THE DARK SIDE OF AI-ENABLED HRM: PERCEIVED ORGANIZATIONAL INJUSTICE, EMPLOYEE PRIVACY CONCERN, AND WELL-BEING

Hima Barima[1*)], Lindawati Sipahutar[2)]

[1,2)] *Management Study Program, Faculty of Economics and Business, Universitas Perwira Purbalingga, Purbalingga 53313, Central Java, Indonesia*

[*]Email: himabarima@gmail.com

ABSTRACT

Purpose: This article develops an employee-centered framework to explain the dark side of AI-enabled human resource management (AI-HRM). Method: Using a structured integrative conceptual review, it synthesizes 52 peer-reviewed articles published between 2019 and 2025, seven foundational sources, and one official report. Findings: Five mechanisms—algorithmic bias, opacity, workplace surveillance, privacy-invasive people analytics, and algorithmic control—influence perceived organizational injustice, employee privacy concern, and work autonomy. These experiences affect psychological, work-related, and relational well-being both directly and indirectly through organizational trust. Four pathway-specific governance capabilities—algorithmic accountability, actionable explainability, privacy-preserving data governance, and meaningful human oversight—can mitigate these risks. The framework is conceptual rather than empirically tested; its ten core propositions therefore require future validation. Contribution: By integrating previously fragmented research on AI-HRM, fairness, privacy, algorithmic management, and well-being, the article offers a parsimonious and testable framework with practical guidance for responsible AI-HRM.

Keywords: AI-enabled HRM; organizational justice; employee privacy; organizational trust; employee well-being

ABSTRAK

Tujuan: Artikel ini mengembangkan kerangka yang berpusat pada karyawan untuk menjelaskan sisi gelap manajemen sumber daya manusia berbasis kecerdasan buatan (AI-HRM). Metode: Dengan menggunakan tinjauan konseptual integratif yang terstruktur, artikel ini menyintesis 52 artikel ilmiah yang terbit pada 2019–2025, tujuh rujukan teoretis fundamental, dan satu laporan resmi. Hasil: Lima mekanisme—bias algoritmik, opasitas, pengawasan di tempat kerja, analitik tenaga kerja yang invasif terhadap privasi, dan kontrol algoritmik—memengaruhi persepsi ketidakadilan organisasi, kekhawatiran privasi karyawan, dan otonomi kerja. Pengalaman tersebut memengaruhi kesejahteraan psikologis, pekerjaan, dan relasional, baik secara langsung maupun tidak langsung melalui kepercayaan organisasi. Empat kapabilitas tata kelola yang spesifik pada jalur—akuntabilitas algoritmik, eksplainabilitas yang dapat ditindaklanjuti, tata kelola data yang menjaga privasi, dan pengawasan manusia yang bermakna—dapat mengurangi risiko tersebut. Kerangka ini bersifat konseptual dan belum diuji secara empiris; karena itu, sepuluh proposisi intinya memerlukan validasi pada penelitian mendatang. Kontribusi: Artikel ini mengintegrasikan riset AI-HRM, keadilan, privasi, manajemen algoritmik, dan kesejahteraan ke dalam kerangka yang ringkas, dapat diuji, dan relevan bagi tata kelola AI-HRM yang bertanggung jawab.

Kata kunci: *AI-HRM; keadilan organisasi; privasi karyawan; kepercayaan organisasi; kesejahteraan karyawan*

1. Introduction

Artificial intelligence has become embedded in human resource management (HRM), where it supports recruitment, selection, performance evaluation, career development, people analytics, employee monitoring, and HR decision-making. AI-enabled HRM (AI-HRM) can improve efficiency, consistency, analytical capacity, and evidence-based decision-making. At the same time, it places algorithms at the center of consequential employment processes. AI-HRM is therefore better understood as a human-centered socio-

technical system than as a neutral technical tool.

Digital transformation and shifting skill requirements make these issues increasingly urgent. The World Economic Forum (2025) identifies AI and related technologies as major drivers of change in jobs, skills, and the workforce. Organizations must therefore look beyond adoption and productivity to consider the social, ethical, and psychological consequences of AI-HRM.

Prior research documents both the benefits of AI-HRM and its unresolved risks. AI can support complex HR processes, yet concerns remain about data quality, accountability, bias, ethics, governance, and employee acceptance (Budhwar et al., 2022; Qamar et al., 2021; Tambe et al., 2019; Vrontis et al., 2022). This article focuses on five dark-side mechanisms: algorithmic bias, opacity, workplace surveillance, privacy-invasive people analytics, and algorithmic control.

These mechanisms shape three dimensions of employee experience. Bias and opacity can threaten distributive, procedural, interpersonal, and informational justice; surveillance and disproportionate data use can heighten employee privacy concern; and algorithmic control can erode work autonomy. In turn, injustice, privacy concern, and diminished autonomy can weaken organizational trust and undermine psychological, work-related, and relational well-being (Kellogg et al., 2020; Leicht-Deobald et al., 2019; Parent-Rocheleau & Parker, 2022; Tursunbayeva et al., 2022).

Research remains fragmented across technological, functional, and employee-centered perspectives (Basu et al., 2023; Giermindl et al., 2022; Hunkenschroer & Luetge, 2022; Pan & Froese, 2023). Design-oriented and multilevel studies consequently call for closer alignment among technology, HR systems, organizational context, and human needs (Einola & Khoreva, 2023; Prikshat et al., 2023; Tinguely et al., 2023). These reviews provide broad maps of AI-HRM capabilities, challenges, and outcomes, but they do not jointly explain how distinct harmful mechanisms are translated into employee experiences, organizational trust, and multidimensional well-being. The present article therefore advances beyond integration alone by specifying a bounded causal architecture with partial mediation through organizational trust and pathway-specific governance moderators that can be tested separately.

The issue has particular practical and regulatory relevance in Indonesia as organizations expand digital recruitment, people analytics, employee monitoring, and data-driven performance management. These applications may process identifiers, behavioral records, communication data, performance scores, or biometric information and therefore intersect with Law No. 27 of 2022 on Personal Data Protection (UU PDP), which establishes data-subject rights and responsibilities for personal-data controllers and processors (Republic of Indonesia, 2022). Relating AI-HRM governance to clear processing purposes, data-subject rights, security, organizational accountability, and meaningful human review strengthens the framework's relevance to Indonesian employers while retaining cross-national applicability.

A key unresolved question is how these mechanisms combine to shape ongoing employment relationships and how governance can reduce pathway-specific risks. This article distinguishes data practices from employee privacy concern, system trust from organizational trust, and workplace surveillance from privacy-invasive people analytics. Evidence drawn from applicants, organizational employees, organization-level datasets, and platform workers is interpreted cautiously because these settings represent different employment relationships.

Three research questions guide the review: RQ1, which dark-side mechanisms

of AI-HRM most consistently shape employee-centered outcomes? RQ2, how do perceived injustice, privacy concern, and work autonomy influence organizational trust and employee well-being? RQ3, which pathway-specific governance capabilities can attenuate these relationships?

To answer these questions, the article develops an employee-centered socio-technical framework, formulates separately testable propositions, and identifies algorithmic accountability, actionable explainability, privacy-preserving data governance, and meaningful human oversight as targeted governance capabilities.

