توسط نهادهای عمومی و دشواری خود نهادها در پاسخگویی به این درخواست‌ها به دلیل واگذاری اختیارات به کارشناسان فنی و شرکت‌های خصوصی است. این دو شکاف نشان می‌دهند که مشکل نظارت بر هیئت مدیره نه تنها یک مسئله ساختاری و حقوقی، بلکه یک مسئله شناختی و ارتباطی نیز می باشد و زمانی که تصمیمات توسط هوش مصنوعی اتخاذ می‌شوند، شکاف پاسخگویی اخلاقی، ناتوانی درک منطق ماشین، و شکاف پاسخگویی عمومی، ناتوانی نهادها در توضیح تصمیمات الگوریتمی، به مانعی جدی برای اعمال نظارت مؤثر تبدیل می‌شوند. پژوهش‌های حوزه مطالعات علم و فناوری نشان داده است که نظام‌های فنی- اجتماعی هنگام ورود به بافت اجتماعی پیرامون خود، با پنج دام جدی مواجه می‌شوند که شامل؛ دام انتزاع نادرست، دام حلقه مفقوده، دام پیامدهای ناخواسته، دام انتقال‌پذیری و دام چهارچوب‌بندی شده و این دام‌ها موجب می‌شوند که مداخلات فنی برای ایجاد انصاف و پاسخگویی، نه تنها ناکارآمد، بلکه در مواردی خطرناک و گمراه‌کننده باشند [۲۴]. در مطالعه‌ای بر روی شرکت‌های حاضر در بورس فلسطین مشاهده شده است که هوش مصنوعی تأثیر مثبت تعداد جلسات هیئت مدیره بر عملکرد مالی را تضعیف می‌کند. بر اساس یافته‌های این مطالعه [۱۲] می‌توان استنباط کرد که هوش مصنوعی بالقوه قادر است برخی از مکانیسم‌های نظارتی سنتی را که مبتنی بر حضور فیزیکی و تعامل مستقیم اعضای هیئت مدیره هستند، کنار بزند یا بی‌اثر سازد. با این حال، با توجه به محدودیت‌های این پژوهش، از جمله کوچک بودن بازار بورس فلسطین و عدم عمق لازم در بازار آن، نمی‌توان نتایج آن را با قاطعیت به سایر نظام‌های راهبری شرکتی تعمیم داد. با وجود این احتیاط روش‌شناختی، اگر هوش مصنوعی بتواند اطلاعات مورد نیاز برای تصمیم‌گیری را بدون نیاز به جلسات مکرر در اختیار اعضا قرار دهد و حتی توصیه‌هایی ارائه کند که نیاز به بحث و تبادل نظر جمعی را کاهش دهد، در آن صورت پرسش بنیادین این است که چه نهادی بر خود فرآیند تصمیم‌گیری الگوریتمی نظارت خواهد کرد و آیا می‌توان به مدیران اجرایی اعتماد کرد که بدون نظارت هیئت مدیره، از سامانه‌های هوش مصنوعی به درستی استفاده کنند. [۹] از منظر حقوق استرالیا به این مسئله اشاره کرده‌اند که مدیران دارای وظیفه امانی مستمر برای نظارت بر استفاده از سیستم‌های مبتنی بر هوش مصنوعی هستند و این نظارت نمی‌تواند به صورت مقطعی و دوره‌ای باشد. آنان همچنین بر نقش مدیران مستقل به عنوان یک مکانیسم نظارتی مؤکد تأکید دارند و استدلال می‌کنند که مدیران مستقل به دلیل نداشتن وابستگی‌های سازمانی، می‌توانند نقش مؤثرتری در نظارت بر استفاده از هوش مصنوعی در هیئت مدیره ایفا کنند (همان منبع قبلی). در این زمینه، تحلیل‌گران حقوقی بر این نکته تأکید دارند که پایداری مشروعیت نظام حاکمیت شرکتی در عصر الگوریتم‌ها، مستلزم تأکید مجدد بر قضاوت انسانی، نظارت مستمر در سطح هیئت مدیره و چهارچوب‌های حاکمیتی تطبیق‌پذیری است که بتوانند قدرت فناوری را با پاسخگویی حقوقی و انتظارات اجتماعی آشتی دهند [۲۵].

در ادبیات موجود در زمینه حسابرسی هوش مصنوعی می‌توان تمایز مهمی میان حسابرسی فنی و حسابرسی فرآیندی قائل شد. به نظر می‌رسد حسابرسی مستمر و مستقل سامانه‌های هوش مصنوعی تا حدی قادر است به عنوان یک لایه نظارتی بر هیئت مدیره عمل کند [۲۶]. با این حال تحقق چنین حسابرسی مستقلی مستلزم وجود استانداردهای روشن، دسترسی کامل به داده‌ها و نیز استقلال حسابرس از نهاد مورد بررسی است. متأسفانه در حال حاضر در بسیاری از حوزه‌های قضایی این پیش‌شرط‌ها فراهم نیست (همان منبع قبلی). از سوی دیگر، حسابرسی مستمر هوش مصنوعی را می‌توان نوعی سیستم پشتیبانی الکترونیک در زمان نزدیک به واقعی تعریف کرد [۲۷]. به نظر می‌رسد این نوع حسابرسی با ارائه گزارش‌های به‌روز و خودکار از عملکرد سامانه‌های هوش مصنوعی می‌تواند خلأ ناشی از نبود نظارت مستقل بر هیئت مدیره را تا حدودی جبران کند. با وجود این، اجرای موفق چنین سیستمی نیازمند معیارهای از پیش تعیین‌شده، خودکارسازی در سطح بالا و الزامات قانونی روشن است. بررسی وضعیت موجود نشان می‌دهد که این الزامات در حال حاضر فقط در تعداد محدودی از چارچوب‌های نظارتی دیده می‌شود (همان منبع قبلی). در یک جمع‌بندی موقت می‌توان گفت که دومین خلأ پاسخگویی ریشه در ساختار دوگانه و ذاتی نظام حاکمیت شرکتی دارد و حل آن نیازمند ایجاد لایه‌های جدید نظارت و پاسخگویی است که در نظام سنتی حاکمیت شرکتی پیش‌بینی نشده بودند. آنچه از این تحلیل به دست می‌آید این است که معمای ناظر بر ناظر نه تنها یک مسئله ساختاری، بلکه یک مسئله شناختی، ارتباطی ،شامل تبادل معنا، بازخورد، توضیح‌خواهی و پاسخگویی، و حقوقی است که برای حل آن باید به طور همزمان به جنبه‌های فنی، سازمانی و قانونی توجه کرد. چهارچوب دوبعدی یکپارچه‌سازی هوش مصنوعی در هیئت مدیره، با تفکیک پنج سطح خودمختاری، از هوش کمکی تا هوش خودپویا، و سه محل یکپارچه‌سازی، سطح مدیران فردی، سطح کمیته‌ها و سطح کل هیئت مدیره، نشان می‌دهد که کاربست هوش مصنوعی با خودمختاری بالا در سطح کل هیئت مدیره، برخلاف کاربست آن در سطح مدیران فردی، می‌تواند عملاً امکان هرگونه نظارت مؤثر انسانی را از میان بردارد، زیرا در چنین وضعیتی ناظری بیرون از چرخه وجود نخواهد داشت که بر خود فرایند تصمیم‌گیری الگوریتمی نظارت کند [۲۸].

