Comparative Studies in Jurisprudence, Law, and Politics

An Introduction to the Theory of Artificial Intelligence Governance in Companies

1. Sajjad Kalantarzadeh Sadabad*: PhD Student, Department of Law, Shahid Beheshti University, Tehran, Iran.
Email: S_kalantarzadeh@sbu.ac.ir (Corresponding Author)
2. Mansour Amini: Associate Professor, Department of Law, Shahid Beheshti University, Tehran, Iran

ABSTRACT

Rapid advances in artificial intelligence are challenging the traditional structure of corporate administration by raising fundamental questions concerning decision-making authority, managerial competence, responsibility, and supervision. The central issue is no longer limited to the use of artificial intelligence as a tool for facilitating corporate activities, but extends to the possibility of transferring part of managerial authority to intelligent systems and recognizing them as effective elements within the structure of corporate governance. This study adopts a descriptive-analytical method with a forward-looking approach to formulate the theory of artificial intelligence governance in companies. Within this framework, a distinction is made among assistive, augmented, and autonomous artificial intelligence, while the proposed theory focuses on the level at which an intelligent system moves beyond a supporting role and participates independently in corporate decision-making and management. Two principal institutional models are conceivable: replacing the board of directors with an artificial intelligence system, or establishing an independent AI governance layer alongside existing corporate organs. The legal recognition of such models raises significant challenges, including the absence of legal personality for artificial intelligence, uncertainty regarding the attribution of liability, the need to define the scope of AI authority, and the requirement for effective human oversight. The implementation of AI governance also depends on principles such as transparency, data governance, continuous auditing, accountability, security, AI literacy, and ongoing legal monitoring. The study concludes that full recognition of this theory is not presently compatible with the existing framework of Iranian company law; however, a gradual legal approach based on a special regulatory framework, limited recognition of the legal capacity of intelligent systems, and experimental implementation in small and low-risk companies could provide the basis for the emergence of a new form of corporate organization referred to as the "smart company."

Keywords: *Company Law, Corporate Governance, Artificial Intelligence, AI Governance, Smart Company, Legal Personality.*

How to cite: Kalantarzadeh Sadabad, S. & Amini, M. (2028). An Introduction to the Theory of Artificial Intelligence Governance in Companies. *Comparative Studies in Jurisprudence, Law, and Politics*, *1*0(1), 1-18.

© 2028 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: 28 June 2026
Revise Date: 23 October 2026
Accept Date: 28 October 2026
Initial Publish Date: 31 October 2026
Final Publish Date: 20 April 2028

پژوهش‌های تطبیقی فقه، حقوق و سیاست

درآمدی بر نظریه حاکمیت هوش مصنوعی در شرکت‌ها

۱. سجاد کلانترزاده سعدآباد*: دانشجوی دکتری، دانشکده حقوق، دانشگاه شهید بهشتی، تهران، ایران.
پست الکترونیک: S_kalantarzadeh@sbu.ac.ir (نویسنده مسئول)
۲. منصور امینی: دانشیار، دانشکده حقوق دانشگاه شهید بهشتی، تهران، ایران

چکیده

تحولات شتابان هوش مصنوعی، ساختار سنتی اداره شرکت‌ها را با پرسش‌های بنیادینی در زمینه جایگاه تصمیم‌گیری، صلاحیت مدیریتی، مسئولیت و نظارت مواجه ساخته است. مسئله اصلی دیگر صرفاً استفاده از هوش مصنوعی به‌عنوان ابزاری برای تسهیل امور شرکت نیست، بلکه امکان انتقال بخشی از اختیارات مدیریتی به سامانه‌های هوشمند و تبدیل آن‌ها به عناصر مؤثر در ساختار حاکمیت شرکتی است. پژوهش حاضر با روش توصیفی ـ تحلیلی و رویکردی آینده‌پژوهانه، به تبیین نظریه حاکمیت هوش مصنوعی در شرکت‌ها می‌پردازد. در این چارچوب، میان هوش مصنوعی کمکی، تقویتی و خودمختار تمایز گذاشته می‌شود و نظریه مورد نظر بر سطحی متمرکز است که در آن سامانه هوشمند از نقش پشتیبان فراتر رفته و به‌طور مستقل در تصمیم‌گیری و اداره شرکت مداخله می‌کند. بر این اساس، دو الگوی اصلی قابل تصور است: جایگزینی هیئت‌مدیره با سامانه هوش مصنوعی و ایجاد لایه‌ای مستقل از حاکمیت هوش مصنوعی در کنار ارکان موجود شرکت. پذیرش این نظریه با چالش‌هایی همچون فقدان شخصیت حقوقی برای هوش مصنوعی، ابهام در انتساب مسئولیت، ضرورت تعیین حدود صلاحیت و لزوم تضمین نظارت مؤثر انسانی روبه‌رو است. همچنین، تحقق این الگو مستلزم رعایت اصولی مانند شفافیت، حاکمیت داده، حسابرسی مستمر، پاسخگویی، امنیت، سواد هوش مصنوعی و پایش حقوقی است. نتیجه پژوهش نشان می‌دهد که پذیرش کامل این نظریه در وضعیت فعلی حقوق ایران با محدودیت‌های تقنینی مواجه است، اما می‌توان با ایجاد چارچوبی خاص، شناسایی تدریجی ظرفیت حقوقی سامانه‌های هوشمند و استفاده آزمایشی در شرکت‌های کوچک و کم‌ریسک، زمینه شکل‌گیری نوع جدیدی از شرکت تحت عنوان «شرکت هوشمند» را فراهم ساخت.

واژگان کلیدی: *حقوق شرکت‌ها، حاکمیت شرکتی، هوش مصنوعی، حاکمیت هوش مصنوعی، شرکت هوشمند، شخصیت حقوقی.*

نحوه استناددهی: کلانترزاده سعدآباد، سجاد، و امینی، منصور. (۱۴۰۷). درآمدی بر نظریه حاکمیت هوش مصنوعی در شرکت‌ها. *پژوهش‌های تطبیقی فقه، حقوق و سیاست*، ۱۰(۱)، ۱۸-۱.

© ۱۴۰۷ 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: ۷ تیر ۱۴۰۵
تاریخ بازنگری: ۱ آبان ۱۴۰۵
تاریخ پذیرش: ۶ آبان ۱۴۰۵
تاریخ چاپ اولیه: ۹ آبان ۱۴۰۵
تاریخ چاپ نهایی: ۱ اردیبهشت ۱۴۰۷

مقدمه

تحولات ناشی از توسعه هوش مصنوعی در سال‌های اخیر، صرفاً به ظهور ابزارهای نوین برای تسهیل فعالیت‌های اقتصادی یا خودکارسازی برخی فرایندهای اداری محدود نمانده است، بلکه به‌تدریج شیوه سازماندهی، اداره، تصمیم‌گیری و نظارت در بنگاه‌های اقتصادی را نیز تحت تأثیر قرار داده است. شرکت تجاری به‌عنوان یکی از مهم‌ترین نهادهای حقوق خصوصی، در ساختار سنتی خود مبتنی بر تفکیک میان ارکان انسانی، توزیع صلاحیت میان مجامع و مدیران و انتساب تصمیمات به اشخاص حقیقی یا حقوقی شناخته‌شده در نظام حقوقی شکل گرفته است. بااین‌حال، توسعه سامانه‌هایی که قادر به تحلیل حجم عظیمی از داده‌ها، استخراج الگوها، پیش‌بینی نتایج، ارزیابی گزینه‌های جایگزین و حتی اتخاذ یا اجرای تصمیمات بدون مداخله مستقیم انسان هستند، این پرسش را مطرح می‌کند که آیا ساختار کلاسیک اداره شرکت همچنان می‌تواند بدون تغییر باقی بماند. مöslein نشان می‌دهد که تحول دیجیتال در حقوق شرکت‌ها دیگر صرفاً مسئله استفاده از ابزارهای الکترونیکی نیست و فناوری‌های مبتنی بر داده و هوش مصنوعی ظرفیت آن را دارند که ساختار تصمیم‌گیری شرکتی را به‌صورت عمیق‌تری متحول کنند (Möslein, 2025). از این منظر، مسئله اساسی نه استفاده یا عدم استفاده شرکت از هوش مصنوعی، بلکه تعیین جایگاه این فناوری در ساختار حقوقی حاکم بر اداره شرکت است.

