Journal of Public Administration
Print ISSN: 2008-5877
Online ISSN: 2423-5342

Structural Modeling of Responsible Artificial Intelligence-Based Human Resource Management and Its Implications for Employee Flourishing

Abbasali Rastgar *
*Corresponding Author, Prof., Department of Management, Faculty of Economics, Management and Administrative Sciences, Semnan University, Semnan, Iran. E-mail: a_rastgar@semnan.ac.ir

Minasadat Mousavi
Ph.D., Department of Management, Faculty of Economics, Management and Administrative Sciences, Semnan University, Semnan, Iran. E-mail: minamousavi@semnan.ac.ir

Abstract

Objective
This study aimed to identify, explain, and model the structural and hierarchical relationships between the components of responsible Artificial Intelligence (AI) in Human Resource Management (HRM) and their subsequent outcomes on employee flourishing within digital organizations. The core research problem addresses a critical gap in the extant literature: with the rapid integration of algorithms and intelligent decision-support systems into HRM, most organizations have disproportionately focused on technical efficiency, processing speed, and cost reduction, thereby neglecting the ethical dimensions and human consequences of these technologies. This study seeks to demonstrate how principles such as algorithmic transparency, accountability, fairness, privacy, and psychological safety can lead to positive work experiences and optimal employee outcomes within the context of human-machine interaction. Consequently, this research aims to explain how abstract AI ethical principles translate into psychological and behavioral outcomes for employees, thereby bridging the theoretical gap between normative AI principles and sustainable HRM practices.

Methods
Utilizing an exploratory-sequential mixed-methods design, this study was conducted within a pragmatic research paradigm. In the qualitative phase, semi-structured interviews were conducted with 17 dual-domain experts, comprising HRM specialists and AI or digital transformation experts, who were selected through purposive criterion and snowball

Citation: Rastgar, Abbasali & Mousavi, Minasadat (2026). Structural Modeling of Responsible Artificial Intelligence-Based Human Resource Management and Its Implications for Employee Flourishing. *Journal of Public Administration*, 18(3), 775-804 (*in Persian*)

Journal of Public Administration, 2026, Vol. 18, No.3, pp. 775-804
Published by University of Tehran, College of Management
https://doi.org/10.22059/jipa.2026.413478.3897
Article Type: Research Paper

Received: March 03, 2026
Revised: April 29, 2026
Accepted: July 12, 2026
Published online: September 14, 2026

sampling techniques. The qualitative data were analyzed using thematic analysis to extract relevant themes and dimensions. In the second (quantitative) phase, Interpretive Structural Modeling (ISM) was employed to map the hierarchical structure and interactions among the identified components. Prior to constructing the Structural Self-Interaction Matrix (SSIM), the dimensions and themes extracted from the qualitative phase were re-evaluated and validated by an expert panel in order to prevent the imposition of researcher bias and ensure the credibility of the findings. Furthermore, the analysis was grounded in Socio-Technical Systems (STS) theory to elucidate the alignment between the technical subsystem, comprising algorithms and infrastructure, and the social subsystem, encompassing humans and organizational interactions, in the implementation of responsible AI.

Results

The qualitative phase identified 15 sub-themes, which were subsequently categorized under the broader dimensions of responsible AI and employee flourishing. The ISM analysis revealed that implementing responsible AI in HR processes is not a one-dimensional technological intervention, but rather a multi-layered, systemic process structured across eight hierarchical levels. At the foundational levels, namely, the governance and technical layers, variables such as ethical policy-making, robust data infrastructure, and fair algorithmic design function as primary drivers. These foundational variables subsequently influence the middle layers, corresponding to the socio-psychological subsystem, by fostering organizational trust, enhancing perceived procedural justice, and mitigating ethical risks. Ultimately, at the highest level of the model, designated as the outcome subsystem, employee flourishing is realized through constructs such as meaning at work, active engagement, personal growth, and optimal performance. The findings collectively confirm that social and ethical components play a crucial mediating role in translating technical capabilities into meaningful human outcomes.

Conclusion

By proposing a three-layered structural-interpretive model, comprising governing-technical, socio-psychological, and outcome layers, this study addresses a significant gap in the literature at the intersection of responsible AI and HRM. The results demonstrate that the success and sustainability of intelligent technologies in organizations rely fundamentally on the joint optimization of both technical and social subsystems. In the absence of adequate attention to the psychological and ethical needs of employees, AI implementation may trigger work alienation and decrease organizational commitment, thereby undermining the very objectives it seeks to achieve. The proposed model serves as a strategic roadmap for organizational policymakers, HR managers, and system developers seeking to design human-centered technological solutions and promote human resource flourishing in the era of digital transformation. These insights offer both theoretical contributions and practical guidance for organizations striving to balance technological advancement with ethical responsibility and employee well-being.

Keywords: Responsible artificial intelligence, Employee thriving, Transparency and accountability, Organizational trust, Socio-technical systems theory.

مدیریت دولتی
شاپا چاپی: ۲۰۰۸-۵۸۷۷
شاپا الکترونیکی: ۲۴۲۳-۵۳۴۲

مدل‌سازی ساختاری مدیریت منابع انسانی مبتنی بر هوش مصنوعی مسئولانه و پیامدهای آن بر شکوفایی کارکنان

عباسعلی رستگار*

* نویسنده مسئول، استاد، گروه مدیریت، دانشکدهٔ اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان، سمنان، ایران. رایانامه: a_rastgar@semnan.ac.ir

میناسادات موسوی

دکتری، گروه مدیریت، دانشکده اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان، سمنان، ایران. رایانامه: minamousavi@semnan.ac.ir

چکیده

هدف: پژوهش حاضر با هدف شناسایی، تبیین و مدل‌سازی روابط ساختاری و سلسله‌مراتبی میان مؤلفه‌های مدیریت منابع انسانی مبتنی بر هوش مصنوعی مسئولانه و پیامدهای آن بر شکوفایی کارکنان در بستر سازمان‌های دیجیتال انجام شد. مسئلهٔ اصلی پژوهش این است که با ورود پُرشتاب الگوریتم‌ها و سامانه‌های تصمیم‌یار هوشمند به قلمرو مدیریت منابع انسانی، بخش عمده‌ای از سازمان‌ها، فقط بر کارایی فنی، سرعت و کاهش هزینه‌های عملیاتی متمرکز شده و از ابعاد اخلاقی و پیامدهای انسانی این فناوری‌ها غفلت ورزیده‌اند. این مطالعه به‌دنبال آن است تا نشان دهد که چگونه اصولی مانند شفافیت الگوریتمی، پاسخ‌گویی، عدالت، حریم خصوصی و امنیت روان‌شناختی، می‌توانند در تعامل انسان ـ ماشین، به ارتقای تجربهٔ مثبت کاری و عملکرد بهینهٔ کارکنان منجر شوند. در واقع، این پژوهش تلاش می‌کند تا نحوهٔ انتقال اصول اخلاقی و فنی هوش مصنوعی به پیامدهای روان‌شناختی و رفتاری کارکنان را تبیین کند و با ارائهٔ چارچوبی منسجم، به پُرکردن شکاف نظری موجود میان اصول انتزاعی هوش مصنوعی مسئولانه و مدیریت منابع انسانی پایدار یاری دهد.

روش: این پژوهش از نظر جهت‌گیری کاربردی و از نظر استراتژی اجرا، با رویکرد آمیخته اکتشافی ـ تبیینی (متوالی) در قالب پارادایم پراگماتیسم انجام شد. در مرحلهٔ کیفی، به‌منظور استخراج مؤلفه‌ها و ابعاد اولیه، مصاحبه‌های نیمه‌ساختاریافته با ۱۷ نفر از خبرگان دوحوزه‌ای (متخصصان مدیریت منابع انسانی و متخصصان هوش مصنوعی و تحول دیجیتال) که به‌روش نمونه‌گیری هدفمند معیارمحور و گلولهٔ برفی انتخاب شده بودند، انجام گرفت. تحلیل داده‌های کیفی با استفاده از روش تحلیل مضمون تأملی انجام شد. در مرحلهٔ دوم (کمی)، برای کشف روابط ساختاریافته و سطح‌بندی مؤلفه‌ها، تکنیک مدل‌سازی ساختاری تفسیری به‌کار گرفته شد. ابعاد و مضامین استخراج‌شده از مرحلهٔ کیفی، پیش از ورود به ماتریس خودتعاملی ساختاری، در پنل خبرگان بار دیگر ارزیابی و تأیید شدند تا از عدم تحمیل پیش‌فرض‌های پژوهشگر اطمینان حاصل شود. همچنین تحلیل‌ها با تکیه بر نظریهٔ سیستم‌های اجتماعی ـ فنی صورت گرفت تا پیوند میان زیرسیستم فنی (الگوریتم‌ها و زیرساخت‌ها) و زیرسیستم اجتماعی (انسان‌ها و تعاملات سازمانی) در استقرار هوش مصنوعی مسئولانه مشخص شود.

استناد: رستگار، عباسعلی و موسوی، میناسادات (۱۴۰۵). مدل‌سازی ساختاری مدیریت منابع انسانی مبتنی بر هوش مصنوعی مسئولانه و پیامدهای آن بر شکوفایی کارکنان. *مدیریت دولتی*، ۱۸(۳)، ۷۷۵-۸۰۴.

مدیریت دولتی، ۱۴۰۵، دوره ۱۸، شماره ۳، صص. ۷۷۵-۸۰۴
ناشر: دانشکدگان مدیریت، دانشگاه تهران
نوع مقاله: علمی پژوهشی

تاریخ دریافت: ۱۴۰۴/۱۲/۱۲
تاریخ ویرایش: ۱۴۰۵/۰۲/۰۹
تاریخ پذیرش: ۱۴۰۵/۰۴/۲۱
تاریخ انتشار: ۱۴۰۵/۰۶/۲۳
doi: https://doi.org/10.22059/jipa.2026.413478.3897

یافته‌ها: یافته‌های بخش کیفی نشان‌دهندهٔ شناسایی ۱۵ مضمون فرعی در قالب ابعاد هوش مصنوعی مسئولانه و شکوفایی کارکنان بود. نتایج حاصل از تحلیل ساختاری تفسیری نشان داد که استقرار هوش مصنوعی مسئولانه در فرایندهای منابع انسانی، نه یک مداخلهٔ تک‌بعدی فناورانه، بلکه فرایندی چندلایه و سیستماتیک است که در ۸ سطح سلسله‌مراتبی سازمان‌دهی می‌شود. در لایه‌های زیرین و ریشه‌ای مدل (سطوح حاکمیتی و فنی)، متغیرهایی مانند خط‌مشی‌گذاری اخلاقی، زیرساخت‌های داده‌ای و طراحی الگوریتمی منصفانه قرار دارند که به‌عنوان محرک‌های اصلی عمل می‌کنند. این متغیرها در لایه‌های میانی (زیرسیستم اجتماعی ـ روان‌شناختی) از طریق تقویت اعتماد سازمانی، ادراک عدالت رویه‌ای و کاهش ریسک‌های اخلاقی، بر لایه‌های بالایی اثر می‌گذارند. در نهایت، در بالاترین سطح مدل (زیرسیستم پیامدی)، شکوفایی کارکنان در قالب سازه‌هایی چون معنا در کار، درگیری فعال، رشد فردی و عملکرد بهینه محقق می‌شود. یافته‌ها مؤید آن است که مؤلفه‌های اجتماعی و اخلاقی، در انتقال اثرهای قابلیت‌های فنی به پیامدهای انسانی نقش کلیدی و واسطه‌ای ایفا می‌کنند.

نتیجه‌گیری: این پژوهش با ارائهٔ یک مدل ساختاری ـ تفسیری سه‌لایه (حاکمیتی ـ فنی، اجتماعی ـ روان‌شناختی و پیامدی)، خلأ موجود در ادبیات هوش مصنوعی مسئولانه و مدیریت منابع انسانی را پوشش می‌دهد. نتایج پژوهش نشان می‌دهد که موفقیت و پایداری فناوری‌های هوشمند در سازمان‌ها، مستلزم هم‌راستایی و بهینه‌سازی مشترک زیرسیستم‌های فنی و اجتماعی است. بدون توجه به نیازهای روان‌شناختی و اخلاقی کارکنان، پیاده‌سازی هوش مصنوعی می‌تواند به کاهش تعهد و احساس بیگانگی شغلی منجر شود. مدل پیشنهادی این پژوهش، به‌عنوان یک نقشه راهبردی، می‌تواند توسط سیاست‌گذاران سازمانی، مدیران منابع انسانی و توسعه‌دهندگان سیستم‌های هوشمند برای طراحی راه‌کارهای فناورانهٔ انسان‌محور و ارتقای شکوفایی منابع انسانی در عصر تحول دیجیتال استفاده شود.

کلیدواژه‌ها: هوش مصنوعی مسئولانه، شکوفایی کارکنان، شفافیت و پاسخ‌گویی، اعتماد سازمانی، نظریه سیستم‌های اجتماعی ـ فنی.

مقدمه

در سال‌های اخیر، یکپارچه‌سازی هوش مصنوعی در فرایندهای مدیریت منابع انسانی، تحولات چشمگیری را در حوزه‌هایی مانند جذب و انتخاب، ارزیابی عملکرد، توسعهٔ کارکنان و تصمیم‌گیری‌های مدیریتی ایجاد کرده است. هوش مصنوعی با قابلیت تحلیل کلان‌داده‌ها، شناسایی الگوهای رفتاری و پشتیبانی از تصمیم‌گیری، کارایی، سرعت و دقت بسیاری از فرایندهای منابع انسانی را ارتقا داده (مؤمنی، یعقوبی، روشن و پورعزت، ۱۴۰۵) و به‌تدریج، از یک ابزار عملیاتی، به یکی از مؤلفه‌های راهبردی مدیریت منابع انسانی تبدیل شده است (بودوار، مالک، دی‌سیلوا و تئویسوتان[1]، ۲۰۲۳). با این حال، گسترش استفاده از سامانه‌های الگوریتمی، چالش‌های اخلاقی، مدیریتی و اجتماعی جدیدی را نیز به همراه داشته است. مطالعات اخیر نشان می‌دهند که کاهش شفافیت در تصمیم‌گیری، بازتولید سوگیری‌های الگوریتمی، نقض حریم خصوصی، محدود شدن قابلیت توضیح‌پذیری تصمیم‌ها و کاهش نقش قضاوت انسانی، از مهم‌ترین چالش‌های استقرار هوش مصنوعی در مدیریت منابع انسانی محسوب می‌شوند. در ادبیات پژوهش، این چالش‌ها با ادراک کارکنان از سامانه‌های هوشمند به‌عنوان فرایندهایی غیرشفاف، فهم‌ناپذیر و تهدیدکنندهٔ استقلال فردی همراه گزارش شده‌اند (بانکینز، فرموسا و اوکین[2]، ۲۰۲۳؛ لیچ دوبالد و همکاران[3]، ۲۰۲۲).