2. Literature Review

2.1 Artificial Intelligence in Human Resource Management

AI-HRM encompasses the use of algorithms, machine learning, automation, and data-driven systems across HR functions such as recruitment, selection, performance appraisal, career development, monitoring, people analytics, and decision-making. It marks a shift from conventional HRM toward digital and algorithmic people management.

AI can strengthen decision support (Qamar et al., 2021; Tambe et al., 2019). However, HR decisions still require contextual judgment and ethical responsibility (Strohmeier, 2020).

Realizing the value of AI-HRM requires appropriate technical and organizational capabilities (Budhwar et al., 2022; Chowdhury et al., 2023; Malik et al., 2023). It also depends on aligning HR practices, governance arrangements, and human needs (Pan & Froese, 2023; Prikshat et al., 2023; Tinguely et al., 2023; Vrontis et al., 2022).

2.2 The Dark Side of AI-Enabled HRM

The dark side of AI-HRM refers to the unintended, harmful, or ethically problematic consequences that can arise when AI is used to manage people, particularly in decisions involving hiring, promotion, evaluation, development, discipline, and employment security.

The principal risks are algorithmic bias, opacity, workplace surveillance, privacy-invasive people analytics, and algorithmic control. These mechanisms can reduce employees to data categories, weaken context and due process, and create asymmetries of information and power (Leicht-Deobald et al., 2019; Varma et al., 2023). Responsible AI-HRM therefore requires fairness, privacy protection, explainability, accountability, contestability, and meaningful human oversight (Bujold et al., 2024; Zhou et al., 2023).

2.3 AI Recruitment, Selection, and Algorithmic Fairness

Recruitment and selection are among the most visible applications of AI-HRM, including resume screening, automated interviews, personality assessment, job-fit prediction, and hiring recommendations. Although these tools can reduce administrative workloads, they can also reproduce discrimination, encourage algorithmic reductionism, and provoke negative applicant reactions (Ore & Sposato, 2022). Here, algorithmic reductionism means representing a person through a narrow set of data points or model scores while omitting relevant context.

Applicant reactions depend on technology quality and social presence (Black & van Esch, 2020; Suen et al., 2019; Van Esch & Black, 2019), as well as transparency, fairness, and the design of digital selection procedures (Folger et al., 2022; Langer et al., 2019, 2020).

Concerns intensify when historical data reproduce disadvantage (Hunkenschroer & Luetge, 2022; Köchling & Wehner, 2020; Mori et al., 2025) or when systems overlook uniqueness, context, and interpersonal treatment (Lavanchy et al., 2023; Newman et al., 2020; Oostrom et al., 2024). Similar concerns extend to AI-

supported career development and performance evaluation (Köchling et al., 2025; Qin et al., 2023).

2.4 Algorithmic HRM and Algorithmic Management

Algorithmic HRM uses digital data and algorithms to support, automate, or guide HR decisions and practices (Meijerink et al., 2021). It overlaps with algorithmic management, through which algorithms allocate tasks, set targets, monitor performance, evaluate workers, and influence rewards or sanctions.

These systems can create value, but continuous monitoring, dashboards, scores, and automated feedback may intensify control and power asymmetries (Duggan et al., 2020; Kellogg et al., 2020). Their effects depend on how they are implemented: supportive use may augment human judgment, whereas rigid or opaque use can restrict agency, autonomy, and meaningful job design (Curchod et al., 2020; Gagné et al., 2022; Meijerink & Bondarouk, 2023; Parent-Rocheleau & Parker, 2022).

2.5 Organizational Justice in AI-HRM

Organizational justice captures employees' perceptions of fairness in outcomes, procedures, interpersonal treatment, and explanations. Colquitt (2001) distinguishes distributive, procedural, interpersonal, and informational justice, each of which becomes relevant when AI influences selection, promotion, compensation, performance ratings, training opportunities, or termination. Indonesian employee evidence likewise indicates that procedural and interactional justice can shape work motivation and job satisfaction (Andriani et al., 2023).

In AI-HRM, biased outcomes may create distributive injustice; opaque or non-contestable processes may create procedural injustice; dehumanized treatment may threaten interpersonal justice; and inadequate explanations may undermine informational justice. Fair AI therefore requires not only accuracy but also dignity, actionable explanation, voice, and human involvement (Pessach & Shmueli, 2022; Robert et al., 2020; Starke et al., 2022). Regional evidence also shows that effective internal communication can improve role clarity, workplace relationships, and employee satisfaction (Ly-Le & Le, 2024), reinforcing the importance of understandable and decision-relevant explanations in AI-HRM.

2.6 Data Privacy, People Analytics, and Workplace Surveillance

AI-HRM depends on employee data, including information about behavior, productivity, communication, performance, engagement, and career potential. Data privacy refers to the organizational rules governing how these data are collected, processed, stored, accessed, retained, and used. Employee privacy concern arises when employees perceive such practices as excessive, illegitimate, non-transparent, or beyond their control.

People analytics and workplace surveillance can support evidence-based decisions, but they also raise concerns about consent, ownership, performance pressure, bias, and misuse (Giermindl et al., 2022; Tursunbayeva et al., 2022). Such concerns tend to be lower when data collection is job-relevant, proportionate, transparent, secure, and accompanied by meaningful opportunities for voice and correction (Culnan & Armstrong, 1999; Glavin et al., 2024; Ravid et al., 2020; Vitak & Zimmer, 2023).

2.7 Employee Well-Being in AI-Enabled Work Systems

Employee well-being is conceptualized as a multidimensional domain with psychological, work-related, and relational aspects (Grant et al., 2007). Stress, insecurity, alienation, and anxiety are treated as indicators of ill-being rather than as interchangeable constructs. Future studies may examine these dimensions

separately or model them as a theoretically justified higher-order construct.

Job Demands-Resources Theory explains how surveillance and algorithmic control can function as job demands, while work autonomy serves as a central job resource (Demerouti et al., 2001). Research grounded in self-determination theory further suggests that algorithmic management can affect autonomy, competence, relatedness, and motivation (Gagné et al., 2022). The effects depend on purpose, intensity, explanation, employee voice, and the quality of human review (Kinowska & Sienkiewicz, 2023; Parent-Rocheleau & Parker, 2022).

2.8 Theoretical Foundations

Five core theories anchor the framework. Organizational Justice Theory explains how algorithmic outcomes, procedures, treatment, and explanations shape distributive, procedural, interpersonal, and informational justice. Job Demands-Resources Theory explains how surveillance and control become demands and how work autonomy functions as a key resource (Colquitt, 2001; Demerouti et al., 2001).

Social Exchange Theory explains why unfair, intrusive, or controlling practices weaken reciprocal relationships, while Organizational Trust Theory identifies competence, integrity, and benevolence as the foundations of trust in the organization that selects and governs AI-HRM (Cropanzano & Mitchell, 2005; Mayer et al., 1995). Socio-Technical Systems Theory connects these effects to the joint design of technology, HR practices, organizational structures, and human needs (Trist & Bamforth, 1951).

Two supporting perspectives further refine the model. The Information Privacy and Procedural Fairness perspective clarifies consent, proportionality, transparency, control, and legitimate data use (Culnan & Armstrong, 1999), while the Employee Well-Being Perspective highlights psychological, work-related, and relational outcomes (Grant et al., 2007). Table 1 shows how the core theories generate the causal pathways and how the supporting perspectives sharpen specific constructs and outcomes.

