

DOI 10.31558/2307-2318.2026.2.8

УДК 005.95:004.8

JEL Classification: M12; M15; O33; C88

Ilich L.

Doctor of Economic Sciences, Professor,
Borys Grinchenko Kyiv Metropolitan University
ORCID: 0000-0002-8594-1824
e-mail: l.ilich@kubg.edu.ua

A NEW PARADIGM OF PERSONNEL EVALUATION: FROM HR METRICS TO AI AND EMPATHY-LED LEADERSHIP

The article examines the limitations of traditional personnel evaluation systems in the context of contemporary organisations characterised by dynamic work environments, digitalisation, and increasing reliance on data-driven decision-making. Conventional appraisal approaches are critically assessed in terms of subjectivity, episodic nature, and insufficient developmental orientation, which reduce their effectiveness and credibility. The study aims to develop and theoretically justify an Integrated Personnel Evaluation Model (IPEM) that combines HR metrics, AI-driven People Analytics, and empathy-led leadership within a unified socio-technical framework. The research is based on an integrative literature review, comparative analysis, and conceptual modelling. The proposed model consists of three interrelated layers: HR metrics and goal architecture, an AI-based analytics engine, and empathy-led leadership. The model also incorporates a governance layer ensuring fairness, transparency, and regulatory compliance. The findings demonstrate that effective evaluation systems should evolve toward continuous, data-informed, and development-oriented processes. The study concludes that the integration of analytical precision with human-centred leadership is essential for building valid, trustworthy, and strategically relevant personnel evaluation systems in the digital economy.

Keywords: personnel evaluation; HR Metrics; artificial intelligence in HRM; empathy-led leadership; performance management; socio-technical systems; algorithmic fairness; psychological safety.

Tab. – 2, Lit. – 28.

Ільїч Л.М.

доктор економічних наук, професор,
Київський столичний університет імені Бориса Грінченка
ORCID: 0000-0002-8594-1824
e-mail: l.ilich@kubg.edu.ua

НОВА ПАРАДИГМА ОЦІНЮВАННЯ ПЕРСОНАЛУ: ВІД HR-МЕТРИК ДО ШІ ТА ЕМПАТІЙНОГО ЛІДЕРСТВА

У статті досліджуються обмеження традиційних систем оцінювання персоналу в умовах сучасних організацій, що характеризуються динамічними робочими середовищами, цифровізацією та зростаючою залежністю від прийняття рішень на основі даних. Традиційні підходи до оцінювання критично проаналізовано з позицій їх

суб'єктивності, епізодичності та недостатньої орієнтації на розвиток, що знижує їх ефективність і довіру до результатів. Метою дослідження є розробка та теоретичне обґрунтування Інтегрованої моделі оцінювання персоналу (ІРЕМ), яка поєднує HR-метрики, аналітику персоналу на основі штучного інтелекту та емпатійне лідерство в межах єдиної соціотехнічної рамки. Дослідження ґрунтується на інтегративному огляді літератури, порівняльному аналізі та концептуальному моделюванні. Запропонована модель складається з трьох взаємопов'язаних рівнів: HR-метрики та архітектура цілей, аналітичний модуль на основі штучного інтелекту та емпатійне лідерство. Модель також включає наскрізний рівень управління, що забезпечує справедливість, прозорість і відповідність нормативним вимогам. Результати дослідження свідчать про те, що ефективні системи оцінювання мають розвиватися у напрямі безперервності, опори на дані та розвитку працівників. У роботі обґрунтовано, що інтеграція аналітичної точності з людиноцентричним лідерством є ключовою умовою формування валідних, надійних і стратегічно релевантних систем оцінювання персоналу в цифровій економіці.

Ключові слова: оцінювання персоналу; HR-метрики; штучний інтелект в управлінні персоналом; емпатійне лідерство; управління результативністю; соціотехнічні системи; алгоритмічна справедливість; психологічна безпека.

Табл. – 2, Літ. – 28.

Problem Statement. Performance evaluation occupies a central position within strategic Human Resource Management, functioning as the primary institutional mechanism through which organisations calibrate talent decisions, allocate developmental resources, and legitimise reward distribution. However, a growing body of research suggests that this function is experiencing not a series of isolated inefficiencies, but a systemic misalignment with the realities of contemporary work. Traditional instruments, including annual reviews, supervisor-driven rating scales, and static goal-setting frameworks, are increasingly incompatible with dynamic, collaborative, and digitally mediated organisational environments.