در پاسخ به این چالش‌ها، مفهوم «هوش مصنوعی مسئولانه»، به‌عنوان پارادایمی نوین برای تضمین اخلاق‌مداری، شفافیت، پاسخ‌گویی، عدالت، حفظ حریم خصوصی و نظارت انسانی در طراحی، توسعه و به‌کارگیری سامانه‌های هوشمند مطرح شده است (استال و همکاران[4]، ۲۰۲۳). در سال‌های اخیر، چارچوب‌های معتبر بین‌المللی از جمله چارچوب مدیریت ریسک هوش مصنوعی مؤسسهٔ ملی استاندارد و فناوری آمریکا، قانون هوش مصنوعی اتحادیه اروپا[5]، اصول هوش مصنوعی مسئولانهٔ گوگل[6] و چارچوب استاندارد هوش مصنوعی مسئولانه مایکروسافت[7] نیز بر اصولی همچون عدالت، شفافیت، پاسخ‌گویی، مدیریت ریسک، قابلیت توضیح‌پذیری و نظارت انسانی تأکید کرده‌اند. با وجود اشتراک این چارچوب‌ها در معرفی اصول حاکم بر توسعهٔ مسئولانهٔ هوش مصنوعی، تمرکز آن‌ها عمدتاً بر الزامات سیاستی، حاکمیتی و طراحی سامانه‌های هوشمند است و دربارهٔ نحوهٔ سازمان‌یافتگی این اصول در بستر مدیریت منابع انسانی و روابط ساختاری میان آن‌ها، تبیین محدودی ارائه شده است. از این رو، با وجودِ تکثرِ این چارچوب‌های هنجاری، فقدان یک مدل ساختاری که بتواند این اصول را به تجربهٔ زیستهٔ کارکنان در محیط کار پیوند بزند، به شکافی بنیادین در ادبیات مدیریت منابع انسانی بدل شده است. این پژوهش با تمرکز بر همین شکاف، به واکاوی روابط ساختاری حاکم بر این مؤلفه‌ها می‌پردازد.

1. Budhwar, Malik, De Silva & Thevisuthan
2. Bankins, Formosa & O'Kane
3. Leicht-Deobald et al.
4. Stahl et al.
5. EU AI Act
6. Google Responsible AI
7. Microsoft Responsible AI Standard

استقرار هوش مصنوعی مسئولانه در مدیریت منابع انسانی، نه تنها ریسک‌های اخلاقی را کاهش می‌دهد، بلکه با ایجاد بستر اعتماد و ادراک عدالت رویه‌ای، می‌تواند نقش کلیدی در بهبود پیامدهای مثبت روان‌شناختی و عملکردی کارکنان، از جمله «شکوفایی کارکنان» ایفا کند (حاجی اسماعیلی، طهماسبی و باباشاهی، ۱۴۰۵؛ مالیک، بودوار و کاظمی[1]، ۲۰۲۳). شکوفایی کارکنان که بر اساس دیدگاه اسپریتزر، ساتکلیف، داتون، سوننشاین و گرانت[2] (۲۰۰۵)، به تجربهٔ هم‌زمان بهزیستی روان‌شناختی و عملکرد اثربخش در محیط کار اشاره دارد، در پژوهش حاضر فقط به‌عنوان یک سازهٔ نظری در نظر گرفته نشده است؛ بلکه ابعاد آن در مرحلهٔ کیفی پژوهش از طریق تحلیل مضمون مصاحبه‌های نیمه‌ساختاریافته با خبرگان شناسایی، پالایش و اعتبارسنجی شده و سپس به‌عنوان ورودی مرحله مدل‌سازی ساختاری ـ تفسیری مورد استفاده قرار گرفته است. از این منظر، شکوفایی کارکنان مفهومی چندبُعدی است که بایستی در بستر سازمان‌های دیجیتال و در تعامل میان فناوری، ساختارهای سازمانی و تجربهٔ کارکنان مطالعه شود.

مرور پیشینه پژوهش نشان می‌دهد که با وجود گسترش مطالعات مرتبط با کاربرد هوش مصنوعی در مدیریت منابع انسانی، بخش عمده‌ای از پژوهش‌های پیشین بر مزایای فنی و اقتصادی این فناوری، از جمله افزایش کارایی، بهبود تصمیم‌گیری و خودکارسازی فرایندها، یا بر شناسایی چالش‌های اخلاقی ناشی از به‌کارگیری آن، نظیر سوگیری الگوریتمی، نقض حریم خصوصی و شفافیت تصمیمات، متمرکز بوده‌اند (چودوری و همکاران[3]، ۲۰۲۳؛ موسوی، رستگار و شفیعی نیک‌آبادی، ۱۴۰۴؛ رستگار و موسوی، ۱۴۰۲). اگرچه این مطالعات شناخت ارزشمندی از ظرفیت‌ها و چالش‌های هوش مصنوعی در مدیریت منابع انسانی ارائه کرده‌اند، همچنان خلأ شایان توجهی در تبیین نحوهٔ سازمان‌یافتگی مؤلفه‌های هوش مصنوعی مسئولانه، جایگاه نسبی هر یک از این مؤلفه‌ها در یک ساختار منسجم و روابط سلسله‌مراتبی میان ابعاد فنی، مدیریتی، اخلاقی و انسانی مشاهده می‌شود. همچنین، چارچوب‌های بین‌المللی هوش مصنوعی مسئولانه بر مجموعه‌ای از اصول و الزامات مشترک تأکید دارند؛ اما نحوهٔ استقرار این اصول در بستر مدیریت منابع انسانی و ارتباط ساختاری آن‌ها با ابعاد شکوفایی کارکنان کمتر در کانون توجه قرار گرفته است.

برای مطالعهٔ این مسئله، پژوهش حاضر از نظریهٔ سیستم‌های اجتماعی ـ فنی[4]، به‌عنوان چارچوب نظری استفاده می‌کند. این نظریه که توسط چرنز[5] (۱۹۷۶) و تریست[6] (۱۹۸۱) توسعه یافته است، بر این اصل استوار است که اثربخشی نظام‌های سازمانی در گرو طراحی و بهینه‌سازی هم‌زمان زیرسیستم فنی و زیرسیستم اجتماعی است. بر اساس اصول طراحی سیستم‌های اجتماعی ـ فنی، از جمله سازگاری[7]، حداقل مشخص‌سازی بحرانی[8]، کنترل انحراف[9]، جریان

1. Malik, Budhwar & Kazmi
2. Spreitzer, Sutcliffe, Dutton, Sonenshein & Grant
3. Chowdhury et al.
4. Socio-Technical Systems Theory
5. Cherns
6. Trist
7. Compatibility
8. Minimal Critical Specification
9. Variance Control

اطلاعات[1] و مشارکت در طراحی[2]، فناوری و عوامل انسانی نباید به‌صورت مستقل طراحی و مدیریت شوند، بلکه لازم است در قالب یک نظام یکپارچه و هماهنگ مورد توجه قرار گیرند (چرنز، ۱۹۷۶؛ تریست، ۱۹۸۱). از این منظر، هوش مصنوعی مسئولانه، فقط مجموعه‌ای از الزامات فنی یا اخلاقی نیست، بلکه سازوکاری برای ایجاد هم‌راستایی میان قابلیت‌های فناورانه، الزامات حاکمیتی و نیازهای انسانی در نظام مدیریت منابع انسانی محسوب می‌شود. بنابراین، نظریهٔ سیستم‌های اجتماعی ـ فنی، بستر مناسبی برای مطالعهٔ روابط ساختاری میان مؤلفه‌های هوش مصنوعی مسئولانه و ابعاد شکوفایی کارکنان فراهم می‌آورد (ورونتیس و همکاران[3]، ۲۰۲۲).

بر این اساس، پژوهش حاضر با هدف شناسایی، تبیین و مدل‌سازی روابط ساختاری میان مؤلفه‌های مدیریت منابع انسانی مبتنی بر هوش مصنوعی مسئولانه و ابعاد شکوفایی کارکنان انجام شده است. این پژوهش با بهره‌گیری از رویکرد آمیختهٔ اکتشافی ـ تبیینی، ابتدا از طریق مصاحبه‌های نیمه‌ساختاریافته با خبرگان، مؤلفه‌های کلیدی را شناسایی و تحلیل می‌کند و سپس با استفاده از تکنیک مدل‌سازی ساختاری ـ تفسیری، جایگاه و روابط سلسله‌مراتبی این مؤلفه‌ها را توضیح می‌دهد. نوآوری نظری پژوهش حاضر در ارائهٔ یک مدل ساختاری از آرایش مؤلفه‌های هوش مصنوعی مسئولانه در مدیریت منابع انسانی، توسعهٔ کاربرد نظریهٔ سیستم‌های اجتماعی ـ فنی در محیط‌های کاری مبتنی بر هوش مصنوعی و تبیین جایگاه نسبی مؤلفه‌های فنی، حاکمیتی، مدیریتی و انسانی در یک چارچوب یکپارچه است. همچنین، نتایج این پژوهش، ضمن غنای ادبیات مدیریت منابع انسانی و هوش مصنوعی مسئولانه، می‌تواند مبنایی برای طراحی سیاست‌ها، سازوکارهای حاکمیتی و برنامه‌های استقرار مسئولانه هوش مصنوعی در سازمان‌ها را فراهم آورد.

پیشینهٔ نظری پژوهش

لنگرگاه نظری: نظریه سیستم‌های اجتماعی ـ فنی

برای تبیین شبکهٔ روابط ساختاری میان هوش مصنوعی و پیامدهای انسانی، پژوهش حاضر بر «نظریهٔ سیستم‌های اجتماعی ـ فنی[4]» (تریست، ۱۹۸۱؛ چرنز، ۱۹۷۶) تمرکز کرده است. بر اساس اصل بنیادین این نظریه، یعنی «بهینه‌سازی مشترک[5]»، کارایی و اثربخشی سیستم‌های سازمانی، فقط زمانی محقق می‌شود که طراحی و استقرار زیرسیستم‌های «فنی» (مانند الگوریتم‌های هوش مصنوعی) همگام، هم‌تراز و در یک پیوند ساختاری با نیازهای زیرسیستم‌های «اجتماعی» (کارکنان، ساختارهای انسانی و ارزش‌های روان‌شناختی) صورت پذیرد. ادبیات پیشین نشان می‌دهد که تمرکز صرف بر کارایی فنیِ هوش مصنوعی (دترمینیسم تکنولوژیک)، غالباً به شکست در پیاده‌سازی، بیگانگی کاری و مقاومت کارکنان منجر می‌شود.

1. Information Flow
2. Participation in Design
3. Vrontis et al.
4. Socio-Technical Systems Theory - STS
5. Joint Optimization

از این جنبه، «هوش مصنوعی مسئولانه»، به‌عنوان یک سازوکار میانجی[1] و پیونددهنده عمل می‌کند که با هم‌راستا کردن منطق الگوریتمیک با ارزش‌های انسانی، زیرسیستم‌های اجتماعی و فنی را به تعادل ساختاری می‌رساند (اسچولته و استریت[2]، ۲۰۲۵). در تبیین دقیق‌تر این مفهوم بر مبنای اصول نه‌گانه طراحی سیستم‌های اجتماعی ـ فنی (چرنز، ۱۹۷۶)، این «تعادل ساختاری» تجلیِ «اصل سازگاری[3]» و بهینه‌سازی مشترک است؛ به این معنا که منطق طراحی الگوریتم‌ها باید با ارزش‌ها و نیازهای روان‌شناختی منابع انسانی هم‌سو باشد. علاوه‌براین، در این مدل، مؤلفه‌هایی نظیر «پیشگیری از سوگیری و رعایت انصاف» پیوند تنگاتنگی با «اصل کنترل انحراف[4]» دارند. بر اساس این اصل، خطاها و انحراف‌ها باید در نزدیک‌ترین نقطه به مبدأ وقوع، شناسایی و اصلاح شوند؛ بنابراین شفافیت الگوریتم‌ها به‌عنوان یک تکنولوژی میانجی، به کاربران اجازه می‌دهد تا سوگیری‌ها را در همان مرحلهٔ پردازش داده مهار کنند. در نهایت، «توانمندسازی مدیریتی و نظارت انسانی» بازتابی از «مشارکت در طراحی[5]» است؛ زیرا مدیران توانمندشده، فقط مصرف‌کنندهٔ فناوری نیستند، بلکه از طریق اعمال کنترل انسانی[6] در تنظیم و بازطراحی پیوستهٔ سیستم در بستر اجتماعی سازمان مشارکت فعال دارند تا به بازتولید «ساختارهای خودسازمان‌ده[7]» کمک کنند.

بر اساس منطق نظریهٔ سیستم‌های اجتماعی ـ فنی، پیامدهای انسانی ناشی از فناوری، نه از ویژگی‌های فنی فناوری به تنهایی و نه از ویژگی‌های فردی کارکنان به‌صورت مستقل ناشی می‌شود، بلکه حاصل کیفیت تعامل و هم‌ترازیِ ساختاری میان این دو زیرسیستم است. در بستر مدیریت منابع انسانی مبتنی بر هوش مصنوعی، مؤلفه‌هایی نظیر شفافیت الگوریتمی، تبیین‌پذیری، عدالت، پاسخ‌گویی و نظارت انسانی به‌عنوان عناصر زیرسیستم فنی عمل می‌کنند. این عناصر از طریق شکل‌دهی ادراک کارکنان نسبت به انصاف، اعتماد، کنترل ادراک‌شده و امنیت روان‌شناختی، بر زیرسیستم اجتماعی اثر می‌گذارند. از این جنبه، شکوفایی کارکنان یک متغیر وابستهِ صرف به فناوری نیست، بلکه نتیجهٔ هم‌راستایی موفق میان الزامات فنی سامانه‌های هوش مصنوعی و نیازهای انسانی کارکنان در یک شبکهٔ سیستمی است.