در مراحل نخست ورود فناوری به حوزه شرکت‌ها، دیجیتالی‌شدن عمدتاً به الکترونیکی‌شدن ارتباطات، ثبت اطلاعات، برگزاری جلسات، تبادل اسناد و تسهیل تعامل میان سهامداران و مدیران مربوط می‌شد. Knapp در بررسی مسائل مرتبط با دیجیتالی‌شدن حقوق شرکت‌ها نیز بخش مهمی از این تحول را در چارچوب ابزارهای ارتباطی و فرایندهای الکترونیکی تحلیل می‌کند (Knapp, 2016). در این مرحله، فناوری ماهیت تصمیم‌گیرنده نداشت و صرفاً ابزاری بود که اشخاص انسانی از آن برای انجام سریع‌تر و دقیق‌تر وظایف خود استفاده می‌کردند. Bertschinger نیز در تحلیل حقوق شرکت‌ها در عصر دیجیتال، تحول ناشی از فناوری را در نسبت با ساختارهای سنتی شرکت و امکان استفاده گسترده‌تر از ابزارهای دیجیتال بررسی کرده است (Bertschinger, 2015). با توسعه هوش مصنوعی، ماهیت مسئله تغییر کرده است؛ زیرا فناوری جدید می‌تواند از نقش منفعل یک ابزار ارتباطی یا پردازشی عبور کند و به عنصری فعال در فرایند تصمیم‌سازی یا حتی تصمیم‌گیری تبدیل شود. در چنین وضعیتی، مرز میان «استفاده شرکت از فناوری» و «اداره شرکت توسط فناوری» اهمیت بنیادین پیدا می‌کند. هوش مصنوعی در سطح ابتدایی می‌تواند در مقام یک ابزار کمکی برای تحلیل اطلاعات، تهیه گزارش، ارزیابی ریسک یا پیش‌بینی وضعیت بازار مورد استفاده قرار گیرد. در این وضعیت، مدیران همچنان تصمیم‌گیرندگان واقعی هستند و خروجی الگوریتم صرفاً یکی از عناصر مؤثر در شکل‌گیری اراده مدیریتی محسوب می‌شود. با افزایش خودمختاری سامانه، رابطه میان مدیر و فناوری پیچیده‌تر می‌شود. Hilb با طرح مفهوم «حکمرانی مصنوعی» نشان می‌دهد که هوش مصنوعی می‌تواند از یک ابزار مدیریتی فراتر رفته و در آینده ساختار حاکمیت شرکتی را تحت تأثیر قرار دهد (Hilb, 2020). از اینجا می‌توان میان دو الگوی متفاوت تمایز قائل شد: در الگوی نخست، «حاکمیت شرکتی با استفاده از هوش مصنوعی» تحقق می‌یابد و فناوری همچنان تحت هدایت مدیران انسانی عمل می‌کند؛ در الگوی دوم، «حاکمیت هوش مصنوعی در شرکت» شکل می‌گیرد و هوش مصنوعی خود در موقعیت اعمال صلاحیت مدیریتی قرار می‌گیرد.

این تمایز مبنای نظری پژوهش حاضر است. مقصود از نظریه حاکمیت هوش مصنوعی در شرکت‌ها وضعیتی نیست که در آن مدیران صرفاً از ابزارهای هوشمند برای انجام بهتر وظایف خود استفاده کنند، بلکه وضعیتی است که در آن بخشی اساسی یا همه صلاحیت‌های مدیریتی، جز اختیاراتی که ماهیتاً یا قانوناً به مجامع عمومی اختصاص دارد، به یک یا چند سامانه هوش مصنوعی واگذار شود. در این مدل، هوش

مصنوعی می‌تواند یا جایگزین هیئت‌مدیره شود یا در قالب یک رکن مستقل و جدید در کنار ارکان موجود شرکت ایفای نقش کند. در هر دو حالت، فناوری از جایگاه «ابزار» به جایگاه «عنصر حکمرانی» منتقل می‌شود. چنین تحولی مستلزم بازاندیشی در مفاهیمی همچون مدیر، اراده شرکت، صلاحیت مدیریتی، مسئولیت، شخصیت حقوقی، نظارت و پاسخگویی است.

اهمیت این تحول هنگامی بیشتر آشکار می‌شود که به ماهیت هوش مصنوعی جدید توجه شود. بسیاری از سامانه‌های معاصر صرفاً بر اجرای دستورات از پیش تعیین‌شده متکی نیستند، بلکه از داده‌ها یاد می‌گیرند، الگوها را استخراج می‌کنند و بر اساس احتمال و پیش‌بینی عمل می‌کنند. Wohlfarth تفاوت میان یادگیری ماشینی و هوش مصنوعی مولد را بر ظرفیت متفاوت این سامانه‌ها در تحلیل، یادگیری و تولید خروجی‌های جدید متمرکز می‌داند Marr .(Wohlfarth, 2024) نیز توضیح می‌دهد که هوش مصنوعی مولد برخلاف بسیاری از سامانه‌های سنتی تنها به دسته‌بندی یا پیش‌بینی محدود نیست و می‌تواند محتوای جدید و راهکارهای نو تولید کند (Marr, 2023). بنابراین، در حوزه مدیریت شرکت نیز می‌توان تصور کرد که سامانه هوشمند نه‌تنها وضعیت مالی یا عملیاتی را تحلیل کند، بلکه گزینه‌های راهبردی تولید کرده، میان آن‌ها انتخاب کند و سپس تصمیم منتخب را به مرحله اجرا برساند.

ورود چنین سامانه‌ای به ساختار شرکت، در عین برخورداری از مزایای بالقوه، مخاطرات ویژه‌ای نیز ایجاد می‌کند. افزایش سرعت تصمیم‌گیری، امکان پردازش حجم گسترده اطلاعات، کاهش برخی سوگیری‌های انسانی، بهبود پیش‌بینی و افزایش قابلیت پایش مستمر می‌تواند به‌عنوان مزیت مطرح شود، اما در مقابل، نبود شفافیت الگوریتمی، دشواری انتساب مسئولیت، سوگیری داده‌ها، تصمیمات غیرقابل توضیح، وابستگی به زیرساخت‌های فنی و خطر فاصله گرفتن تصمیمات از معیارهای حقوقی یا منافع شرکت نیز از چالش‌های مهم به‌شمار می‌روند. Nersessian و Mancha نشان داده‌اند که گذار از اتوماسیون به خودمختاری می‌تواند شکاف‌های مهمی در حوزه مسئولیت حقوقی و اخلاقی ایجاد کند (Nersessian & Mancha, 2020). این شکاف در حقوق شرکت‌ها اهمیت مضاعف دارد؛ زیرا نظام مسئولیت مدیران اساساً بر این فرض بنا شده است که تصمیم مدیریتی به یک شخص قابل شناسایی و دارای اراده حقوقی منتسب است.

در نظام حقوقی ایران نیز ساختار اداره شرکت‌ها اساساً بر حضور اشخاص استوار است. حقوق شرکت‌ها وظایف و اختیارات را میان ارکانی توزیع می‌کند که اعضای آن‌ها از شخصیت حقوقی شناخته‌شده برخوردارند. Eskini در تحلیل ساختار شرکت‌های تجاری، شخصیت حقوقی شرکت و جایگاه ارکان مدیریتی را از عناصر بنیادین سازمان حقوقی شرکت می‌داند (Eskini, 2014). ماده ۱۱۰ لایحه اصلاح قسمتی از قانون تجارت نیز امکان عضویت شخص حقوقی در هیئت‌مدیره شرکت سهامی را به رسمیت می‌شناسد، اما در وضعیت کنونی، سامانه هوش مصنوعی نه شخص حقیقی است و نه در شمار اشخاص حقوقی قرار می‌گیرد. بنابراین، ورود آن به هیئت‌مدیره یا اعطای موقعیت مستقل مدیریتی به آن در وضع موجود با مانع مفهومی و تقنینی مواجه است.

پژوهش حاضر با رویکردی توصیفی ـ تحلیلی و آینده‌پژوهانه، در پی تبیین یک نظریه حقوقی برای مواجهه با چنین وضعیتی است. هدف اصلی آن بررسی این مسئله است که حاکمیت هوش مصنوعی در شرکت‌ها از حیث مفهومی به چه معناست، بر چه ساختاری استوار می‌شود، تحت چه شرایطی امکان پذیرش حقوقی آن وجود دارد و چه اصولی باید برای کنترل مخاطرات ناشی از آن در نظر گرفته شود. پرسش محوری این است که آیا می‌توان از ساختار سنتی حاکمیت شرکتی که در آن صلاحیت اداره به اشخاص انسانی یا حقوقی شناخته‌شده اعطا می‌شود، به ساختاری گذر کرد که در آن سامانه هوش مصنوعی بخشی از صلاحیت‌های مدیریتی را به‌طور مستقل اعمال کند. پاسخ به این

پرسش، مستلزم تفکیک دقیق سطوح مداخله هوش مصنوعی، تبیین مدل‌های ممکن حاکمیت، بررسی مسئله شخصیت و صلاحیت حقوقی و در نهایت ارائه مجموعه‌ای از اصول برای تضمین شفافیت، نظارت و پاسخگویی است.

مبانی نظری حاکمیت هوش مصنوعی در شرکت‌ها

برای تبیین نظریه حاکمیت هوش مصنوعی در شرکت‌ها، نخست باید میان استفاده ابزاری از هوش مصنوعی و واگذاری صلاحیت حکمرانی به آن تمایز قائل شد. بسیاری از کاربردهایی که امروزه تحت عنوان هوش مصنوعی در شرکت‌ها مطرح می‌شوند، در حقیقت تغییری در ساختار حقوقی اداره شرکت ایجاد نمی‌کنند. استفاده از سامانه‌های تحلیل داده، مدیریت اسناد، شناسایی ریسک، پیش‌بینی بازار، ارزیابی عملکرد کارکنان یا ارائه توصیه‌های سرمایه‌گذاری، تا زمانی که تصمیم نهایی توسط مقام انسانی اتخاذ شود، در قلمرو استفاده ابزاری باقی می‌ماند. Hilb بر این نکته تأکید می‌کند که کاربرد هوش مصنوعی در حاکمیت شرکتی می‌تواند طیفی از پشتیبانی محدود تا نقش‌آفرینی عمیق در فرایندهای حکمرانی را دربرگیرد (Hilb, 2020). بنابراین، صرف حضور یک سامانه پیشرفته در فرایند مدیریتی به معنای تحقق حاکمیت هوش مصنوعی نیست.

مبنای مناسب برای تفکیک این وضعیت‌ها، سطح خودمختاری سامانه است. در سطح نخست می‌توان از هوش مصنوعی کمکی سخن گفت. در این سطح، سامانه تنها وظایف مشخص و محدودی را انجام می‌دهد و خروجی آن ماهیت پشتیبان دارد. برای مثال، ممکن است سامانه مجموعه‌ای از قراردادها را تحلیل، داده‌های مالی را طبقه‌بندی یا ریسک‌های احتمالی را شناسایی کند، اما تصمیم درباره انعقاد قرارداد، تخصیص منابع یا تغییر راهبرد همچنان به مدیر انسانی تعلق دارد. در این وضعیت، فاصله میان تحلیل ماشینی و اراده حقوقی شرکت روشن است و انتساب تصمیم نیز با دشواری جدی مواجه نمی‌شود. چنین کاربردی را می‌توان ادامه منطقی اتوماسیون در محیط شرکت دانست.

