Score : Jurnal Lentera Manajemen Pemasaran
Volume 03 No 02 2025
E ISSN : 2987-7903

https://lenteranusa.id/

Employees' Perceptions of Fairness and Ethical Implications of Artificial Intelligence Algorithms in Global Performance Assessment

Agus Supriyanto*[1]
[1]Universitas Islam Negeri Sunan Kudus
E-mail: agussupriyanto@iainkudus.ac.id

Submitted 18-09-2025 Reviewed : 15-10-2025 Accepted : 10-11-2025 Published : 30-12-2025

Abstract
The massive adoption of Artificial Intelligence (AI) in algorithmic management within multinational corporations has triggered a radical transformation in the landscape of International Human Resource Management (IHRM). This qualitative-exploratory study aims to investigate professional employees' perceptions of the ethical and fairness dimensions of AI-based performance appraisal systems. Primary data were collected through online semi-structured in-depth interviews with white-collar professionals selected using purposive sampling techniques. The interview transcripts were subsequently analyzed using Braun and Clarke's Thematic Analysis approach, supported by ATLAS.ti software to ensure systematic coding and the validity of findings. The results reveal that algorithmic autonomy systematically undermines the core pillars of organizational justice. The "black-box" nature of AI systems weakens procedural justice, data reductionism neglects qualitative contributions and compromises distributive justice, while the absence of empathy erodes interactional justice. This ethical crisis manifests in the form of workplace dehumanization through hyper-surveillance and diminished employee autonomy, contributing to phenomena such as quiet quitting and the emergence of algorithmic resistance. The study recommends a critical transition toward Human-Centric AI Management through the implementation of Human-in-the-Loop (HITL) principles to restore human dignity and ethical considerations within the global workplace.

Keyword: Algorithmic Management, Organizational Justice, Algorithmic Dehumanization, Human-Centric AI, IHRM.

Introduction

The rapid advancement of digital technology has fundamentally transformed the nature of work, ushering in an era in which geographical boundaries no longer constrain corporations from operating on a global scale. Within the landscape of International Human Resource Management (IHRM), one of the most significant shifts has been marked by the adoption of Artificial Intelligence (AI) and automation, which have increasingly penetrated strategic organizational functions, including employee performance evaluation systems. Multinational corporations (MNCs) are under growing pressure to operate swiftly, efficiently, and objectively while managing thousands of employees distributed across diverse geographical locations and cultural contexts. This situation has encouraged global organizations to transition from conventional management approaches to Algorithmic Management, a governance system in which supervisory, evaluative, and managerial decision-making functions are delegated to mathematical algorithms and computational codes. The implementation of this technology is

widely believed to overcome the inherent limitations of human judgment, which is often perceived as subjective, inconsistent, and susceptible to personal and cultural biases. Through the utilization of real-time computerized performance metrics, multinational corporations seek to establish globally standardized evaluation systems to achieve distributive fairness for all employees, regardless of their physical work locations (Cheng & Hackett, 2021).

As AI systems increasingly assume control over monitoring digital activities, analyzing response times, and even recommending critical decisions such as compensation adjustments, promotions, and employee termination, the humanistic essence of people management has begun to erode (Kellogg et al., 2020). This process contributes to what has been described as workplace dehumanization, whereby employees are no longer regarded as dynamic intellectual assets but are instead reduced to productivity metrics and lines of code within management dashboards. The "black-box" nature of algorithmic systems renders their decision-making processes highly opaque and difficult for employees to understand (Parker & Boeing, 2023). Such opacity restricts employees' rights to receive reasonable explanations regarding performance outcomes, thereby generating widespread anxiety, prolonged work-related stress, and feelings of excessive monitoring or hyper-surveillance that ultimately undermine employees' psychological well-being across global workplaces (De Stefano, 2020).

