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

# Strategies for Mitigating Cognitive Bias in the Application of Artificial Intelligence to Human Resource Recruitment and Selection

**Mohammad Bashokouh Ajirloo***
*Corresponding Author, Prof., Department of Business Management, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: bashokouh@uma.ac.ir

**Mehri Maghami Dolatabad**
PhD Candidate, Department of Business Management, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: mehri.maghamii@gmail.com

**Bahman Khodapanah**
Assistant Prof., Department of Business Management, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: khodapanah@ut.ac.ir

## Abstract

### Objective
This study sought to identify and explain strategies for mitigating cognitive bias in the application of artificial intelligence (AI) to human resource recruitment and selection. As intelligent systems have become more prevalent in human resource management, algorithms are increasingly used to screen résumés, analyze competencies, rank candidates, and support hiring decisions. When properly designed and used, such systems can improve the speed, consistency, and traceability of decisions. Yet their reliance on historical data, proxy variables, ambiguous criteria, and biased human judgments carries the risk of reproducing or even intensifying existing inequalities. In addition, the cognitive biases of decision-makers, in selecting data, interpreting model outputs, and endorsing algorithmic recommendations, can undermine the fairness of the recruitment process. Accordingly, the central problem addressed by this study was to identify the human, data-related, algorithmic, and organizational mechanisms for reducing cognitive bias in AI-based recruitment.

### Methods
The study adopted a qualitative approach and used thematic analysis. Data were gathered through 12 in-depth, semi-structured interviews with human resource experts, technology

**Citation:** Bashokouh Ajirloo, Mohammad; Maghami Dolatabad, Mehri & Khodapanah, Bahman (2026). Strategies for Mitigating Cognitive Bias in the Application of Artificial Intelligence to Human Resource Recruitment and Selection. *Journal of Public Administration*, 18(3), 723-751. (*in Persian*)

Journal of Public Administration, 2026, Vol. 18, No.3, pp. 723-751
Published by University of Tehran, College of Management
https://doi.org/10.22059/jipa.2026.417273.3944
Article Type: Research Paper
© Authors
Received: April 23, 2026
Revised: July 15, 2026
Accepted: August 20, 2026
Published online: September 14, 2026

specialists, and individuals familiar with the design or use of intelligent systems in Iranian organizations. Participants were selected through purposive sampling, and interviews continued until theoretical saturation was reached. After the interviews were transcribed, the data were examined repeatedly to identify meaningful units related to the sources, manifestations, and strategies for controlling bias. The analysis proceeded through initial coding, extraction of basic themes, comparison and consolidation of similar codes, organization of themes, and formation of overarching (global) themes, and was carried out using MAXQDA version 20. To strengthen the credibility of the analysis, codes were reviewed continuously, raw data were revisited, participants' perspectives were compared, and coherence between themes was examined. This process ultimately yielded 130 basic themes, 18 organizing themes, and 6 global themes.

**Results**

The final thematic structure indicated that strategies for mitigating cognitive bias in intelligent recruitment can be explained along six main dimensions: integrating interactive human–system decision-making; standardizing human-centered assessment processes; ensuring the transparency and explainability of intelligent systems; algorithmic monitoring and validation; identifying and quantitatively reducing data bias; and establishing ethical and operational protocols for intelligent systems. The findings indicate that human intervention reduces bias only when the human role extends beyond the formal approval of algorithmic recommendations to include critical evaluation of outputs, the possibility of reviewing decisions, documented justification for accepting or rejecting recommendations, and ultimate accountability. Standardizing competency indicators, training and cognitively retraining evaluators, ensuring the quality and representativeness of recruitment data, assessing differences in outcomes across applicant groups, continuously monitoring model performance, and conducting independent audits were likewise recognized as effective measures. Explainability of the system's decision logic and the provision of mechanisms for applicant appeals also reinforce the trustworthiness and accountability of the process.

**Conclusion**

Bias in AI-based recruitment and selection is a multidimensional phenomenon that arises from the entanglement of humans, data, algorithms, and organizational structures. Its management, therefore, cannot be achieved through technical model refinement alone or by simply removing sensitive variables; rather, it calls for a sociomaterial approach that analyzes humans and technology within an interactive, decision-making network. Establishing meaningful human oversight, standardizing judgments, improving data quality, ensuring system transparency and validation, conducting periodic audits, and developing ethical and operational protocols can enhance recruitment fairness, applicant trust, and organizational accountability, thereby laying the groundwork for responsible AI governance in Iranian human resource management.

**Keywords:** Cognitive bias, Digital human resource management, Human-AI interaction, Intelligent recruitment.

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

# راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی، به‌منظور جذب و استخدام منابع انسانی

**محمد باشکوه اجیرلو***

* نویسنده مسئول، استاد، گروه مدیریت بازرگانی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: bashokouh@uma.ac.ir

**مهری مقامی دولت‌آباد**

دانشجوی دکتری، گروه مدیریت بازرگانی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: mehri.maghamii@gmail.com

**بهمن خداپناه**

استادیار، گروه مدیریت بازرگانی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: khodapanah@ut.ac.ir

## چکیده

**هدف:** پژوهش حاضر با هدف شناسایی و تبیین راه‌کارهای کاهش سوگیری شناختی برای به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام منابع انسانی انجام شد. گسترش کاربرد سامانه‌های هوشمند در مدیریت منابع انسانی، استفاده از الگوریتم‌ها را در غربالگری رزومه‌ها، تحلیل شایستگی‌ها، رتبه‌بندی متقاضیان و پشتیبانی از تصمیم‌های استخدامی افزایش داده است. این سامانه‌ها در صورت طراحی و استفادهٔ مناسب می‌توانند سرعت، یکپارچگی و قابلیت پیگیری تصمیم‌ها را ارتقا دهند؛ با این‌حال، اتکای آن‌ها به داده‌های تاریخی، شاخص‌های جانشین، معیارهای مبهم و قضاوت‌های انسانی سوگیرانه، خطر بازتولید یا تشدید نابرابری‌ها را به همراه دارد. افزون‌براین، سوگیری‌های شناختی تصمیم‌گیران در انتخاب داده‌ها، تفسیر خروجی‌های مدل و تأیید توصیه‌های الگوریتمی، می‌تواند عدالت فرایند استخدام را مخدوش کند. از این رو، مسئلهٔ اصلی پژوهش، شناسایی سازوکارهای انسانی، داده‌ای، الگوریتمی و سازمانی برای کاهش سوگیری شناختی در استخدام مبتنی‌بر هوش مصنوعی بود.

**روش:** پژوهش با رویکرد کیفی و روش تحلیل مضمون انجام شد. داده‌ها از طریق ۱۲ مصاحبه عمیق نیمه‌ساختاریافته با خبرگان منابع انسانی، متخصصان فناوری و افراد آشنا با طراحی یا به‌کارگیری سامانه‌های هوشمند در سازمان‌های ایرانی گردآوری شد. مشارکت‌کنندگان با نمونه‌گیری هدفمند انتخاب شدند و مصاحبه‌ها تا دستیابی به اشباع نظری ادامه یافت. پس از پیاده‌سازی مصاحبه‌ها، داده‌ها بارها مطالعه و واحدهای معنایی مرتبط با منشأها، نمودها و راه‌کارهای کنترل سوگیری شناسایی شد. سپس، کدگذاری اولیه، استخراج مضامین پایه، مقایسه و ادغام کدهای مشابه، سازمان‌دهی مضامین و شکل‌دهی مضامین فراگیر انجام گرفت. تحلیل داده‌ها با نرم‌افزار مکس‌کیودا نسخهٔ ۲۰ انجام شد. به‌منظور تقویت اعتبار تحلیل، بازبینی مستمر کدها، رجوع مکرر به داده‌های خام، مقایسه دیدگاه مشارکت‌کنندگان و بررسی

**استناد:** باشکوه اجیرلو، محمد؛ مقامی دولت‌آباد، مهری و خداپناه، بهمن (۱۴۰۵). راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی، به‌منظور جذب و استخدام منابع انسانی. *مدیریت دولتی*، ۱۸(۳)، ۷۲۳ – ۷۵۱.

مدیریت دولتی، ۱۴۰۵، دوره ۱۸، شماره ۳، صص. ۷۲۳ – ۷۵۱
ناشر: دانشکدگان مدیریت، دانشگاه تهران
نوع مقاله: علمی پژوهشی
© نویسندگان
تاریخ دریافت: ۱۴۰۵/۰۲/۰۳
تاریخ ویرایش: ۱۴۰۵/۰۴/۲۴
تاریخ پذیرش:۱۴۰۵/۰۵/۲۹
تاریخ انتشار: ۱۴۰۵/۰۶/۲۳
doi: https://doi.org/10.22059/jipa.2026.417273.3944

انسجام میان مضامین صورت گرفت. برایند فرایند تحلیل به استخراج ۱۳۰ مضمون پایه، ۱۸ مضمون سازمان‌دهنده و ۶ مضمون فراگیر منتهی شد.

**یافته‌ها:** ساختار نهایی مضامین نشان داد که راه‌کارهای کاهش سوگیری شناختی در استخدام هوشمند، در شش محور اصلی تبیین می‌شود: یکپارچه‌سازی تصمیم‌گیری تعاملی انسان و سامانه‌های هوشمند، استانداردسازی فرایندهای ارزیابی انسان‌محور، شفافیت و توضیح‌پذیری سامانه‌های هوشمند، نظارت و اعتبارسنجی الگوریتمی، شناسایی و کاهش کمّی سوگیری‌های داده‌ای و استقرار پروتکل‌های اخلاقی و عملیاتی سامانه‌های هوشمند. یافته‌ها بیانگر آن است که مداخلهٔ انسانی تنها زمانی به کاهش سوگیری منجر می‌شود که نقش انسان به تأیید صوری پیشنهاد الگوریتم محدود نشود، بلکه شامل ارزیابی انتقادی خروجی، امکان بازبینی تصمیم، ثبت مستند دلایل رد یا پذیرش و پذیرش مسئولیت نهایی باشد. همچنین، استانداردسازی شاخص‌های شایستگی، آموزش و بازآموزی شناختی ارزیابان، کنترل کیفیت و نمایندگی داده‌های استخدامی، سنجش تفاوت نتایج میان گروه‌های متقاضی، پایش مستمر عملکرد مدل و ممیزی مستقل، از اقدامات مؤثر شناخته شد. توضیح‌پذیری منطق تصمیم سامانه و پیش‌بینی سازوکار اعتراض متقاضیان نیز اعتمادپذیری و پاسخ‌گویی فرایند را تقویت می‌کند.

**نتیجه‌گیری:** سوگیری در جذب و استخدام مبتنی‌بر هوش مصنوعی، پدیده‌ای چندبُعدی است که از درهم‌تنیدگی انسان، داده، الگوریتم و ساختارهای سازمانی پدید می‌آید. بنابراین، مدیریت آن فقط با اصلاح فنی مدل یا حذف متغیرهای حساس محقق نمی‌شود، بلکه به رویکردی جامعهٔ مادی نیاز دارد که انسان و فناوری را در شبکه‌ای تعاملی و تصمیم‌ساز تحلیل کند. استقرار کنترل انسانی معنادار، استانداردسازی قضاوت‌ها، ارتقای کیفیت داده‌ها، شفافیت و اعتبارسنجی سامانه‌ها، ممیزی دوره‌ای و تدوین پروتکل‌های اخلاقی و عملیاتی، می‌تواند عدالت استخدامی، اعتماد متقاضیان و پاسخ‌گویی سازمانی را ارتقا دهد و زمینهٔ حکمرانی مسئولانه هوش مصنوعی در مدیریت منابع انسانی ایران را فراهم سازد.

**کلیدواژه‌ها:** استخدام هوشمند، تعامل انسان و هوش مصنوعی، سوگیری شناختی، منابع انسانی دیجیتال.

## مقدمه

در سال‌های اخیر، ورود هوش مصنوعی به حوزه‌های مختلف مدیریتی، به‌ویژه مدیریت منابع انسانی، شیوه تصمیم‌گیری سازمان‌ها را تا حد زیادی تغییر داده است. این فناوری با امکان پردازش حجم زیادی از داده‌ها، شناسایی الگوهای پنهان و پشتیبانی از تصمیم‌های پیچیده، می‌تواند در فرایندهایی مانند غربالگری رزومه‌ها، ارزیابی شایستگی‌ها و پیش‌بینی تناسب فرد با شغل نقش مؤثری داشته باشد. با این حال، هوشمندسازی منابع انسانی، تنها به‌معنای افزودن ابزارهای فناورانه به فرایندهای موجود نیست، بلکه نیازمند هماهنگی میان عوامل مدیریتی، فرایندی و انسانی است؛ موضوعی که در پژوهش‌های مرتبط با تحول دیجیتال منابع انسانی نیز بر آن تأکید شده است (رستگار، ابراهیمی، شفیعی نیک‌آبادی و کلاهی، ۱۴۰۱). در بسیاری از سازمان‌ها، به‌ویژه در مراحل اولیه جذب، هنوز بخش مهمی از تصمیم‌ها بر پایه روش‌های سنتی، تجربه شخصی و قضاوت فردی انجام می‌شود. این شیوه‌ها در شرایطی که تعداد متقاضیان زیاد، مشاغل متنوع و معیارهای شایستگی پیچیده‌تر شده‌اند، بیش از گذشته در معرض خطاهای شناختی قرار می‌گیرند. به همین دلیل، در سطح جهانی استفاده از ابزارهای هوشمند برای غربال اولیه، تحلیل رزومه‌ها و کمک به تصمیم‌گیری استخدامی گسترش یافته است. در ایران نیز، اگرچه به‌کارگیری این ابزارها هنوز به سطحی فراگیر نرسیده است، برخی سازمان‌ها و پلتفرم‌های کاریابی در سال‌های اخیر به‌صورت محدود یا آزمایشی از سامانه‌های هوشمند در برخی مراحل جذب و استخدام بهره گرفته‌اند. البته استفاده از هوش مصنوعی لزوماً به‌معنای حذف خطاهای انسانی نیست. یکی از چالش‌های جدی در این زمینه آن است که سامانه‌های هوشمند می‌توانند سوگیری‌های موجود در داده‌ها و تصمیم‌های گذشته را بازتولید کنند. بسیاری از داده‌هایی که برای آموزش یا تنظیم این سامانه‌ها به‌کار می‌روند، پیش‌تر توسط انسان‌ها جمع‌آوری، انتخاب، تفسیر یا برچسب‌گذاری شده‌اند. بنابراین اگر در تعریف شایستگی، انتخاب معیارها، ثبت اطلاعات متقاضیان یا تصمیم‌های قبلی استخدامی سوگیری وجود داشته باشد، این سوگیری می‌تواند وارد سامانه شود و در خروجی آن ادامه پیدا کند.