سطح دوم، هوش مصنوعی تقویتی است. در این وضعیت، فناوری به‌صورت مستمر در فرایند تصمیم‌گیری حضور دارد و تصمیم مدیر تا حد قابل توجهی بر خروجی الگوریتم متکی است. سامانه ممکن است چند سناریوی راهبردی را تولید، آثار مالی هر یک را محاسبه، ریسک‌های احتمالی را پیش‌بینی و گزینه مطلوب را پیشنهاد کند. بااین‌حال، انسان همچنان مرجع نهایی تصمیم است. Koulu در بحث کنترل انسانی بر اتوماسیون نشان می‌دهد که حتی در ساختارهای پیشرفته، حفظ ظرفیت مداخله انسانی یکی از عناصر کلیدی برای جلوگیری از تبدیل اتوماسیون به خودمختاری مهارنشده است (Koulu, 2020). در این سطح، رابطه انسان و ماشین به‌صورت مشارکتی شکل می‌گیرد؛ هوش مصنوعی کیفیت اطلاعات و تحلیل را افزایش می‌دهد و مدیر مسئول پذیرش یا رد توصیه است.

سطح سوم زمانی تحقق می‌یابد که سامانه از استقلال در اتخاذ تصمیم برخوردار شود. در این سطح، هوش مصنوعی نه‌فقط داده تولید می‌کند و نه‌فقط تصمیم پیشنهاد می‌دهد، بلکه خود بر اساس معیارهای تعریف‌شده یا یادگرفته‌شده تصمیم می‌گیرد و ممکن است آن را اجرا کند. تفاوت بنیادی این سطح با دو سطح پیشین در انتقال نقطه نهایی اختیار است. در مدل کمکی و تقویتی، اراده حقوقی از انسان ناشی می‌شود؛ در مدل خودمختار، تصمیم عملیاتی یا مدیریتی می‌تواند بدون تأیید موردی انسان اتخاذ شود. Nersessian و Mancha این گذار را نقطه‌ای می‌دانند که در آن مسئله مسئولیت پیچیده‌تر می‌شود، زیرا رفتار سامانه ممکن است دیگر به‌سادگی به دستور مستقیم یک شخص خاص تقلیل‌پذیر نباشد (Nersessian & Mancha, 2020).

نظریه حاکمیت هوش مصنوعی دقیقاً با سطح سوم پیوند دارد. این نظریه بر این فرض استوار است که در آینده می‌توان سامانه‌هایی را تصور کرد که توانایی انجام بخش عمده‌ای از وظایف سنتی هیئت‌مدیره را داشته باشند. چنین سامانه‌ای می‌تواند وضعیت شرکت را به‌طور پیوسته

پایش کند، تصمیمات سرمایه‌گذاری را ارزیابی کند، برنامه‌های عملیاتی را تغییر دهد، درباره تأمین مالی تصمیم بگیرد، ریسک‌های قراردادی را بسنجند و حتی در محدوده اختیارات خود قراردادهایی را منعقد یا اجرا کند. در چنین ساختاری، سامانه از ابزاری در اختیار مدیر به عامل مستقلی در ساختار حکمرانی شرکت تبدیل می‌شود.

این تحول مستلزم تمایز میان «حکمرانی با هوش مصنوعی» و «حکمرانی هوش مصنوعی» است. حکمرانی با هوش مصنوعی به معنای استفاده مدیران از فناوری برای ارتقای کیفیت تصمیم‌گیری است. در این مدل، ساختار حقوقی سنتی حفظ می‌شود و تنها ابزارهای تصمیم‌گیری تغییر می‌کنند. در مقابل، حکمرانی هوش مصنوعی به معنای قرار گرفتن خود سامانه در جایگاه اعمال صلاحیت است. Möslein تحول فناوری در حقوق شرکت‌ها را زمینه‌ای برای بازاندیشی در ساختار سنتی تصمیم‌گیری می‌داند (Möslein, 2025). بر این اساس، نظریه حاکمیت هوش مصنوعی را می‌توان مرحله‌ای فراتر از دیجیتالی‌شدن یا حتی استفاده گسترده از الگوریتم در شرکت دانست.

از لحاظ فنی، تحقق چنین نظریه‌ای به وجود سامانه‌ای نیازمند است که بتواند همزمان چند نوع قابلیت را ترکیب کند. هوش مصنوعی تحلیلی امکان پردازش و ارزیابی داده‌های گذشته و حال را فراهم می‌کند، در حالی که سامانه‌های مولد می‌توانند سناریوهای جدید، پیشنهادهای راهبردی و راه‌حل‌های بدیل تولید کنند. Marr بر همین تمایز میان ظرفیت تحلیلی سامانه‌های سنتی و قدرت تولید محتوای جدید در هوش مصنوعی مولد تأکید دارد (Marr, 2023). اگر این دو قابلیت با امکان اجرای مستقل تصمیم ترکیب شوند، می‌توان مدلی را تصور کرد که از مرحله مشاوره مدیریتی فراتر رفته و به اداره عملی شرکت نزدیک شود.

از دیدگاه حقوقی، خودمختاری فنی به‌تنهایی برای تحقق حاکمیت کافی نیست. سامانه باید در ساختار صلاحیت شرکت نیز جایگاه مشخصی داشته باشد. ممکن است یک سامانه از نظر فنی توانایی اتخاذ تصمیم داشته باشد، اما تا زمانی که حقوق شرکت آن را صاحب صلاحیت نشناسد، تصمیم آن از منظر حقوقی صرفاً نتیجه یک فرایند فنی است که باید در نهایت به شخص انسانی یا حقوقی دیگری منتسب شود. ازاین‌رو نظریه حاکمیت هوش مصنوعی، علاوه بر بعد فناورانه، دارای یک بعد نهادی است. مسئله اصلی این نیست که ماشین «می‌تواند» تصمیم بگیرد، بلکه آن است که آیا نظام حقوقی «اجازه می‌دهد» تصمیم ماشینی مستقیماً به‌عنوان تصمیم یک رکن شرکتی شناخته شود.

این موضوع با مفهوم شخصیت و صلاحیت ارتباط مستقیم پیدا می‌کند. Amini و Kalantarzadeh Sadabad با بازاندیشی مفهوم سنتی شخصیت حقوقی در پرتو پیشرفت‌های هوش مصنوعی، نشان می‌دهند که توسعه سامانه‌های مستقل می‌تواند نظام حقوقی را با ضرورت بازنگری در معیارهای برخورداری از شخصیت مواجه کند (Amini & Kalantarzadeh Sadabad, 2026). در چارچوب سنتی، شخصیت حقوقی ابزار انتساب حق، تکلیف و مسئولیت است. اگر سامانه هوش مصنوعی قرار باشد صاحب اختیارات مدیریتی شود، پرسش آن است که آیا می‌توان بدون اعطای نوعی وضعیت حقوقی مستقل، این صلاحیت را به آن واگذار کرد.

برخی رویکردها به‌جای شناسایی شخصیت کامل، به دنبال طراحی اشکال محدودتر هستند. Boudraa ایده شخصیت حقوقی موقت یا محدود برای هوش مصنوعی را به‌عنوان ابزاری برای ایجاد چارچوب پاسخگویی مطرح می‌کند (Boudraa, 2025). چنین الگویی می‌تواند برای حقوق شرکت‌ها جذاب باشد؛ زیرا لازم نیست سامانه هوش مصنوعی در تمام حوزه‌های حقوقی همانند انسان یا شرکت دارای شخصیت باشد. می‌توان شخصیت یا صلاحیت آن را تنها در محدوده وظایف شرکتی تعریف کرد و حقوق و مسئولیت‌های آن را نیز محدود به همان قلمرو ساخت.

با وجود این، نظریه حاکمیت هوش مصنوعی نباید با ایده حذف کامل انسان از شرکت یکسان دانسته شود. حتی در پیشرفته‌ترین مدل، انسان همچنان در سطح مالکیت، تعیین اهداف بنیادی، تصویب موضوعاتی که در صلاحیت مجامع قرار دارد و نظارت بر عملکرد سامانه نقش خواهد داشت. Laux اصل نظارت انسانی را یکی از اجزای اساسی حکمرانی مسئولانه هوش مصنوعی می‌داند (Laux, 2024). بنابراین، مدل مطلوب نه شرکت فاقد انسان، بلکه شرکتی است که در آن نقش انسان از «تصمیم‌گیر روزمره» به «طراح اهداف، ناظر و مداخله‌کننده نهایی» تغییر می‌کند.

این تمایز همچنین مانع از آن می‌شود که هر شرکت دارای فرایندهای خودکار، «شرکت هوشمند» نامیده شود. ممکن است شرکتی تقریباً تمام عملیات حسابداری، بازاریابی، فروش و لجستیک خود را خودکار کند، اما اگر تصمیمات نهایی همچنان به هیئت‌مدیره انسانی منتسب باشند، ساختار حقوقی آن تغییری نکرده است. شرکت هوشمند در معنای نظری حاضر، شرکتی است که سامانه هوش مصنوعی در آن صلاحیت حقوقی مشخص برای اداره امور دارد. از این جهت، نظریه حاکمیت هوش مصنوعی نه نظریه‌ای درباره ابزارهای فناوری اطلاعات، بلکه نظریه‌ای درباره بازتوزیع صلاحیت در ساختار حقوقی شرکت است.