تکامل هوش مصنوعی در مدیریت منابع انسانی و ضرورت رویکرد مسئولانه

بررسی سیر تحول فناوری در سازمان‌ها نشان می‌دهد که ورود هوش مصنوعی به حوزهٔ مدیریت منابع انسانی، موجب دگرگونی اساسی در منطق تصمیم‌گیری، طراحی فرایندها و شیوه‌های تعامل با نیروی کار شده است. پژوهشگران (مانند خایر، ماهاداسا، تولی و آنده[8]، ۲۰۲۰) تأیید کرده‌اند که کاربرد هوش مصنوعی در جذب، ارزیابی عملکرد و آموزش، سرعت و دقت تصمیم‌ها را به‌شدت افزایش می‌دهد. با این حال، نسل جدید مطالعات هشدار می‌دهند که پدیده‌هایی نظیر «جعبهٔ سیاه الگوریتمی» باعث شده‌اند که کارکنان، تصمیم‌های ماشین را فرایندهایی فهم‌ناپذیر و چالش‌ناپذیر ادراک کنند که این امر به‌طور مستقیم عاملیت انسانی و اعتماد سازمانی را هدف قرار می‌دهد.

1. Mediating Mechanism
2. Schulte & Streit
3. Compatibility Principle
4. Variance Control
5. Design Participation
6. Human-in-the-loop
7. Self-organizing Structures
8. Khair, Mahadasa, Tuli & Ande

در پاسخ به این بحران، مفهوم «هوش مصنوعی مسئولانه» ظهور کرده است. هوش مصنوعی مسئولانه در مدیریت منابع انسانی به مجموعه‌ای از سازوکارها، اصول و رویه‌های حاکمیتی اطلاق می‌شود که تضمین می‌کنند طراحی، استقرار و استفاده از سامانه‌های هوش مصنوعی با معیارهای شفافیت، پاسخ‌گویی، تبیین‌پذیری، عدالت، حفظ حریم خصوصی، نظارت انسانی و قابلیت اصلاح و کنترل هم‌سو باشد. این مفهوم بر مبنای چارچوب‌های معتبر بین‌المللی از جمله چارچوب مدیریت ریسک هوش مصنوعی مؤسسه ملی استاندارد و فناوری آمریکا[1]، قانون هوش مصنوعی اتحادیه اروپا[2] و ادبیات هوش مصنوعی مسئولانه در مدیریت منابع انسانی توسعه یافته است. ترپیاتی و کومار[3] (۲۰۲۵) هوش مصنوعی مسئولانه را نه یک افزونهٔ فنی، بلکه یک «الزام اخلاقی ـ استراتژیک» می‌دانند که چهار رکن اساسی دارد: عدالت الگوریتمی، شفافیت تصمیم‌ها، پاسخ‌گویی و حفظ کرامت انسانی. استقرار این الگو، تضمین‌کنندهٔ ایجاد تعادل میان مزایای فناورانه و مسئولیت‌های اجتماعی در اکوسیستم منابع انسانی است (چانگ و که[4]، ۲۰۲۴).

شکوفایی کارکنان در عصر الگوریتم‌ها

ادبیات نوین مدیریت منابع انسانی، از تمرکز سنتی بر متغیرهای تقلیلی نظیر «رضایت شغلی» عبور کرده است (واعظی و پاشایی، ۱۴۰۴) و بر مفهوم کلان‌تر «شکوفایی کارکنان» تأکید دارد. شکوفایی، سازه‌ای چندبُعدی است که احساس معنا در کار، سلامت روان، درگیری فعال و رشد فردی را دربرمی‌گیرد (والتونن، کیمپیماکی و ساولا[5]، ۲۰۲۶). در واقع، شکوفایی کارکنان، به‌عنوان یک وضعیت روان‌شناختی مثبت و پویا در نظر گرفته شده است که در آن، افراد علاوه‌بر عملکرد اثربخش، احساس رشد، یادگیری، معناداری و پیشرفت مستمر را در محیط کار تجربه می‌کنند. این تعریف با دیدگاه «شکوفایی در کار» ارائه‌شده توسط اسپریتزر و همکاران (۲۰۰۵) هم‌سو است که شکوفایی را حاصل ترکیب یادگیری و سرزندگی در محیط کار می‌دانند. بر این اساس، مؤلفه‌هایی نظیر رشد حرفه‌ای، معنا در کار، امنیت روان‌شناختی و مشارکت فعال به‌عنوان بازتاب‌های کلیدی شکوفایی کارکنان در بستر هوش مصنوعی مسئولانه شناسایی شدند. در پژوهش حاضر نیز، این ابعاد چهارگانه (رشد حرفه‌ای، معنا در کار، امنیت روان‌شناختی و مشارکت فعال) بر اساس یکپارچه‌سازی چارچوب‌های نظری اسپریتزر و همکاران (۲۰۰۵) و والتونن و همکاران (۲۰۲۶) به‌عنوان شالوده مفهومیِ سازه شکوفایی کارکنان صورت‌بندی شده‌اند.

مطالعات نشان می‌دهد که صرفِ پیاده‌سازی فناوری‌های هوشمند، تضمین‌کنندهٔ شکوفایی و رفاه کارکنان نیست؛ بلکه نحوهٔ ادراک آن‌ها از «عدالت» و «حریم خصوصی» در تعامل با الگوریتم‌هاست که تعیین می‌کند آیا فناوری به شکوفایی منجر می‌شود یا به بیگانگی کاری می‌انجامد (صادقی[6]، ۲۰۲۴؛ سومرو، فن، سوهو، سومرو و شیخ[7]، ۲۰۲۴). در

1. NIST AI RMF
2. EU AI Act
3. Tripathi & Kumar
4. Chang & Ke
5. Valtonen, Valtonen, Kimpimäki & Savela
6. Sadeghi
7. Soomro, Fan, Sohu, Soomro & Shaikh

چارچوب نظریهٔ سیستم‌های اجتماعی ـ فنی، هوش مصنوعی مسئولانه از طریق ایجاد شفافیت در منطق تصمیم‌گیری الگوریتم‌ها، کاهش ادراک سوگیری، افزایش قابلیت پاسخ‌گویی و حفظ نظارت انسانی، شرایطی را فراهم می‌کند که کارکنان احساس اعتماد، انصاف و امنیت روان‌شناختی بیشتری داشته باشند. این سازوکارها به‌عنوان حلقه اتصال میان زیرسیستم فنی و زیرسیستم اجتماعی عمل کرده و زمینهٔ لازم برای شکل‌گیری مشارکت فعال، یادگیری مستمر، معنا در کار و در نهایت شکوفایی کارکنان را فراهم می‌سازند.

مرور انتقادی مطالعات پیشین و تبیین شکاف پژوهشی

مرور نظام‌مند ادبیات در تقاطع هوش مصنوعی و مدیریت منابع انسانی نشان می‌دهد که با وجود رشد فزایندهٔ مطالعات، پیکرهٔ دانشی موجود با سه نارسایی بنیادین نظری و متدولوژیک مواجه است:

غلبه پارادایم تکنومحور و سوگیری کارکردگرا[1]: جریان غالب پژوهش‌ها تاکنون، هوش مصنوعی را در چارچوب مکتب کارکردگرایی و فقط به‌عنوان ابزاری برای بهینه‌سازی فرایندها و ارتقای بهره‌وری سازمانی ارزیابی کرده است (بودوار و همکاران، ۲۰۲۳؛ ورونتیس و همکاران، ۲۰۲۲) و تأثیرهای عمیق این فناوری بر تجربهٔ زیسته، امنیت روانی و شکوفایی روان‌شناختی کارکنان تا حد زیادی نادیده گرفته شده است (اسچولته و استریت، ۲۰۲۵).

خلأ در تبیین شبکهٔ نومولوژیک و سازوکارهای ساختاری[2]: بخش عمده‌ای از مطالعاتی که به پیامدهای اخلاقی و انسانی هوش مصنوعی پرداخته‌اند، در سطح مرور مفهومی متوقف شده یا به آزمون‌های هم‌بستگی خطیِ ساده بسنده کرده‌اند (پریکشات، مالک و بودوار[3]، ۲۰۲۳). بر این اساس، خلأ تئوریکِ جدی در درک چگونگیِ اثرگذاریِ ابعاد «هوش مصنوعی مسئولانه» بر «شکوفایی کارکنان» وجود دارد. در واقع، هنوز مدلی که بتواند این ارتباط را به‌صورت نظام‌مند، در قالب یک «سلسله‌مراتب ساختاری» تبیین کند، توسعه نیافته است (تریپاتی و کومار، ۲۰۲۵).

مطالعات پیشین به‌ندرت سازوکارهای میانجی و فرایندهای انتقال را که از طریق آن‌ها ویژگی‌های فنی سامانه‌های هوش مصنوعی به پیامدهای انسانی و روان‌شناختی تبدیل می‌شوند، مورد توجه قرار داده‌اند. در نتیجه، نحوهٔ تعامل زیرسیستم‌های فنی و اجتماعی که هستهٔ مرکزی نظریهٔ سیستم‌های اجتماعی ـ فنی و اصول نه‌گانه چرنز را تشکیل می‌دهد، همچنان به‌طور کافی در مدل‌های ساختاری تبیین نشده است.

نکتهٔ مهم دیگر آن است که اگرچه چارچوب‌های معتبر بین‌المللی، نظیر چارچوب مدیریت ریسک هوش مصنوعی مؤسسهٔ ملی استاندارد و فناوری آمریکا، قانون هوش مصنوعی اتحادیهٔ اروپا، اصول هوش مصنوعی گوگل و چارچوب هوش مصنوعی مسئولانه شرکت مایکروسافت، سهم مهمی در توسعهٔ ادبیات هوش مصنوعی مسئولانه داشته‌اند؛ اما تمرکز اصلی آن‌ها بر طراحی، حاکمیت، مدیریت ریسک و کنترل پیامدهای نامطلوب سامانه‌های هوش مصنوعی است. این چارچوب‌ها مجموعه‌ای از اصول هنجاری و الزامات اجرایی نظیر عدالت، شفافیت، پاسخ‌گویی، امنیت، حفظ حریم

1. Techno-Centric Paradigm & Functionalist Bias
2. Paucity of Causal & Structural Modeling
3. Prikshat, Malik & Budhwar

خصوصی و نظارت انسانی را ارائه می‌کنند؛ اما به‌طور عمده در سطح «آنچه باید انجام شود» (اصول هنجاری) باقی می‌مانند.

جدول ۱. تقابل رویکرد چارچوب‌های بین‌المللی با رویکرد پژوهش حاضر

شاخص مقایسه	چارچوب‌های بین‌المللی	مدل پیشنهادی پژوهش حاضر
تمرکز اصلی	مدیریت ریسک و حاکمیت	مدیریت منابع انسانی و شکوفایی
ماهیت رویکرد	هنجاری و الزامی	ساختاری و عملکردی
هدف نهایی	پیشگیری از آسیب و کنترل ریسک	ارتقای تجربهٔ زیسته و شکوفایی کارکنان
واحد فنی	سیستم و زیرساخت فنی	تعامل سیستم و انسان
پرسش کلیدی	چه اصولی باید رعایت شوند؟	این اصول چگونه به پیامد انسانی تبدیل می‌شوند؟

همان‌طور که در جدول ۱ تبیین شد، شکاف اصلی نه در کمبود اصول، بلکه در فقدان «سازوکار انتقال» نهفته است. در واقع، چارچوب‌های مذکور بر «محتوای» اصول تمرکز دارند، در حالی که پژوهش حاضر بر «ساختار» اثرگذاری آن‌ها متمرکز است.

در مقابل، هنوز در ادبیات موجود توضیح روشنی دربارهٔ این موضوع وجود ندارد که این اصول از طریق چه سازوکارهای سازمانی، اجتماعی و روان‌شناختی به پیامدهای انسانی مطلوب در محیط کار منجر می‌شوند. به‌ویژه در حوزهٔ مدیریت منابع انسانی، بخش عمدهٔ پژوهش‌های موجود یا بر شناسایی اصول هوش مصنوعی مسئولانه متمرکز بوده‌اند یا اثرهای مستقیم فناوری را بر پیامدهای سازمانی بررسی کرده‌اند. در نتیجه، دانش موجود هنوز فاقد یک مدل ساختاری منسجم است که بتواند مسیرهای اثرگذاری مؤلفه‌های فنی هوش مصنوعی مسئولانه بر متغیرهای انسانی را در قالب یک نظام فنی ـ اجتماعی تبیین کند. این خلأ زمانی اهمیت بیشتری پیدا می‌کند که شکوفایی کارکنان به‌عنوان یک سازه چندبُعدی و پیچیده، نه پیامد مستقیم فناوری، بلکه حاصل تعامل میان ویژگی‌های فنی سامانه‌های هوش مصنوعی و ادراکات، نگرش‌ها و تجربهٔ زیستهٔ کارکنان تلقی شود. بر این اساس، نوآوری نظری پژوهش حاضر در معرفی اصول جدید هوش مصنوعی مسئولانه نیست، بلکه در تبیین روابط ساختاری میان اصول شناخته‌شده هوش مصنوعی مسئولانه و پیامدهای انسانی آن در بستر مدیریت منابع انسانی نهفته است. پژوهش حاضر با اتکا بر نظریه سیستم‌های اجتماعی ـ فنی و بهره‌گیری از مدل‌سازی ساختاری ـ تفسیری، می‌کوشد تا شکاف موجود میان چارچوب‌های حاکمیتی هوش مصنوعی مسئولانه و پیامدهای رفتاری و روان‌شناختی کارکنان را پُر کند و با ترسیم «زنجیرهٔ تعاملات»، سازوکارهای انتقال اثرهای مؤلفه‌های فنی به شکوفایی کارکنان را آشکار سازد.

یگانه‌انگاری روش‌شناختی[1] و فقر مطالعات اکتشافی ــ بومی

پدیدهٔ هوش مصنوعیِ مسئولانه ماهیتی به‌شدت نوظهور، چندبعدی و وابسته به بافتار[2] دارد. با این حال، پیشینهٔ تجربی غالباً بر رویکردهای کمّی ـ قیاسی تکیه کرده است. این رویکرد از کشف مؤلفه‌های پنهانِ مرتبط با ارزش‌ها و نگرش‌های بومیِ نیروی کار عاجز است (بونجاک، سرنا و پوپوویچ[3]، ۲۰۲۳).