These developments have generated critical ethical concerns regarding employees' perceptions of organizational justice. Workplace justice is no longer measured solely by the outcomes employees receive (distributive justice), but also by the procedures through which decisions are made (procedural justice) and the manner in which employees are treated and communicated with throughout the process (interactional justice). When human managers are replaced by AI systems, the channels of emotional interaction and empathy that traditionally facilitate workplace conflict resolution become significantly weakened. Employees who disagree with AI-generated evaluations often have limited opportunities to defend themselves, provide contextual explanations, or negotiate outcomes because algorithms lack the capacity to understand emotional nuances and external circumstances affecting human performance. This deficiency in interactional justice is further exacerbated by the fact that many AI systems are developed using historical datasets that may inadvertently replicate and reinforce pre-existing social, gender, and racial biases embedded within past organizational practices, resulting in decisions that are ethically and morally problematic.

The urgency of investigating this phenomenon has become increasingly evident in order to preserve sustainable industrial relations in the digital era. If perceptions of injustice resulting from AI implementation continue to be overlooked, the consequences may be highly detrimental to the internal stability of global corporations. Deep-seated distrust toward management systems can undermine organizational commitment, reduce employee motivation, encourage productivity sabotage, and increase turnover intentions among highly skilled employees (Glikson & Woolley, 2020). At the international level, policymakers are actively developing regulatory frameworks governing the ethical use of AI in workplaces, as reflected in initiatives such as the European Union's AI Act. Consequently, multinational corporations face substantial pressure to strike an appropriate balance between leveraging technology for business efficiency and upholding their moral responsibility to protect employees' rights, dignity, and workplace justice while minimizing potential legal disputes in the future.

A review of the current scholarly literature reveals several significant research gaps that warrant further investigation. The first gap concerns the dominance of macro-organizational

perspectives in studies examining AI within human resource management. Most existing research has primarily focused on organizational benefits, exploring how algorithms can reduce operational costs, accelerate big data processing, and optimize workplace efficiency metrics. Comparatively little scholarly attention has been devoted to positioning employees as the primary subjects of investigation and examining their micro-psychological experiences, particularly how they subjectively perceive the ethical fairness of decisions made by intelligent machines (Budhwar et al., 2022). Without a deeper understanding of employees' voices and perceptions, contemporary HRM theories may lose their explanatory power in capturing the realities of organizational behavior in an increasingly automated environment.

The second research gap lies in the limited industrial contexts and participant groups represented in the existing algorithmic management literature. Most empirical studies have concentrated on the gig economy, focusing predominantly on blue-collar platform workers such as ride-hailing drivers, delivery couriers, and freelance digital workers. However, AI-driven performance evaluation systems have now been widely adopted by multinational corporations to assess white-collar professionals operating within complex formal organizational structures across national boundaries (Meijerink et al., 2021). Professional employees possess distinct expectations regarding fairness, work autonomy, and career development opportunities that differ substantially from those of platform workers. Furthermore, the interaction between employees' local cultural interpretations and globally centralized HR control systems creates rich and complex social dynamics that remain largely underexplored in management research (Van Veen et al., 2024).

Against this backdrop of practical challenges and theoretical gaps, the present qualitative study adopts an exploratory approach to address the existing shortcomings in the literature. By focusing on professional employees' lived experiences, ethical perspectives, and interpretations of algorithm-based performance evaluation systems, this study seeks to uncover the realities underlying technological implementation in global workplaces. Through this exploration, the study aims to identify emerging patterns regarding how employees redefine procedural, distributive, and interactional justice in the era of artificial intelligence. Ultimately, the findings are expected to contribute to the advancement of contemporary International Human Resource Management (IHRM) literature while providing strategic recommendations for policymakers and managers in multinational corporations. Specifically, the study advocates the development of human-centric AI management systems in which technological sophistication is balanced with ethical principles, fairness, and the preservation of human dignity in employment relationships.

Method

This study employed a qualitative approach with an exploratory-descriptive research design to gain an in-depth understanding of employees' psychological experiences and perceptions regarding the ethical and fairness dimensions of AI-based performance evaluation systems (Saunders et al., 2019). The exploratory nature of the study enabled the researchers to uncover subjective interpretations, concerns, and hidden forms of resistance among professional employees that could not be adequately captured through quantitative survey methods. Participants were selected using purposive sampling based on specific inclusion criteria. Eligible informants were active white-collar professionals employed in multinational corporations (MNCs) where performance evaluation processes had been integrated with

artificial intelligence systems. In addition, participants were required to have at least one year of experience working under such systems to ensure sufficient exposure and depth of empirical insight.