چارچوب نظری حاکمیت هوش مصنوعی و امکان پذیرش آن در حقوق شرکت‌ها

چارچوب نظری حاکمیت هوش مصنوعی را می‌توان بر مبنای دو الگوی نهادی اصلی تصور کرد. در الگوی نخست، هوش مصنوعی مستقیماً جایگزین هیئت‌مدیره می‌شود. در این مدل، سامانه یا مجموعه‌ای از سامانه‌ها وظایفی را بر عهده می‌گیرند که در ساختار سنتی بر عهده مدیران انسانی بوده است. تصمیمات مدیریتی، تخصیص منابع، ارزیابی قراردادها، انتخاب راهبردها، کنترل عملیات و سایر صلاحیت‌های مرتبط با اداره شرکت در محدوده قانونی به سامانه واگذار می‌شود. مجمع عمومی همچنان جایگاه خود را در موضوعاتی که ذاتاً در صلاحیت سهامداران قرار دارد حفظ می‌کند، اما اداره جاری و راهبردی شرکت به هوش مصنوعی منتقل می‌شود. این مدل بیشترین فاصله را با ساختار سنتی دارد و به همین دلیل بیشترین اصلاح حقوقی را نیز طلب می‌کند.

در الگوی دوم، هیئت‌مدیره انسانی حذف نمی‌شود، بلکه یک لایه مستقل حاکمیت هوش مصنوعی در کنار آن ایجاد می‌شود. در این ساختار، بخشی از صلاحیت‌هایی که در وضعیت سنتی در اختیار هیئت‌مدیره است، به رکن هوشمند منتقل می‌شود و هیئت‌مدیره انسانی عمدتاً نقش نظارتی پیدا می‌کند. مزیت این مدل آن است که گذار به حاکمیت هوش مصنوعی می‌تواند به‌صورت تدریجی انجام شود و مسئولیت نظارتی انسان همچنان حفظ گردد. Hilb در بررسی آینده حاکمیت شرکتی، نقش ترکیبی انسان و هوش مصنوعی را از سناریوهای مهم تحول سازمانی می‌داند (Hilb, 2020). چنین ساختاری می‌تواند مرحله میانی میان حاکمیت کاملاً انسانی و مدل کاملاً خودمختار تلقی شود.

در هر دو الگو، مسئله تعیین حدود صلاحیت اهمیت اساسی دارد. هیچ دلیلی وجود ندارد که پذیرش هوش مصنوعی در ساختار حاکمیت به معنای انتقال همه اختیارات شرکت به آن باشد. برخی صلاحیت‌ها می‌توانند بنا بر قانون، اساسنامه یا ماهیت تصمیم همچنان برای اشخاص انسانی محفوظ بمانند. برای مثال، اصلاح اساسنامه، افزایش یا کاهش سرمایه، انحلال شرکت یا تصمیماتی که به تغییر بنیادین ساختار حقوقی شرکت منجر می‌شوند، می‌توانند همچنان در صلاحیت مجمع عمومی باقی بمانند. در مقابل، مدیریت عملیات، انتخاب پیمانکار، مدیریت سرمایه در گردش، تخصیص بودجه داخلی یا تصمیمات تجاری روزمره می‌توانند در قلمرو صلاحیت سامانه قرار گیرند.

پرسش مهم‌تر آن است که آیا چنین صلاحیتی را می‌توان در نظام حقوقی موجود به یک سامانه هوش مصنوعی اعطا کرد. در بسیاری از نظام‌های حقوقی، عضویت در هیئت‌مدیره به اشخاص حقیقی محدود شده است. Gehrlein در تحلیل امکان اداره یک شخص حقوقی توسط

شخص حقوقی دیگر نشان می‌دهد که حتی در خصوص اشخاص حقوقی نیز مسئله عضویت در ارکان مدیریتی با محدودیت‌های مفهومی و تقنینی مواجه بوده است (Gehrlein, 2016). در حقوق ایران وضعیت تا حدی متفاوت است؛ زیرا قانون، عضویت شخص حقوقی در هیئت‌مدیره شرکت سهامی را به رسمیت شناخته است. این امر نشان می‌دهد که حقوق ایران الزاماً مدیریت را به «انسان» محدود نکرده است، اما همچنان «شخص بودن» مدیر را مفروض گرفته است.

این نکته مسیر تحلیل را به مسئله شخصیت حقوقی هوش مصنوعی هدایت می‌کند. اگر سامانه هوشمند هیچ موقعیت حقوقی مستقلی نداشته باشد، نمی‌توان به‌آسانی آن را حامل صلاحیت‌های مدیریتی دانست. Amini و Kalantarzadeh Sadabad بحث شخصیت حقوقی را دقیقاً از همین زاویه با تحولات هوش مصنوعی مرتبط می‌کنند و ضرورت بازنگری در برداشت سنتی از شخصیت را مطرح می‌سازند (Amini & Kalantarzadeh Sadabad, 2026). چنانچه قانون‌گذار نوعی شخصیت یا ظرفیت حقوقی محدود برای سامانه‌های واجد شرایط به رسمیت بشناسد، امکان طراحی قواعد خاص برای حضور آن‌ها در ساختار شرکت نیز فراهم می‌شود.

این شخصیت لزوماً نباید مشابه شخصیت حقوقی شرکت یا شخصیت حقیقی انسان باشد. Boudraa از امکان طراحی نوعی شخصیت موقت یا محدود برای سامانه‌های هوشمند سخن می‌گوید که هدف آن ایجاد امکان انتساب مسئولیت و تنظیم روابط حقوقی است (Boudraa, 2025). در حوزه شرکت‌ها می‌توان چنین ایده‌ای را به‌صورت «شخصیت مدیریتی محدود» تصور کرد؛ یعنی سامانه فقط در محدوده مأموریت شرکتی خود قابلیت برخورداری از صلاحیت داشته باشد. در این صورت، موجودیت حقوقی آن از شرکت جدا اما به وظایف خاص آن وابسته خواهد بود.

راهکار دیگر آن است که قانون‌گذار بدون اعطای شخصیت کامل، مفهوم «مدیر» یا «رکن مدیریتی» را توسعه دهد و سامانه‌های واجد شرایط را صراحتاً مشمول آن بداند. چنین روشی از حیث تقنینی ساده‌تر به نظر می‌رسد، اما با مشکل مسئولیت روبه‌رو است؛ زیرا اگر سامانه خود شخص نباشد، باید مشخص شود آثار حقوقی تصمیم آن به چه کسی منتسب خواهد شد. انتساب خودکار همه تصمیمات به شرکت می‌تواند حقوق اشخاص ثالث را تا حدی تضمین کند، اما مسئولیت داخلی ناشی از تصمیم معیوب همچنان نیازمند قواعد خاص خواهد بود.

مسئله مسئولیت یکی از جدی‌ترین موانع پذیرش نظریه است. Nersessian و Mancha توضیح می‌دهند که افزایش خودمختاری الگوریتمی می‌تواند فاصله میان رفتار سامانه و کنترل مستقیم انسان را افزایش دهد و در نتیجه شکاف مسئولیت ایجاد کند (Nersessian & Mancha, 2020). در ساختار هیئت‌مدیره انسانی، در صورت نقض وظایف قانونی یا اساسنامه‌ای، می‌توان رفتار مدیر را ارزیابی و حسب مورد مسئولیت او را احراز کرد. اما اگر سامانه در نتیجه یادگیری و پردازش مستقل به تصمیمی برسد که طراح، توسعه‌دهنده یا مدیر ناظر آن را مستقیماً پیش‌بینی نکرده است، معیار انتساب مسئولیت پیچیده می‌شود.

یکی از راهکارها، توزیع مسئولیت میان چند سطح است. شرکت می‌تواند در برابر اشخاص ثالث مسئول تصمیمات سامانه شناخته شود، در حالی که در روابط داخلی امکان رجوع به توسعه‌دهنده، تأمین‌کننده داده یا مدیر ناظر بر اساس منشأ خطا وجود داشته باشد. چنین نظامی باید میان خطای طراحی، خطای داده، خطای نظارت، نقص امنیتی و رفتار غیرقابل پیش‌بینی سامانه تفکیک کند. Kazim و همکاران بر اهمیت ارزیابی اثرات و حسابرسی هوش مصنوعی برای شناسایی منابع خطر و کاهش مخاطرات تصمیمات خودکار تأکید می‌کنند (Kazim et al., 2021). بدون چنین سازوکارهایی، واگذاری اختیار مدیریتی به سامانه می‌تواند نظام مسئولیت شرکت را با ابهام جدی مواجه سازد.

قابلیت توضیح تصمیم نیز در این چارچوب اهمیت دارد. مدیر انسانی را می‌توان ملزم کرد دلایل تصمیم خود را ارائه کند، اما برخی سامانه‌های پیچیده ممکن است خروجی‌هایی تولید کنند که منطق دقیق آن‌ها به‌آسانی قابل بازسازی نباشد. این وضعیت در محیط شرکت می‌تواند حقوق سهامداران، طلبکاران، کارکنان و سایر ذی‌نفعان را تحت تأثیر قرار دهد. اگر سامانه تصمیم به رد یک قرارداد، قطع همکاری با یک تأمین‌کننده یا تخصیص منابع به یک پروژه بگیرد، باید امکان ارزیابی حقوقی تصمیم وجود داشته باشد. ازاین‌رو، شفافیت و قابلیت ممیزی باید بخشی از شروط اعطای صلاحیت باشد.

مسئله دیگری که باید مورد توجه قرار گیرد، تعارض میان اهداف اقتصادی الگوریتم و الزامات حقوقی است. سامانه ممکن است برای بیشینه‌سازی سود، کاهش هزینه یا افزایش ارزش شرکت طراحی شود، اما تصمیمات شرکتی نمی‌توانند صرفاً بر اساس یک تابع اقتصادی اتخاذ شوند. الزامات قانونی، حقوق کارکنان، قواعد رقابت، تعهدات زیست‌محیطی و قواعد حمایت از داده نیز باید در فرآیند تصمیم‌گیری وارد شوند. Mirishli بر اهمیت چارچوب‌های حقوقی در شکل‌دهی استفاده اخلاقی و مسئولانه از هوش مصنوعی در حاکمیت شرکتی تأکید می‌کند (Mirishli, 2024). بنابراین، سامانه حاکم نمی‌تواند تنها «کارآمد» باشد؛ بلکه باید «قانون‌مند» نیز باشد.