پژوهش‌های پیشین نیز نشان داده‌اند که بخش مهمی از سوگیری‌های الگوریتمی ریشه در داده‌ها و قضاوت‌های انسانی پیشین دارد (مهرابی، مورستاتر، ساکسنا، لرمان و گالستیان[1]، ۲۰۲۱؛ لپری، الیور، لتوزه، پنتلند و وینک[2]، ۲۰۱۸؛ رَگاون، باروکاس، کلینبرگ و لوی[3]، ۲۰۲۰). از این منظر، مسئله فقط به عملکرد فنی الگوریتم محدود نمی‌شود. بخش مهمی از سوگیری پیش از آن شکل می‌گیرد که داده وارد سامانه شود؛ جایی‌که انسان‌ها درباره معیارهای شایستگی، نوع داده‌های قابل ثبت، نحوه برچسب‌گذاری و اهمیت هر ویژگی تصمیم می‌گیرند. اگر این مراحل با دقت و نظارت کافی انجام نشود، حتی سامانه‌های پیشرفته نیز ممکن است به نتایجی ناعادلانه برسند. به همین دلیل، در مطالعات جدیدتر بر توجه به اخلاق، شفافیت و نظارت بر سامانه‌های خودکار جذب تأکید شده و نقش فرایندهای انسانی پیش از ورود داده به

1. Mehrabi, Morstatter, Saxena, Lerman & Galstyan
2. Lepri, Oliver, Letouzé, Pentland & Vinck
3. Raghavan, Barocas, Kleinberg & Levy

الگوریتم برجسته‌تر شده است (سونی، آرمین، اشرف، اسلام و دبنات[1]، ۲۰۲۵؛ ریگوتی و فوش ویلارونگا[2]، ۲۰۲۴). همچنین، نقش کارشناسان منابع انسانی، تحلیلگران داده و توسعه‌دهندگان سامانه‌ها در شکل‌گیری یا کاهش این سوگیری‌ها اهمیت زیادی دارد (البارودی، منصوری و الامیر[3]، ۲۰۲۴). با وجود این، بخش چشمگیری از پژوهش‌های موجود، بیشتر بر جنبه‌های فنی سامانه‌ها، کیفیت مدل‌ها و شاخص‌های آماری عدالت تمرکز کرده‌اند. در مقابل، منشأهای شناختی و انسانی سوگیری، به‌ویژه در مراحل اولیه جذب و در شرایطی که داده‌های استخدامی تولید و آماده استفاده می‌شوند، کمتر در کانون است. این مسئله در زمینه سازمان‌های ایرانی اهمیت بیشتری پیدا می‌کند؛ زیرا استفاده از هوش مصنوعی در جذب و استخدام هنوز در مرحله توسعه و تجربه است و لازم است پیش از گسترش وسیع آن، راه‌کارهای مدیریتی، انسانی، داده‌ای و الگوریتمی کاهش سوگیری شناسایی شود. با توجه به ماهیت پیچیده و زمینه‌مند این مسئله، پژوهش حاضر با رویکرد کیفی و روش تحلیل مضمون انجام شده است تا تجربه‌ها و دیدگاه‌های خبرگان منابع انسانی، متخصصان فناوری و افراد آشنا با طراحی یا به‌کارگیری سامانه‌های هوشمند در سازمان‌های ایرانی بررسی شود. هدف پژوهش، شناسایی و تبیین راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام منابع انسانی است. بر این اساس، پژوهش حاضر به‌دنبال پاسخ به دو پرسش کلیدی است: نخست، چگونه می‌توان خطاهای شناختی را در مراحل اولیه فرایند جذب شناسایی و کاهش داد؟ دوم، چه راه‌کارهایی باید طراحی و اجرا شود تا احتمال انتقال این خطاها به داده‌ها، معیارهای ارزیابی و خروجی سامانه‌های هوشمند کاهش یابد و تصمیم‌های استخدامی از عدالت و دقت بیشتری برخوردار شوند؟

## پیشینهٔ نظری پژوهش

### رویکرد اول. سیر تحول روش‌ها در جذب و گزینش

روش‌های سنتی جذب و گزینش، از جمله انتشار آگهی در رسانه‌های چاپی، مصاحبه‌های حضوری، غربالگری دستی رزومه‌ها و اتکا به شبکه‌های ارتباطی، برای سال‌ها مسیر اصلی ورود افراد به سازمان‌ها بوده‌اند و هنوز نیز در بسیاری از سازمان‌ها کنار گذاشته نشده‌اند. مزیت این روش‌ها آن است که تا حدی امکان شناخت نزدیک‌تر و کیفی از داوطلبان را فراهم می‌کنند؛ اما همین وابستگی به قضاوت انسانی، شهود مدیران و روابط غیررسمی، آن‌ها را با محدودیت‌هایی جدی روبه‌رو می‌سازد. هزینه و زمان بالای اجرا، نبود معیارهای یکسان در ارزیابی، تأثیر برداشت‌های ذهنی مصاحبه‌گران و امکان بازتولید نابرابری، از مهم‌ترین چالش‌های این رویکردها به‌شمار می‌رود (هایگهاوس[4]، ۲۰۰۸؛ تیورسکی و کاهنمن[5]، ۱۹۷۴؛ رایان و پلیهارت[6]، ۲۰۱۴). افزون‌براین، غربالگری اولیه رزومه‌ها و اتکا به معرفی‌های درون‌سازمانی می‌تواند زمینهٔ

1. Sony, Armin, Ashraf, Islam & Debnath
2. Rigotti & Fosch-Villaronga
3. Albaroudi, Mansouri & Alameer
4. Highhouse
5. Tversky & Kahneman
6. Ryan & Ployhart

بروز سوگیری‌های قومی، جنسیتی، طبقاتی و شبکه‌ای را تقویت کند و به کاهش تنوع نیروی کار و توزیع نابرابر فرصت‌های شغلی بینجامد (دِروس و رایان[1]، ۲۰۱۹؛ برایو[2]، ۲۰۱۳؛ اعتمادی، چیت‌ساز، کوشکی و جعفری، ۱۴۰۳). در چنین زمینه‌ای، حرکت سازمان‌ها به‌سمت فناوری‌های دیجیتال و سامانه‌های مبتنی‌بر هوش مصنوعی را باید پاسخی به برخی از همین محدودیت‌ها دانست. این گذار، فقط جابه‌جایی ابزارهای کاغذی با ابزارهای دیجیتال نیست، بلکه تلاشی برای استانداردسازی ارزیابی، افزایش کارایی، کاهش مداخله‌های سلیقه‌ای و نزدیک‌شدن به ایده گزینش فراگیرتر است. ابزارهای هوشمند می‌توانند با پردازش حجم بالایی از داده‌ها، غربالگری سریع‌تر متقاضیان و ایجاد معیارهای نسبتاً یکنواخت برای ارزیابی، ظرفیت تصمیم‌گیری در مدیریت منابع انسانی را افزایش دهند (تامبه، کاپیلی و یاکوبوویچ[3]، ۲۰۱۹؛ کوشلینگ و وهنر[4]، ۲۰۲۰؛ هونکنشروئر و لوته[5]، ۲۰۲۲). با این حال، این امیدواری نباید به‌معنای بی‌طرف دانستن فناوری تلقی شود. بخش مهمی از ادبیات پژوهش نشان می‌دهد که سامانه‌های هوشمند، اگر بر داده‌های تاریخی مغرضانه آموزش ببینند یا بر پایهٔ معیارهای محدود و ناعادلانه طراحی شوند، می‌توانند همان تبعیض‌های پیشین را در سطحی تازه بازتولید کنند (باروکاس و سلبست[6]، ۲۰۱۶؛ رَگاون و همکاران، ۲۰۲۰؛ مهرابی و همکاران، ۲۰۲۱). این مسئله در سازمان‌های دولتی و عمومی اهمیت بیشتری پیدا می‌کند؛ زیرا جذب و گزینش در چنین بسترهایی تنها یک تصمیم عملیاتی در واحد منابع انسانی نیست، بلکه با پاسخ‌گویی، شفافیت، ارزیابی نظام‌مند و حکمرانی سازمانی پیوند دارد. از این منظر، تعالی مدیریت منابع انسانی در دستگاه‌های اجرایی مستلزم آن است که تصمیم‌های استخدامی و سازوکارهای ارزیابی افراد در چارچوبی شفاف، قابل‌پایش و سنجش‌پذیر طراحی شوند. بنابراین، به‌کارگیری فناوری‌های هوشمند در جذب زمانی می‌تواند به عدالت و کارآمدی کمک کند که با الزامات پاسخ‌گویی و ارزیابی نظام‌مند همراه باشد (قلی‌پور، ۱۴۰۳). با وجود گسترش این فناوری‌ها، ادبیات مربوط به پیامدهای هوش مصنوعی در جذب و استخدام یکدست نیست. بخشی از پژوهش‌ها با نگاهی خوش‌بینانه استدلال می‌کنند که ابزارهای هوش مصنوعی، از طریق کاهش نویز تصمیم‌گیری، استانداردسازی ارزیابی و محدود کردن قضاوت‌های سلیقه‌ای مدیران، می‌توانند به بی‌طرفی بیشتر و تقویت تنوع در استخدام کمک کنند (میلر[7]، ۲۰۲۲). در مقابل، جریان انتقادی هشدار می‌دهد که الگوریتم‌ها ممکن است تبعیض‌های موجود در داده‌های آموزشی را آموخته و حتی تشدید کنند. از این زاویه، سازمان‌ها با چالشی دوگانه روبه‌رو می‌شوند؛ زیرا تلاش برای افزایش تنوع از طریق مداخله الگوریتمی می‌تواند با دغدغه‌هایی درباره روایی پیش‌بینی، دقت ارزیابی و عدالت تصمیم‌گیری همراه شود (چن، ما، هناک و ویلسون[8]، ۲۰۱۸؛ رائو و ژائو[9]، ۲۰۲۵؛

1. Derous & Ryan
2. Breaugh
3. Tambe, Cappelli & Yakubovich
4. Köchling & Wehner
5. Hunkenschroer & Lütge
6. Barocas & Selbst
7. Miller
8. Chen, Ma, Hannák & Wilson
9. Rao & Zhao

هربوش[1]، ۲۰۲۵). در کنار این دو نگاه، جریان سومی نیز شکل گرفته است که مبنای نظری پژوهش حاضر را فراهم می‌کند. این جریان، تمرکز را از فناوری به‌عنوان یک ابزار مستقل، به تعامل میان انسان، داده، الگوریتم و زمینه سازمانی منتقل می‌کند. در این نگاه، پرسش اصلی این نیست که هوش مصنوعی به‌تنهایی عادلانه‌تر از انسان تصمیم می‌گیرد یا خیر؛ بلکه مسئله آن است که پیش‌فرض‌های انسانی، کیفیت داده‌ها، منطق طراحی سامانه و نحوه استفاده مدیران از خروجی‌های الگوریتمی چگونه در کنار یکدیگر تصمیم استخدامی را شکل می‌دهند (میلر، ۲۰۲۲؛ ویلکنز، لوتزایر، ژنگ، بسر و پریلا[2]، ۲۰۲۵). به همین دلیل، رویکردهای جدید به‌جای کنارگذاشتن انسان از فرایند استخدام، بر «تقویت» قضاوت انسانی از طریق فناوری تأکید می‌کنند؛ یعنی وضعیتی که در آن، هوش مصنوعی و داوری انسانی در تعامل با یکدیگر می‌توانند به تصمیم‌هایی دقیق‌تر، شفاف‌تر و قابل‌بازبینی‌تر منجر شوند (فیشر[3]، ۲۰۲۲؛ ویلکنز و همکاران، ۲۰۲۵). برای فهم این تعامل، پژوهش حاضر از نظریه «جامعهٔ مادی» بهره می‌گیرد. بر پایهٔ این دیدگاه، فناوری عنصری جدا از سازمان و کنشگران انسانی نیست، بلکه در شبکه‌ای از رویه‌های اجتماعی، ادراکات انسانی، داده‌ها و ساختارهای فنی معنا پیدا می‌کند. از این رو، سوگیری را نمی‌توان فقط به ذهن انسان یا فقط به کد الگوریتم نسبت داد؛ بلکه باید آن را حاصل پیوند انتخاب‌های انسانی، داده‌های تاریخی، قواعد فنی و رویه‌های سازمانی دانست. بنابراین، داشتن یک ابزار هوشمند دقیق، برای ارتقای عدالت در جذب لازم است، اما کافی نیست. کاهش سوگیری زمانی امکان‌پذیر می‌شود که رفتار کاربران انسانی، کیفیت داده‌ها، شفافیت الگوریتمی، حکمرانی داده و نظارت بر تعامل انسان و سامانه به‌صورت هم‌زمان موردتوجه قرار گیرد (اسکات و اورلیکوسکی[4]، ۲۰۲۵). در امتداد همین بحث، پژوهش‌های اخیر درباره پذیرش هوش مصنوعی در حکمرانی هوشمند نیز نشان می‌دهند که موفقیت این فناوری فقط به دقت فنی آن وابسته نیست. زیرساخت‌های دیجیتال، حکمرانی و امنیت داده، شفافیت الگوریتمی، فرهنگ سازمانی نوآورانه، آموزش و توانمندسازی کارکنان و نظارت بر فرایندهای هوش مصنوعی، همگی در پذیرش و کارکرد مؤثر این فناوری نقش دارند (مؤمنی، یعقوبی، روشن و پورعزت، ۱۴۰۵).

## رویکرد دوم. واکاوی انواع سوگیری‌ها؛ از خطای انسانی تا بازتولید الگوریتمی

علاوه‌بر ظهور سامانه‌های مبتنی‌بر هوش مصنوعی در استخدام، در ابتدا نویدبخش افزایش کارایی و عینیت در تصمیم‌گیری بود؛ اما شواهد نشان داد که این سامانه‌ها نه‌تنها از سوگیری‌های انسانی مصون نیستند، بلکه می‌توانند آن‌ها را در مقیاسی گسترده بازتولید یا تقویت کنند. این سوگیری‌ها که به نتایج ناعادلانه و تبعیض‌آمیز منجر می‌شوند، عموماً در سه دسته اصلی طبقه‌بندی می‌شوند:

- سوگیری داده: این نوع سوگیری، رایج‌ترین شکل سوگیری در سامانه‌های یادگیری ماشین است و زمانی رخ می‌دهد که الگوریتم بر داده‌های تاریخیِ حاوی الگوهای تبعیض‌آمیز آموزش ببیند. برای مثال، اگر سازمانی در

1. Herbosch
2. Wilkens, Lutzeyer, Zheng, Beser & Prilla
3. Fischer
4. Scott & Orlikowski

گذشته برای نقش‌های فنی بیشتر مردان را استخدام کرده باشد، داده‌های استخدامی این الگو را بازتاب می‌دهد و سامانه نیز ممکن است رزومه‌های مشابه مردان را در اولویت قرار دهد؛ حتی اگر جنسیت به‌عنوان متغیر صریح وارد مدل نشده باشد. در نتیجه، سامانه به‌جای شناسایی شایستگی واقعی، الگوهای نابرابر گذشته را تداوم می‌بخشد (رَگاون و همکاران، ۲۰۱۹؛ البارودی و همکاران، ۲۰۲۴؛ کوشلینگ و وهنر، ۲۰۲۰).