موضع پژوهش و معماری روش‌شناختی[4]

در پاسخ مستقیم به نارسایی‌های فوق، پژوهش حاضر با عبور از طرح‌های تک‌روشی، یک استراتژی « آمیخته اکتشافی ـ تبیینی متوالی[5]» را اتخاذ کرده است. با استفاده از مطالب فوق، این معماری پژوهشی ابتدا در مرحلهٔ کیفی، از طریق تحلیل مضمون عمیق، ابعاد پنهان و بومی هوش مصنوعی مسئولانه را شناسایی می‌کند. سپس در مرحله کمی، با بهره‌گیری از تکنیک مدل‌سازی ساختاری تفسیری، از تحلیل‌های آماریِ تخت رایج فراتر رفته و امکان ترسیم شبکه روابط ساختاری و سطح‌بندی مؤلفه‌ها را فراهم می‌آورد. این رویکرد، گامی بدیع در جهت گذار از «توصیف پدیده» به «تبیین سازوکارهای ساختاریِ» هوش مصنوعی مسئولانه در ارتقای شکوفایی کارکنان محسوب می‌شود.

روش‌شناسی پژوهش

پژوهش حاضر در چارچوب پارادایم پراگماتیسم و با بهره‌گیری از طرح آمیخته اکتشافی ـ تبیینی متوالی انجام شده است (کرسول و پلانوکلارک[6]، ۲۰۱۸). با توجه به فقدان چارچوب‌های نظری و مدل‌های ساختاری جامع برای تبیین روابط میان هوش مصنوعی مسئولانه و شکوفایی کارکنان، این رویکرد امکان می‌دهد تا در مرحلهٔ نخست، ابعاد، مؤلفه‌ها و سازوکارهای مرتبط با پدیده مورد مطالعه به صورت اکتشافی شناسایی شوند و در مرحلهٔ دوم، ساختار روابط و الگوی تعامل میان این مؤلفه‌ها مورد بررسی قرار گیرد. در مطالعاتی با ماهیت اکتشافی و توسعه مدل، دانش تخصصی و تجربیات خبرگان، یکی از معتبرترین منابع برای شناسایی، پالایش و سازمان‌دهی مفاهیم پیچیده و چندبُعدی محسوب می‌شود. از این رو، در مرحلهٔ مدل‌سازی، از قضاوت گروهی خبرگان آشنا با حوزه‌های مدیریت منابع انسانی، هوش مصنوعی و اخلاق فناوری استفاده شد تا ساختار روابط میان مؤلفه‌های شناسایی‌شده استخراج شود. این رویکرد با مبانی روش مدل‌سازی ساختاری ـ تفسیری سازگار بوده و امکان تبیین روابط سلسله‌مراتبی، الگوهای تعامل و وابستگی متقابل میان متغیرها را در حوزه‌های نوظهور و دارای پیچیدگی مفهومی فراهم می‌سازد.

مرحلهٔ اول (اکتشاف کیفی و استخراج مؤلفه‌ها)

جامعه آماری این مرحله، خبرگان دوحوزه‌ای (بخش منابع انسانی و توسعه‌دهندگان/سیاست‌گذاران هوش مصنوعی) بود.

1. Methodological Monism
2. Context-dependent
3. Bunjak, Černe & Popovič
4. Methodological Positioning
5. Sequential Exploratory Mixed-Methods Design
6. Creswell & Plano Clark

انتخاب مشارکت‌کنندگان بر اساس رویکرد نمونه‌گیری «هدفمند معیارمحور[1]» سپس «گلولهٔ برفی» انجام شد. معیارهـای ورود شامل داشتن حداقل ۱۵ سال سابقهٔ کار تخصصی و درگیری عملی/پژوهشی با سیسـتم‌های هوشـمند بـود. فراینـد جمع‌آوری داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافته صورت گرفت و پس از انجام ۱۷ مصاحبه، پـژوهش بـه «اشـباع نظری» رسید؛ نقطه‌ای که داده‌های جدید، کد مفهـومی تـازه‌ای تولیـد نکردنـد. داده‌هـای متنـی بـا اسـتفاده از رویکـرد بازنگری‌شده و تأملی تحلیل مضمون استخراج شدند (براون و کلارک[2]، ۲۰۲۱). برای تضمین «اعتبارپذیری[3]»، از تکنیک «بررسی توسط اعضا» استفاده شد. همچنین، به‌منظور کنتـرل سـوگیری پژوهشـگر و تأییـد «پایـایی بین‌کدگـذاران» از ضریب کاپای کوهن استفاده شد. مقدار محاسباتی این شاخص برای دو کدگذار مستقل برابر بـا ۰/۸۳ به‌دسـت آمـد کـه نشان‌دهندهٔ توافق بالا و مطلوب است.

مرحلهٔ دوم (مدل‌سازی ساختاری تفسیری)

به‌منظور گذار یکپارچه از فاز کیفی به کمّی و عملیاتی‌سازی دقیق متغیرهای پژوهش، ابعاد اولیه استخراج‌شـده از ادبیـات نظری (به‌ویژه مؤلفه‌های چندگانه شکوفایی کارکنان)، فقط به‌عنوان فهرست اولیه متغیرها مطرح شدند و پیش از ورود به مدل ساختاری ـ تفسیری، در پنل خبرگان و در فرایند تحلیل مضمون کیفی، بررسی، تعدیل و تأیید شدند. بر این اسـاس، این متغیرها به‌صورت پیش‌فرض در نظر گرفته نشدند؛ بلکه ابتدا در فاز کیفـی و از طریـق تحلیـل مضـمون مصـاحبه‌ها، توسط پنل خبرگان مورد ارزیابی، بومی‌سازی و غربالگری قرار گرفتند. تنها پس از تأییـد نهـایی ایـن مؤلفـه‌ها، بـه‌عنوان ورودی‌های قطعی تکنیک «مدل‌سازی ساختاری تفسیری» مورد اسـتفاده قـرار گرفتنـد ایـن تکنیـک، یـک متـدولوژی سیستماتیک برای شناسایی روابط ساختاریافته مستقیم و غیرمستقیم میان متغیرهای یک سیستم پیچیده است (سوشیل[4]، ۲۰۱۲). شایان ذکر است که روابط حاصل از این تکنیک بیانگر ساختار ارتباطات و سطح‌بندی میان متغیرها بـوده و نبایـد به‌عنوان روابط علّی میان متغیرها تفسیر شوند. در این مرحله، یک پنل خبـره متشـکل از ۱۷ نفـر از متخصصـان ارشـد، ماتریس خودتعاملی ساختاری را تکمیل کرد. به‌منظور حفظ انسجام مفهـومی میـان مرحلـهٔ شناسـایی مؤلفـه‌ها و مرحلـهٔ مدل‌سازی ساختاری، بخشی از خبرگان در هر دو مرحله پژوهش مشارکت داشتند. این رویکـرد در مطالعـات اکتشـافی و مدل‌سازی مبتنی بر خبرگان با هدف اطمینان از تداوم دانش تخصصی و تفسیر منسجم مفاهیم به‌کار گرفته می‌شـود. بـا این حال، برای کاهش احتمال سوگیری ناشی از توافق گروهی، تلاش شد از خبرگان دارای سـوابق علمـی، تخصصـی و تجربیات حرفه‌ای متنوع در حوزه‌های مدیریت منابع انسانی، هوش مصنوعی و اخلاق فناوری استفاده شود. این مـاتریس بر اساس مقایسه‌های زوجی مفهومی برای تعیین جهت روابط (V, A, X, O) میان ابعـاد هـوش مصـنوعی مسـئولانه و شکوفایی کارکنان تنظیم شد. برای اطمینان از هم‌گرایی و اجماع نظرهـای خبرگـان در تکمیـل ماتریس‌هـا، از «ضـریب

1. Criterion-based Purposive Sampling
2. Braun & Clarke
3. Credibility
4. Sushil

هماهنگی کندال[1]» استفاده شد و مقدار آن در پژوهش حاضر ۰/۸۴ به‌دست آمد. از آنجا که مقادیر بالای ۰/۷ نشان‌دهندهٔ توافق قوی میان ارزیابان است (فیلد[2]، ۲۰۱۸)، پایایی ابزار در مرحلهٔ مدل‌سازی به‌طور قطعی تأیید شد. در نهایت، با تشکیل ماتریس‌های دسترسی اولیه و اعمال «قانون تراگذری» در ریاضیات، ماتریس دسترسی نهایی شکل گرفت و مبنای ترسیم گراف جهت‌دار و سطح‌بندی سلسله‌مراتبی متغیرها شد.

تجزیه‌وتحلیل داده‌ها

در گام نخست، داده‌های میدانی از طریق مصاحبه‌های عمیق و نیمه‌ساختاریافته با ۱۷ نفر از خبرگان گردآوری شد. این پنل متشکل از متخصصانی با میانگین ۲۱ سال سابقه حرفه‌ای متمرکز و تحصیلات عالی (۲۹ درصد دارای مدرک دکتری) بود. همان‌طور که در جدول ۲ مشاهده می‌شود، توزیع متوازن تخصص‌ها در حوزه‌های مدیریت منابع انسانی، هوش مصنوعی، و اخلاق فناوری، امکان تثلیث[3] منابع داده و بررسی چندوجهی پدیده را فراهم ساخت. فرایند نمونه‌گیری تا رسیدن به «اشباع نظری» ادامه یافت؛ به‌طوری‌که از مصاحبه پانزدهم به بعد، هیچ کد مفهومی جدیدی که تغییر معناداری در ساختار مدل ایجاد کند، شناسایی نشد.

جدول ۲. اطلاعات جمعیت‌شناختی خبرگان پژوهش

متغیر	طبقه‌بندی	تعداد	درصد
جنسیت	مرد	۱۱	۶۴/۷
	زن	۶	۳۵/۳
سطح تحصیلات	کارشناسی ارشد	۱۲	۷۰/۶
	دکتری	۵	۲۹/۴
حوزه تخصصی اصلی	مدیریت منابع انسانی	۷	۴۱/۲
	مدیریت رفتار سازمانی	۳	۱۷/۶
	اخلاق فناوری و سیاست‌گذاری	۴	۲۳/۵
	هوش مصنوعی و تحلیل داده	۳	۱۷/۶
نوع تجربه حرفه‌ای	دانشگاهی	۶	۳۵/۳
	اجرایی/ مدیریتی	۵	۲۹/۴
	دانشگاهی ـ اجرایی	۶	۳۵/۳

تحلیل متون پیاده‌سازی‌شده با اتکا بر رویکرد شش‌مرحله‌ای و تأملی تحلیل مضمون (براون و کلارک، ۲۰۲۱) انجام گرفت. خروجی این فرایند رفت‌وبرگشتی، استخراج ۱۹۶ کد اولیه بود که پس از پالایش مستمر، تقطیر معنایی و حذف هم‌پوشانی‌ها، در قالب ۱۵ مضمون فرعی و ۵ مضمون اصلی (تم) ساختاریافته شدند (جدول ۳).

1. Kendall’s W
2. Field
3. Triangulation

جدول ۳. مضامین اصلی، فرعی و پایه مدل پژوهش

مضامین پایه	مضامین فرعی	مضامین اصلی
وجود سوگیری تاریخی در داده‌های استخدام، بازتولید تبعیض جنسیتی در الگوریتم‌ها، بازتولید تبعیض سنی در ارزیابی عملکرد، وابستگی مدل به داده‌های ناقص، نبود پایش عدالت الگوریتمی، ضرورت تست برابری فرصت، بررسی پیامدهای ناهمگون تصمیمات، کنترل متغیرهای تبعیض‌آمیز، اصلاح دیتاست‌های قدیمی، تحلیل خطای سیستماتیک، مقایسه نتایج بین گروه‌ها، حساسیت مدل به ویژگی‌های جمعیت‌شناختی، ضرورت ممیزی الگوریتم، شفاف‌سازی منبع داده‌ها، استانداردسازی داده‌های ورودی، حذف داده‌های غیرمرتبط، اعتبارسنجی داده‌های آموزشی، بازنگری دوره‌ای مدل	پیشگیری از سوگیری داده‌ای	حاکمیت و تنظیم‌گری اخلاقی
مشخص بودن مسئول انسانی تصمیم، امکان پیگیری تصمیم الگوریتمی، ثبت مستندات تصمیمات، سازوکار اعتراض کارکنان، کمیتهٔ نظارت اخلاقی، گزارش‌دهی شفاف مدیریتی، تعیین مرجع پاسخ‌گو، جلوگیری از فرافکنی مسئولیت به ماشین، تعریف حدود اختیارات سیستم، وجود نظام رسیدگی به شکایات، مستندسازی فرایند تصمیم، ارزیابی اثرهای اجتماعی تصمیمات	پاسخ‌گویی و مسئولیت‌پذیری	
تدوین آیین‌نامه استفاده از هوش مصنوعی، تعریف اصول اخلاقی سازمان، انطباق با قوانین کار، رعایت مقررات حفاظت داده، وجود سیاست محرمانگی، آموزش اخلاق هوش مصنوعی به مدیران، تدوین کد رفتاری، پایش انطباق با استانداردها، ارزیابی ریسک فناوری، تحلیل اثرهای بلندمدت، طراحی چارچوب کنترل داخلی، تعریف خطوط قرمز تصمیم‌گیری، تعیین سطح مداخله انسانی، هم‌راستاسازی با ارزش‌های سازمان، ارزیابی تبعات شهرت سازمان، توجه به عدالت بین‌نسلی، در نظر گرفتن پیامدهای فرهنگی، پایش اثرهای نابرابر فناوری، بررسی اکثریت‌ها بر اقلیت‌ها، مستندسازی تغییرات الگوریتم، تعیین مالکیت داده‌ها، تعیین سطح دسترسی کاربران، ارزیابی امنیت سایبری، کنترل نشت اطلاعات، پیشگیری از سوءاستفاده داده، تعیین چرخه عمر داده، حذف داده‌های مازاد، به‌روزرسانی دوره‌ای سیاست‌ها، بازبینی سیاست‌های داخلی، ممیزی بیرونی سیستم، گزارش اثرات اخلاقی سالانه، ارزیابی ریسک تبعیض، تحلیل اثرات تصمیم بر رفاه کارکنان، تعریف شاخص‌های اخلاقی عملکرد	چارچوب سیاست‌گذاری	
قابلیت توضیح منطق تصمیم‌گیری، امکان مشاهده معیارهای ارزیابی، وضوح شاخص‌های رتبه‌بندی داوطلبان، قابلیت ارائه گزارش تحلیلی به کارکنان، نمایش وزن متغیرها در مدل، تبیین علت رد یا پذیرش فرد، توضیح دلایل کاهش امتیاز عملکرد، امکان مشاهده روند تصمیم، استفاده از مدل‌های قابل‌توضیح، اجتناب از الگوریتم‌های جعبهٔ سیاه، ترجمهٔ خروجی فنی به زبان فهمیدنی، طراحی داشبورد شفاف، مستندسازی منطق مدل، آموزش کارکنان دربارهٔ نحوهٔ تحلیل خروجی، امکان مقایسه نتایج انسانی و ماشینی	تبیین‌پذیری الگوریتم	طراحی و معماری فناورانه مسئولانه
امکان اصلاح دستی خروجی، توقف سیستم در شرایط خطا، تنظیم سطح خودکارسازی، تعریف آستانه‌های هشدار، شناسایی نتایج غیرعادی، امکان بازتنظیم پارامترها، بررسی خطای پیش‌بینی، بازبینی تصمیم‌های حساس، آزمایش سناریوهای جایگزین، ارزیابی دقت الگوریتم، اعتبارسنجی متقابل مدل، آزمون دوره‌ای عملکرد، امکان لغو تصمیم نهایی	قابلیت کنترل و اصلاح	