پذیرش این نظریه در حقوق ایران در شرایط فعلی با محدودیت جدی روبه‌رو است. ساختار قانونی شرکت‌ها بر مبنای ارکان شناخته‌شده و صلاحیت‌های مشخص طراحی شده و جایگاهی برای یک رکن الگوریتمی مستقل پیش‌بینی نکرده است. Eskini نشان می‌دهد که ساختار شرکت تجاری و اختیارات ارکان آن بر مبنای قواعد قانونی نسبتاً صریح سازمان یافته است (Eskini, 2014). بنابراین، ایجاد رکن جدید یا جایگزینی هیئت‌مدیره با سامانه هوش مصنوعی صرفاً از طریق توافق خصوصی سهامداران قابل تحقق نیست و به مداخله قانون‌گذار نیاز دارد. بااین‌حال، این مانع را نباید به معنای ناممکن بودن نظریه دانست. قواعد حقوق شرکت‌ها در طول تاریخ بارها در پاسخ به تحولات اقتصادی تغییر کرده‌اند و می‌توانند در برابر تحول فناورانه نیز تطبیق یابند. راهکار مناسب آن است که پذیرش حاکمیت هوش مصنوعی به‌صورت تدریجی، محدود و آزمون‌پذیر انجام شود. قانون‌گذار می‌تواند ابتدا آن را در شرکت‌های کوچک با موضوع فعالیت محدود یا در قالب محیط‌های آزمایشی حقوقی مجاز بداند و پس از ارزیابی نتایج، دامنه کاربرد را توسعه دهد. چنین رویکردی امکان شناسایی مخاطرات واقعی را فراهم می‌کند و از تحمیل یک تحول دفعی به کل نظام حقوق شرکت‌ها جلوگیری خواهد کرد.

اصول و الزامات حقوقی حاکم بر حاکمیت هوش مصنوعی و الگوی شرکت هوشمند

پذیرش حاکمیت هوش مصنوعی بدون طراحی مجموعه‌ای از اصول الزام‌آور، می‌تواند به جای بهبود ساختار اداره شرکت، مخاطرات حقوقی و سازمانی جدیدی ایجاد کند. نخستین اصل، شفافیت است. شرکت باید مشخص کند که در چه بخش‌هایی از فرایند مدیریتی از هوش مصنوعی استفاده می‌شود، سطح خودمختاری سامانه چقدر است، چه نوع داده‌هایی در تصمیم‌گیری مورد استفاده قرار می‌گیرد و کدام تصمیمات مستقیماً توسط سامانه اتخاذ می‌شوند. Paunov و همکاران نشان داده‌اند که شفاف‌سازی سازوکار تصمیم می‌تواند بر پذیرش و انطباق با سیاست‌ها اثرگذار باشد (Paunov et al., 2019). در زمینه حاکمیت شرکتی نیز شفافیت پیش‌شرط نظارت مؤثر سهامداران و سایر ذی‌نفعان بر عملکرد سامانه است.

شفافیت در اینجا تنها به اعلام استفاده از هوش مصنوعی محدود نمی‌شود. باید امکان درک حدود صلاحیت سامانه، اهداف برنامه‌ریزی‌شده، معیارهای اصلی تصمیم‌گیری و روش ارزیابی عملکرد آن وجود داشته باشد. اگر یک سامانه اختیار اتخاذ تصمیمات مالی عمده را داشته باشد، سهامداران باید بدانند چه محدودیت‌هایی برای آن تعیین شده است و تحت چه شرایطی تصمیم آن قابل توقف یا بازنگری است. در نتیجه،

اساسنامه یا اسناد حاکمیت شرکتی باید نقش هوش مصنوعی را با دقت تعریف کنند و مرز میان صلاحیت سامانه و اختیار انسان را مشخص سازند.

اصل دوم، حاکمیت داده است. عملکرد سامانه هوش مصنوعی به کیفیت، صحت، جامعیت و مشروعیت داده‌هایی وابسته است که بر اساس آن‌ها آموزش می‌بیند یا تصمیم می‌گیرد. Kazim و همکاران در بحث حسابرسی هوش مصنوعی بر ضرورت بررسی کیفیت داده و آثار آن بر تصمیمات خودکار تأکید کرده‌اند (Kazim et al., 2021). داده ناقص یا سوگیرانه می‌تواند سامانه را به تصمیمی سوق دهد که از نظر تجاری نادرست یا از نظر حقوقی تبعیض‌آمیز باشد. بنابراین، شرکت باید سازوکارهایی برای تعیین منشأ داده، ارزیابی کیفیت، حفاظت امنیتی و کنترل دسترسی ایجاد کند.

حاکمیت داده همچنین با حفاظت از اطلاعات محرمانه شرکت پیوند دارد. سامانه حاکم ممکن است به گسترده‌ترین مجموعه داده‌های شرکت، از اطلاعات مالی و قراردادها تا اطلاعات کارکنان و مشتریان، دسترسی داشته باشد. چنین تمرکزی از اطلاعات، خطر امنیتی قابل توجهی ایجاد می‌کند. هرگونه نفوذ، دستکاری داده یا دسترسی غیرمجاز می‌تواند مستقیماً کیفیت تصمیمات مدیریتی را تحت تأثیر قرار دهد. بنابراین، امنیت سایبری باید به‌عنوان جزء جدایی‌ناپذیر حاکمیت هوش مصنوعی تلقی شود.

اصل سوم، حسابرسی مستمر است. سامانه‌ای که در زمان استقرار عملکرد مطلوبی دارد ممکن است در نتیجه تغییر محیط کسب‌وکار، داده‌های جدید، تغییر رفتار کاربران یا تحول الگوریتم دچار انحراف شود. Mökander و همکاران حسابرسی مبتنی بر اخلاق سامانه‌های تصمیم‌گیری خودکار را ابزاری برای ارزیابی مستمر ماهیت و آثار عملکرد آن‌ها می‌دانند (Mökander et al., 2021). در شرکت هوشمند، حسابرسی باید شامل ارزیابی دقت، قابلیت اتکا، انطباق حقوقی، امنیت، عدم تبعیض و هماهنگی عملکرد سامانه با منافع شرکت باشد.

این حسابرسی نمی‌تواند صرفاً به واحد فناوری اطلاعات واگذار شود. ماهیت حاکمیت هوش مصنوعی میان‌رشته‌ای است و نیازمند همکاری حقوق‌دانان، متخصصان فناوری، حسابرسان، مدیران ریسک و کارشناسان حوزه فعالیت شرکت است. ارزیابی فنی ممکن است نشان دهد که سامانه با دقت بالایی پیش‌بینی می‌کند، اما ارزیابی حقوقی می‌تواند آشکار سازد که برخی معیارهای تصمیم‌گیری آن با الزامات قانونی ناسازگار است. بنابراین، حسابرسی باید همزمان فنی، حقوقی و سازمانی باشد.

اصل چهارم، نظارت انسانی است. واگذاری تصمیم‌گیری به سامانه نباید به معنای حذف کامل ظرفیت مداخله انسان باشد. Koulu تأکید می‌کند که کنترل انسانی باید به‌گونه‌ای طراحی شود که شخص ناظر عملاً قادر به فهم و اصلاح عملکرد خودکار باشد، نه اینکه صرفاً به صورت صوری در ساختار باقی بماند Laux .(Koulu, 2020) نیز نظارت انسانی را عنصر بنیادین طراحی حکمرانی هوش مصنوعی می‌داند (Laux, 2024). بنابراین، شرکت هوشمند باید دارای سازوکاری باشد که امکان توقف، تعلیق یا بازنگری تصمیم سامانه را در شرایط خاص فراهم کند.

این نظارت باید مستقل باشد. اگر همان افرادی که مسئول طراحی یا استقرار سامانه هستند بر عملکرد آن نیز نظارت کنند، امکان بروز تعارض منافع افزایش می‌یابد. از این رو می‌توان ایجاد یک لایه مستقل نظارت انسانی را در شرکت هوشمند ضروری دانست. وظیفه این لایه نه اداره روزمره شرکت، بلکه کنترل انطباق تصمیمات هوش مصنوعی با قانون، اساسنامه، سیاست‌های شرکت و معیارهای ایمنی است. گزارش‌های دوره‌ای نیز باید به مجمع عمومی ارائه شود تا سهامداران از نحوه عملکرد سامانه آگاه باشند.

اصل پنجم، سواد هوش مصنوعی است. حتی اگر سامانه بخش عمده تصمیمات را اتخاذ کند، اشخاص انسانی همچنان باید توانایی فهم محدودیت‌ها و مخاطرات آن را داشته باشند. ناظر انسانی که فاقد دانش کافی درباره نحوه عملکرد الگوریتم باشد، نمی‌تواند نظارت مؤثری اعمال کند. این موضوع به‌ویژه در مواردی اهمیت دارد که سامانه خروجی پیچیده یا احتمالاتی تولید می‌کند. بنابراین، شرکت باید آموزش مستمر اعضای هیئت‌مدیره ناظر، مدیران حقوقی، حسابرسان و سایر اشخاص مرتبط را در نظر بگیرد.

اصل ششم، پایش مستمر تحولات حقوقی است. حوزه هوش مصنوعی به‌سرعت در حال تغییر است و قواعد ناظر بر آن ممکن است در دوره‌های کوتاه اصلاح شوند. Mirishli نقش چارچوب‌های حقوقی را در شکل‌دهی استفاده مسئولانه از هوش مصنوعی در محیط شرکتی برجسته می‌داند (Mirishli, 2024). بر این مبنا، واحد حقوقی شرکت نباید تنها در زمان استقرار سامانه، انطباق آن را بررسی کند، بلکه باید سازوکاری برای ارزیابی دوره‌ای تغییرات قانونی و تنظیم عملکرد سامانه با الزامات جدید داشته باشد.