This misalignment is further exacerbated by well-documented cognitive and structural limitations inherent in conventional appraisal systems. Performance evaluations remain vulnerable to persistent biases, including halo effects, recency bias, and leniency bias, which compromise rating validity and reduce the credibility of evaluation outcomes. More critically, the coexistence of evaluative and developmental objectives within a single episodic event creates a fundamental contradiction: systems intended to support employee growth often generate anxiety, impression management behaviours, and disengagement. These outcomes point not to implementation failures, but to deeper design limitations embedded within traditional performance management architectures.

Tambe et al. (2019) and van den Broek et al. (2021) argue that the rapid advancement of AI and the expansion of workplace data infrastructures have opened new possibilities for rethinking personnel evaluation. In particular, contemporary developments in people analytics enable organisations to move beyond retrospective assessments toward continuous, data-driven, and predictive evaluation processes [1, 2]. These capabilities offer the potential to enhance objectivity, increase evaluation frequency, and provide more actionable developmental insights. However, while the technical capacities of AI-enabled evaluation are increasingly well understood, their integration into systems that are trustworthy, transparent, and psychologically sustainable remains insufficiently theorised.

Raghavan et al. (2020) and Wachter et al. (2021) highlight that this gap becomes particularly evident when considering the fundamental tensions that define modern evaluation systems. Specifically, a key tension emerges between algorithmic objectivity and relational legitimacy: while AI promises standardisation and bias reduction, algorithmic decision-making often lacks transparency and may be perceived by employees as less fair and less trustworthy

than human judgment. [3, 4]. Kellogg et al. (2020) and Duggan et al. (2020) demonstrate that a critical tension in modern evaluation systems arises between continuous data capture and trust. While advanced data infrastructures enable real-time performance monitoring, they simultaneously intensify concerns related to surveillance, privacy, and autonomy, potentially undermining psychological safety and employee engagement [5, 6]. Third, there is a tension between evaluation as control and evaluation as development. Data-driven systems are frequently deployed to monitor compliance and efficiency rather than to support learning and growth, thereby reinforcing instrumental rather than developmental logics of performance management.

Despite the growing body of literature on AI in HRM, algorithmic governance, and performance management transformation, existing research remains fragmented. Prior studies tend to examine technological, behavioural, or ethical dimensions in isolation, without offering a coherent framework capable of integrating these elements into a unified evaluation system. As a result, organisations often face persistent trade-offs: metrics without trust lead to formalism, AI without governance risks unfairness, and human-centred approaches without data may result in subjectivity and inconsistency.

To address this gap, the present study proposes the Integrated Personnel Evaluation Model (IPEM) – a socio-technical framework that combines HR metrics, AI-driven analytics, and empathy-led leadership into a coherent system of personnel evaluation. By reconciling analytical rigor with human-centred leadership and psychological sustainability, the model contributes to the development of more valid, trusted, and development-oriented evaluation practices in contemporary organisations.

Analysis of Recent Studies and Publications. Modern approaches to personnel evaluation are grounded in the long-standing evolution of performance measurement paradigms. One of the foundational contributions is the Balanced Scorecard framework proposed by Kaplan and Norton, which introduced a multidimensional perspective on performance by integrating financial and non-financial indicators, including internal processes, learning and growth, and customer-related outcomes. Within this framework, personnel evaluation shifts from a purely administrative function toward a strategic management tool [7].

This perspective is further reinforced by the goal-setting theory developed by Locke and Latham (2002), who demonstrated that specific and challenging goals, when accepted by employees, significantly enhance motivation and performance. This implies that effective personnel evaluation systems must be closely aligned with goal-setting processes and organizational objectives [8]. Consequently, contemporary evaluation systems increasingly focus not only on final ratings but also on goal alignment, continuous performance tracking, and adaptive feedback mechanisms.

The limitations of traditional annual appraisal systems have been critically examined by Cappelli and Tavis (2016), who argue that conventional performance management models are increasingly misaligned with the dynamic nature of modern organizations [9]. Annual reviews are often retrospective, rigid, and insufficiently responsive to rapidly changing roles and tasks. As a result, organizations are transitioning toward continuous performance management systems characterized by frequent feedback, shorter evaluation cycles, and greater flexibility in performance criteria.