خلأ پاسخگویی در ناهمگونی زمانی نظارت بر تصمیمات الگوریتمی

سومین خلأ پاسخگویی که ادبیات موجود تقریباً به طور کامل از کنار آن عبور کرده است به ناهمگونی بازه‌های زمانی تصمیم‌گیری میان هیئت مدیره و سامانه‌های هوش مصنوعی بازمی‌گردد، ناهمگونی که به دلیل ماهیت جلسات دوره‌ای و غیرمستمر هیئت مدیره در یک سو و عملکرد لحظه‌ای و مستمر سامانه‌های الگوریتمی در سوی دیگر ایجاد می‌شود و امکان نظارت مؤثر و پاسخگویی به موقع را از میان می‌برد. هیئت مدیره در ساختار سنتی حاکمیت شرکتی، نهادی است که به صورت دوره‌ای تشکیل جلسه می‌دهد و تصمیمات راهبردی را در مقیاس هفته‌ها و ماه‌ها اتخاذ و بررسی می‌کند، در حالی که سامانه‌های هوش مصنوعی قادر هستند در سطح میلی‌ثانیه تصمیم بگیرند، توصیه ارائه دهند، یا حتی به صورت خودگردان عمل کنند [۱۴][۲۰]. این ناهمگونی زمانی دو پیامد جدی برای پاسخگویی ایجاد می‌کند. پیامد نخست آنکه هنگامی که یک سامانه هوش مصنوعی در فاصله بین دو جلسه هیئت مدیره تصمیمی می‌گیرد یا توصیه‌ای ارائه می‌دهد که منجر به خسارت می‌شود، هیئت مدیره تا جلسه بعدی قادر به شناسایی، ارزیابی و پاسخگویی به آن خطا نخواهد بود و در این فاصله زمانی، خسارت ممکن است به مراتب بیشتر از آن شود که بتوان آن را جبران کرد. پیامد دوم آنکه حتی اگر هیئت مدیره در جلسه بعدی از خطا مطلع شود، بازسازی زنجیره علت و معلولی که به آن خطا منجر شده است به دلیل حجم بالای داده‌ها و سرعت بالای تصمیمات الگوریتمی عملاً غیرممکن یا بسیار دشوار خواهد بود [۱۳]. [۱۴] در چهارچوب سناریوهای پنج‌گانه خود، به این ناهمگونی زمانی اشاره کرده و نشان داده است که در سطوح بالاتر هوش مصنوعی (هوش خودگردان و هوش خودپویا)، مشکل عدم تطابق زمانی به حدی حاد می‌شود که عملاً امکان هرگونه نظارت مؤثر انسانی را از میان می‌برد [۱۴]. در سطح هوش خودگردان که سامانه‌های هوش مصنوعی می‌توانند بدون نیاز به تأیید انسانی و در چارچوبی از پیش تعیین‌شده، تصمیمات مستقلی اتخاذ کنند، هیئت مدیره عملاً قادر نیست در زمان واقعی بر عملکرد این سامانه‌ها نظارت داشته باشد و تنها می‌تواند پس از وقوع خسارت به بررسی علل آن بپردازد که در آن هنگام نیز، دیگر دیر شده است و خسارت جبران‌ناپذیر خواهد بود [۱۴]. وی به درستی خاطرنشان می‌کند که در چنین شرایطی، نقش هیئت مدیره از یک ناظر فعال و پیشگیرانه به نهادی تنزل می‌یابد که صرفاً پس از وقوع خسارت و با فاصله زمانی قابل توجه وارد فرآیند بررسی می‌شود، در حالی که به دلیل فقدان دسترسی به داده‌های لحظه‌ای و ناتوانی در بازسازی زنجیره تصمیم‌گیری الگوریتمی اساساً نمی‌تواند مسئولیت واقعی را به عامل خاصی منتسب کند. این به معنای آن است که خلأ پاسخگویی زمانی، در عمل، نظام نظارتی سنتی را به کلی بی‌اثر می‌سازد. در مقابل انواع دیگر مسئولیت که صرفاً ناظر بر رویدادهای گذشته هستند، نوع دیگری از مسئولیت تحت عنوان مسئولیت فعال وجود دارد که آینده‌نگر است. این نوع مسئولیت به اهداف، ارزش‌ها و هنجارهایی مربوط می‌شود که متخصصان و مدیران موظف به ترویج و رعایت آنها هستند و نیز به پیامدهایی که باید از آنها جلوگیری کنند. به اعتقاد [۸]، معرفی هوش مصنوعی می‌تواند دو دسته مسئله مرتبط با مسئولیت فعال ایجاد کند. دسته نخست مسئله آگاه نبودن مهندسان، مدیران و سایر عوامل درگیر در توسعه و استفاده از فناوری نسبت به وظایف اخلاقی و اجتماعی خود است. دسته دوم مسئله ناتوانی یا بی‌انگیزگی آنان در ایفای تعهداتی است که از آنها آگاه هستند. این شکاف مسئولیت فعال ارتباط مستقیمی با ناهمگونی زمانی دارد، زیرا سرعت بالای تصمیمات الگوریتمی و ناتوانی هیئت مدیره در همگامی با این سرعت، مانع از ایفای مسئولیت فعال مدیران می‌شود. برای صورتبندی دقیق‌تر این چالش، می‌توان به تحلیل [۶] اشاره کرد که بین دو نوع پاسخگویی یعنی پاسخگویی پیش‌فعال و پاسخگویی واکنش‌گرا تمایز قائل شده‌اند. پاسخگویی پیش‌فعال به تدوین استانداردها و الزامات پیش از وقوع هر رویدادی مربوط می‌شود و به پرسش یک سامانه پاسخگو باید چگونه باشد، پاسخ می‌دهد، در حالی که پاسخگویی واکنش‌گرا به اجرای اقدامات پس از وقوع رویداد مربوط می‌شود و به پرسش چگونه می‌توان سامانه را پاسخگو ساخت، پاسخ می‌دهد. این تمایز برای تحلیل ناهمگونی زمانی سودمند بوده زیرا نشان می‌دهد که اتکای صرف به پاسخگویی واکنش‌گرا در مواجهه با سامانه‌های پرسرعت و خودگردان کافی نیست و باید با تقویت پاسخگویی پیش‌فعال، الزامات نظارتی را پیشاپیش در طراحی و استقرار سامانه‌ها لحاظ کرد [۶]. این ناهمگونی زمانی را می‌توان از زاویه دیگری نیز بررسی کرد. در مطالعه فراترکیبی که [۱۶] انجام داده‌اند، یکی از ویژگی‌های اصلی تصمیم‌گیری دادهمحور سرعت معرفی شده است. بر اساس یافته‌های آنان، هوش مصنوعی توانایی واکنش سریع به تغییرات بازار، افزایش نمایی سرعت پردازش داده‌ها و امکان تجزیه و تحلیل مستمر و لحظه‌ای را برای سازمان‌ها فراهم می‌آورد [۱۶]. این ویژگی هرچند از نظر کارایی عملیاتی و مزیت رقابتی بسیار ارزشمند است، اما از منظر پاسخگویی و نظارت یک چالش جدی محسوب می‌شود. دلیل آن است که اگر هیئت مدیره نتواند با همین سرعت به تصمیمات الگوریتمی واکنش نشان دهد و آنها را ارزیابی کند، در عمل کنترل خود را بر فرآیند تصمیم‌گیری از دست خواهد داد [۱۶]. به عبارت دیگر، همان سرعت بالا و قابلیت پاسخ‌دهی لحظه‌ای که هوش مصنوعی را برای تصمیم‌گیری راهبردی جذاب می‌کند، دقیقاً همان ویژگی است که نظام نظارتی سنتی را با بحران مواجه می‌سازد.