Primary data were collected through semi-structured in-depth interviews. The interviews were conducted online via video conferencing platforms to facilitate participation from informants located across different geographical regions. Each interview lasted approximately 45 to 60 minutes. A flexible interview guide was utilized to encourage participants to openly discuss their experiences and perceptions regarding procedural, distributive, and interactional justice while interacting with AI-driven performance evaluation systems. All interviews were audio-recorded with participants' written informed consent. To uphold research ethics and protect professional privacy, both participants' identities and the names of their organizations were anonymized throughout the study. Data collection and participant recruitment were terminated once data saturation had been achieved, indicating that subsequent interviews yielded repetitive information and no longer generated new thematic insights.

All interview recordings were transcribed verbatim and analyzed using Braun and Clarke's Thematic Analysis framework. The analytical process was supported by ATLAS.ti qualitative data analysis software to ensure a systematic, rigorous, and well-structured research procedure (Miles et al., 2014). The analysis began with data familiarization through repeated readings of the interview transcripts, followed by line-by-line coding to identify concepts related to fairness, ethics, and employees' experiences. The resulting codes were subsequently organized into broader thematic categories grounded in organizational justice theory. Finally, the trustworthiness of the study was ensured through the application of credibility-enhancing strategies, particularly member checking. This process involved returning transcript drafts and preliminary interpretations to participants for verification and confirmation, thereby enhancing the accuracy, credibility, and scientific accountability of the research findings.

Result and Discussion

Ethical Crisis and the Manifestation of Dehumanization in the Global Workplace

The expansion of algorithmic management into professional white-collar environments within multinational corporations has not merely transformed operational structures but has also generated a profound ethical crisis on a global scale. As artificial intelligence systems are granted autonomous authority to continuously monitor, evaluate, and regulate human behavior, the ethical boundary between productivity optimization and human exploitation becomes increasingly blurred. This phenomenon manifests itself through systemic dehumanization, whereby employees are compelled to conform to machine-driven work rhythms, experience diminished autonomy, and endure pervasive surveillance practices that undermine the fundamental essence of individual freedom in the workplace (De Stefano & Taes, 2023).

Hyper-Surveillance and the Digital Panopticon

One of the primary mechanisms underpinning AI-based performance evaluation systems is the implementation of ubiquitous and continuous digital surveillance. To generate accurate and real-time performance assessments, AI systems require vast quantities of employee activity data. Consequently, multinational corporations have increasingly adopted sophisticated employee monitoring technologies, often referred to as "bossware." These systems track virtually every aspect of employees' digital behavior, including keystroke activity, periodic

screen captures, AI-powered webcam monitoring to assess eye movement and attention levels, and even the analysis of vocal tone and emotional expressions during client interactions through Natural Language Processing (NLP) technologies. Such extensive monitoring creates a psychological structure resembling a digital panopticon within the contemporary workplace. Employees become acutely aware that every moment of their workday even brief pauses to rest, stretch, or attend to personal needs is recorded, analyzed, and evaluated by an algorithmic entity that never ceases to observe. Research (De Stefano & Taes, 2023) argues that AI-driven surveillance of this magnitude crosses ethical boundaries surrounding employees' fundamental privacy rights and transforms workplaces into digitally oppressive environments. The persistent awareness of being continuously monitored generates chronic psychological stress and anxiety among employees. Rather than being motivated by intrinsic satisfaction, creativity, or professional commitment, workers increasingly become driven by fear of automated performance penalties imposed by AI systems.