- سوگیری طراحی و الگوریتم: این سوگیری از انتخاب‌ها و مفروضات انسانی در مرحله طراحی و پیاده‌سازی الگوریتم ناشی می‌شود. توسعه‌دهندگان و متخصصان منابع انسانی هنگام تعریف «موفقیت»، انتخاب ویژگی‌های ارزیابی و تعیین وزن متغیرها، ناگزیر پیش‌فرض‌های ذهنی خود را وارد فرایند می‌کنند. بنابراین، تصمیم‌های فنیِ به‌ظاهر عینی می‌توانند حامل سوگیری‌های پنهان باشند و به تعریف محدودی از «کاندیدای ایده‌آل» منجر شوند که به نفع گروه‌های خاص عمل می‌کند (کردزاده و قاسم‌آقایی[1]، ۲۰۲۱).
- سوگیری تعامل و حلقه بازخورد: این سوگیری زمانی شکل می‌گیرد که میان خروجی سامانه و تصمیم‌گیرنده انسانی، چرخه‌ای معیوب ایجاد شود. حتی اگر سامانه خروجی نسبتاً بی‌طرفی ارائه کند، مدیران استخدام ممکن است براساس پیش‌داوری‌های خود برخی پیشنهادها را بپذیرند یا نادیده بگیرند. سامانه نیز با دریافت این بازخورد، ترجیحات سوگیرانه مدیر را به‌عنوان نتیجه صحیح یاد می‌گیرد و در چرخه‌های بعدی خروجی‌های خود را با همان ترجیحات تنظیم می‌کند؛ فرایندی که به تقویت تکرارشونده سوگیری می‌انجامد (دیفنهارت[2]، ۲۰۲۵).

چارچوب تحلیلی فوق، این سه نوع سوگیری را به‌صورت مجزا اما مرتبط بررسی می‌کند و بر مدل مفهومی توسعه‌یافته در پژوهش سلیمانی، انتظاری، اروسمیت، پائولین و تاسکین[3] (۲۰۲۵) استوار است؛ مدلی که در ادامه به‌عنوان یکی از مبانی اصلی پژوهش حاضر مورد استفاده قرار می‌گیرد.

## ریشهٔ انسانی سوگیری الگوریتمی: سوگیری شناختی

هرچند سوگیری در سامانه‌های هوشمند استخدامی معمولاً در سطح داده، طراحی الگوریتم یا نحوه تعامل انسان با سامانه دیده می‌شود، ریشه بسیاری از این خطاها را باید در قضاوت انسانی جست‌وجو کرد. به بیان دیگر، آنچه در خروجی الگوریتم به‌صورت سوگیری ظاهر می‌شود، در بسیاری از موارد ادامه همان خطاهای شناختی است که پیش‌تر در تصمیم‌های انسانی، معیارهای ارزیابی یا داده‌های استخدامی شکل گرفته است. سوگیری‌های شناختی به الگوهای نسبتاً پایدار خطا در قضاوت گفته می‌شود که از اتکای انسان به میان‌برهای ذهنی یا هیوریستیک‌ها ناشی می‌شوند. این میان‌برها به افراد کمک می‌کنند در شرایط پیچیده و با اطلاعات محدود سریع‌تر تصمیم بگیرند، اما همین ویژگی در موقعیت‌هایی مانند استخدام، که با ابهام، فشار زمانی و حجم بالای اطلاعات همراه است، می‌تواند به قضاوت‌های

1. Kordzadeh & Ghasemaghaei
2. Diefenhardt
3. Soleimani, Intezari, Arrowsmith, Pauleen & Taskin

نادرست و تکرارشونده منجر شود (کورتلینگ و توئت[1]، ۲۰۲۱). در چنین شرایطی، مدیران و ارزیابان گاه بیش از آنکه بر شواهد عینی تکیه کنند، به برداشت اولیه، تجربه شخصی یا قضاوت شهودی خود متکی می‌شوند؛ امری که زمینه بروز خطاهای شناختی را در فرایند جذب افزایش می‌دهد (برثت[2]، ۲۰۲۲). در فرایند جذب و استخدام، برخی سوگیری‌ها بیشتر از سایر موارد اثرگذارند. برای نمونه، در سوگیری شباهت، ارزیاب ناخودآگاه افرادی را مطلوب‌تر می‌بیند که از نظر پیشینه، علایق یا ویژگی‌های شخصیتی به او نزدیک‌ترند (ویسال[3]، ۲۰۱۸). در سوگیری کلیشه‌ای، ویژگی‌های فرد نه براساس شایستگی واقعی، بلکه با تکیه بر عضویت او در یک گروه اجتماعی، جنسیتی یا آموزشی ارزیابی می‌شود (مارینوسی، مازوکا و گانگمی[4]، ۲۰۲۳). سوگیری تأیید نیز باعث می‌شود ارزیاب پس از شکل‌گیری برداشت اولیه، بیشتر به‌دنبال شواهدی باشد که همان برداشت را تأیید می‌کند و نشانه‌های مخالف را کمتر ببیند. در اثر هاله‌ای و شاخی نیز یک ویژگی مثبت یا منفی، به سایر ابعاد شخصیت و عملکرد فرد تعمیم داده می‌شود (شائو[5]، ۲۰۲۳؛ رادیکه و ستاهلسکی[6]، ۲۰۲۰). پیوند میان خطای انسانی و سوگیری الگوریتمی را می‌توان در سه مرحله اصلی مدل سلیمانی و همکاران (۲۰۲۵) دنبال کرد. در مرحلهٔ نخست، تصمیم‌های استخدامی گذشته که تحت‌تأثیر سوگیری‌هایی مانند شباهت، کلیشه یا برداشت اولیه اتخاذ شده‌اند، به داده‌های تاریخی تبدیل می‌شوند و همین داده‌ها مبنای آموزش الگوریتم قرار می‌گیرند. در مرحلهٔ دوم، سوگیری در سطح طراحی ظاهر می‌شود؛ زیرا طراحان و متخصصان منابع انسانی هنگام تعریف شاخص‌های موفقیت، انتخاب متغیرها و تعیین اهمیت آن‌ها، ممکن است ناخواسته پیش‌فرض‌های ذهنی خود را وارد مدل کنند. به همین دلیل، انتخاب‌های فنیِ ظاهراً بی‌طرف می‌توانند ساختار الگوریتم را به سمت تعریف خاصی از «کاندیدای مطلوب» هدایت کنند (کردزاده و قاسم‌آقایی، ۲۰۲۱). در مرحلهٔ سوم، تعامل مدیران با خروجی سامانه اهمیت پیدا می‌کند. اگر مدیران فقط آن دسته از پیشنهادهای هوش مصنوعی را بپذیرند که با شهود و برداشت قبلی آنان سازگار است، سامانه نیز از همین بازخوردها می‌آموزد و در چرخه‌های بعدی همان الگوهای ترجیحی را تقویت می‌کند (ویلکنز و همکاران، ۲۰۲۵). بر این اساس، کاهش سوگیری در سامانه‌های هوشمند استخدامی تنها با اصلاح فنی الگوریتم‌ها ممکن نیست و باید ریشه‌های شناختی و انسانی آن نیز مورد توجه قرار گیرد. از این منظر، تعامل انسان و هوش مصنوعی باید به‌گونه‌ای طراحی شود که امکان بازبینی، مقایسه چندمنبعی و اعتبارسنجی تصمیم‌ها وجود داشته باشد. پژوهش‌های روش‌شناختی اخیر نیز نشان می‌دهند که تلفیق قضاوت خبرگان انسانی با قابلیت‌های مدل‌های زبانی می‌تواند به تحلیل چندلایه‌تر مسائل مدیریتی، افزایش دقت، تنوع دیدگاه‌ها و پایداری داده‌ها کمک کند؛ موضوعی که ضرورت سازوکارهای انسان‌درحلقه و زاویه‌بندی تصمیم‌ها را در مواجهه با خروجی‌های الگوریتمی برجسته می‌سازد (دهقانان، پورامینی، یزدان‌شناس و رئیسی وانانی، ۱۴۰۵).

1. Korteling & Toet
2. Berthet
3. Whysall
4. Marinucci, Mazzuca & Gangemi
5. Shao
6. Radeke & Stahelski

## پیشینهٔ تجربی

روش‌های تصمیم‌گیری در فرایند جذب و استخدام نیروی انسانی همواره یکی از حساس‌ترین حوزه‌های مدیریت منابع انسانی بوده است؛ چراکه این تصمیم‌ها افزون بر اثرگذاری بر عملکرد کوتاه‌مدت سازمان، پیامدهای مهمی برای پایداری منابع انسانی، عدالت سازمانی و سرمایهٔ اجتماعی به‌همراه دارند. مطالعات انجام‌شده در حوزهٔ روان‌شناسی تصمیم‌گیری و مدیریت منابع انسانی نشان می‌دهد که قضاوت‌ها و انتخاب‌های مدیران، حتی در محیط‌های حرفه‌ای و نسبتاً ساخت‌یافته، لزوماً عقلایی و بی‌خطا نیستند و به‌طور معناداری تحت‌تأثیر سوگیری‌های شناختی قرار می‌گیرند. این سوگیری‌ها به‌ویژه در شرایط عدم‌قطعیت اطلاعات، فشار زمانی و پیچیدگی گزینه‌های تصمیم‌گیری تشدید می‌شوند و می‌توانند پیامدهای بلندمدت و گاه جبران‌ناپذیری برای سازمان به‌دنبال داشته باشند. بر این اساس، پژوهش‌های داخلی و خارجی متعددی به بررسی ابعاد مختلف تصمیم‌گیری در جذب و استخدام، نقش ابزارهای فناورانه و هوشمند و چالش‌های مرتبط با عدالت و سوگیری پرداخته‌اند. خلاصه‌ای از مهم‌ترین این پژوهش‌ها، همراه با محورهای تمرکز، روش‌شناسی و نتایج کلیدی آن‌ها، در جدول ۱ ارائه شده است.

**جدول ۱. پیشینهٔ پژوهش**

| نویسنده/ سال پژوهش | موضوع/ هدف پژوهش | روش، یافته کلیدی و وجه افتراق با پژوهش حاضر |
|---|---|---|
| سلیمانی و همکاران، ۲۰۲۵ | کاهش سوگیری هوش مصنوعی در جذب و انتخاب؛ تبیین چگونگی شکل‌گیری و کاهش سوگیری | با رویکرد کیفی و نظریه داده‌بنیاد نشان دادند که سوگیری حاصل برهم‌کنش تصمیم‌گیرندگان انسانی، داده‌های ورودی و منطق الگوریتمی است. تمایز مطالعه حاضر در تمرکز ریزبینانه‌تر بر منشأهای شناختیِ پیشاالگوریتمی، به‌ویژه در تعریف شایستگی و برچسب‌گذاری داده‌های استخدامی است. |
| سونی و همکاران، ۲۰۲۵ | سوگیری در سیستم‌های مدیریت منابع انسانی مبتنی‌بر هوش مصنوعی؛ بررسی ریسک‌های تبعیض | در مطالعه‌ای تحلیلی ـ کاربردی، خطر بازتولید تبعیض‌های جنسیتی، سنی و اجتماعی در سامانه‌های منابع انسانی مبتنی بر هوش مصنوعی برجسته شده است. اگرچه آن پژوهش بیشتر بر پیامدهای تبعیض‌آمیز متمرکز است، اما مطالعه حاضر علاوه بر آن، مسیر انتقال سوگیری را از قضاوت انسانی به داده و سپس به الگوریتم واکاوی می‌کند. |
| یوآن، ژینگ و ژائو[1]، ۲۰۲۵ | نقش حضور منابع انسانی در چرخهٔ تصمیم‌گیری برای ارتقای ادراک عدالت در سیستم‌های انتخاب مبتنی‌بر هوش مصنوعی | با اتکا به نظریه عدالت سازمانی، حضور منابع انسانی را در مراحل پیشاارزیابی، ارزیابی و پس‌ارزیابی عامل افزایش اعتماد و پذیرش سامانه دانستند. تفاوت اصلی در آن است که مطالعه حاضر به‌جای ارائه چارچوب مفهومی، بر داده‌های مصاحبه و تحلیل مضمون متکی است. |