در کنار نظارت انسانی، استفاده از «هوش مصنوعی ناظر» نیز می‌تواند به‌عنوان ابزار تکمیلی مطرح شود. در این مدل، یک سامانه مستقل عملکرد سامانه مدیریتی را پایش می‌کند و در صورت شناسایی رفتار غیرعادی، تخطی از قواعد یا افزایش ریسک، هشدار می‌دهد یا عملکرد را متوقف می‌سازد. تجربه‌هایی از طراحی سامانه‌های ایمنی برای کنترل سیستم‌های خودمختار در پروژه‌های آزمایشگاهی نشان می‌دهد که استفاده از لایه‌های خودکار نظارتی می‌تواند به افزایش ایمنی کمک کند (Johns Hopkins Applied Physics, 2023). بااین‌حال، هوش مصنوعی ناظر نباید جایگزین ناظر انسانی شود، بلکه باید مکمل آن باشد.

مسئله نظارت میان سامانه‌های هوشمند همچنین با بحث «ضعیف به قوی» ارتباط پیدا می‌کند. Burns در پژوهش خود درباره تعمیم از نظارت ضعیف به سامانه‌های قوی‌تر نشان می‌دهد که کنترل عملکرد مدل‌های پیشرفته توسط ناظری که از نظر توان تحلیلی ضعیف‌تر است، خود یک مسئله پیچیده محسوب می‌شود (Burns, 2023). در حوزه شرکت‌ها نیز ممکن است سامانه مدیریتی در تحلیل داده‌ها توانایی‌هایی فراتر از ناظران انسانی داشته باشد. بنابراین، نظارت انسانی صرفاً با افزودن یک شخص به ساختار تحقق نمی‌یابد و باید با ابزارهای فنی، گزارش‌دهی قابل فهم و سازوکارهای توقف همراه شود.

اصل بعدی، پاسخگویی است. حاکمیت هوش مصنوعی نباید به ایجاد «منطقه بی‌مسئولیتی» منجر شود. اگر سامانه تصمیمی اتخاذ کند که موجب زیان شرکت یا اشخاص ثالث شود، باید از پیش مشخص باشد که چه اشخاصی در چه شرایطی مسئول خواهند بود. Nersessian و Mancha هشدار می‌دهند که خودمختاری سامانه‌ها می‌تواند شکاف میان قدرت تصمیم‌گیری و مسئولیت را افزایش دهد (Nersessian & Mancha, 2020). بنابراین، قواعد مسئولیت باید به موازات افزایش خودمختاری توسعه یابند.

در این چارچوب، مفهوم «شرکت هوشمند» معنا پیدا می‌کند. شرکت هوشمند شرکتی نیست که صرفاً از نرم‌افزارهای پیشرفته یا الگوریتم‌های متعدد استفاده می‌کند. ویژگی اصلی آن این است که هوش مصنوعی در ساختار حقوقی اداره شرکت دارای صلاحیت مستقل است و می‌تواند بخشی قابل توجه از تصمیمات مدیریتی را اتخاذ کند. Mirishli تحول در حاکمیت شرکتی را در ارتباط با چارچوب حقوقی استفاده از هوش مصنوعی تحلیل می‌کند و نشان می‌دهد که استفاده اخلاقی و قابل اعتماد از این فناوری نیازمند طراحی نهادی است (Mirishli, 2024). از این جهت، شرکت هوشمند یک قالب سازمانی جدید و نه صرفاً یک شرکت فناورانه است.

در چنین شرکتی، نقش انسان از تصمیم‌گیری روزمره به تعیین هدف، نظارت و کنترل تغییر می‌کند. ممکن است مجمع عمومی اهداف کلی و حدود ریسک قابل پذیرش را تعیین کند، سامانه هوش مصنوعی تصمیمات عملیاتی را اتخاذ کند و یک هیئت نظارت انسانی بر انطباق عملکرد

سامانه با قانون و منافع شرکت نظارت داشته باشد. این ساختار با مفهوم «شرکت خودران» نزدیک است، اما خودران بودن نباید به معنای فقدان هرگونه عنصر انسانی فهمیده شود. خودمختاری در اینجا به محل اتخاذ تصمیمات مدیریتی مربوط است، نه به حذف کامل انسان از سازمان. اجرای عملی چنین الگویی باید تدریجی باشد. در شرایط فعلی، مناسب‌ترین محیط برای آزمون این نظریه شرکت‌های کوچک با موضوع فعالیت مشخص و فرایندهای نسبتاً قابل پیش‌بینی هستند. در این شرکت‌ها دامنه تصمیمات محدودتر است و می‌توان عملکرد سامانه را با سهولت بیشتری ارزیابی کرد. پس از شناسایی مزایا و مخاطرات، امکان توسعه مدل به شرکت‌های پیچیده‌تر فراهم خواهد شد. چنین رویکردی حقوق شرکت‌ها را قادر می‌سازد بدون پذیرش دفعی یک ساختار ناشناخته، از طریق تجربه تدریجی به چارچوبی پایدار برای حاکمیت هوش مصنوعی دست یابد.

نتیجه‌گیری

تحول هوش مصنوعی در محیط شرکت‌ها، حقوق شرکت را با پرسشی بنیادین مواجه ساخته است که پاسخ به آن از حدود استفاده از ابزارهای فناورانه فراتر می‌رود. مسئله اصلی دیگر این نیست که شرکت‌ها تا چه اندازه از هوش مصنوعی برای تحلیل داده، پیش‌بینی بازار یا تسهیل فعالیت‌های مدیریتی استفاده خواهند کرد، بلکه مسئله تعیین جایگاه این فناوری در ساختار حقوقی تصمیم‌گیری و اداره شرکت است. با افزایش سطح خودمختاری سامانه‌ها، امکان انتقال بخشی از وظایفی که به‌طور سنتی در صلاحیت مدیران انسانی قرار داشته‌اند، به سامانه‌های هوشمند تقویت می‌شود و از این رهگذر ضرورت بازاندیشی در مفاهیم بنیادی حاکمیت شرکتی آشکار می‌گردد.

نظریه حاکمیت هوش مصنوعی در شرکت‌ها بر تمایز میان استفاده از هوش مصنوعی به‌عنوان ابزار و شناسایی آن به‌عنوان عنصر حاکمیتی استوار است. در مرحله نخست، هوش مصنوعی می‌تواند نقش کمکی یا تقویتی داشته باشد و اطلاعات، تحلیل یا توصیه در اختیار مدیر قرار دهد. در چنین وضعیتی، تصمیم نهایی همچنان به شخص انسانی منتسب است و ساختار حقوقی سنتی حاکمیت شرکتی حفظ می‌شود. اما در سطح خودمختار، سامانه می‌تواند در محدوده‌ای از صلاحیت‌های از پیش تعیین‌شده، تصمیم را بدون تأیید موردی انسان اتخاذ و اجرا کند. نقطه آغاز نظریه حاکمیت هوش مصنوعی نیز همین انتقال اختیار از انسان به سامانه است.

در چارچوب این نظریه می‌توان دو الگوی نهادی اصلی را تصور کرد. در الگوی نخست، سامانه هوش مصنوعی جایگزین هیئت‌مدیره می‌شود و اداره شرکت، جز در اموری که قانوناً یا ماهیتاً در صلاحیت مجمع عمومی است، به سامانه واگذار می‌گردد. در الگوی دوم، یک لایه مستقل حاکمیت هوش مصنوعی در کنار ارکان سنتی شرکت ایجاد می‌شود و هیئت‌مدیره انسانی از یک رکن تصمیم‌گیرنده به رکن ناظر تبدیل می‌شود. الگوی دوم از منظر گذار تدریجی عملی‌تر به نظر می‌رسد؛ زیرا امکان حفظ کنترل انسانی و ارزیابی آثار حقوقی تصمیمات سامانه را فراهم می‌کند.

مانع اصلی پذیرش این نظریه در شرایط فعلی، فقدان موقعیت حقوقی مشخص برای سامانه هوش مصنوعی است. حقوق شرکت‌ها برای توزیع صلاحیت، انتساب آثار تصمیم و تعیین مسئولیت به مفهوم شخص متکی است، در حالی که سامانه هوشمند در وضعیت کنونی نه شخص حقیقی است و نه در شمار اشخاص حقوقی شناخته‌شده قرار دارد. از این رو، اعطای مستقیم سمت مدیریتی به آن بدون اصلاح قانونی با دشواری مواجه خواهد شد. حل این مسئله می‌تواند از طریق شناسایی نوعی شخصیت حقوقی محدود، ایجاد ظرفیت حقوقی خاص برای سامانه‌های هوشمند یا بازتعریف مفهوم مدیر و رکن مدیریتی صورت گیرد. انتخاب هر یک از این مسیرها باید همراه با قواعد روشن درباره حدود صلاحیت، آثار تصمیمات و مسئولیت باشد.

پذیرش حاکمیت هوش مصنوعی بدون استقرار اصول حقوقی مکمل قابل دفاع نیست. شفافیت باید تضمین کند که نقش سامانه و حدود اختیار آن برای اشخاص ذی‌نفع قابل شناسایی باشد. حاکمیت داده باید کیفیت، صحت، امنیت و مشروعیت داده‌های مورد استفاده را تضمین کند. حسابرسی مستمر باید امکان شناسایی انحراف عملکرد، خطا، سوگیری و عدم انطباق را فراهم سازد. نظارت انسانی نیز باید واقعی و مؤثر باشد و ناظر بتواند در صورت ضرورت عملکرد سامانه را متوقف، اصلاح یا محدود کند. در کنار این اصول، آموزش اشخاص ناظر، پایش مستمر تحولات قانونی، امنیت سایبری و طراحی سازوکارهای روشن مسئولیت از عناصر ضروری یک نظام حقوقی قابل اعتماد خواهند بود.