Human Reductionism: Transforming Individuals into Data Commodities

A second ethical concern lies in the inherent tendency of commercial AI systems toward human reductionism. Humanistic management philosophies regard employees as holistic individuals possessing social identities, emotions, values, creativity, and unique developmental potential. In contrast, algorithmic management systems are fundamentally incapable of understanding employees as complex human beings. AI systems can only process digital information, thereby reducing employees to quantifiable datasets and measurable indicators. As a result, employees undergo a process of datafication, whereby their workplace existence is translated into efficiency metrics, productivity charts, and numerical performance scores displayed on managerial dashboards. This reductionist approach fosters a particularly impersonal form of dehumanization in which personal uniqueness, years of loyalty, ethical commitment, and professional dedication become irrelevant unless they are reflected in positive data visualizations. Many employees consequently experience profound identity conflicts because they no longer feel valued as complete human beings but rather as replaceable components within a digital profit-generating system. Research (Kellogg et al., 2020), argue that the commodification of human beings into digital data representations erodes workers' intrinsic dignity and legitimizes exploitative organizational practices under the guise of scientific objectivity. Once employees are viewed merely as numerical indicators, organizations may become more willing to implement ethically questionable actions, such as automated mass layoffs or performance-related sanctions, without meaningful human deliberation or moral accountability.

Loss of Work Autonomy and the Standardization of Human Behavior

Work autonomy the freedom to determine how tasks are performed, manage work schedules, and make independent professional decisions has long been recognized as a critical determinant of job satisfaction and employee well-being. However, AI-driven performance evaluation systems increasingly threaten this autonomy. Because algorithmic systems prioritize standardized and predictable work patterns that are easier to measure and process, they frequently prescribe how employees should perform their tasks rather than merely evaluating outcomes.Algorithms not only assess final performance results but also influence the methods and processes through which work is completed. Employees who attempt to innovate, exercise

professional judgment, or adopt alternative approaches may find their behaviors classified as anomalies that negatively affect their performance metrics. Consequently, many professionals feel compelled to suppress creativity, intuition, and independent decision-making in favor of conforming to algorithmically prescribed routines. This phenomenon has been described as "algorithmic Taylorism," whereby workers are encouraged to behave in increasingly standardized, repetitive, and machine-compatible ways. Research by (Van Veen et al., 2024) suggest that the erosion of autonomy under global algorithmic control weakens employees' sense of agency and contributes to feelings of learned helplessness, as workers increasingly perceive themselves as subordinate to technological systems over which they have little influence or control. Collectively, these findings suggest that while AI-based performance management systems may enhance organizational efficiency and standardization, they simultaneously generate significant ethical challenges related to surveillance, autonomy, privacy, and human dignity. These concerns reinforce the necessity of developing human-centered governance frameworks capable of balancing technological innovation with ethical responsibility and respect for employees as individuals rather than merely sources of data.

Global Industrial Relations Dynamics: Trust, Commitment, and Employee Resistance

The implementation of AI driven performance evaluation systems that neglect principles of organizational justice and human centered ethics generates far reaching consequences for industrial relations within multinational corporations. Beyond their operational implications, these systems fundamentally reshape the psychological relationship between employers and employees. The most significant consequence is the erosion of two foundational pillars that sustain organizational stability: organizational trust and employees' affective commitment. As these foundations weaken, new and increasingly sophisticated forms of employee resistance emerge within digitally mediated workplaces. The growing reliance on algorithmic management creates perceptions that organizations prioritize efficiency, surveillance, and data driven control over employee well being and dignity. As a result, employees increasingly question whether organizational decisions genuinely reflect fairness and concern for human interests. This erosion of trust contributes to declining emotional attachment to the organization, reduced willingness to engage in discretionary efforts, and the emergence of resistance behaviors aimed at reclaiming autonomy and influence within technologically controlled work environments. Consequently, the adoption of AI based performance management systems should not be evaluated solely in terms of efficiency gains or productivity improvements. Equally important is their impact on organizational trust, employee commitment, and the quality of industrial relations. Without adequate safeguards to ensure transparency, fairness, and meaningful human involvement, algorithmic management may inadvertently create organizational environments characterized by distrust, disengagement, and persistent employee resistance (Bélanger et al., 2020).

