Unveiling the Benefits, Limitations, and Mitigation of Bias in Artificial Intelligence within Organizational Contexts-A Systematic Review

Xinyi Li

The University of Hong Kong, Hong Kong, 999077, China

m15264257168@163.com

Abstract. This essay offers an extensive discussion of the advantages, drawbacks, and potential remedies for reducing bias in artificial intelligence (AI) systems used in corporate settings. As AI technology continues to permeate various sectors, concerns regarding bias have emerged as a critical issue. The presence of bias in AI systems can perpetuate discrimination and social disparities, leading to ethical, legal, and social implications for organizations. By examining existing literature and exploring real-world examples, this study aims to provide a comprehensive understanding of the advantages and drawbacks of AI in organizational settings. Furthermore, it delves into potential strategies and solutions to mitigate bias. Through this systematic review, organizations can gain valuable insights into the impact of AI and interactions with human capital, thus developing informed strategies to address this pressing challenge.

Keywords: Artificial intelligence, machine learning, organization, bias

1 Introduction.

The applications of artificial intelligence (AI) has transformed many facets of contemporary civilization, including decision-making and managerial processes. As AI systems become increasingly integrated into organizational frameworks, concerns surrounding potential limitations have gained significant attention. Bias in AI systems can perpetuate unfairness, discrimination, and social disparities, raising ethical, legal, and social implications. In spite of the increasing prominence of the benefits brought to companies, a relatively limited number of research offers a comprehensive analysis of the impact on organizations and employees. This paper aims to explore the benefits from multiple perspectives, including job application, recruitment & selection, interview & assessment, management & control, and innovation process—additionally, imitations of current AI technology in organizations while delving into potential solutions to mitigate bias. By examining current research and industry practices systemically, this review seeks to shed light on the complexities of AI implementation and provide insights into how organizations can navigate this critical challenge.

H. Kassim et al. (eds.), *Proceedings of the 2023 8th International Conference on Modern Management and Education Technology (MMET 2023)*, Advances in Social Science, Education and Humanities Research 798,
https://doi.org/10.2991/978-2-38476-146-3_9

2 Definition and Categories of Artificial Intelligence

Artificial intelligence is a subject of research in computer science. The phrase "artificial intelligence" first came into use in the 1950s to describe the basic concept that machines could display human intelligence [1]. The objective is to develop machines that can think as humans execute, including learning, reasoning, and self-correction [2]. Most artificial intelligence (AI) performs at least one of the following seven tasks, according to Castro and New [3]: monitoring, discovering, predicting, interpreting, interacting with the physical environment, interacting with humans, and interacting with other machines.

Machine learning, or the capability of software to improve its own performance by analyzing interactions with the outside world, appears to be the branch of AI that is expanding the most rapidly [4]. In this segment, a computer learns from data sets to perform functions rather than just carrying out the specific tasks it was programmed to do. The main driver of the AI boom has been machine learning, which has been used in everything from search and product recommendation engines to speech recognition, fraud detection, and image interpretation systems [5]. There is usually an overlap between robots and AI, blurring the lines between different technologies. When AI is stated, robotics is frequently the first thing that comes to mind as a sub-segment of AI [4].

3 General Review of AI in Organization

One of the early studies examining artificial intelligence was a computer simulation of human cognitive and social processes [6]. Later, Dean, Yoon, and Susman [7] reported findings on how organizational structure and advanced manufacturing technology (AMT) interact. Research carried out by George and his team [8] considered how corporations manage massive data. They make comments on service level agreements (SLAs), which specify the type and caliber of IT services, and they point out that massive data-sharing agreements typically have a shoddy structure and are informal. They discussed the approaches for extensive data analysis and asserted that it is simple to obtain erroneous correlations when utilizing conventional statistical tools for extensive data analysis. The study also discussed the application of big data to management and behavioral science. One book covered a range of topics discussing the data science revolution and organizational psychology [9]. People consider slowly generated forecasts from algorithms less accurate and are less ready to depend on them [10]. Slower response times indicate that both humans and algorithms are working harder. For humans and machines, the link between perceived effort and prediction skill varies. These findings highlight the intricate dynamics and processes that underlie people's perceptions of the reliability of algorithmic (and human) predictions. Integrating complementarity and role theory, a study [11] demonstrated that working with intelligent machines has advantages and disadvantages. However, more importantly, conscientious (i.e., orderly) employees are less likely to benefit from using them. Through the use of ML, one research intends to improve the precision of interest inventory-based job choice prediction [12]. According to the results, the machine-learning augmented method was more accurate overall than the standard profile method for predicting occupational membership (for employed participants) and vocational goals (for jobless participants). The machine-learning enhanced technique underpredicted job categories with low base rates while being particularly predictive of job categories with high base rates.