در این زمینه، تحلیل‌گران حقوق اداری بر این نکته تأکید دارند که دشواری درک شهودی نتایج الگوریتم‌های یادگیری ماشین، بهای نسبتاً کوچکی است که در برابر مزایای دقت و کارایی بیشتر تصمیمات در نظام اداری پرداخت می‌شود و نباید این دشواری، دولت‌ها و سازمان‌ها را از بهره‌مندی از پیشرفت‌های آماری و محاسباتی بازدارد،

بلکه بایستی به‌جای کناره‌گیری از این فناوری، به‌دنبال ایجاد سازوکارهای نظارتی متناسب با سرعت و پیچیدگی آن بود [۲۹]. در این زمینه، [۲۷] در مقاله خود با عنوان حسابرسی مستمر هوش مصنوعی، راهکار اصلی برای حل مشکل ناهمگونی زمانی را در پیاده‌سازی سیستم‌های حسابرسی مستمر می‌دانند. چنین سیستم‌هایی به صورت خودکار و در زمان نزدیک به واقعی، عملکرد سامانه‌های هوش مصنوعی را پایش کرده و انحرافات از معیارهای از پیش تعیین‌شده را گزارش می‌دهند. به اعتقاد آنان، حسابرسی مستمر نیازمند شش معیار اصلی است که عبارتند از زمان واقعی، تکرارپذیری، جمع‌آوری داده، خودکارسازی، معیارهای از پیش تعیین‌شده و الزام قانونی. بررسی آنان از ۳۵ چارچوب و ابزار موجود نشان می‌دهد که فقط تعداد محدودی از آنها مستقیماً برای حسابرسی مستمر هوش مصنوعی مناسب هستند و بیشتر این ابزارها در یک بخش یا حوزه مشکل خاص محدود شده‌اند. در همین زمینه، [۲۶] در تحلیل خود از انواع رویکردهای حسابرسی هوش مصنوعی، تمایزی میان رویکردهای حقوقی، اخلاقی و فنی معرفی کرده و تأکید دارد که حسابرسی صرفاً زمانی می‌تواند به خلأ پاسخگویی پایان دهد که مستقل، شفاف و مبتنی بر معیارهای روشن باشد. این تأکید بر استقلال و شفافیت، در واقع به ناتوانی نظام‌های نظارتی سنتی در همگام شدن با سرعت تصمیمات الگوریتمی اشاره دارد، زیرا بدون وجود حسابرس مستقل و معیارهای شفاف، هیئت مدیره نمی‌تواند به موقع و مؤثر بر فرآیندهای الگوریتمی نظارت کند.

در برخی از مدل‌های حکمرانی که برای نظام‌های هوش مصنوعی تدوین شده‌اند، میان دو مقوله اختیار تصمیم‌گیری و مالکیت پیامدهای ناشی از آن تصمیم تمایز قائل می‌شوند. بر این اساس، در سیستم‌های خودگردان هوش مصنوعی، مسئولیت به صورت شبکه‌ای و در طول زنجیره‌ای از عوامل فنی و انسانی توزیع می‌شود. از این منظر، مدل انتساب مسئولیت مبتنی بر شبکه و مسیرهای مسئولیت می‌تواند به تعیین پاسخگوی نهایی در مواجهه با خطاهای الگوریتمی کمک کند [۳۰]. همچنین یافته‌های یک مطالعه کیفی نشان می‌دهد که ۶۰ درصد از مدیران ارشد شرکت‌کننده در پژوهش، پیچیدگی یکپارچه‌سازی را به عنوان یکی از موانع اصلی پیاده‌سازی هوش مصنوعی ذکر کرده‌اند. این پیچیدگی تا حد زیادی به ناهماهنگی میان سرعت تصمیمات الگوریتمی و سرعت فرآیندهای تصمیم‌گیری انسانی بازمی‌گردد [۳۱]. برای تکمیل بحث درباره راهکارهای نظارتی، می‌توان به تحلیل [۲۶] اشاره کرد که انواع رویکردهای حسابرسی هوش مصنوعی را در سه دسته رویکردهای حقوقی، اخلاقی و فنی طبقه‌بندی کرده است. به اعتقاد وی، حسابرسی هوش مصنوعی صرفاً زمانی می‌تواند به خلأ پاسخگویی پایان دهد که دارای سه ویژگی اساسی باشد. این ویژگی‌ها عبارت از استقلال حسابرس از نهاد مورد بررسی، شفافیت فرآیند حسابرسی و مبتنی بودن بر معیارهای روشن و از پیش تعیین‌شده می باشند. این شروط در واقع همان معیارهایی هستند که [۲۷] نیز بر آنها تأکید داشتند، به ویژه الزام به خودکارسازی و وجود معیارهای از پیش تعیین‌شده در سیستم‌های حسابرسی مستمر.در یک جمع‌بندی از این بخش می‌توان گفت که سومین خلأ پاسخگویی، یعنی ناهمگونی زمانی میان بازه‌های تصمیم‌گیری انسان و ماشین، شاید از همه حادتر و دشوارتر باشد، زیرا نه یک نقص در طراحی نظام نظارتی، بلکه یک ویژگی ذاتی و اجتناب‌ناپذیر در تعامل انسان با سامانه‌های سریع و خودگردان است. آنچه از این تحلیل به دست می‌آید این است که حل این خلأ نیازمند گذار از پاسخگویی واکنش‌گرا به پاسخگویی پیش‌فعال و طراحی سیستم‌های نظارتی مستمر و خودکار است که بتوانند با سرعت سامانه‌های الگوریتمی همگام شوند.

راهکارهای پیشنهادی برای پر کردن سه خلأ پاسخگویی

با توجه به تحلیل سه خلأ پاسخگویی که در بخش‌های پیشین تشریح شد، این مقاله سه راهکار مشخص و به هم پیوسته پیشنهاد می‌کند که هر یک به طور مستقیم یکی از این خلأها را هدف قرار می‌دهد. علاوه بر این، چهارچوب چهار هدف پاسخگویی یعنی انطباق، گزارش، نظارت و اجرا [۶] به عنوان سازوکاری فراتحلیلی برای ساماندهی این راهکارها پیشنهاد می‌شود. بر این اساس، هر یک از راهکارهای زیر باید بتوانند هر چهار هدف پاسخگویی را به درجات مختلف پوشش دهند. راهکار نخست، قاعده تفویض ناپذیری مسئولیت نهایی انسانی است که به طور مستقیم نخستین خلأ را هدف قرار می‌دهد. بر اساس این قاعده، در تمامی سطوح هوش مصنوعی، همواره یک عامل انسانی به عنوان پاسخگوی نهایی در قبال خطاهای الگوریتمی تعیین می‌شود. این قاعده از دو مسیر قابل پیاده‌سازی است. مسیر نخست، مسئولیت مستقیم هیئت مدیره است. بر این اساس، هیچ تصمیم الگوریتمی که پیامدهای مالی یا غیرمالی چشمگیری برای شرکت و ذی‌نفعان آن دارد، بدون تأیید صریح و مستند حداقل یک عضو هیئت مدیره که از سواد الگوریتمی کافی برخوردار است، نهایی نخواهد شد. این قاعده به این معنا نیست که هوش مصنوعی نمی‌تواند به صورت مستقل عمل کند و توصیه ارائه دهد، بلکه در این چهارچوب، سامانه‌های هوش مصنوعی نقش پیشنهاددهنده را ایفا می‌کنند و نه تصمیم‌گیرنده نهایی. مسئولیت تصمیم همواره بر عهده عامل انسانی باقی می‌ماند و او باید بتواند مبنای تصمیم الگوریتمی را درک کند و در صورت لزوم، آن را رد یا اصلاح نماید [۱۳][۲۰][۲۱].