In plain terms, the theories jointly address five questions: Was the AI-supported outcome and process fair? Did the system add demands or remove autonomy? What message did the practice send about the employment relationship? Did the organization remain competent, honest, and benevolent? And were technology and work design aligned with human needs? Together, these questions connect AI-HRM design to employee experience, organizational trust, and well-being; the supporting perspectives specify the legitimacy of data practices and the forms of well-being at stake.

Table 1. Theoretical foundations and their role in the conceptual framework

Foundation	Role in this article	Main conceptual link	Supported propositions
Organizational Justice Theory (core)	Explains fairness of outcomes, procedures, treatment, and explanations.	Bias/opacity → perceived injustice	P1, P4, P5, P7, P8
Job Demands-Resources Theory (core)	Frames control as a demand and autonomy as a job resource.	Control → autonomy → trust/well-being	P3, P4, P5, P6, P10
Social Exchange Theory (core)	Explains relational responses to unfair, intrusive, or controlling practices.	Employee experiences → organizational trust	P4, P6
Organizational Trust Theory (core)	Defines competence, integrity, and benevolence as bases of organizational trust.	Organizational trust → well-being/mediation	P4, P6
Socio-Technical Systems Theory (core)	Links technology, HR practices, organizational context, and human needs.	Integrated contingent framework	P1–P10

Foundation	Role in this article	Main conceptual link	Supported propositions
Information Privacy and Procedural Fairness (supporting)	Explains legitimate, proportionate, transparent, and controllable data use.	Surveillance/analytics → privacy concern	P2, P4, P5, P6, P9
Employee Well-Being Perspective (supporting)	Treats well-being as multidimensional and potentially subject to trade-offs.	Experiences/trust → well-being	P5, P6

Source: Authors' theoretical synthesis (2026).

3. Research Method

3.1 Research Design

This article adopts a structured integrative conceptual review to connect fragmented evidence and develop a theory-informed framework for the dark side of AI-HRM. Its purpose is to synthesize concepts, compare mechanisms and boundary conditions, and develop propositions rather than estimate statistical effects.

The review does not claim to provide the exhaustive coverage or formal risk-of-bias assessment expected of a PRISMA-based systematic review, nor is it a meta-analysis. Numerical screening is reported to make the selection process transparent, while inclusion decisions reflect conceptual relevance, theoretical contribution, publication quality, recency, and usefulness for explaining employee-centered AI-HRM risks.

3.2 Literature Sources

The evidence base comprises 52 peer-reviewed journal articles published from 2019 to 2025, seven foundational theoretical sources, and the World Economic Forum (2025) report. The recent articles capture current debates, the foundational sources anchor the model, and the report provides workforce context. Ending the primary search window in 2025 avoids comparing a partial 2026 publication year with seven complete publication years.

Appendix A provides the complete study-level coding record for the 52 recent journal articles.

3.3 Literature Search Strategy

The literature search was completed on July 1, 2026, before manuscript submission. Scopus and Web of Science were the primary bibliographic databases; targeted publisher-platform searches and backward and forward citation checking were used only to identify conceptually central studies not consistently captured by the database searches. Table 2 provides a concise summary of the search, eligibility, selection, appraisal, and coding process.

Table 2. Concise review protocol and selection summary

Element	Concise specification
Review type	Structured integrative conceptual review for theory development; not a PRISMA-based systematic review, meta-analysis, or formal risk-of-bias assessment.
Search completion and window	July 1, 2026, before submission; recent peer-reviewed journal articles from 2019–2025; foundational sources retained without a date limit.
Sources	Scopus and Web of Science as primary databases; targeted publisher searches and backward/forward citation checking as supplements.
Eligibility and appraisal	English-language peer-reviewed sources directly informing a construct, mechanism, boundary condition, or governance pathway; assessed for design/evidence clarity, scope, construct clarity, and conceptual relevance.
Selection summary	100 retrieval events → 92 records after deduplication → 69 full texts assessed → 52 recent journal articles retained.
Evidence base	52 recent journal articles, seven foundational sources, and one World Economic Forum report; two DeReMa regional studies and the Indonesian PDP Law are contextual only.

Element	Concise specification
Coding and review process	The first author searched, screened, and coded; the second author reviewed eligibility, classifications, definitions, and framework coherence. Decisions were checked against explicit criteria and publisher metadata.

Source: Authors' structured integrative review protocol (completed July 1, 2026).

3.4 Inclusion, Exclusion, and Quality Appraisal Criteria

Eligible sources were English-language, peer-reviewed conceptual, empirical, review, or theory-building articles that directly informed at least one construct, mechanism, boundary condition, or governance pathway. Purely technical studies were excluded unless they provided foundational fairness concepts that could be explicitly translated into HRM, as in the retained review by Pessach & Shmueli (2022). Quality was appraised for fit with the review purpose rather than through a formal risk-of-bias score. Empirical studies were assessed for an identifiable design, unit or context, and relevant evidence; reviews for transparent scope and synthesis; and conceptual papers for construct clarity, theoretical grounding, and a coherent contribution. Bibliographic details, DOI records, and Scopus and/or Web of Science indexing were verified where available.

3.5 Literature Selection Process

The search produced 100 retrieval events before cross-platform deduplication. After eight duplicates were removed, 92 records were screened by title and abstract; 69 full texts were assessed; and 52 recent journal articles were retained because they directly informed the framework's constructs, mechanisms, contexts, or governance pathways.

These counts are reported for transparency rather than to claim an exhaustive systematic review. Appendix A provides the complete study-level record and distinguishes primary thematic cluster, evidence design, and employment context.

Appendix A reports each retained article by its primary thematic cluster, decision or employment context, and evidence design. The synthesis also considered cross-theme relevance because a single study could inform more than one pathway.

The first author conducted the initial database search, screening, and coding. The second author reviewed eligibility decisions, thematic classifications, construct definitions, and framework coherence, and both authors approved the final synthesis. Because this was a fit-for-purpose conceptual review conducted by a two-author team rather than a systematic review requiring duplicate independent coding, a formal inter-coder reliability statistic was not computed. Consistency was instead supported through explicit eligibility criteria, shared coding definitions, cross-checking of disputed classifications, bibliographic verification, and publisher metadata.

3.6 Data Analysis and Conceptual Synthesis

The analysis combined structured coding, thematic comparison, and conceptual synthesis. Study-level coding captured author-year, evidence design, primary thematic cluster, and employment context. Mechanisms, outcomes, and contributions were synthesized at the thematic and framework levels and are reported in Table 3 and Sections 4.3–4.5. Operationalization guidance follows Proposition 10, while Appendix A contains the complete study-level coding matrix.

Rather than simply summarizing individual studies, the analysis compared evidence for convergence, contradiction, context, and boundary conditions. The synthesis produced five dark-side mechanisms, three employee experiences, direct and trust-mediated pathways to well-being, four pathway-specific governance capabilities, and ten core propositions. Component paths remain visible in Figure

1 without being treated as separate subpropositions.

4. Results and Discussion

4.1 Thematic Synthesis of the Literature

Table 3 summarizes eight interconnected research streams: AI-HRM/digital HRM, AI recruitment and selection, algorithmic HRM/management, algorithmic fairness and justice, people analytics/privacy, workplace surveillance, employee well-being, and responsible AI governance. By study design, the corpus comprises 17 empirical or mixed-empirical studies, 21 reviews or syntheses, and 14 conceptual or theoretical articles. By decision or employment context, it includes 21 cross-context or HRM-wide studies, 15 organizational-workplace studies (including organization-level data), 13 recruitment or selection studies, and three platform or gig-work studies.