At the same time, the rise of HR analytics has introduced new expectations regarding the role of data in evaluation processes. Angrave et al. (2016) caution that the mere availability of large HR datasets does not automatically translate into better decision-making. Without strong analytical capabilities, theoretical grounding, and integration into organizational strategy, HR analytics risks becoming fragmented and ineffective. Therefore, HR metrics should be understood not as technical instruments alone but as part of a broader intellectual and strategic infrastructure for performance evaluation [10].

Overall, personnel evaluation has evolved from a static administrative procedure into a dynamic system that integrates strategic metrics, goal alignment, development, and continuous feedback. However, measurement alone does not resolve issues of objectivity, interpretation, and trust, which opens the space for the integration of AI and people analytics.

A significant transformation in personnel evaluation is driven by the development of people analytics and the integration of AI into HRM. Tambe, Cappelli, and Yakubovich conceptualize AI in HR as a qualitatively new stage in which analytical systems move beyond descriptive functions to predictive and prescriptive capabilities, enabling organizations to identify patterns, forecast outcomes, and support decision-making in recruitment, development, retention, and evaluation [1]. In the context of personnel evaluation, this implies a shift from episodic, subjective assessments toward continuous, data-driven, and predictive evaluation systems.

However, Angrave et al. (2016) emphasize that the value of HR analytics depends on its meaningful integration into organizational decision-making processes rather than its technical sophistication alone [10]. Data without context, interpretation, and strategic alignment may fail to improve evaluation outcomes.

The ethnographic study by van den Broek, Sergeeva, and Huysman (2021) provides deeper insight into the human–AI interaction in HR processes. Their findings suggest that AI systems are not autonomous or neutral but are shaped by expert assumptions, organizational contexts, and interpretive practices. Rather than replacing human judgment, AI creates hybrid decision-making environments where human expertise and algorithmic outputs are intertwined [3]. This is particularly relevant for personnel evaluation, where interpretation and contextual understanding remain essential.

Kellogg, Valentine, and Christin (2020) further demonstrate that the increasing use of algorithms in organizational settings introduces a new form of control, where work is monitored, evaluated, and regulated through digital systems. Algorithmic systems not only coordinate tasks but also discipline behavior, creating tensions between efficiency and autonomy [6]. From a performance evaluation perspective, this leads to a paradox: while AI enhances precision and frequency of evaluation, it may also contribute to surveillance, depersonalization, and reduced employee agency.

Similarly, Duggan et al. (2020) analyse algorithmic management in the gig economy, highlighting how digital systems reshape power relations, transparency, and employee experiences. Although their focus is on platform work, their insights are broadly applicable to HRM. Algorithmic management can increase efficiency and standardization but may simultaneously undermine autonomy, perceived fairness, and employee engagement [6].

Thus, while AI and people analytics offer significant opportunities to enhance the accuracy, scalability, and predictive value of personnel evaluation, they also introduce critical challenges related to control, trust, interpretation, and fairness. These challenges necessitate a closer examination of bias, governance, and ethical considerations.

The integration of AI into personnel evaluation raises fundamental concerns regarding bias, fairness, and governance. Raghavan et al. (2020) critically examines claims about the neutrality of algorithmic hiring systems and demonstrate that such systems may reproduce or even amplify existing biases if trained on historically biased data [3]. Their findings challenge the assumption that algorithmic decision-making inherently leads to more objective outcomes.

Wachter et al. (2021) extend this discussion by arguing that fairness cannot be fully automated. They highlight a fundamental gap between technical definitions of fairness in AI systems and legal or ethical interpretations of non-discrimination. Even when algorithms satisfy formal fairness criteria, they may still produce outcomes that are considered unjust in broader social or legal contexts [5]. This insight is particularly important for personnel

evaluation, where decisions have direct implications for careers, well-being, and organizational trust.

Moreover, the increasing reliance on AI systems introduces governance challenges, especially when decision-making processes become opaque. In such cases, managers may rely on algorithmic outputs without fully understanding their underlying logic, while employees may lack transparency regarding evaluation criteria. This creates risks not only of bias but also of diminished trust in evaluation systems. Therefore, effective AI-enabled evaluation requires robust governance mechanisms, including data quality control, model auditing, transparency, and the preservation of human oversight.