مضامین پایه	مضامین فرعی	مضامین اصلی
رمزگذاری داده‌های کارکنان، محدودسازی دسترسی به اطلاعات، تفکیک داده‌های شخصی و شغلی، ذخیره‌سازی ایمن اطلاعات، اطلاع‌رسانی دربارهٔ نحوهٔ استفاده از داده، اخذ رضایت آگاهانه کارکنان، حذف داده‌های غیرضروری، پایش دسترسی غیرمجاز، تعیین مدت نگهداری داده، شفافیت در انتقال داده، رعایت مقررات حفاظت داده، جلوگیری از اشتراک‌گذاری غیرمجاز	حفاظت داده و حریم خصوصی	
اطلاع‌رسانی قبل از اجرای سیستم، توضیح اهداف استفاده از هوش مصنوعی، جلسات پرسش و پاسخ کارکنان، شفاف‌سازی پیامدهای شغلی، اطلاع‌رسانی درباره تغییرات سیستم، ارائه گزارش‌های دوره‌ای، پاسخ‌گویی مدیران به نگرانی‌ها، ایجاد کانال ارتباطی مستقیم، انتشار راهنمای کاربری، اطلاع از معیارهای ارتقا، شفافیت در ارزیابی عملکرد	ارتباطات شفاف	اجرای سازمانی و فرایندی
آموزش سواد داده‌ای مدیران، آموزش تحلیل گزارش‌های هوش مصنوعی، توسعهٔ مهارت تصمیم‌گیری داده‌محور، آشنایی مدیران با محدودیت‌های الگوریتم، درک ریسک‌های فناوری، تقویت مهارت نظارت انسانی، آموزش اخلاق هوش مصنوعی، ارتقای شایستگی دیجیتال، تقویت تفکر انتقادی نسبت به خروجی سیستم	توانمندسازی مدیریتی	
استفاده از هوش مصنوعی به‌عنوان ابزار کمکی، تأیید نهایی انسانی تصمیمات حساس، ترکیب قضاوت انسانی و داده، بررسی شرایط خاص فردی، امکان نادیده گرفتن خروجی سیستم، انعطاف در تصمیم‌گیری، ارزیابی کیفی در کنار ارزیابی کمی، مشورت گروهی قبل از تصمیم نهایی، استفاده مرحله‌ای از هوش مصنوعی، مداخله انسانی در موارد بحرانی، حفظ نقش سرپرست مستقیم، توازن بین اتوماسیون و قضاوت انسانی	تصمیم ترکیبی انسان ـ ماشین	
احساس عدالت رویه‌ای، احساس عدالت توزیعی، عدالت تعاملی، پذیرش تصمیمات سازمان، ادراک بی‌طرفی سیستم، احساس برابری فرصت، کاهش احساس تبعیض، مقایسه منصفانه عملکرد	ادراک انصاف	سازوکارهای روان‌شناختی
اعتماد به فناوری، اعتماد به مدیریت، اعتماد به صحت داده‌ها، باور به خیرخواهی سازمان، کاهش بدبینی نسبت به سیستم، اطمینان از صحت تصمیم، پذیرش نتایج الگوریتم، تمایل به استفاده از سیستم	اعتماد سازمانی	
کاهش اضطراب فناورانه، کاهش ترس از جایگزینی، احساس امنیت شغلی، اطمینان از حفظ کرامت انسانی، کاهش نگرانی درباره نظارت افراطی، احساس کنترل بر مسیر شغلی، آرامش در مواجهه با فناوری، کاهش استرس ناشی از ارزیابی دیجیتال	امنیت روان‌شناختی	پیامدهای توسعه‌ای و شکوفایی
یادگیری شخصی‌سازی‌شده، پیشنهاد دوره‌های متناسب با مهارت، شناسایی شکاف مهارتی، برنامه‌ریزی مسیر شغلی، دریافت بازخورد تحلیلی، افزایش مهارت‌های تخصصی، توسعه شایستگی‌های نرم، تسهیل پیشرفت شغلی	رشد حرفه‌ای	
احساس پیشرفت حرفه‌ای، درک ارزش کار، احساس تأثیرگذاری، هم‌راستایی با اهداف سازمان، تعلق سازمانی، انگیزش درونی، رضایت از مسیر رشد	معنا در کار	
افزایش تعامل شغلی، مشارکت در تصمیم‌گیری، تمایل به ارائه ایده، افزایش انرژی کاری، تعهد سازمانی، رفتار شهروندی سازمانی، همکاری بین واحدی، تمایل به ماندگاری در سازمان، احساس مالکیت نسبت به کار، انگیزه برای بهبود عملکرد، اشتیاق به یادگیری مستمر، تعامل مثبت با مدیر، احساس دیده شدن، احساس ارزشمندی، ابتکار عمل فردی، افزایش مسئولیت‌پذیری، تمایل به نوآوری، مشارکت در پروژه‌های توسعه‌ای، بازخوردگیری فعال، تلاش برای خودبهبودگری، هم‌سویی با فرهنگ دیجیتال سازمان	درگیری و مشارکت فعال	

تحلیل عمیق یافته‌های کیفی نشان می‌دهد که استقرار هوش مصنوعی مسئولانه، از یک مداخله صرفاً فناورانه فراتر رفته و به‌مثابهٔ یک دگردیسی عمیق در اکوسیستم منابع انسانی سازمان عمل می‌کند. نتایج حاکی از آن است که غلبه بر پدیدهٔ «جعبهٔ سیاه» از طریق شفافیت و تبیین‌پذیری الگوریتم‌ها، شرط لازم برای اعتمادسازی است؛ اما شرط کافی نیست. این شفافیت تنها زمانی به «امنیت روان‌شناختی» منجر می‌شود که با پاسخ‌گویی اخلاقی و صیانت از حریم خصوصی درهم‌آمیزد. در چنین بستر امنی است که سازوکارهای روان‌شناختی کارکنان فعال شده و با احیای احساس عاملیت و ادراک انصاف، مسیر برای شکوفایی، رشد حرفه‌ای و مشارکت فعال آنان هموار می‌شود.

یافته‌های پژوهش

به‌منظور کشف روابط ساختاریافته و سطح‌بندی ۱۵ مضمون فرعی استخراج‌شده، از تکنیک مدل‌سازی ساختاری تفسیری استفاده شد. در گام اول، ماتریس خودتعاملی ساختاری بر اساس خرد جمعی پنل خبرگان تشکیل شد (جدول ۴). در این ماتریس، جهت روابط میان جفت‌متغیرهای i و j با استفاده از نمادهای استاندارد مشخص شد: V (متغیر i بر j تأثیر دارد)، A (متغیر j بر i تأثیر دارد)، X (رابطه دوسویه) و O (عدم وجود رابطه).

جدول ٤. ماتریس ساختار روابط درونی متغیرها

	۱	۲	۳	٤	٥	٦	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱٤	۱٥
۱	-	A	A	X	A	O	O	A	V	V	V	O	O	O	O
۲	V	-	A	A	A	V	X	A	X	V	V	V	O	O	V
۳	V	V	-	V	V	V	V	V	V	V	V	V	V	V	V
۴	X	V	A	-	A	O	V	A	V	V	V	O	O	O	O
۵	V	V	A	V	-	O	O	A	V	V	V	O	O	O	O
۶	O	A	A	O	O	-	O	A	O	O	V	V	O	O	O
۷	O	X	A	A	O	O	-	A	X	V	V	V	O	O	V
۸	V	V	A	V	V	V	V	-	V	V	V	V	V	V	V
۹	A	X	A	A	A	O	X	A	-	V	V	O	O	O	V
۱۰	A	A	A	A	A	O	A	A	A	-	V	V	O	V	V
۱۱	A	A	A	A	A	A	A	A	A	A	-	X	O	V	V
۱۲	O	A	A	O	O	A	A	A	O	A	X	-	O	O	V
۱۳	O	O	A	O	O	O	O	A	O	O	O	O	-	V	V
۱۴	O	O	A	O	O	O	O	A	O	A	A	O	A	-	V
۱۵	O	A	A	O	O	O	A	A	A	A	A	A	A	A	-

در گام بعد، بر اساس قواعد تبدیل باینری، ماتریس خودتعاملی ساختاری به «ماتریس دسترسی اولیه» با درایه‌های صفر و یک تبدیل شد؛ بدین صورت که قطر اصلی ماتریس برابر با ۱ درنظر گرفته شد.

جدول ۵. ماتریس دسترسی اولیه

	۱	۲	۳	٤	٥	٦	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱٤	۱٥
۱	۱	۰	۰	۱	۰	۰	۰	۰	۱	۱	۱	۰	۰	۰	۰
۲	۱	۱	۰	۰	۰	۱	۱	۰	۱	۱	۱	۱	۰	۰	۱
۳	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱
۴	۱	۱	۰	۱	۰	۰	۱	۰	۱	۱	۱	۰	۰	۰	۰
۵	۱	۱	۰	۱	۱	۰	۰	۰	۱	۱	۱	۰	۰	۰	۰
۶	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۱	۱	۰	۰	۰
۷	۰	۱	۰	۰	۰	۰	۱	۰	۱	۱	۱	۱	۰	۰	۱
۸	۱	۱	۰	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱
۹	۰	۱	۰	۰	۰	۰	۱	۰	۱	۱	۱	۰	۰	۰	۱
۱۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۱	۱	۰	۱	۱
۱۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۱	۰	۱	۱
۱۲	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۱
۱۳	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۱	۱
۱۴	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۱
۱۵	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱

مهم‌ترین گام در مدل‌سازی ساختاری ـ تفسیری، کشف روابط پنهان سیستمی است. بر همین اساس، ماتریس اولیه با استفاده از الگوریتم محاسباتی «فلوید ـ وارشال[1]» در نرم‌افزار اکسل مورد تحلیل قرار گرفت تا «اصل تراگذری» بر آن اعمال شود. منطق ریاضی این اصل بیان می‌کند که اگر عامل F_i بر عامل F_j اثرگذار باشد و عامل F_j نیز بر عامل F_k تأثیر بگذارد، آنگاه عامل F_i نیز به‌صورت غیرمستقیم بر عامل F_k اثر خواهد داشت. با تکرار این الگوریتم تا رسیدن به ثبات ساختاری، ماتریس دسترسی نهایی شکل گرفت که دربردارندهٔ کلیهٔ روابط مستقیم و غیرمستقیم متغیرهاست (جدول۵) (آلاوامله، الطوال، لحلوح و جامع[2]، ۲۰۲۳؛ داس، کیرون و پرامود[3]، ۲۰۲۴). به‌منظور سطح‌بندی متغیرها و ترسیم ساختار سلسله‌مراتبی مدل، مجموعه‌های «دسترسی[4]» و «پیشایندی[5]» برای هر متغیر استخراج شد. مجموعه دسترسی شامل تمامی عواملی است که یک عامل می‌تواند به آن‌ها دسترسی مستقیم یا غیرمستقیم داشته باشد و مجموعه پیشایندی شامل تمامی عواملی است که بر آن عامل اثرگذار هستند. بر اساس الگوریتم کانونی مدل‌سازی ساختاری ـ تفسیری، در هر تکرار، متغیرهایی که «مجموعه دسترسی» آن‌ها با «مجموعه اشتراک» کاملاً برابر بود، به‌عنوان عناصر بالاترین سطح سیستم شناسایی شدند (وبر[6]، ۲۰۲۱). دلیل ریاضی این امر آن است که این متغیرها، تأثیرپذیرترین عناصر سیستم هستند و توانایی اثرگذاری بر سایر سطوح را ندارند. پس از شناسایی این متغیرها، آن‌ها از فهرست حذف شدند و فرایند تا سطح‌بندی تمامی ۱۵ متغیر ادامه یافت (جدول ۶).

