



## From Traditional to Smart Assessment Centers: A Human-in-the-Loop Framework for AI-Based Decision Support in Talent Assessment

Mir Ali Seyed Naqavi<sup>1</sup>  | Niloufar Mozafari<sup>2</sup> 

### Article Info

#### Article type:

Research Article

#### Article history:

Received 01 January 2026

Received in revised form 19

January 2026

Accepted 08 February 2026

Published online 27 February  
2026

#### Keywords:

*Smart Assessment Center,  
Digital Transformation,  
Human Resource  
Management, Artificial  
Intelligence, Professional  
Ethics.*

### Abstract


**Background and Objective:** The digital transformation and the proliferation of Artificial Intelligence (AI) in Human Resource Management (HRM) have necessitated a fundamental rethinking of Assessment Center (AC) design, leading to a gradual transition toward "Smart Assessment Centers." This study aims to elucidate this evolutionary process and develop a conceptual framework for the transformation of ACs within the context of contemporary HRM.

**Methodology:** his research employs a qualitative and exploratory approach using content analysis. Data were gathered from 42 credible scientific sources published between 2020 and 2025, selected from the Wiley and Taylor & Francis databases based on their relevance to digital transformation and HRM. The data were analyzed through three levels of coding: open, axial, and selective. The scientific rigor and trustworthiness of the findings were ensured based on Lincoln and Guba's (1985) criteria.

**Findings:** The findings indicate that the evolution of ACs can be explained in three distinct waves: the first wave involves the digitization of tools; the second wave focuses on augmented intelligence with a Human-in-the-Loop (HITL) approach; and the third wave centers on the emergence of ecosystems driven by generative models and interactive simulations. While smart ACs enhance operational efficiency and the richness of behavioral evidence, they pose significant challenges regarding validity, fairness, explainability, and legal compliance.

**Conclusion:** The synthesis of the findings suggests that Smart Assessment Centers are more than mere technical tools; they are, in fact, "socio-technical systems" in which technology serves to augment, rather than replace, professional human judgment. Consequently, the future of this field lies not in pure automation, but in responsible, ethical, and human-centric intelligence.

**Cite this article:** Seyed Naqavi, M. and Mozafari, N. (2026). From Traditional to Smart Assessment Centers: A Human-in-the-Loop Framework for AI-Based Decision Support in Talent Assessment. *Intelligent Management of Human Capital*, 2(7), 1-50

 DOI: <https://doi.org/10.22034/imhr.2026.574320.1073>

**Publisher:** Human Capital institute, Command and Staff University of I.R.I Army, <https://www.imhr.ir>

© "Authors retain the copyright and full publishing rights."

DOI: 10.22034/imhr.2026.574320.1073



1. Full Professor, Faculty of Management and Accounting, Department of Public Administration, Allameh Tabataba'i University, Tehran, Iran. E-mail: [asnaghavi@atu.ac.ir](mailto:asnaghavi@atu.ac.ir)

2. Corresponding Author, Ph.D. Student in Human Resource Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran. E-mail: [n\\_mozafari@atu.ac.ir](mailto:n_mozafari@atu.ac.ir)



## From Traditional to Smart Assessment Centers: A Human-in-the-Loop Framework for AI-Based Decision Support in Talent Assessment

### Extended Abstract

**Background and Objective:** Over the past five decades, Assessment Centers (ACs) have established themselves as the cornerstone of talent assessment, leadership development, and succession planning within Human Resource Management (HRM). Traditionally, these centers have relied heavily on structured observations, behavioral simulations, and the expert judgment of human assessors to evaluate managerial potential. While the validity and reliability of traditional methods are well-documented, the rapid digitization of organizational processes and the proliferation of Artificial Intelligence (AI) in HR ecosystems have confronted these traditional models with increasing challenges regarding scalability, cost, and efficiency. The emergence of Smart Assessment Centers (SACs) represents a profound paradigm shift, moving from observation-based evaluations to data-driven, algorithm-augmented ecosystems. This transformation aligns with the broader evolution of HRM 5.0, where cloud computing, machine learning, and digital analytics redefine the operational and strategic landscape of talent management. However, despite the promising prospects of enhanced precision and predictive power, the transition to SACs introduces complex challenges related to algorithmic bias, data privacy, legal compliance, and the "black box" nature of AI decision-making. Consequently, the primary objective of this study is to systematically elucidate this evolutionary process—from traditional to smart assessment—and to extract a comprehensive conceptual model for the transformation of ACs. The study aims to map how digital transformation redefines assessment design, execution, and governance, offering a framework that integrates technological capabilities with human-centric values and professional standards.

**Methodology:** This research was conducted using a qualitative and exploratory design, employing a directed content analysis approach to synthesize the existing body of knowledge. The study methodology was structured based on the "Research Onion" model by Saunders et al. (2019), adopting an interpretivist philosophy and an inductive approach to generate new theoretical insights from emerging phenomena. The data corpus consisted of 42 high-impact, peer-reviewed scientific articles and specialized reports published between 2020 and 2025. These sources were systematically selected from reputable academic databases, specifically Wiley and Taylor & Francis, based on rigorous inclusion criteria such as conceptual relevance to digital transformation in assessment, HRM technology adoption, and psychometric validity. The data analysis was executed through a three-level coding process: Open Coding (identifying initial semantic units and concepts), Axial Coding (grouping codes into core categories and defining relationships), and Selective Coding (abstracting categories to form the final causal-conditional-consequential model). To ensure the scientific rigor and trustworthiness of the study, Lincoln and Guba's (1985) criteria were applied, including credibility (through data triangulation), transferability (via thick description of the context), dependability (through transparent documentation of the analysis path), and confirmability.

**Findings:** The synthesis of the literature reveals that the evolution of Assessment Centers can be delineated into three distinct and progressive waves, each characterized by specific technologies, psychometric focuses, and institutional challenges.

- **The First Wave (Digitization of Tools):** This initial phase focuses on the migration of traditional assessment tools (e.g., role-plays, in-basket exercises, interviews) to digital platforms. The primary concern here is "psychometric equivalence"—ensuring that virtual execution does not compromise validity compared to face-to-face settings. Findings indicate that while digital tools can achieve comparable reliability, the focus shifts heavily towards managing "applicant reactions," "perceived justice," and the "candidate experience" as critical conditions for legitimacy.



- **The Second Wave (Augmented Intelligence):** This wave marks the entry of machine learning and semi-smart systems into the scoring process. The logic shifts from simple digitization to "data-driven behavior analysis." Technologies such as Natural Language Processing (NLP) are used to automate the scoring of open-ended responses and interpersonal exercises by analyzing micro-behavioral signals. The findings highlight a "Human-in-the-loop" model, where AI acts as a reliability support to human assessors rather than a total replacement. Key challenges in this wave include controlling for "construct drift" and ensuring that algorithmic scoring remains aligned with theoretical constructs.
- **The Third Wave (Generative Ecosystems):** The most recent and advanced wave involves the emergence of ecosystems driven by Generative AI (LLMs) and interactive agents. In this stage, ACs evolve from static measurement tools to dynamic, adaptive environments where AI not only scores but also generates scenarios and interacts with candidates in real-time. This wave offers the potential for richer, more realistic behavioral evidence but imposes strict requirements regarding "governance," "fairness auditing," and "legal compliance." The study identifies a causal model where the legitimacy of these advanced systems is contingent upon completing three loops: the Validity Loop (data-construct links), the Reliability Loop (model stability), and the Governance Loop (monitoring adverse impact).

**Conclusion:** The study concludes that Smart Assessment Centers should not be viewed merely as technological upgrades but as complex "Socio-Technical Systems." The findings argue that the future of talent assessment lies not in maximum automation or the complete replacement of human assessors, but in a "Responsible AI" framework where technology serves to reinforce and augment professional human judgment. The research implies a significant role shift for Human Resource departments: moving from operational executors of assessment tools to strategic "System Architects" and "Risk Auditors." Furthermore, the proposed conceptual model emphasizes that the legitimacy of SACs depends on a hybrid approach—integrating the scalability and pattern-recognition capabilities of AI with the explainability, ethical judgment, and contextual understanding of humans. The study recommends that organizations and policymakers prioritize the establishment of robust governance frameworks, including continuous validation mechanisms, algorithmic fairness audits, and transparency protocols. Ultimately, the successful transition to Smart Assessment Centers requires a balance where data analytics and AI are subservient to fair, explainable, and evidence-based decision-making, ensuring that the pursuit of efficiency does not compromise institutional trust or social legitimacy.

**Keywords:** *Smart Assessment Center, Digital Transformation, Human Resource Management, Artificial Intelligence, Professional Ethics.*



# از کانون‌های ارزیابی سنتی تا کانون‌های ارزیابی هوشمند: ارائه چارچوبی انسان محور در چرخه برای پشتیبانی تصمیم‌گیری مبتنی بر هوش مصنوعی در ارزیابی استعداد

میرعلی سید نقوی<sup>۱</sup> | نیلوفر مظفری<sup>۲</sup>

## چکیده

تحول دیجیتال و گسترش هوش مصنوعی در مدیریت منابع انسانی، موجب بازاندیشی در طراحی کانون‌های ارزیابی و گذار تدریجی آن‌ها به سوی کانون‌های ارزیابی هوشمند شده است. هدف پژوهش، تبیین فرایند این تحول و استخراج چارچوب مفهومی برای دگرگونی کانون‌های ارزیابی در چارچوب مدیریت منابع انسانی معاصر است. پژوهش حاضر با رویکرد کیفی و اکتشافی، با استفاده از تحلیل محتوا انجام شده است. داده‌ها از ۴۲ منبع علمی معتبر که در فاصله زمانی ۲۰۲۰ تا ۲۰۲۵ منتشر شده‌اند، گردآوری شده است. این منابع از پایگاه‌های علمی Wiley و Taylor & Francis انتخاب و بر اساس معیارهای مرتبط با تحول دیجیتال و مدیریت منابع انسانی بررسی شدند. تحلیل داده‌ها در سه سطح کدگذاری باز، محوری و انتخابی صورت گرفته است. دقت علمی و اعتمادپذیری یافته‌ها بر اساس معیارهای لینکلن و گوبا (۱۹۸۵) تضمین گردید.

یافته‌ها نشان می‌دهد که تکامل کانون‌های ارزیابی در سه موج قابل تبیین است: موج نخست، دیجیتالی‌سازی ابزارها؛ موج دوم، هوشمندسازی تقویتی با حضور انسان در حلقه تصمیم؛ و موج سوم، شکل‌گیری اکوسیستم‌های مبتنی بر مدل‌های مولد و شبیه‌سازی‌های تعاملی. اگرچه کانون‌های هوشمند موجب ارتقای کارایی و غنای شواهد رفتاری می‌شوند، اما چالش‌های جدی در حوزه‌های روابی، عدالت، تبیین‌پذیری و انطباق حقوقی ایجاد می‌کنند. در نهایت، سنتز یافته‌های پژوهش تبیین می‌کند که کانون‌های ارزیابی هوشمند فراتر از یک ابزار فنی، در واقع سامانه‌هایی «اجتماعی-فناورانه» هستند که در آن‌ها فناوری نقش مقوم قضاوت حرفه‌ای انسان را ایفا می‌کند. بر این اساس، آینده این حوزه نه در خودکارسازی محض، بلکه در گرو هوشمندسازی مسئولانه، اخلاق مدار و انسان‌محور نهفته است.

## اطلاعات مقاله

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۴/۱۰/۱۱

تاریخ بازنگری:

۱۴۰۴/۱۰/۲۹

تاریخ پذیرش:

۱۴۰۴/۱۱/۱۹

تاریخ انتشار:

۱۴۰۴/۱۲/۰۸

## کلیدواژه‌ها:

کانون ارزیابی هوشمند، تحول دیجیتال، مدیریت منابع انسانی، هوش مصنوعی، اخلاق حرفه‌ای.

**استناد:** سیدنقوی، میرعلی؛ مظفری، نیلوفر. (۱۴۰۴). از کانون‌های ارزیابی سنتی تا کانون‌های ارزیابی هوشمند: ارائه چارچوبی انسان محور در چرخه برای پشتیبانی تصمیم‌گیری مبتنی بر هوش مصنوعی در ارزیابی استعداد. مدیریت هوشمند سرمایه انسانی، ۲(۷)، ۵۰-۱.



DOI: <https://doi.org/10.22034/imhr.2026.574320.1073>

ناشر: پژوهشکده سرمایه انسانی دانشگاه فرماندهی و ستاد آجا، <https://www.imhr.ir>

© «حق نشر (کپی رایت) و کلیه حقوق انتشار برای نویسندگان محفوظ است.»



DOI: 10.22034/imhr.2026.574320.1073

۱. استاد، گروه مدیریت دولتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران، [asnaghavi@atu.ac.ir](mailto:asnaghavi@atu.ac.ir)

۲. نویسنده مسئول، دانشجوی دکتری مدیریت منابع انسانی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

رایانامه: [n\\_mozafari@atu.ac.ir](mailto:n_mozafari@atu.ac.ir)

## مقدمه

در پنج دهه گذشته، کانون‌های ارزیابی به عنوان رکن بنیادین ارزیابی استعداد، توسعه رهبری و برنامه‌ریزی جانشین‌پروری در مدیریت منابع انسانی شناخته شده‌اند. کانون‌های ارزیابی به طور سنتی برای سنجش پتانسیل مدیریتی بر شبیه‌سازی‌های رفتاری، مشاهدات ساختاریافته و قضاوت‌های تخصصی استوار بوده‌اند. اگرچه روایی و پایایی این روش‌ها به اثبات رسیده است اما دیجیتالی شدن سریع فرایندهای سازمانی و گسترش فناوری‌های هوش مصنوعی در زیست‌بوم‌های منابع انسانی آن‌ها را با چالش‌های فزاینده‌ای مواجه ساخته است (Yilmaz, 2023). ظهور کانون‌های ارزیابی هوشمند بیانگر تحولی چشمگیر است تحولی که با ادغام تحلیل‌های دیجیتال، یادگیری ماشین و سیستم‌های مبتنی بر رایانش ابری، موجب ارتقای دقت، مقیاس‌پذیری و تجربه کاربری در ارزیابی استعدادها می‌شود.

یکپارچه‌سازی هوش مصنوعی با مدیریت منابع انسانی، از طریق فعال‌سازی تحلیل‌های پیش‌بینی‌کننده، برنامه‌ریزی هوشمند نیروی کار و ارزیابی‌های خودکار، تصمیم‌گیری‌های سنتی را دگرگون کرده است (Pandey & Mishra, 2025). این فناوری‌ها ضمن تسهیل فرایند ارزیابی نامزدها، شایستگی‌های پنهان و روندهای رفتاری را از طریق بینش‌های داده‌محور آشکار می‌سازند. کانون‌های ارزیابی هوشمند با بهره‌گیری از هوش مصنوعی، سوگیری‌های انسانی را به حداقل رسانده، ثبات در سنجش عملکرد را بهبود بخشیده و حلقه‌های بازخورد آنی را تسهیل می‌کنند. پیشرفت‌هایی که دستیابی به آن‌ها برای کانون‌های سنتی دشوار می‌نماید (Nalini, 2024). این دگرگونی همسو با تکامل دیجیتال گسترده‌تر در منابع انسانی است، جایی که رایانش ابری و هوش مصنوعی مشترکاً چشم‌انداز عملیاتی و استراتژیک مدیریت استعداد را بازتعریف می‌کنند.

با وجود چشم‌انداز امیدوارکننده کانون‌های ارزیابی هوشمند، به‌کارگیری آن‌ها چالش‌های پیچیده‌ای را به همراه دارد. نگرانی‌های اخلاقی پیرامون حریم خصوصی داده‌ها، سوگیری‌های الگوریتمی و شفافیت، به طور فزاینده‌ای برجسته شده‌اند (Purba, 2024). افزون بر این، گذار به این سیستم‌ها نیازمند ارتقای اساسی مهارت‌های متخصصان منابع انسانی، بازآرایی ساختارهای سازمانی و همسوسازی سیستم‌های هوش مصنوعی با ارزش‌های انسان‌محور است (Khan & Khan, 2024). این تحولات نشان‌دهنده تغییر پارادایم از ارزیابی‌های مبتنی بر مشاهده‌گر سنتی به زیست‌بوم‌های ارزیابی تقویت‌شده با الگوریتم است.

بر این اساس، نوآوری کلیدی پژوهش حاضر، گذر از رویکردهای توصیفی پراکنده و ارائه یک چارچوب مفهومی انسان‌محور در چرخه برای کانون‌های ارزیابی است. برخلاف مطالعات پیشین که غالباً فناوری را به مثابه ابزاری صرفاً خودکارساز و جانشین عاملیت انسانی می‌پندارند، این پژوهش با صورت‌بندی تکامل این کانون‌ها در قالب یک گونه‌شناسی سه‌موجی متمایز، کانون‌های ارزیابی هوشمند را به عنوان

سیستم‌هایی اجتماعی-فناورانه بازتعریف می‌کند. در این چارچوب نوآورانه، هوش مصنوعی نه جایگزینی برای ارزیاب، بلکه بازوی توانمندساز و مقوم ساختاری قضاوت حرفه‌ای وی تبیین می‌گردد. در نهایت، این مقاله با تلفیق سنت‌های ارزیابی رفتاری و ظرفیت‌های تحلیل هوشمند، نقشه راهی نظری و کاربردی برای گذار سازمان‌ها به سیستم‌های ارزیابی عادلانه، انطباق‌پذیر و مبتنی بر هم‌افزایی هوش مصنوعی و انسان ارائه می‌دهد.

## مبانی نظری پژوهش

### تحول نظری از مراکز ارزیابی سنتی به مراکز ارزیابی هوشمند

#### ۱. بنیان‌های رفتاری و روان‌سنجی مراکز ارزیابی سنتی

گذار از مراکز ارزیابی سنتی<sup>۱</sup> به مراکز ارزیابی هوشمند<sup>۲</sup> نشان‌دهنده یک دگرگونی پارادایمی در منطق نظری، طراحی روش‌شناختی و معماری تصمیم‌گیری ارزیابی استعداد در مدیریت منابع انسانی است. مراکز ارزیابی کلاسیک به‌عنوان رویکردهایی ساختارمند، مبتنی بر رفتار و مشاهده، بر شبیه‌سازی‌های استاندارد، ارزیابی‌های چندمنبعی و قضاوت تخصصی ارزیابان انسانی استوار بوده‌اند. بنیان نظری این رویکرد بر نظریه ثبات رفتاری و الگوهای روانی ملاک‌محور و رویکرد شایستگی‌محور مبتنی است؛ نظریه‌هایی که فرض می‌کنند رفتار مشاهده‌شده در موقعیت‌های شبیه‌سازی‌شده می‌تواند عملکرد آینده فرد را در محیط واقعی سازمان پیش‌بینی کند. این چارچوب نظری، مشروعیت علمی مراکز ارزیابی را به‌عنوان ابزارهایی معتبر، پایا و عادلانه برای سنجش شایستگی‌های مدیریتی تثبیت کرده است (Yilmaz, 2023).

با این حال، رشد فناوری‌های داده‌محور، نفوذ هوش مصنوعی و گسترش زیرساخت‌های دیجیتال، بنیان‌های معرفت‌شناختی ارزیابی انسانی را با چالش مواجه ساخته و منجر به بازتعریف مفاهیم بنیادین در پیش‌بینی عملکرد، قضاوت تخصصی و تصمیم‌گیری سازمانی شده است. این تحول، ارزیابی استعداد را از یک فرایند صرفاً مشاهده‌محور به یک نظام داده‌محور، الگوریتم‌پایه و تقویت‌شده با فناوری منتقل کرده است.

#### ۲. نظریه سیستم‌های اجتماعی-فناورانه و منطق هوش ترکیبی

هسته نظری تحول به سوی SACها در نظریه سیستم‌های اجتماعی-فناورانه<sup>۳</sup> قرار دارد. این نظریه بیان می‌کند که عملکرد بهینه سازمانی تنها زمانی حاصل می‌شود که نظام‌های انسانی و فناورانه

<sup>1</sup> Assessment Centers: ACs

<sup>2</sup> Smart Assessment Centers: SACs

<sup>3</sup> Socio-Technical Systems Theory

به صورت هم‌زمان و هم‌تراز تکامل یابند. در زمینه مراکز ارزیابی، این رویکرد بر ضرورت هم‌افزایی میان قضاوت انسانی و دقت الگوریتمی تأکید دارد، نه جایگزینی یکی با دیگری (Cherep, Voronkova, & Androsova, 2022).

در این چارچوب، مراکز ارزیابی هوشمند به‌عنوان اکوسیستم‌های اجتماعی-فناورانه مفهوم‌پردازی می‌شوند که در آن‌ها پلتفرم‌های دیجیتال، الگوریتم‌های یادگیری ماشینی و زیرساخت‌های ابری، عملکرد شناختی ارزیابان انسانی را تقویت می‌کنند. این تعامل نظری با مفهوم هوش ترکیبی<sup>۱</sup> هم‌راستا است؛ مدلی که شهود انسانی و استدلال محاسباتی را در یک نظام تصمیم‌گیری یکپارچه می‌سازد (Raime et al., 2025). در این الگو، هوش مصنوعی نه به‌عنوان جایگزین ارزیاب، بلکه به‌عنوان توانمندساز شناختی عمل می‌کند که ظرفیت تحلیل الگوهای پیچیده رفتاری، نشانه‌های ارتباطی، شاخص‌های زبان‌شناختی روانی و داده‌های چندوجهی را فراهم می‌سازد؛ ظرفیتی که فراتر از توان پردازشی انسان قرار دارد.