AI-HRM reviews emphasize capability and transformation; recruitment studies document applicant reactions and fairness; algorithmic management research highlights control, power, and autonomy; and people analytics studies focus on consent, privacy, and data misuse. Although complementary, these streams are not interchangeable. Applicant reactions relate to organizational entry, internal HR decisions shape ongoing employment, and platform studies reflect distinct contractual and control arrangements.

Across these contexts, findings depend on how AI-HRM is designed and used rather than being uniformly negative. Automation may improve consistency, access, or decision support, but its fairness benefits diminish when historical data reproduce disadvantage or when employees cannot understand or contest consequential decisions. Monitoring may be accepted when it is job-relevant and proportionate; hidden, continuous, or punitive monitoring, by contrast, is associated with greater privacy and well-being concerns. Human review protects employees only when reviewers have the competence, authority, and accountability to do more than endorse algorithmic outputs.

Accordingly, the synthesis conceptualizes the dark side of AI-HRM as a contingent socio-technical process. Technical features interact with organizational choices, employment context, data governance, employee voice, and the authority of human reviewers to shape injustice, privacy concern, work autonomy, trust, and well-being.

Table 3. Thematic synthesis of the AI-HRM literature

Thematic cluster	Main focus	Key tension	Role in the framework
AI-HRM and digital HRM	Automation, decision support, transformation	Efficiency may obscure employee-centered risks	Defines the socio-technical context
AI recruitment and selection	Screening, interviews, assessment, hiring	Bias, reductionism, opacity, applicant reactions	Links bias/opacity to injustice
Algorithmic HRM/management	Task allocation, evaluation, monitoring, control	Value creation versus power asymmetry and lost autonomy	Defines control and autonomy pathways
Algorithmic fairness/justice	Fairness of AI-supported decisions	Distributive, procedural, interpersonal, informational injustice	Defines perceived injustice
People analytics/data privacy	Consent, access, retention, secondary use	Evidence-based HRM versus intrusive data use	Distinguishes practices from privacy concern
Workplace surveillance	Digital monitoring and performance tracking	Effects vary by intensity, purpose, visibility, proportionality	Supports surveillance–privacy/well-being paths
Employee well-being	Psychological, work-related, relational outcomes	Strain, insecurity, alienation, reduced meaningfulness	Defines the multidimensional outcome
Responsible AI governance	Accountability, explainability, privacy governance, oversight	Safeguards may be substantive or symbolic	Provides pathway-specific moderators

Source: Authors' synthesis based on 52 recent journal articles (2019–2025).

4.2 Research Gap and Conceptual Contribution

The main gap is that these research streams remain weakly integrated. AI-HRM reviews map adoption and capability (Basu et al., 2023; Pan & Froese, 2023; Vrontis et al., 2022); recruitment research concentrates on applicant fairness (Hunkenschroer & Luetge, 2022; Mori et al., 2025); algorithmic management examines control and agency (Kellogg et al., 2020; Parent-Rocheleau & Parker, 2022); and people analytics research foregrounds surveillance and privacy (Giermindl et al., 2022; Tursunbayeva et al., 2022).

Organizations often deploy several AI-HRM applications simultaneously, so employees may experience unfair outcomes, opaque procedures, intrusive data practices, and constrained autonomy together. The framework integrates these experiences while retaining important distinctions between data privacy and privacy concern, system trust and organizational trust, and surveillance and privacy-invasive analytics.

The framework makes three contributions. First, it connects five mechanisms to three employee experiences and to both direct and trust-mediated pathways to well-being. Second, it assigns explicit causal roles to the core and supporting theories. Third, it links four governance capabilities to specific risk pathways instead of offering a generic ethical checklist.

To preserve parsimony, the model focuses on primary pathways. Cross-path effects—effects outside the model's main arrows, such as surveillance reducing autonomy, control increasing injustice, or opacity directly weakening trust—remain theoretically plausible and should be tested as extensions. Boundary conditions are contextual factors that may change the strength or direction of a relationship; relevant examples include the AI-HRM function, decision stakes, employment context, data intensity, employee voice, and the competence and authority of human reviewers.

4.3 Proposed Conceptual Framework

Figure 1 presents the proposed framework by distinguishing direct effects, mediation through organizational trust, and pathway-specific moderation. It links five dark-side mechanisms—algorithmic bias, opacity, workplace surveillance, privacy-invasive people analytics, and algorithmic control—to perceived organizational injustice, employee privacy concern, work autonomy, organizational trust, and employee well-being.

Algorithmic bias refers to systematically unfair outcomes arising from data, design choices, proxies, or weak contextual understanding. Opacity denotes limited clarity about data, logic, uncertainty, responsibility, or avenues for appeal. Workplace surveillance involves the digital monitoring of behavior, communication, productivity, or performance, whereas privacy-invasive people analytics involves data use that exceeds legitimate and proportionate HR purposes. Algorithmic control occurs when systems allocate work, set targets, evaluate performance, or influence rewards and sanctions.

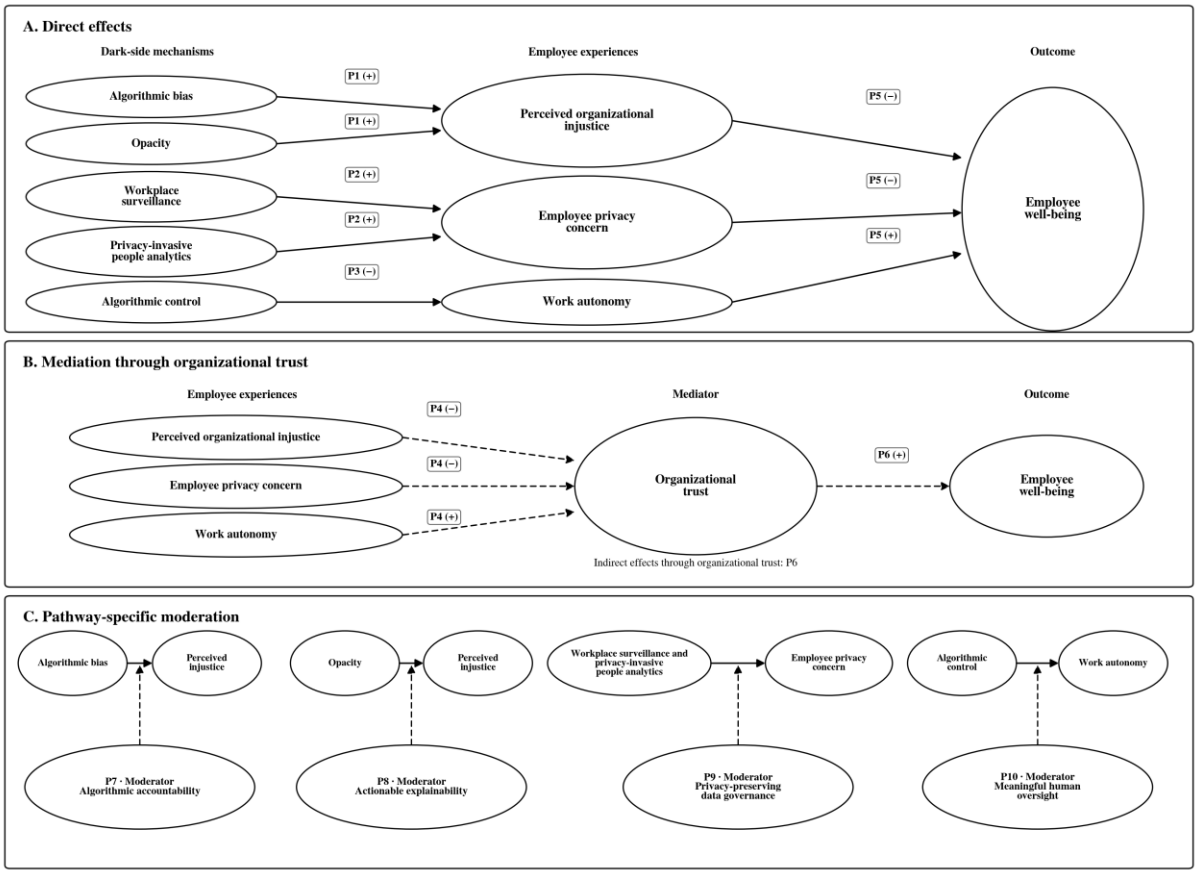
Together, these mechanisms give rise to three primary employee experiences: bias and opacity increase perceived organizational injustice; surveillance and privacy-invasive analytics heighten employee privacy concern; and algorithmic control reduces work autonomy. Perceived organizational injustice is treated as a higher-order domain whose distributive, procedural, interpersonal, and informational dimensions can also be tested separately. Injustice and privacy concern are expected to undermine well-being, while work autonomy is expected to support it, both directly and through organizational trust.