مسیر دوم، مسئولیت مبتنی بر بیمه و صندوق غرامت است. این مسیر برای سامانه‌های خودگردانی پیشنهاد می‌شود که به دلیل سرعت یا پیچیدگی، امکان مداخله پیشینی و تأیید انسانی در آنها وجود ندارد. در چنین مواردی، ترکیبی از مسئولیت مدنی توسعه‌دهنده، بیمه اجباری و صندوق غرامت که از محل عوارض صنعت هوش مصنوعی تأمین می‌شود، می‌تواند جایگزین پاسخگویی سنتی گردد. در این رویکرد، توسعه‌دهندگان سامانه‌های هوش مصنوعی موظف به خرید

بیمه مسئولیت حرفه‌ای با سقف مشخص هستند و برای خسارت‌های فراتر از سقف بیمه، صندوق غرامت ملی یا صنفی پاسخگو خواهد بود [۷][۱۳] . [۱۷] با تأکید بر امکان کدگذاری تعهدات امانی در هوش مصنوعی، نشان داده‌اند که قاعده تفویض ناپذیری مسئولیت نهایی انسانی از نظر فنی امکان‌پذیر است. [۱۸] نیز با نقد نظریه اعطای شخصیت حقوقی به هوش مصنوعی، استدلال کرده‌اند که ترکیب مسئولیت مدنی، بیمه و صندوق غرامت می‌تواند جایگزین کارآمدتری باشد. راهکار دوم، ایجاد کمیته تخصصی هوش مصنوعی در هیئت مدیره با اختیارات حسابرسی مستمر و نه صرفاً جلسات دوره‌ای است. این راهکار به طور مستقیم دومین خلأ، یعنی فقدان سازوکاری برای نظارت بر خود هیئت مدیره در مقام ناظر بر سامانه‌های هوش مصنوعی را هدف قرار می‌دهد. بر اساس این راهکار، یک کمیته فرعی از اعضای هیئت مدیره تشکیل می‌شود که متخصصان فنی مستقل نیز در آن حضور دارند. وظیفه این کمیته، پایش مستمر و در زمان واقعی عملکرد سامانه‌های هوش مصنوعی مورد استفاده در شرکت، ثبت و تحلیل گزارش‌های استثنائات، و ارائه گزارش حداقل ماهانه به هیئت مدیره است [۱۵][۲۰][۲۲] . برای ارزیابی کارآمدی این راهکار، می‌توان آن را در چهارچوب چهار هدف پاسخگویی شامل انطباق، گزارش، نظارت و اجرا تحلیل کرد [۶]. از این منظر، کمیته تخصصی هوش مصنوعی عمدتاً دو هدف گزارش و نظارت را پوشش می‌دهد. هدف گزارش از طریق جمع‌آوری و ثبت اطلاعات مربوط به عملکرد سامانه‌های الگوریتمی و ارائه گزارش‌های منظم به هیئت مدیره محقق می‌شود. هدف نظارت نیز از طریق فراهم آوردن ابزارهای لازم برای ارزیابی مستمر و مداخله به موقع در عملکرد سامانه‌ها تأمین می‌گردد. این کمیته باید دارای اختیارات حسابرسی مستقیم و دسترسی بی‌واسطه به داده‌های ورودی و خروجی سامانه‌های الگوریتمی باشد و بتواند در صورت تشخیص نقص یا خطای سیستماتیک، اجرای سامانه را به حالت تعلیق درآورد یا تغییر در پارامترهای آن را الزامی کند [۱۴]. در خصوص ترکیب اعضای این کمیته، تأکید بر حضور اعضای مستقل هیئت مدیره حائز اهمیت است [۹]. همچنین، [۲۶] بر اهمیت استقلال حسابرس و شفافیت رویه‌ای تأکید دارد و نشان می‌دهد که ترکیب حسابرسی فنی و فرآیندی می‌تواند الگویی مناسب برای این کمیته تخصصی فراهم آورد. [۳۲] نیز در چارچوب حکمرانی مسئولانه خود، سه لایه حکمرانی شامل هیئت مدیره و مدیران اجرایی، دفتر حکمرانی هوش مصنوعی، و مسئولیت واحدهای کسب‌وکار را پیشنهاد می‌دهد. ساختار پیشنهادی این کمیته، در واقع تلفیقی از لایه نظارتی هیئت مدیره و لایه تخصصی مستقل در چهارچوب بول است. بانک تسویه بین‌المللی نیز بر ضرورت تشکیل "جامعه عملی" شامل نهادهای نظارتی، بانک‌های مرکزی و شرکت‌های پیشرو برای اشتراک دانش، داده‌ها و بهترین شیوه‌های اجرایی در زمینه هوش مصنوعی تأکید کرده است. این رویکرد می‌تواند الگویی برای ایجاد شبکه همکاری میان هیئت‌های مدیره شرکت‌های مختلف و نهادهای ناظر بر آنها فراهم آورد [۳۳].

راهکار سوم، بازتعریف وظیفه امانتداری در عصر الگوریتم‌ها است که به طور مستقیم هر سه خلأ را به طور همزمان هدف قرار می‌دهد. بر اساس این راهکار، وظیفه امانتداری اعضای هیئت مدیره باید به گونه‌ای بازتعریف شود که شامل سواد الگوریتمی و مسئولیت نظارت بر چرخه حیات مدل‌های هوش مصنوعی نیز بشود [۱۳][۱۵][۲۳]. در این راستا، نظریه امانتداری مصنوعی نشان می‌دهد که سنت حقوقی موجود، فاقد راه‌حل‌های کافی برای مواجهه با یکپارچه‌سازی هوش مصنوعی در نظام حاکمیت شرکتی است و حوزه‌های قضایی با نظام حقوقی پیشرفته در حوزه شرکت‌ها، ناگزیر به اصلاح قوانین شرکت‌های خود برای هماهنگ‌سازی وظایف امانی سنتی ، به‌ویژه وظیفه مراقبت و وظیفه وفاداری، با پیشرفت‌های سریع فناوری خواهند بود [۳۴]. اعضای هیئت مدیره نه تنها موظف به نظارت بر عملکرد مالی و راهبردی شرکت هستند، بلکه موظف‌اند از نحوه طراحی، آموزش، استقرار، و بازآموزی سامانه‌های هوش مصنوعی مورد استفاده در شرکت آگاهی کافی داشته باشند و در صورت قصور در این نظارت در معرض مسئولیت شخصی قرار گیرند [۲۰]. از منظر چهار هدف پاسخگویی، این بازتعریف به طور همزمان هر چهار هدف را پوشش می‌دهد بطوری که انطباق را از طریق الزام به رعایت استانداردهای الگوریتمی، گزارش را از طریق مستندسازی تصمیمات الگوریتمی، نظارت را از طریق ایجاد سازوکارهای حسابرسی مستمر و اجرا را از طریق تعیین ضمانت‌های اجرایی برای موارد قصور تأمین می‌کند [۶]. نقض این وظیفه، به ویژه در موارد اتکای کور و غیرمنتقدانه به خروجی سامانه‌های هوش مصنوعی بدون درک محدودیت‌ها و سوگیری‌های احتمالی آنها، باید به عنوان مصداقی از قصور در ایفای وظایف قانونی و امانی مدیر تلقی شده و مستوجب مسئولیت مدنی در قبال شرکت و سهامداران ، از جمله قابلیت طرح دعوای مسئولیت توسط سهامداران و همچنین مبنایی برای رأی عدم کفایت و عزل توسط مجمع عمومی باشد [۱۳]. [۳۵] نیز بر ضرورت بازتعریف وظایف امانی برای شامل شدن نظارت بر سامانه‌های الگوریتمی تأکید کرده و پیشنهاد می‌دهند که الزام به حسابرسی الگوریتمی اجباری و مسیرهای شفاف پاسخگویی می‌تواند به عنوان مکمل این بازتعریف عمل کند. [۳۶] در تحلیل خود از قوانین مختلف حوزه‌های قضایی نشان می‌دهد که قواعد موجود شاغل بودن مدیران، برای انسان‌ها طراحی شده است و در صورت پذیرش مدیران مصنوعی، نیاز به بازنگری اساسی در قواعد نمایندگی، تفویض اختیار، و مسئولیت وجود دارد [۳۶]. سه راهکار پیشنهادی فوق برای آنکه به یک بسته سیاستی منسجم تبدیل شوند نیازمند یک چارچوب بالادستی وحدت‌بخش هستند که [۸] آن را تحت عنوان کنترل معنادار انسانی ارائه داده‌اند، چارچوبی که مستلزم تحقق دو شرط اساسی یعنی شرط ردیابی به معنای همسویی سیستم با دلایل و