### ۳. تحول دیجیتال، HRM 5.0 و منطق راهبردی SACها

نظریه تحول دیجیتال در مدیریت منابع انسانی، SACها را در چارچوب گذار گسترده‌تر سازمان‌ها به سوی پارادایم‌های صنعت ۴.۰ و HRM ۵.۰ تبیین می‌کند. HRM ۵.۰ به‌عنوان نظامی اجتماعی-فناورانه و انسان‌محور تعریف می‌شود که نوآوری دیجیتال را با ارزش‌های اخلاقی، شمول‌گرایی و رفاه کارکنان تلفیق می‌کند (Kravchuk, 2025). در این بستر، مراکز ارزیابی هوشمند تجسمی از این هم‌افزایی هستند زیرا با بهره‌گیری از هوش مصنوعی، تحلیل داده و زیرساخت‌های ابری، چارچوبی نوین برای ارزیابی استعداد فراهم می‌آورند که هم‌زمان دقت تحلیلی و تجربه انسانی داوطلبان را ارتقا می‌دهد. افزون بر این، دیدگاه منبع‌محور سازمان امکان تحلیل SACها به‌عنوان قابلیت راهبردی را فراهم می‌سازد. قابلیت‌هایی نظیر تحلیل‌های پیش‌بین، الگوریتم‌های تطبیقی استعداد و سامانه‌های بازخورد هوشمند می‌توانند به منابع کمیاب، تقلیدناپذیر و ارزش‌آفرین تبدیل شوند که مزیت رقابتی پایدار ایجاد می‌کنند. از این منظر، سازمان‌هایی که به‌طور مؤثر SACها را پیاده‌سازی می‌کنند، ظرفیت ارزیابی استعداد خود را از یک فعالیت عملیاتی به یک شایستگی محوری راهبردی ارتقا می‌دهند. این منطق راهبردی در هم‌راستایی با معماری کاربست فناوری‌های هم‌گرا در فرآیند جذب و گزینش است که بر اساس آن، ابزارهای هوشمند دیجیتال فراتر از رویکردهای ابزاری و سنتی، بنیانی برای بازآفرینی الگوهای کلان مدیریت سرمایه انسانی فراهم می‌آورند (گلشاهی و خوشنویس زارچ، ۱۴۰۴).

<sup>۱</sup> هوش ترکیبی (Hybrid Intelligence): هوش ترکیبی به مدلی اطلاق می‌شود که در آن توانمندی‌های مکمل انسان (مانند درک زمینه، قضاوت اخلاقی، شهود) و ظرفیت‌های هوش مصنوعی (مانند پردازش حجم بالا، تشخیص الگوهای پیچیده، خستگی‌ناپذیری) در یک نظام تصمیم‌گیری یکپارچه و مشارکتی گرد هم می‌آیند

#### ۴. داده‌محوری، تصمیم‌گیری الگوریتمی و بازتعریف روایی پیش‌بین

نظریه داده‌محور شدن<sup>۱</sup> و پارادایم تصمیم‌گیری الگوریتمی تبیین می‌کنند که چگونه کنش‌ها و رفتارهای انسانی که پیش‌تر عمدتاً کیفی و تفسیری قلمداد می‌شدند، اکنون از طریق سیستم‌های هوش مصنوعی و کلان‌داده به اطلاعات کمی، ساختاریافته و قابل تحلیل تبدیل می‌شود. سیستم‌های مدیریت استعداد الگوریتم‌پایه، ظرفیت منحصربه‌فردی در تبدیل این کلان‌داده‌های رفتاری به شاخص‌های بسط‌یافته و پیش‌بینی‌کننده عملکرد و پتانسیل شغلی دارند؛ امری که موجبات ارتقای روایی پیش‌بین<sup>۲</sup> ارزیابی‌ها را نسبت به رویکردهای سنتی فراهم می‌آورد (Faqihi & Miah, 2023).

در همین راستا، شواهد پژوهشی در ادبیات این حوزه نشان می‌دهد که گذار به سوی تصمیم‌گیری عینی و عادلانه، مستلزم عبور از تحلیل‌های تک‌بعدی است؛ به‌طوری‌که بهره‌گیری از سامانه‌های هوشمند چندوجهی<sup>۳</sup> که به‌طور هم‌زمان سیگنال‌های کلامی را از طریق پردازش زبان طبیعی و نشانه‌های غیرکلامی را از طریق بینایی ماشین و تحلیل حرکتی چهره رمزگشایی می‌کنند می‌تواند با تلفیق این لایه‌های چندبعدی اطلاعاتی، مدل‌سازی روان‌شناختی دقیق‌تر و عاری از سوگیری از متقاضیان ارائه دهد (میرزانژاد و حاجی مولانا، ۱۴۰۴).

بر این اساس، کانون‌های ارزیابی هوشمند با تعبیه مدل‌های یادگیری ماشین تطبیقی، بستر لازم را برای شناسایی الگوهای رفتاری پیچیده، سنجش شایستگی‌های چندبعدی و ارائه تحلیل‌های بلادرنگ را فراهم می‌سازند. این منطق ساختاری با مفهوم مدیریت منابع انسانی تقویت‌شده با هوش مصنوعی هم‌گرا است؛ مدلی پویا که در آن، سیستم‌های هوشمند از طریق حلقه‌های بازخورد مداوم و یادگیری مستمر، معیارهای ارزیابی را به‌صورت ارگانیک اصلاح کرده و از این طریق، ضمن پیشینه‌سازی دقت پیش‌بینی، خطاهای شناختی و سوگیری‌های اسنادی ارزیابان انسانی را به کمترین میزان ممکن کاهش می‌دهند (Panigrahi, 2025).

#### ۵. پذیرش فناوری، مشروعیت ادراکی و آمادگی سازمانی

نظریه یکپارچه پذیرش و استفاده از فناوری<sup>۴</sup> چارچوبی رفتاری برای تبیین پذیرش و اعتماد متخصصان منابع انسانی به SACها فراهم می‌کند. پژوهش‌های تجربی نشان می‌دهند که ادراک سودمندی، سهولت استفاده، نفوذ اجتماعی و شرایط تسهیل‌کننده، نقش تعیین‌کننده‌ای در پذیرش فناوری‌های دیجیتال در حوزه HR ایفا می‌کنند (Theres & Strohmeier, 2023). در زمینه مراکز ارزیابی هوشمند، پذیرش سازمانی تنها به کارایی فنی سیستم‌ها وابسته نیست، بلکه به میزان قابلیت تفسیر،

<sup>1</sup> Datafication

<sup>2</sup> Predictive Validity

<sup>3</sup> Multimodal Intelligent Systems

<sup>4</sup> Unified Theory of Acceptance and Use of Technology (UTAUT)

شفافیت و مشروعیت ادراک‌شده تصمیم‌های الگوریتمی نیز بستگی دارد. از این منظر، اعتماد نهادی و مقبولیت اجتماعی، پیش‌شرط‌های اساسی موفقیت SACها به‌شمار می‌آیند.

### ۶. بنیان‌های انسان‌محور و اخلاقی SACها

نظریه مدیریت انسان‌محور و اصول اخلاق دیجیتال، موازنه‌ای هنجاری در برابر گرایش‌های فناورانه افراطی فراهم می‌آورد. بر اساس این رویکرد، تحول دیجیتال در HR باید در خدمت کرامت انسانی، خودمختاری فردی و توانمندسازی کارکنان قرار گیرد (Cherep et al., 2022). در SACها، این منطق از طریق به‌کارگیری اصل انسان در چرخه<sup>۱</sup> تحقق می‌یابد؛ به‌گونه‌ای که الگوریتم‌ها وظیفه پردازش داده‌ها و تولید پیشنهادهای تحلیلی را بر عهده دارند، اما تفسیر نهایی و تصمیم‌گیری مسئولانه همچنان در اختیار ارزیابان انسانی باقی می‌ماند (Makedon & Myachin, 2025).

این رویکرد از مفهوم هوش افزوده<sup>۲</sup> حمایت می‌کند و بر این اصل استوار است که فناوری باید نقش شریک مکمل را ایفا کند، نه جایگزین تصمیم‌گیر انسانی. در این چارچوب، عدالت الگوریتمی، شفافیت مدل‌ها و حفاظت از حریم خصوصی داده‌ها به مؤلفه‌های بنیادین طراحی SACها تبدیل می‌شوند (Adabi, 2025).

در مجموع، مبانی نظری تحول از مراکز ارزیابی سنتی به مراکز ارزیابی هوشمند مبتنی بر هم‌نشینی چند جریان نظری مکمل است: نظریه‌های رفتاری کلاسیک، سیستم‌های اجتماعی-فناورانه، تحول دیجیتال و HRM 5.0، داده‌محوری و تصمیم‌گیری الگوریتمی، نظریه پذیرش فناوری و رویکردهای انسان‌محور و اخلاقی. این هم‌گرایی نظری امکان طراحی چارچوبی مفهومی را فراهم می‌آورد که در آن SACها به‌عنوان نظام‌هایی ترکیبی، تطبیق‌پذیر و اخلاق‌محور تبیین می‌شوند.

بر این اساس، پژوهش حاضر SACها را به‌مثابه اکوسیستم‌هایی اجتماعی-فناورانه در نظر می‌گیرد که از طریق هم‌افزایی قضاوت انسانی و تحلیل‌های هوش مصنوعی، دقت پیش‌بینی، عدالت ارزیابی و کارایی عملیاتی را به‌طور هم‌زمان ارتقا می‌دهند. چارچوب مفهومی پیشنهادی این پژوهش، مبنایی تحلیلی برای تبیین ابعاد طراحی، منطق تصمیم‌گیری و الزامات اخلاقی کانون‌های ارزیابی هوشمند فراهم ساخته و مسیر گذار نظام‌مند سازمان‌ها از الگوهای سنتی به سامانه‌های ارزیابی هوشمند مبتنی بر شواهد را ترسیم می‌کند.

### روش‌شناسی پژوهش

پژوهش حاضر با به‌کارگیری طرح تحقیق کیفی و اکتشافی به بررسی فرایند گذار از مراکز ارزیابی

<sup>1</sup> Human-in-the-Loop

<sup>2</sup> Augmented Intelligence

سننتی (ACs) به مراکز ارزیابی هوشمند (SACs) در چارچوب مدیریت منابع انسانی معاصر می‌پردازد. روش‌شناسی این پژوهش بر اساس مدل پیاز پژوهش ساندروز و همکاران (۲۰۱۹) طراحی شده است که رویکردی نظام‌مند و لایه‌ای برای طراحی و اجرای تحقیق ارائه می‌دهد. شمای کلی و لایه‌های این رویکرد در شکل ۱ ارائه شده است.

در لایه بیرونی مدل پیاز پژوهش، فلسفه حاکم بر این مطالعه تفسیرگرایی است. این رویکرد بر این باور است که تحولات سازمانی از جمله گذار از مراکز ارزیابی سننتی به مراکز ارزیابی هوشمند پدیده‌هایی اجتماعی و وابسته به بافت هستند. در لایه دوم، پژوهش از رویکرد استقرایی بهره گرفته است. رویکرد استقرایی برای بررسی پدیده‌های نوظهوری که نظریه‌های موجود قادر به تبیین کامل آن‌ها نیستند، مناسب است. در این پژوهش، هدف آزمون فرضیه‌های ازپیش تعیین شده نیست، بلکه تولید بینش‌های مفهومی و تدوین مدلی نظری برای تبیین تحول از AC‌های سننتی به SAC‌هاست.

پژوهش حاضر از استراتژی کیفی تحلیل محتوا بهره می‌برد. داده‌های پژوهش از منابع شامل مقالات علمی داوری شده، گزارش‌های تخصصی منتشر شده با در نظر گرفتن افق زمانی مقطعی فاصله سال‌های ۲۰۲۰ تا ۲۰۲۵ گردآوری شده است. این انتخاب از آن جهت اتخاذ شده است که پدیده SAC هنوز نوظهور است و هرچند ادبیات علمی و حرفه‌ای غنی در این زمینه وجود دارد، پژوهش‌های میدانی محدودی انجام شده است.

داده‌ها به صورت نظام‌مند از پایگاه‌های علمی معتبر مانند Wiley و Taylor & Francis گردآوری شده‌اند. معیار انتخاب منابع شامل ارتباط محتوایی با تحول دیجیتال در ارزیابی و HRM، تازگی انتشار و ارزش مفهومی در تبیین پذیرش فناوری در سیستم‌های HR بوده است.

تحلیل داده‌ها با استفاده از تحلیل محتوای استقرایی کیفی طبق مراحل پیشنهادی مایرینگ (۲۰۱۴) انجام گرفت:

آشنایی با متن از طریق مطالعه مکرر، شناسایی واحدهای معنایی و کدهای اولیه، گروه‌بندی کدها در طبقات مفهومی بالاتر و انتزاع یافته‌ها برای تولید بینش‌های نظری.



شکل ۱. نگاهت لایه‌های روش‌شناسی پژوهش مبتنی بر پیاز پژوهش

به‌منظور تضمین دقت علمی و کیفیت یافته‌های کیفی، از چارچوب معیارهای چهارگانه لینکلن و گوبا (۱۹۸۵) استفاده شد که جزئیات اجرایی آن به شرح زیر است:

- **قابلیت اعتبار:** جهت اطمینان از انطباق یافته‌ها با واقعیت متون تحلیل شده، از راهبرد مثلث‌سازی منابع استفاده گردید. در این راستا، داده‌ها نه تنها از مقالات اصیل پژوهشی، بلکه از گزارش‌های راهبردی و اسناد کنفرانسی در پایگاه‌های متنوع (Taylor & Wiley و Francis) استخراج شدند تا جامعیت ابعاد تحول دیجیتال در کانون‌های ارزیابی تضمین شود.

- **قابلیت انتقال:** پژوهشگران با تشریح دقیق ویژگی‌های ۴۲ منبع منتخب، بازه زمانی (۲۰۲۵-۲۰۲۰) و معیارهای ورود و خروج، بستر لازم را برای داوران و سایر محققان فراهم آورد تا بتوانند میزان تناسب یافته‌ها با بافت‌های سازمانی مشابه را ارزیابی کنند.

- **قابلیت اطمینان/وابستگی:** جهت پایش ثبات در تحلیل‌ها، تمامی مراحل کدگذاری از شناسایی واحدهای معنایی تا انتزاع تم‌های نهایی به صورت گام‌به‌گام مستندسازی شد.

- **قابلیت تأیید:** برای به حداقل رساندن سوگیری‌های ذهنی پژوهشگر و اطمینان از اینکه یافته‌ها برخاسته از داده‌ها هستند و نه پیش‌فرض‌های فردی، فرایند بازبینی همتایان<sup>۱</sup> به کار گرفته شد. در این مرحله، کدهای اولیه و ساختار طبقات محوری با دو تن از متخصصان مدیریت منابع انسانی به اشتراک گذاشته شد و پس از دریافت نظرات اصلاحی، مدل نهایی تعدیل و تأیید گردید.

پیکره نهایی پژوهش شامل ۴۲ سند علمی معتبر که پس از غربالگری نظام‌مند به‌عنوان واحدهای اصلی تحلیل انتخاب شده و مبنای استخراج کدها و مضامین قرار گرفته‌اند، به تفکیک در جدول ۱ ارائه شده است.

جدول ۱. منابع

ردیف	منابع
۱	Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2020). Personnel selection in the digital age: A review of validity and applicant reactions, and future research challenges. <i>European Journal of work and organizational psychology</i> , 29(1), 64-77.
۲	Alrababah, S. A., & Molnár, G. (2021). The evolution of technology-based assessment: Past, present, and future. <i>International Journal of Learning Technology</i> , 16(2), 134-157.
۳	Liou, G., Bonner, C. V., & Tay, L. (2022). A psychometric view of technology-based assessments. <i>International Journal of Testing</i> , 22(3-4), 216-242.
۴	Ariesthiawati, J. (2022, October). The effectiveness of virtual assessment center platform. In <i>2022 international conference on assessment and learning (ICAL)</i> , 1-3.
۵	Avni, E., & Luria, G. (2022). Validity and reliability of virtual versus face-to-face assessment center. In <i>Academy of Management Proceedings</i> (Vol. 2022, No. 1, p. 14415). Briarcliff Manor, NY 10510: Academy of Management.
۶	Usmani, Y. Z., Petruzzello, G., Rizzo, B., & Mariani, M. G. (2025). Assessment centers in the virtual age: validity and fairness in gender and age. <i>Central European Management Journal</i> .
۷	Avni, E., & Luria, G. (2025). Virtual assessment centres versus face-to-face assessment centres: psychometric properties. <i>European Journal of Work and Organizational Psychology</i> , 34(3), 339-347.
۸	Baranyi, V. (2025). Systematic literature review on the digital transformation of the personnel selection process. <i>German Journal of Human Resource Management</i> , 23970022251363012.
۹	Sarumathi, S., Gowthaman, R. L., Sabareesh, M. B., & Sultana, A. (2025, February). AI-Enhanced HR Interview Simulation for Realistic Candidate Assessment. In <i>2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)</i> (pp. 1089-1095). IEEE. 10.1109/IDCIOT64235.2025.10915193
۱۰	Rajeswari, A. M., Ahamed, R., & Vasanth, P. (2025, February). InterviewIQ: AI-Powered Chatbot for Upskilling Candidates in Technical and HR Interviews. In <i>2025 International Conference on Intelligent Control, Computing and Communications (IC3)</i> (pp. 392-397). IEEE. 10.1109/IC363308.2025.10956589
۱۱	Saputra, I., Kurniawan, A., Yanita, M., Putri, E. Y., & Mahniza, M. (2024). The evolution of educational assessment: How artificial intelligence is shaping the trends and future of learning evaluation. <i>The Indonesian Journal of Computer Science</i> , 13(6).

<sup>1</sup> Peer Debriefing

۱۲	Chang, K., Carrim, N. M., Gupta, M., Cheng, K., & Sandland, S. (2024). Digitalisation of Personnel Recruitment and Selection. In <i>HRM 5.0: Unpacking the Digitalisation of Human Resource Management</i> (pp. 87-111). Cham: Springer Nature Switzerland.
۱۳	Mihaljević, H., Müller, I., Dill, K., Yollu-Tok, A., & von Grafenstein, M. (2024). More or less discrimination? Practical feasibility of fairness auditing of technologies for personnel selection. <i>AI &amp; SOCIETY</i> , 39(5), 2507-2523.
۱۴	Simelane, R. M., & Pillay, P. (2024). The Evolution of Assessment Methods in Higher Education Due to the Shift to Remote Learning: A Case Study. <i>Research in Educational Policy and Management</i> , 6(2), 393-414.
۱۵	Lumdee, N. (2025, April). Needs Analysis for Developing Intelligent Digital Platform for Course Learning Outcome Assessment of Information Technology Program in Higher Education. In <i>2025 IEEE International Conference on Cybernetics and Innovations (ICCI)</i> (pp. 1-6). IEEE.
۱۶	Guenole, N., Svensson, C., Wille, B., Aloyan, K., & Saville, P. (2022). A European perspective on psychometric measurement technology. <i>Technology and Measurement around the Globe</i> .
۱۷	Thompson, I., Koenig, N., Mracek, D. L., & Tonidandel, S. (2023). Deep learning in employee selection: Evaluation of algorithms to automate the scoring of open-ended assessments. <i>Journal of Business and Psychology</i> , 38(3), 509-527. (17 Thompson et al, 2023, Pos. 1)
۱۸	Hickman, L., Herde, C. N., Lievens, F., & Tay, L. (2023). Automatic scoring of speeded interpersonal assessment center exercises via machine learning: Initial psychometric evidence and practical guidelines. <i>International Journal of Selection and Assessment</i> , 31(2), 225-239.
۱۹	Koenig, N., Tonidandel, S., Thompson, I., Albritton, B., Koohifar, F., Yankov, G., ... & Newton, C. (2023). Improving measurement and prediction in personnel selection through the application of machine learning. <i>Personnel Psychology</i> , 76(4), 1061-1123.
۲۰	Breil, S. M., Lievens, F., Forthmann, B., & Back, M. D. (2023). Interpersonal behavior in assessment center role-play exercises: Investigating structure, consistency, and effectiveness. <i>Personnel Psychology</i> , 76(3), 759-795. (20 Breil et al. 2023, Pos. 1)
۲۱	Campion, M. A., & Campion, E. D. (2023). Machine learning applications to personnel selection: Current illustrations, lessons learned, and future research. <i>Personnel Psychology</i> , 76(4), 993-1009. (21 Campion & Campion, 2023, Pos. 1)
۲۲	Grunenberg, E., Stachl, C., Breil, S. M., Schäpers, P., & Back, M. D. (2025). Predicting and Explaining Assessment Center Judgments: A Cross-Validated Behavioral Approach to Performance Judgments in Interpersonal Assessment Center Exercises. <i>Human Resource Management</i> , 64(2), 423-445.
۲۳	Campion, M. A. (2025). Can Legal and Professional Personnel Selection Principles be Met With Machine Learning (Artificial Intelligence)?. <i>Human Resource Management</i> .
۲۴	Pathak, G., & Pandey, D. (2025). AI Agents in Recruitment: A Multi-Agent System for Interview, Evaluation, and Candidate Scoring. <i>Evaluation, and Candidate Scoring</i> (May 01, 2025).
۲۵	Chakraborty, T., & Sharada, V. S. (2025). Optimizing Talent Acquisition: The Synergy of AI and Human Expertise in Creating Enhanced Candidate Experience. In <i>Modern Trends and Future Innovation in Human Resource Management</i> (pp. 259-276). IGI Global.
۲۶	Grunenberg, E., Peters, H., Francis, M. J., Back, M. D., & Matz, S. C. (2024). Machine learning in recruiting: predicting personality from CVs and short text responses. <i>Frontiers in Social Psychology</i> , 1, 1290295.
۲۷	Rezzani, A., Caputo, A., & Cortese, C. G. (2020). An analysis of the literature about the application of Artificial Intelligence to the Recruitment and Personnel Selection. <i>Bollettino di Psicologia Applicata</i> , 25-33.