One research focuses on a specific time period, which used discontinuous growth modeling (DGM) to examine how COVID-19 and the ensuing instructions to stay at home altered the public's emotional trajectory toward working from home [13] The findings showed that stay-at-home orders had an impact on the public's emotional trajectories both immediately (i.e., intercept change) and longitudinally (i.e., slope change). The emotion trajectories were not considerably altered by the daily fresh COVID-19 case counts. Despite a number of other research exploring the application of artificial intelligence during the pandemic [14][15], most of them focused on its usage for diagnosis, medication, and other medical purposes. There are limited studies on the influence of using artificial intelligence on companies during that time.

There is research focusing attention on specific industries. A study specializing in the medical field demonstrates the hindering effects of opacity that experts encountered when utilizing AI techniques [16]. Professionals in all three departments examined reported feeling more uncertain as a result of this opacity. Barrett's team [17] investigates the impact of robotic advances on the border dynamics of three distinct occupational groups working in a hospital pharmacy: pharmacists, technicians, and assistants. They discovered that interaction with the hybrid and dynamic materiality of the robot over time altered border relations between the three occupational groups, with significant and paradoxical ramifications for the abilities, jurisdictions, positions, and visibility of pharmacy workers. Another research focuses on the architecture industry. Boland's team [18] studied the use of digital three-dimensional (3-D) representations in the architecture of Frank O. Gehry's buildings and found that numerous heterogeneous firms developed a variety of breakthroughs, each of which left a trail of innovation. The fire engineering case study demonstrates how such programs leverage simulation technology as a boundary object to stimulate innovation in a new organizational sector [19]. Engineers employ simulation technologies to bring about significant improvements in fire control and management, such as the use of elevators to evacuate buildings during emergencies. A framework is constructed to investigate how decisions might be achieved and tensions resolved among various diverse and discordant parties attempting to share a common appreciation of negotiated futures. Lyu and Liu [20] computed a systematic measure of the rising digital technology intensity in work skill needs. They demonstrated that Artificial Intelligence is the most valuable in the energy sector from both the employee's and the employer's perspective. Artificial intelligence, in particular, increases the average pay of the adopted energy firm and the local labor market. Meanwhile, Artificial Intelligence has the most significant impact on the performance of energy companies. Consequently, findings suggest that energy companies should deliberately enhance the necessity for Artificial Intelligence when employing new employees.

4 Benefits of Employing Artificial Intelligence

Kellogg et al. [21] synthesize the interdisciplinary research on algorithms at work using Edwards' [22] perspective of "contested terrain," in which managers apply production technology to maximize the value of labor while workers resist. They discover that algorithmic control in the workplace occurs via six major processes, which we refer to as the "6 Rs"—employers can use algorithms to direct workers by restricting and suggesting, evaluate workers by recording and rating, and discipline workers by replacing and rewarding.

4.1 Job Application

One project uses machine learning techniques to analyze information from job application forms [23], such as past job descriptions and stated reasons for changing jobs, in order to create interpretable measures of the value of work experience, tenure history, involuntary turnover history, history of avoiding bad jobs, and history of approaching better jobs. According to a simulation of the effects of machine learning [24] the potential of machine learning for staff selection is unlikely to be fulfilled in selection procedures that prioritize scale composites from previously validated psychometric tests. Instead, it will be implemented in novel data forms, including text, image, audio, video, and behavioral traces, or in unorthodox design circumstances like using a single object to infer numerous traits.