Organizational trust is modeled as a partial rather than an exclusive mediator. The organization selects the system, sets

data boundaries and decision authority, and determines whether employees receive explanations, correction, and opportunities to appeal. Trust in the AI-HRM system remains a distinct supporting concept based on technical accuracy and reliability; the central mediator is trust in the organization's competence, integrity, and benevolence.

Four pathway-specific governance capabilities—algorithmic accountability, actionable explainability, privacy-preserving data governance, and meaningful human oversight—can work in two ways. They may prevent risks from arising through better design and governance, and they may attenuate employee consequences when a focal mechanism is already present. Section 4.5.6 develops these distinct roles.

Figure 1. Conceptual framework of the dark side of AI-HRM



Source: Authors' conceptual framework (2026). Solid arrows indicate direct or focal paths; dashed arrows indicate mediation paths and moderator-to-path relationships.

4.4 Organizational Trust as a Partial Mediator

Employees form judgments about both the AI-HRM system and the organization behind it. Trust in the system rests on technical accuracy and reliability. Organizational trust, by contrast, reflects the perceived competence, integrity, and benevolence of the organization that selects, uses, audits, and communicates the system (Mayer et al., 1995).

Organizational trust functions as a partial mediator because the organization determines data boundaries, decision authority, appeal rights, and corrective action. Injustice and privacy concern may directly undermine well-being, while autonomy may directly support it. At the same time, these experiences convey broader signals about organizational intent, which are transmitted through trust.

4.5 Development of Propositions

4.5.1 Algorithmic Bias, Opacity, and Perceived Organizational Injustice

AI-HRM decisions affect access to employment, rewards, promotion, evaluation, training, and career opportunities. Employees consequently judge not only the outcomes of AI-supported decisions but also the procedures, explanations, opportunities for voice, and interpersonal treatment surrounding them (Bankins et al., 2022; Colquitt, 2001; Starke et al., 2022).

Algorithmic bias can create distributive injustice when outcomes systematically disadvantage groups or individuals and procedural injustice when models use invalid proxies, inconsistent rules, or decontextualized data (Köchling & Wehner, 2020; Newman et al., 2020; Pessach & Shmueli, 2022). Opacity can create informational and procedural injustice when employees cannot understand the basis of a decision, identify responsibility, or obtain review (Langer & König, 2023; Leicht-Deobald et al., 2019).

Bias and opacity remain analytically distinct within Proposition 1 because they create different justice concerns and require different governance responses; their component paths are retained in Figure 1 and in the operationalization guidance following Proposition 10.

Proposition 1: Algorithmic bias and opacity in AI-enabled HRM are likely to increase perceived organizational injustice: bias primarily through distributive and procedural injustice, and opacity primarily through informational and procedural injustice.

4.5.2 Workplace Surveillance, People Analytics, and Privacy Concern

AI-HRM may draw on data about productivity, attendance, communication, work patterns, platform activity, and performance. Workplace surveillance involves monitoring behavior or performance, whereas privacy-invasive people analytics uses personal, secondary, or disproportionate data beyond a clearly legitimate HR purpose. People analytics is therefore not assumed to be inherently equivalent to surveillance (Giermindl et al., 2022; Tursunbayeva et al., 2022).

Employee privacy concern arises when monitoring or analytics are hidden, excessive, poorly secured, unrelated to legitimate work purposes, or used without clear limits on access, retention, secondary use, correction, and appeal (Glavin et al., 2024; Ravid et al., 2020; Vitak & Zimmer, 2023).

Proposition 2: Workplace surveillance and privacy-invasive people analytics are likely to heighten employee privacy concern, particularly when data practices are intensive, hidden, continuous, disproportionate, or weakly controllable.

4.5.3 Algorithmic Control and Work Autonomy

Algorithmic control occurs when systems allocate tasks, set targets, monitor performance, provide feedback, or influence rewards and sanctions. Such systems may operate continuously through dashboards, scores, and automated recommendations, thereby redistributing discretion and power (Duggan et al., 2020; Kellogg et al., 2020).

Work autonomy declines when employees have little discretion over how they perform their jobs or cannot explain, negotiate, or challenge algorithmic directions. Because autonomy is both a central job resource and a basic motivational need, losing it can weaken motivation and increase strain (Gagné et al., 2022; Meijerink & Bondarouk, 2023; Parent-Rocheleau & Parker, 2022).

Proposition 3: Algorithmic control in AI-enabled HRM is likely to reduce employees' work autonomy.

4.5.4 Perceived Organizational Injustice, Privacy Concern, Work Autonomy, and Trust

Organizational trust depends on employees' belief that the organization is competent, acts with integrity, respects data boundaries, treats them fairly, and preserves meaningful human involvement (Mayer et al., 1995). Unfair, intrusive, or controlling AI-HRM can therefore signal harmful intentions on the part of the organization, not simply a technical error.

The three experiences are conceptually distinct and should be modeled separately within Proposition 4. A single opaque performance system may simultaneously create injustice, privacy concern, and constrained autonomy, but each experience can contribute independently to organizational trust.

Social exchange theory suggests that employees withdraw relational trust when organizational practices violate expectations of fairness, legitimate data use, and autonomy (Cropanzano & Mitchell, 2005; Culnan & Armstrong, 1999).

Proposition 4: Perceived organizational injustice and employee privacy concern are likely to weaken organizational trust, whereas work autonomy is likely to strengthen it.

4.5.5 Direct Effects, Organizational Trust, and Employee Well-Being

Employee well-being comprises psychological conditions such as emotional balance and security, work-related conditions such as meaningfulness and manageable strain, and relational conditions such as belonging and supportive workplace relationships (Grant et al., 2007). Organizational trust can promote well-being by reducing uncertainty about how AI-HRM will be used and whether harmful outcomes will be corrected (Kinowska & Sienkiewicz, 2023; Mayer et al., 1995).