۲۸	Tippins, N. T., Oswald, F. L., & McPhail, S. M. (2021). Scientific, legal, and ethical concerns about AI-based personnel selection tools: a call to action. <i>Personnel Assessment and Decisions</i> , 7(2), 1.
۲۹	Bronkhorst, P. V., & Becker, J. (2024). Use of artificial intelligence in leadership competency development and selection: An empirical study. <i>Consulting Psychology Journal</i> .
۳۰	Tuitt, T. A., Addison, L. M., & Hosein, P. (2025, February). Generative AI and Multi-Agent Systems Approach to Psychometric Evaluation for Human Resource Management and Talent Acquisition. In 2025 7th International Symposium on Computational and Business Intelligence (ISCBI) (pp. 108-112). IEEE.
۳۱	Dorsey, D. W., & Michaels, H. R. (2022). Validity arguments meet artificial intelligence in innovative educational assessment. <i>Journal of Educational Measurement</i> , 59(3), 267-271.
۳۲	Raza, H. (2024). Ai-driven assessment: Reliability, bias, and ethical implications. <i>AI EDIFY Journal</i> , 1(2), 36-47.
۳۳	Burgues, M., Goujet, R., & Zarak, J. (2024). Learning Soft Skills with an AI-Based Simulation Role-Play: A Literature Review. <i>EDULEARN24 Proceedings</i> , 6285-6293.
۳۴	Kaldaras, L., Akaeze, H. O., & Reckase, M. D. (2024, August). Developing valid assessments in the era of generative artificial intelligence. In <i>Frontiers in education</i> (Vol. 9, p. 1399377). Frontiers Media SA.
۳۵	Grunenberg, E., Stachl, C., Breil, S. M., Schäpers, P., & Back, M. D. (2025). Predicting and Explaining Assessment Center Judgments: A Cross-Validated Behavioral Approach to Performance Judgments in Interpersonal Assessment Center Exercises. <i>Human Resource Management</i> , 64(2), 423-445.
۳۶	Zhang, T., Koutsoumpis, A., Oostrom, J. K., Holtrop, D., Ghassemi, S., & De Vries, R. E. (2024). Can large language models assess personality from asynchronous video interviews? A comprehensive evaluation of validity, reliability, fairness, and rating patterns. <i>IEEE Transactions on Affective Computing</i> , 15(3), 1769-1785.
۳۷	Hickman, L., Saef, R., Ng, V., Woo, S. E., Tay, L., & Bosch, N. (2024). Developing and evaluating language-based machine learning algorithms for inferring applicant personality in video interviews. <i>Human Resource Management Journal</i> , 34(2), 255-274.
۳۸	Modi, T. B. (2023). Artificial Intelligence Ethics and Fairness: A study to address bias and fairness issues in AI systems, and the ethical implications of AI applications. <i>Revista Review Index Journal of Multidisciplinary</i> , 3(2), 24-35.
۳۹	Anthamatten, A., & Holt, J. E. (2024). Integrating artificial intelligence into virtual simulations to develop entrustable professional activities. <i>The Journal for Nurse Practitioners</i> , 20(9), 105192.
۴۰	Behrend, T. S., & Landers, R. N. (2025). Participant Interactions with Artificial Intelligence: Using Large Language Models to Generate Research Materials for Surveys and Experiments. <i>Journal of Business and Psychology</i> , 1-23.
۴۱	Ringelband, O., & Warneke, C. (2025). Some ethical and legal issues in using artificial intelligence in personnel selection. <i>Consulting Psychology Journal</i> .
۴۲	Weinert, S., Günther, E., Rüger-Muck, E., & Raab, G. (2020). ARTIFICIAL INTELLIGENCE IN PERSONNEL SELECTION AND ITS INFLUENCE ON EMPLOYER ATTRACTIVENESS. <i>Marketing Science &amp; Inspirations</i> , 15(3).

## یافته‌های پژوهش

سنتز مفهومی ادبیات پژوهش در حوزه هوش مصنوعی، ارزیابی و انتخاب نیروی انسانی

۱. بازتعریف انتخاب نیروی انسانی در بستر تحول دیجیتال و هوش مصنوعی

مرور جامع ادبیات علمی نشان می‌دهد که فرایند انتخاب و ارزیابی نیروی انسانی در دهه اخیر، از

یک تغییر تدریجی فناوری عبور کرده و وارد مرحله‌ای از تحول ساختاری و مفهومی شده است. این تحول صرفاً به معنای جایگزینی ابزارهای کاغذی با نسخه‌های دیجیتال یا انتقال مصاحبه‌ها به بسترهای آنلاین نیست، بلکه بیانگر دگرگونی در منطق قضاوت، ماهیت شواهد ارزیابی، نقش ارزیاب انسانی و مرزهای مسئولیت حرفه‌ای در تصمیمات استخدامی است (Woods et al., 2020; Chang et al., 2024). در این چارچوب، انتخاب نیروی انسانی دیگر یک فرایند صرفاً فنی یا اداری تلقی نمی‌شود، بلکه به یک سازوکار پیچیده داده‌محور تبدیل شده است که پیامدهای راهبردی، اخلاقی، حقوقی و اجتماعی گسترده‌ای دارد.

تحلیل مفهومی بدنه ۴۲ منبع علمی گردآوری شده نشان می‌دهد که این تحول را می‌توان در قالب سه موج تکاملی مرتبط اما متمایز صورت‌بندی کرد. هر یک از این موج‌ها نه تنها از نظر فناوری‌های مورد استفاده، بلکه از حیث نوع داده‌های ارزیابی، سطح خودمختاری الگوریتم‌ها، و میزان چالش‌های اخلاقی و نهادی تفاوت‌های اساسی با موج‌های پیشین دارند. در واقع، ادبیات موجود نشان می‌دهد که حرکت از دیجیتالی‌سازی به هوشمندسازی و سپس به ارزیابی‌های مولد و تعاملی، هم‌زمان با افزایش قدرت تحلیلی سیستم‌ها، موجب افزایش حساسیت نسبت به موضوعاتی نظیر انصاف، شفافیت و مشروعیت اجتماعی شده است (Tippins et al., 2021; Ringelband & Warneke, 2025).

### موج نخست: دیجیتالی‌سازی ابزارهای سنتی و تثبیت دغدغه‌های روایی و واکنش داوطلبان

یافته‌های این پژوهش در موج نخست نشان می‌دهد که گذار از کانون‌های ارزیابی سنتی به کانون‌های ارزیابی دیجیتال/مجازی در گام اول، بیش از آنکه به معنای تحول در منطق سنجش و قضاوت باشد، ناظر بر تغییر بستر اجرا و دیجیتالی‌کردن رویه‌های موجود است. در این مرحله، معماری کلاسیک کانون ارزیابی شامل تمرین‌های شبیه‌سازی (role-play)، تمرین‌های گروهی، مطالعه موردی، آزمون‌های روان‌سنجی و مصاحبه‌های ساختاریافته محیط‌های آنلاین منتقل می‌شود اما فرض اصلی همچنان پابرجاست که شایستگی‌ها از طریق مشاهده/امتیازدهی انسان و بر پایه چارچوب‌های شایستگی سنجیده می‌شوند. به بیان دیگر، فناوری در موج نخست عمدتاً نقش کانال اجرا و تسهیل‌گر عملیاتی را ایفا می‌کند، نه عامل تصمیم‌ساز یا امتیازده خودکار. همین ویژگی باعث شده محور غالب ادبیات در این موج، به جای تمرکز بر هوشمندسازی، بر دو دغدغه بنیادین متمرکز شود: (۱) حفظ کیفیت روان‌سنجی (روایی/پلیایی) در انتقال به محیط دیجیتال و (۲) مدیریت واکنش‌ها و تجربه داوطلبان به‌عنوان شرط مشروعیت و اثربخشی کانون‌های ارزیابی دیجیتال (Woods et al., 2020; Avni & Luria, 2022).

در بعد روان‌سنجی، یافته‌ها نشان می‌دهد که مطالعات موج نخست معمولاً با منطق هم‌ارزی پیش می‌روند؛ یعنی سؤال کلیدی این است که آیا اجرای مجازی کانون ارزیابی می‌تواند سطوح قابل قبول

روایی سازه، روایی ملاکی و پایایی را مشابه نسخه حضوری حفظ کند یا خیر. شواهد اولیه حاکی از آن است که در بسیاری از ابزارهای دیجیتال هم‌تراز ابزارهای سنتی به‌ویژه آزمون‌های روان‌سنجی آنلاین و مصاحبه‌های دیجیتال ساختاریافته می‌توان به روایی قابل قبول نزدیک شد اما این نتیجه به‌شدت وابسته به کیفیت طراحی، استانداردسازی اجراء و کنترل عوامل مزاحم است (Woods et al., 2020). در حوزه کانون ارزیابی، مطالعات مقایسه‌ای مجازی در برابر حضوری نشان می‌دهند که همبستگی‌ها و شاخص‌های پایایی می‌تواند در سطح قابل دفاع قرار گیرد، اما هم‌زمان هشدار می‌دهند که تغییر کلنل اجراء، احتمالاً ترکیب خطاها را تغییر می‌دهد: خطاهای مشاهده‌گری و سوگیری‌های ناشی از تعامل رودررو ممکن است کاهش یابد، اما خطاهای جدیدی مانند کیفیت اتصال، شرایط محیطی داوطلب، تفاوت تجهیزات، یا تفاوت‌های سواد دیجیتال وارد مدل اندازه‌گیری می‌شود (Avni & Luria, 2022; Ariesthiawati, 2022). بنابراین، موج نخست به جای آنکه به حل مسئله روایی منجر شود، به نوعی جابجایی کانون مسئله روایی ختم می‌گردد. در رویکرد سنتی، روایی عمدتاً تابع طراحی تمرین‌ها (مطابقت تمرین با شایستگی) و کیفیت مشاهده و قضاوت ارزیابان بود. اما پس از دیجیتالی‌سازی، روایی دیگر فقط تابع این دو عامل نیست، بلکه تابع اکولوژی دیجیتال اجراء نیز هست. منظور از اکولوژی دیجیتال اجراء مجموعه‌ای از عوامل زمینه‌ای شامل پلتفرم مورد استفاده (همزمان یا ناهمزمان)، کیفیت اتصال و سخت‌افزار، مهارت‌های دیجیتال داوطلب و ارزیاب، طراحی محیط مجازی (میزان شبیه‌سازی رفتارهای غیرکلامی)، و شرایط محیطی داوطلب (نور، صدا، حریم خصوصی) است. این عوامل می‌توانند به طور نظام‌مند بر نمره‌های مشاهده‌شده تأثیر بگذارند بدون آنکه ربطی به شایستگی هدف داشته باشند. در نتیجه، دفاع از روایی یک کانون ارزیابی مجازی نیازمند مستندسازی نه‌تنها روابط درون‌آزمونی، بلکه ثبات این روابط در طیف گسترده‌ای از شرایط اجرایی است.

تحلیل ادبیات نشان می‌دهد یکی از نقاط کور رایج در موج نخست این است که بسیاری از پروژه‌های دیجیتالی‌سازی، بدون بازاندیشی نظری درباره چه چیز و چگونه در کانون ارزیابی سنجیده می‌شود، صرفاً همان منطق سنتی را به فضای آنلاین منتقل می‌کنند. Woods و همکاران (۲۰۲۰) به‌طور کلی در حوزه رویه‌های انتخاب دیجیتال نشان می‌دهند که بخش مهمی از پژوهش‌ها، نظریه‌های سنتی انتخاب را به ابزارهای دیجیتال الحاق می‌کنند، درحالی‌که فضای دیجیتال می‌تواند سازوکارهای جدیدی از رفتار، تعامل و ادراک عدالت ایجاد کند که با نظریه‌های کلاسیک به‌طور کامل تبیین نمی‌شود. همین استدلال در سطح کانون ارزیابی نیز مصداق دارد. هنگامی که تمرین‌ها آنلاین

می‌شوند، نشانه‌های غیرکلامی، دینامیک تعامل<sup>۱</sup>، مدیریت نوبت‌گیری، و حتی حضور اجتماعی تغییر می‌کند و در نتیجه، نسبت مشاهده رفتاری به سازه‌ها می‌تواند دستخوش تغییر گردد. از این منظر، دستاورد موج نخست آن است که برابری ظاهری نمره‌ها لزوماً به معنای برابری سازه‌ای نیست و مطالعات آینده باید فراتر از مقایسه میانگین‌ها و همبستگی‌های ساده، به سمت آزمون‌های دقیق‌تر مانند ناهمسانی اندازه‌گیری، تحلیل خطای منبع‌محور، و مطالعه اثرات زمینه‌ای حرکت کنند (Liu et al., 2022؛ Woods et al., 2020).

در بعد واکنش داوطلبان، یافته‌ها نشان می‌دهد که دیجیتالی‌سازی کانون ارزیابی از همان آغاز با مسئله پذیرش و عدالت ادراک‌شده گره خورده است. ادبیات انتخاب دیجیتال به‌طور مستمر تأکید می‌کند که حتی اگر ابزار از منظر روان‌سنجی قابل دفاع باشد، شکست در مدیریت تجربه داوطلب می‌تواند به کاهش مشارکت، افزایش انصراف، کاهش اعتماد، و لطمه به برند کارفرمایی منجر شود (Woods et al., 2020). در همین راستا، Weinert و همکاران (۲۰۲۰) با تمرکز بر اشاره صریح به استفاده از فناوری‌های هوشمند (مانند تشخیص گفتار رایانه‌ای) در پیام‌های استخدامی نشان می‌دهند که سیگنال فناوریانه می‌تواند دو اثر هم‌زمان و متعارض ایجاد کند: از یک سو نوآوری و جذابیت کارفرما را افزایش دهد، و از سوی دیگر با فعال‌سازی نگرانی‌های حریم خصوصی، کنترل‌پذیری و ابهام در تصمیم‌گیری، جذابیت را تضعیف کند. این اثرات به‌صورت معنادار توسط پذیرش فناوری و تفاوت‌های فردی/زمینه‌ای تعدیل می‌شوند (Weinert et al., 2020). ترجمان مستقیم این نتیجه برای کانون ارزیابی دیجیتال آن است که طراحی تجربه داوطلب در موج نخست به یک مؤلفه کیفیت تبدیل می‌شود، نه یک موضوع حاشیه‌ای؛ زیرا کانون ارزیابی برخلاف آزمون‌های کوتاه یک تجربه چندمرحله‌ای و پرتنش است و کوچک‌ترین ابهام در معیارها، بازخورد یا نحوه استفاده از داده‌ها می‌تواند عدالت ادراک‌شده را کاهش دهد.

در سطح حکمرانی حرفه‌ای و ریسک، موج نخست اگرچه هنوز هوش مصنوعی تصمیم‌ساز را محور قرار نمی‌دهد، اما بذر دغدغه‌های اخلاقی و حقوقی از همین مرحله کاشته می‌شود. مرورهای اولیه درباره کاربرد AI در جذب و انتخاب نشان می‌دهند که ورود فناوری حتی در سطح دیجیتالی‌سازی موضوعاتی نظیر حریم خصوصی، رضایت آگاهانه، شفافیت و پایداری انسانی-سازمانی را فعال می‌کند (Rezzani et al., 2020). از طرف دیگر، ادبیات هنجاری‌تر در سال‌های بعد یادآور می‌شود که هر چه تصمیمات استخدامی بیشتر به داده‌های دیجیتال و الگوریتم‌ها متکی شوند، دفاع‌پذیری علمی و

<sup>۱</sup> دینامیک یا پویایی تعامل (Interaction Dynamics) در محیط‌های واسطه‌گری شده توسط فناوری، به تغییر در الگوهای ساختاری و محتوایی ارتباط میان افراد اشاره دارد. این مفهوم شامل تغییر در سرعت و توالی گفتگو، نحوه تفسیر نشانه‌های غیرکلامی و میزان نفوذ یا تسلط واسط دیجیتال بر کیفیت بازخوردی‌های دوسویه است که می‌تواند بر نحوه ادراک و قضاوت ارزیاب از رفتارهای داوطلب تأثیر بگذارد.

حقوقی مستلزم اتکا به تحلیل شغل، روایی سنجی مستمر، و مستندسازی دقیق است؛ زیرا ابزارهای انتخاب حتی اگر فقط دیجیتال باشند در عمل به عنوان ابزار سنجش تلقی می‌شوند و باید استانداردهای حرفه‌ای را رعایت کنند (Tippins et al., 2021). بنابراین، موج نخست را باید مرحله ایجاد حساسیت نهادی دانست: سازمان‌ها می‌آموزند که دیجیتالی کردن کانون ارزیابی، صرفاً یک پروژه فناوری اطلاعات نیست، بلکه پروژه‌ای است که باید با اصول روان‌سنجی، عدالت رویه‌ای و ارتباطات سازمانی هم‌تراز شود.

به منظور استخراج نظام‌مند یافته‌های این موج، متون منتخب با استفاده از روش تحلیل محتوا مورد واکاوی قرار گرفتند. در مرحله نخست (کدگذاری باز)، مفاهیم کلیدی مرتبط با ابزارهای دیجیتال و چالش‌های اولیه آن‌ها شناسایی و برچسب‌گذاری شدند. جزئیات این کدها، تعاریف عملیاتی و منابع مستخرج در جدول ۲ ارائه شده است.

جدول ۲. کدگذاری باز منابع موج نخست

منبع	تعریف عملیاتی کد	کد باز	کد
Woods et al., 2020	شناسایی انواع ابزارهای انتخاب دیجیتال شامل درخواست آنلاین، آزمون، مصاحبه دیجیتال، بازی‌وارسازی و شبکه‌های اجتماعی	دسته‌بندی رویه‌های دیجیتال انتخاب کارکنان	x1
Woods et al., 2020	ثبت، ارسال و غربالگری اولیه اطلاعات داوطلب در سامانه‌های اینترنتی	درخواست شغلی آنلاین	x2
Woods et al., 2020	آزمون‌های شناختی، شخصیتی و توانایی در بستر دیجیتال	آزمون‌های روان‌سنجی آنلاین	x3
Woods et al., 2020 Tippins et al., 2021	مصاحبه ویدئویی و تغییر منطبق مشاهده و قضاوت در بستر دیجیتال	مصاحبه دیجیتال (همزمان/غیرهمزمان)	x4
Woods et al., 2020	استفاده از تکالیف شبه‌بازی با هدف سنجش شایستگی‌ها	ارزیابی بازی‌وارسازی شده	x5
Woods et al., 2020	استفاده از داده‌های رسانه‌های اجتماعی داوطلبان در انتخاب	غربالگری مبتنی بر شبکه‌های اجتماعی	x6
Woods et al., 2020	میزان انطباق سنج‌های دیجیتال با سازه‌های نظری موردنظر	روایی سازه‌ای ابزارهای دیجیتال	x7
Woods et al., 2020	توان پیش‌بینی عملکرد شغلی و پیامدهای کاری	روایی ملاکی / پیش‌بین	x8
Woods et al., 2020	تفاوت سطح پشتوانه تجربی میان انواع ابزارهای دیجیتال	ناهمگنی شواهد روایی	x9
Woods et al., 2020	فقدان چارچوب نظری اختصاصی برای DSPها	ابهام مفهومی ابزارهای دیجیتال	x10

کد	کد باز	تعریف عملیاتی کد	منبع
x11	ضرورت تحلیل شغل و ارتباط شغلی	وابستگی مشروعیت ابزارهای دیجیتال به تحلیل شغل	Tippins et al., 2021
x12	خطر سنجش ویژگی‌های نامرتب شغلی	استفاده از داده‌های سهل‌الوصول به‌جای KSAOهای مرتبط	Tippins et al., 2021
x13	عدالت ادراک‌شده داوطلبان	ادراک انصاف و نبود تبعیض در فرایند انتخاب دیجیتال	Woods et al., 2020
x14	شفافیت فرایند انتخاب	وضوح اینکه چه چیزی، چگونه و چرا سنجیده می‌شود	Woods et al., 2020
x15	پذیرش و تمایل به مشارکت داوطلب	میزان تمایل داوطلبان برای شرکت در ابزارهای دیجیتال	Woods et al., 2020
x16	واکنش‌های منفی داوطلبان	بی‌اعتمادی، مقاومت، نگرانی و ترک فرایند	Woods et al., 2020
x17	حریم خصوصی و رضایت آگاهانه	کنترل داده‌ها، اطلاع‌رسانی و رضایت داوطلب	Tippins et al., 2021
x18	سوگیری الگوریتمی و اثر نامطلوب	خطر تبعیض سیستماتیک ناشی از داده یا مدل	Tippins et al., 2021
x19	جعبه سیاه و تبیین‌ناپذیری تصمیم	نا توانی در توضیح منطق تصمیم‌های الگوریتمی	Tippins et al., 2021
x20	جذابیت کارفرما و تصویر سازمان	اثر استفاده از ابزارهای دیجیتال بر برند کارفرما	Weinert et al., 2020
x21	شکاف اجرا و آمادگی سازمانی	توان منابع انسانی برای پیاده‌سازی، پایش و حکمرانی ابزار دیجیتال	Rezzani et al., 2020

پس از استخراج کدهای باز، در مرحله دوم تحلیل، از طریق شناسایی وجوه اشتراک و پیوندهای معنایی میان کدها، مقولات انتزاعی تری تحت عنوان مقولات محوری شکل گرفتند. این فرآیند سنتز، به دسته‌بندی چالش‌ها در ۶ محور اصلی منجر شد که در جدول ۳ نمایش داده شده است. این جدول در واقع زیربنای تبیین چرایی و چگونگی تحولات در موج نخست را تشکیل می‌دهد.