In sum, the literature suggests that the challenge of AI in personnel evaluation is not merely technical but fundamentally socio-technical, requiring the integration of analytical capability with ethical, legal, and organizational legitimacy.

Despite the increasing role of metrics and AI, effective personnel evaluation cannot be achieved without a strong human and relational dimension. A central concept in this context is psychological safety. Edmondson (1999) defines psychological safety as a shared belief that the team environment is safe for interpersonal risk-taking, allowing individuals to speak up, admit mistakes, and share ideas without fear of negative consequences [11]. In evaluation contexts, psychological safety is critical, as it determines whether feedback is perceived as a threat or as an opportunity for growth.

A meta-analytic review by Frazier et al. (2017) confirms that psychological safety is positively associated with learning behaviour, knowledge sharing, engagement, and performance. This indicates that psychological safety is not merely a “soft” factor but a key mechanism through which performance and development are achieved [12].

In parallel, emerging research on employee well-being highlights its strategic importance for organizational performance. De Neve, Kaats, and Ward (2024) provide empirical evidence of a positive relationship between workplace well-being and firm performance, including profitability, return on assets, and market valuation [13]. Their findings suggest that employee well-being should be considered not as a peripheral concern but as a central component of performance evaluation systems.

Within this context, empathy-led leadership can be conceptualized as a managerial approach that enables the integration of evaluation, development, and well-being. It supports psychological safety, facilitates constructive feedback, and ensures that evaluation processes are perceived as fair and meaningful. Importantly, empathy-led leadership does not replace metrics or AI but complements them by providing the human interpretive layer necessary for responsible and developmental evaluation.

The reviewed literature reveals three major, yet largely fragmented, streams of research. The first stream focuses on HR metrics and performance management, emphasizing measurement, goal alignment, and strategic evaluation systems [7, 8, 9, 10]. The second stream examines AI, people analytics, and algorithmic management, highlighting both the opportunities and risks associated with data-driven evaluation [1, 2, 3, 4, 5, 6]. The third stream emphasizes psychological safety, well-being, and developmental evaluation, focusing on trust, learning, and employee experience [11, 12, 13].

However, a critical gap emerges at the intersection of these streams. Existing studies provide valuable insights into metrics, AI, or human-centred leadership separately, but they rarely offer an integrated framework that combines analytical rigor, technological capability, and human-centred leadership within a unified evaluation system. As a result, organizations often face three persistent tensions: metrics without trust lead to formalism; AI without governance risks unfairness; and empathy without data may result in subjectivity.

Therefore, there is a clear need for a comprehensive approach to personnel evaluation that conceptualizes it as a socio-technical system, where HR metrics, AI, and empathy-led

leadership are aligned and mutually reinforcing. Addressing this gap constitutes the core contribution of the present study.

Research Aim. The primary objective of this article is to develop and theoretically justify an IPEM that synthesises the quantitative rigour of HR Metrics, the predictive and diagnostic power of AI-driven People Analytics, and the relational and developmental orientation of empathy-led leadership. Secondary objectives include: (1) providing a systematic comparative analysis of traditional and modern evaluation approaches; (2) identifying the mechanisms through which each IPEM component addresses specific limitations of legacy systems; and (3) articulating the practical conditions and organisational capabilities required for effective model implementation.

Presentation of the Main Material. Traditional performance evaluation systems, most commonly operationalised through annual appraisals, graphic rating scales, and forced distribution rankings, exhibit structural limitations that are increasingly problematised in contemporary HRM research. While historically viable in relatively stable organisational contexts, these systems demonstrate a growing misalignment with the complexity, dynamism, and data-rich nature of modern work environments.

Scullen et al. (2000) demonstrated that performance ratings are influenced not only by employees' actual job performance but also by rater biases and measurement error; notably, a substantial proportion of variance in ratings is explained by rater-specific effects rather than "true" performance [14]. Subsequent research confirms that performance appraisal remains vulnerable to systematic distortions, including leniency bias, and that such evaluations are often shaped by social context and procedural factors. At the same time, more recent studies on assessment center judgments suggest that behavioral signals may explain a meaningful share of rating variance more effectively than bias-based models alone. Therefore, it is more appropriate to conceptualise biases not as dominant, but as significant factors influencing rating decisions.