1. Yuan, Xing & Zhao

| نویسنده/ سال پژوهش | موضوع/ هدف پژوهش | روش، یافته کلیدی و وجه افتراق با پژوهش حاضر |
|---|---|---|
| ریگوتی و فوش ویلارونگا، ۲۰۲۴ | انصاف، هوش مصنوعی و استخدام؛ بررسی الزامات عدالت، شفافیت و پاسخ‌گویی | در قالب مرور دامنه‌دار میان‌رشته‌ای، چالش‌هایی مانند نبود شفافیت، دشواری انتساب مسئولیت و محدودیت پاسخ‌گویی در استخدام خودکار را شناسایی کردند. نسبت این مطالعه با پژوهش حاضر بیشتر در سطح عدالت الگوریتمی است؛ درحالی‌که این پژوهش الگوهای ذهنی و رفتاری متخصصان منابع انسانی و فنی را در تولید سوگیری بررسی می‌کند. |
| البارودی و همکاران، ۲۰۲۴ | هوش مصنوعی برای پرداختن به سوگیری الگوریتمی در استخدام؛ شناسایی تکنیک‌های کاهش سوگیری | مرور نظام‌مند آنان نشان داد که ادبیات موجود عمدتاً بر پیش‌پردازش داده، اصلاح مدل و تکنیک‌های کاهش سوگیری الگوریتمی متمرکز است. خلأ مورد توجه مطالعه حاضر، نقش فرایندهای انسانی و قضاوت‌های اولیه در شکل‌گیری داده‌های استخدامی است. |
| برثت، ۲۰۲۲ | تأثیر سوگیری‌های شناختی بر تصمیم‌گیری متخصصان در حوزه‌های مدیریت، مالی، پزشکی و حقوق | در یک مرور تلفیقی، اثر سوگیری‌های شناختی، به‌ویژه اعتمادبه‌نفس بیش‌ازحد، بر تصمیم‌گیری حرفه‌ای را گزارش کرد. در قیاس با آن، مطالعه حاضر سوگیری شناختی را نه در تصمیم‌گیری عمومی متخصصان، بلکه در بستر خاص جذب و تولید داده برای سامانه‌های هوشمند بررسی می‌کند. |
| فضل‌طلب و جمشیدی، ۱۴۰۴ | منابع انسانی الگوریتمی؛ بررسی فرصت‌ها و چالش‌های تصمیم‌گیری الگوریتمی در منابع انسانی | در پژوهشی مروری ـ تحلیلی، افزایش کارایی در کنار چالش‌هایی مانند تبعیض الگوریتمی، کاهش شفافیت و تضعیف عدالت ادراک‌شده را مطرح کردند. نوآوری مطالعه حاضر عبور از سطح پیامدهای کلی و تمرکز بر مرحله پیشاالگوریتمی، یعنی جایی است که معیارها، داده‌ها و برچسب‌ها تحت‌تأثیر قضاوت انسانی ساخته می‌شوند. |
| محمدی، ۱۴۰۴ | کاربرد الگوریتم‌های هوش مصنوعی در بهینه‌سازی استخدام شهرداری‌ها؛ مطالعه موردی شهرداری صدرا | در مطالعه‌ای کاربردی ـ موردی، کارایی الگوریتم‌هایی مانند SVM، درخت تصمیم، شبکه عصبی، جنگل تصادفی و رگرسیون را در کاهش زمان و هزینه جذب بررسی کرد. برخلاف رویکرد فنی این مطالعه، پژوهش حاضر به لایه شناختی و انسانیِ پشت داده‌های ورودی سامانه می‌پردازد. |
| سرایی، صراف و حمیدیان، ۱۴۰۳ | مدل‌سازی خطاهای سوگیری بر تصمیم‌گیری مالی مدیران با نگرش چندسطحی | با روش دلفی فازی نشان دادند که خطاهای سوگیری و متغیرهای زمینه‌ای در سطوح مختلف تصمیم‌گیری مالی اثر متفاوت دارند. دامنه این پژوهش حوزه مالی و مدیران بورسی است؛ اما مطالعه حاضر مسئله سوگیری را در تصمیم‌های جذب، انتخاب و ارزیابی داوطلبان دنبال می‌کند. |
| بذرکار، مرادزاد و شایگان، ۱۴۰۲ | اثر عوامل تکنولوژیکی، سازمانی و محیطی بر به‌کارگیری هوش مصنوعی در جذب کارکنان | با رویکرد توصیفی ـ پیمایشی و مدل‌سازی کمّی نشان دادند عوامل تکنولوژیکی، سازمانی و محیطی بر پذیرش هوش مصنوعی در جذب اثر معنادار دارند. وجه متمایز پژوهش حاضر، رویکرد کیفی و استخراج مضامین مرتبط با کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی است. |

## ارزیابی انتقادی از پیشینه پژوهش

در بررسی پیشینهٔ پژوهش روشن می‌شود که بخش عمده‌ای از مطالعات انجام‌شده در حوزهٔ به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام، چه در ادبیات داخلی و چه در ادبیات خارجی، عمدتاً بر شناسایی مزایا، چالش‌ها، الزامات فنی و ملاحظات اخلاقی این فناوری تمرکز داشته‌اند. محور اصلی این پژوهش‌ها بر عملکرد سامانه‌های هوشمند، مدیریت تبعیض الگوریتمی و پیامدهای سازمانی و نهادی تصمیم‌گیری مبتنی‌بر هوش مصنوعی استوار بوده و به چگونگی شکل‌گیری داده‌ها در مراحل نخستین فرایند جذب توجه کمتری شده است. اگرچه این مطالعات در تبیین ظرفیت‌ها و مخاطرات سامانه‌های هوشمند نقش قابل‌توجهی ایفا کرده‌اند، اما در اغلب موارد، سوگیری به‌مثابهٔ پدیده‌ای فنی و وابسته به سازوکار الگوریتم‌ها یا داده‌های موجود تحلیل شده است. در نتیجه، منشأهای انسانی شکل‌گیری سوگیری ـ به‌ویژه قضاوت‌ها، انتخاب‌ها و نحوهٔ تعریف معیارهای ارزیابی توسط تصمیم‌گیرندگان انسانی ـ به‌صورت محدود مورد واکاوی قرار گرفته‌اند. این در حالی است که بخش شایان توجهی از داده‌های ورودی سامانه‌های هوش مصنوعی، مستقیماً بازتاب برداشت‌های ذهنی، پیش‌فرض‌ها و الگوهای تصمیم‌گیری منابع انسانی در مراحل اولیهٔ فرایند جذب است. از سوی دیگر، حتی در پژوهش‌هایی که به مفاهیمی همچون عدالت، انصاف ادراک‌شده یا تعامل انسان و سیستم پرداخته‌اند، تمرکز عمدتاً معطوف به نتایج نهایی و پیامدهای کاربرد سامانه‌های هوشمند بوده است، نه به سازوکارهای شناختی و فرایندهای ذهنی‌ای که پیش از ورود اطلاعات به این سامانه‌ها شکل می‌گیرند. رویکرد غالب در این مطالعات را می‌توان رویکردی واکنشی و پیامدمحور دانست؛ بدین‌معنا که سوگیری پس از ظهور در داده یا خروجی الگوریتم شناسایی شده و راه‌کارها نیز بیشتر ناظر بر اصلاحات فنی یا مداخلات نهادی بوده‌اند. در مجموع، با وجود غنای نظری و فنی ادبیات موجود، می‌توان گفت که تبیین فرایند انتقال سوگیری از ذهن تصمیم‌گیرندگان انسانی به داده‌های ورودی سامانه‌های هوش مصنوعی به‌صورت نظام‌مند موردتوجه قرار نگرفته است. به‌ویژه نقش خطاها و سوگیری‌های شناختی در مراحل پیشاالگوریتمی جذب و استخدام همچنان با یک خلأ پژوهشی جدی مواجه است. بر این اساس، پژوهش حاضر با تمرکز بر سازوکارهای شناختی تصمیم‌گیرندگان انسانی در مراحل اولیهٔ جذب و انتخاب و با اتخاذ رویکردی فرایندمحور، می‌کوشد از سطح توصیف پیامدها فراتر رفته و به شناسایی منشأ شکل‌گیری سوگیری‌ها پرداخته و زمینهٔ ارائهٔ راه‌کارهای عملی برای کاهش سوگیری‌های شناختی در تعامل میان انسان و هوش مصنوعی را فراهم سازد.

## روش‌شناسی پژوهش

در این پژوهش، به‌منظور استخراج الگوهای معنایی و تبیین نظام‌مند راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام، از روش تحلیل مضمون با رویکرد داده‌محور و استقرایی استفاده شد. انتخاب این روش به‌دلیل ماهیت اکتشافی پژوهش و ضرورت فهم عمیق تجربه‌ها، ادراکات و قضاوت‌های خبرگان در زمینهٔ تعامل انسان و سیستم‌های هوشمند در تصمیم‌گیری‌های استخدامی صورت گرفت. داده‌های کیفی از طریق مصاحبه‌های عمیق نیمه‌ساختاریافته گردآوری شد. سؤالات مصاحبه براساس اهداف پژوهش و مرور اولیهٔ ادبیات موضوع طراحی شد.

در این مرحله، نظریه جامعهٔ مادی به‌عنوان یک لنز نظری و حساس‌کننده نیز موردتوجه قرار گرفت تـا طراحـی سـؤالات، علاوه‌بر تمرکز بر قضاوت‌ها و سوگیری‌های شناختی افـراد، بـه نقـش و تعامـل داده‌هـای اسـتخدامی، ابزارهـای هـوش مصنوعی، فرایندهای منابع انسانی و زمینه‌های سازمانی در تصمیم‌گیری‌های استخدامی نیز جهت یابد و در جریان انجـام مصاحبه‌ها، متناسب با پاسخ‌های مشارکت‌کنندگان، پرسش‌های تکمیلی و کاوشی نیز برای تعمیـق مفـاهیم مطـرح شـد. مصاحبه‌ها به‌صورت آنلاین انجام گرفت که این امر امکان دسترسی به خبرگان شاغل در سازمان‌ها و حوزه‌های مرتبط با منابع انسانی، تصمیم‌گیری مدیریتی و طراحی یا به‌کارگیری سیستم‌های هوش مصنوعی در جـذب و اسـتخدام را فـراهم ساخت. نمونه‌گیری در بخش کیفی به‌صورت هدفمند (قضاوتی) انجام شد. در انتخاب مشارکت‌کنندگان، تمرکز اصـلی بـر تلفیق دانش نظری و تجربه حرفه‌ای در حوزه‌های منابع انسانی، تحول دیجیتال و سامانه‌های هوشمند بود. بـا توجـه بـه نوظهور بودن کاربرد هوش مصنوعی در جذب و استخدام، معیار انتخاب افراد صرفاً جایگاه سازمانی آنان نبود، بلکه میزان آشنایی و تجربه آن‌ها در مواجهه با چالش‌های پیاده‌سازی فناوری در فرایندهای تصـمیم‌گیری انسـانی نیـز مـدنظر قـرار گرفت. بر این اساس، مشارکت‌کنندگان شامل اساتید متخصص در حوزهٔ هوش مصنوعی و منابع انسانی، پژوهشـگران، و کارشناسان متخصص در حوزهٔ جذب، استخدام و سامانه‌های هوشمند سازمانی بودند. این پژوهش با هدف تعمـیم آمـاری یافته‌ها به صنایع خاص یا مقایسه نظام‌مند میان صنایع انجام نشده است. به همین دلیل، نوع صـنعت مشـارکت‌کنندگان به‌صورت طبقه‌بندی‌شده ثبت نشد. تمرکز پـژوهش بـر دسـتیابی بـه درکـی عمیـق از چالش‌هـای سـوگیری شـناختی و راه‌کارهای کاهش آن در فرایندهای جذب و استخدام مبتنی‌بر هوش مصنوعی بـود. بنـابراین، یافتـه‌ها بیـانگر الگوهـای مشترک در تصمیم‌گیری‌های استخدامی هوشمند هستند؛ با این حال، استفاده از آن‌ها در هر بستر سازمانی مستلزم توجـه به ویژگی‌ها، ساختارها و شرایط خاص همان بستر است. گردآوری داده‌ها تا رسیدن به اشباع نظری ادامـه یافـت. اشـباع نظری پس از مصاحبه یازدهم حاصل شد و مصاحبه دوازدهم برای اطمینان از تثبیت مضامین و بررسی احتمال ظهور کد یا مضمون تازه انجام شد. در مصاحبهٔ پایانی، کد یا مضمون تازه‌ای شناسایی نشد. مدت زمان هر مصاحبه بین ۳۰ تـا ۶۰ دقیقه متغیر بود. تحلیل داده‌های کیفی براساس مراحل تحلیل مضمون انجام شد. در این فراینـد، ابتـدا مـتن مصـاحبه‌ها به‌طور کامل پیاده‌سازی و چندین بار مطالعه شد تا آشنایی عمیق با داده‌ها حاصل شود. سپس کدهای اولیه از بخش‌هـای معنادار متن به‌صورت داده‌محور استخراج شد و کدهای مشابه از نظر مفهومی در قالب مضامین پایـه گروه‌بنـدی شـد. در ادامه با بازبینی و پالایش مستمر، مضامین سازمان‌دهنده و در نهایت مضامین فراگیـر شناسـایی شـدند کـه بـا اتکـا بـه داده‌های مصاحبه و با توجه به پیوندهای انسانی، داده‌ای، فناورانه و سازمانی مورد تأکید در نظریه جامعهٔ مادی، چـارچوب مفهومی راه‌کارهای کاهش سوگیری شناختی در جذب مبتنی‌بر هـوش مصـنوعی را تبیـین می‌کننـد. تحلیـل داده‌هـا بـا استفاده از نسخهٔ ۲۰ نرم‌افـزار مکس‌کیـودا انجـام شـد. بـه‌منظور افـزایش کیفیـت و اعتمادپـذیری یافتـه‌های کیفـی، از راهبردهایی نظیر تنوع مشارکت‌کنندگان از نظر پیشینهٔ حرفه‌ای، بازبینی مستمر کـدها و مضـامین و مستندسـازی دقیـق مراحل تحلیل استفاده شد. مشخصات کلی مشارکت‌کنندگان پژوهش در جدول ۲ ارائه شده است.

**جدول ۲. جمعیت‌شناختی خبرگان**

| سن | فراوانی | جنسیت | فراوانی | سابقه کار (سال) | فراوانی |
|---|---|---|---|---|---|
| ۲۵–۳۵ | ۱ | زن | ۳ | ۵–۱۰ | ۲ |
| ۳۵–۴۵ | ۵ | مرد | ۹ | ۱۰–۲۰ | ۶ |
| ۴۵ و بیشتر | ۶ | – | – | ۲۰ و بیشتر | ۴ |
| کل | ۱۲ | کل | ۱۲ | کل | ۱۲ |

منبع: (یافته‌های پژوهش)

به‌منظور استخراج مضامین کلیدی مرتبط با راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام، داده‌های حاصل از مصاحبه‌های عمیق نیمه‌ساختاریافته موردتحلیل قرار گرفت. در این راستا، پروتکل مصاحبه‌ای طراحی شد تا دیدگاه خبرگان پیرامون ظرفیت‌ها، چالش‌ها و الزامات کاهش سوگیری شناختی در تعامل انسان و سیستم‌های هوش مصنوعی در حوزهٔ جذب و استخدام به‌صورت ساختارمند موردبررسی قرار گیرد. برای تأیید روایی داده‌های کیفی، در این پژوهش از دو روش مکمل افزایش اعتبار شامل «مثلث‌سازی» و «بازبینی توسط مشارکت‌کنندگان» استفاده شد. مطابق با دیدگاه مریام[1] (۲۰۰۹)، مثلث‌سازی به‌معنای بهره‌گیری از منابع داده‌ای یا روش‌های گردآوری متفاوت به‌منظور بررسی همگرایی یافته‌ها و افزایش اعتمادپذیری نتایج است. در این پژوهش، داده‌های حاصل از مصاحبه‌های نیمه‌ساختاریافته با خبرگان دارای پیشینه‌های تخصصی متنوع گردآوری و در کنار اسناد و گزارش‌های مرتبط به‌صورت مقایسه‌ای موردتحلیل قرار گرفتند.