جدول ۳. کدگذاری محوری منابع موج نخست

کد	مقوله محوری	تعریف تحلیلی مقوله محوری	کدهای باز مرتبط
AC1	دیجیتالی‌سازی رویه‌های سنتی انتخاب و ارزیابی	انتقال ابزارها و رویه‌های کلاسیک انتخاب کارکنان به بسترهای دیجیتال، بدون تغییر بنیادین در منطق نظری سنجش	دسته‌بندی رویه‌های دیجیتال انتخاب کارکنان؛ درخواست شغلی آنلاین؛ آزمون‌های روان‌سنجی آنلاین؛ مصاحبه دیجیتال (همزمان/غیرهمزمان)؛ ارزیابی بازی‌وارسازی‌شده؛ غربالگری مبتنی بر شبکه‌های اجتماعی
AC2	تثبیت دغدغه‌های روانی و تحلیل شغل	برجسته‌شدن روانی سازه‌ای، روانی ملاکی و تحلیل شغل به‌عنوان	روایی سازه‌ای ابزارهای دیجیتال؛ روانی ملاکی / پیش‌بین؛ ناهمگنی شواهد روانی؛ ابهام مفهومی

کد	مقوله محوری	تعریف تحلیلی مقوله محوری	کدهای باز مرتبط
		معیارهای مشروعیت علمی و حقوقی ابزارهای دیجیتال	ابزارهای دیجیتال؛ ضرورت تحلیل شغل و ارتباط شغلی؛ خطر سنجش ویژگی‌های نامرتب شغلی
AC3	واکنش داوطلب و تجربه کاربر در انتخاب دیجیتال	اهمیت تجربه داوطلب، عدالت ادراک‌شده و شفافیت به‌عنوان شرط پذیرش و اثربخشی رویه‌های دیجیتال انتخاب	عدالت ادراک‌شده داوطلبان؛ شفافیت فرایند انتخاب؛ پذیرش و تمایل به مشارکت داوطلب؛ واکنش‌های منفی داوطلبان
AC4	چالش‌های اخلاقی و الگوریتمی ابزارهای دیجیتال	بروز ریسک‌های اخلاقی، حقوقی و حرفه‌ای ناشی از داده‌های حساس، الگوریتم‌های غیرشفاف و سوگیری‌های پنهان	حریم خصوصی و رضایت آگاهانه؛ سوگیری الگوریتمی و اثر نامطلوب؛ جعبه سیاه و تبیین‌ناپذیری تصمیم
AC5	پیامدهای انتخاب دیجیتال بر برند کارفرما	تأثیر استفاده از ابزارهای دیجیتال انتخاب بر تصویر ذهنی سازمان و جذابیت کارفرما در بازار کار	جذابیت کارفرما و تصویر سازمان
AC6	آمادگی سازمانی و ظرفیت نهادی اجرا	نقش قابلیت‌ها، دانش و حکمرانی واحد منابع انسانی در تبدیل دیجیتالی‌سازی به فرصت یا ریسک سازمانی	شکاف اجرا و آمادگی سازمانی

در نهایت، برای درک پویایی‌های این موج و تبیین روابط میان متغیرها، مقولات محوری در قالب یک مدل «علی-شرایطی-پیامدی» ترکیب شدند. در این مرحله (کدگذاری انتخابی)، کدهای CM استخراج گردیدند تا نشان دهند که چگونه گذار فناورانه (CM1) تحت شرایطی خاص به پیامدهایی نظیر قطبی شدن واکنش داوطلبان (CM6) منجر می‌شود. خلاصه این زنجیره منطقی در جدول ۴ تدوین شده است.

جدول ۴. مدل علی-شرایطی-پیامدی (با کدهای انتخابی) منابع موج نخست

مؤلفه مدل	کد	کد انتخابی	پشتوانه‌های محوری
علی	CM1	گذار فناورانه کارایی‌محور از ابزارهای سنتی به انتخاب دیجیتال و ناهمگنی بلوغ علمی ابزارهای دیجیتال در غیاب چارچوب نظری مستقل	دیجیتالی‌سازی رویه‌های سنتی انتخاب/ارزیابی
شرایطی	CM2	مشروط‌بودن مشروعیت انتخاب دیجیتال به روایی، تحلیل شغل و قابلیت دفاع	تثبیت مسئله روایی و تحلیل شغل
	CM3	وابستگی پذیرش داوطلب به عدالت ادراک‌شده، شفافیت و تبیین‌پذیری تصمیم	واکنش داوطلب و تجربه کاربر + اخلاق/حقوق/حکمرانی داده
	CM4	تشدید حساسیت اخلاقی با افزایش استفاده از داده‌های ضمنی و غیرقابل کنترل داوطلب	اخلاق/حقوق/حکمرانی داده

مؤلفه مدل	کد	کد انتخابی	پشتوانه‌های محوری
	CM5	نقش ظرفیت نهادی منابع انسانی در تبدیل دیجیتال سازی به فرصت یا ریسک سازمانی	آمادگی سازمان و ظرفیت اجرا
پیامدی	CM6	قطبی شدن واکنش داوطلبان و اثرگذاری مستقیم بر جذابیت کارفرما	واکنش داوطلب و تجربه کاربر + برند کارفرما و ارتباطات استخدامی
	CM7	نهادینه شدن روایی، انصاف و قابلیت دفاع به عنوان معیارهای مسلط ارزیابی DSPها	تثبیت مسئله روایی و تحلیل شغل + اخلاق/حقوق/حکمرانی داده

در جمع‌بندی موج نخست دیجیتال سازی کانون‌های ارزیابی، گرچه در ظاهر یک تغییر بستر اجرا است اما در عمل یک آزمون نهادی برای سنجش‌گری حرفه‌ای محسوب می‌شود؛ زیرا نشان داد انتقال تمرین‌ها به فضای مجازی تنها زمانی قابل دفاع است که زنجیره شواهد روایی و پایایی حفظ شود و هم‌زمان، تجربه داوطلب از منظر عدالت ادراک‌شده، شفافیت و کنترل داده‌ها مدیریت گردد. بر اساس مدل علی - شرایطی - پیامدی، فشارهای کارایی محور و گسترش فناوری، محرک دیجیتال سازی شد (CM1)، اما مشروعیت و اثربخشی این گذار به شرایطی چون اتکای صریح به تحلیل شغل و مستندسازی روایی (CM2)، تبیین‌پذیری و عدالت رویه‌ای در مواجهه با داوطلب (CM3)، حساسیت‌های اخلاقی ناشی از داده‌های ضمنی و غیرقابل کنترل (CM4) و ظرفیت نهادی واحد منابع انسانی برای اجرا و حکمرانی (CM5) وابسته ماند. پیامد این موج، از یک سو قطبی شدن واکنش داوطلبان و اثر مستقیم بر برند کارفرما (CM6) و از سوی دیگر نهادینه شدن روایی، انصاف و قابلیت دفاع به عنوان معیارهای مسلط ارزیابی انتخاب دیجیتال (CM7) بود؛ بنابراین موج نخست را باید مرحله‌ای دانست که در آن، پیش از ورود به هوشمندسازی تصمیم، چارچوب‌های روان‌سنجی و اخلاقی به عنوان پیش‌نیازهای غیرقابل اغماض برای گذار به موج‌های بعدی تثبیت می‌شوند.

### موج دوم: هوشمندسازی ارزیابی و گذار به تحلیل داده محور رفتار

موج دوم تحول، نقطه گسست از دیجیتال سازی به مثابه تغییر کانال اجرا و ورود به هوشمندسازی به مثابه تغییر منطق سنجش است. در این موج، فناوری از نقش پلتفرم و زیرساخت (اجرای مجازی تمرین‌ها) فراتر می‌رود و به یک عامل تحلیلی و نمره‌ساز تبدیل می‌شود؛ به گونه‌ای که خروجی ارزیابی نه صرفاً حاصل جمع‌بندی ادراک ارزیابان انسانی، بلکه برآمده از مدل‌های محاسباتی، یادگیری ماشین و تحلیل داده‌های رفتاری (به‌ویژه زبان و تعامل) است (Liou et al., 2022؛ Campion & Campion, 2023؛ Koenig et al., 2023). بنابراین، اگر موج نخست عمدتاً دغدغه «هم‌ارزی روان‌سنجی در محیط آنلاین» را تثبیت می‌کرد، موج دوم با طرح پرسش‌های عمیق‌تر مواجه می‌شود: «چه چیزی به عنوان داده معتبر تلقی می‌شود؟»، «چگونه داده‌های غیرساختاریافته به سازه‌های شغلی نگاشت می‌شوند؟» و «آیا نمره الگوریتمی می‌تواند همان استنباط‌های معتبر نمره انسانی را پشتیبانی کند؟» (Liou et al., )

(2022).

یافته‌های روان‌سنجی موج دوم نشان می‌دهد که ورود داده‌های ارگانیک/غیرواکنشی و داده‌های غیرساختاریافته (متن، گفتار، تعاملات) فرصت و تهدید را هم‌زمان ایجاد کرده است. از یک‌سو، این داده‌ها ظرفیت افزایش تفکیک‌پذیری، مقیاس‌پذیری و حساسیت سنجش را فراهم می‌کنند؛ از سوی دیگر، «پیوند داده-سازه» را پیچیده‌تر و شکننده‌تر می‌سازند، زیرا مشاهده‌پذیری یک نشانه رفتاری لزوماً به معنای سازه‌مندی آن نیست و بدون مدل مفهومی و شواهد روایی، خطر تبدیل ارزیابی به «پیش‌بینی، بدون نظریه» افزایش می‌یابد (Liou et al., 2022). در همین راستا، Liou و همکاران نشان می‌دهند که اصول کلاسیک روایی و پایایی در ارزیابی‌های فناورانه نه تنها حذف نمی‌شود، بلکه باید با دقت بیشتری بازتفسیر شود؛ زیرا پویایی الگوریتم‌ها، به‌روزرسانی مدل‌ها و تغییر ترکیب داده‌ها می‌تواند ثبات اندازه‌گیری را تهدید کرده و «رانس سازه‌ای» (construct drift) ایجاد کند؛ یعنی مدلی که امروز ظاهراً همان سازه را می‌سنجد، در نسخه‌های بعدی عملاً به سمت پیش‌بینی‌های جانشین و کم‌ربط حرکت کند (Liou et al., 2022). بر این مبنا، موج دوم از نظر روش‌شناختی با یک الزام تازه همراه است: اعتبارسنجی باید «مستمر و چرخه‌ای» باشد، نه یک‌بار برای همیشه.

در سطح کانون‌های ارزیابی، مهم‌ترین دستاورد موج دوم، گذار از فرآیندهای پرهزینه و زمان‌بر نمره‌دهی انسانی به پارادایم نمره‌دهی خودکار پاسخ‌های باز<sup>۱</sup> است. این تحول مستقیماً بنیان‌های اقتصادی و عملیاتی کانون ارزیابی را تحت تأثیر قرار می‌دهد؛ چرا که بخش عمده‌ای از هزینه‌های اجرایی در رویکرد سنتی، معطوف به فرآیند مشاهده، کدگذاری و استخراج نمره توسط ارزیابان خیره است. شواهد تجربی متأخر نشان می‌دهد الگوریتم‌های یادگیری عمیق به‌ویژه مدل‌های مبتنی بر معماری ترنسفورمر<sup>۲</sup> به سطحی از کفایت دست یافته‌اند که می‌توانند در ارزیابی پاسخ‌های متنی و استخراج شایستگی‌های شغلی، به همگرایی آماری قابل‌توجهی با ارزیابان انسانی دست یابند (Thompson et al., 2023). اهمیت این دستاورد صرفاً در جنبه‌های فنی خلاصه نمی‌شود؛ بلکه به لحاظ روان‌سنجی به این معناست که اگر سقف توافق انسانی محدود باشد رسیدن الگوریتم به نزدیکی آن سقف، امکان طراحی سیستم‌های هیبریدی را فراهم می‌کند که در آن، الگوریتم نقش «ارزیاب سوم» یا «پشتیبان پایایی» را بازی کند و بار شناختی ارزیاب انسانی را کاهش دهد، بدون آنکه الزاماً روایی

<sup>۱</sup> Automated Scoring of Constructed Responses (ASCR)

<sup>۲</sup> مدل‌های ترنسفورمر (Transformer Models): نسل نوینی از معماری‌های یادگیری عمیق در پردازش زبان طبیعی (NLP) هستند که با بهره‌گیری از مکانیزم خود-توجهی (Self-Attention)، برخلاف مدل‌های خطی قدیمی، قادر به درک روابط میان‌واژه‌ای در کل بافتار متن (Context) می‌باشند. این ویژگی به مدل اجازه می‌دهد تا مفاهیم پیچیده و معنای نهفته در پاسخ‌های باز داوطلبان را با دقت مشابه ارزیاب انسانی کدگذاری کند. مدل‌هایی نظیر BERT و خانواده GPT از این معماری بهره می‌برند.

ملاکی قربانی شود (Thompson et al., 2023؛ Campion & Campion, 2023). با این حال، یک نکته تعیین‌کننده در همین یافته‌ها آشکار است: کیفیت داده‌های معیار آموزشی (نمرات SMEs و توافق بین ارزیابان) زیربنای کیفیت الگوریتم است؛ به بیان دقیق‌تر، الگوریتم نمی‌تواند از داده‌ای که خود از نظر سازه‌ای یا پایایی ضعیف است، یک سنجه معتبر بسازد. بنابراین، موج دوم برخلاف تصور رایج، جایگزینی ساده انسان با ماشین نیست، بلکه نوعی دیالوگ فنی میان قضاوت انسانی و دقت ماشین یا به عبارتی وابستگی کیفیت مدل به غنای قضاوت انسانی را تثبیت می‌کند (Hickman et al., 2023؛ Campion & Campion, 2023).

در امتداد همین تحول، پژوهش‌های موج دوم به جای تمرکز صرف بر «متن‌های طولانی و تشریحی»، به سمت تحلیل تمرین‌های بین‌فردی و نقش‌آفرینی‌های سریع نیز حرکت کرده‌اند؛ یعنی دقیقاً همان قلب رفتاری کانون ارزیابی. Hickman و همکاران (۲۰۲۳) نشان می‌دهند که با استفاده از پردازش زبان طبیعی و یادگیری ماشین می‌توان بخشی از واریانس نمرات انسانی را در تمرین‌های بین‌فردی ارزیابی کرد و حتی سناریوهایی را آزمون کرد که در آن جایگزینی یک یا چند ارزیاب انسانی با نمره‌دهی الگوریتمی، روایی ملاکی را در سطح قابل قبول حفظ می‌کند (Hickman et al., 2023). نکته ظریف اما راهبردی اینجاست که این پژوهش‌ها صرفاً دقت الگوریتم را گزارش نمی‌کنند، بلکه به یک مسئله کلاسیک یعنی واریانس نامرتب با سازه نظیر حساسیت نمره به نشانه‌های سطحی مانند شمار واژگان یا پرحرفی در کانون ارزیابی حساس‌اند. وقتی مدل‌های یادگیری ماشین از داده زبانی تغذیه می‌شوند، ریسک آن وجود دارد که به جای کیفیت رفتاری و محتوایی، بر شاخص‌های کمی سهل‌الوصول متمرکز شوند. بنابراین، موج دوم با یک معیار جدید کیفیت همراه است: مدل مطلوب مدلی است که همگرایی بالا با خبره را با پالایش سوگیری‌های زبانی<sup>۱</sup> و شفافیت ویژگی‌های پیش‌بین ترکیب کند (Hickman et al., 2023).

هم‌زمان با نمره‌دهی خودکار، موج دوم یک جابه‌جایی مفهومی عمیق را تثبیت می‌کند: حرکت از نمرات کلی و برداشت‌های تجمیعی به سمت نشانه‌های رفتاری خرد و ابعاد میانجی رفتار. پژوهش Breil و همکاران (۲۰۲۳) نشان می‌دهد اگرچه رفتار، جوهره‌ی اصلی کانون ارزیابی است، اما در بسیاری از سنت‌های پژوهشی، تحلیل آن به‌صورت نظام‌مند و در سطح خرد انجام نشده بود. این مطالعه با کدگذاری نشانه‌های رفتاری در سطح خرد طی تمرینات ایفای نقش در کانون ارزیابی و استخراج ابعاد کلان (نظیر عاملیت، پیوندجویی<sup>۲</sup>، آرامش بین‌فردی و شایستگی ذهنی)، نشان می‌دهد که رفتارهای مشاهده‌شده از ثبات و قدرت تبیین قابل‌توجهی برخوردارند. از این رو، این داده‌ها

<sup>1</sup> Linguistic Bias Mitigation

<sup>2</sup> Communion

ورودی‌های کارآمدتری برای مدل‌های داده‌محور فراهم می‌سازند (Breil et al., 2023). پیام موج دوم در اینجا کاملاً شفاف است: هوشمندسازی در کانون ارزیابی فراتر از صرف خودکارسازی نمره‌دهی است و در واقع بازطراحی معرفت‌شناسی سنجش به شمار می‌رود؛ یعنی گذار از اتکا به نمرات کلی و مبهم، به سوی قضاوت عملکرد مبتنی بر شواهد رفتاری قابل‌ردیابی و تحلیل‌پذیر.

در سطح سازمانی و اکوسیستمی، موج دوم با واقعیتی همراه است که Guenole و همکاران (۲۰۲۲) آن را تبیین می‌کنند: در عمل، مرز میان ابزارها (پرسشنامه، آزمون، مصاحبه، کانون ارزیابی) برداشته شده و متخصصان مدیریت استعداد، مجموعه‌ای پیوسته و تلفیقی از فناوری‌های سنجش را به کار می‌گیرند. بنابراین هوشمندسازی کانون ارزیابی در این موج به معنای توسعه یک ابزار منفرد نیست؛ بلکه به معنای انتظام‌بخشی به داده‌های حاصل از منابع چندگانه (متن، رفتار و امتیازات انسانی) در قالب یک پلتفرم تحلیل پیش‌بین است (Koenig et al., 2023؛ Guenole et al., 2022). تفاوت ظریف اما بنیادین این رویکرد با موج سوم در این است که در موج دوم، ابزارها (مانند مصاحبه یا آزمون) کماکان ساختار مستقل و از پیش تعیین‌شده خود را حفظ می‌کنند، اما خروجی‌های آن‌ها در یک معماری آماری-تصمیم‌گیری برای پیش‌بینی دقیق‌تر پیامدهای شغلی ادغام می‌شوند این دیدگاه جامع توضیح می‌دهد که چرا نقش یادگیری ماشین در این موج، از نمره‌دهی پاسخ‌ها فراتر رفته و به سطوح راهبردی‌تری همچون بهینه‌سازی وزن هر شایستگی در پیش‌بینی عملکرد نزدیک شده است؛ تلاشی که هدف نهایی آن، افزایش پایایی و روایی تصمیمات استخدامی از طریق تجمع داده‌های ساختاریافته در مقیاس انبوه است (Campion & Campion, 2023؛ Koenig et al., 2023).

با وجود دستاوردهای یادشده، تحلیل یافته‌های موج دوم آشکار می‌سازد که کانون توجه پژوهش‌ها عمدتاً معطوف به کارایی، دقت فنی و هم‌سویی با قضاوت انسانی بوده و یکپارچه‌سازی نظام‌مند ملاحظات عدالت و حکمرانی، غالباً مغفول مانده است. این در حالی است که از منظر روان‌سنجی فناوری‌محور، سوگیری اندازه‌گیری و اثرات نامطلوب گروهی به‌عنوان چالش‌های ساختاری مدل‌های داده‌محور شناخته می‌شوند؛ چالش‌هایی که مدیریت آن‌ها مستلزم تمهیدات دقیق در طراحی داده، گزینش ویژگی‌ها و پایش مستمر است (Liou et al., 2022؛ Koenig et al., 2023). برآیند تحلیل‌ها در این بخش، نقشه راهی را پیش روی مقله ترسیم می‌کند. اگرچه موج دوم، امکان‌پذیری فنی کانون ارزیابی داده‌محور را تثبیت کرده است، اما مشروعیت علمی و حرفه‌ای آن را در گرو تکمیل سه حلقه اصلی می‌داند: (۱) حلقه روایی (استقرار پیوند منطقی میان داده-سازه و استنباط شواهد)، (۲) حلقه پایایی (تضمین ثبات مدل و تکرارپذیری خروجی در برابر نوسانات داده) و (۳) حلقه حکمرانی (پایش مداوم، ممیزی و کنترل اثرات نامطلوب گروهی). بر همین اساس، تجویز ضمنی این موج پرهیز از حذف کامل عامل انسانی و گذار به سمت مدل‌های هیبریدی و رویکرد انسان در چرخه است؛

معماری دقیقی که در آن الگوریتم، بار پایایی و مقیاس‌پذیری را بر دوش می‌کشد و انسان، مسئولیت تبیین‌پذیری، قضاوت حرفه‌ای و نظارت بر کاربرد را حفظ می‌کند (Campion & Campion, 2023؛ Hickman et al., 2023).