1. Floyd-Warshall Algorithm
2. Alawamleh, Al-Twal, Lahlouh & Jame
3. Das, Kiron & Pram
4. Reachability Set
5. Antecedent Set
6. Weber

جدول ٦. مجموعه‌های دستیابی، پیشایندی و اشتراک شاخص‌ها برای تعیین سطوح

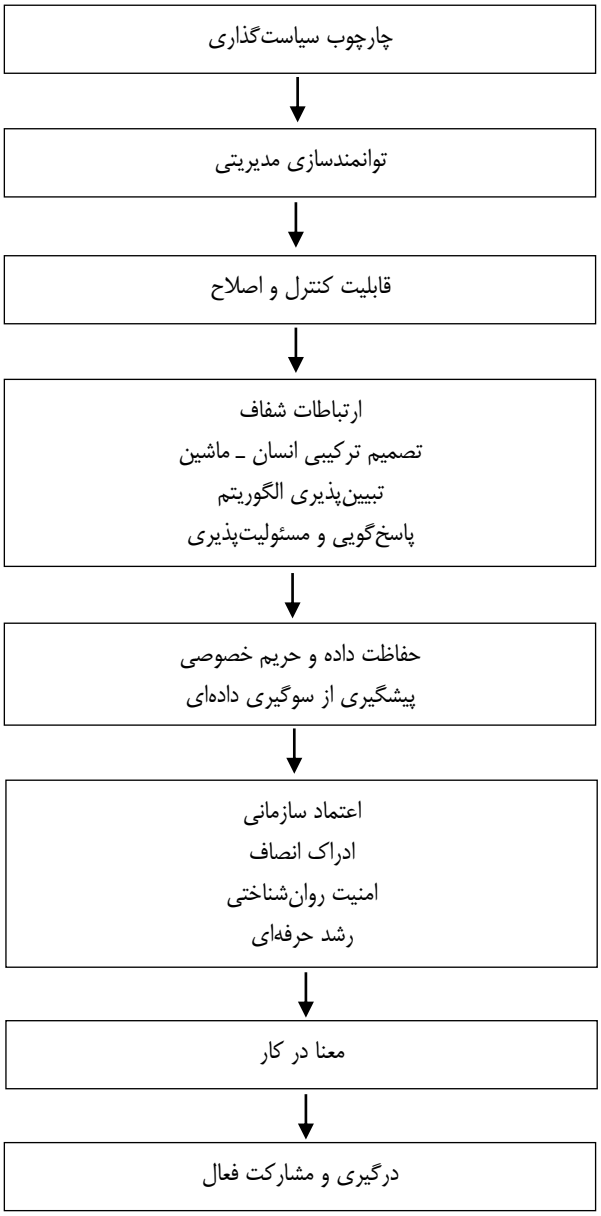
شاخص	مجموعه دستیابی	مجموعه پیشایندی	اشتراک
۱	F۱, F۲, F۴, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۱, F۲, F۳, F۴, F۵, F۷, F۸, F۹	F۱, F۲, F۴, F۷, F۹
۲	F۱, F۲, F۴, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۱, F۲, F۳, F۴, F۵, F۷, F۸, F۹	F۱, F۲, F۴, F۷, F۹
۳	F۱–F۱۵	F۳	F۳
۴	F۱, F۲, F۴, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۱, F۲, F۳, F۴, F۵, F۷, F۸, F۹	F۱, F۲, F۴, F۷, F۹
۵	F۱, F۲, F۴, F۵, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۳, F۵, F۸	F۵
۶	F۶, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۲, F۳, F۴, F۵, F۶, F۷, F۸, F۹	F۶
۷	F۱, F۲, F۴, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۱, F۲, F۳, F۴, F۵, F۷, F۸, F۹	F۱, F۲, F۴, F۷, F۹
۸	F۱, F۲, F۴, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۳, F۸	F۸
۹	F۱, F۲, F۴, F۶, F۷, F۹, F۱۰, F۱۱, F۱۲, F۱۴, F۱۵	F۱, F۲, F۳, F۴, F۵, F۶, F۷, F۸, F۹	F۹
۱۰	F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۱– F۱۲	F۱۰, F۱۱, F۱۲
۱۱	F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۱– F۱۲	F۱۰, F۱۱, F۱۲
۱۲	F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۱– F۱۲	F۱۰, F۱۱, F۱۲
۱۳	F۱۳, F۱۴, F۱۵	F۳, F۸, F۱۳	F۱۳
۱۴	F۱– F۱۴	F۹, F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۱۴
۱۵	F۱– F۱۵	F۸, F۹, F۱۰, F۱۱, F۱۲, F۱۳, F۱۴, F۱۵	F۱۵

جدول ۷. تعیین سطح شاخص‌های مدل‌سازی ساختاری ـ تفسیری

شماره سطح	شاخص‌های هر سطح	کارکرد اصلی
سطح ۱	درگیری و مشارکت فعال (۱۵)	وابسته: تأثیرپذیری بسیار بالا و تأثیرگذاری بسیار کم. این شاخص خروجی نهایی مدل است.
سطح۲	معنا در کار (۱۴)	وابسته: متغیرهای هدف که بیشتر تحت تأثیر سطوح پایین‌تر قرار دارند.
سطح۳	ادراک انصاف (۱۰)، اعتماد سازمانی (۱۱)، امنیت روان‌شناختی (۱۲)، رشد حرفه‌ای (۱۳)	رابط / پیوندی: متغیرهای میانی که هم تأثیرپذیری و هم تأثیرگذاری نسبتاً بالایی دارند و نقش انتقال‌دهنده را ایفا می‌کنند.
سطح۴	حفاظت داده و حریم خصوصی (۶)، پیشگیری از سوگیری داده‌ای (۱)	رابط / پیوندی: متغیرهای واسطه‌ای در زنجیره تأثیرات.
سطح۵	ارتباطات شفاف (۷)، تصمیم ترکیبی انسان–ماشین (۹)، تبیین‌پذیری الگوریتم (۴)، پاسخ‌گویی و مسئولیت‌پذیری (۲)	رابط / پیوندی: متغیرهای واسطه‌ای که به سمت استقلال میل می‌کنند.
سطح۶	قابلیت کنترل و اصلاح (۵)	مستقل / نفوذی: تأثیرگذاری بالا و تأثیرپذیری کم.
سطح۷	توانمندسازی مدیریتی (۸)	مستقل / نفوذی: متغیرهای پایه‌ای با قدرت هدایت بالا.
سطح۸	چارچوب سیاست‌گذاری (۳)	مستقل: سنگ بنای مدل. این شاخص کلیدی‌ترین محرک سیستم است که بر تمام شاخص‌های دیگر تأثیر می‌گذارد اما از هیچ‌کدام تأثیر نمی‌پذیرد.

خروجی نهایی این فرایند، استخراج یک مدل ساختاری هشت‌سطحی است (جدول ۷).

مدل نهایی پژوهش که با استفاده از روش مدل‌سازی ساختاری ـ تفسیری توسعه یافته است، ساختار سلسله‌مراتبی مؤلفه‌های مرتبط با اخلاق هوش مصنوعی را در ۸ سطح ترسیم می‌کند (شکل ۱). این مدل، الگوی ارتباطات و تعاملات ساختاری میان متغیرها را در قالب یک نظام سطح‌بندی‌شده نشان می‌دهد؛ به‌گونه‌ای که مؤلفه‌های زیربنایی در سطوح پایین‌تر و مؤلفه‌های وابسته‌تر در سطوح بالاتر قرار گرفته‌اند.



شکل ۱. مدل نهایی پژوهش

در قاعده مدل (سطح ۸)، «چارچوب سیاست‌گذاری» به‌عنوان پیشران ساختاری و بنیادی‌ترین مؤلفه قرار دارد. این مؤلفه در نظام روابط، دارای بیشترین قدرت نفوذ بوده و بستر اولیه برای تعاملات سایر مؤلفه‌ها در سیستم را فراهم

می‌آورد. در سطوح ۷ و ۶، مؤلفه‌های «توانمندسازی مدیریتی» و «قابلیت کنترل و اصلاح» قرار گرفته‌اند که در شبکه روابط، نقش پیونددهنده سیاست‌های کلان به الزامات عملیاتی و فنی را ایفا می‌کنند. سطوح میانی (۵، ۴ و ۳)، به‌عنوان کانون تعاملات ساختاری در مدل شناخته می‌شوند؛ به‌گونه‌ای که مؤلفه‌هایی همچون «پاسخ‌گویی» و «تبیین‌پذیری» در پیوند با سازوکارهای فنی نظیر «پیشگیری از سوگیری» قرار دارند. این سازوکارها در ساختار مدل، با مؤلفه‌های ادراکی ـ روان‌شناختی مانند «ادراک انصاف» و «اعتماد سازمانی» در یک شبکه ارتباطی متقابل قرار می‌گیرند. در نهایت، در سطوح ۲ و ۱، مؤلفه‌های وابسته به سایر متغیرها جای گرفته‌اند. «معنا در کار» به‌عنوان یک پیامد روان‌شناختی کلیدی، در بالاترین سطح مدل، جایگاه وابسته‌تری نسبت به سایر مؤلفه‌ها دارد و با مؤلفه «درگیری و مشارکت فعال» در یک سطح از تحلیل ساختاری قرار می‌گیرد. این مدل نشان می‌دهد که استقرار هوش مصنوعی در سازمان‌ها مستلزم یک رویکرد جامع و چندلایه است که از تدوین سیاست‌های ساختاری در سطح پایه آغاز شده و تا برقراری ارتباط با ادراک انصاف و اعتماد در سطح فردی امتداد می‌یابد.

نتیجه‌گیری و پیشنهادها

پژوهش حاضر با هدف ارائه یک مدل یکپارچه برای پیاده‌سازی هوش مصنوعی مسئولانه در مدیریت منابع انسانی و تبیین جایگاه آن در شکوفایی کارکنان انجام شد. یافته کلیدی پژوهش، تدوین یک مدل ساختاری ـ تفسیری هشت‌سطحی است که نشان می‌دهد دستیابی به شکوفایی کارکنان، از طریق هوش مصنوعی نه یک پیامد مستقیم فناوری، بلکه نتیجهٔ زنجیرهٔ تعاملات ساختاریِ پیچیده است که از سطح حاکمیتی آغاز شده و از طریق سازوکارهای میانجی به پیامدهای روان‌شناختی و رفتاری در سطوح بالاتر منتهی می‌شود. این یافته نشان می‌دهد که رابطه میان هوش مصنوعی و شکوفایی کارکنان، رابطه‌ای خطی و فناورانه نیست، بلکه رابطه‌ای چندلایه و وابسته به هم‌ترازی ساختاری میان زیرسیستم فنی و زیرسیستم اجتماعی سازمان است. این یافته، دقیقاً همان شکاف تبیینی مطرح شده در بخش پیشینه را پُر می‌کند؛ چرا که نشان می‌دهد چگونه اصول هنجاریِ هوش مصنوعی مسئولانه، از طریق این زنجیره تعاملات، از سطح «الزامات سیاستی» فراتر می‌رود و به «تجربیات زیسته و پیامدهای انسانی» در محیط کار تبدیل می‌شوند.

بر خلاف برخی پژوهش‌ها که نقطهٔ شروع را مستقیماً فناوری می‌دانند (برای نمونه تامبه، کاپلی و یاکوبوویچ[1]، ۲۰۱۹)، مدل حاضر نشان می‌دهد که «چارچوب سیاست‌گذاری» (سطح ۸) به‌عنوان یک متغیر مستقل، بنیادی‌ترین عنصر در ساختار سیستم است. در واقع، برخلاف رویکردهای صرفاً هنجاریِ چارچوب‌های بین‌المللی (مانند مایکروسافت و گوگل) که بر «محتوای» اصول تمرکز دارند، یافته‌های این پژوهش نشان می‌دهد که اهمیت اصلی در «ساختارِ عملیاتی‌سازی» این اصول نهفته است؛ به این معنا که سیاست‌گذاری تنها زمانی به تغییر در رفتار کارکنان منجر می‌شود که به‌عنوان ریشهٔ یک زنجیرهٔ ساختاری عمل کند. این نتیجه، از جنبهٔ نظریهٔ سیستم‌های اجتماعی ـ فنی قابل تبیین

1. Tambe, Cappelli & Yakubovich

است؛ زیرا در این چارچوب، پیش‌نیاز بهینه‌سازی مشترک میان فناوری و انسان، وجود سازوکارهای حاکمیتی است که جهت‌گیری ارزشی و اخلاقی سیستم را تعیین می‌کنند. این یافته با ادبیات حاکمیت فناوری نیز هم‌سو است که استدلال می‌کند بدون وجود خط‌مشی‌های روشن، شفاف و اخلاقی، استقرار فناوری‌های پیشرفته با بی‌نظمی نهادی و تضعیف اعتماد سازمانی همراه خواهد بود (وون کروگ[1]، ۲۰۱۲) در این معنا، چارچوب سیاست‌گذاری فقط یک سند اداری نیست، بلکه یک پیشران ساختاری است که بستر نهادی لازم برای «توانمندسازی مدیریتی» (سطح ۷) و «قابلیت کنترل و اصلاح» (سطح ۶) را فراهم می‌کند. پژوهش حاضر بدین ترتیب استدلال می‌کند که مسئولیت‌پذیری هوش مصنوعی پیش از آنکه در سطح الگوریتم‌ها تعریف شود، باید در سطح استراتژی و سیاست‌های کلان سازمان نهادینه شود.

قلب مدل در سطوح میانی قرار دارد؛ جایی که متغیرهای رابط، انتقال الزامات حاکمیتی به تجربهٔ زیسته کارکنان را ممکن می‌سازند. از جنبهٔ نظری، این بخش از مدل با اصول طراحی سیستم‌های اجتماعی ـ فنی چرنز (۱۹۷۶) و ادبیات عدالت سازمانی پیوند می‌خورد. در چارچوب عدالت سازمانی، کارکنان فقط به نتایج تصمیمات سازمانی توجه نمی‌کنند، بلکه شیوهٔ اتخاذ این تصمیمات و امکان مشارکت یا بازبینی آن‌ها را نیز در ارزیابی خود دخیل می‌دانند. در محیط‌های مبتنی بر هوش مصنوعی، زمانی که منطق تصمیمات الگوریتمی قابل درک و تبیین باشد، امکان بازبینی و اصلاح وجود داشته باشد و سازوکارهای کنترل سوگیری فعال باشند، عدالت رویه‌ای تقویت می‌شود و ادراک انصاف در میان کارکنان شکل می‌گیرد. در ساختار مدل، این سازوکارها پیش‌نیازهای سطح ۳، یعنی «اعتماد سازمانی»، «امنیت روان‌شناختی» و «ادراک انصاف» را تشکیل می‌دهند. این نتیجه با نظریهٔ تبادل اجتماعی (بلاو[2]، ۱۹۵۴) نیز هم‌خوان است؛ نظریه‌ای که بر مبادلهٔ متقابل میان رفتار منصفانه سازمان و واکنش‌های مثبت کارکنان تأکید دارد. در این چارچوب، شفافیت، پاسخ‌گویی و پیشگیری از سوگیری، به‌منزلهٔ سیگنال‌هایی از احترام و ارزش‌گذاری سازمان به کارکنان عمل می‌کنند و به شکل‌گیری اعتماد، تعهد و مشارکت منجر می‌شوند. بنابراین، اعتماد در مدل حاضر نه یک پیش‌فرض ایستا، بلکه یک سازه برساخته در شبکه روابط است که از طریق عدالت، شفافیت و پاسخ‌گویی تولید و بازتولید می‌شود.