Trust is unlikely to be the sole pathway. Injustice can cause distress, privacy concern can generate vigilance and insecurity, and work autonomy can directly sustain motivation and meaningfulness (Demerouti et al., 2001; Glavin et al., 2024; Ravid et al., 2020). The framework therefore specifies direct and indirect effects separately.

Proposition 5: Perceived organizational injustice and employee privacy concern are likely to reduce employee well-being, whereas work autonomy is likely to improve it.

Proposition 6: Organizational trust is likely to improve psychological, work-related, and relational well-being and to partially mediate the effects of perceived injustice, privacy concern, and work autonomy on employee well-being.

4.5.6 Pathway-Specific Responsible AI-HRM Governance

Design and governance shape the negative consequences of AI-HRM. The four capabilities discussed below can reduce risk through better system and process design; the propositions examine how these capabilities moderate the focal relationships once a mechanism is present. This distinction separates the underlying mechanism from the organization's response to it (Bankins, 2021; Bujold et al., 2024; Robert et al., 2020).

Algorithmic accountability is particularly important in determining whether bias persists and how employees interpret it. Organizations cannot transfer responsibility to vendors; they must assign ownership, audit outcomes across groups, document decisions, and remedy disadvantage promptly (Bankins, 2021; Rodgers et al., 2023; Soleimani et al., 2025).

Proposition 7: Algorithmic accountability is likely to weaken the positive relationship between algorithmic bias and perceived

organizational injustice by requiring ownership, auditing, and timely remediation.

Opacity is not simply the opposite of transparency. The relevant governance capability is actionable explainability: employees need understandable, timely, and decision-relevant explanations, together with accessible procedures for review or appeal. Simply disclosing more information does not help when that information is technical, late, or impossible to act on (Langer & König, 2023; Leicht-Deobald et al., 2019).

Proposition 8: Actionable explainability is likely to weaken the positive relationship between AI-HRM opacity and perceived organizational injustice.

Privacy-preserving data governance should minimize collection, restrict access and retention, limit secondary use, ensure security, and keep data use proportionate to legitimate work purposes. Meaningful human oversight, in turn, should preserve employee voice and contextual judgment. To be substantive rather than symbolic, reviewers need the competence, authority, and accountability to challenge or change algorithmic outcomes (Bujold et al., 2024; Soleimani et al., 2025; Tursunbayeva et al., 2022).

Proposition 9: Privacy-preserving data governance is likely to weaken the positive relationships of workplace surveillance and privacy-invasive people analytics with employee privacy concern.

Proposition 10: Meaningful human oversight is likely to weaken the negative relationship between algorithmic control and work autonomy by preserving employee voice, contextual judgment, and genuine decision discretion.

Empirical tests of Propositions 1–6 should keep bias, opacity, justice dimensions, privacy concern, autonomy, organizational trust, and the three well-being dimensions distinct, with Proposition 6 estimated through specific indirect effects. Propositions 7–10 require separate interaction tests using indicators of accountability, actionable explainability, privacy governance, and meaningful human oversight.

4.6 Discussion

The framework moves AI-HRM research beyond technology-centered assessment toward an employee-centered socio-technical explanation. Efficiency, predictive accuracy, and speed are insufficient criteria when systems also redistribute information, discretion, and power. Standardized decisions are not necessarily fair when data reproduce disadvantage, models rely on inappropriate proxies, or affected people lack meaningful explanations and avenues for appeal (Köchling & Wehner, 2020; Newman et al., 2020; Starke et al., 2022).

Actionable explainability is therefore more precise than generic transparency. Similarly, monitoring may be accepted when it is job-relevant, proportionate, secure, and openly discussed, but it can undermine privacy and well-being when it is continuous, hidden, punitive, or repurposed (Glavin et al., 2024; Ravid et al., 2020; Vitak & Zimmer, 2023). Distinguishing surveillance from privacy-invasive analytics also avoids treating all people analytics as inherently intrusive.

Meaningful human oversight cannot be reduced to appointing a nominal reviewer. Reviewers must be competent, have the authority to alter outcomes, provide documented reasons, and remain accountable for their decisions. The model applies most directly to applicants and organizational employees. Evidence from platform work informs the control mechanism but should be tested separately because contractual and power relations differ (Kellogg et al., 2020; Parent-Rocheleau & Parker, 2022).

4.7 Practical Implications for Responsible AI-HRM

Before deployment, organizations should conduct impact assessments, map data flows, define legitimate purposes, test vendors and outcomes across groups, assign accountability, minimize data collection, disclose limitations, involve employees, and provide explanation, correction, and contestation channels.

After deployment, organizations should audit outcomes and model drift, review privacy and retention, investigate incidents, track appeals, and maintain accountable governance. In Indonesia, these practices should align with the UU PDP through documented purposes and accountability, relevant data use, protection, and applicable data-subject rights (Republic of Indonesia, 2022). Compliance is a floor rather than proof of fairness; high-impact employment decisions should remain AI-supported and provide explanation, appeal, and meaningful human review.

4.8 Limitations and Future Research Directions

This review is limited by its English-language, selected-database scope; non-exhaustive publisher searches; interpretive coding without formal inter-coder reliability; and combined applicant, workplace, organization-level, and platform evidence. Indonesian and Southeast Asian studies were scarce, so regional applicability remains theoretically informed. The fit-for-purpose appraisal was not a formal risk-of-bias assessment, and the framework remains conceptual and untested through 2025.

Future studies should test focused pathways through surveys, experiments, qualitative cases, SEM, longitudinal, or multilevel designs; distinguish justice and well-being dimensions and organizational from system trust; and compare Indonesian sectors, AI-HRM functions, and employment arrangements. Future systematic reviews should use multiple independent coders, database-specific logs, reason-specific exclusions, and PRISMA-based reporting.

5. Conclusion

This article explains how five AI-HRM dark-side mechanisms shape injustice, privacy concern, and autonomy, which may affect psychological, work-related, and relational well-being directly and indirectly through organizational trust.

Its contribution is an employee-centered socio-technical framework with clear construct boundaries and separately testable direct, mediated, and moderated paths, while recognizing cross-path effects and contextual limits.

Governance should match safeguards to each pathway. Initial studies should prioritize Propositions 2, 4, and 6: whether surveillance or privacy-invasive analytics raise privacy concern, whether that concern weakens organizational trust, and whether trust transmits the effect to well-being. Later studies can test Propositions 7–10 across HR functions and employment contexts.

ACKNOWLEDGEMENTS

The authors acknowledge the constructive academic feedback received during manuscript development.