جمع‌بندی موج دوم این است که هوشمندسازی، در این مرحله به معنای داده‌محور کردن رفتار است: رفتار به داده تبدیل می‌شود، داده به ویژگی، ویژگی به نمره، و نمره به استنباط. اما هر حلقه از این زنجیره، اگر با منطق روان‌سنجی و اصول طراحی شایستگی/تحلیل شغل همراه نشود، می‌تواند به تولید نمرات دقیق اما نامعتبر یا منصفانه‌نما اما تبعیض‌آفرین منجر گردد (Liou et al., 2022؛ Thompson et al., 2023؛ Koenig et al., 2023). بنابراین، موج دوم در عین تثبیت امکان‌های محاسباتی، ضرورت «بازتعریف معیار کیفیت» را نیز تثبیت می‌کند: کیفیت دیگر فقط به اجرای صحیح تمرین‌ها یا آموزش ارزیاب محدود نیست، بلکه به معماری داده، کیفیت برچسب‌های انسانی، کنترل آلودگی‌های زبانی/رفتاری، و پایش مستمر عملکرد مدل وابسته است (Breib et al., 2023؛ Hickman et al., 2023).

موج دوم با همه دستاوردهای خود در نمره‌دهی خودکار و تحلیل داده‌محور رفتار، سه محدودیت ساختاری را آشکار کرد که به‌طور مستقیم، مسیر را به سوی موج سوم گشود. نخست، وابستگی شدید مدل‌های یادگیری ماشین به داده‌های برچسب‌خورده انسانی و هزینه تولید برچسب معتبر باعث شد بسیاری از سازمان‌ها به دنبال ابزارهایی باشند که بتوانند بخشی از محتوا، سناریو و حتی تعامل را با هزینه کمتر تولید یا مقیاس‌پذیر کنند؛ این دقیقاً همان وعده مدل‌های مولد است (Thompson et al., 2023؛ Behrend & Landers, 2025). دوم، پویایی و تغییرپذیری مدل‌ها در موج دوم، ضرورت «اعتبارسنجی مستمر» را تثبیت کرد و نشان داد که سیستم‌های هوشمند بدون چارچوب حکمرانی و استانداردهای روایی، آسیب‌پذیرند؛ در موج سوم که تولید و تطبیق محتوا نیز به سیستم افزوده می‌شود، این نیاز به چارچوب‌های روایی/انصاف چند برابر می‌گردد (Dorsey & Michaels, 2022؛ Liou et al., 2022؛ Kaldaras et al., 2024). سوم، تمرکز موج دوم بر نمره‌دهی نشان داد که حتی اگر نمره‌دهی خودکار به سطحی نزدیک به انسان برسد، مسئله «چه چیزی را می‌سنجیم» و «چگونه از نمره به تصمیم می‌رسیم» همچنان باقی است. موج سوم با افزودن قابلیت تبیین، تعامل، و طراحی سناریو، تلاش می‌کند شکاف میان سنجش و تصمیم را با سازوکارهای جدید پر کند. هرچند خود این سازوکارها مسائل تازه‌ای در پایایی و انصاف ایجاد می‌کنند (Zhang et al., 2024؛ Ringelband & Warneke, 2025).

در گام نخست تحلیل موج دوم، با بررسی متون منتخب، مفاهیم اولیه شناسایی شدند. حاصل این واکاوی، استخراج ۲۱ کد باز بود که زیربنای مفهومی تحول به سمت هوشمندسازی را تشکیل می‌دهند؛ جزئیات این کدها در جدول ۵ آورده شده است.

جدول ۵. کدگذاری باز منابع موج دوم

منبع	تعریف عملیاتی کد	کد باز	کد
Liou et al., 2022	تأکید بر اینکه اعتبار استنباط‌های نمره‌ای در TBA <sup>1</sup> همچنان تابع اصول روان‌سنجی است	رویکرد روان‌سنجی به ارزیابی‌های مبتنی بر فناوری	x1
Liou et al., 2022	استفاده از داده‌های غیرواکنشی/رفتاری <sup>2</sup> به‌عنوان ورودی سنجش	داده‌های ارگانیک و ردپاهای رفتاری دیجیتال در سنجش	x2
Liou et al., 2022 Koenig et al., 2023	به‌کارگیری یادگیری ماشین و یادگیری عمیق <sup>3</sup> برای نمره‌دهی، طبقه‌بندی و پیش‌بینی پیامدها	نمره‌دهی و پیش‌بینی مبتنی بر یادگیری ماشین	x3
Liou et al., 2022	نیاز به بازتعریف شواهد روایی و شفاف‌سازی پیوند «داده→سازه» در داده‌های غیرسنتی	پسچیده‌شدن روایی در سنجش فناوریانه	x4
Liou et al., 2022	اثر به‌روزرسانی مدل‌ها و سیستم‌های تطبیقی بر ثبات اندازه‌گیری	پویایی الگوریتم و چالش پایایی	x5
Liou et al., 2022 Modi, 2023	بازتولید/تشدید نابرابری‌ها از طریق داده یا مدل و پیامدهای متفاوت برای گروه‌ها	سوگیری داده‌الگوریتم و اثرات گروهی	x6
Ariesthiawati, 2022	استقرار VAC <sup>4</sup> با زیرساخت فناوری و فرآیندهای اجرایی/نظارتی در مقیاس	کانون ارزیابی مجازی به‌مثابه پلتفرم سازمانی	x7
Ariesthiawati, 2022	نقش همه‌گیری و فشار تقاضا در جهش پروژه‌های VAC	تسریع مجازی‌سازی به‌واسطه کووید-۱۹ و رشد تقاضا	x8
Ariesthiawati, 2022	ارزیابی تجربه کاربران منابع انسانی/مدیران از منظر سهولت استفاده و سودمندی درک‌شده	شاخص‌های اثربخشی VAC: سهولت استفاده و مفید بودن ادراک‌شده	x9
Ariesthiawati, 2022 Avni & Luria, 2022	نبود تفاوت معنادار/هم‌ارزی نتایج شایستگی‌ها بین VAC و FTF-AC <sup>5</sup> در برخی شواهد	همسانی نتایج VAC با کانون حضوری	x10
Avni & Luria, 2022	حفظ سطح قابل اتکا از توافق ارزیابان در VAC	پایایی بین‌ارزیابان در کانون ارزیابی مجازی	x11
Avni & Luria, 2022	روایی پیش‌بین در VAC برای همه شایستگی‌ها یکسان نیست	ناهمگنی روایی پیش‌بین شایستگی‌ها در VAC	x12

<sup>1</sup> TBA (Technology-Based Assessment): ارزیابی‌های مبتنی بر فناوری؛ شامل هرگونه فرآیند سنجش که در آن فناوری نقشی محوری در ارائه، اجرا یا نمره‌دهی ایفا می‌کند.

<sup>2</sup> Digital Traces (ردپاهای دیجیتال): داده‌های غیرواکنشی که از تعاملات کاربر در محیط‌های دیجیتال (مانند سرعت تایپ، مسیر حرکت ماوس یا لاگ‌های سیستم) به‌جا می‌ماند.

<sup>3</sup> ML/DL (Machine/Deep Learning): یادگیری ماشین و یادگیری عمیق؛ زیرمجموعه‌هایی از هوش مصنوعی که با استفاده از الگوریتم‌ها و شبکه‌های عصبی پیچیده، الگوهای موجود در داده‌ها را شناسایی می‌کنند.

<sup>4</sup> VAC (Virtual Assessment Center): کانون ارزیابی مجازی؛ نسخه‌ای از کانون ارزیابی که تمام فرآیندهای آن به صورت آنلاین و از راه دور انجام می‌شود.

<sup>5</sup> FTF-AC (Face-to-Face Assessment Center): کانون ارزیابی حضوری؛ شیوه سنتی اجرای کانون که مستلزم حضور فیزیکی ارزیاب و داوطلب در یک مکان واحد است.

کد	کد باز	تعریف عملیاتی کد	منبع
x13	اکوسیستم سنجش و هم‌پوشانی ابزارها در عمل	مرزهای قراردادی بین آزمون، پرسشنامه و AC و استفاده ترکیبی در محیط واقعی	Guenole et al., 2022
x14	خودکارسازی نمره‌دهی پاسخ‌های تشریحی در تمرین‌های AC	نمره‌دهی خودکار داده متنی/روایی (constructed responses) در تمرین‌های کانون ارزیابی	Thompson et al., Koenig et al., 2023 2023
x15	هم‌ارزی دقت الگوریتم با ارزیاب انسانی در نمره‌دهی متنی	همبستگی بالا/نزدیک به پایایی انسانی در مدل‌های پیشرفته (مثل ترنسفورمرها)	Thompson et al., 2023
x16	نمره‌دهی خودکار تمرین‌های نقش‌آفرینی بین‌فردی با NLP/ML	استخراج ویژگی‌های زبانی و نمره‌دهی تمرین‌های role-play (خصوصاً speeded <sup>2</sup> )	Hickman et al., 2023
x17	آلودگی شمار واژگان و کنترل متغیرهای مزاحم در نمره‌دهی ML	خطر وابستگی نمره ML به word count و ضرورت بهینه‌سازی ویژگی‌ها	Hickman et al., 2023
x18	پایش مستمر و اعتبارسنجی مداوم مدل‌ها	پایش و اعتبارسنجی مستمر <sup>۳</sup> به‌عنوان شرط استقرار مسئولانه	Hickman et al., Campion & 2023 Campion, 2023
x19	ارزش داده‌های خرد، رفتاری به‌عنوان ورودی ML	کدگذاری نشانه‌های رفتاری خرد و استفاده از آن‌ها برای افزایش روایی/پیش‌بینی	Breil et al., 2023
x20	تقویت قضاوت انسانی به‌جای جایگزینی کامل	رویکرد تقویت <sup>۴</sup> و حفظ حضور انسان در چرخه <sup>۵</sup> انتخاب AC/	Campion & Campion, 2023 Modi, 2023
x21	حکمرانی اخلاقی: شفافیت، تبیین‌پذیری و پاسخ‌گویی	ضرورت XAI/traceability <sup>۶</sup> ، حساسرسی اخلاقی و مسئولیت‌پذیری سازمانی	Modi, 2023 Campion & Campion, 2023

<sup>۱</sup> Constructed Responses (پاسخ‌های برساخته/باز): نوعی از سوالات ارزیابی که داوطلب باید به جای انتخاب گزینه، خود به تولید متن یا تحلیل (پاسخ تشریحی) بپردازد.

<sup>۲</sup> Speeded Exercises (تمرین‌های زمانی سریع): تمرین‌هایی که در بازه زمانی بسیار کوتاه طراحی می‌شوند تا واکنش‌های آنی و ظرفیت پردازش سریع داوطلب را بسنجند.

<sup>۳</sup> Continuous Monitoring & Ongoing Validation: پایش و اعتبارسنجی مستمر؛ فرآیند نظارت دائم بر خروجی الگوریتم‌ها برای اطمینان از اینکه دقت و روایی آن‌ها در طول زمان دچار زوال (Drift) نشود.

<sup>۴</sup> Augmentation (تقویت): رویکردی که در آن هوش مصنوعی برای بهبود و تکمیل توانمندی‌های انسانی به‌کار می‌رود، نه برای حذف جایگزینی انسان.

<sup>۵</sup> Human-in-the-loop (HITL): حضور انسان در چرخه؛ مدلی که در آن الگوریتم‌ها پیشنهاد می‌دهند اما تصمیم نهایی یا نظارت بر فرآیند بر عهده انسان است.

<sup>۶</sup> XAI (Explainable AI) & Traceability: هوش مصنوعی تبیین‌پذیر و قابلیت ردیابی؛ به معنای توانایی مدل برای ارائه دلایل منطقی پشت تصمیمات و امکان پیگیری مسیر استنتاج الگوریتم از ورودی تا خروجی.

در مرحله کدگذاری محوری، کدهای باز جدول ۵ بر پایه هم‌پوشانی معنایی تجمیع شدند که حاصل آن شناسایی ۶ مقوله محوری (AC1 تا AC6) است. منطبق پیوند میان این کدهای خرد و مقولات کلان در جدول ۶ مستند شده است.

جدول ۶. کدگذاری محوری منابع موج دوم

کد	مقوله محوری	تعریف تحلیلی مقوله محوری	کدهای باز مرتبط
AC1	پلتفرمی‌سازی کانون ارزیابی و تثبیت VAC	تبدیل کانون ارزیابی به پلتفرم دیجیتال، قابل اجرا در مقیاس، با شاخص‌های تجربه کاربری و اثربخشی عملیاتی	کانون ارزیابی مجازی به‌مثابه پلتفرم سازمانی؛ تسریع مجازی‌سازی به‌واسطه کووید-۱۹ و رشد تقاضا؛ شاخص‌های اثربخشی: VAC؛ سهولت استفاده و مفید بودن ادراک‌شده
AC2	هم‌ارزی روان‌سنجی VAC و کانون حضوری با ناهمگنی شایستگی‌ها	وجود شواهد همسانی کلی در توزیع/میانگین نمرات و پایایی بین‌ارزیابان، همراه با تفاوت در روایی پیش‌بین برخی شایستگی‌ها	همسانی نتایج VAC با کانون حضوری؛ پایایی بین‌ارزیابان در کانون ارزیابی مجازی؛ ناهمگنی روایی پیش‌بین شایستگی‌ها در VAC
AC3	بازپیکربندی روایی/پایایی در سنجش فناورانه و الگوریتم‌های پویا	پیش‌پسینه‌شدن مفاهیم کلاسیک روایی و پایایی در حضور داده‌های ارگانیک، مدل‌های ML و به‌روزرسانی‌های الگوریتمی	رویکرد روان‌سنجی به ارزیابی‌های مبتنی بر فناوری؛ داده‌های ارگانیک و ردپاهای رفتاری دیجیتال در سنجش؛ پیش‌پسینه‌شدن روایی در سنجش فناورانه؛ پویایی الگوریتم و چالش پایایی
AC4	خودکارسازی نمره‌دهی و داده‌محور شدن مشاهده در کانون ارزیابی	حرکت از مشاهده انسانی، صرف به سمت نمره‌دهی خودکار داده‌های متنی/گفتاری و بهره‌گیری از نشانگان خرد رفتاری برای پیش‌بینی	خودکارسازی نمره‌دهی پاسخ‌های تشریحی در تمرین‌های AC؛ هم‌ارزی دقت الگوریتم با ارزیاب انسانی در نمره‌دهی متنی؛ نمره‌دهی خودکار تمرین‌های نقش‌آفرینی بین‌فردی با NLP/ML؛ ارزش داده‌های خرد، رفتاری به‌عنوان ورودی ML
AC5	کنترل خطاهای فناورانه و کیفیت داده در ML	شناسایی منابع خطای جدید (آلودگی، کیفیت برچسب‌ها، متغیرهای مزاحم) و تبدیل آن‌ها به الزامات طراحی/اعتبارسنجی	آلودگی شمار واژگان و کنترل متغیرهای مزاحم در نمره‌دهی ML؛ پایش مستمر و اعتبارسنجی مداوم مدل‌ها
AC6	اخلاق و حکمرانی انسان‌محور در کانون‌های نیمه‌هوشمند	تثبیت شفافیت/تبیین‌پذیری/پاسخ‌گویی و human-in-the-loop به‌عنوان شرایط مشروعیت و عدالت	سوگیری داده الگوریتم و اثرات گروهی؛ تقویت قضاوت انسانی به‌جای جایگزینی کامل؛ حکمرانی اخلاقی: شفافیت، تبیین‌پذیری و پاسخ‌گویی

برای درک پویایی‌های حاکم بر این موج و ترسیم مسیر حرکت از محرک‌ها به سمت نتایج، مقولات محوری در قالب یک الگوی فرآیندی انتظام یافتند. در این مرحله که منطبق بر کدگذاری انتخابی است، روابط میان متغیرهای علی، شرایط میانجی و پیامدهای نهایی تحلیل شد. نحوه شکل‌گیری این

زنجیره منطقی و کدهای انتخابی حاصل از آن (CM1 تا CM7) به همراه پشتوانه‌های محوری هر یک، در جدول ۷ به شکلی صریح بیان شده است.

جدول ۷. مدل علی-شرایطی-پیامدی (با کدهای انتخابی) منابع موج دوم

مؤلفه مدل	کد	کد انتخابی	پشتوانه‌های محوری
علی	CM1	پلتفرمی‌سازی کانون ارزیابی تحت فشار مقیاس‌پذیری و اختلال‌های محیطی (مثل کووید-۱۹) و تبدیل VAC به قالب غالب اجرا در بسیاری از کاربردها	پلتفرمی‌سازی کانون ارزیابی و تثبیت VAC
علی	CM2	ورود یادگیری ماشین به نمره‌دهی تمرین‌های کانون ارزیابی (متنی/گفتاری/بین‌فردی) به‌عنوان پاسخ به هزینه و محدودیت مشاهده انسانی و تلاش برای هم‌ارزی دقت با ارزیاب	خودکارسازی نمره‌دهی و داده‌محور شدن مشاهده در کانون ارزیابی
شرایطی	CM3	مشروط بودن اعتبار کانون‌های مجازی/نیمه‌هوشمند به بازپیکربندی شواهد روایی و پایایی در محیط‌های پویا (داده ارگانیک، الگوریتم‌های به‌روزرسان)	بازپیکربندی روایی/پایایی در سنجش فناوریانه و الگوریتم‌های پویا
شرایطی	CM4	وابستگی کیفیت نمره‌دهی خودکار به کیفیت داده و کنترل خطاهای فناوریانه (آلودگی‌ها، متغیرهای مزاحم، کیفیت برچسب‌های انسانی) و الزام پایش/اعتبارسنجی مستمر	کنترل خطاهای فناوریانه و کیفیت داده در ML
شرایطی	CM5	ضرورت حکمرانی اخلاقی و طراحی انسان‌محور: کاهش سوگیری، تقویت شفافیت و تبیین‌پذیری، و استقرار انسان در حلقه تصمیم برای دفاع‌پذیری حرفه‌ای	اخلاق و حکمرانی انسان‌محور در کانون‌های نیمه‌هوشمند
پیامدی	CM6	پیامد موج دوم: تثبیت VAC به‌عنوان «زیرساخت اجرا» و حرکت از همسانی، صرف به سمت سنجش اثربخشی/روایی پیش‌بین، تفکیکی در سطح شایستگی‌ها	هم‌ارزی روان‌سنجی VAC و کانون حضوری با ناهمگنی شایستگی‌ها + پلتفرمی‌سازی کانون ارزیابی و تثبیت VAC
پیامدی	CM7	پیامد موج دوم: نهادینه شدن «هوشمندسازی محدود اما عملیاتی» در قالب خودکارسازی نمره‌دهی و تحلیل داده‌های رفتاری، همراه با استاندارد شدن پایش مداوم، حکمرانی و human-in-the-loop به‌عنوان شرط مشروعیت	خودکارسازی نمره‌دهی و داده‌محور شدن مشاهده در کانون ارزیابی + کنترل خطاهای فناوریانه و کیفیت داده در ML + اخلاق و حکمرانی انسان‌محور در کانون‌های نیمه‌هوشمند

در موج دوم، کانون ارزیابی از دیجیتالی‌شدن ابزارها به سمت پلتفرم ارزیابی و سپس نیمه‌هوشمندسازی سنجش حرکت می‌کند؛ یعنی VAC به‌عنوان زیرساخت غالب اجرا تثبیت می‌شود و هم‌زمان، یادگیری ماشین برای خودکارسازی نمره‌دهی، داده‌های متنی/گفتاری و بهره‌گیری از نشانگان خرد رفتاری وارد قلب فرآیند مشاهده می‌گردد. با این حال، ادبیات موج دوم نشان می‌دهد این گذار صرفاً فنی نیست و مشروعیت آن به شرایط سخت‌گیرانه‌ای وابسته است: بازتعریف و انباشت شواهد روایی/پایایی در محیط‌های پویا، کنترل منابع خطا و کیفیت داده، و استقرار حکمرانی اخلاقی با محوریت شفافیت، پاسخ‌گویی و human-in-the-loop. بنابراین، دستاورد موج دوم را می‌توان واقعی شدن

امکان‌های هوشمندسازی در سطح نمره‌دهی و تحلیل داده دانست، اما با این قید که استانداردهای روان‌سنجی و اخلاقی به‌عنوان پیش‌نیازهای استقرار و دفاع‌پذیری حرفه‌ای، بیش از گذشته در کانون توجه قرار می‌گیرند (Liou et al., 2022؛ Avni & Luria, 2022؛ Thompson et al., 2023؛ Hickman et al., 2023؛ Modi, 2023؛ Campion & Campion, 2023).