افزون‌براین، به‌منظور اطمینان از دقت و پذیرش‌پذیری یافته‌ها از روش «بازبینی توسط مشارکت‌کنندگان» استفاده شد؛ بدین‌صورت که خلاصه‌ای از یافته‌های کیفی در اختیار چهار نفر از خبرگان قرار گرفت و اصلاحات پیشنهادی آنان در تحلیل نهایی اعمال شد. این اقدامات، اعتبار درونی و اعتمادپذیری نتایج کیفی پژوهش را تقویت نمود. برای تحلیل داده‌های کیفی حاصل از مصاحبه‌های نیمه‌ساختاریافته، از روش تحلیل مضمون براساس رویکرد براون و کلارک[2] (۲۰۰۶) استفاده شد. مصاحبه‌ها با طرح پرسش اصلی پژوهش مبنی‌بر «راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی در فرایند جذب و استخدام منابع انسانی چیست؟» آغاز شد و در ادامه، پرسش‌های تکمیلی متناسب با پاسخ‌های مصاحبه‌شوندگان و به‌منظور تعمیق درک پژوهشگر از تجربه‌ها و دیدگاه‌های آنان مطرح شد. پس از تحلیل داده‌های حاصل از مصاحبه‌ها با استفاده از روش تحلیل مضمون، در مجموع ۱۳۰ مضمون پایه شناسایی شد که در قالب ۱۸ مضمون سازمان‌دهنده و ۶ مضمون فراگیر دسته‌بندی شد. ساختار مضامین استخراج‌شده و نحوهٔ طبقه‌بندی آن‌ها در جدول‌های ۳ و ۴ ارائه شده است.

1. Merriam
2. Braun & Clarke

**جدول ۳. جدول فراوانی کدهای مستخرج از مصاحبه‌ها**

| ردیف | کد مفهومی | کدهای مصاحبه‌شوندگان |
|---|---|---|
| ۱ | مدل‌های تصمیم‌گیری مشارکتی انسان ـ سامانه | q1, q2, q3, q4, q6, q7, q9, q11 |
| ۲ | همکاری‌های بین‌رشته‌ای و ساختارهای تصمیم جمعی | q1, q3, q5, q6, q7, q8, q10, q12 |
| ۳ | بازآموزی شناختی و توانمندسازی ارزیابان در کار با هوش مصنوعی | q2, q3, q4, q8, q9, q10 |
| ۴ | مصاحبه‌های ساختاریافته مبتنی‌بر معیارهای شغلی | q1, q4, q5, q6, q8, q9, q11 |
| ۵ | چارچوب‌های شایستگی شغلی برای هدایت قضاوت ارزیابان | q2, q3, q5, q6, q7, q9, q11, q12 |
| ۶ | ابزارهای ارزیابی و چک‌لیست‌های استاندارد | q1, q2, q4, q6, q8, q9, q10 |
| ۷ | تفسیرپذیری الگوریتم و شفاف‌سازی مدل | q1, q3, q4, q5, q7, q8, q9, q11 |
| ۸ | مستندسازی و ردیابی تصمیم‌های انسان ـ سامانه | q2, q3, q4, q6, q7, q8, q10, q12 |
| ۹ | شفافیت برای متقاضیان و ذی‌نفعان | q1, q3, q5, q7, q8, q9, q11 |
| ۱۰ | پایش عملکرد سامانه در بستر منابع انسانی | q2, q4, q5, q7, q9, q10 |
| ۱۱ | کنترل کیفیت داده‌های استخدامی | q1, q2, q4, q6, q8, q9, q11, q12 |
| ۱۲ | ممیزی و اعتبارسنجی مستقل | q2, q3, q5, q6, q7, q8, q10 |
| ۱۳ | مدیریت داده برای عدالت الگوریتمی | q1, q3, q4, q6, q7, q9, q11 |
| ۱۴ | حذف ویژگی‌های تبعیض‌آمیز از داده‌ها | q2, q4, q5, q8, q10, q12 |
| ۱۵ | پایش عدالت در خروجی‌ها و تصمیم‌های استخدامی | q1, q2, q3, q6, q7, q8, q9 |
| ۱۶ | سیاست‌ها و چارچوب‌های اخلاقی برای کاربرد هوش مصنوعی | q3, q4, q5, q6, q8, q9, q10, q11 |
| ۱۷ | سازوکارهای پاسخ‌گویی در تصمیم‌های استخدامی | q2, q3, q5, q6, q7, q9, q12 |
| ۱۸ | زیرساخت‌های عملیاتی و حاکمیتی سامانه‌های جذب هوشمند | q1, q4, q5, q6, q8, q10, q11 |

منبع: (یافته‌های پژوهش)

**جدول ٤. نمونه‌هایی از گزاره‌های مصاحبه‌شوندگان**

| کدهای مصاحبه‌کنندگان | نمونه‌ای از متن مصاحبه‌ها | مفاهیم مستخرج از کدها |
|---|---|---|
| q1 | به نظر من بهتر است غربالگری اولیهٔ رزومه‌ها توسط سامانه‌های هوشمند انجام شود تا معیارهای نسبتاً عینی مانند مهارت‌ها و تجربه‌های مرتبط مبنای ارزیابی قرار گیرند. این کار می‌تواند دخالت سلیقه‌های فردی را در مراحل اولیه کاهش دهد. البته در برخی موارد لازم است رزومه‌های خاص توسط کارشناسان دوباره بررسی شوند تا از حذف احتمالی استعدادها توسط سامانه جلوگیری شود. | غربالگری اولیهٔ رزومه با هوش مصنوعی؛ معیارهای عینی؛ بازبینی انسانی |

| کدهای مصاحبه‌کنندگان | نمونه‌ای از متن مصاحبه‌ها | مفاهیم مستخرج از کدها |
|---|---|---|
| q2 | اگر آگهی شغلی و شرح وظایف به‌صورت دقیق و استاندارد تدوین شوند، متقاضیان مناسب‌تری جذب می‌شوند. در چنین شرایطی هم ارزیابان انسانی و هم سامانه‌های هوشمند با ابهام کمتری در ارزیابی داوطلبان مواجه خواهند شد. | استانداردسازی آگهی و شرح شغل |
| q3, q4 | – q3: به‌کارگیری آزمون‌های مهارتی استاندارد می‌تواند ارزیابی داوطلبان را منصفانه‌تر کند، زیرا همهٔ افراد براساس شاخص‌های یکسان سنجیده می‌شوند.<br>– q4: اگر مصاحبه‌ها به‌صورت ساختاریافته طراحی شوند و پرسش‌های مشابهی از همه داوطلبان پرسیده شود، امکان مقایسهٔ دقیق‌تر و عادلانه‌تر میان آن‌ها فراهم می‌شود. | آزمون‌های مهارتی استاندارد؛ مصاحبه ساختاریافته |
| q5 | پیش از تحلیل رزومه‌ها لازم است داده‌ها پاک‌سازی شوند. برخی اطلاعات ممکن است ناخواسته موجب سوگیری شوند؛ بنابراین حذف این متغیرها کمک می‌کند ارزیابی بیشتر بر شایستگی‌های واقعی داوطلبان متمرکز شود. | پاک‌سازی داده‌ها؛ حذف متغیرهای سوگیری‌زا |
| q6 | برای افزایش دقت در تصمیم‌گیری‌های استخدامی، لازم است معیارهای شایستگی به‌صورت شفاف تعریف شوند. به نظر من این معیارها باید مبتنی‌بر رفتارهای قابل‌مشاهده باشند تا دخالت قضاوت‌های سلیقه‌ای در ارزیابی کاهش یابد. | تعریف شفاف معیارهای شایستگی؛ شاخص‌های رفتاری |
| q7, q8 | – q7: سامانه‌های هوش مصنوعی پیش از به‌کارگیری در فرایند استخدام باید بر روی گروه‌های مختلف آزمون شوند تا مشخص شود برای همهٔ داوطلبان به‌شکل عادلانه عمل می‌کنند.<br>– q7: همچنین لازم است شاخص‌های عدالت و تنوع به‌صورت مستمر پایش شوند تا در صورت مشاهدهٔ نشانه‌هایی از سوگیری، امکان اصلاح آن فراهم باشد. | آزمون الگوریتم روی گروه‌های مختلف؛ پایش عدالت |
| q9 | اگر سامانه‌های هوش مصنوعی بدون طراحی مناسب و نظارت مستمر مورداستفاده قرار گیرند، ممکن است سوگیری‌های موجود را تقویت کنند. بنابراین شفافیت و قابلیت تبیین منطق تصمیم‌گیری باید از مرحلهٔ طراحی در سامانه لحاظ شود و عملکرد آن نیز به‌صورت دوره‌ای مورد ممیزی قرار گیرد. | شفافیت و قابلیت توضیح الگوریتم |

منبع: (یافته‌های پژوهش)

کدها حاصل بررسی دقیق متن مصاحبه‌ها و شناسایی مفاهیم مطرح‌شده در آن‌ها بودند. در بسیاری از بخش‌های مصاحبه‌ها، مصاحبه‌شوندگان به مطالب مشترک و مفاهیم مشابهی اشاره کرده بودند که با کدهای استخراج‌شده ارتباط داشتند. بر همین اساس، کدهای اولیه پس از بررسی و مقایسه با یکدیگر در سطحی بالاتر سازمان‌دهی شدند و در لایهٔ بعدی، مؤلفه‌ها و مضامین مرتبط با آن‌ها شناسایی و ارائه شدند.

**جدول ۵. دسته‌بندی مضامین استخراج‌شده**

| مضمون فراگیر | مضمون سازمان‌دهنده | مضامین پایه |
|---|---|---|
| یکپارچه‌سازی تصمیم‌گیری تعاملی انسان و سامانه‌های هوشمند | مدل‌های تصمیم‌گیری مشارکتی انسان ـ سامانه | تعریف شفاف نقش انسان در کنار سیستم هوش مصنوعی در تصمیم‌گیری؛ تفکیک غربال اولیه بین هوش مصنوعی و انسان؛ غربالگری هوشمند و تصمیم نهایی انسانی پاسخ‌گو؛ تصمیم نهایی انسان؛ پذیرش پیشنهاد سیستم با مسئولیت نهایی انسان؛ تصمیم نهایی ترکیبی الگوریتم و کمیته متنوع؛ هم‌افزایی انسان و الگوریتم؛ تقسیم مرحله‌ای تصمیم جذب؛ تقسیم تصمیم جذب؛ هشدار سیستم به جای حکم؛ هشدارگری سیستم به جای تصمیم؛ محدودسازی نقش سیستم در تصمیم‌گیری؛ جایگاه مشورتی الگوریتم؛ پشتیبانی سیستم بدون جایگزینی انسان؛ تفکیک پیشنهاد سیستم از تصمیم نهایی؛ تصمیم قابل‌بازگشت؛ تصمیم قابل‌توضیح |
| | همکاری‌های بین‌رشته‌ای و ساختارهای تصمیم جمعی | کمیته جمعی به‌جای تصمیم فردی در استخدام؛ تصمیم‌گیری جمعی چندتخصصی؛ ایجاد کمیته‌های بین‌رشته‌ای؛ میان‌رشته‌ای‌سازی تیم توسعه؛ مشارکت ذی‌نفعان متنوع در تعریف نیازها؛ حضور چند ارزیاب با پیشینه‌های متفاوت؛ مشارکت فعال منابع انسانی در تصمیم‌های داده‌ای؛ مشورت ساختارمند با خبرگان فناوری؛ طراحی چندتخصصی سامانه جذب؛ تیم مشترک برای ممیزی سوگیری |
| | بازآموزی شناختی و توانمندسازی ارزیابان در کار با هوش مصنوعی | استفاده از آموزش سوگیری‌های شناختی در استخدام؛ آموزش رسمی سوگیری‌های شناختی؛ آموزش تکنیک‌های کاهش سوگیری مبتنی‌بر چک‌لیست؛ آموزش سوگیری‌های شناختی به کارکنان جذب؛ کارگاه‌های شناخت سوگیری؛ کارگاه‌های تخصصی سوگیری؛ آموزش مبتنی‌بر تجربه؛ آموزش مبتنی‌بر مصادیق؛ آموزش خطر اعتماد بیش از حد به اتوماسیون؛ آموزش خطر رد شهودی خروجی سیستم؛ آموزش تفسیر گزارش‌های الگوریتمی؛ پرورش تفکر انتقادی؛ توسعهٔ سواد داده‌ای؛ فهم محدودیت‌های شناخت انسانی؛ خودآگاهی شناختی در تصمیم‌گیری |
| استانداردسازی فرایندهای ارزیابی انسان‌محور | مصاحبه‌های ساختاریافته مبتنی‌بر معیارهای شغلی | مصاحبه‌های ساختاریافته با معیارهای یکسان؛ طراحی مصاحبه‌های ساختاریافته با سؤال‌های استاندارد؛ یکسان‌سازی سؤال‌های مصاحبه؛ مهار سلیقه‌محوری در تصمیم؛ تعلیق قضاوت سریع؛ جداسازی حس اولیه در ارزیابی؛ حذف رتبه‌بندی قطعی؛ استفاده از چند ارزیاب |
| | چارچوب‌های شایستگی شغلی برای هدایت قضاوت ارزیابان | تعریف دقیق معیارهای شایستگی؛ تعریف شایستگی‌های شغل‌محور؛ تمرکز بر شایستگی‌های ضروری؛ عینی‌سازی و چندمنبعی کردن ارزیابی شایستگی؛ استفاده از آزمون‌های مهارتی استاندارد؛ اولویت دادن به شواهد عینی شایستگی؛ سنجش تناسب فرهنگی؛ طراحی چارچوب استاندارد شغل؛ تدوین شرح شغل دقیق و واقعی |
| | ابزارهای ارزیابی و چک‌لیست‌های استاندارد | چک‌لیست شفاف مهارت، تجربه و دانش شغل؛ استفاده اجباری از چک‌لیست پیش از تصمیم استخدامی؛ چک‌لیست ثبت نقاط قوت و ضعف؛ طراحی جداول امتیازدهی؛ امتیازدهی جداگانه به هر شایستگی؛ امتیازدهی ساختاریافته؛ استفاده از فرم‌های امتیازدهی عددی |
| شفافیت و توضیح‌پذیری سامانه‌های هوشمند | تفسیرپذیری الگوریتم و شفاف‌سازی مدل | شفافیت اهداف و سازوکار سامانه؛ شفاف‌سازی منطق تصمیم‌گیری الگوریتم؛ توضیح‌پذیری خروجی سیستم؛ گزارش شفاف منطق تصمیم سامانه؛ نمایش بصری وزن معیارهای شغلی؛ استفاده از ابزارهای توضیح‌پذیری تصمیم؛ طراحی رابط کاربری شفاف درباره نقش متغیرها |