The Collapse of Organizational Trust

Trust represents one of the most valuable psychological resources within an organization. When employees believe that management acts fairly, honestly, and with genuine concern for their well being, they are more likely to contribute beyond their formal job responsibilities through organizational citizenship behaviors. However, when organizations delegate performance evaluation, a process that directly influences employees' careers and livelihoods, to opaque and inflexible AI algorithms, the foundation of trust begins to deteriorate. Employees

often perceive that management has abdicated its moral responsibility as a leader by hiding behind the perceived objectivity of technological systems. This erosion of trust operates in both directions. Employees no longer trust the intentions of the organization, while organizations that rely heavily on AI implicitly signal a lack of trust in their employees by implementing extensive surveillance mechanisms. Research by (Bélanger et al., 2020) argue that organizations that fail to establish transparency regarding the operation of artificial intelligence systems inevitably weaken both interpersonal and institutional trust. As a result, employees begin to interpret new organizational policies with skepticism and cynicism, viewing them as concealed attempts to strengthen algorithmic control and intensify work demands without providing equitable compensation.

Declining Affective Commitment and the Global Quiet Quitting Phenomenon

A direct consequence of declining trust and perceived injustice is the erosion of employees' affective commitment, defined as their emotional attachment to and identification with the organization. Employees who feel respected, valued, and treated fairly typically develop strong organizational commitment. In contrast, when they perceive themselves as mere data commodities within AI driven management systems, this emotional connection is significantly weakened. The employment relationship gradually shifts from a social and psychological partnership to a purely transactional arrangement characterized by detachment and resignation. This decline in affective commitment has become increasingly visible through the global phenomenon known as quiet quitting, or working strictly according to minimum job requirements. Professional white collar employees consciously choose to withhold discretionary effort, creativity, and innovation from organizations. Instead, they adopt highly calculated work behaviors, focusing only on activities monitored by algorithms and meeting minimum performance thresholds required to avoid negative evaluations. At the end of the workday, they disengage entirely from organizational concerns. Employees become unwilling to invest their intellectual and emotional energy in organizations that treat them as data generating mechanisms rather than as human beings. This phenomenon poses substantial challenges for multinational corporations because it suppresses innovation, weakens organizational learning, and reduces long term competitiveness in increasingly dynamic global markets (Dominique-Ferreira et al., 2023).

The Evolution of Employee Resistance: The Rise of Algorithmic Resistance

Despite operating under sophisticated AI surveillance systems and the extensive power of multinational corporations, employees are not passive recipients of algorithmic control. Recent developments have revealed the emergence of new forms of industrial resistance in which traditional labor protests and collective actions are increasingly complemented by technology based resistance strategies commonly referred to as algorithmic resistance. Employees use their collective intelligence to identify loopholes within organizational AI systems and strategically exploit them to influence or manipulate performance evaluations. These forms of resistance range from subtle individual tactics to coordinated collective actions in digital environments. At the individual level, employees engage in what is often described as gaming the algorithm. Examples include using automated cursor movement software to simulate computer activity or strategically incorporating keywords favored by natural language processing systems into reports and communications in order to obtain more favorable evaluations regardless of the

actual quality of their work. At the collective level, employees establish encrypted communication networks beyond organizational oversight, such as private discussion groups and digital communities, where they exchange strategies for navigating and circumventing algorithmic control systems. These practices represent a form of collective response to perceived injustices and the loss of workplace autonomy. The emergence of algorithmic resistance demonstrates that regardless of technological sophistication, organizational systems that fail to incorporate principles of fairness and human dignity will continue to encounter resistance from employees seeking to preserve their autonomy and agency (Bonini & Treré, 2024).

Toward Human Centric AI Management in International Human Resource Management

The empirical evidence concerning the detrimental consequences of unchecked algorithmic management highlights the urgent need to reconsider how artificial intelligence is implemented within International Human Resource Management. Multinational corporations can no longer rely on purely technocentric approaches that sacrifice human dignity in pursuit of short term efficiency gains. Both scholars and practitioners must embrace a new paradigm of Human Centric AI Management that seeks to balance technological sophistication with ethical responsibility, organizational justice, and respect for human values (Fenwick et al., 2024).