### **موج سوم: ظهور مدل‌های مولد و عامل‌های هوشمند و گذار از نمره‌دهی خودکار به طراحی/تعامل تطبیقی در کانون‌های ارزیابی هوشمند**

یافته‌های این پژوهش نشان می‌دهد که موج سوم تحول کانون‌های ارزیابی، نقطه‌ای گسست‌گونه نسبت به دو موج پیشین محسوب می‌شود؛ به‌گونه‌ای که تمرکز از دیجیتالی‌سازی ابزارها و حتی خودکارسازی نمره‌دهی مبتنی بر داده فراتر رفته و به بازطراحی بنیادین منطق تولید شواهد سنجش، معماری ارزیابی و نقش فناوری در قضاوت انسانی منتهی می‌گردد. در این موج، مفهوم هوشمندسازی از برداشت‌های تقلیل‌گرایانه پیشین فاصله می‌گیرد و دیگر نمی‌توان آن را صرفاً به استخراج نشانه‌های رفتاری از داده‌های غیرساختاریافته یا جایگزینی محدود قضاوت ارزیاب انسانی با الگوریتم‌های محاسباتی فروکاست. بلکه هوشمندسازی در این مرحله به ظهور کانون ارزیابی به‌عنوان یک سامانه ارزیابی هوشمند، مولد و تطبیقی دلالت دارد؛ سامانه‌ای که منطق آن علاوه بر اجرای ایستای تمرین‌ها بر تولید پویا و موقعیت‌مند سناریوهای ارزیابی، تنظیم بلادرنگ مسیر تعامل بر مبنای پاسخ‌ها و تصمیم‌های داوطلب، و تحلیل یکپارچه داده‌های رفتاری در سطوح مختلف استوار است. در چنین چارچوبی، کانون ارزیابی از یک بستر سنجش ازپیش تعریف‌شده فراتر رفته و به محیطی تعاملی تبدیل می‌شود که خود در فرایند شکل‌گیری شواهد سنجش نقش فعال و سازنده ایفا می‌کند. این تحول با آنچه Baranyi (۲۰۲۵) از آن به‌عنوان گذار به فرایندهای تحول‌یافته دیجیتال در انتخاب کارکنان یاد می‌کند هم‌راستاست؛ سطحی که در آن فناوری دیگر ابزار پشتیبان تصمیم نیست بلکه خود در منطق تصمیم‌سازی و طراحی فرایند مداخله می‌کند.

در این چارچوب، ورود مدل‌های مولد و به‌ویژه مدل‌های زبانی بزرگ به کانون‌های ارزیابی، افق جدیدی از سنجش مبتنی بر تعامل مولد را می‌گشاید. اگر موج دوم عمدتاً با داده‌محور کردن رفتار و نمره‌دهی مبتنی بر مدل شناخته می‌شد، موج سوم سنجش را از قالب مشاهده پاسخ‌های نسبتاً ثابت خارج کرده و آن را در بستر تعاملات تطبیقی با محیط‌های شبیه‌سازی‌شده‌ای قرار می‌دهد که خود توسط AI تولید و تنظیم می‌شوند (Tuit et al., 2025؛ Behrend & Landers, 2025). در این سطح، کانون ارزیابی دیگر صرفاً اجراکننده تمرین‌های ازپیش‌طراحی‌شده نیست، بلکه به سامانه‌ای تبدیل می‌شود که می‌تواند نقش‌آفرینی‌ها، سناریوهای کاری و حتی کنش متقابل سایر بازیگران (همکار، مدیر یا مشتری مجازی) را به‌صورت پویا و متناسب با شغل، شایستگی هدف و سطح پیچیدگی شناختی تنظیم کند.

بدین ترتیب، شواهد سنجش در کنار پاسخ داوطلب از نحوه تطبیق، تصمیم‌گیری و تعامل او در مواجهه با عدم قطعیت و تغییر استخراج می‌شود.

از منظر نظری، موج سوم بر این فرض استوار است که رفتار داوطلب در محیط‌های تعاملی مولد، بالقوه کمتر قابل دستکاری و نزدیک‌تر به رفتار واقعی شغلی است؛ زیرا داوطلب با یک مسیر از پیش قابل پیش‌بینی مواجه نیست بلکه باید به کنش‌های غیرمنتظره عامل‌های مجازی واکنش نشان دهد (Tuit et al., 2025). با این حال، یافته‌ها به‌روشنی نشان می‌دهد که این مزیت بالقوه، خود مستلزم آزمون‌های سخت‌گیرانه روایی و پایایی است. انعطاف‌پذیری و تولیدگری که نقطه قوت مدل‌های مولد محسوب می‌شود، در محیط‌های پرریسک (high-stakes) مانند انتخاب و ارتقای شغلی می‌تواند به تهدیدی برای استانداردسازی، تکرارپذیری و انصاف تبدیل شود، مگر آنکه منطق سنجش، سازه هدف و زنجیره استنباط به‌دقت مستندسازی گردد.

در همین راستا، Kaldaras و همکاران (۲۰۲۴) هشدار می‌دهند که استفاده از ابزارهای مولد در طراحی یا نمره‌دهی ارزیابی‌ها، بدون اطمینان از هم‌ترازی سازه‌ای با قضاوت انسانی آموزش‌دیده، خطر جایگزینی سازه را به‌طور جدی افزایش می‌دهد. در چنین وضعیتی، نمرات ممکن است به‌جای شایستگی‌های هدف، به ویژگی‌های جانشین مانند فصاحت زبانی، سبک بیان یا الگوهای رایج پاسخ‌دهی حساس شوند. از این‌رو، موج سوم فشار نهادی تازه‌ای برای بازآفرینی چارچوب‌های اعتبارسنجی ایجاد می‌کند؛ چارچوب‌هایی که باید بر مبنای نظریه روایی و استانداردهای آزمون‌های آموزشی و روان‌شناختی، متناسب با ماهیت تطبیقی و مولد سیستم‌ها بازطراحی شوند (Kaldaras et al., 2024؛ Dorsey & Michaels, 2022). به بیان دیگر، هرچه سیستم ارزیابی مولدتر و پویاتر می‌شود، بار اثبات روایی و انصاف نیز سنگین‌تر خواهد شد.

یافته‌ها در حوزه سنجش ویژگی‌های فردی با استفاده از مدل‌های زبان‌محور، این تنش را به‌صورت تجربی برجسته می‌سازد. Zhang و همکاران (۲۰۲۴) نشان می‌دهند که LLMها در مصاحبه‌های ویدئویی غیرهمزمان می‌توانند از نظر روایی صفر-شات در پیش‌بینی برخی صفات شخصیتی، به عملکردی هم‌سطح یا حتی بالاتر از مدل‌های وظیفه‌محور دست یابند و توضیحاتی تولید کنند که تا حدی با نظریه‌های شخصیت قابل تفسیر است. با این حال، ضعف در پایایی بازآزمایی و ظهور الگوهای سوگیرانه در برخی ابعاد، محدودیت جدی این مدل‌ها برای استفاده مستقل در تصمیم‌های سرنوشت‌ساز منابع انسانی را آشکار می‌سازد (Zhang et al., 2024). یافته‌های مکمل Hickman و همکاران (۲۰۲۴) نیز نشان می‌دهد که دقت الگوریتم‌های زبان‌محور به‌شدت به منبع معیار وابسته است؛ به‌گونه‌ای که این مدل‌ها در بازتولید ارزیابی‌های مصاحبه‌گران موفق‌تر از پیش‌بینی خودگزارش‌دهی هستند. این نتیجه از منظر موج سوم اهمیت مفهومی دارد، زیرا نشان می‌دهد که بخشی از کارکرد

LLMها نه سنجش مستقیم سازه‌های روان‌شناختی، بلکه مدل‌سازی قضاوت‌های اجتماعی و ادراکی انسانی است؛ امری که لزوماً معادل سنجش «ویژگی واقعی فرد» نیست.

در سطح روش‌شناختی، موج سوم پیامدهای مستقیمی برای طراحی ابزارها و مواد ارزیابی نیز به همراه دارد. Landers و Behrend (۲۰۲۵) نشان می‌دهند که استفاده از LLMها در تولید محتوای ارزیابی، مستلزم تصمیم‌گیری آگاهانه درباره نقش مدل (دستیار، تولیدکننده تطبیقی یا شریک تعاملی)، انتخاب الگوریتم، پیش‌آزمون، پایش و کنترل اخلاقی است. تعمیم این یافته به کانون‌های ارزیابی حاکی از آن است که تولید سناریوهای نقش‌آفرینی، سؤالات پس از تمرین یا مکالمات تعاملی بدون مهندسی دقیق و نظارت مستمر، می‌تواند اعتبار درونی و بیرونی ارزیابی را تضعیف کرده و مقایسه منصفانه داوطلبان را مخدوش سازد. از این منظر، موج سوم اگرچه «قدرت تولید» را به کانون ارزیابی می‌افزاید، اما هم‌زمان ضرورت حکمرانی تولید را به‌عنوان یک الزام نوظهور مطرح می‌کند.

در نهایت، یافته‌ها نشان می‌دهد که موج سوم، بیش از هر زمان دیگری مسئله اخلاق، حقوق و شفافیت را به کانون توجه می‌آورد. با گسترش استفاده از داده‌های متنوع آزمون‌های روان‌سنجی تا رده‌های دیجیتال و مشاهده‌های مستقیم و غیرمستقیم مسئله اصلی دیگر صرفاً وجود الگوریتم نیست، بلکه ماهیت داده‌های آموزشی، مبنای نظری مدل‌ها و میزان تبیین‌پذیری تصمیم‌هاست (Ringelband & Warneke, 2025). پژوهش‌های اخیر تأکید می‌کنند که بدون حسابرسی عدالت، شفافیت الگوریتمی و استقرار سازوکارهای Human-in-the-loop، هوشمندسازی کانون‌های ارزیابی با خطر کاهش مشروعیت علمی و حقوقی مواجه خواهد شد (Mihaljević et al., 2024).

در مرحله کدگذاری باز موج سوم، با واکاوی منابع منتخب، ۳۷ کد پایه (X1 تا X37) شناسایی شد که ابعاد ظهور مدل‌های مولد را تبیین می‌کنند. جزئیات این کدها، تعاریف عملیاتی و منابع استخراج آن‌ها در جدول ۸ ارائه شده است.

جدول ۸. کدگذاری باز منابع موج سوم

کد	کد باز	تعریف عملیاتی کد	منبع
X1	روایی و عدالت جنسیتی/سنی در VAC	ارزیابی هم‌زمان روایی و عدالت ارزیابی در VAC برحسب جنسیت و سن	Usmani et al., 2025
X2	بسته تحلیلی جامع برای روان‌سنجی VAC	به‌کارگیری همبستگی، PCA <sup>۱</sup> ، آمگا، MANOVA <sup>۲</sup> /کای‌دوالجستیک برای سنجش روایی/ساختار/پایایی	Usmani et al., 2025

<sup>۱</sup> PCA (Principal Component Analysis): تحلیل مؤلفه‌های اصلی؛ یک روش آماری برای کاهش ابعاد داده‌ها و شناسایی ساختارهای

زیربنایی در آزمون‌های روان‌سنجی.

<sup>۲</sup> MANOVA (Multivariate Analysis of Variance): تحلیل واریانس چندمتغیری؛ برای بررسی تفاوت‌های معنادار بین گروه‌ها (مثلاً

جنسیتی یا سنی) در چندین متغیر وابسته به‌طور هم‌زمان.

کد	کد باز	تعریف عملیاتی کد	منبع
X3	نبود اثر نامطلوب با معیار ۵/۴	استفاده از قاعده ۵/۴ برای بررسی Adverse Impact و گزارش نبود شواهد معنادار	Usmani et al., 2025
X4	هم‌راستایی عدالت VAC با کانون حضوری	نتیجه‌گیری مبنی بر قابلیت جایگزینی منصفانه VAC نسبت به FTF-AC	Usmani et al., 2025
X5	تفاوت در «ضبط قابلیت‌ها» در VAC و حضوری	VAC و FTF-AC قابلیت‌ها را به شیوه متفاوت بازنمایی می‌کنند (شیوه سنجش متفاوت)	Avni & Luria, 2025
X6	هم‌ارزی الگوی روابط بیرونی در VAC و حضوری	با وجود تفاوت شیوه سنجش، الگوهای ارتباطی با متغیرهای بیرونی مشابه گزارش می‌شود	Avni & Luria, 2025
X7	پایایی بین‌ارزیابان بالا در VAC	حفظ سطح بالای توافق ارزیابان در VAC (هم‌سطح‌انزدیک به حضوری)	Avni & Luria, 2025
X8	بالا تر بودن نمرات VAC نسبت به حضوری	مشاهده تفاوت سطح نمرات (VAC بالاتر) و طرح ضرورت تبیین بستر اجرا/ادراک ارزیاب	Avni & Luria, 2025
X9	طبقه‌بندی سه‌سطحی تحول دیجیتال انتخاب	تفکیک «انتخاب دستی»، «دیجیتالی‌شده»، «تحول یافته دیجیتال» به‌عنوان چارچوب تحلیل گذار	Baranyi, 2025
X10	تمایز فناوری دیجیتال و افزونه‌های AI	تفکیک Digital Technologies از AI Enhancements <sup>1</sup> و تأکید بر پیامدهای متفاوت	Baranyi, 2025
X11	منافع تحول دیجیتال در انتخاب	افزایش کارایی/سرعت، کاهش هزینه، بهبود تجربه داوطلب، و امکان کاهش سوگیری انسانی	Baranyi, 2025
X12	چالش‌ها و ریسک‌های تحول دیجیتال	نگرانی‌های اخلاقی/حقوقی، سوگیری الگوریتمی، موانع سازمانی، و پذیرش/ادراک منفی داوطلبان	Baranyi, 2025
X13	شبیه‌سازی مصاحبه HR مبتنی بر AI	معرفی سامانه مصاحبه شبیه‌سازی‌شده به‌عنوان مؤلفه کانون ارزیابی هوشمند برای ارزیابی واقع‌گرایانه‌تر	Sarumathi et al., 2025
X14	تولید سؤال پویا/تطبیقی با LLM و RL <sup>2</sup>	تولید پرسش‌های متناسب با رزومه و شرح شغل و تطبیق سطح پرسش‌ها با یادگیری تقویتی	Sarumathi et al., 2025
X15	ارزیابی چندبعدی شناختی-عاطفی	سنجش هم‌زمان ابعاد شناختی و عاطفی با تحلیل لحن/هیجان گفتار و حالات چهره	Sarumathi et al., 2025
X16	چت‌بات مصاحبه برای توانمندسازی داوطلب	طراحی ابزارهای شبیه‌ساز مصاحبه با سؤال‌سازی پویا و بازخورد تفصیلی برای ارتقای آمادگی	Rajeswari et al., 2025

<sup>1</sup> AI Enhancements (افزونه‌های هوش مصنوعی): قابلیت‌های پیشرفته‌ای که فراتر از دیجیتالی‌سازی ساده، قدرت تحلیل، پیش‌بینی و تولید محتوا را به سامانه‌های موجود اضافه می‌کنند.

<sup>2</sup> RL (Reinforcement Learning): یادگیری تقویتی؛ نوعی از یادگیری ماشین که در آن سیستم از طریق تعامل با محیط و دریافت پاداش/جریمه، مجینه‌ترین مسیر (مثلاً بهترین سؤال) را انتخاب می‌کند.  
LLM (Large Language Models): مدل‌های زبانی بزرگ؛ مدل‌های هوش مصنوعی (مانند GPT) که قادر به درک و تولید متن‌های انسانی با دقت بالا هستند.

کد	کد باز	تعریف عملیاتی کد	منبع
X17	تغییر کارکرد از گزینش به توسعه (Upskilling)	برجسته شدن نقش ابزارهای هوشمند در «توسعه/آمادگی» نه صرفاً «قضاوت استخدامی»	Rajeswari et al., 2025
X18	رفتارمحوری در تبیین قضاوت ارزیاب	استفاده از نشانه‌های رفتاری کدگذاری شده برای پیش‌بینی قضاوت ارزیابان در تمرین‌های بین‌فردی	Grunenberg et al., 2025
X19	اعتبارسنجی برون‌نمونه‌ای و کنترل بیش‌برازش	اعتبارسنجی متقاطع <sup>۱</sup> و مقایسه درون‌ابرون‌نمونه‌ای برای پرهیز از overfitting در مدل‌های هوشمند	Grunenberg et al., 2025
X20	تبیین‌پذیری قضاوت با الگوهای رفتاری	ارائه الگوهای خاص تمرین و الگوهای پایدار بین‌تمرینی برای شفاف‌سازی مبنای قضاوت	Grunenberg et al., 2025
X21	نمردهی متن شبیه‌سازی‌های AC با AI	استفاده از AI برای نمردهی پاسخ‌های نوشتاری شبیه‌سازی‌های AC و سنجش شایستگی‌ها	Bronkhorst & Becker, 2024
X22	روایی همگرا/ملاکی قابل دفاع AI در AC متن‌محور	گزارش روایی همگرا و ملاکی قابل قبول و هم‌ترازی با ارزیاب انسانی	Bronkhorst & Becker, 2024
X23	بهبود استفاده از دامنه نمردهی (Range Utilization)	AI دامنه کامل مقیاس را بهتر از ارزیاب انسانی به کار می‌گیرد (کاهش Range Restriction انسانی)	Bronkhorst & Becker, 2024
X24	استنباط شخصیت از داده‌های درون‌فرآیندی استخدام	پیش‌بینی Big Five از CV و پاسخ‌های متنی کوتاه بدون اتکا به ردپاهای بیرونی	Grunenberg et al., 2024
X25	ارزش اخلاقی/حقوقی داده‌های درون‌فرآیندی	ترجیح داده‌های تولیدشده در فرایند استخدام نسبت به داده‌های بیرونی از منظر اخلاق اقلون	Grunenberg et al., 2024
X26	LLM در LLM: رویای قابل قبول با پایایی بازآزمایی پایین	ارزیابی جامع LLM ها در مصاحبه ویدیویی غیرهمزمان و برجسته‌شدن ضعف test-retest <sup>3</sup>	Zhang et al., 2024
X27	ریسک بازتولید «ادراک مصاحبه‌گر» به جای «شخصیت واقعی»	الگوریتم‌های زبانی ممکن است بیشتر ادراک اجتماعی مصاحبه‌گران را بازتولید کنند تا سازه پایدار شخصیت	Hickman et al., 2024
X28	امکان انطباق اصول حرفه‌ای/حقوقی با ML (مشروط)	استدلال درباره قابلیت رعایت اصول انتخاب با ML همراه با نیاز به حکمرانی/اعتبارسنجی مستمر	Campion, 2025
X29	ریسک تقلب و تضعیف اعتبار در ارزیابی‌های بدون ناظر	آسیب‌پذیری ارزیابی‌های unproctored و متن‌های تولیدی AI به‌عنوان تهدید اعتبار اطلاعات	Campion, 2025
X30	اتوماسیون سرتاسری فرایند ارزیابی با معماری چندعامله	معماری چندعامل (جذب/مصاحبه/ارزیابی/تصمیم) برای خودکارسازی و مقیاس‌پذیری جذب	Pathak & Pandey, 2025

<sup>۱</sup> Cross-validation (اعتبارسنجی متقاطع): تکنیکی در یادگیری ماشین برای سنجش قدرت تعمیم مدل به داده‌های جدید و جلوگیری از حفظ کردن داده‌های آموزشی (Overfitting).

<sup>۲</sup> AVI (Asynchronous Video Interview): مصاحبه ویدیویی غیرهمزمان؛ مصاحبه‌ای که داوطلب پاسخ‌ها را ضبط کرده و ارزیابی آن در زمان دیگری توسط انسان یا هوش مصنوعی انجام می‌شود.

<sup>۳</sup> Test-retest Reliability (پایایی بازآزمایی): شاخصی که نشان می‌دهد اگر یک ارزیابی در دو زمان مختلف برای یک فرد تکرار شود، نتایج تا چه حد ثبات دارند.

کد	کد باز	تعریف عملیاتی کد	منبع
X31	استانداردسازی نمره‌دهی با معیارها و بازخورد خودکار	نمره‌دهی روبریک‌محور و تولید بازخورد خودکار برای افزایش ثبات و کاهش تفاوت‌های ارزیابی	Pathak & Pandey, 2025
X32	حسابرسی عدالت: بافت‌مندی و مشارکت ذی‌نفعان	عدالت‌سنجی عملیاتی باید داده‌افزاینده-بافت‌مند و مشارکتی (فراتر از فروشنده/مشتری) باشد	Mihaljević et al., 2024
X33	کارگزار معتمد داده و دوراهی عدالت-حریم خصوصی	پیشنهاد کارگزار معتمد داده <sup>۱</sup> برای امکان‌میزی مستقل عدالت بدون نقض حریم خصوصی	Mihaljević et al., 2024
X34	ریسک‌های پایایی/سوگیری/اخلاق و نیاز به HITL	وابستگی پایایی به کیفیت داده/تبیین‌پذیری/ثبات؛ خطر سوگیری؛ ضرورت human-in-the-loop	Raza, 2024
X35	مسائل اخلاقی و حقوقی: انواع داده/اجعه‌سیاه/تبعیض غیرمستقیم	طبقه‌بندی داده‌ها، ریسک تبعیض حتی بدون نیت، دشواری دفاع حقوقی در غیاب شفافیت و مستندسازی	Ringelband & Warneke, 2025
X36	هم‌افزایی انسان-AI و ضرورت «لمس انسانی»	تأکید بر اینکه کارایی AI باید با قضاوت انسانی/همدلی برای تناسب فرهنگی و تجربه داوطلب تکمیل شود	Chakraborty & Sharada, 2025
X37	نقش‌های چندگانه LLM در طراحی مواد ارزیابی/پژوهش	مفهوم‌پردازی نقش‌های متنوع LLM و تأکید بر کنترل روایی/کیفیت و مهندسی پرامپت	Behrend & Landers, 2025

در مرحله کدگذاری محوری، کدهای باز بر اساس تشابه موضوعی در ۸ مقوله کلان (AC1 تا AC8) دسته‌بندی شدند. منطق این گذار از کدهای خرد به مقولات محوری در جدول ۹ مستند شده است.