شرکت هوشمند را باید نتیجه نهادی پذیرش نظریه حاکمیت هوش مصنوعی دانست. در این شرکت، هوش مصنوعی صرفاً زیرساخت فنی یا ابزار مدیریتی نیست، بلکه بخشی از سازمان حقوقی اداره شرکت به‌شمار می‌رود. بااین‌حال، شرکت هوشمند لزوماً شرکتی بدون انسان نیست. نقش انسان در چنین ساختاری از تصمیم‌گیرنده روزمره به تعیین‌کننده اهداف، ناظر، کنترل‌کننده و مداخله‌گر نهایی تغییر می‌کند. این تمایز اهمیت زیادی دارد؛ زیرا هدف نظریه، حذف انسان از ساختار شرکت نیست، بلکه بازتوزیع منطقی و کنترل‌شده صلاحیت میان انسان و سامانه هوشمند است.

در حقوق ایران، پذیرش فوری و کامل چنین ساختاری با قواعد موجود سازگار نیست و نیازمند مداخله قانون‌گذار است. بااین‌حال، به‌جای رد کامل این الگو یا پذیرش دفعی آن، راهکار مناسب‌تر ایجاد یک چارچوب آزمایشی و تدریجی است. قانون‌گذار می‌تواند ابتدا امکان تشکیل یا فعالیت شرکت‌های هوشمند را در حوزه‌های محدود، شرکت‌های کوچک و فعالیت‌هایی با ساختار عملیاتی ساده مورد شناسایی قرار دهد. در این مرحله می‌توان الزامات فنی، سرمایه‌ای، نظارتی و بیمه‌ای ویژه‌ای در نظر گرفت و عملکرد این شرکت‌ها را تحت نظارت مستمر قرار داد.

تجربه حاصل از چنین محیط آزمایشی می‌تواند به شناسایی دقیق‌تر مخاطرات و مزایای مدل کمک کند. اگر مشخص شود که سامانه‌های هوشمند در برخی حوزه‌ها می‌توانند با سطح قابل قبول ایمنی، شفافیت و انطباق حقوقی اداره شرکت را بر عهده گیرند، دامنه فعالیت آن‌ها به‌تدریج قابل توسعه خواهد بود. در مقابل، حوزه‌هایی که تصمیمات آن‌ها دارای آثار اجتماعی گسترده یا ریسک بالا هستند می‌توانند همچنان تحت محدودیت شدیدتر قرار گیرند. چنین رویکرد مبتنی بر ریسک، از یک سو مانع توقف نوآوری می‌شود و از سوی دیگر اجازه نمی‌دهد توسعه فناوری بدون چارچوب حقوقی مناسب صورت گیرد.

در نهایت، نظریه حاکمیت هوش مصنوعی در شرکت‌ها را باید پیشنهادی برای بازاندیشی در یکی از مفروضات بنیادین حقوق شرکت‌ها دانست؛ فرضی که اداره شرکت را همواره وابسته به تصمیم‌گیرنده انسانی یا شخص حقوقی متعارف تلقی کرده است. پیشرفت فناوری این امکان را ایجاد کرده است که در آینده سامانه‌هایی بتوانند بخشی از وظایف مدیریتی را با سرعت، پیوستگی و ظرفیت تحلیلی بسیار بیشتر از مدیران انسانی انجام دهند. بااین‌حال، توانایی فنی به‌تنهایی برای ایجاد صلاحیت حقوقی کافی نیست. تبدیل هوش مصنوعی از ابزار مدیریت به رکن حاکمیتی تنها زمانی قابل پذیرش است که قانون، حدود اختیار، نظارت، مسئولیت، شفافیت و سازوکار پاسخگویی آن را به‌صورت دقیق تنظیم کند.

بر این اساس، نظریه حاکمیت هوش مصنوعی در شرکت‌ها نه توصیف یک واقعیت حقوقی موجود، بلکه چارچوبی آینده‌نگر برای مواجهه با تحولی است که احتمالاً ساختار سنتی اداره شرکت‌ها را تحت تأثیر قرار خواهد داد. ارزش این نظریه در آن است که به‌جای انتظار برای تحقق

کامل شرکت‌های خودمختار و سپس تلاش برای حل مسائل حقوقی ناشی از آن، از هم‌اکنون مبانی لازم برای تنظیم این تحول را مورد بررسی قرار می‌دهد. در صورت ایجاد چارچوب تقنینی مناسب، شناسایی محدود و تدریجی موقعیت حقوقی سامانه‌های هوشمند و تضمین نظارت انسانی مؤثر، می‌توان از ظرفیت هوش مصنوعی برای ایجاد شکل جدیدی از حاکمیت شرکتی بهره گرفت؛ شکلی که در آن فناوری نه صرفاً ابزار مدیر، بلکه یکی از عناصر سازمان حقوقی شرکت خواهد بود.

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

Rapid advances in artificial intelligence have begun to transform not only the technological infrastructure of companies but also the conceptual foundations of corporate administration, decision-making, and governance. The growing ability of AI systems to process large volumes of data, identify patterns, generate forecasts, recommend strategic alternatives, and increasingly perform autonomous tasks creates a need to reconsider the traditional assumption that corporate management must necessarily be exercised by human directors or conventional legal persons. Recent developments in European company law indicate that digital transformation can no longer be reduced to the electronic transmission of information or the automation of administrative procedures, because artificial intelligence increasingly affects the substantive architecture of corporate governance and managerial decision-making (Möslein, 2025). Earlier discussions on the digitalization of company law largely focused on electronic communications, digital procedures, and the modernization of interactions among corporate organs (Knapp, 2016), while analyses of company law in the digital age likewise examined the gradual incorporation of technological tools into traditional corporate structures (Bertschinger, 2015). Artificial intelligence, however, introduces a qualitatively different challenge because advanced systems may gradually move from assisting corporate decision-makers to exercising functions that resemble autonomous managerial authority. This distinction provides the central theoretical foundation of the present study. The article introduces the theory of artificial intelligence governance in companies as a framework for examining the transition from “corporate governance with artificial intelligence,” in which AI remains an instrument supporting human directors, to “artificial intelligence governance,” in which an AI system itself occupies a legally recognized position in the corporate decision-making structure. The objective is therefore not merely to examine the growing use of AI tools in companies, but to determine whether corporate law can conceptually and legally accommodate a future structure in which significant managerial powers are delegated to autonomous systems and whether such a transformation would require the recognition of a new type of corporate organization.

The conceptual foundation of the proposed theory rests on differentiating levels of artificial intelligence according to the degree of autonomy exercised in corporate decision-making. At the first level, assistive artificial intelligence performs limited supporting functions such as information retrieval, document analysis, classification, risk detection, financial analysis, or administrative automation, while the ultimate decision remains entirely with human corporate organs. At the

second level, augmented artificial intelligence becomes more deeply integrated into managerial processes by continuously producing recommendations, forecasts, risk assessments, strategic scenarios, and decision alternatives, although human directors retain final authority. Hilb's analysis of artificial governance demonstrates that AI may progressively influence the future structure of corporate governance and that its role can extend well beyond simple technological assistance (Hilb, 2020). Human control remains especially significant at these two levels because the legal act of decision-making continues to be attributable to a human director or another recognized corporate organ; Koulu emphasizes that meaningful human control over automated systems must involve the real capacity to understand, evaluate, and intervene in automated processes rather than merely preserving nominal human participation (Koulu, 2020). The third level, however, is autonomous artificial intelligence, in which a system is authorized to select and implement decisions without case-by-case human approval. This stage fundamentally changes the legal analysis because autonomy separates the generation of a decision from direct human authorization and may create substantial gaps in attribution and responsibility (Nersessian & Mancha, 2020). The technological basis of such autonomy is increasingly plausible because contemporary AI combines advanced machine learning with generative capabilities. Machine-learning systems can identify complex patterns and generate predictive outputs, while generative AI can create new solutions, proposals, and strategic content on the basis of available information (Wohlfarth, 2024). Generative systems therefore move beyond traditional predictive tools by producing new outputs and alternative courses of action rather than merely classifying existing data (Marr, 2023). The theory developed in this article concerns primarily this third level, because only when AI acquires independent decision-making authority does it become possible to speak of a genuine transition from governance supported by artificial intelligence to governance exercised by artificial intelligence.

On this basis, the theory of artificial intelligence governance in companies may be institutionalized through two principal models. In the first model, artificial intelligence replaces the conventional board of directors and exercises those managerial powers that are not reserved exclusively for the general meeting of shareholders. Under such an arrangement, one integrated AI system or a coordinated group of intelligent systems could continuously monitor the company's financial, operational, contractual, and strategic environment, evaluate alternative courses of action, adopt managerial decisions, and initiate their implementation. In the second model, the traditional board is not formally abolished, but an independent AI governance layer is established alongside existing corporate organs, with substantial managerial powers transferred to that layer while the human board is transformed primarily into a supervisory institution. The second model may offer a more realistic transitional mechanism because it permits gradual delegation while maintaining a clearly identifiable human oversight structure. However, both models require a distinction between technical autonomy and legal authority. A system may be technically capable of making decisions, but unless corporate law recognizes its competence, its outputs remain legally attributable to another person or corporate organ. This problem leads directly to the issue of legal personality. Traditional corporate law is constructed around persons who can hold rights, exercise powers, incur duties, and bear responsibility. In Iranian corporate law, although legal persons may under certain circumstances serve as directors, the legal framework nevertheless presupposes that the director must possess legally recognized personhood, and the conventional structure of commercial companies remains organized around legally defined corporate organs (Eskini, 2014). Comparative analysis also demonstrates that even the participation of legal persons in corporate management has historically generated conceptual difficulties, showing how deeply the concept of the director is tied

to conventional legal personality (Gehrlein, 2016). Accordingly, the recognition of autonomous AI as a corporate governor would require either an extension of existing concepts or the creation of a new legal status specifically adapted to intelligent systems.