4.2 Recruitment & Selection

Several papers are investigating the use of artificial intelligence during recruitment. The narrative data gathered from candidates in written or oral responses to assessment questions (referred to as constructed responses) can be scored by ML with the same accuracy and reliability as human judges but much more quickly [25]. This makes it easier to include such responses in personnel selection and frequently improves validity with minimal adverse effects. Additionally, algorithms can be developed to simultaneously predict various outcomes (such as productivity and turnover) and generalize across assessment questions. Even job analysis using ML has been shown to be more effective by identifying knowledge and skill requirements from job descriptions. Another two-year ethnographic research [26] focuses on how programmers dealt with tension when creating a machine learning system to assist a substantial international company's hiring procedure for job hopefuls. Contrary to widespread perceptions that place domain knowledge and machine learning in opposition to one another, this study emphasizes their interconnectedness and, as a result, demonstrates the dialectical nature of developing ML. One of the perennial problems in hiring is the diversity-validity conundrum. However, new data analysis methods and technological developments in the areas of machine learning and industrial-organizational psychology are presenting creative solutions to this conundrum. Given these quick developments, Rottman and his colleagues [27] provide a framework for reducing group differences by combining analytical techniques frequently employed in these two disciplines. Then, the findings show the efficacy of two methods—iterative predictor removal and multispecialty optimization—for lowering group differences while preserving validity.

4.3 Interview & Assessment

Despite the lack of studies on the validity, reliability, and generalizability of automated video interviews (AVIs), businesses are increasingly using them to screen job applications. In the study conducted by Hickman has his team [28], researchers created AVIs that assess the Big Five personality traits using verbal, paraverbal, and nonverbal actions taken from video interviews. Training the AVI personality assessments on interviewer reports instead of self-reports resulted in better evidence of validity. Another study uses collaborative filtering recommender system algorithms, a type of machine learning technique [29], to examine the comparison between satisfaction with the existing job and potential job alternatives, which is inherent in theories about individual turnover decisions. The quantity of more desirable job alternatives and the quality of job alternatives are two operationalizations produced when those anticipated ratings are compared to the employee's current job satisfaction.

4.4 Management & Control

Data are now more than just a part of administrative and management tasks; they are a widespread resource and a channel through which organizations learn about and respond to the uncertainties they face. Alaimo and Kallinikos [30] hypothesize that current technology advancements not only reframe and broaden the traditional roles of data as management and control tools but also strengthen those roles. Leonardi [31] investigated what issues can occur when interpreting this implicit knowledge over time and location and how humans might solve these issues.

4.5 Innovation

A recent study on the impact of AI on employee creativity explored how AI can help with the sequential division of labor inside enterprises [32]. According to the study, AI assistance alters work design by enhancing employees' interactions with more important clients. Although AI can enhance creativity in workers, this desirable result is skill-biased. This result aligns with the economic theory of skill-biased technology in which production technologies allow skilled labor to become more productive than unskilled labor [33]. For businesses to learn successfully as a whole, Sturm and his colleagues [34] investigate how they may coordinate human learning and ML. Based on a number of agent-based simulations, they discover that ML can decrease an organization's need for human explorative learning that aims to uncover new ideas, that human modifications to ML systems are generally beneficial, but that this effect can diminish or even become harmful under certain circumstances; and that reliance on ML system knowledge can facilitate organizational learning in turbulent environments. However, doing so necessitates making significant initial investments in these systems and adequately integrating them with people.

5 Limitations of Artificial Intelligence

A growing number of research literature have considered the potential limitations of artificial intelligence. One paper pointed out that AI is a prediction technology, and predictions cannot be valued without knowing how payoffs occur [35]. Another study shows that while ML may help with finding minor improvements, it is unlikely to be able to tackle the problem of the unfavorable impact caused by subgroup differences in personnel selection [36]. Though machine learning appears to be unrestrained by human cognitive constraints and inflexibility, it is not genuinely sentient learning and instead relies on formal statistical analysis for decision-making [37]. A study contends that the distinctions between human learning and machine learning risk reducing within-organizational variation in organizational routines and the quantity of causal, contextual, and generic knowledge connected with routines, increasing learning myopia. Autor [38] suggested that the incorporation of machine and human comparative advantage enables computers to substitute for employees in executing repetitive, codifiable tasks while amplifying workers' comparative advantage in providing problem-solving abilities, adaptability, and creativity.