| مضمون فراگیر | مضمون سازمان‌دهنده | مضامین پایه |
|---|---|---|
| | مستندسازی و ردیابی تصمیم‌های انسان؛ سامانه | مستندسازی دلایل تصمیم استخدامی؛ ثبت و مستندسازی مسیر تصمیم؛ ثبت تغییرات داده؛ ردیابی نسخه تصمیم؛ ردیابی تغییرات تصمیم؛ ثبت اختلاف نظر در تصمیم‌گیری |
| | شفافیت برای متقاضیان و ذی‌نفعان | شفاف‌سازی فرایند جذب؛ شفاف‌سازی معیارهای جذب؛ شفاف‌سازی دلایل رد متقاضی؛ ایجاد پورتال شفاف اطلاع‌رسانی مراحل جذب؛ ارائه گزارش قابل‌فهم به متقاضیان |
| نظارت و اعتبارسنجی الگوریتمی | پایش عملکرد سامانه در بستر منابع انسانی | پایش مستمر نتایج استخدام؛ پایش پویای داده‌ها؛ تحلیل خروجی سیستم؛ تحلیل خطا برای شناسایی تبعیض؛ مقایسهٔ عملکرد سیستم با نتایج واقعی؛ پایش الگوهای غیرمعمول |
| | کنترل کیفیت داده‌های استخدامی | پاکسازی داده‌های ورودی؛ حذف داده‌های مزاحم و نامرتبط؛ کنترل کیفیت اطلاعات پیش از ثبت؛ بازبینی داده‌های ردشده؛ بازبینی داده‌های آموزشی؛ بازنگری مستمر داده |
| | ممیزی و اعتبارسنجی مستقل | ممیزی دوره‌ای عملکرد مدل؛ ممیزی مستقل داده‌ها؛ ممیزی منظم سوگیری؛ اعتبارسنجی دوره‌ای الگوریتم؛ ارزیابی داخلی به‌روزرسانی مدل؛ ممیزی ویژگی‌های تبعیض‌آمیز |
| شناسایی و کاهش کمی سوگیری‌های داده‌ای | مدیریت داده برای عدالت الگوریتمی | مدیریت تنوع داده برای کاهش سوگیری؛ تنوع و توازن داده‌های ورودی؛ متعادل‌سازی داده‌های آموزشی؛ افزایش وزن گروه‌های کمترنمایان؛ کاهش وزن داده‌های تاریخی سوگیرانه؛ استفاده از دادهٔ مصنوعی برای گروه‌های کم‌نمایان |
| | حذف ویژگی‌های تبعیض‌آمیز از داده‌ها | حذف ویژگی‌های حساس از داده‌ها؛ حذف متغیرهای جانشین؛ اجتناب از داده‌های تبعیض‌آمیز و جانبدارانه؛ تفکیک متغیرهای ضروری و غیرضروری |
| | پایش عدالت در خروجی‌ها و تصمیم‌های استخدامی | ممیزی توزیع پذیرش و رد بین گروه‌ها؛ تنظیم آستانه‌های تصمیم‌گیری؛ استفاده از معیارهای عدالت؛ آزمون پادواقعی مدل؛ بازطراحی مدل در صورت مشاهده سوگیری |
| پروتکل‌های اخلاقی و عملیاتی سامانه‌های هوشمند | سیاست‌ها و چارچوب‌های اخلاقی برای کاربرد هوش مصنوعی | تدوین سیاست سازمان درباره استفاده از سامانه‌های هوشمند در استخدام؛ تدوین چارچوب اخلاقی استخدام؛ مسئولیت اخلاقی در برابر سوگیری؛ مسئولیت اجتماعی سامانه‌های هوشمند |
| | سازوکارهای پاسخ‌گویی در تصمیم‌های استخدامی | ایجاد سازوکار اعتراض برای متقاضیان ردشده؛ گزارش‌دهی دوره‌ای درباره عدالت سامانه؛ الزام به مستندسازی تصمیم‌ها؛ پاسخ‌گویی مدیریتی در قبال سوگیری |
| | زیرساخت‌های عملیاتی و حاکمیتی سامانه‌های جذب هوشمند | سامانه یکپارچه استخدام؛ داشبورد نظارتی سوگیری؛ سیستم گزارش‌دهی خودکار سوگیری؛ احراز هویت چندعاملی برای ردیابی تصمیم‌ها؛ ایجاد مخزن تصمیمات استخدامی برای تحلیل سوگیر |

منبع: (یافته‌های پژوهش)

در پژوهش حاضر، نظریه جامعهٔ مادی به‌عنوان لنز نظری و چارچوبی حساس‌کننده برای تفسیر درهم‌تنیدگی عوامل انسانی، فناورانه و سازمانی در فرایند جذب و استخدام مبتنی‌بر هوش مصنوعی به‌کار گرفته شد. بر این مبنا، مضامین استخراج‌شده پس از کدگذاری استقرایی، در مرحله تفسیر از منظر کنش انسانی، سامانه‌های هوشمند، داده‌ها، الگوریتم‌ها،

فرایندهای سازمانی و سازوکارهای پاسخ‌گویی بازخوانی شدند تا پیوند میان مادیت فناوری و عمل سازمانی آشکار شود. جدول ۶ این انطباق تحلیلی را در سطح مضامین فراگیر و سازمان‌دهنده نمایش می‌دهد.

**جدول ٦. تحلیلِ انطباقیِ مضامین پژوهش با ابعاد نظریه جامعهٔ مادی (لنز تحلیل یافته‌ها)**

| بُعد تحلیلی جامعهٔ مادی | مضامین فراگیر مرتبط | مضامین سازمان‌دهنده مرتبط | دلالت تحلیلی |
|---|---|---|---|
| درهم‌تنیدگی کنش انسانی و فناوری | یکپارچه‌سازی تصمیم‌گیری تعاملی انسان و سامانه‌های هوشمند | مدل‌های تصمیم‌گیری مشارکتی انسان ـ سامانه؛ همکاری‌های بین‌رشته‌ای و ساختارهای تصمیم جمعی؛ بازآموزی شناختی و توانمندسازی ارزیابان در کار با هوش مصنوعی | سوگیری نه حاصل خطای صرف انسان است و نه نقص صرف فناوری؛ بلکه در نقطه تعامل این دو پدید می‌آید. |
| مادی‌شدن قضاوت در داده و معیار | شناسایی و کاهش کمی سوگیری‌های داده‌ای | مدیریت داده برای عدالت الگوریتمی؛ حذف ویژگی‌های تبعیض‌آمیز از داده‌ها؛ پایش عدالت در خروجی‌ها و تصمیم‌های استخدامی؛ کنترل کیفیت داده‌های استخدامی | سوگیری‌های انسانی زمانی ماندگار می‌شوند که در داده‌ها و معیارها تثبیت شوند. |
| اجراگری و شفافیت تصمیم | شفافیت و توضیح‌پذیری سامانه‌های هوشمند؛ نظارت و اعتبارسنجی الگوریتمی | تفسیرپذیری الگوریتم و شفاف‌سازی مدل؛ پایش عملکرد سامانه در بستر منابع انسانی؛ مستندسازی و ردیابی تصمیم‌های انسان ـ سامانه؛ شفافیت برای متقاضیان و ذی‌نفعان | شفافیت، سازوکاری برای مشاهده‌پذیر کردن منطق تصمیم و امکان نقد آن است. |
| حکمرانی و پاسخ‌گویی شبکه‌ای | پروتکل‌های اخلاقی و عملیاتی سامانه‌های هوشمند | ممیزی و اعتبارسنجی مستقل؛ سازوکارهای پاسخ‌گویی در تصمیم‌های استخدامی؛ سیاست‌ها و چارچوب‌های اخلاقی برای کاربرد هوش مصنوعی؛ زیرساخت‌های عملیاتی و حاکمیتی سامانه‌های جذب هوشمند | پاسخ‌گویی باید از سطح فردی به سطح شبکه انسان ـ فناوری ـ سازمان منتقل شود. |

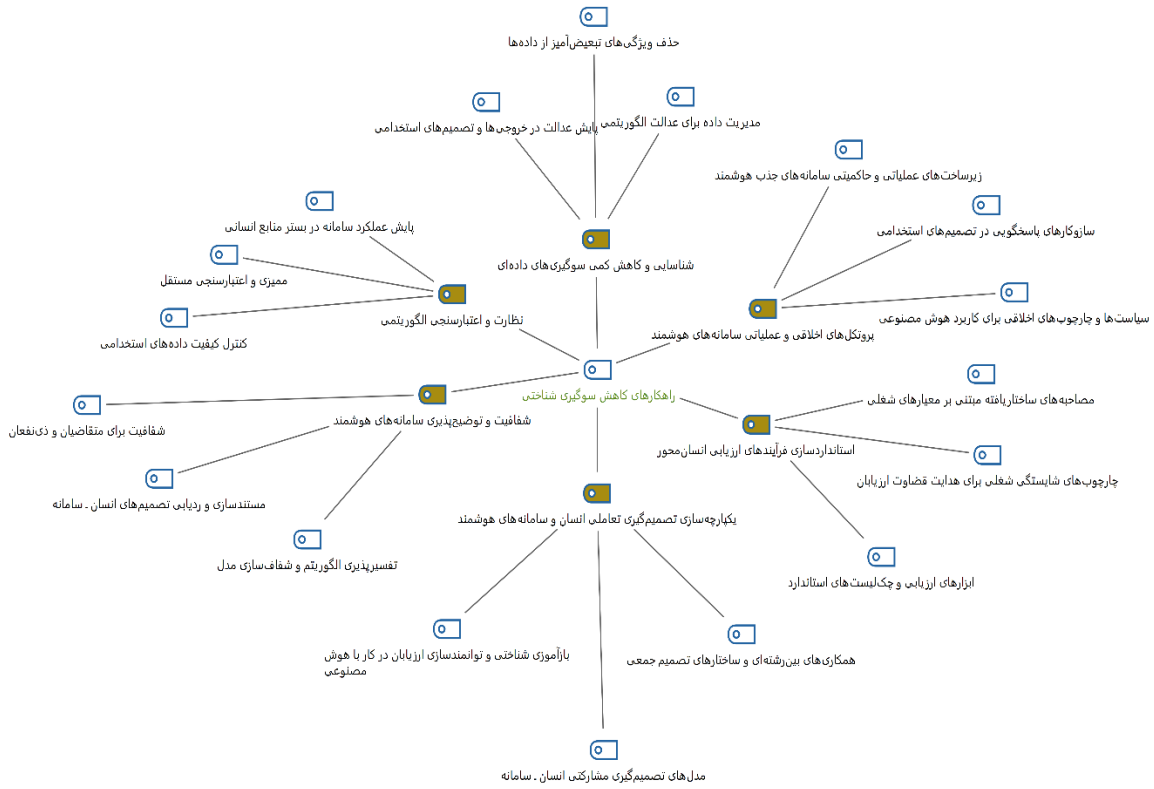
منبع: (یافته‌های پژوهش)

یکی از محورهای اصلی یافته‌ها، یکپارچه‌سازی تصمیم‌گیری تعاملی انسان و سامانه‌های هوشمند است که از خلال مضامینی مانند تصمیم‌گیری مشارکتی، همکاری بین‌رشته‌ای و بازآموزی شناختی ارزیابان صورت‌بندی شد. مشارکت‌کنندگان تأکید داشتند که هوش مصنوعی نباید جایگزین کامل قضاوت انسانی شود، بلکه باید در نقش ابزار پشتیبان تصمیم عمل کند. یکی از خبرگان در این زمینه بیان کرد: «باید دقیق مشخص کنیم کجا انسان تصمیم‌گیرنده است و کجا ابزار فقط کمک‌کننده؛ چراکه عدم‌شفافیت در این مرزبندی منجربه تکیه بیش‌ازحد یا طرد کامل سیستم می‌شود.» از منظر جامعهٔ مادی، تصمیم استخدامی محصول «عاملیت توزیع‌شده» میان ارزیاب، سامانه و قواعد سازمانی است؛ بنابراین، کاهش سوگیری مستلزم تنظیم دقیق این نسبت در یک شبکه پاسخ‌گو است. بخش دیگری از یافته‌ها به

استانداردسازی فرایندهای ارزیابی انسان‌محور از طریق ساختاریافته‌سازی مصاحبه‌ها و تعریف شایستگی‌های دقیق اختصاص داشت.

مشارکت‌کنندگان، استانداردسازی را راهی برای مهار قضاوت‌های شهودی و سوگیری‌های کلیشه‌ای می‌دانستند. یکی از مشارکت‌کنندگان گفت: «اگر بخواهیم کیفیت تصمیم‌گیری بالا برود، باید معیارهای شایستگی را شفاف تعریف کنیم. تا وقتی برداشت‌ها ذهنی باشد، احتمال سوگیری بالا می‌رود.» در خوانش جامعهٔ مادی، چک‌لیست‌ها و چارچوب‌های شایستگی، «واسطه‌های مادی» هستند که قضاوت انسانی را از سطح برداشت‌های پراکنده به سمت ارزیابی مبتنی‌بر شواهد هدایت کرده و تعامل میان شناخت انسانی و فرایند سازمانی را بازتنظیم می‌کنند. تحلیل داده‌ها همچنین بر اهمیت شفافیت و توضیح‌پذیری سامانه‌های هوشمند تأکید داشت. یافته‌ها نشان داد که خروجی سامانه زمانی قابل‌اعتماد است که منطق آن برای ارزیابان فهم‌پذیر باشد. یکی از مشارکت‌کنندگان اظهار داشت: «برای هر شایستگی باید سطح‌بندی مشخص همراه با مثال‌های رفتاری داشته باشیم تا همه ارزیاب‌ها برداشت یکسانی از آن داشته باشند.» مستندسازی مسیر تصمیم و ردیابی تغییرات، تصمیم استخدامی را از یک «جعبه سیاه» خارج کرده و آن را به‌مثابه محصولی از شبکه روابط انسانی و فنی مرئی می‌سازد. در لایه نظارت و اعتبارسنجی الگوریتمی، مصاحبه‌ها نشان دادند که اعتبار سامانه باید در جریان استفاده واقعی و براساس پیامدهای استخدامی سنجیده شود. یکی از مصاحبه‌شونده‌ها بیان کرد: «نظارت بر سیستم‌های هوشمند کاملاً به عدالت سازمانی وصل است؛ چیزی که پایه اعتماد کارکنان محسوب می‌شود.» از منظر جامعهٔ مادی، فناوری در بستر سازمان ایستا نیست و در جریان «اجراگری» دائماً بازتولید می‌شود؛ از این رو پایش مستمر نتایج و اعتبارسنجی مستقل، شرط لازم برای جلوگیری از تثبیت سوگیری در چرخهٔ عمر فناوری است.