Human in the Loop and Human in Command Principles

A fundamental pillar of human centric AI governance is the recognition that artificial intelligence should function as an augmentative tool designed to support human managers rather than as an autonomous substitute that determines employees' professional futures. To achieve this objective, organizations should adopt Human in the Loop (HITL) and Human in Command (HITC) principles throughout the performance management process. These principles emphasize that no critical decision affecting employees' careers or well being, including promotion decisions, disciplinary actions, bonus allocations, demotions, or employment termination, should be made exclusively by an AI system without meaningful human review and final approval.

Within a human centered performance management cycle, AI systems initially collect and analyze real time productivity data before generating recommendations and preliminary performance assessments. However, these recommendations should not be implemented automatically. Instead, they must pass through a stage of critical human intervention in which managers evaluate contextual information, consider non digital factors, and provide employees with opportunities to explain or challenge evaluation outcomes. The final decision remains under the authority of human managers, ensuring that accountability and moral responsibility are preserved within organizational decision making processes. Human managers therefore serve as ethical filters who assess whether algorithmic recommendations adequately reflect principles of fairness, contextual realities, and organizational values.research by (Budhwar et al., 2022) suggest that combining the analytical capabilities of AI with human emotional intelligence represents one of the most effective approaches for developing performance management systems that are fair, credible, and trustworthy. By restoring human oversight and decision making authority, organizations can alleviate concerns regarding dehumanization and strengthen employees' sense of psychological safety.

Score : Jurnal Lentera Manajemen Pemasaran
Volume 03 No 02 2025
E ISSN : 2987-7903

https://lenteranusa.id/

Radical Algorithmic Transparency and the Right to Explanation

Employees should have the right to understand the technological systems that evaluate their performance. Organizations must provide clear and accessible information regarding the indicators monitored by AI systems, the weighting of performance variables, and the computational logic underlying performance assessments. Furthermore, in line with emerging global AI governance frameworks such as the European Union AI Act, organizations should recognize employees' right to explanation. When employees receive unfavorable evaluations from AI systems, they should be entitled to request detailed explanations regarding how those outcomes were generated. Organizations should also establish independent appeal mechanisms through which employees can present contextual evidence and challenge algorithmic decisions that may have overlooked important qualitative aspects of performance. As emphasized by (Tambe et al., 2019) such transparency is essential for reducing suspicion, strengthening organizational accountability, and restoring the moral legitimacy of AI based performance evaluation systems.

Periodic Ethical Auditing and Global Bias Mitigation

Artificial intelligence systems are not inherently free from bias; they often reproduce the values and assumptions embedded within the data used to train them. Consequently, a third pillar of human centric AI management involves conducting periodic ethical and bias audits of algorithmic performance management systems. These audits should be carried out by independent external bodies to identify whether AI systems unintentionally generate discriminatory outcomes based on race, gender, age, nationality, or cultural background. For example, productivity monitoring systems developed within Western cultural contexts may produce biased evaluations when applied rigidly to employees in Asian societies, where indirect communication styles and collective harmony are often prioritized over individual self promotion. Regular ethical audits ensure that algorithms remain adaptive to diverse cultural norms and organizational contexts within international human resource management environments. Research by (Kocielnik et al., 2019) argue that responsible AI development requires continuous organizational commitment to identifying, evaluating, and mitigating the social risks associated with technological systems. Through rigorous ethical auditing, multinational corporations can ensure that AI serves as a mechanism for promoting fairness, equality, and human dignity rather than becoming a new instrument of digital oppression in the global workplace.

Table 1. Comparative Analysis of Traditional Management, Algorithmic Management, and Human Centric AI Management

Analysis Dimension	Traditional Management (Past Model)	AI Based Algorithmic Management (Current Reality)	Human Centric AI Management (Future Solution 2027)
Decision Making Authority	Human managers make all decisions, making the process vulnerable to	Autonomous computer systems and algorithms make decisions through	Hybrid collaboration in which AI provides recommendations while human managers