انگیزه‌های عوامل انسانی ذیربط و شرط ردیابی ظرفیت‌ها به معنای همسویی سیستم با ظرفیت‌های فنی، انگیزشی و اخلاقی عوامل انسانی است و در واقع نقش اصول حاکم بر سه راهکار پیشنهادی را ایفا می‌کند به نحوی که قاعده تفویض‌ناپذیری مسئولیت نهایی انسانی شرط ردیابی را تضمین می‌کند و کمیته تخصصی هوش مصنوعی ظرفیت نظارت مستمر را فراهم می‌آورد و بازتعریف وظیفه امانتداری هر دو شرط را در سطح حقوقی نهادینه می‌سازد، بنابراین کنترل معنادار انسانی نه یک راهکار مجزا و اضافی بلکه مبنای نظری و ضامن انسجام سه راهکار عملی این مقاله به شمار می‌رود. این سه راهکار نه به صورت مستقل و مجزا بلکه به عنوان یک بسته سیاستی منسجم و به هم پیوسته باید اجرا شوند. قاعده تفویض ناپذیری مسئولیت نهایی انسانی بدون وجود کمیته تخصصی هوش مصنوعی و بازتعریف وظیفه امانتداری صرفاً یک توصیه اخلاقی بی‌ضمانت اجرا خواهد بود. کمیته تخصصی هوش مصنوعی بدون قاعده تفویض ناپذیری مسئولیت نهایی انسانی و بازتعریف وظیفه امانتداری به نهادی تشریفاتی و بی‌تأثیر تبدیل خواهد شد. بازتعریف وظیفه امانتداری بدون دو راهکار دیگر بار سنگینی بر دوش مدیران خواهد گذاشت بدون آنکه ابزارهای لازم برای ایفای این وظیفه را در اختیار آنها قرار دهد. بنابراین اجرای همزمان و هماهنگ این سه راهکار ضروری است [۲۰]. [۲۷] نیز تأکید دارند که حسابرسی مستمر هوش مصنوعی به عنوان یک لایه نظارتی اضافی می‌تواند مکمل این سه راهکار باشد و با فراهم کردن گزارش‌های به‌روز و خودکار امکان نظارت مؤثرتر هیئت مدیره بر سامانه‌های الگوریتمی را فراهم آورد.

بحث و نتیجه گیری

هوش مصنوعی با ورود به فرآیندهای تصمیم‌گیری هیئت مدیره، نه تنها دقت، سرعت و کارایی را افزایش داده، بلکه نظام سنتی پاسخگویی در حاکمیت شرکتی را با سه خلأ ساختاری و بنیادین مواجه ساخته است که در ادبیات موجود تاکنون به طور نظام‌مند و جامع به آنها پرداخته نشده است. نخستین خلأ، نبود قاعده روشن برای تعیین مسئول در هنگام بروز خطا از سوی سامانه‌های هوش مصنوعی است که به دلیل ماهیت بسته‌ای و چندمؤلفه‌ای که خود مفهوم مسئولیت دارد، پیچیده‌تر نیز می‌شود. برخی پژوهشگران به درستی اشاره کرده‌اند که استفاده از اصطلاح مسئولیت به عنوان ابزاری تحلیلی برای مسائل اخلاقی هوش مصنوعی گمراه‌کننده است، زیرا این اصطلاح مؤلفه‌های متعددی از جمله پاسخگویی به عنوان توانایی ارائه دلیل، قابل سرزنش بودن و مسئولیت قانونی را در خود جمع کرده است که هر یک الزامات متفاوتی برای عامل ایجاد می‌کنند [۷]. این خلأ در سطوح هوش افزایشی و هوش تقویتی که تصمیم‌گیری به شکلی متوازن و غیرقابل تفکیک میان انسان و ماشین توزیع می‌شود، به شکل شدیدتری بروز می‌کند. [۸] این وضعیت را شکاف قابلیت سرزنش می‌نامند و نشان می‌دهند که خطر اصلی این شکاف در کاهش انگیزه عوامل انسانی برای پیشگیری از رفتارهای نادرست و کاهش احتمال جبران خسارت قربانیان است.

دومین خلأ، فقدان سازوکاری برای نظارت بر خود هیئت مدیره در مقام ناظر بر سامانه‌های هوش مصنوعی بوده که از آن با عنوان معمای ناظر بر ناظر یاد شد و نشان داده شد که در نظام‌های سهامدارمحور به مراتب حادتر از نظام‌های ذی‌نفع‌گرا است. [۸] در تحلیل خود، دو نوع شکاف پاسخگویی را شناسایی کرده‌اند که مستقیماً با این معمای نظارتی مرتبط هستند: نخست، شکاف پاسخگویی اخلاقی که عبارت است از دشواری افراد در درک، تبیین و بازتاب منطق حاکم بر رفتار خود و دیگران به دلیل پیچیدگی و ابهام فرآیندهای تصمیم‌گیری الگوریتمی و دوم، شکاف پاسخگویی عمومی که ناظر بر دشواری شهروندان در دریافت توضیح درباره تصمیمات اتخاذشده توسط نهادهای عمومی و دشواری خود نهادها در پاسخگویی به این درخواست‌ها به دلیل واگذاری اختیارات به کارشناسان فنی و شرکت‌های خصوصی است [۸]. افزون بر این، از منظر حقوق استرالیا بر وظیفه امانی مستمر مدیران برای نظارت بر استفاده از سیستم‌های هوش مصنوعی و نقش کلیدی مدیران مستقل به عنوان یک مکانیسم نظارتی مؤکد تأکید شده است [۹].