A central concern of employing this technology is bias and fairness. Bias is defined as a lack of internal validity or an improper assessment of the relationship between an exposure and an effect in the target population, where the statistic estimated has an expectation that does not equal the actual value [39] Cowgill and Tucker [40] developed an economic perspective on algorithmic bias. One of the findings in this study reveals the disproportionate failure made by prediction for minority groups. AI was found to induce gender and racial bias in other studies. Buolamwini and Gebru [41] found that all classifiers have a biased performance in favor of lighter individuals and men. One breakthrough was the recognition of strategic input incompleteness as a source of bias [42]. The bias of ML is toward locating previous art that is textually comparable to the focus claims, and domain expertise is required to find the most relevant prior art.

For enterprises, the perceived impartiality of the decision-making processes is of utmost importance. By removing biases, algorithms have made it possible to increase fairness. However, the procedure could appear reductionistic to those examined [43]. This may cast doubt on their perceptions of the procedural justice of utilizing HR algorithms to assess performance. Another study emphasizes how crucial it is to take stakeholder strategy into account when creating and assessing fair machine-learning algorithms [44]. The findings show that, if made into legislation, fair ML algorithms that demand effect parity might not be able to provide some of the desired advantages.

6 Mitigating Bias and Challenges

Bias mitigation approaches can be classified into the following three categories: (a) data-focused preprocessing methods, (b) ML algorithm-focused in-processing methods, and (c) ML model-focused postprocessing methods [45].

The results of one study that compared human decisions and machine predictions [46] suggest that, while machine learning can be useful, achieving this value necessitates integrating these tools into an economic framework. This framework needs to establish the connection between predictions and decisions, determine the scope of payoff functions, and create objective decision counterfactuals. In the healthcare industry, there are potential biases across the AI life cycle, including data collection, annotation, machine learning model construction, evaluation, deployment, operationalization, monitoring, and feedback integration. A recent study proposed involving diverse stakeholders and applying human-centered AI concepts to reduce these biases [47]. Human-centered AI can assist in ensuring that AI systems are built and deployed in ways that benefit patients and society, hence reducing health disparities and inequities. Therefore, AI specialists and everyone participating in decision-making should be mindful of bias issues and the impact of their choices regarding design and assumptions.

A study specializes in investigating the technological and administrative challenges of business transformation brought about by the wise adoption and creative use of data sciences in the company [48]. A project aims to address the problem of shrinkage in the Pareto trade-off curve (i.e., diversity shrinkage and validity shrinkage) [49]. A shrinkage approximation formula and a new approach for regularization of Pareto-optimal predictor weights are developed. The dependence of employees on one another and physical items changed as iconic simulation models became more realistic, leading management to confuse operating within representations with acting with oron representations. Managers structured simulation work in virtual teams due to this flawed understanding and their attraction to the virtual, which distanced employees from the real-world examples that served as the basis for their models and made it more challenging to empirically test them. By analyzing how changes to work organization vary by kind of virtual labor, researchers [50] can draw conclusions for the study of virtual work.

7 Conclusion

This paper reviewed the studies on the use of artificial intelligence in the organizational context. It contributes to a systemic understanding of the definition of artificial intelligence, benefits in job applications, recruitment & selection, interview & assessment, management & control, and innovation process. Despite the advantages brought by this technology in these business processes, it also contains limitations, specifically bias. Since most of the existing studies focus on a specific industry like health care, or high-tech, the different ways employees interact with machine learning and other artificial intelligence-related techniques and the degree of involvement should be considered. Government policies regarding the implementation of artificial intelligence vary, causing extra caution multinational firms need to take when devising strategies in different branches. In addition, company-specific factors are important to be taken into consideration, such as financial resources available to invest in and develop AI applications and the skills and attitudes of employees.