	personal bias and subjectivity.	opaque and highly restrictive processes.	retain final decision making authority through Human in the Loop mechanisms.
Nature of Workplace Monitoring	Periodic evaluation conducted through semiannual or annual performance reviews and face to face meetings.	Continuous and pervasive monitoring characterized by real time surveillance of employee activities.	Balanced monitoring that utilizes limited data tracking while maintaining strong protection of employee privacy.
Performance Measurement Focus	Combination of qualitative and quantitative indicators, including target achievement and social behavior.	Extreme data reductionism that primarily values quantifiable digital outputs and measurable activities.	Holistic evaluation that integrates AI generated performance metrics with recognition of emotional labor, creativity, and contextual contributions.
Communication Characteristics	Two way communication that is dialogical, empathetic, and provides opportunities for contextual clarification.	One way communication delivered through automated system notifications with limited opportunities for explanation or appeal.	Digitally enabled dialogue supported by transparent system explanations and formal appeal channels to human managers.
Psychological Impact on Employees	Stress may arise from organizational politics, but employees generally feel recognized and valued as human beings.	Increased risks of dehumanization, acute anxiety, workplace alienation, and the emergence of quiet quitting behaviors.	Enhanced psychological safety, stronger intrinsic motivation, and higher levels of organizational commitment.
Forms of Employee Resistance	Physical demonstrations, labor strikes, and resistance to managerial directives.	Gaming the algorithm, cyber based data manipulation, and coordinated underground digital resistance networks.	Active employee participation in the development of ethical AI policies through collaborative approaches such as algorithmic unionism.

Source: Authors' own elaboration based on the literature review and research findings.

Score : Jurnal Lentera Manajemen Pemasaran
Volume 03 No 02 2025
E ISSN : 2987-7903

https://lenteranusa.id/

A New Theoretical Framework: Integrating Organizational Justice Theory with Artificial Intelligence Ethics

Through the comprehensive analysis presented in this study, this exploratory qualitative research contributes a novel theoretical perspective that extends the boundaries of knowledge within the field of International Human Resource Management (IHRM). Specifically, the study proposes an integrated conceptual framework that combines Greenberg's classical Organizational Justice Theory with contemporary principles of Artificial Intelligence (AI) Ethics. This theoretical synthesis is particularly significant because it demonstrates that, in the twenty-first century, workplace justice can no longer be separated from the ways in which technological architectures are designed, governed, and implemented within organizations. The proposed framework suggests that achieving comprehensive organizational justice in the era of automation requires multinational corporations to move beyond mere legal and regulatory compliance. Instead, organizations must address three interrelated dimensions of algorithmic ethical justice:

1. Procedural (Technological Justice)
   In the contemporary workplace, procedural justice is increasingly determined by the transparency of algorithmic decision-making systems. Procedural fairness can only be achieved when organizations adopt the principles of Explainable Artificial Intelligence (XAI), whereby the formulas, variable weightings, and computational logic used to evaluate professional employees' performance are accessible, understandable, and auditable by human stakeholders. Employees must be provided with sufficient information to understand how performance outcomes are generated, thereby reducing uncertainty and enhancing trust in organizational decision-making processes.
2. Distributive (Contextual Justice)
   Distributive justice should no longer be narrowly defined as mathematical equality based solely on the volume of digital performance data. Contemporary distributive justice requires the development of inclusive algorithmic systems capable of recognizing and rewarding qualitative contributions, abstract thinking, creativity, knowledge sharing, and emotional labor that are essential to organizational effectiveness but may not generate extensive digital footprints. Therefore, fair AI-driven performance management systems must incorporate contextual sensitivity and acknowledge the diverse forms of value creation contributed by professional employees.
3. Interactional (Emotional Justice)
   The adoption of artificial intelligence should not sever the human relationships that constitute the foundation of organizational life. Interactional justice in the digital era requires organizations to preserve spaces for meaningful interpersonal interaction characterized by empathy, respect, and dignity. Human managers must continue to serve as mediators, final decision-makers, and trusted communication partners through whom employees can express concerns, provide contextual explanations, and participate in organizational dialogue. Maintaining a human presence within AI-supported management systems is essential to preserving employees' sense of dignity and belonging.