Murphy and Cleveland (1995) argue that when evaluative and developmental functions are combined within the same performance management procedure, employees tend to interpret the process primarily as control-oriented, which reduces openness to feedback and limits learning-oriented responses [15]. Building on this perspective, Parent-Rocheleau and Parker (2022) demonstrate that traditional performance appraisal systems are often misaligned with contemporary work arrangements, particularly in contexts characterised by short project cycles, fluid roles, and continuous feedback [16]. In such environments, infrequent review cycles fail to capture meaningful fluctuations in performance and reduce the timeliness and developmental value of feedback.

HR metrics provide the quantitative scaffolding upon which systematic personnel evaluation must be built. Contemporary frameworks distinguish between input metrics (measuring resources and capabilities brought to a role), process metrics (capturing behavioural quality and efficiency), and output metrics (assessing tangible results). An effective evaluation system integrates all three categories rather than relying exclusively on output indicators, which are often influenced by contextual and systemic factors beyond individual control.

Doerr (2018) conceptualises the OKR framework as a dynamic architecture for performance evaluation that has gained increasing prominence in contemporary organisations [17]. Unlike traditional KPI systems, OKRs combine aspirational goal setting with time-bound and measurable key results, thereby enabling continuous tracking of progress and periodic recalibration. Marler and Boudreau (2017) and van den Broek et al. (2021) argue that the effectiveness of such systems depends not only on the formal design of metrics but also on their integration into broader digital and organisational ecosystems [18, 2]. Kellogg et al. (2020) and Pulakos et al. (2015) further demonstrate that, in digitally enabled performance

management environments, real-time monitoring and adaptive goal alignment shift evaluation away from static measurement toward continuous performance calibration [5, 19].

However, the effectiveness of HR metrics remains contingent upon several prerequisites, including participatory goal-setting (Latham & Locke, 2007), rigorous data quality governance, and sufficient analytical literacy among HR professionals and line managers [20]. Bankins et al. (2022, 2024) and Chowdhury et al. (2023) suggest that, in the absence of such conditions, metric systems are more likely to reinforce compliance-oriented behaviours than to promote developmental learning [21-23]. This risk is particularly pronounced when evaluation tools are perceived primarily as instruments of control rather than as mechanisms for fair and meaningful feedback. More broadly, their work indicates that the effectiveness of metrics depends not only on technical design but also on their embedding within organisational cultures that support transparency, trust, and responsible decision-making.

People analytics represents an important extension of the evidential basis for personnel evaluation by enabling organisations to integrate multiple data sources into HR decision-making processes. Rigamonti et al. (2024) suggest that the value of such systems depends not only on their technical sophistication but also on the organisation's ability to translate data into coherent HR practices and decision routines [23]. Compared with traditional episodic appraisal systems, mature analytics infrastructures support more continuous and data-informed management processes. More broadly, van den Broek et al. (2021) indicate that the effectiveness of AI-enabled HR processes is shaped by the interaction between algorithms, human expertise, and organisational context [2].

Raghavan et al. (2020), Wachter et al. (2021), Bankins et al. (2022), and Madad et al. (2024) emphasise that natural language processing (NLP) applications represent a particularly significant extension of AI-enabled evaluation systems. NLP-based tools enable the analysis of communication content, sentiment, and interaction patterns, thereby substantially enriching the informational basis of performance assessment [3, 4, 24, 25]. However, their effectiveness remains highly contingent upon data quality, representativeness, and careful contextual interpretation. These studies further highlight that the explainability challenge is especially acute in this domain. When algorithmic systems generate behavioural inferences, managers require explanations that are both interpretable and actionable in order to translate these insights into meaningful developmental interventions.

Moreover, the responsible deployment of AI in personnel evaluation necessitates a multidimensional governance approach encompassing algorithmic fairness, transparency, and employee trust. These dimensions are deeply interdependent: failures in bias mitigation undermine perceived fairness and erode employee trust, while insufficient transparency constrains interpretability and limits the practical usability of AI-generated insights. Taken together, these considerations demonstrate that effective AI-enabled evaluation systems must integrate technical robustness with relational legitimacy, a challenge directly addressed by the IPEM Governance Layer.