در این فرایند، تقویت سطوح میانی به ایجاد «ساختارهای خودسازمان‌ده[3]» کمک می‌کند؛ ساختارهایی که امکان حفظ عاملیت انسانی و تنظیم تعادل میان سیستم فنی و سیستم انسانی را فراهم می‌سازند. اهمیت این سازوکارها در تجربه‌های واقعی سازمانی نیز قابل مشاهده است. برای مثال، تجربهٔ شرکت آمازون در توسعهٔ سامانهٔ غربالگری رزومه مبتنی بر هوش مصنوعی نشان داد که آموزش الگوریتم بر داده‌های تاریخی سوگیرانه می‌تواند به بازتولید تبعیض جنسیتی منجر شود. این تجربهٔ آشکار ساخت که حتی سامانه‌هایی با عملکرد فنی مناسب، در صورت فقدان سازوکارهای شفافیت، نظارت انسانی و کنترل سوگیری، با کاهش ادراک انصاف و تضعیف اعتماد ذی‌نفعان همراه خواهند بود. از جنبهٔ عدالت سازمانی، چنین وضعیتی به تضعیف عدالت رویه‌ای می‌انجامد؛ از جنبهٔ نظریه تبادل اجتماعی، پیام بی‌توجهی سازمان به دغدغه‌های کارکنان را منتقل می‌کند؛ و از جنبهٔ سیستم‌های اجتماعی ـ فنی، نشانه شکست در بهینه‌سازی

1. von Krogh
2. Blau
3. Self-organizing Structures

مشترک و غلبه زیرسیستم فنی بر زیرسیستم اجتماعی است. یافته‌های پژوهش حاضر نیز نشان می‌دهد که پیشگیری از سوگیری، تبیین‌پذیری و پاسخ‌گویی، پیش‌نیازهای اساسی شکل‌گیری اعتماد سازمانی و امنیت روان‌شناختی در شبکه روابط هستند. بنابراین، تجربهٔ آمازون به‌عنوان یک شاهد تجربی، دقیقاً تأییدکننده مدل حاضر است: فقدان سازوکارهای انتقال در سطوح میانی (عدالت و شفافیت)، باعث می‌شود که حتی پیشرفته‌ترین زیرسیستم فنی، نتواند به شکوفایی زیرسیستم اجتماعی منجر شود و در نهایت در شکستِ بهینه‌سازی مشترک متوقف شود.

در نهایت، متغیرهای وابسته مدل یعنی «معنا در کار» (سطح ۲) و «درگیری و مشارکت فعال» (سطح ۱) به‌عنوان پیامدهای نهایی این زنجیرهٔ ساختاری ظاهر می‌شوند. این نتیجه دیدگاه‌هایی را که هوش مصنوعی را ذاتاً تهدیدی برای عاملیت انسانی می‌دانند، به چالش می‌کشد. پژوهش حاضر نشان می‌دهد که اگر هوش مصنوعی از طریق سطوح ۸ تا ۳ و در چارچوب یک معماری مسئولانه پیاده‌سازی شود، می‌تواند با خودکارسازی وظایف تکراری، ارائه بازخوردهای توسعه‌ای و افزایش قابلیت تصمیم‌گیری آگاهانه، نه‌تنها عاملیت انسانی را تضعیف نکند، بلکه به شکوفایی آن کمک کند (اسپریتزر و همکاران، ۲۰۰۵).

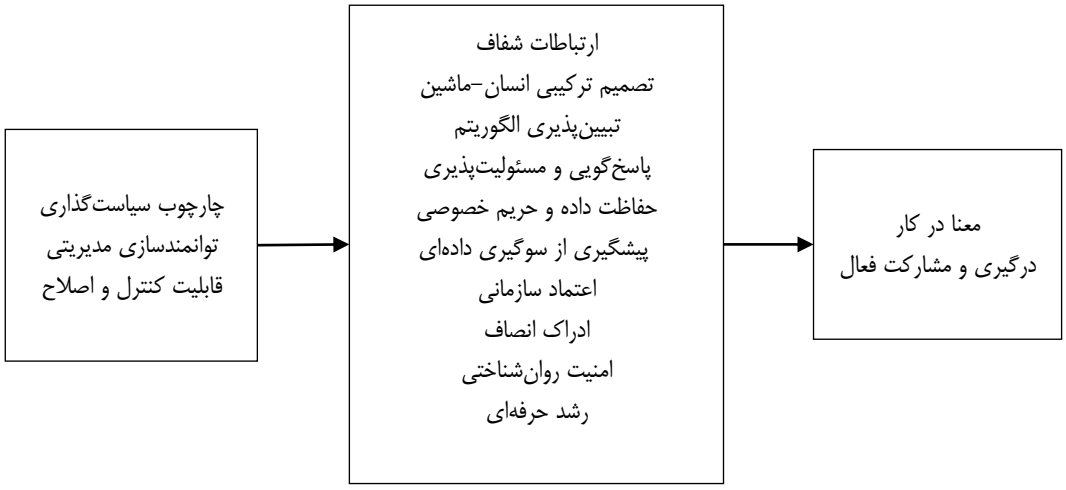
اگرچه مدل نهایی در قالب هشت سطح سلسله‌مراتبی ارائه شده است، به‌منظور ارتقای قابلیت تفسیر نظری و تسهیل کاربرد مدیریتی، این سطوح در قالب سه لایه مفهومی کلان بازنمایی شده‌اند. مبنای این لایه‌بندی، منطق نظری «نظریه سیستم‌های اجتماعی ـ فنی» است که بر ضرورت هم‌ترازی و بهینه‌سازی مشترک میان زیرسیستم فنی و زیرسیستم اجتماعی در سازمان تأکید دارد. بر این اساس، سطوح پایین‌تر مدل که نقش پیشران‌های ساختاری را ایفا می‌کنند در قالب لایه «حاکمیتی ـ فنی» (سطوح ۸ تا ۶) صورت‌بندی شده‌اند. این لایه به‌عنوان زیرسیستم فنی شامل مؤلفه‌هایی نظیر چارچوب‌های سیاست‌گذاری، سازوکارهای حاکمیت داده و الگوریتم، توانمندسازی مدیریتی و قابلیت کنترل، نظارت و اصلاح سامانه‌های هوش مصنوعی است که بستر نهادی و فناورانه استقرار هوش مصنوعی مسئولانه را فراهم می‌سازند.

در سطح بعدی، لایه «اجتماعی ـ روان‌شناختی» (سطوح ۵ تا ۳) قرار دارد که نمایانگر زیرسیستم اجتماعی سازمان است. این لایه بیانگر سازوکارهایی است که از طریق آن‌ها ویژگی‌های زیرسیستم فنی به ادراکات، نگرش‌ها و تجربه‌های کارکنان منتقل می‌شود. مؤلفه‌هایی نظیر عدالت ادراک‌شده، شفافیت تصمیم‌های الگوریتمی، اعتماد به سامانه‌های هوش مصنوعی و امنیت روان‌شناختی در این سطح قرار می‌گیرند و نقش میانجی را در تبدیل الزامات فنی و حاکمیتی به پیامدهای انسانی ایفا می‌کنند.

در نهایت، لایه «پیامدی» (سطوح ۲ تا ۱) قرار دارد که منعکس‌کنندهٔ برون‌دادهای انسانی و سازمانی حاصل از هم‌ترازی موفق زیرسیستم‌های فنی و اجتماعی است. در این سطح، پیامدهایی نظیر معنا در کار، مشارکت فعال کارکنان و در نهایت شکوفایی کارکنان به‌عنوان نتایج نهایی پیاده‌سازی مسئولانه هوش مصنوعی در مدیریت منابع انسانی ظهور می‌یابد.

بنابراین، بازنمایی سه‌لایه‌ای ارائه‌شده در این پژوهش تفسیری نظری از ساختار روابط نفوذ استخراج‌شده از مدل‌سازی ساختاری ـ تفسیری است که با اتکا به منطق نظریه سیستم‌های اجتماعی ـ فنی صورت‌بندی شده است. این

تفسیر مفهومی، بدون ایجاد تغییر در ساختار سلسله‌مراتبی مدل، امکان درک بهتر منطق علّی میان سطوح مختلف را فراهم ساخته و در مرحله نهایی تحلیل نیز با در نظر گرفتن دیدگاه‌های خبرگان حوزه مدیریت منابع انسانی و هوش مصنوعی مورد تأیید قرار گرفته است. از این رو، این لایه‌بندی علاوه بر تقویت انسجام نظری مدل، کاربردپذیری آن را برای سیاست‌گذاران و مدیران در مسیر پیاده‌سازی مسئولانه هوش مصنوعی در سازمان‌ها افزایش می‌دهد.



شکل ۲. مدل تسهیل‌یافته نهایی پژوهش

این پژوهش چندین مشارکت نظری مهم دارد:

- یکپارچه‌سازی ادبیات: این مطالعه با ارائه یک مدل سلسله‌مراتبی، سه حوزه مجزای ادبیات را به هم پیوند می‌زند: حاکمیت فناوری اطلاعات (سطوح بالا)، عدالت سازمانی (سطوح میانی)، و رفتار سازمانی مثبت‌گرا (سطوح پایین). این مدل تبیین می‌کند که چگونه مفاهیم کلان حاکمیتی در یک پیوند ساختاری با پیامدهای خرد روان‌شناختی قرار می‌گیرند.
- عمق‌بخشی به نقش اعتماد: در حالی که بسیاری از مطالعات بر اهمیت «اعتماد به هوش مصنوعی» تأکید می‌کنند، این پژوهش با تعیین جایگاه «اعتماد سازمانی» در سطوح میانی مدل، نشان می‌دهد که اعتماد، بخشی از یک شبکه تعاملی گسترده‌تر است که سازمان برای مدیریت منصفانه فناوری ایجاد می‌کند.
- ارائه مدل ساختاری برای هوش مصنوعی مسئولانه: این پژوهش فراتر از فهرست کردن اصول اخلاقی، یک مدل ساختاری ارائه می‌دهد که جایگاه و اولویت ساختاریِ هر یک از این اصول را در مسیر تحقق نتایج مرتبط با کارکنان مشخص می‌کند.
- بسط چارچوب‌های هوش مصنوعی مسئولانه به حوزه پیامدهای انسانی: یکی از مشارکت‌های نظری مهم این پژوهش، انتقال کانون توجه از اصول حاکمیتی هوش مصنوعی مسئولانه به جایگاه آن‌ها در تجربیات انسانی

در محیط کار است. مدل ارائه‌شده آشکار می‌سازد که مؤلفه‌های فنی هوش مصنوعی مسئولانه زمانی با شکوفایی کارکنان هم‌راستا می‌شوند که از طریق اعتماد سازمانی، ادراک انصاف و امنیت روان‌شناختی در تجربه کارکنان بازتولید شوند. از این منظر، پژوهش حاضر فهم موجود از هوش مصنوعی مسئولانه را از سطح طراحی و حاکمیت فناوری به سطح تعاملات انسانی و پیامدهای رفتاری در سازمان گسترش می‌دهد.

یافته‌های این پژوهش راهنمایی‌های روشنی برای مدیران و رهبران سازمان‌ها فراهم می‌کند:

- از سطوح بنیادین شروع کنید: مدیران ارشد باید پیش از استقرار سیستم‌های هوش مصنوعی، ابتدا یک «چارچوب سیاست‌گذاری» جامع، شفاف و اخلاقی تدوین کنند که به‌عنوان پیشران اصلی سیستم عمل کند.
- روی سازوکارهای ساختاری سرمایه‌گذاری کنید: سازمان‌ها باید به دنبال ابزارهایی باشند که در ساختار عملیاتی خود دارای قابلیت «تبیین‌پذیری» و «کنترل انسانی» هستند.
- اعتماد را در شبکه روابط مدیریت کنید: هدف اصلی باید ایجاد بستری برای پذیرش و اعتماد کارکنان باشد. مدیران باید از هوش مصنوعی به‌عنوان ابزاری برای توانمندسازی و رشد حرفه‌ای استفاده کنند.

علاوه‌براین، یافته‌های پژوهش نشان می‌دهد که مسئولیت استقرار هوش مصنوعی مسئولانه صرفاً بر عهده واحد فناوری اطلاعات نیست و نیازمند همکاری میان مدیران منابع انسانی، مدیران فناوری و مدیران ارشد سازمان است. از جنبهٔ اجرایی، سازمان‌ها می‌توانند کمیته‌های حاکمیت هوش مصنوعی تشکیل دهند، ممیزی‌های دوره‌ای سوگیری الگوریتمی را انجام دهند، سازوکارهای اعتراض و بازبینی تصمیمات مبتنی بر هوش مصنوعی را برای کارکنان فراهم سازند و شاخص‌های اعتماد، ادراک انصاف و امنیت روان‌شناختی را به‌عنوان بخشی از نظام پایش عملکرد استقرار هوش مصنوعی مورد ارزیابی قرار دهند. چنین اقداماتی می‌تواند به کاهش مقاومت کارکنان در برابر فناوری و افزایش پذیرش و اثربخشی سامانه‌های هوش مصنوعی کمک کند.

هر پژوهشی با محدودیت‌هایی همراه است که باید در تفسیر یافته‌ها مورد توجه قرار گیرند.

نخست، مدل پیشنهادی این پژوهش بر پایه قضاوت و دانش تخصصی خبرگان توسعه یافته است. اگرچه بهره‌گیری از خبرگان در مطالعات اکتشافی و روش مدل‌سازی ساختاری ـ تفسیری رویکردی پذیرفته‌شده برای شناسایی و ساختاردهی پدیده‌های پیچیده و نوظهور محسوب می‌شود، اما ممکن است همه پیچیدگی‌ها و تفاوت‌های موجود در محیط‌های واقعی سازمانی را منعکس نکند. از این رو، یافته‌های پژوهش باید در چارچوب یک مدل ساختاری مبتنی بر ادراکات و قضاوت‌های تخصصی تفسیر شوند.