یافته‌ها در لایه شناسایی و کاهش کمی سوگیری‌های داده‌ای نشان داد که داده‌های استخدامی، بازتابی از ترجیحات تاریخی سازمان هستند. یکی از مشارکت‌کنندگان گفت: «اگر داده‌های قبلی سازمان سوگیری داشته باشد، مدل همان را یاد می‌گیرد و حتی ممکن است آن را دقیق‌تر تکرار کند.» داده‌ها در اینجا «حافظه رسوب‌یافته» سازمان تلقی می‌شوند؛ بنابراین، عدالت الگوریتمی نیازمند مدیریت فعال داده، حذف ویژگی‌های تبعیض‌آمیز و مداخله در پیوند میان تاریخچه سازمانی و منطق ماشین است. در نهایت، ضرورت استقرار پروتکل‌های اخلاقی و عملیاتی برجسته شد. مشارکت‌کنندگان معتقد بودند مسئولیت تصمیم نباید پشت سامانه پنهان شود. یکی از آن‌ها بیان کرد: «حتی وقتی ابزار هوشمند پیشنهاد می‌دهد، سازمان باید پاسخ‌گو باشد که چرا آن پیشنهاد پذیرفته یا رد شده است.» از منظر جامعهٔ مادی، مسئولیت در یک عامل منفرد خلاصه نمی‌شود، بلکه در شبکه‌ای از نقش‌ها، سیاست‌ها و زیرساخت‌ها توزیع می‌شود. از این منظر، مدیریت سوگیری نه یک اصلاح فنی صرف، بلکه فرایند بازتنظیم مستمر پیوندهای میان ارزیابان، داده‌ها، منطق الگوریتمی و قواعد سازمانی برای دستیابی به پیامدهای عادلانه است.




**شکل ۱. شبکهٔ مضامینِ راه‌کارهای کاهش سوگیری شناختی در به‌کارگیری هوش مصنوعی در جذب و استخدام؛ در پرتو نظریه جامعهٔ مادی (خروجی نرم‌افزار مکس کیودا)**

## نتیجه‌گیری و پیشنهادها

یافته‌های این پژوهش نشان می‌دهد که سوگیری در استخدام مبتنی‌بر هوش مصنوعی را نمی‌توان فقط به‌عنوان پیامد نقص‌های فنی الگوریتم‌ها یا ناکارآمدی سامانه‌های هوشمند توضیح داد. شکل‌گیری سوگیری در استخدام هوشمند در بستری پیچیده از عوامل اجتماعی، شناختی و فناورانه رخ می‌دهد و ریشه در سازوکارهای سازمانی و نحوهٔ حکمرانی بر فناوری دارد. این برداشت با منطق نظریهٔ جامعهٔ مادی هم‌سو است که بر شکل‌گیری پدیده‌های سازمانی در پیوند متقابل میان عناصر انسانی، فناورانه و مادی تأکید دارد. در این بستر، داده‌های ورودی، مفروضات طراحان، قضاوت‌های ارزیابان انسانی، نحوهٔ تفسیر خروجی‌های الگوریتمی و سازوکارهای نظارتی سازمان، به‌صورت درهم‌تنیده بر فرایند تصمیم‌گیری اثر می‌گذارند. بر این اساس، عدالت در استخدام هوشمند، نه محصول فناوری به خودی‌خود، بلکه نتیجهٔ نحوهٔ طراحی، استقرار، نظارت و پاسخ‌گویی به این فناوری در بستر سازمانی است.

از جنبهٔ نظری، این پژوهش سهمی در توسعهٔ ادبیات منابع انسانی دیجیتال و مطالعات مرتبط با هوش مصنوعی در استخدام دارد. نخست آنکه نتایج این مطالعه، توجه را به مرحلهٔ پیشاالگوریتمی شکل‌گیری سوگیری جلب می‌کند؛ مرحله‌ای که در آن قضاوت‌های انسانی، سوگیری‌های شناختی و معیارهای انتخاب شغلی پیش از ورود داده‌ها به مدل‌های الگوریتمی در ساختار تصمیم‌گیری نهادینه می‌شوند. درحالی‌که بسیاری از پژوهش‌های پیشین تمرکز خود را بر

سوگیری در خروجی الگوریتم‌ها یا کیفیت داده‌ها قرار داده‌اند، یافته‌های این پژوهش نشان می‌دهد که بخش مهمی از تبعیض پیش‌تر و در فرایندهایی مانند تعریف شایستگی‌های شغلی، انتخاب شاخص‌های ارزیابی، تدوین آگهی‌های استخدامی و نحوهٔ برچسب‌گذاری داده‌ها شکل می‌گیرد. بنابراین، سوگیری الگوریتمی در بسیاری از موارد بازتاب سوگیری‌های شناختی انسانی است که به‌صورت نظام‌مند به سامانه‌های هوشمند منتقل می‌شود. دوم آنکه یافته‌های این پژوهش با فاصله‌گرفتن از نگاه دوگانهٔ «انسان در برابر ماشین»، درکی تعاملی‌تر از فرایند استخدام هوشمند ارائه می‌دهد. از این جنبه، نظریهٔ جامعهٔ مادی امکان می‌دهد که رابطهٔ انسان و سامانهٔ هوشمند، نه به‌صورت تقابل، بلکه به‌عنوان نوعی درهم‌تنیدگی میان قضاوت انسانی، داده، الگوریتم و قواعد سازمانی فهم شود. در این چارچوب، هوش مصنوعی می‌تواند به‌عنوان ابزار پشتیبان تصمیم‌گیری عمل کند، درحالی‌که مسئولیت نظارت، تفسیر و پاسخ‌گویی همچنان برعهدهٔ انسان باقی می‌ماند. این برداشت با برخی مطالعات اخیر در حوزهٔ تعامل انسان و هوش مصنوعی هم‌سو است که بر نقش نظارتی انسان در کنار سامانه‌های الگوریتمی تأکید دارند (سلیمانی و همکاران، ۲۰۲۵). تحقق این هم‌افزایی، مستلزم اقداماتی مانند استانداردسازی فرایندهای ارزیابی، افزایش شفافیت در تصمیم‌گیری و ممیزی مستمر سامانه‌های الگوریتمی است.

از جنبهٔ نظریه‌پردازی نیز نتایج این پژوهش به فهم سوگیری در استخدام هوشمند به‌عنوان پدیده‌ای چندسطحی کمک می‌کند. در این چارچوب، سوگیری نه در یک نقطهٔ منفرد، بلکه در زنجیره‌ای از تعاملات و تصمیم‌ها شکل می‌گیرد؛ از قضاوت انسانی تا تولید داده، از داده تا الگوریتم، از الگوریتم تا تفسیر انسانی و در نهایت تا سیاست‌ها و رویه‌های سازمانی. چنین نگاهی با رویکردهایی که در ادبیات عدالت الگوریتمی بر بررسی کل چرخهٔ تصمیم‌گیری تأکید دارند، هم‌خوانی دارد (ریگوتی و فوش ویلارونگا، ۲۰۲۴). بر این اساس، اصلاح الگوریتم‌ها اگرچه ضروری است؛ اما برای تحقق عدالت در استخدام کافی نیست و لازم است کل فرایند تصمیم‌گیری، از مرحلهٔ تولید داده تا سازوکارهای پاسخ‌گویی سازمانی، موردتوجه قرار گیرد.

نتایج این پژوهش پیامدهای مهمی برای سازمان‌ها و مدیران منابع انسانی دارد. یافته‌ها نشان می‌دهد که استقرار سامانه‌های هوش مصنوعی در استخدام، بدون بازنگری در فرایندهای انسانی، می‌تواند به‌جای کاهش سوگیری، به تثبیت یا حتی تشدید آن منجر شود. بنابراین، استفادهٔ مسئولانه از هوش مصنوعی در استخدام مستلزم توجه هم‌زمان به حاکمیت و کیفیت داده‌ها، استانداردسازی فرایندهای تصمیم‌گیری انسانی و ایجاد سازوکارهای شفافیت، توضیح‌پذیری و ممیزی مستمر برای نظارت بر عملکرد سامانه‌های الگوریتمی است. در مقایسه با مطالعات پیشین نیز می‌توان گفت نتایج این پژوهش در امتداد مطالعاتی قرار می‌گیرد که به مخاطرات تبعیض الگوریتمی در مدیریت منابع انسانی اشاره کرده‌اند (البارودی و همکاران، ۲۰۲۴)، اما درعین‌حال نشان می‌دهد که منشأ این تبعیض را باید در تعامل میان تصمیم‌گیری انسانی، داده‌ها و طراحی سازمانی جست‌وجو کرد. از این رو، این پژوهش می‌کوشد با ارائهٔ یافته‌ها و پیشنهادهای مبتنی‌بر تحلیل کیفی، فاصلهٔ میان مباحث نظری عدالت الگوریتمی و سازوکارهای عملی کاهش سوگیری در فرایندهای استخدام را کاهش دهد. در به‌کارگیری سامانه‌های هوشمند جذب و استخدام، آنچه اهمیت بنیادین دارد، حفظ بازبینی انسانی

مستند است. در سال‌های اخیر برخی سازمان‌ها برای مدیریت رزومه‌ها، ارتباط با متقاضیان و غربالگری اولیه به ابزارهـای دیجیتال روی آورده‌اند، اما پژوهش حاضر بر این باور است که چنین فناوری‌هایی، نباید جایگزین قضـاوت نهـایی انسـان شوند؛ بلکه باید در نقش پشتیبان تصمیم ظاهر شوند؛ یعنی با مرتب‌سازی نخستین رزومه‌ها و کـاهش حجـم اطلاعـات، کارشناسان جذب را یاری کنند؛ اما مُهر نهایی پذیرش یا رد، همچنـان از آنِ ارزیـاب انسـانی باشـد. از همـین رو، بـرای پیشگیری از انتقال سوگیری‌های شناختی نهفته در داده‌ها، ضروری است که خروجی این سـامانه‌ها پـیش از نهایی‌شـدن توسط انسان به‌دقت بررسی و در صورت نیاز اصلاح شود. دربارهٔ مسیرهای پژوهشی آینده، چند محور قابـل طـرح اسـت. نخست، از آنجا که مشارکت‌کنندگان این پژوهش طیف گسترده‌ای از گروه‌های سنی، از نیروی کار جوان تا میان‌سـال، را دربرمی‌گرفتند، این تنوع، غنـای یافتـه‌ها را از جنبـهٔ تجربـه‌های متفـاوت شـغلی، ممکـن سـاخت و تصـویری جـامع از دغدغه‌های نسل‌های مختلف در مواجهه با استخدام هوشمند ارائه داد. با این حال، بررسی عمیـق‌تر و متمرکـز اثـر ایـن فناوری بر فرصت‌های شغلی گروه‌های خاص، از جمله نیروی کار جوان، می‌تواند در پژوهش‌های آینده پی گرفتـه شـود؛ به‌ویژه اگر این مطالعات به داده‌های واقعی سازمانی نیز دسترسی داشته باشند؛ یعنی به آمار دقیـق میـزان خودکارسـازی فرایند جذب، نحوهٔ مواجهه مدیران با خروجی سامانه، معیارهای پـذیرش و رد متقاضـیان و امکـان اعتـراض یـا بـازبینی تصمیم. در این مسیر، تدوین یک چارچوب تحلیلی برای سنجش پیامدهای اجتماعی این فناوری، به‌ویژه در سـازمان‌های بزرگ و بوروکراتیک، می‌تواند به فهم عمیق‌تر چالش‌های استخدامی کمک کند. افزون‌بـراین، پیشـنهاد می‌شـود کـه در مطالعات بعدی، شاخص‌های عملیاتی نظارت انسانی به‌طور تجربی آزمـون شـوند؛ شـاخص‌هایی همچـون نـرخ بـازبینی پرونده‌ها توسط انسان، میزان مغایرت میان نظر ارزیاب و خروجی سامانه، آمـار تصـمیم‌هایی کـه پـس از بـازبینی تغییـر کرده‌اند و درجهٔ مستندسازی دلایل هر تصمیم. این سنجه‌ها گرچه در پژوهش حاضـر فقـط بـه‌عنوان مفـاهیمی نظـری مطرح شدند، می‌توانند زمینه‌ای برای عینیت بخشیدن به مباحث کیفی این حوزه فراهم آورند. بررسی تفاوت‌های اقتضایی میان انواع سازمان‌ها نیز از ضرورت‌های پژوهش‌های آینده است؛ زیرا سازوکار بروز و کنترل سـوگیری در یـک سـازمان چابک با یک نهاد بوروکراتیک تفاوت‌های ماهوی دارد. یافته‌ها نشان می‌دهد که مداخلات کاهش سوگیری باید با بسـتر سازمانی تناسب داشته باشند. برای نمونه، در کسب‌وکارهای دیجیتال، پایش مستمر عملکرد مـدل و ارزیـابی عـدالت آن اهمیت بیشتری دارد؛ در صنایع تولیدی، بازنگری معیارهای شایسـتگی و اسـتفاده از مصـاحبه‌های سـاختاریافته می‌توانـد مناسب‌تر باشد؛ در سازمان‌های دولتی، شفافیت در زنجیره تصمیم‌گیری و فراهم‌کردن امکان اعتراض اهمیـت می‌یابـد؛ و در سازمان‌های کوچک، محدودسازی اتوماسیون و اتکا به ارزیابی انسانیِ مستند، رویکردی واقع‌بینانه‌تر اسـت. ایـن نگـاه اقتضایی، چارچوب اولیه‌ای برای مطالعات تطبیقی فراهم می‌آورد تا بتوان به الگوهای دقیق‌تری برای استفاده مسئولانه از هوش مصنوعی در جذب و استخدام دست یافت. در به‌کارگیری این سامانه‌ها، نظارت انسانی هنگامی معنـا پیـدا می‌کنـد که عملیاتی و مستند باشد. سازمان‌ها باید با پایشِ شاخص‌های کلیدی نظیر نرخ انتخـاب نسـبی گروه‌هـا، شـکاف دقـت مدل، میزان ردپای ممیزی و نرخ مداخلات انسانی، ریسک‌های پنج‌گانه‌ای را مدیریت کنند کـه عبارت‌انـد از: تبعـیض در غربالگری، اتکای افراطی به خروجی مدل، تغییر نامتوازن ساختار نیروی کار، عدم شـفافیت در فراینـد اعتـراض و نشـت