جدول ۹. کدگذاری محوری منابع موج سوم

کد	مقوله محوری	تعریف تحلیلی مقوله محوری	کدهای باز مرتبط
AC1	بلوغ VAC با شواهد روایی/پایایی و عدالت جمعیت‌شناختی	موج سوم، VAC را از «گزینه اضطراری» به «گزینه قابل دفاع روان‌سنجی» ارتقا می‌دهد و عدالت جنسیتی/سنی را به معیار اصلی مشروعیت تبدیل می‌کند	روایی و عدالت جنسیتی/سنی در VAC + بسته تحلیلی جامع برای روان‌سنجی VAC + نبود اثر نامطلوب با معیار ۵/۴ + هم‌راستایی عدالت VAC با کانون حضوری + تفاوت در «ضبط قابلیت‌ها» در VAC و حضوری + هم‌ارزی الگوی روابط بیرونی در VAC و حضوری + پایایی بین‌ارزیابان بالا در VAC + بالاتر بودن نمرات VAC نسبت به حضوری
AC2	چارچوب‌گذاری تحول دیجیتال انتخاب و مدیریت منافع/ریسک‌ها	صورت‌بندی گذار از دیجیتالی‌سازی تا تحول دیجیتال، همراه با توازن میان کارایی/تجربه داوطلب و ریسک‌های اخلاقی/پذیرشی	طبقه‌بندی سه‌سطحی تحول دیجیتال انتخاب + تمایز فناوری دیجیتال و افزونه‌های AI + منافع تحول دیجیتال در انتخاب + چالش‌ها و ریسک‌های تحول دیجیتال

<sup>۱</sup> Data Trusted Broker (کارگزار معتمد داده): یک نهاد یا ساختار واسطه که وظیفه مدیریت امن داده‌های حساس داوطلبان را بر عهده دارد تا تمیزی‌های عدالت بدون افشای هویت افراد انجام شود.

کد	مقوله محوری	تعریف تحلیلی مقوله محوری	کدهای باز مرتبط
AC3	شبیه‌سازی‌های تعاملی AI و چرخش به ارزیابی/توسعه چندبعدی	ورود شبیه‌سازی‌های هوشمند (مصاحبه/عامل) با تولید سؤال تطبیقی و سنجش چندبعدی (شناختی-عاطفی) و همچنین کارکرد توسعه‌ای <sup>۱</sup>	شبیه‌سازی مصاحبه HR مبتنی بر AI + تولید سؤال پویا/تطبیقی با LLM و RL + ارزیابی چندبعدی شناختی-عاطفی + چت‌بات مصاحبه برای توانمندسازی داوطلب + تغییر کارکرد از گزینش به توسعه (Upskilling) + هم‌افزایی انسان-AI و ضرورت «لمس انسانی»
AC4	رفتار محوری و تبیین‌پذیری قضاوت به‌عنوان زیرساخت Smart AC	تبدیل «مشاهده انسانی» به «مدل‌سازی رفتاری قابل تبیین» با الزام اعتبارسنجی برون‌نمونه‌ای و کاهش بیش‌برازش	رفتار محوری در تبیین قضاوت ارزیاب + اعتبارسنجی برون‌نمونه‌ای و کنترل بیش‌برازش + تبیین‌پذیری قضاوت با الگوهای رفتاری
AC5	خودکارسازی نمره‌دهی و استنباط سازه‌ها از متن/ویدئو با محدودیت‌های روان‌سنجی	گسترش نمره‌دهی متن‌محور و استنباط شخصیت از داده‌های درون‌فرآیندی و AVI؛ همراه با ریسک ناپایداری، بازتولید ادراک، و نیاز به تفکیک «سازه واقعی» از «برداشت اجتماعی»	نمره‌دهی متن شبیه‌سازی‌های AC با AI + روایی همگرا/املاکی قابل دفاع AI در AC متن‌محور + بهبود استفاده از دامنه نمره‌دهی + استنباط شخصیت از داده‌های درون‌فرآیندی استخدام + ارزش اخلاقی/حقوقی داده‌های درون‌فرآیندی + LLM در AVI: روایی قابل قبول با پایایی بازآزمایی پایین + ریسک بازتولید «ادراک مصاحبه‌گر» به‌جای «شخصیت واقعی»
AC6	اتوماسیون عامل‌محور و استانداردسازی تصمیم در اکوسیستم جذب	شکل‌گیری معماری‌های چندعامله برای خودکارسازی انتها به انتها و اتکای بیشتر بر روبریک‌های استاندارد و بازخورد خودکار برای ثبات	اتوماسیون سرتاسری فرایند ارزیابی با معماری چندعامله + استانداردسازی نمره‌دهی با معیارها و بازخورد خودکار
AC7	حکمرانی، ممیزی عدالت، حریم خصوصی و انطباق حقوقی/اخلاقی	موج سوم، «کنترل‌پذیری و پاسخ‌گویی» را هم‌سنگ فناوری می‌نشانند: ممیزی عدالت، موازنه عدالت-حریم خصوصی، نظارت انسانی، و خوانایی حقوقی تصمیم	امکان انطباق اصول حرفه‌ای/حقوقی با ML (مشروط) + ریسک تقلب و تضعیف اعتبار در ارزیابی‌های بدون ناظر + حساسی عدالت: بافت‌مندی و مشارکت ذی‌نفعان + امین داده و دوراهی عدالت - حریم خصوصی + ریسک‌های پایایی/سوگیری/اخلاق و نیاز به HITL + مسائل اخلاقی و حقوقی: انواع داده/جعبه‌سیاه/تبعیض غیرمستقیم
AC8	روش‌مندی تولید مواد ارزیابی با LLM و کنترل روایی	استفاده از LLM در طراحی مواد سنجش/پژوهش، تنها با مهندسی پرامپت، کنترل کیفیت، و ممیزی ادعاهای روایی قابل دفاع است	نقش‌های چندگانه LLM در طراحی مواد ارزیابی/پژوهش

<sup>۱</sup> Upskilling (کارکرد توسعه‌ای): فرآیند یادگیری مهارت‌های جدید یا بهبود مهارت‌های فعلی برای تطبیق با نیازهای تغییر یافته شغل.

در گام نهایی، از طریق کدگذاری انتخابی، روابط میان عوامل علی، شرایطی و پیامدی در قالب یک مدل فرآیندی تبیین شد. کدهای انتخابی حاصل (CM1 تا CM8) و زنجیره منطقی حاکم بر آن‌ها در جدول ۱۰ ارائه شده است.

جدول ۱۰. مدل علی-شرایطی-پیامدی (با کدهای انتخابی) منابع موج سوم

مؤلفه مدل	کد	کد انتخابی	پشتوانه‌های محوری
علی	CM1	بلوغ VAC و تمرکز نهادی بر عدالت جمعیت‌شناختی؛ تبدیل «روایی+عدالت» به شرط پذیرش VAC در استخدام‌های پربریک	بلوغ VAC با شواهد روایی/پایایی و عدالت جمعیت‌شناختی
علی	CM2	جهش ابزارهای تعاملی AI (LLM/RL) /تحلیل هیجانی) و حرکت از سنجش ایستا به شبیه‌سازی‌های پویا و چندبعدی	شبیه‌سازی‌های تعاملی AI و چرخش به ارزیابی/توسعه چندبعدی
علی	CM3	گذار ساختاری انتخاب کارکنان از دیجیتال‌سازی به تحول دیجیتال و افزایش فشار کارایی/مقیاس/تجربه داوطلب	چارچوب‌گذاری تحول دیجیتال انتخاب و مدیریت منافع/ریسک‌ها
شرایطی	CM4	مشروط‌بودن اعتبار موج سوم به رفتارمحوری، تبیین‌پذیری و اعتبارسنجی برون‌نمونه‌ای برای کنترل بیش‌برازش و دفاع‌پذیری علمی	رفتارمحوری و تبیین‌پذیری قضاوت + الزام اعتبارسنجی
شرایطی	CM5	ناهمگنی روان‌سنجی LLM/متن‌آویدنو و ریسک بازتولید ادراک یا ناپایداری؛ ضرورت تفکیک سازه از برداشت و کنترل پایایی	خودکارسازی نمره‌دهی و استنباط سازه‌ها از متن‌آویدنو با محدودیت‌های روان‌سنجی
شرایطی	CM6	ضرورت حکمرانی؛ ممیزی عدالت، موازنه عدالت-حریم خصوصی، شفافیت/مستندسازی، HITL، و مدیریت ریسک تقلب در ارزیابی‌های بدون ناظر	حکمرانی، ممیزی عدالت، حریم خصوصی و انطباق حقوقی/اخلاقی
پیامدی	CM7	پیامد موج سوم: شکل‌گیری اکوسیستم کانون ارزیابی هوشمند (VAC) معتبر و منصفانه + شبیه‌سازی‌های AI + نمره‌دهی داده‌محور + اتوماسیون عامل‌محور)	اتوماسیون عامل‌محور و استانداردسازی تصمیم در اکوسیستم جذب
پیامدی	CM8	پیامد موج سوم: بازتعریف نقش HR از مجری ارزیابی به معمار سیستم و ناظر ریسک/عدالت؛ توسعه مهارت‌های حکمرانی داده، ممیزی و طراحی تجربه داوطلب	روش‌مندی تولید مواد ارزیابی با LLM و کنترل روایی

یافته‌های موج سوم نشان می‌دهد مسیر گذار از کانون‌های ارزیابی سنتی به کانون‌های ارزیابی هوشمند وارد مرحله اکوسیستم‌محور شده است. در این موج، VAC نه فقط به عنوان جایگزین فناورانه، بلکه به عنوان روش سنجش، قابل دفاع از منظر روایی/پایایی و عدالت جمعیت‌شناختی تثبیت می‌شود (Usmani et al., 2025؛ Avni & Luria, 2025). هم‌زمان، توسعه شبیه‌سازی‌های تعاملی مبتنی بر AI (تولید سؤال تطبیقی، سنجش چندبعدی شناختی - عاطفی، و ابزارهای توسعه‌محور مانند چت‌بات‌های مصاحبه) نشان می‌دهد کانون‌های هوشمند در حال حرکت از ارزیابی‌های ایستا به سمت سناریوهای

پویا و نزدیک‌تر به واقعیت کار هستند (Rajeswari et al., 2025؛ Sarumathi et al., 2025). با این وجود، موج سوم به موازات این تحولات، الزامات سخت‌گیرانه‌ای را برای مشروعیت‌بخشی مطرح می‌سازد: رفتارمحوری و تبیین‌پذیری مبتنی بر نشانه‌های رفتاری، اعتبارسنجی برون‌نمونه‌ای برای کنترل بیش‌برازش، و حساسیت به محدودیت‌های روان‌سنجی LLMها (ناپایداری و بازتولید ادراک) (Hickman et al., 2024؛ Zhang et al., 2024؛ Grunenberget al., 2025). در سطح حکمرانی، تمرکز از توصیه‌های کلی اخلاقی به سازوکارهای عملیاتی مانند ممیزی عدالت (بافت‌مند و مشارکتی)، مدیریت دوراهی عدالت-حریم خصوصی با ایده امین داده، نظارت انسانی و الزامات حقوقی/دفاع‌پذیری تصمیم‌ها منتقل می‌شود (Ringelband & Warneke, 2025؛ Raza, 2024؛ Mihaljević et al., 2024). نتیجه نهایی موج سوم، ظهور کانون ارزیابی هوشمند به‌عنوان سامانه تصمیم‌گیری سازمانی، تنظیم‌پذیر است: ترکیبی از VACهای معتبر، شبیه‌سازی‌های AI، نمره‌دهی داده‌محور و اتوماسیون عامل‌محور، اما مشروط به حکمرانی، ممیزی و نقش فعال HR در طراحی و نظارت (Campion, 2025؛ Baranyi, 2025؛ Chakraborty & Sharada, 2025؛ Pathak & Pandey, 2025).

### نتیجه‌گیری و پیشنهادها

پژوهش حاضر با هدف تبیین فرایند گذار از مراکز ارزیابی سنتی به مراکز ارزیابی هوشمند و ارائه چارچوبی مفهومی برای مدیریت این تحول انجام شد. یافته‌ها نشان داد که این دگرگونی در سه موج متمایز قابل مشاهده است. موج نخست، که بر دیجیتالی‌سازی ابزارها و مهاجرت فرایندها به بسترهای برخط استوار بود، اگرچه موجبات ارتقای دسترسی‌پذیری و کارایی عملیاتی را فراهم آورد، اما چالش‌های بنیادینی را در باب هم‌ارزی روان‌سنجی و کیفیت تجربه داوطلب پدیدار ساخت. در موج دوم، با استقرار تحلیل‌های داده‌محور و مکانیزم‌های نمره‌دهی خودکار، تلاش شد تا ضمن بهینه‌سازی هزینه‌ها، دقت سنجش افزایش یابد. با این وجود، این رویکرد پیچیدگی‌های نوظهوری را در ابعاد روایی سازه، پایایی و مدیریت خطاهای فناورانه به همراه داشت. شواهد تجربی نشان می‌دهد که کارآمدی این سامانه‌های نیمه‌هوشمند، همبستگی معناداری با کیفیت داده‌های ورودی، فرایند برچسب‌گذاری انسانی، معماری مدل‌ها و پایش مستمر دارد. بر این اساس، یافته‌ها اثبات می‌کند که جایگزینی تام و تمام قضاوت انسانی نه امکان‌پذیر است و نه مطلوب؛ بلکه پارادایم غالب، حرکت به سوی هوشمندسازی تقویتی با محوریت رویکرد انسان در حلقه تصمیم است.

در موج سوم، سیر تکامل کانون‌های ارزیابی به ظهور یک زیست‌بوم پیشرفته منجر می‌شود. جایی که هم‌گرایی مدل‌های مولد، عامل‌های هوشمند و شبیه‌سازی‌های تعاملی، منطق سنجش را از نمره‌دهی ایستا به طراحی و تعامل تطبیقی ارتقا می‌دهد. شواهد پژوهش بیانگر آن است که اگرچه این رویکرد نوین، ظرفیت استخراج شواهد رفتاری غنی‌تر و دارای وفاداری بالاتر به واقعیت شغلی را

داراست، اما هم‌زمان الزامات سخت‌گیرانه‌تری را در ابعاد روایی، پلایی، عدالت و انطباق حقوقی تحمیل می‌کند. در این پارادایم، دغدغه کانونی از صرف کارایی یا دقت الگوریتم فراتر رفته و بر قابلیت دفاع علمی، حقوقی و اجتماعی تصمیمات هوش‌محور متمرکز می‌گردد. فراتحلیل این سه موج نشان می‌دهد که کانون ارزیابی هوشمند باید به‌عنوان یک سامانه اجتماعی- فناورانه مفهوم‌پردازی شود؛ سامانه‌ای که در آن فناوری نقش افزاینده و مقوم قضاوت حرفه‌ای انسان را ایفا می‌کند، نه جایگزین آن.

چارچوب مفهومی ارائه‌شده در این مطالعه خروجی نظام‌مند تحلیل کیفی متون پیشین است. این چارچوب از طریق فرآیند کدگذاری سه‌مرحله‌ای (باز، محوری و انتخابی) و سنتز مقولات مستخرج از فراتلفیق سه موج تحولی کانون‌های ارزیابی تدوین شده است. به‌گونه‌ای که مقولات محوری مرتبط با فشارهای دیجیتال‌سازی و یادگیری ماشینی، بلوک پیشران‌ها و ورودی‌ها را شکل دادند؛ چالش‌های شناسایی‌شده در زمینه پایایی، عدالت الگوریتمی و تبیین‌پذیری، هسته‌ی مکانیسم پردازش و حکمرانی را برساختند؛ و در نهایت، پیامدهای استخراج‌شده از متون در سه سطح داوطلب، استانداردها و اکوسیستم، بلوک نتایج راهبردی را پیکربندی کردند. بر پایه این سنتز روش‌مند، چارچوب مفهومی یکپارچه کانون ارزیابی هوشمند انسان‌محور (HITL-SAC) مطابق شکل ۱ پیشنهاد می‌گردد.



شکل ۱. چارچوب مفهومی یکپارچه کانون ارزیابی هوشمند انسان‌محور

این اکوسیستم در قالب یک جریان پردازشی پویا و دارای بازخورد، از سه بلوک اصلی تشکیل شده است که در ادامه با جزئیات تشریح می‌گردند:

### ۱. پیشران‌ها و ورودی‌های سیستم

این بلوک تبیین‌کننده نیروهای محرکه‌ای است که سازمان‌ها را به سمت گذار از ارزیابی سنتی به هوشمند سوق می‌دهند. این پیشران‌ها در سه بعد قابل واکاوی هستند:

- **فشارهای محیطی و گذار ساختاری:** کانون‌های ارزیابی تحت فشار فزاینده برای مقیاس‌پذیری و تاب‌آوری در برابر اختلالات محیطی، ناگزیر به پلتفرمی‌سازی شده‌اند. این گذار ساختاری به سمت تحول دیجیتال، پاسخی راهبردی به افزایش فشارهای مرتبط با کارایی عملیاتی و لزوم ارتقای تجربه داوطلب در محیط‌های رقابتی جذب استعداد است.
- **موتور تکنولوژیک و ابزاری:** پیشرفت‌های فناوریانه به‌عنوان نیروی محرکه اصلی عمل می‌کنند. ورود الگوریتم‌های یادگیری ماشین (ML) به فرآیند نمردهی تمرین‌ها، پاسخی مستقیم به محدودیت‌های شناختی و هزینه‌های بالای ارزیابان انسانی است. هم‌زمان، جهش ابزارهای

تعاملی مبتنی بر هوش مصنوعی (نظیر مدل‌های زبانی بزرگ LLM و یادگیری تقویتی RL)، امکان حرکت از آزمون‌های ایستا به شبیه‌سازی‌های پویای رفتاری را فراهم آورده است.

- بلوغ و پذیرش: پیاده‌سازی این سیستم‌ها نیازمند بلوغ پیشینی کانون‌های ارزیابی مجازی و تمرکز نهادی بر تحقق عدالت جمعیت‌شناختی است که به‌عنوان شرط اساسی پذیرش شناختی و نهادی این فناوری‌ها در سازمان عمل می‌کند.

## ۲. مکانیسم پردازش و حکمرانی

این بلوک، هسته مرکزی و فیلتر مشروعیت‌بخش چارچوب است. ورودی‌های سیستم تنها در صورتی به نتایج راهبردی مطلوب منجر می‌شوند که از چهار لایه پردازشی و نظارتی زیر با موفقیت عبور کنند:

- الزامات روان‌سنجی نوین: در زیست‌بوم هوشمند، مشروعیت سیستم قویاً مشروط به حفظ روایی، تطابق با تحلیل شغل و قابلیت دفاع حقوقی است. این امر نیازمند بازپیکربندی شواهد روایی و پایایی در محیط‌های پویا (مبتنی بر داده‌های ارگانیک) است. همچنین، سیستم باید قادر باشد بر چالش ناهمگنی روان‌سنجی در مواجهه با متون تولیدشده توسط LLM غلبه کرده و سازه اصلی مورد سنجش را از برداشت‌های الگوریتمی تفکیک نماید.
- کیفیت داده و تبیین‌پذیری: کیفیت نمره‌دهی تماماً به کیفیت داده‌های ورودی و توانمندی سیستم در کنترل خطاهای فناورانه وابسته است. در این لایه، اعتبار ارزیابی‌ها مشروط به رفتار محوری، تبیین‌پذیری منطق تصمیم‌گیری الگوریتم و اعتبارسنجی برون‌نمونه‌ای است تا از تعمیم‌پذیری نتایج اطمینان حاصل شود.
- حکمرانی اخلاقی و حقوقی: با توجه به تشدید حساسیت‌های اخلاقی ناشی از پردازش داده‌های ضمنی و غیرقابل کنترل، پذیرش نهایی سیستم در گرو تأمین عدالت ادراک‌شده، شفافیت و تبیین‌پذیری است. این امر استقرار مکانیزم‌های سخت‌گیرانه حکمرانی، شامل ممیزی عدالت، حفظ حریم خصوصی و مدیریت ریسک تقلب را گریزناپذیر می‌سازد.
- ظرفیت و مداخله انسانی (انسان در چرخه): نقطه ثقل چارچوب مفهومی حاضر در این بخش تبلور می‌یابد. استقرار این مدل مشروط به طراحی انسان‌محور و حفظ حضور انسان در چرخه تصمیم است. نقش ظرفیت نهادی منابع انسانی در اینجا، تبدیل دیجیتال‌سازی صرف به یک فرصت راهبردی و دارای قابلیت دفاع است که مشروعیت نهایی ابزارها را تضمین می‌کند.