سومین خلأ، ناهمگونی بازه‌های زمانی تصمیم‌گیری میان هیئت مدیره و سامانه‌های هوش مصنوعی است؛ ناهمگونی که امکان هرگونه نظارت مؤثر و پاسخگویی به موقع را از میان می‌برد و در پژوهش‌های مربوط به تصمیم‌گیری در زمان واقعی به طور سیستماتیک نادیده گرفته شده است. [۸] این وضعیت را ذیل شکاف مسئولیت فعال مورد بررسی قرار داده‌اند. مسئولیت فعال، آینده‌نگر است و به اهداف، ارزش‌ها و هنجارهایی مربوط می‌شود که متخصصان و مدیران موظف به ترویج و رعایت آنها هستند و نیز پیامدهایی که باید از آنها جلوگیری کنند. معرفی هوش مصنوعی می‌تواند دو دسته مسئله مرتبط با مسئولیت فعال ایجاد کند که مورد اول، آگاه نبودن مهندسان، مدیران و سایر عوامل درگیر نسبت به وظایف اخلاقی و اجتماعی خود بوده و مورد دوم، ناتوانی یا بی‌انگیزگی آنان در ایفای تعهداتی که از آنها آگاه هستند می باشد [۸]. [۶] با تمایز میان پاسخگویی پیش‌فعال و پاسخگویی واکنش‌گرا نشان داده‌اند که اتکای صرف به پاسخگویی واکنش‌گرا در مواجهه با سامانه‌های پرسرعت و خودگردان کافی نیست و باید با تقویت پاسخگویی پیش‌فعال، الزامات نظارتی را پیشاپیش در طراحی و استقرار سامانه‌ها لحاظ کرد.

تحلیل سه خلأ پاسخگویی ارائه‌شده در این مقاله و راهکارهای پیشنهادی برای پر کردن آنها، اگرچه مبتنی بر پژوهش‌های تجربی و نظری معتبر است، اما با محدودیت‌ها و چالش‌های عملی متعددی مواجه است که باید مورد توجه قرار گیرد. نخستین و مهم‌ترین محدودیت، وابستگی شدید راهکارهای پیشنهادی به شفافیت و تفسیرپذیری سامانه‌های هوش مصنوعی است. در حال حاضر، بسیاری از الگوریتم‌های پیشرفته به ویژه در حوزه یادگیری عمیق، ذاتاً غیرقابل

تفسیر یا جعبه سیاه هستند و حتی طراحان آنها نیز نمی‌توانند به طور کامل توضیح دهند که چرا یک ورودی خاص به یک خروجی خاص منجر شده است [۱۴][۲۰]. در چنین شرایطی، نمی‌توان از اعضای هیئت مدیره انتظار داشت که منطق پشت توصیه الگوریتمی را درک کنند و سپس در قبال آن تصمیم مسئولیت بپذیرند. بنابراین، پیش از اجرایی شدن این راهکار، باید سرمایه‌گذاری گسترده‌ای در حوزه هوش مصنوعی تفسیرپذیر انجام شود تا سامانه‌های هوش مصنوعی به گونه‌ای طراحی شوند که خروجی آنها برای انسان‌های غیرمتخصص نیز قابل درک باشد [۱۶]. تا زمانی که این پیش‌شرط فنی برآورده نشود، قاعده تفویض ناپذیری مسئولیت نهایی انسانی یا غیرممکن خواهد بود یا به یک تشریفات اداری بی‌محتوای کاهش می‌یابد. افزون بر این، تجربه حوزه اخلاق پزشکی نشان می‌دهد که اتکای صرف به اصول کلی، بدون ایجاد سازوکارهای نظارت مشارکتی، ساختارهای پاسخگویی الزام‌آور، فرایندهای شفاف بازبینی و حسابرسی اخلاقی مستقل در سطح سازمانی و صنفی، نمی‌تواند از بروز خطاها و سوءرفتارهای الگوریتمی جلوگیری کند و اخلاق هوش مصنوعی را باید به‌عنوان یک فرایند مستمر و نه یک مقصد نهایی در نظر گرفت که در آن اختلافات هنجاری عمیق نه نشانه شکست، بلکه نشانه جدی‌گیری اخلاقی و تنوع اندیشه است [۳۷].

دومین محدودیت، تفاوت‌های چشمگیر میان نظام‌های حقوقی و سنت‌های حاکمیت شرکتی در حوزه‌های قضایی مختلف است. همان طور که [۱۵] در تحلیل تطبیقی خود نشان داده است، نظام‌های سهامدارمحور مانند ایالات متحده، در مقایسه با نظام‌های ذی‌نفع‌گرا مانند اتحادیه اروپا، ظرفیت بسیار کمتری برای یکپارچه‌سازی سازوکارهای پاسخگویی الگوریتمی دارند و هرگونه تلاش برای بازتعریف وظیفه امانتداری در این نظام‌ها با مقاومت ساختاری و سیاسی مواجه خواهد شد [۱۵]. قانون هوش مصنوعی اتحادیه اروپا که در سال ۲۰۲۴ به تصویب رسید، سامانه‌های هوش مصنوعی پرخطر را موظف به رعایت الزامات سختگیرانه‌ای از جمله ایجاد نظام مدیریت ریسک، حاکمیت بر داده‌ها، مستندسازی فنی، نظارت انسانی، و دقت و امنیت سایبری کرده است، اما در کشورهایی که فاقد چنین قوانین جامعی هستند، خلأ پاسخگویی به قوت خود باقی خواهد ماند [۱۵]. از سوی دیگر، [۹] با تأکید بر ظرفیت‌های موجود در حقوق استرالیا نشان داده‌اند که حتی در نظام‌های حقوقی پیشرفته نیز برای مواجهه با چالش‌های هوش مصنوعی در حاکمیت شرکتی، نیاز به اصلاحات قانونی اساسی وجود دارد.

سومین محدودیت، فقدان یک نظام بین‌المللی هماهنگ برای تنظیم‌گری هوش مصنوعی در سطح حاکمیت شرکتی است. شرکت‌های چندملیتی با مجموعه‌ای از قوانین پراکنده و گاه متناقض در حوزه‌های قضایی مختلف مواجه هستند و این پراکندگی، امکان اجرای یکسان راهکارهای پیشنهادی را دشوار می‌سازد [۳۶][۳۵]. [۶] نیز بر این نکته تأکید دارند که پاسخگویی در هوش مصنوعی یک مسئله صرفاً حقوقی یا فنی نیست، بلکه یک مسئله اجتماعی و سیاسی است که نیازمند هماهنگی میان سیاست‌گذاران، تنظیم‌گران و فناوران در سطح ملی و فراملی است. با وجود این محدودیت‌ها، نادیده گرفتن سه خلأ پاسخگویی و عدم اقدام برای پر کردن آنها، هزینه‌های بسیار سنگین‌تری برای نظام حاکمیت شرکتی در پی خواهد داشت. تجربه خطاهای الگوریتمی در سال‌های اخیر، از سوگیری‌های جنسیتی و نژادی در سامانه‌های استخدام تا تصمیمات اشتباه در سامانه‌های تشخیص پزشکی و خودروهای خودران، نشان داده است که اتکای کور به سامانه‌های هوش مصنوعی بدون ایجاد سازوکارهای پاسخگویی شفاف، می‌تواند به خسارت‌های جبران‌ناپذیری برای افراد، سازمان‌ها و جوامع منجر شود. [۷] با تأکید بر لزوم استفاده از جعبه ابزار اخلاقی تخصصی به جای اتکای صرف به مفهوم کلی مسئولیت، نشان می‌دهد که راه حل مشکلات اخلاقی هوش مصنوعی در تفکیک مؤلفه‌های مختلف مسئولیت و طراحی سازوکارهای متناسب با هر یک نهفته است.