داده‌ها. برای مهار این ریسک‌ها، راه‌کارهایی نظیر ممیزی مستقل، بازبینی موارد مرزی، ایجاد کانال‌های رسمی اعتراض و سیاست‌های حداقل‌سازی داده ضروری است. این چارچوبِ کنترلی، یک الگوی اقتضایی است که باید متناسب با بلوغ دیجیتال سازمان پیاده‌سازی شود. نکته حائز اهمیت این است که حضور انسان، تضمین‌کنندهٔ بی‌طرفی نیست؛ اگر این نظارت بر معیارهای شایستگی‌محور و مستند استوار نباشد، صرفاً به پوششی برای بازتولید سوگیری‌های شناختی و سلیقه‌ای ارزیاب تبدیل خواهد شد؛ ازاین‌رو بهره‌گیری از چک‌لیست‌های ارزیابی و ثبت دقیقِ دلایلِ تصمیم، برای جلوگیری از این خطاها الزامی است. شفافیت نیز در این میان نقشی کلیدی ایفا می‌کند؛ سازمان‌ها باید برای متقاضیان و تصمیم‌گیرندگان تبیین کنند که سامانه‌ها در کدام مرحله از فرایند وارد می‌شوند، خروجی آن‌ها چه سهمی در تصمیم نهایی دارد و در صورت بروز خطا یا اعتراض، چه مسیری برای بازبینی وجود دارد. بدیهی است که این شفافیت هم اعتماد متقاضیان را جلب می‌کند و هم مسئولیت‌پذیری سازمان را در برابر تصمیم‌های استخدامی تقویت می‌بخشد. در نهایت، به‌کارگیری هوش مصنوعی در استخدام زمانی بی‌طرفانه و اخلاقی خواهد بود که در چارچوب اصول شفافیت، انصاف و پاسخ‌گویی قرار گیرد؛ اصولی چون پرهیز از تبعیض، حفاظت از داده‌های شخصی، توضیح‌پذیری تصمیم‌ها و امکان اعتراض. بر اساس یافته‌های این پژوهش، هوش مصنوعی نه جانشین کامل قضاوت انسانی، بلکه ابزاری در خدمت تصمیم‌گیری است؛ ابزاری که بار پردازش اطلاعات را سبک می‌کند؛ اما تصمیم نهایی همچنان با مسئولیت انسانی و بر پایه معیارهای شغلی اتخاذ می‌شود. این رویکرد هم از جنبهٔ عملیاتی برای سازمان‌ها معنا دارد و هم از منظر اخلاقی برای محققان این حوزه. مواجهه مؤثر با سوگیری در استخدام هوشمند تنها با اصلاح الگوریتم‌ها یا افزودن ابزارهای فناورانه میسر نمی‌شود؛ بلکه نیازمند سامان‌دهی هم‌زمان داده‌ها، معیارهای شایستگی، نقش ارزیابان انسانی، شفافیت فرایندها و سازوکارهای حکمرانی فناوری است. به عبارت دیگر، تصمیم استخدامی یک فرایند مشترک میان انسان و سامانه است؛ جایی که فناوری خطاهای احتمالی را آشکار می‌سازد و انسان، با تکیه بر خرد و تجربه، مسئولیت نهایی پذیرش یا اصلاح را برعهده می‌گیرد. چنین نگاهی می‌تواند مسیر استفاده محتاطانه، منصفانه و اتکاپذیرتری را از هوش مصنوعی در فرایند جذب و استخدام هموار سازد.

## منابع


اعتمادی، محمد؛ چیت‌ساز، احسان؛ کوشکی، سحر و جعفری، سیدمحمدعلی (۱۴۰۳). هوش مصنوعی در مقابل روش‌های هدایت انسانی در ارزیابی استخدام منابع انسانی: فراترکیب. *فصلنامهٔ علمی مدیریت منابع انسانی پایدار*، ۶(۱)، ۱۹۱-۲۱۴.

بذرکار، ادشیر؛ مرادزاد، مهرداد و شایگان، شادی (۱۴۰۲). تحلیل اثر عوامل تکنولوژیکی، سازمانی و محیطی بر به‌کارگیری هوش مصنوعی در فرایند جذب کارکنان. *پژوهش‌های نوین در ارزیابی عملکرد*، ۲(۱)، ۵۲-۳۲.

دهقانان، حامد؛ پورامینی، زهرا؛ یزدان‌شناس، مهدی و رئیسی وانانی، ایمان (۱۴۰۴). روش‌شناسی دلفی مبتنی‌بر هوش مصنوعی: رویکردی نوین در زاویه‌بندی پژوهش‌های مدیریتی. *مدیریت دولتی*، ۱۷(۴)، ۹۶۳- ۹۹۱.

رستگار، عباسعلی؛ ابراهیمی، سیدعباس؛ شفیعی نیک‌آبادی، محسن و کلاهی، بهاره (۱۴۰۱). معماری منابع انسانی هوشمند: رویکردی ساختاری بر تحول دیجیتال شرکت‌های دانش‌بنیان. *مدیریت دولتی*، ۱۴(۲)، ۲۱۵-۲۳۴.

سرایی، سمیرا؛ صراف، فاطمه و حمیدیان، محسن (۱۴۰۳). مدل‌سازی خطاهای سوگیری بر تصمیم‌گیری مالی مدیران: یک نگرش چندسطحی. *فصلنامه تحقیقات حسابداری و حسابرسی*، ۶۱، ۴۵-۶۶.

فضل‌طلب حمیدرضا و جمشیدی، فاطمه (۱۴۰۴). منابع انسانی الگوریتمی. *نشریه علمی رویکردهای پژوهشی نوین مدیریت و حسابداری*، ۹(۳۲)، ۱۴۴۳-۱۴۴۸.

قلی‌پور، آرین (۱۴۰۳). سخن سردبیر: مدل ارزیابی تعالی مدیریت منابع انسانی در بخش دولتی. *مدیریت دولتی*، ۱۶(۲)، ۲۸۲-۳۱۵.

محمدی، محمدجواد (۱۴۰۳). کاربرد الگوریتم‌های هوش مصنوعی در بهینه‌سازی فرایند استخدام شهرداری: مطالعه موردی. *فصلنامه پژوهش‌های نوین در شهر هوشمند*، ۳(۳)، ۱۹-۴۱.

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.

## References

Albaroudi, E., Mansouri, T. & Alameer, A. (2024). A comprehensive review of AI techniques for addressing algorithmic bias in job hiring. *AI (Switzerland),* 5(1), 383-404. https://doi.org/10.3390/ai5010019.

Barocas, S. & Selbst, A.D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671-732. https://doi.org/10.15779/Z38BG31.

Bazrkar, A., Moradzad, M. & Shayegan, S. (2023). Analysis of the Effect of Technological, Organizational and Environmental Factors on the Use of Artificial Intelligence in the Recruitment Process of Employees. *Modern Research in Performance Evaluation*, 2(1), 32-52. https://doi.org/10.22105/mrpe.2023.174080. (*in Persian*)

Berthet, V. (2022). The cognitive mechanisms of human decision-making. *Rationality and Society*, 34(3), 290-323. https://doi.org/10.1177/10434631221109040

Berthet, V. (2022). The impact of cognitive biases on professionals' decision-making: A review of four occupational areas. *Frontiers in Psychology*, 12, 802439. https://doi.org/10.3389/fpsyg.2021.802439

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Breaugh, J.A. (2013). Employee recruitment. *Annual Review of Psychology*, 64, 389-416. https://doi.org/10.1146/annurev-psych-113011-143757

Chen, L., Ma, R., Hannák, A. & Wilson, C. (2018). Investigating the impact of gender on rank in resume search engines. In *R. Mandryk, M. Hancock, M. Perry & A. Cox (Eds.), Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM.https://doi.org/10.1145/3173574.3174225

Dehghanan, H., Pouramini, Z., Yazdanshenas, M. & Raeesi Vanani, I. (2025). AI-based Delphi methodology: A novel approach to triangulation in managerial research. *Management of Government*. https://jipa.ut.ac.ir/article_105475.html (*in Persian*)

Derous, E. & Ryan, A.M. (2019). When your resume is (not) turning you down: Modelling ethnic bias in resume screening. *Human Resource Management Journal*, 29(2), 113-130. https://doi.org/10.1111/1748-8583.12217.

Diefenhardt, A. (2025). Automating the managerial gaze: critical and genealogical notes on machine learning in personnel assessment. *The International Journal of Human Resource Management*, *36*(14), 2516-2549.

Etemadi, M., Chitsaz, E., Koushki, S., Jafari, S.M. (2024). Artificial Intelligence (AI) vs. Human-Led Approaches in Human Resource Recruitment Assessment: A Meta-Synthesis of Advantages and Disadvantages. *Journal of Sustainable Human Resource Management*, 6(11), 191-214. Doi: 10.22080/shrm.2024.5100. (*in Persian*)

Fazltalab, H. & Jamshidi, F. (2024). Algorithmic human resources. Journal of New Research Approaches in Management and Accounting, 96, 1443-1448. https://majournal.ir/index.php/ma/article/view/3206. (*in Persian*)

Fischer, G. (2022). A research framework focused on AI and humans instead of AI versus humans. In B.R. Barricelli, G. Fischer, D. Fogli, A. Morch, A. Piccinno & S. Valtolina (Eds.), Proceedings of CoPDA2022 - Sixth International Workshop on Cultures of Participation in the Digital Age: AI for Humans or Humans for AI? (pp. 1-8). *CEUR Workshop Proceedings*. https://ceur-ws.org/Vol-3136/paper-1.pdf

Gholipour, A. (2024). Model for evaluating the excellence of human resource management in the public sector. *Management of Government*, 16(2), 282–315. https://jipa.ut.ac.ir/article_97893.html (*in Persian*)

Herbosch, M. (2025). To err is human: Managing the risks of contracting AI systems. *Computer Law & Security Review*, 56, 106110.https://doi.org/10.1016/j.clsr.2025.106110

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, 1(3), 333-342. https://doi.org/10.1111/j.1754-9434.2008.00058.x.

Hunkenschroer, A.L. & Lütge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178(4), 977-1007. https://doi.org/10.1007/s10551-022-05049-6.

Köchling, A. & Wehner, M.C. (2020). Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, 13(3), 795-848. https://doi.org/10.1007/s40685-020-00134-w

Kordzadeh, N. & Ghasemaghaei, M. (2021). Algorithmic bias: a path to realizing AI's promised benefits? *Journal of Business Analytics*, 5(1), 104-123. https://doi.org/10.1080/2573234X.2021.1946633

Korteling, J.E. & Toet, A. (2021). *Cognitive biases. In Handbook of Human Factors and Ergonomics* (pp. 231-266). John Wiley & Sons: https://doi.org/10.1002/9781119636113.ch10

Lepri, B., Oliver, N., Letouzé, E., Pentland, A. & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology,* 31, 611-627. https://doi.org/10.1007/s13347-017-0279-x

Marinucci, L., Mazzuca, C. & Gangemi, A. (2023). Exposing implicit biases and stereotypes in human and artificial intelligence: State of the art and challenges with a focus on gender. *AI & Society*, 38(2), 747-761. https://doi.org/10.1007/s00146-022-01474-3

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys,* 54(6), 1-35. https://doi.org/10.1145/3457607

Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Ba.

Miller, J. (2022). *Inclusion at Work 2022. Chartered Institute of Personnel and Development*. https://www.cipd.org/globalassets/media/knowledge/knowledge-hub/reports/2022-pdfs/2022-inclusion-at-work-report.pdf

Mohammadi, M.J. (2024). Application of Artificial Intelligence Algorithms in Optimizing the Recruitment Process of Municipalities: A Case Study of Sadra Municipality. *Journal: New Researches in The Smart City*, 3(3), 19-41. https://www.sid.ir/paper/1696613/en. (*in Persian*)

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. *Management of Government*, 18(1), 263–295. https://jipa.ut.ac.ir/article_106486.html (*in Persian*)

Radeke, M. & Stahelski, A.J. (2020). The halo and horns effect: A review of the literature and its implications for performance appraisal. *Journal of Business and Psychology*, 35(1), 1-15.

Raghavan, M., Barocas, S., Kleinberg, J. & Levy, K. (2019). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Conference on Fairness, Accountability, and Transparency (FAT '20)* (pp. 469-481). https://doi.org/10.1145/3351095.3372828

Raghavan, M., Barocas, S., Kleinberg, J. & Levy, K. (2020). Mitigating bias in algorithmic employment screening: Evaluating claims and practices. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469-481. https://doi.org/10.1145/3351095.3372828.

Rao, S. & Zhao, T. (2025). Ethical AI in HR: A case study of tech hiring. Journal of Computer Information Systems. *Advance online publication*. https://doi.org/10.1080/08874417.2024.2446954

Rastgar, A., Ebrahimi, S. A., Shafiei Nikabadi, M. & Kalahi, B. (2022). Smart human resource architecture: A structural approach to digital transformation in knowledge-based companies. *Management of Government*, 14(2), 215–234. https://jipa.ut.ac.ir/article_88922.html (*in Persian*)

Rigotti, C. & Fosch-Villaronga, E. (2024). Algorithmic fairness in recruitment: Legal and conceptual challenges. *Computer Law & Security Review*, 53, 105966.https://doi.org/10.1016/j.clsr.2024.105966

Rigotti, J. & Fosch-Villaronga, E. (2024). Algorithmic fairness and HR recruitment: Challenges of transparency and accountability in automated decision-making. *AI and Ethics*. https://doi.org/10.1007/s43681-023-00286-7

Ryan, A.M. & Ployhart, R.E. (2014). A century of selection. *Annual Review of Psychology*, 65, 693-717. https://doi.org/10.1146/annurev-psych-010213-115134.

Saraei S., Sarraf, F., Hamidian, M. (2024). Modeling bias errors on managers' financial decision making a multi-level approach. *ISC E-journals*, 61, 45-66. https://ecc.isc.ac/showJournal/4207/284622/3585306. (*in Persian*)

Scott, S.V. & Orlikowski, W.J. (2025). Exploring AI-in-the-making: Sociomaterial genealogies of AI performativity. *Information and Organization*, 35(1), 100558.https://doi.org/10.1016/j.infoandorg.2025.100558

Shao, R. (2023). The persistence of the halo effect in sequential evaluations: A study in hiring contexts. *Organizational Behavior and Human Decision Processes*, 175, 104231. https://doi.org/10.1016/j.obhdp.2023.104231

Soleimani, M., Intezari, A., Arrowsmith, J., Pauleen, D.J. & Taskin, N. (2025). Reducing AI bias in recruitment and selection: An integrative grounded approach. *The International Journal of Human Resource Management*, 36(14), 2480-2515.https://doi.org/10.1080/09585192.2025.2480617

Sony, A.A.M., Armin, M.B., Ashraf, A., Islam, K.M.A. & Debnath, G.C. (2025). Bias in AI-driven HRM systems: Investigating discrimination risks embedded in AI recruitment tools and HR analytics. *Social Sciences & Humanities Open,* 12, 102082. https://doi.org/10.1016/j.ssaho.2025.102082

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

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Whysall, Z. (2018). AI and the future of recruitment: The role of cognitive bias. In *The Oxford Handbook of Recruitment* (pp. 315-330).

Wilkens, U., Lutzeyer, I., Zheng, C., Beser, A. & Prilla, M. (2025). Augmenting diversity in hiring decisions with artificial intelligence tools. *The International Journal of Human Resource Management*, 36(14), 2585–2622.https://doi.org/10.1080/09585192.2025.2492867

Yuan, S., Xing, L. & Zhao, D. (2025). A multi stage HR in the loop approach to enhance fairness perceptions of AI selection systems. *The International Journal of Human Resource Management*, 36(14), 2623-2658.https://doi.org/10.1080/09585192.2025.2564235