## ۳. نتایج و پیامدهای راهبردی

در صورت استقرار صحیح مکانیسم‌های پردازش و حکمرانی، خروجی‌های این اکوسیستم در سه سطح کلان متبلور خواهند شد:

- پیامدهای سطح داوطلب و برند: پیاده‌سازی هوشمندسازی می‌تواند به قطبی‌شدن واکنش‌های داوطلبان منجر شود؛ با این حال، استقرار شفاف و عادلانه آن اثری مستقیم و مثبت بر ارتقای جذابیت کارفرما خواهد داشت.
  - پیامدهای سطح استانداردسازی: در این سطح، کانون ارزیابی مجازی به‌عنوان زیرساخت اصلی اجرا و سنجش تفکیکی تثبیت می‌گردد. همچنین روایی، انصاف و قابلیت دفاع سیستم نهادینه شده و هوشمندسازی محدود اما عملیاتی (تحت نظارت انسانی) به یک استاندارد پذیرفته‌شده سازمانی تبدیل می‌شود.
  - پیامدهای نهایی اکوسیستم: در نهایت، هم‌گرایی VAC، هوش مصنوعی و اتوماسیون، به شکل‌گیری اکوسیستم یکپارچه کانون ارزیابی هوشمند منتج می‌شود. مهم‌ترین دستاورد این دگرگونی، بازتعریف نقش واحد منابع انسانی از یک مجری عملیاتی وظیفه‌محور، به معمار سیستم و ناظر ارشد ریسک و عدالت در سازمان است.
- در نهایت، همان‌گونه که در مسیرهای بازخورد شکل (۱) نیز مشهود است، این چارچوب ماهیتی خطی ندارد، بلکه یک چرخه یادگیری تکاملی است که در آن پیامدهای راهبردی، به‌طور مستمر بلوغ نهادی و فشارهای محیطی را بازتعریف کرده و موجب ارتقای پویای سیستم در گذر زمان می‌گردند.
- تیین جایگاه چارچوب پیشنهادی در ادبیات پژوهش نشان می‌دهد که این مدل در عین نوآوری، پیوندهای نظری وثیقی با تحقیقات پیشین داشته و هم‌زمان در مرزهای مشخصی از آن‌ها متمایز می‌گردد. در لایه اشتراکات نظری، این چارچوب هم‌سو با بنیان‌های رفتاری کلاسیک (Yilmaz, 2023)، همچنان مشروعیت سیستم را در گرو صیانت از روایی ملاک‌محور و تحلیل دقیق شغل دانسته و توان پیش‌بینی عملکرد را به‌عنوان کارکرد اصلی کانون حفظ می‌کند. افزون بر این، در تطابق با نظریه سیستم‌های اجتماعی-فناورانه (Cherep et al., 2022) و پارادایم HRM 5.0 (Kravchuk, 2025)، کانون ارزیابی هوشمند در این مطالعه نه یک ابزار فنی صرف، بلکه یک اکوسیستم اجتماعی-فناورانه و انسان‌محور مفهوم‌پردازی شده است که موفقیت آن نیازمند هم‌ترازی دقیق میان منطق محاسباتی ماشین و ارزش‌های اخلاقی انسان است.
- با این حال، نقطه افتراق و چرخش پارادایمیک این چارچوب با رویکردهای سنتی، در گذار از نظریه ثبات رفتاری مشاهده‌محور به نظریه داده‌محوری و تصمیم‌گیری الگوریتمی (Faqihi & Miah, 2023) نهفته است؛ جایی که ارزیابی از یک رویداد مقطعی و متکی بر شهود، به جریانی مستمر، ساختار یافته و مبتنی بر هوش ترکیبی بدل می‌شود. برخلاف رویکردهای تقلیل‌گرایانه در موج‌های اولیه دیجیتال‌سازی که بر حذف ارزیاب انسانی اصرار داشتند، چارچوب HITL-SAC با اتکا بر اصول اخلاق

دیجیتال، رویکرد انسان در چرخه (Makedon & Myachin, 2025) را به‌عنوان هسته مرکزی حکمرانی خود معرفی می‌کند. در این چارچوب، فناوری در نقش هوش افزوده ظاهر شده و با بازتعریف هویت مدیریت منابع انسانی، این واحد را از یک نهاد مجری به متولی راهبردی زیست‌بوم، دیده‌بان انصاف و مدیر مخاطرات اخلاقی در تعامل انسان و ماشین بدل می‌سازد.

در جمع‌بندی، نتایج این پژوهش نشان می‌دهد که مسیر آینده کانون‌های ارزیابی نه در خودکارسازی حداکثری، بلکه در هوشمندسازی مسئولانه و انسان‌محور نهفته است. توسعه و استقرار کانون‌های ارزیابی هوشمند بدون توجه هم‌زمان به روایی، انصاف و حکمرانی، می‌تواند به تضعیف اعتماد عمومی و مشروعیت نهادی منجر شود، حتی اگر کارایی کوتاه‌مدت افزایش یابد. بر این اساس، توصیه می‌شود سیاست‌گذاران، نهادهای اجرایی و سازمان‌های بزرگ، پیش از گسترش کاربرد این ابزارها، چارچوب‌های مشخصی برای اعتبارسنجی، ممیزی عدالت و نظارت انسانی تدوین و اجرا کنند. پژوهش‌های آینده می‌توانند با تمرکز بر مطالعات میدانی، آزمون تجربی چارچوب مفهومی ارائه‌شده و بررسی اثرات بلندمدت کانون‌های ارزیابی هوشمند بر کیفیت تصمیم‌های منابع انسانی، این مسیر نظری را تکمیل و غنی‌تر سازند.

## قدردانی

این بخش شامل قدردانی از افراد و نهادهایی است که محقق را در نوشتن مقاله یاری نموده‌اند.

## Acknowledgments

This section includes acknowledgments for individuals and institutions that have assisted the researcher in writing the article.

## تعارض منافع

تعارض منافع نداریم.

## Conflict of interest

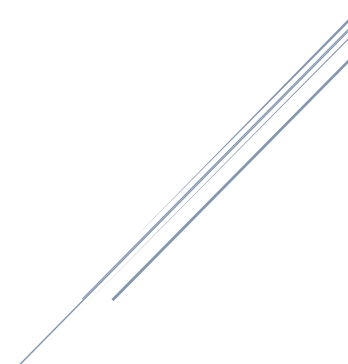
The authors declare that they have a conflict of interest.

## حمایت مالی

نویسنده(گان) هیچ‌گونه حمایت مالی برای انجام این پژوهش، نگارش و یا انتشار این مقاله دریافت نکرده‌اند.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.



## منابع

- گلشاهی، بهنام و خوشنویس زارچ، اکبر. (۱۴۰۴). معماری کاربست فناوری واقعیت ترکیبی در فرایند جذب و گزینش سرمایه انسانی ارتش جمهوری اسلامی ایران. مدیریت هوشمند سرمایه انسانی، ۲(۵)، ۳۳-۶۲. <https://doi.org/10.22034/imhr.2025.548889.1040>
- میرزانژاد، امیرحسین و حاجی مولانا، عماد. (۱۴۰۴). مدل‌سازی و تحلیل روان‌شناختی متقاضیان شغل با استفاده از پردازش زبان طبیعی و تحلیل چهره. مدیریت هوشمند سرمایه انسانی، ۲(۵)، ۳۲-۱. <https://doi.org/10.22034/imhr.2025.535383.1033>
- Adabi, M. (2025). Providing a model of diversity and inclusion in human resource management with a focus on digital transformation tools. *International Journal of Innovation Management and Organizational Behavior*. <https://doi.org/10.61838/kman.ijimob.5.4.16>
- Alrababah, S. A., & Molnár, G. (2021). The evolution of technology-based assessment: Past, present, and future. *International Journal of Learning Technology*, 16(2), 134-157. <https://doi.org/10.1504/IJLT.2021.117765>
- Anthamatten, A., & Holt, J. E. (2024). Integrating artificial intelligence into virtual simulations to develop entrustable professional activities. *The Journal for Nurse Practitioners*, 20(9), 105192. <https://doi.org/10.1016/j.nurpra.2024.105192>
- Ariesthiawati, J. (2022, October). The effectiveness of virtual assessment center platform. In *2022 international conference on assessment and learning (ICAL) (pp. 1-3)*. IEEE. <https://doi.org/10.1109/ICAL50372.2022.10075569>
- Avni, E., & Luria, G. (2022). Validity and reliability of virtual versus face-to-face assessment center. In *Academy of Management Proceedings*. 1. <https://doi.org/10.5465/AMBPP.2022.14415abstract>
- Avni, E., & Luria, G. (2025). Virtual assessment centres versus face-to-face assessment centres: psychometric properties. *European Journal of Work and Organizational Psychology*, 34(3), 339-347. <https://doi.org/10.1080/1359432X.2025.2457780>
- Baranyi, V. (2025). Systematic literature review on the digital transformation of the personnel selection process. *German Journal of Human Resource Management*. <https://doi.org/10.1177/23970022251363012>
- Behrend, T. S., & Landers, R. N. (2025). Participant Interactions with Artificial Intelligence: Using Large Language Models to Generate Research Materials for Surveys and Experiments. *Journal of Business and Psychology*, 1-23. <https://doi.org/10.1007/s10869-025-10035-6>
- Breil, S. M., Lievens, F., Forthmann, B., & Back, M. D. (2023). Interpersonal behavior in assessment center role-play exercises: Investigating structure, consistency, and effectiveness. *Personnel Psychology*, 76(3), 759-795. <https://doi.org/10.1111/peps.12507>

- Bronkhorst, P. V., & Becker, J. (2024). Use of artificial intelligence in leadership competency development and selection: An empirical study. *Consulting Psychology Journal*. <https://psycnet.apa.org/doi/10.1037/cpb0000288>
- Burgues, M., Goujet, R., & Zarak, J. (2024). Learning Soft Skills with an AI-Based Simulation Role-Play: A Literature Review. *EDULEARN24 Proceedings*, 6285-6293. <https://doi.org/10.21125/edulearn.2024.1484>
- Campion, M. A. (2025). Can Legal and Professional Personnel Selection Principles be Met With Machine Learning (Artificial Intelligence)?. *Human Resource Management*. <https://doi.org/10.1002/hrm.70025>
- Campion, M. A., & Campion, E. D. (2023). Machine learning applications to personnel selection: Current illustrations, lessons learned, and future research. *Personnel Psychology*, 76(4), 993-1009. <https://doi.org/10.1111/peps.12621>
- Chakraborty, T., & Sharada, V. S. (2025). Optimizing Talent Acquisition: The Synergy of AI and Human Expertise in Creating Enhanced Candidate Experience. In *Modern Trends and Future Innovation in Human Resource Management* (pp. 259-276). IGI Global. <https://doi.org/10.4018/979-8-3693-6402-4.ch009>
- Chang, K., Carrim, N. M., Gupta, M., Cheng, K., & Sandland, S. (2024). Digitalisation of Personnel Recruitment and Selection. In *HRM 5.0: Unpacking the Digitalisation of Human Resource Management* (pp. 87-111). [https://doi.org/10.1007/978-3-031-58912-6\\_5](https://doi.org/10.1007/978-3-031-58912-6_5)
- Cherep, A., Voronkova, V., & Androsova, O. (2022). Transformational changes in organizational management and human resources in the digital age. *Baltic Journal of Economic Studies*. <https://doi.org/10.30525/2256-0742/2022-8-3-210-219>
- Dorsey, D. W., & Michaels, H. R. (2022). Validity arguments meet artificial intelligence in innovative educational assessment. *Journal of Educational Measurement*, 59(3), 267-271. <https://doi.org/10.1111/jedm.12330>
- Faqih, A., & Miah, S. (2023). Artificial intelligence-driven talent management system: Exploring the risks and options for constructing a theoretical foundation. *Journal of Risk and Financial Management*. <https://doi.org/10.3390/jrfm16010031>
- Golshahi, B., & Khoshnevis Zarch, A. (2025). Architecture of applying mixed reality technology in the human capital recruitment and selection process of the Islamic Republic of Iran Army (AJA). *Intelligent Management of Human Capital*, 2(5), 33-62. [In Persian] <https://doi.org/10.22034/imhr.2025.548889.1040>
- Grunenberg, E., Peters, H., Francis, M. J., Back, M. D., & Matz, S. C. (2024). Machine learning in recruiting: predicting personality from CVs and short text responses. *Frontiers in Social Psychology*, 1. <https://doi.org/10.3389/frsps.2023.1290295>
- Grunenberg, E., Stachl, C., Breil, S. M., Schäpers, P., & Back, M. D. (2025). Predicting and Explaining Assessment Center Judgments: A Cross-Validated Behavioral Approach to Performance Judgments in Interpersonal Assessment Center

Exercises. *Human Resource Management*, 64(2), 423-445.

<https://psycnet.apa.org/doi/10.1002/hrm.22252>

- Guenole, N., Svensson, C., Wille, B., Aloyan, K., & Saville, P. (2022). A European perspective on psychometric measurement technology. *Technology and Measurement around the Globe*. <https://doi.org/10.1017/9781009099813.013>

- Hickman, L., Herde, C. N., Lievens, F., & Tay, L. (2023). Automatic scoring of speeded interpersonal assessment center exercises via machine learning: Initial psychometric evidence and practical guidelines. *International Journal of Selection and Assessment*, 31(2), 225-239. <https://doi.org/10.1111/ijasa.12418>

- Hickman, L., Saef, R., Ng, V., Woo, S. E., Tay, L., & Bosch, N. (2024). Developing and evaluating language-based machine learning algorithms for inferring applicant personality in video interviews. *Human Resource Management Journal*, 34(2), 255-274. <https://doi.org/10.1111/1748-8583.12356>

- Kaldaras, L., Akaeze, H. O., & Reckase, M. D. (2024, August). Developing valid assessments in the era of generative artificial intelligence. In *Frontiers in education*, 9. <https://doi.org/10.3389/feduc.2024.1399377>

- Khan, S. N., & Khan, S. U. (2024). A review of industry and AI impact on human resource management. *Journal of Advancement in Computing*, 2(1). <https://doi.org/10.36755/jac.v2i1.60>

- Koenig, N., Tonidandel, S., Thompson, I., Albritton, B., Koohifar, F., Yankov, G., ... & Newton, C. (2023). Improving measurement and prediction in personnel selection through the application of machine learning. *Personnel Psychology*, 76(4), 1061-1123. <https://doi.org/10.1111/peps.12608>

- Kravchuk, O. (2025). *The digital ecosystem of human resource management 5.0: A new paradigm of human-centered development*. <https://doi.org/10.54929/2786-5738-2025-20-04-07>

- Liou, G., Bonner, C. V., & Tay, L. (2022). A psychometric view of technology-based assessments. *International Journal of Testing*, 22(3-4), 216-242. <https://doi.org/10.1080/15305058.2022.2070757>

- Lumdee, N. (2025, April). Needs Analysis for Developing Intelligent Digital Platform for Course Learning Outcome Assessment of Information Technology Program in Higher Education. In *2025 IEEE International Conference on Cybernetics and Innovations (ICCI)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICCI64209.2025.10987370>

- Makedon, V., & Myachin, V. (2025). Transformation of personnel management under the influence of artificial intelligence technologies. *Market Infrastructure*. <https://doi.org/10.32782/infrastructure83-38>

- Mayring, P. (2014). Qualitative content analysis: theoretical foundation, basic procedures and software solution.

- Mihaljević, H., Müller, I., Dill, K., Yollu-Tok, A., & von Grafenstein, M. (2024). More or less discrimination? Practical feasibility of fairness auditing of technologies

- for personnel selection. *AI & SOCIETY*, 39(5), 2507-2523. <https://doi.org/10.1007/s00146-023-01726-w>
- Mirzanezhad, A. H., & Haji Maulana, E. (2025). Psychological modeling and analysis of job applicants using natural language processing and facial analysis. *Intelligent Management of Human Capital*, 2(5), 1–32. [In Persian] <https://doi.org/10.22034/imhr.2025.535383.1033>
- Modi, T. B. (2023). Artificial Intelligence Ethics and Fairness: A study to address bias and fairness issues in AI systems, and the ethical implications of AI applications. *Revista Review Index Journal of Multidisciplinary*, 3(2), 24-35. <https://doi.org/10.31305/rrijm2023.v03.n02.004>
- Nalini, M. (2024). Transformative horizons: Navigating the evolution of HR through AI and cloud technologies. *Shanlax International Journal of Management*, 11(Special Issue 1), 1–12. <https://doi.org/10.34293/management.v11i1S1-Jan.7138>
- Pandey, S., & Mishra, A. (2025). The AI-driven evolution of human resource management: Trends, challenges, and future directions. *International Journal for Multidisciplinary Research*, 7(3). <https://doi.org/10.36948/ijfmr.2025.v07i03.46823>
- Panigrahi, S. (2025). AI in HR: Impact of artificial intelligence on transforming human resources. *The American Journal of Management and Economics Innovations*. <https://doi.org/10.37547/tajmei/Volume07Issue07-06>
- Pathak, G., & Pandey, D. (2025). AI Agents in Recruitment: A Multi-Agent System for Interview, Evaluation, and Candidate Scoring. *Evaluation, and Candidate Scoring*. <https://dx.doi.org/10.2139/ssrn.5242372>
- Purba, Y. Y. (2024). Implications of technological evolution on human resource management strategies and legal compliance. *Journal Minfo Polgan*, 12(2), 101–115. <https://doi.org/10.33395/jmp.v12i2.13337>
- Raime, S., Rahman, N. A., Hashim, R. A., & Shamsudin, M. F. (2025). AI management and emotional disengagement: A conceptual overview. *International Journal of Academic Research in Business and Social Sciences*. <http://dx.doi.org/10.6007/IJARBSS/v15-i8/26186>
- Rajeswari, A. M., Ahamed, R., & Vasanth, P. (2025, February). InterviewIQ: AI-Powered Chatbot for Upskilling Candidates in Technical and HR Interviews. In *2025 International Conference on Intelligent Control, Computing and Communications (IC3)* (pp. 392-397). IEEE. <https://doi.org/10.1109/IC363308.2025.10956589>
- Raza, H. (2024). Ai-driven assessment: Reliability, bias, and ethical implications. *AI EDIFY Journal*, 1(2), 36-47.
- Rezzani, A., Caputo, A., & Cortese, C. G. (2020). An analysis of the literature about the application of Artificial Intelligence to the Recruitment and Personnel Selection. *Bollettino di Psicologia Applicata*, 25-33. <https://doi.org/10.26387/bpa.289.3>
- Ringelband, O., & Warneke, C. (2025). Some ethical and legal issues in using artificial intelligence in personnel selection. *Consulting Psychology Journal*. <https://doi.org/10.1037/cpb0000289>

- Saputra, I., Kurniawan, A., Yanita, M., Putri, E. Y., & Mahniza, M. (2024). The evolution of educational assessment: How artificial intelligence is shaping the trends and future of learning evaluation. *The Indonesian Journal of Computer Science*, 13(6). <https://doi.org/10.33022/ijcs.v13i6.4465>
- Sarumathi, S., Gowthaman, R. L., Sabareesh, M. B., & Sultana, A. (2025, February). AI-Enhanced HR Interview Simulation for Realistic Candidate Assessment. *In 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 1089-1095). IEEE. <https://doi.org/10.1109/IDCIoT64235.2025.10915193>
- Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students (8th ed.)*. Pearson Education.
- Simelane, R. M., & Pillay, P. (2024). The Evolution of Assessment Methods in Higher Education Due to the Shift to Remote Learning: A Case Study. *Research in Educational Policy and Management*, 6(2), 393-414. <https://doi.org/10.46303/repam.2024.41>
- Theres, C., & Strohmeier, S. (2023). Consolidating the theoretical foundations of digital human resource management acceptance and use research: A meta-analytic validation of UTAUT. *Management Review Quarterly*, 74(8), 2683–2715. <https://doi.org/10.1007/s11301-023-00367-z>
- Thompson, I., Koenig, N., Mracek, D. L., & Tonidandel, S. (2023). Deep learning in employee selection: Evaluation of algorithms to automate the scoring of open-ended assessments. *Journal of Business and Psychology*, 38(3), 509-527. <https://doi.org/10.1007/s10869-023-09874-y>
- Tippins, N. T., Oswald, F. L., & McPhail, S. M. (2021). Scientific, legal, and ethical concerns about AI-based personnel selection tools: a call to action. *Personnel Assessment and Decisions*, 7(2), 1. <https://doi.org/10.31234/osf.io/6gczw>
- Tuitt, T. A., Addison, L. M., & Hosein, P. (2025, February). Generative AI and Multi-Agent Systems Approach to Psychometric Evaluation for Human Resource Management and Talent Acquisition. *In 2025 7th International Symposium on Computational and Business Intelligence (ISCBI)* (pp. 108-112). <https://doi.org/10.1109/ISCBI64586.2025.11015385>
- Usmani, Y. Z., Petruzzello, G., Rizzo, B., & Mariani, M. G. (2025). Assessment centers in the virtual age: validity and fairness in gender and age. *Central European Management Journal*. <https://doi.org/10.1108/CEMJ-04-2024-0118>
- Weinert, S., Gunther, E., Ruger-Muck, E., & Raab, G. (2020). Artificial intelligence in personnel selection and its influence on employer attractiveness. *Marketing Science & Inspirations*, 15(3), 22–35. <https://doi.org/10.46286/msi.2020.15.3.2>
- Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2020). Personnel selection in the digital age: A review of validity and applicant reactions, and future research challenges. *European Journal of work and organizational psychology*, 29(1), 64-77. <https://doi.org/10.1080/1359432X.2019.1681401>

- Yilmaz, A. (2023). The evolution of human resource management in managerial thinking. *Toplum Ekonomi ve Yönetim Dergisi*, 3(2), 45–60. <https://doi.org/10.58702/teyd.1289550>
- Zhang, T., Koutsoumpis, A., Ostrom, J. K., Holtrop, D., Ghassemi, S., & De Vries, R. E. (2024). Can large language models assess personality from asynchronous video interviews? A comprehensive evaluation of validity, reliability, fairness, and rating patterns. *IEEE Transactions on Affective Computing*, 15(3), 1769-1785. <https://doi.org/10.1109/TAFFC.2024.3374875>