محدودیت دیگر، به ویژگی‌های پنل خبرگان پژوهش برمی‌گردد. نخست اینکه حجم نمونه خبرگان نسبتاً محدود بوده و مشارکت بخشی از آن‌ها در هر دو مرحله پژوهش، اگرچه به حفظ انسجام مفهومی کمک کرده، اما ممکن است احتمال شکل‌گیری «سوگیری توافقی[1]» را افزایش داده باشد. همچنین، تمرکز نمونه‌گیری در این پنل بیشتر بر

1. Consensus Bias

تخصص‌های کارکردی (مدیریت منابع انسانی، توسعه هوش مصنوعی و اخلاق فناوری) بوده و تنوع صنایع مورد بررسی در آن محدود است. از آنجا که استقرار هوش مصنوعی در صنایع مختلف (نظیر بانکداری، نظام سلامت، تولید و...) می‌تواند چالش‌ها و پیامدهای متفاوتی برای شکوفایی کارکنان به همراه داشته باشد، این مسئله تعمیم‌پذیری دقیق نتایج به بافتارهای خاص صنعتی را با محدودیت مواجه می‌سازد. ازاین‌رو، پیشنهاد می‌شود در پژوهش‌های آتی علاوه بر استفاده از گروه‌های مستقل خبرگان در مراحل مختلف، مدل پیشنهادی با مشارکت طیف گسترده‌تری از ذی‌نفعان در صنایع مختلف مورد بررسی و آزمون قرار گیرد.

دوم، روش مدل‌سازی ساختاری ـ تفسیری با هدف شناسایی، ساختاردهی و سطح‌بندی روابط میان متغیرهای یک سیستم پیچیده به کار می‌رود و ماهیت آن مبتنی بر تحلیل قضاوت خبرگان است. بنابراین، روابط شناسایی‌شده در این پژوهش بیانگر الگوهای ساختاری، جهت‌دار و نفوذی میان مؤلفه‌ها بوده و نباید به منزله اثبات روابط علّی در معنای آماری یا تجربی آن تلقی شوند. افزون بر این، مدل نهایی در مرحله کنونی مبتنی بر دانش و قضاوت خبرگان بوده و با داده‌های میدانی کارکنان یا سازمان‌ها مورد آزمون قرار نگرفته است؛ ازاین‌رو، فاقد اعتبارسنجی آماری و تحلیل برازش بوده و بیشتر ماهیتی اکتشافی و مفهومی دارد. همچنین، روش مدل‌سازی ساختاری–تفسیری کلاسیک امکان تحلیل حساسیت مبتنی بر تغییر وزن متغیرها را فراهم نمی‌کند، زیرا وزن عددی برای مؤلفه‌ها تعریف نمی‌شود. در نتیجه، پایداری ساختار مدل در برابر تغییرات احتمالی در شدت روابط میان مؤلفه‌ها مورد ارزیابی قرار نگرفت. پیشنهاد می‌شود پژوهش‌های آتی با استفاده از داده‌های تجربی و بهره‌گیری از روش‌هایی نظیر مدل‌سازی معادلات ساختاری، مدل‌سازی ساختاری ـ تفسیری فازی، میک‌مک فازی یا نگاشت شناختی فازی، اعتبار تجربی، برازش، پایداری و حساسیت مدل پیشنهادی را مورد بررسی قرار دهند.

سوم، پژوهش حاضر ماهیتی مقطعی دارد و روابط شناسایی‌شده را در یک مقطع زمانی مشخص بررسی کرده است. در حالی که فرایند استقرار و پذیرش هوش مصنوعی در سازمان‌ها ماهیتی پویا و تکاملی دارد، مطالعات طولی می‌توانند درک عمیق‌تری از چگونگی تحول این روابط در طول زمان ارائه دهند.

چهارم، مدل پیشنهادی در یک بستر فرهنگی و سازمانی خاص توسعه یافته است. از این رو، تعمیم نتایج به سایر صنایع، کشورها و زمینه‌های فرهنگی باید با احتیاط صورت گیرد. اجرای پژوهش‌های مشابه در محیط‌های سازمانی و فرهنگی متفاوت می‌تواند به ارزیابی پایداری و تعمیم‌پذیری ساختار پیشنهادی کمک کند.

بر این اساس، پیشنهاد می‌شود پژوهش‌های آینده با بهره‌گیری از داده‌های میدانی کارکنان و مدیران در سازمان‌های مختلف، مدل پیشنهادی را از طریق روش‌های کمی نظیر مدل‌سازی معادلات ساختاری مورد آزمون و تحلیل برازش قرار دهند تا جنبه‌های کاربردی و قابلیت تعمیم مدل ارتقا یابد. همچنین بررسی نقش متغیرهای زمینه‌ای و تعدیل‌گر، از جمله سواد دیجیتال کارکنان، آمادگی سازمانی برای پذیرش هوش مصنوعی، و سبک‌های رهبری، می‌تواند به توسعه و غنای بیشتر مدل ارائه‌شده کمک کند.

در مجموع، یافته‌های این پژوهش نشان می‌دهد که استقرار مسئولانه هوش مصنوعی در مدیریت منابع انسانی صرفاً یک اقدام فناورانه نیست، بلکه فرایندی چندلایه و فنی ـ اجتماعی است که موفقیت آن به کیفیت تعامل میان

زیرسیستم فنی و زیرسیستم اجتماعی سازمان وابسته است. در چارچوب نظریه سیستم‌های اجتماعی ـ فنی، مؤلفه‌هایی نظیر شفافیت، تبیین‌پذیری، عدالت الگوریتمی، پاسخ‌گویی و نظارت انسانی از طریق شکل‌دهی ادراک کارکنان نسبت به اعتماد، انصاف و امنیت روان‌شناختی عمل می‌کنند. این سازوکارهای اجتماعی به نوبه خود زمینه لازم برای مشارکت فعال، معنا در کار و شکوفایی کارکنان را فراهم می‌سازند. از این منظر، شکوفایی کارکنان پیامد مستقیم فناوری نیست، بلکه حاصل هم‌ترازی موفق میان الزامات فنی سامانه‌های هوش مصنوعی و نیازهای انسانی کارکنان است. بنابراین، مدل سلسله‌مراتبی ارائه‌شده می‌تواند به‌عنوان چارچوبی راهنما برای طراحی و استقرار نظام‌های هوش مصنوعی انسان‌محور و مسئولانه در سازمان‌ها مورد استفاده قرار گیرد.

منابع

حاجی اسماعیلی، حسین؛ طهماسبی، رضا و باباشاهی، جبار (۱۴۰۵). طراحی چارچوب توسعهٔ استعداد فراگیر درسازمان (مورد مطالعه: سازمان تأمین اجتماعی). *مدیریت دولتی*، ۱۸(۱)، ۲۰۳-۲۳۳.

رستگار، عباسعلی و موسوی، میناسادات (۱۴۰۲). بررسی اثر تنوع قومی اعضای تیم استارتاپ‌های ایرانی بر حجم سرمایه‌گذاری جمع‌آوری‌شده (مورد مطالعه: شرکت‌های حاضر در نمایشگاه GITEX دبی). *مدیریت دولتی*، ۱۵(۴)، ۷۵۱-۷۶۵.

موسوی، میناسادات؛ رستگار؛ عباسعلی و شفیعی نیکآبادی، محسن (۱۴۰۴). هوش مصنوعی در مدیریت منابع انسانی با رویکرد مثبت‌گرا: چارچوبی برای شکوفایی کارکنان. *مطالعات مدیریت کسب‌وکار هوشمند*، ۱۴(۵۴)، ۱۷۵-۲۱۶.

مؤمنی، امیررضا؛ یعقوبی، نورمحمد؛ روشن، سیدعلیقلی و پورعزت، علی اصغر (۱۴۰۵). طراحی الگوی پذیرش هوش مصنوعی در حکمرانی هوشمند با استفاده از رویکرد فراترکیب. *مدیریت دولتی*، ۱۸(۱)، ۲۶۳-۲۹۵.

واعظی، سید کمال و پاشایی، فرانک (۱۴۰۴). کاربرد هوش مصنوعی برای تولید تلنگرهای رفتاری هوشمند. *مدیریت دولتی*، ۱۷(۴)، ۹۶۲-۹۳۸.

References

Alawamleh, M., Al-Twal, A., Lahlouh, L. & Jame, R. O. (2023). Interpretive structural modelling of organizational innovation factors: An emerging market perspective. *Journal of Open Innovation: Technology, Market, and Complexity, 9*(2), 100067.

Bankins, S., Formosa, P. & O'Kane, P. (2024). The ethical implications of artificial intelligence (AI) in human resource management: A review and framework. *Human Resource Management Review, 34*(1), 100940.

Blau, P. M. (1964). *Exchange and power in social life*. Wiley.

Braun, V. & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology, 18*(3), 328-352

Budhwar, P., Malik, A., De Silva, M. T. & Thevisuthan, P. (2023). Artificial intelligence – challenges and opportunities for international HRM: a review and research agenda. *The International Journal of Human Resource Management, 34*(10), 1968-2007.

Bunjak, A., Černe, M. & Popovič, A. (2023). Artificial intelligence and employees' well-being: A systematic literature review and future research agenda. *Journal of Business Research, 162*, 113886.

Chang, Y. L. & Ke, J. (2024). Socially responsible artificial intelligence empowered people analytics: a novel framework towards sustainability. *Human Resource Development Review, 23*(1), 88-120.

Cherns, A. (1976). The principles of sociotechnical design. *Human relations, 29*(8), 783-792.

Chowdhury, S., Dey, P. K., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A. & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review,* 33(1), 100899.

Creswell, J. W. & Plano Clark, V. L. (2023). Revisiting mixed methods research designs twenty years later. *Handbook of mixed methods research designs*, *1*(1), 21-36.

Das, S. S., Pramod, V. R., & Kiron, K. (2024). Interpretive structural modeling (ISM): A literature review. *International Journal of Creative Research Thoughts*, *12*(5), 863-868.

European Commission. (2024). *Artificial Intelligence Act: Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence*. Official Journal of the European Union.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage.

Haji Esmaeili, H., Tahmasebi, R. & Babashahi, J. (2026). Designing an Inclusive Talent Development Framework in the Organization (Case Study: Social Security Organization). *Journal of Public Administration*, 18(1), 203-233. https://doi.org/10.22059/jipa.2025.399679.3749 (*in Persian*)

Khair, M. A., Mahadasa, R., Tuli, F. A. & Ande, J. R. P. K. (2020). Beyond human judgment: exploring the impact of artificial intelligence on HR decision-making efficiency and fairness. *Global Disclosure of Economics and Business, 9*(2), 163-176.

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Scherer, S. & Schildhauer, T. (2022). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics, 179*(3), 661-683.

Malik, A., Budhwar, P. & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HR decision making: A multi-level framework. *Human Resource Management Review, 33*(1), 100910.

Momeni, A., Yaghoubi, N. M., Roshan, S. A. & Pourezzat, A. A. (2026). The Design of an Artificial Intelligence Adoption Model in Smart Governance Using a Meta-Synthesis Approach. *Journal of Public Administration*, *18*(1), 263-295. doi: 10.22059/jipa.2025.398124.3737 (*in Persian*)

Mousavi, M., Rastgar, A.A. & Shafiei Nikabadi, M. (2025). Artificial Intelligence in Human Resource Management with a Positive Approach: A Framework for Employee Flourishing. *Business Intelligence Management Studies*, *14*(54), 175-216. doi: 10.22054/ims.2025.86244.2626 (*in Persian*)

National Institute of Standards and Technology (NIST). (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

Prikshat, V., Malik, A. & Budhwar, P. (2023). AI-augmented HRM: Antecedents, assimilation and multilevel consequences. *Human Resource Management Review, 33*(1), 100886.

Rastgar, A. & Mousavi, M. (2023). Investigating the Effect of Ethnic Diversity of Iranian Startup Team Members on the Investment Capital Raised (Study Case: Startups at Dubai Gitex Exhibition). *Journal of Public Administration*, *15*(4), 751-765. doi: 10.22059/jipa.2023.355429.3295 (*in Persian*)

Sadeghi, S. (2024). *Employee Well-being in the Age of AI: Perceptions, Concerns, Behaviors, and Outcomes*. arXiv preprint arXiv:2412.04796.

Schulte, P. A. & Streit, J. M. (2025). Psychosocial risks and ethical implications of technology: considerations for decent work. *Annals of Work Exposures and Health, 69*(4), 360-376.

Soomro, S., Fan, M., Sohu, J. M., Soomro, S. & Shaikh, S. N. (2024*). AI adoption: a bridge or a barrier? The moderating role of organizational support in the path toward employee well-being.* Kybernetes.

Spreitzer, G., Sutcliffe, K., Dutton, J., Sonenshein, S. & Grant, A. M. (2005). A socially embedded model of thriving at work. *Organization Science, 16*(5), 537–549. https://doi.org/10.1287/orsc.1050.0153

Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., … & Way, C. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review, 56*(3), 2097-2127.

Sushil. (2012). Interpreting the interpretive structural model. Global Journal of Flexible *Systems Management, 13*(2), 87-106.

Tambe, P., Cappelli, P. & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42. https://doi.org/10.1177/0008125619867910

Tripathi, A. & Kumar, V. (2025). Ethical practices of artificial intelligence: A management framework for responsible AI deployment in businesses. *AI and Ethics, 1-12.*

Trist, E. L. (1981). The evolution of socio-technical systems: A conceptual framework and an action research program. *Occasional paper,* 2(1981), 1981.

Vaezi, S. K. & Pashaei, F. (2025). Application of artificial intelligence for Generating Smart Behavioral Nudges. *Journal of Public Administration*, *17*(4), 964-988. doi: 10.22059/jipa.2025.395752.3700 (*in Persian*)

Valtonen, A., Kimpimäki, J. P. & Savela, N. (2026). Employee wellbeing: a computational review on the consequences of workplace automation. *Technovation, 152*, 103424.

von Krogh, G. (2012). How does social software change knowledge management? Toward a strategic research agenda. *The Journal of Strategic Information Systems, 21*(2), 154–164. https://doi.org/10.1016/j.jsis.2012.04.003

Vrontis, D., Christofi, M., Pereira, V., Tarba, S., Makrides, A. & Trichina, E. (2023). Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. *Artificial intelligence and international HRM*, 172-201. 55.

Weber, R. (2021). Constructs and Indicators: An Ontological Analysis. *MIS Quarterly, 45*(4).