By implementing these three dimensions of algorithmic ethical justice, multinational corporations can transform algorithmic management from a system often perceived as impersonal, intimidating, and dehumanizing into an intelligent, ethical, and human-centered

model of global talent governance. Within this framework, technology is no longer positioned as a mechanism of digital control that restricts employee autonomy. Instead, it evolves into a strategic partner that supports human flourishing, fosters creativity and innovation, and promotes the holistic well-being of employees within the future global workplace. The proposed framework contributes to the growing literature on AI-enabled human resource management by advancing the concept of Human-Centric AI Management. It further highlights that the long-term sustainability of AI-driven organizational systems depends not only on technological sophistication and operational efficiency but also on their ability to uphold fairness, ethical responsibility, and respect for human dignity.

Conclusion

An ideal AI based performance evaluation cycle should not place the entire decision making process under algorithmic automation. Instead, it should combine the analytical strengths of artificial intelligence with human ethical judgment and professional expertise. The use of AI in collecting, monitoring, and analyzing performance data can significantly enhance efficiency, consistency, and objectivity in employee evaluation processes. However, algorithmic outputs are often unable to fully capture the complexity of workplace situations, individual circumstances, and contextual factors that influence employee performance. Therefore, the presence of a Human in the Loop mechanism becomes essential for reviewing, verifying, and interpreting the recommendations generated by AI systems. Furthermore, providing employees with opportunities for clarification and appeal serves as an important safeguard against potential bias, analytical errors, and the risk of dehumanization in human resource management practices. Decisions related to promotion, rewards, career development, and corrective actions should remain under human authority through the Human in Command principle. By adopting a balanced hybrid approach, organizations can benefit from the efficiency and analytical capabilities of advanced technologies while simultaneously preserving transparency, accountability, fairness, and human dignity within increasingly digitalized work environments.

References

Bélanger, J., Edwards, P. K., & Haiven, L. (2020). *Workplace industrial relations and the global challenge*. Cornell University Press.

Bonini, T., & Treré, E. (2024). *Algorithms of resistance: The everyday fight against platform power*. Mit Press.

Budhwar, P., Malik, A., De Silva, M. T. T., & Thevisuthan, P. (2022). Artificial intelligence–challenges and opportunities for international HRM: a review and research agenda. *The InTernaTIonal Journal of Human Resource ManagemenT*, *33*(6), 1065–1097.

Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, *31*(1), 100698.

De Stefano, V. (2020). 'Masters and servers': Collective labour rights and private government in the contemporary world of work. *International Journal of Comparative Labour Law and Industrial Relations*, *36*(4).

De Stefano, V., & Taes, S. (2023). Algorithmic management and collective bargaining. *Transfer: European Review of Labour and Research*, *29*(1), 21–36.

Dominique-Ferreira, S., Oliveira, M., & Prentice, C. (2023). Organizational commitment: the

role of organizational happiness and quiet quitting. *International Conference on Tourism, Technology and Systems*, 61–70.

Fenwick, A., Molnar, G., & Frangos, P. (2024). The critical role of HRM in AI-driven digital transformation: a paradigm shift to enable firms to move from AI implementation to human-centric adoption. *Discover Artificial Intelligence*, *4*(1), 34.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660.

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.

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will you accept an imperfect ai? exploring designs for adjusting end-user expectations of ai systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–14.

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: Synthesizing developments and cross-disciplinary insights on digital HRM. In *The InTernaTIonal Journal of human resource managemenT* (Vol. 32, Issue 12, pp. 2545–2562). Taylor & Francis.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis*. sage.

Parker, S. K., & Boeing, A. A. (2023). Workplace innovation in the digital era: a role for SMART work design. In *A research agenda for workplace innovation* (pp. 91–112). Edward Elgar Publishing.

Saunders, H., Gallagher-Ford, L., Kvist, T., & Vehviläinen-Julkunen, K. (2019). Practicing healthcare professionals' evidence-based practice competencies: An overview of systematic reviews. *Worldviews on Evidence-Based Nursing*, *16*(3), 176–185.

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

Van Veen, D., Van Uden, C., Blankemeier, L., Delbrouck, J.-B., Aali, A., Bluethgen, C., Pareek, A., Polacin, M., Reis, E. P., & Seehofnerová, A. (2024). Adapted large language models can outperform medical experts in clinical text summarization. *Nature Medicine*, *30*(4), 1134–1142.

.