The integration of AI and HR metrics into personnel evaluation creates a risk that is rarely acknowledged in the technical literature: the potential displacement of human judgement and relational quality in favour of algorithmic efficiency. Research on algorithmic management documents consistent negative effects on employee autonomy, trust, and professional identity, particularly where algorithmic oversight is pervasive (Duggan et al., 2020; Parent-Rocheleau & Parker, 2022). Bankins et al. (2022) demonstrate that perceived algorithmic surveillance significantly reduces psychological safety, the condition identified by Edmondson (1999) and Frazier et al. (2017) as a key enabler of authentic performance and developmental learning [6, 16, 21].

Empathy-led leadership provides a critical relational and developmental counterbalance that prevents AI-augmented evaluation systems from reproducing, at greater scale and speed, the dehumanising dynamics of earlier paradigms. Kniffin et al. (2021) demonstrate that compassionate leadership moderates the relationship between workplace stressors and both performance and well-being outcomes [26], indicating that empathy functions not only as a relational mechanism but also as a strategic performance driver.

Well-being integration into performance evaluation represents a specific application of the empathy-led approach with growing empirical support. De Neve et al. (2023) demonstrate that employee well-being is a significant predictor of both individual and organisational performance, providing a strong evidence base for incorporating well-being indicators into evaluation frameworks [13]. Dewe and Cooper (2021) further show that conventional evaluation systems are structurally incapable of addressing the well-being challenges of contemporary knowledge work, as they are designed to measure output rather than sustainable human capacity [27]. These findings provide the empirical foundation for the IPeM's incorporation of well-being monitoring as a core component of Layer 3.

Drawing on the preceding analysis and relevant theoretical frameworks, this section proposes the IPeM as a conceptual framework designed to address the documented limitations of traditional evaluation systems through the joint optimisation of three interdependent layers.

The model integrates HR metrics and goal architecture, AI-driven people analytics, and empathy-led leadership and human dialogue. Together, these layers constitute a socio-technical system in which quantitative measurement, algorithmic processing, and human-centered interpretation are systematically aligned.

Table 1 presents the structural architecture of the model and outlines the functional contribution of each layer to the overall evaluation process.

Layer 1 (HR-metrics and goal architecture) encompasses the quantitative data systems through which employee performance is captured and represented. This layer integrates HR metrics (KPIs, OKRs, productivity indicators) with continuous data streams from People Analytics platforms, including communication analytics, learning system engagement data, project management records, and 360-degree feedback aggregates. Its primary function is to create a multi-dimensional, temporally granular representation of employee contribution, consistent with the evidence-based HRM imperative for rigorous and internally valid measurement [28]. Crucially, metrics in this layer are co-developed with employees through participatory goal-setting processes [8], reducing the risk of Goodhartian distortion and strengthening goal commitment.

Layer 2 (AI / People Analytics Engine) applies machine learning and NLP algorithms to the data infrastructure generated by Layer 1, performing three analytical functions. Descriptive analytics provide real-time performance dashboards summarising current trajectories at individual and team levels. Predictive analytics identify leading indicators of attrition risk, performance decline, and skill development needs, enabling proactive talent management [1]. Prescriptive analytics generate contextualised development recommendations and intervention triggers, translating analytical patterns into actionable HR guidance. The layer is designed in accordance with human-centred AI principles, incorporating explainability, contestability, and fairness auditing mechanisms. Explainability ensures that managers can understand the basis of algorithmic recommendations, contestability enables employees to challenge algorithmic conclusions, and fairness audits apply demographic parity testing and bias detection to all model outputs [3].

Table 1- Integrated Personnel Evaluation Model (IPEM): Structural Architecture and Functional Logic