Although there is research focusing on the impacts of artificial intelligence on idea generation and innovation of companies, future studies can expand the scope of discussion to investigate the influence of different types of innovation. This review has summarized several findings about the interactions between human capital and artificial intelligence. However, it calls for the need for future research to get a more comprehensive picture of the interactions at individual, team, and organizational levels.

References

1. Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean, and how will they impact health care?. The Journal of Arthroplasty, 33(8), 2358–2361.
2. Kok, J. N., Boers, E. J., Kosters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. Artificial intelligence, 1, 270-299.
3. Castro, D., & New, J. (2016). The promise of artificial intelligence: 70 Real-World examples. Information Technology and Innovation Foundation.
4. Simon, J. P. (2019). Artificial intelligence: scope, players, markets and geography. Digital Policy, Regulation and Governance, 21(3), 208–237.
5. Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., ... & Teller, A. (2022). Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. arXiv preprint arXiv:2211.06318.
6. Meinhart, W. A. (1966). Artificial intelligence, computer simulation of human cognitive and social processes, and management thought. Academy of Management Journal, 9(4), 294-307.
7. Dean Jr, J. W., Yoon, S. J., & Susman, G. I. (1992). Advanced manufacturing technology and organization structure: empowerment or subordination?. Organization Science, 3(2), 203-229.
8. George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. Academy of Management Journal, 57(2), 321-326.
9. Kausel, E. (2018). Big Data at Work: The Data Science Revolution and Organizational Psychology (Book Review). Personnel Psychology. 71. 135–136.
10. Efendić, E., Van de Calseyde, P. P., & Evans, A. M. (2020). Slow response times undermine trust in algorithmic (but not human) predictions. Organizational Behavior and Human Decision Processes, 157, 103-114.
11. Man Tang, P., Koopman, J., McClean, S. T., Zhang, J. H., Li, C. H., De Cremer, D., ... & Ng, C. T. S. (2022). When conscientious employees meet intelligent machines: An integrative approach inspired by complementarity theory and role theory. Academy of Management Journal, 65(3), 1019–1054.
12. Song, Q. C., Shin, H. J., Tang, C., Hanna, A., & Behrend, T. (2022). Investigating machine learning's capacity to enhance the prediction of career choices. Personnel Psychology.
13. Min, H., Peng, Y., Shoss, M., & Yang, B. (2021). Using machine learning to investigate the public's emotional responses to work from home during the COVID-19 pandemic. Journal of Applied Psychology, 106(2), 214.
14. Mohanty, S., Rashid, M. H. A., Mridul, M., Mohanty, C., & Swayamsiddha, S. (2020). Application of Artificial Intelligence in COVID-19 drug repurposing. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14(5), 1027-1031.
15. Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14(4), 337-339.
16. Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. Organization Science, 33(1): 126-148.