در نهایت، این مقاله نشان داد که پاسخگویی در عصر هوش مصنوعی نه یک مفهوم ساده و یک‌بعدی، بلکه شبکه‌ای پیچیده از روابط، مؤلفه‌ها و اهداف به هم پیوسته است. در این مقاله از چارچوب سودمندی که [۶] بر اساس چهار هدف انطباق، گزارش، نظارت و اجرا، برای ساماندهی این روابط پیچیده فراهم آورده‌اند نیز بهره گرفته شد. از سوی دیگر، [۸] با شناسایی چهار شکاف قابلیت سرزنش، پاسخگویی اخلاقی، پاسخگویی عمومی و مسئولیت فعال، نقشه راه روشنی برای شناسایی و درمان آسیب‌پذیری‌های نظام حاکمیت شرکتی در مواجهه با هوش مصنوعی ترسیم کرده‌اند که ترکیب این دو چارچوب با سه راهکار عملی پیشنهادی در این مقاله یعنی قاعده تفویض ناپذیری مسئولیت نهایی انسانی، ایجاد کمیته تخصصی هوش مصنوعی در هیئت مدیره، و بازتعریف وظیفه امانتداری، می‌تواند مبنایی مستحکم برای اصلاح نظام حاکمیت شرکتی در عصر الگوریتم‌ها فراهم آورد. همچنین، تجربه چالش‌های پیاده‌سازی فناوری‌های نوین در سازمان‌ها، از جمله هوش تجاری، نشان می‌دهد که نظام حاکمیت شرکتی در مواجهه با هوش مصنوعی نیازمند طراحی سازوکارهای نظارتی و پاسخگویی متناسب است [۳۸]. پیشنهاد می گردد که پژوهش‌های آینده باید به طور خاص بر روی روش‌های عملی برای افزایش شفافیت و تفسیرپذیری سامانه‌های هوش مصنوعی، طراحی نهادهای نظارتی مستقل و تخصصی در سطح ملی و فراملی و بررسی تطبیقی اثربخشی راهکارهای پیشنهادی در حوزه‌های قضایی مختلف و صنایع گوناگون متمرکز شوند. در این راستا، چهارچوب هوش مصنوعی مسئول و نقشه‌راه تحقیقاتی ارائه‌شده در منابع [۳۹][۴۰]، چهار پرسش کلیدی را برای جهت‌دهی به پژوهش‌های آتی پیشنهاد می‌کنند: نخست، چگونه می‌توان روش‌های تبیین‌پذیری هوش مصنوعی را برای اعضای غیرمتخصص هیئت مدیره قابل‌درک و کاربردی ساخت و چه سازوکارهایی از

سوءاستفاده از آن جلوگیری می‌کند؟ دوم، سازمان‌ها چگونه مسئولیت خطاهای الگورتیمی را تخصیص می‌دهند و آیا نظارت بر هوش مصنوعی باید در کمیته‌های موجود یا کمیته جدیدی مستقر شود؟ سوم، نظام‌های کنترل داخلی مؤثر بر هوش مصنوعی چه ویژگی‌هایی باید داشته باشند؟ و چهارم، تفاوت‌های نظارتی میان حوزه‌های قضایی چگونه بر شیوه‌های حاکمیت شرکتی تأثیر می‌گذارد؟ پاسخ به این پرسش‌ها، زمینه‌ساز تحقق نظام حاکمیت شرکتی خواهد بود که در آن، هوش مصنوعی نه به‌عنوان جعبه‌سیاهی غیرقابل‌کنترل، بلکه به‌عنوان ابزاری شفاف، عادلانه و پاسخگو در خدمت تصمیم‌گیری راهبردی قرار می‌گیرد.

با توجه به اینکه مواجهه با چالش‌های اخلاقی و حاکمیتی هوش مصنوعی، نیازمند برنامه‌ریزی جامع و مشارکت همهٔ ذی‌نفعان شامل سیاست‌گذاران، فناوران، نهادهای مدنی و شهروندان است و این رویکرد مشارکتی می‌تواند الگویی برای اصلاح نظام حاکمیت شرکتی در مواجهه با سه خلأ پاسخگویی فراهم آورد [۴۱]، پیشنهاد می‌شود که پژوهش‌های آتی به‌طور خاص به بررسی نگرش‌ها، انتظارات و اولویت‌های ذی‌نفعان مختلف در خصوص نحوه تخصیص مسئولیت در خطاهای الگورتیمی و سازوکارهای نظارت بر هیئت مدیره بپردازند تا از این طریق، زمینه برای طراحی چهارچوب‌های مشارکتی و عملیاتی‌تر پاسخگویی در نظام حاکمیت شرکتی فراهم شود. بدون چنین تلاش‌هایی، نظام حاکمیت شرکتی در عصر هوش مصنوعی یا به کلی بی‌اعتبار خواهد شد و یا به شکلی فزاینده به سمت خودگردانی کامل و حذف نظارت انسانی حرکت خواهد کرد که پیامدهای آن برای دموکراسی اقتصادی و پاسخگویی شرکتی، در حال حاضر قابل پیش‌بینی نیست اما بی‌تردید ژرف و دگرگون‌کننده خواهد بود.

تعارض منافع

نویسندگان این مقاله اعلام می دارند که هیچگونه تعارض منافع توسط نویسندگان بیان نشده است.

مراجع

[1] National Association of Corporate Directors. AI Oversight Activities Conducted by the Board: 2025 Public Company Board Practices Oversight Survey. 2025.

[2] Heidrick & Struggles. AI focus: How boards find expertise to chart the unknown [#Internet#] . 2025 [#cited 2026 Jun 2#]. Available from: https://www.heidrick.com/en/insights/frontier-tech/ai-focus_how-boards-find-expertise

[3] Bayat, Z. (2023). The Functions of Artificial Intelligence in the Field of Education and Electronic Knowledge Transfer. *Arman Process Journal (APJ)*, *3*(4), 43-49

[4] MIT Sloan Management Review. Successful companies now have AI-savvy boards [#Internet#]. 2025 [#cited 2026 Jun 2#]. Available from: https://mitsloan.mit.edu/ideas-made-to-matter/successful-companies-now-have-ai-savvy-boards

[5] OECD. Artificial Intelligence in Corporate Governance: Risks and Opportunities. OECD Publishing, Paris, 2025.

[6] Novelli C, Taddeo M, Floridi L. Accountability in artificial intelligence: what it is and how it works. Ai & Society. 2024 Aug;39(4):1871-82.

[7] Heinrichs JH. Responsibility assignment won't solve the moral issues of artificial intelligence. AI and Ethics. 2022 Nov;2(4):727-36.

[8] Santoni de Sio F, Mecacci G. Four responsibility gaps with artificial intelligence: Why they matter and how to address them. Philosophy & technology. 2021 Dec;34(4):1057-84.

[9] Rahim M, Dey P. Directors in Artificial Intelligence-Based Corporate Governance-An Australian Perspective. Corporate Governance and Artificial Intelligence–a Conflicting or Complementary Approach. 2022 Jan 1.

[10] Wang L, Chen T, Wang X, Cheng M. The Impact of Artificial Intelligence on Corporate Strategic Decision-Making: Convergence or Differentiation? In Proceedings of the 2025 International Conference on Digital Economy and Intelligent Computing 2025 May 23 (pp. 187-191).

[11] Gan, K. C. J., Wilczek-Stronczek, B., & Papasava, A. (2026). Governing the algorithm: a conceptual review of strategic AI oversight in global corporations. *AI & SOCIETY*, 1-24.