REFERENCES

Andriani, P., Widia, E., & Patiro, S. P. S. (2023). Determination of procedural justice, interactional justice, distributive justice with work motivation as an intervening variable on employee job satisfaction at Raja Ahmad Tabib Hospital. *DeReMa (Development Research of Management): Jurnal Manajemen, 18*(1), 62–81. https://doi.org/10.19166/derema.v18i1.6574

Bankins, S. (2021). The ethical use of artificial intelligence in human resource management: A decision-making framework. *Ethics and Information Technology, 23*, 841–854. https://doi.org/10.1007/s10676-021-09619-6

Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context. *Information Systems Frontiers, 24*(3), 857–875. https://doi.org/10.1007/s10796-021-10223-8

Basu, S., Majumdar, B., Mukherjee, K., Munjal, S., & Palaksha, C. (2023). Artificial intelligence-HRM interactions and outcomes: A systematic review and causal configurational explanation. *Human Resource Management Review, 33*(1), 100893. https://doi.org/10.1016/j.hrmr.2022.100893

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons, 63*(2), 215–226. https://doi.org/10.1016/j.bushor.2019.12.001

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. https://doi.org/10.1080/09585192.2022.2035161

Bujold, A., Roberge-Maltais, I., Parent-Rocheleau, X., Boasen, J., Sénécal, S., & Léger, P.-M. (2024). Responsible artificial intelligence in human resources management: A review of the empirical literature. *AI and Ethics, 4*(4), 1185–1200. https://doi.org/10.1007/s43681-023-00325-1

Chowdhury, S., Dey, P. K., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review, 33*(1), 100899. https://doi.org/10.1016/j.hrmr.2022.100899

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386–400. https://doi.org/10.1037/0021-9010.86.3.386

Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management, 31*(6), 874–900. https://doi.org/10.1177/0149206305279602

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science, 10*(1), 104–115. https://doi.org/10.1287/orsc.10.1.104

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2020). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly, 65*(3), 644–676. https://doi.org/10.1177/0001839219867024

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal, 30*(1), 114–132. https://doi.org/10.1111/1748-

8583.12258
Einola, K., & Khoreva, V. (2023). Best friend or broken tool? Exploring the co-existence of humans and artificial intelligence in the workplace ecosystem. *Human Resource Management, 62*(1), 117–135. https://doi.org/10.1002/hrm.22147
Folger, N., Brosi, P., Stumpf-Wollersheim, J., & Welpe, I. M. (2022). Applicant reactions to digital selection methods: A signaling perspective on innovativeness and procedural justice. *Journal of Business and Psychology, 37*(4), 735–757. https://doi.org/10.1007/s10869-021-09770-3
Gagné, M., Parent-Rocheleau, X., Bujold, A., Gaudet, M.-C., & Lirio, P. (2022). How algorithmic management influences worker motivation: A self-determination theory perspective. *Canadian Psychology/Psychologie Canadienne, 63*(2), 247–260. https://doi.org/10.1037/cap0000324
Giermindl, L. M., Strich, F., Christ, O., Leicht-Deobald, U., & Redzepi, A. (2022). The dark sides of people analytics: Reviewing the perils for organisations and employees. *European Journal of Information Systems, 31*(3), 410–435. https://doi.org/10.1080/0960085X.2021.1927213
Glavin, P., Bierman, A., & Schieman, S. (2024). Private eyes, they see your every move: Workplace surveillance and worker well-being. *Social Currents, 11*(4), 327–345. https://doi.org/10.1177/23294965241228874
Grant, A. M., Christianson, M. K., & Price, R. H. (2007). Happiness, health, or relationships? Managerial practices and employee well-being tradeoffs. *Academy of Management Perspectives, 21*(3), 51–63. https://doi.org/10.5465/amp.2007.26421238
Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics, 178*(4), 977–1007. https://doi.org/10.1007/s10551-022-05049-6
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. https://doi.org/10.5465/annals.2018.0174
Kinowska, H., & Sienkiewicz, L. J. (2023). Influence of algorithmic management practices on workplace well-being: Evidence from European organisations. *Information Technology & People, 36*(8), 21–42. https://doi.org/10.1108/ITP-02-2022-0079
Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research, 13*, 795–848. https://doi.org/10.1007/s40685-020-00134-w
Köchling, A., Wehner, M. C., & Rühle, S. A. (2025). This (AI)n't fair? Employee reactions to artificial intelligence (AI) in career development systems. *Review of Managerial Science, 19*(4), 1195–1228. https://doi.org/10.1007/s11846-024-00789-3
Langer, M., & König, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency and strategies to reduce opacity in algorithm-based human resource management. *Human Resource Management Review, 33*(1), 100881. https://doi.org/10.1016/j.hrmr.2021.100881
Langer, M., König, C. J., & Hemsing, V. (2020). Is anybody listening? The impact of automatically evaluated job interviews on impression management and applicant reactions. *Journal of Managerial Psychology, 35*(4), 271–284. https://doi.org/10.1108/JMP-03-2019-0156
Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment, 27*(3), 217–234. https://doi.org/10.1111/ijsa.12246
Lavanchy, M., Reichert, P., Narayanan, J., & Savani, K. (2023). Applicants' fairness

perceptions of algorithm-driven hiring procedures. *Journal of Business Ethics, 188*(1), 125–150. https://doi.org/10.1007/s10551-022-05320-w

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics, 160*(2), 377–392. https://doi.org/10.1007/s10551-019-04204-w

Ly-Le, T.-M., & Le, H. C. P. (2024). The impact of internal communication on Vietnamese Gen Z employees' work satisfaction. *DeReMa (Development Research of Management): Jurnal Manajemen, 19*(2), 113–130. https://doi.org/10.19166/derema.v19i2.8165

Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review, 33*(1), 100940. https://doi.org/10.1016/j.hrmr.2022.100940

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

Meijerink, J., & Bondarouk, T. (2023). The duality of algorithmic management: Toward a research agenda on HRM algorithms, autonomy and value creation. *Human Resource Management Review, 33*(1), 100876. https://doi.org/10.1016/j.hrmr.2021.100876

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: Synthesizing developments and cross-disciplinary insights on digital HRM. *The International Journal of Human Resource Management, 32*(12), 2545–2562. https://doi.org/10.1080/09585192.2021.1925326

Mori, M., Sassetti, S., Cavaliere, V., & Bonti, M. (2025). A systematic literature review on artificial intelligence in recruiting and selection: A matter of ethics. *Personnel Review, 54*(3), 854–878. https://doi.org/10.1108/PR-03-2023-0257

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, 160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Oostrom, J. K., Holtrop, D., Koutsoumpis, A., van Breda, W., Ghassemi, S., & de Vries, R. E. (2024). Applicant reactions to algorithm- versus recruiter-based evaluations of an asynchronous video interview and a personality inventory. *Journal of Occupational and Organizational Psychology, 97*(1), 160–189. https://doi.org/10.1111/joop.12465

Ore, O., & Sposato, M. (2022). Opportunities and risks of artificial intelligence in recruitment and selection. *International Journal of Organizational Analysis, 30*(6), 1771–1782. https://doi.org/10.1108/IJOA-07-2020-2291

Pan, Y., & Froese, F. J. (2023). An interdisciplinary review of AI and HRM: Challenges and future directions. *Human Resource Management Review, 33*(1), 100924. https://doi.org/10.1016/j.hrmr.2022.100924

Parent-Rocheleau, X., & Parker, S. K. (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review, 32*(3), 100838. https://doi.org/10.1016/j.hrmr.2021.100838

Pessach, D., & Shmueli, E. (2022). A review on fairness in machine learning. *ACM Computing Surveys, 55*(3), 51. https://doi.org/10.1145/3494672

Prikshat, V., Islam, M., Patel, P., Malik, A., Budhwar, P., & Gupta, S. (2023). AI-augmented HRM: Literature review and a proposed multilevel framework for future research. *Technological Forecasting and Social Change, 193*, 122645. https://doi.org/10.1016/j.techfore.2023.122645

Qamar, Y., Agrawal, R. K., Samad, T. A., & Chiappetta Jabbour, C. J. (2021). When technology meets people: The interplay of artificial intelligence and human resource

management. *Journal of Enterprise Information Management, 34*(5), 1339–1370. https://doi.org/10.1108/JEIM-11-2020-0436