17. Barrett, M., Oborn, E., Orlikowski, W. J., & Yates, J. (2012). Reconfiguring boundary relations: Robotic innovations in pharmacy work. Organization Science, 23(5), 1448-1466.
18. Boland Jr, R. J., Lyytinen, K., & Yoo, Y. (2007). Wakes of innovation in project networks: The case of digital 3-D representations in architecture, engineering, and construction. Organization Science, 18(4), 631-647.
19. Dodgson, M., Gann, D. M., & Salter, A. (2007). "In case of fire, please use the elevator": Simulation technology and organization in fire engineering. Organization Science, 18(5), 849–864.
20. Lyu, W., & Liu, J. (2021). Artificial Intelligence and emerging digital technologies in the energy sector. Applied Energy, 303, 117615.
21. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. Academy of Management Annals, 14(1), 366–410.
22. Edwards, R. 1979. Contested terrain : the transformation of the workplace in the twentieth century: New York : Basic Books, c1979.
23. Sajjadiani, S., Sojourner, A. J., Kammeyer-Mueller, J. D., & Mykerezi, E. (2019). Using machine learning to translate applicant work history into predictors of performance and turnover. Journal of Applied Psychology, 104(10), 1207.
24. Landers, R. N., Auer, E. M., Dunk, L., Langer, M., & Tran, K. N. (2023). A simulation of the impacts of machine learning to combine psychometric employee selection system predictors on performance prediction, adverse impact, and number of dropped predictors. Personnel Psychology.
25. 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.
26. Van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. MIS Quarterly, 45(3).
27. Rottman, C., Gardner, C., Liff, J., Mondragon, N., & Zuloaga, L. (2023). New strategies for addressing the diversity–validity dilemma with big data. Journal of Applied Psychology.
28. Hickman, L., Bosch, N., Ng, V., Saef, R., Tay, L., & Woo, S. E. (2022). Automated video interview personality assessments: Reliability, validity, and generalizability investigations. Journal of Applied Psychology, 107(8), 1323.
29. Min, H., Yang, B., Allen, D. G., Grandey, A. A., & Liu, M. (2022). Wisdom from the crowd: Can recommender systems predict employee turnover and its destinations?. Personnel Psychology.
30. Alaimo, C., & Kallinikos, J. (2022). Organizations decentered: Data objects, technology and knowledge. Organization Science, 33(1), 19-37.
31. Leonardi, P. M., & Bailey, D. E. (2008). Transformational technologies and the creation of new work practices: Making implicit knowledge explicit in task-based offshoring. MIS Quarterly, pp. 411–436.
32. Jia, N., Luo, X., Fang, Z., & Liao, C. (2023). When and how artificial intelligence augments employee creativity. Academy of Management Journal.
33. Card, D., & DiNardo, J. E. 2002. Skill-biased technological change and rising wage inequality: Some problems and puzzles. Journal of Labor Economics, 20(4): 733-783.
34. Sturm, T., Gerlach, J. P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., ... & Buxmann, P. (2021). Coordinating Human and Machine Learning for Effective Organizational Learning. MIS Quarterly, 45(3).
35. Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction, judgement, and uncertainty. NBER Working Paper, 24243.

36. Zhang, N., Wang, M., Xu, H., Koenig, N., Hickman, L., Kuruzovich, J., ... & Kim, Y. (2023). Reducing subgroup differences in personnel selection through the application of machine learning. Personnel Psychology.
37. Balasubramanian, N., Ye, Y., & Xu, M. 2022. Substituting human decision-making with machine learning: Implications for organizational learning. Academy of Management Review, 47(3): 448-465.
38. Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of economic perspectives, 29(3), 3–30.
39. Delgado-Rodriguez, M., & Llorca, J. (2004). Bias. Journal of Epidemiology & Community Health, 58(8), 635–641.
40. Cowgill, B., & Tucker, C. E. (2019). Economics, fairness, and algorithmic bias. preparation for: Journal of Economic Perspectives.
41. Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency (pp. 77–91). PMLR.
42. Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. Strategic Management Journal, 41(8), 1381–1411.
43. Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. Organizational Behavior and Human Decision Processes, pp. 160, 149–167.
44. Fu, R., Aseri, M., Singh, P. V., & Srinivasan, K. (2022). "Un" fair machine learning algorithms. Management Science, 68(6), 4173-4195.
45. Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1356.
46. Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). Human decisions and machine predictions. The Quarterly Journal of Economics, 133(1), 237-293.
47. Chen, Y., Clayton, E. W., Novak, L. L., Anders, S., & Malin, B. (2023). Human-centered design to address biases in artificial intelligence. Journal of Medical Internet Research, p. 25, e43251.
48. Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational issues of big data and analytics in networked business. MIS Quarterly, 40(4), 807-818.
49. Song, Q., Tang, C., Newman, D. A., & Wee, S. (2023). Adverse impact reduction and job performance optimization via pareto-optimal weighting: A shrinkage formula and regularization technique using machine learning. Journal of Applied Psychology.
50. Bailey, D. E., Leonardi, P. M., & Barley, S. R. (2012). The lure of the virtual. Organization Science, 23(5), 1485–1504.