[12] Amarna AH, Venkateswarlu P, Venkateswarlu P. Investigating the Moderating Role of AI in the Relationship Between Corporate Governance and Financial Performance: Evidence from Palestine. An-Najah University Journal for Research-B (Humanities). 2026 Mar 24;9999(9999): None-.

[13] Petrin M. Corporate Management in the Age of AI. Colum. Bus. L. Rev. 2019:965.

[14] Hilb M. Toward artificial governance? The role of artificial intelligence in shaping the future of corporate governance: M. Hilb. Journal of Management and Governance. 2020 Dec;24(4):851-70.

[15] UNESCO. Who governs AI in the EU? A breakdown of authorities in the EU AI Act [#Internet#]. 2025 [#cited 2026 Jun 2#]. Available from: https://www.unesco.org/en/articles/who-governs-ai-eu-breakdown-authorities-eu-ai-act

[16] Riahi M., hashangholipoor thyasory T., Divandari A., Kalantari A. Strategic Decision-Making in the Age of Artificial Intelligence. SSQ, ۲۰۲۵; ۲۷(۴): ۱۷۷-۱۸۶. doi:۱۰/۲۲۰۳۴/srq.۲۰۲۴/۴۷۴۴۶۵/۴۱۹۲

Decision Precision and Risk Mitigation. OPUS Journal of Society Research. 2025 Apr 8;22(4):580-92.
[32] Bollu VK. Responsible AI governance in Enterprise Systems: A Risk and Compliance Framework. Acta Sci. 2026; 27:1.
[33] Bank for International Settlements (BIS). AI and financial stability: governance challenges for boards. BIS Working Papers, No 1215, 2025.
[34] Li, Z. (2024). Artificial fiduciaries. *Washington and Lee Law Review*, *81*(4), 1299
[35] Ahmed FA, Gul S, Shahzad S. Ensuring accountability and transparency in AI-driven corporate governance. International Journal of Social Sciences Bulletin. 2025 May 10;3(5):330-41.
[36] Möslein F. Robots in the boardroom: artificial intelligence and corporate law. In Research handbook on the law of artificial intelligence 2025 Jun 19 (pp. 614-635). Edward Elgar Publishing.
[37] Mittelstadt B. Principles alone cannot guarantee ethical AI. Nature Machine Intelligence. 2019; 1:501-507.

[38] Mohammadi. (2023). Business Intelligence and Industry 4.0: Opportunities & Challenges. (e1-7). *Arman Process Journal (APJ)*, *4*(1), e1-7 doi: 10.22034/apj.2023.706756
[39] Arrieta AB, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020; 58:82-115.

[40] Pittz, T. G. (2026). The new accountability frontier: AI governance, ESG assurance, and the future of an accountability infrastructure. *Cogent Business & Management*, *13*(1), 2685990.
[41] Floridi L, Cowls J, Beltrametti M, et al. AI4People—An Ethical Framework for a Good AI Society. Minds & Machines. 2018; 28:689-707.

[17] Pasban, M., Toosi, A., Mazaheri, M. Using Artificial Intelligence as a Corporate Director. *Private Law Research*,
[18] Hosseini S. A., Hashemizade, S. A. Study on the Granting of Legal Personality (Corporate) to Artificial Intelligence. *MTL*, 2025; 6(12): 147-167. doi: 10.22133/mtlj.2025.441446.1297 [#۱۵#]
[19] Kalkan G. The impact of artificial intelligence on corporate governance. Корпоративные финансы. 2024;18(2):17-25.
[20] Chaudhry FA. AI-Powered Decision-Making: Balancing Automation and Human Oversight in Corporate Governance. International Journal of Business & Computational Science. 2024;1(1).
[21] Humberd BK, Latham SF. When AI becomes an agent of the firm: Examining the evolution of AI in organizations through an agency theory lens. Journal of Management Studies. 2026 Mar;63(2):668-94.
[22] Sithipolvanichgul J. Board of directors' effectiveness and enterprise risk management: Do effective boards improve risk oversight? Thammasat Review. 2021 Apr 27;24(1):133-67.
[23] Nurzianti R, Rahmaddian T, Hadi MY. The Impact of Corporate Governance Mechanisms on Strategic Risk Management and Firm Performance. Journal Management & Economics Review (JUMPER). 2026;3(7):367-82.
[24] Selbst *AD, Boyd D, Friedler SA, et al. Fairness and Abstraction in Sociotechnical Systems. In Proceedings of FAT 2019: 59-68.*
[25] O'Malley, P. J., & Underwood, P. (2026). Reimagining corporate governance in the AI era: Law, legitimacy and algorithmic power. *Cambridge Forum on AI: Law and Governance, 2*(e2), 1–17. https://doi.org/10.1017/cfl.2026.10051
[26] Mökander, J. (2023). Auditing of AI: Legal, ethical and technical approaches. DISO 2, 49.
[27] Minkkinen M, Laine J, Mäntymäki M. Continuous auditing of artificial intelligence: a conceptualization and assessment of tools and frameworks. Digital Society. 2022 Dec;1(3):21.
[28] Ahdadou, M., Ahdadou, A., Aajly, A., & Tahrouch, M. (2025). Artificial intelligence in corporate boards: a dual-dimensional framework for integration across autonomy and structural levels. *Data Science and Management*.
[29] Coglianese C, Lehr D. Regulating by robot: Administrative decision making in the machine-learning era. Georgetown Law Journal. 2019; 105:1147-1223.
[30] Chinnaraju A. When AI Agents Act: Governance, Accountability, and Strategic Risk in Autonomous Organizations. International Journal of Research and Scientific Innovation. 2026; 12:12.
[31] Aktürk EB. Strategic Decision-Making and Artificial Intelligence: Exploring the Impact of AI Applications on

معرفی نویسندگان

مازیار قاسمی استادیار گروه مدیریت، مؤسسه آموزش عالی ادیبان گرمسار است. ایشان مدرک دکتری تخصصی خود را در رشته مدیریت مالی از دانشگاه پوترا مالزی(UPM) دریافت نموده است. حوزه تخصصی ایشان شامل مدیریت ریسک، حاکمیت شرکتی، مالی رفتاری و فینتک می‌باشد. علاقه‌مندی پژوهشی ایشان بررسی چالش‌های نوظهور پاسخگویی و حاکمیتی در عصر تحول دیجیتال و کاربست هوش مصنوعی در نظام راهبری شرکتی و مدیریت ریسک در عصر دیجیتال است.

Ghasemi, M., Assistant Professor, Department of Management, Adiban Garmsar Institute of Higher Education,Garmsar, Semnan, Iran
maziarghasemi@adiban.ac.ir

محمد جعفرپورجلالی استادیار گروه مهندسی برق، موسسه آموزش عالی ادیبان گرمسار است. ایشان مدرک دکتری تخصصی خود را در رشته مخابرات از دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران) دریافت نموده است. حوزه تخصصی ایشان شامل آنتن های هوشمند و کنترل پراکندگی با AI میباشد. علاقه مندی پژوهشی ایشان، تحقیق بر روی هوش مصنوعی مولد و یادگیری اپراتور (FNO) و حل بلادرنگ مسائل پیچیده با تمرکز بر توسعه مدل‌های فیزیک‌محور (PINN) در عصر دیجیتال است.

Jafarpour Jalali, M., Assistant Professor, Department of Electrical Engineering, Adiban Garmsar Institute of Higher Education,Garmsar, Semnan, Iran
Mjp.jalali@yahoo.com