Qin, M. S., Jia, N., Luo, X., Liao, C., & Huang, Z. (2023). Perceived fairness of human managers compared with artificial intelligence in employee performance evaluation. *Journal of Management Information Systems, 40*(4), 1039–1070. https://doi.org/10.1080/07421222.2023.2267316

Ravid, D. M., Tomczak, D. L., White, J. C., & Behrend, T. S. (2020). EPM 20/20: A review, framework, and research agenda for electronic performance monitoring. *Journal of Management, 46*(1), 100–126. https://doi.org/10.1177/0149206319869435

Republic of Indonesia. (2022). *Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi*. https://peraturan.bpk.go.id/Details/229798/uu-no-27-tahun-2022

Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human-Computer Interaction, 35*(5–6), 545–575. https://doi.org/10.1080/07370024.2020.1735391

Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review, 33*(1), 100925. https://doi.org/10.1016/j.hrmr.2022.100925

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

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society, 9*(2), 20539517221115189. https://doi.org/10.1177/20539517221115189

Strohmeier, S. (2020). Digital human resource management: A conceptual clarification. *German Journal of Human Resource Management, 34*(3), 345–365. https://doi.org/10.1177/2397002220921131

Suen, H.-Y., Chen, M. Y.-C., & Lu, S.-H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior, 98*, 93–101. https://doi.org/10.1016/j.chb.2019.04.012

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

Tinguely, P. N., Lee, J., & He, V. F. (2023). Designing human resource management systems in the age of AI. *Journal of Organization Design, 12*(4), 263–269. https://doi.org/10.1007/s41469-023-00153-x

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting. *Human Relations, 4*(1), 3–38. https://doi.org/10.1177/001872675100400101

Tursunbayeva, A., Pagliari, C., Di Lauro, S., & Antonelli, G. (2022). The ethics of people analytics: Risks, opportunities and recommendations. *Personnel Review, 51*(3), 900–921. https://doi.org/10.1108/PR-12-2019-0680

Van Esch, P., & Black, J. S. (2019). Factors that influence new generation candidates to engage with and complete digital, AI-enabled recruiting. *Business Horizons, 62*(6), 729–739. https://doi.org/10.1016/j.bushor.2019.07.004

Varma, A., Dawkins, C., & Chaudhuri, K. (2023). Artificial intelligence and people management: A critical assessment through the ethical lens. *Human Resource*

Management Review, 33(1), 100923. https://doi.org/10.1016/j.hrmr.2022.100923

Vitak, J., & Zimmer, M. (2023). Surveillance and the future of work: Exploring employees' attitudes toward monitoring in a post-COVID workplace. *Journal of Computer-Mediated Communication, 28*(4), zmad007. https://doi.org/10.1093/jcmc/zmad007

Vrontis, D., Christofi, M., Pereira, V., Tarba, S. Y., Makrides, A., & Trichina, E. (2022). Artificial intelligence, robotics, advanced technologies and human resource management: A systematic review. *The International Journal of Human Resource Management, 33*(6), 1237–1266. https://doi.org/10.1080/09585192.2020.1871398

World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum. https://www.weforum.org/publications/the-future-of-jobs-report-2025/

Zhou, Y., Wang, L., & Chen, W. (2023). The dark side of AI-enabled HRM on employees based on AI algorithmic features. *Journal of Organizational Change Management, 36*(7), 1222–1241. https://doi.org/10.1108/JOCM-10-2022-0308

APPENDIX A. Study-Level Coding Matrix (52 Journal Articles, 2019–2025)

No.	Study	Code	No.	Study	Code	No.	Study	Code
1	Bankins (2021)	C|FAIR-X	19	Köchling et al. (2025)	E|CAREER-W	37	Qin et al. (2023)	E|PERF-W
2	Bankins et al. (2022)	E|FAIR-W	20	Langer & König (2023)	C|OPAQ-X	38	Ravid et al. (2020)	R|SURV-W
3	Basu et al. (2023)	R|AHRM-X	21	Langer et al. (2020)	E|REC-A	39	Robert et al. (2020)	R|FAIR-W
4	Black & van Esch (2020)	R|REC-A	22	Langer et al. (2019)	E|REC-A	40	Rodgers et al. (2023)	C|FAIR-X
5	Budhwar et al. (2022)	R|AHRM-X	23	Lavanchy et al. (2023)	E|REC-A	41	Soleimani et al. (2025)	E-G|REC-A
6	Bujold et al. (2024)	R|FAIR-X	24	Leicht-Deobald et al. (2019)	C|FAIR-W	42	Starke et al. (2022)	R|FAIR-X
7	Chowdhury et al. (2023)	C|AHRM-X	25	Malik et al. (2023)	C|AHRM-X	43	Strohmeier (2020)	C|AHRM-X
8	Curchod et al. (2020)	E-Q|CTRL-P	26	Meijerink & Bondarouk (2023)	C|CTRL-P	44	Suen et al. (2019)	E|REC-A
9	Duggan et al. (2020)	C|CTRL-P	27	Meijerink et al. (2021)	R|AHRM-X	45	Tambe et al. (2019)	R|AHRM-X
10	Einola & Khoreva (2023)	C|AHRM-W	28	Mori et al. (2025)	R|REC-A	46	Tinguely et al. (2023)	C|AHRM-X
11	Folger et al. (2022)	E|REC-A	29	Newman et al. (2020)	M|FAIR-W	47	Tursunbayeva et al. (2022)	R|PRIV-W
12	Gagné et al. (2022)	C|CTRL-X	30	Oostrom et al. (2024)	E|REC-A	48	Van Esch & Black (2019)	E|REC-A
13	Giermindl et al. (2022)	R|PRIV-W	31	Ore & Sposato (2022)	R|REC-A	49	Varma et al. (2023)	C|FAIR-X
14	Glavin et al. (2024)	E|SURV-W	32	Pan & Froese (2023)	R|AHRM-X	50	Vitak & Zimmer (2023)	E|SURV-W
15	Hunkenschroer & Luetge (2022)	R|REC-A	33	Parent-Rocheleau & Parker (2022)	C|CTRL-W	51	Vrontis et al. (2022)	R|AHRM-X
16	Kellogg et al. (2020)	R|CTRL-X	34	Pessach & Shmueli (2022)	R|FAIR-X	52	Zhou et al. (2023)	M|AHRM-W
17	Kinowska & Sienkiewicz (2023)	E-S|WELL-W	35	Prikshat et al. (2023)	R|AHRM-X			
18	Köchling & Wehner (2020)	R|REC-A	36	Qamar et al. (2021)	R|AHRM-X			

Design codes: E = empirical, C = conceptual/theoretical, R = review/synthesis, M = mixed, E-Q = qualitative, E-S = SEM-CFA, and E-G = grounded theory. Thematic codes: AHRM = AI-HRM/digital HRM, REC = recruitment/selection, CTRL = algorithmic control/management, FAIR = fairness/justice, PRIV = privacy/people analytics, SURV = surveillance, WELL = well-being, OPAQ = opacity, CAREER = career development, and PERF = performance evaluation. Context codes: A = applicants/recruitment, W = workplace, X = cross-context, and P = platform/gig. This three-block table preserves the complete 52-study coding record with fewer columns and a clearer layout than the original matrix. Source: Authors' synthesis (2026).