The legal-personality problem does not necessarily require the recognition of AI as a complete legal person equivalent to a human being or a corporation. A more limited and functional approach may be preferable. Recent scholarship has questioned whether traditional understandings of legal personality remain fully adequate in light of increasingly autonomous artificial intelligence and has suggested that legal systems may eventually need to reconsider the criteria through which legal capacity is allocated (Amini & Kalantarzadeh Sadabad, 2026). One possible approach would be to recognize a form of limited or special-purpose legal capacity that allows a qualifying AI system to exercise defined managerial powers without granting it the full bundle of rights traditionally associated with natural or legal persons. Boudraa's concept of temporary legal personality for artificial intelligence illustrates how legal status might be used functionally to facilitate accountability without necessarily recognizing complete personhood (Boudraa, 2025). Within corporate law, such an approach could take the form of limited managerial personality, under which the AI system would possess only those rights, competencies, and duties necessary for its defined corporate role. Alternatively, legislation could amend the legal definition of a corporate director or governance organ so that qualifying artificial intelligence systems may exercise specifically enumerated powers without being recognized as full legal persons. Both approaches would nevertheless require a carefully designed liability framework because increasing autonomy can create a gap between those who design, deploy, supervise, and benefit from an AI system and the system that actually produces the disputed decision (Nersessian & Mancha, 2020). Liability may therefore need to be distributed among the company, AI provider, developers, data suppliers, and human supervisors according to the source of the failure. AI auditing and impact assessment become particularly important in this context because they can help determine whether harm arose from defective data, faulty system design, inadequate supervision, discriminatory outputs, or failures of compliance (Kazim et al., 2021). The theory proposed in this article thus rejects the view that technological capacity alone can justify the transfer of managerial authority; such transfer is legally acceptable only when the institutional status, scope of competence, attribution of decisions, and responsibility structure of the AI system are clearly defined.

The recognition of AI governance also requires a coherent set of substantive principles capable of controlling the risks associated with autonomous corporate decision-making. Transparency constitutes the first of these principles because shareholders and other affected stakeholders must know where artificial intelligence is used, what degree of authority it possesses, what objectives govern its decisions, and how its outputs affect corporate conduct. Research on transparency indicates that disclosure of decision mechanisms can significantly influence compliance and trust, suggesting that transparency should be treated as a functional element of accountability rather than merely as an informational formality (Paunov et al., 2019). A second principle is data governance because the quality, legality, security, and representativeness of data directly affect the reliability of AI decisions; auditing and impact-assessment frameworks consequently emphasize the need to evaluate the data environment as an integral part of responsible AI governance (Kazim et al., 2021). A third principle is continuous auditing. Autonomous systems should not be evaluated only before deployment because changing datasets, market environments, user behavior, or system updates may alter performance over time. Ethics-based auditing of automated decision-making systems therefore provides an important mechanism for detecting emerging errors, bias, and non-compliance

(Mökander et al., 2021). A fourth principle is effective human oversight. Even where an AI system is legally empowered to make decisions, human supervisors should retain the practical ability to intervene, suspend, correct, or terminate the system under defined circumstances; Laux's analysis of institutionalized distrust and human oversight supports the design of governance structures in which oversight is treated as a substantive safeguard rather than a symbolic requirement (Laux, 2024). Human oversight may also be reinforced by technological supervision in the form of a separate monitoring AI. Research and experimental projects concerning autonomous-system safety demonstrate the potential value of independent automated safety layers capable of identifying abnormal or dangerous behavior (Johns Hopkins Applied Physics, 2023). Nevertheless, monitoring one advanced system with another raises additional supervisory difficulties, especially when the monitored model may possess capabilities exceeding those of its human supervisors, a challenge explored in work on weak-to-strong generalization (Burns, 2023). For this reason, AI-based supervision should supplement rather than eliminate human review. Legal compliance must likewise remain dynamic because corporate AI systems operate in a rapidly changing regulatory environment, and legal frameworks have a central role in shaping ethical and responsible artificial intelligence use within corporate governance (Mirishli, 2024). Taken together, transparency, data governance, continuous auditing, human oversight, AI literacy, regulatory monitoring, cybersecurity, and clearly allocated responsibility constitute the minimum normative architecture within which autonomous corporate governance could become legally defensible.

In conclusion, the theory of artificial intelligence governance in companies should be understood as a forward-looking proposal for rethinking the traditional allocation of managerial authority within corporate law rather than as a description of a legal reality that has already been fully realized. Its central claim is that the development of increasingly autonomous systems may eventually make it possible to transfer substantial corporate decision-making powers from human directors to artificial intelligence, but such a transfer cannot be justified solely by technological capability. A meaningful distinction must be maintained between companies that merely automate their operations and companies in which artificial intelligence possesses legally recognized managerial authority. The latter may appropriately be described as smart or self-driving companies because AI becomes part of the company's legal governance structure rather than merely part of its technological infrastructure. Two principal institutional models are conceivable: replacement of the board of directors by an AI governance system, or creation of an independent AI governance layer operating under human supervision. In both models, human participation should not disappear completely; rather, its function would shift from continuous operational decision-making toward defining corporate objectives, supervising compliance, setting risk limits, and exercising emergency intervention powers. Under current Iranian law, the immediate recognition of such a structure remains difficult because artificial intelligence lacks an independent legal status and existing company law does not expressly recognize an algorithmic governance organ. Accordingly, any future adoption of the theory would require legislative intervention addressing legal capacity, managerial competence, liability, transparency, data governance, auditing, security, and supervision. A gradual model of legal experimentation would be preferable to immediate unrestricted recognition. Smart companies could initially be permitted only in narrowly defined sectors, small enterprises, or activities with relatively predictable operational structures, allowing regulators to observe their legal and economic effects before extending the model more broadly. The ultimate significance of the theory lies in identifying a possible future stage in corporate evolution in which artificial intelligence moves from the periphery of corporate administration to the legal core of corporate governance,

while remaining subject to a carefully designed framework that preserves accountability, human oversight, and the protection of corporate and stakeholder interests.

References

Amini, M., & Kalantarzadeh Sadabad, S. (2026). Reconsidering the Traditional Concept of Legal Personality in Light of Developments in Artificial Intelligence. *Legal Studies in Digital Age*.

Bertschinger, U. (2015). Aktienrecht im digitalen Zeitalter. In *Recht im digitalen Zeitalter: Festgabe Schweizerischer Juristentag 2015 in St. Gallen* (pp. 167-202).

Boudraa, K. (2025). Temporary Legal Personality for Artificial Intelligence: A Novel Legal Construct for Accountability. https://ssrn.com/abstract=5349502

Burns, C. (2023). Weak-to-Strong Generalization. https://openai.com/research/weak-to-strong-generalization

Eskini, R. (2014). *Commercial Law: Commercial Companies* (20th ed., Vol. 1). SAMT.

Gehrlein, M. (2016). Leitung einer juristischen Person durch juristische Person? *Neue Zeitschrift für Gesellschaftsrecht (NZG)*.

Hilb, M. (2020). Toward Artificial Governance? The Role of Artificial Intelligence in Shaping the Future of Corporate Governance. *Journal of Management and Governance*, *24*, 851-870.

Johns Hopkins Applied Physics, L. (2023). Johns Hopkins APL Creates Autonomy Safety Net for AFWERX Autonomy Prime Program. https://www.jhuapl.edu/news/news releases/230815-afwerx-autonomy-prime safety

Kazim, E., Denny, D. M. T., & Koshiyama, A. (2021). AI Auditing and Impact Assessment: According to the UK Information Commissioner's Office. *Ai and Ethics*, *1*.

Knapp, V. (2016). *What Are the Issues Relating to Digitalisation in Company Law?*

Koulu, R. (2020). Human Control over Automation: EU Policy and AI Ethics. *European Journal of Legal Studies*, *12*(1), 9-46.

Laux, J. (2024). Institutionalised Distrust and Human Oversight of Artificial Intelligence: Towards a Democratic Design of AI Governance under the European Union AI Act. *Ai & Society*, *39*(6), 2853-2866.

Marr, B. (2023). The Difference between Generative AI and Traditional AI: An Easy Explanation for Anyone. https://www.forbes.com/sites/bernardmarr/2023/07/24/the-difference-between-generative-ai-and-traditional-ai-an-easy-explanation-for-anyone/

Mirishli, S. (2024). The Role of Legal Frameworks in Shaping Ethical Artificial Intelligence Use in Corporate Governance. *International Law and Integration Problems*(68), 52-65.

Mökander, J., Morley, J., Taddeo, M., & Floridi, L. (2021). Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations. *Science and Engineering Ethics*, *27*(4).

Möslein, F. (2025). European Company Law in the Age of Digital Transformation: Artificial Intelligence, Data Economy, and Digital Corporate Governance. *Era Forum*, *26*, 629-646.

Nersessian, D., & Mancha, R. (2020). From Automation to Autonomy: Legal and Ethical Responsibility Gaps in Artificial Intelligence Innovation. *Michigan Technology Law Review*, *27*(1).

Paunov, Y., Wänke, M., & Vogel, T. (2019). Transparency Effects on Policy Compliance: Disclosing How Defaults Work Can Enhance Their Effectiveness. *Behavioural Public Policy*, *3*(2).

Wohlfarth, A. (2024). Machine Learning vs. Generative AI: A Detailed Look. https://www.ourcrowd.com/learn/machine-learning-vs-generative-ai