Model layer	Components	Function
LAYER 1 HR-metrics and goal architecture	KPIs (input, process, output) OKRs Balanced Scorecard indicators Engagement indices 360° feedback aggregates Competency ratings	Establishes the quantitative measurement foundation; translates multi-dimensional performance into governable data constructs; provides the raw evidential basis for AI-layer processing.
▼ Bidirectional Feedback and Data Flow ▲		
LAYER 2 AI / People Analytics Engine	Descriptive real-time dashboards Predictive attrition and performance models NLP sentiment analysis Skill-gap identification algorithms Algorithmic fairness audits Prescriptive development recommendations	Processes and enriches Layer 1 data; generates descriptive, predictive, and prescriptive insights; applies bias-mitigation and explainability protocols; surfaces non-obvious performance patterns for managerial action.
▼ Bidirectional Feedback and Data Flow ▲		
LAYER 3 Empathy-Led Leadership and Human Dialogue	Continuous micro-feedback conversations Employee self-reflection against OKRs Well-being & psychological safety monitoring Manager empathy accountability Developmental planning dialogues	Translates analytical insights into relational, developmental interactions; ensures evaluation is experienced as growth-oriented; maintains human-centred governance of AI-generated outputs.

Note: The governance layer is cross-cutting across all levels of the model and includes algorithmic fairness audits, employee privacy protection, compliance with GDPR and the EU AI Act, as well as institutional mechanisms such as privacy and ethics committees and equity monitoring systems.

Source: Developed by the author based on theoretical synthesis.

Layer 3 (Empathy-Led Leadership and Human Dialogue) constitutes the relational and developmental dimension of the IPEM, operationalising empathy-led leadership through three structured mechanisms: (1) continuous micro-feedback conversations, replacing annual appraisal with frequent, low-stakes developmental dialogues grounded in Layer 2 analytical insights; (2) structured employee self-reflection against OKRs and well-being indicators, cultivating metacognitive awareness and developmental ownership; and (3) manager empathy accountability, embedding empathy competencies within managerial evaluation criteria and equipping managers with AI-generated well-being and engagement insights to inform coaching conversations. This layer is empirically grounded in research on psychological safety [11; 12], employee well-being [13], and compassionate leadership [26].

The distinctive architectural feature of the IPEM is the bidirectional integration between its three layers. Data flows upward from Layer 1 to Layer 2 and from Layer 2 to Layer 3; simultaneously, insights generated through empathic dialogue feed back into metric selection and goal-setting in Layer 1, while employee feedback on algorithmic outputs informs fairness auditing and model refinement in Layer 2. This bidirectional feedback architecture operationalises the STS joint optimisation principle: neither the technical nor the social subsystem is treated as primary; each is continuously informed and improved through interaction with the other. The cross-cutting Governance Layer ensures GDPR and EU AI Act compliance, employee rights protections, algorithmic equity monitoring, and ethical oversight throughout all three layers.

The IPEM's configurational logic, the principle that the effectiveness of each layer is contingent upon its alignment with the others, represents its central theoretical contribution. HR Metrics without AI analytics generate data that exceeds human processing capacity; AI analytics without empathy-led leadership produce insights that cannot be translated into developmental action; empathy-led leadership without data infrastructure produces relationship quality that is not grounded in evidence-based performance understanding. Only through their structured integration does the IPEM generate an evaluation system that is simultaneously more objective, more predictive, and more human than any single-component approach.

Table 2 provides a systematic comparison of traditional personnel evaluation approaches and the modern integrated approach embodied in the IPEM across fourteen evaluatively significant criteria.

Table 2 – Comparison of Traditional and Modern Integrated Evaluation Approaches

Criterion	Traditional Approaches	Modern Integrated Approach (IPEM)
Evaluation Frequency	Annual or semi-annual	Continuous, real-time monitoring
Data Sources	Supervisor ratings, self-reports	Multi-source: AI analytics, 360°, HR metrics
Objectivity	Low; prone to rater bias	Augmented by algorithmic fairness mechanisms
Primary Focus	Past performance review	Future potential and development trajectory
Employee Well-being	Rarely considered	Central: psychological safety, empathy, balance
Feedback Type	Evaluative, summative	Developmental, continuous, dialogic
Metrics Used	Simple KPIs, rating scales	OKRs, predictive analytics, engagement indices
System Adaptability	Static, standardised forms	Dynamic, context-sensitive frameworks
Manager Role	Judge and evaluator	Coach, mentor, empathy-led leader
Technology Integration	Minimal or absent	AI, NLP, dashboards, People Analytics
Bias Risk	High (halo, recency, leniency)	Reduced via algorithmic fairness audits
Governance	Absent or informal	Formal: Privacy and Fairness Committee, GDPR/EU AI Act
Developmental Value	Limited; compliance-driven	High; growth-oriented and strategic

Source: Compiled by the author based on review of relevant literature

The comparative analysis demonstrates that modern integrated approaches differ from traditional systems not incrementally but categorically, fundamentally reconceptualising the evaluation function itself. The transformation in evaluation temporality reflects a shift from static state measurement to dynamic trajectory-based assessment, with corresponding improvements in developmental utility and the quality of strategic talent decisions. The inclusion of an explicit governance criterion within the extended framework reflects a key insight from recent research: the legitimacy and effectiveness of AI-augmented evaluation are not solely technical properties, but relational and institutional ones, contingent upon transparent governance structures perceived by employees as procedurally fair [4; 21].

The redefinition of the manager's role—from evaluator to empathy-led coach—deserves particular emphasis. Within traditional evaluation systems, the manager's primary function is to produce an accurate and defensible assessment of past performance. In contrast, within the IPEM, the manager's role shifts toward leveraging insights generated by the measurement and analytics layers to facilitate development-oriented, empathy-driven conversations. These interactions enable employees to interpret their performance trajectories, identify growth opportunities, and experience evaluation as an investment in their long-term potential, consistent with the cultural transformations observed in leading digital organisations.

Conclusions. This article advances a theoretically grounded and practically applicable integrated framework for modern personnel evaluation, addressing a well-documented yet insufficiently resolved gap in contemporary HRM literature. The findings generate conclusions of relevance to both scholarly and practitioner audiences, structured around three interrelated domains of contribution.

The IPEM advances HRM theory by offering a systematically integrated framework that unites HR metrics, AI-driven people analytics, and empathy-led leadership within a coherent socio-technical architecture grounded in evidence-based HRM and human-centred AI principles. The model's contribution lies not in the novelty of its individual components, each of which has been extensively theorised, but in their structured integration and the explicit resolution of three fundamental tensions that have fragmented performance management research: (1) the tension between algorithmic objectivity and relational authenticity; (2) the tension between continuous data capture and employee trust; and (3) the tension between administrative control and developmental purpose. By addressing these tensions through a bidirectional and mutually reinforcing system of measurement, analytics, and human dialogue, the IPEM provides a novel explanatory lens for understanding why isolated investments in metrics, AI technologies, or leadership development frequently fail to deliver anticipated improvements in evaluation quality. In doing so, the model contributes to the broader development of socio-technical theorising in AI-augmented organisational contexts, extending current conceptualisations of performance management toward more integrative and system-level explanations.

The article demonstrates the analytical value of integrative conceptual modelling as a rigorous methodological approach for synthesising fragmented research streams in emerging interdisciplinary domains. By combining integrative review logic with theory-building through conceptual modelling, the study develops a structured and internally consistent framework that generates clear, empirically testable propositions. This approach contributes to advancing methodological practices in HRM research by illustrating how conceptual integration can move beyond descriptive synthesis toward explanatory theory development.

The proposed model opens avenues for future empirical research, particularly longitudinal and multi-level studies examining the independent and joint effects of HR metrics maturity, AI analytics capability, and empathy-led leadership on evaluation quality, employee outcomes, and organisational performance across diverse institutional contexts.

For HR practitioners and organisational leaders, the IPEM offers a structured and actionable implementation pathway grounded in the best available scholarly evidence and informed by analytically interpreted organisational practices. The findings identify managerial capability development as the critical bottleneck in successful implementation, with direct implications for organisational investment priorities. More broadly, the article demonstrates that effective personnel evaluation systems in the digital economy must be simultaneously more data-driven and more human-centred. Systems that fail to integrate these dimensions risk either technical optimisation without legitimacy or relational engagement without analytical rigour. The IPEM addresses this dual imperative by embedding fairness, transparency, and employee well-being into the core architecture of evaluation, thereby redefining evaluation not merely as a mechanism of control, but as a strategic and developmental organisational capability.

REFERENCES

1. 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/10.1177/0008125619867>
2. van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly*, 45(3), 1557–1582. <https://doi.org/10.25300/MISQ/2021/16559>
3. Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency, 469–481. <https://doi.org/10.1145/3351095.3372828>
4. Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, 41, 105567. <https://doi.org/10.1016/j.clsr.2021.105567>
5. 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>
6. 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>
7. Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70(1), 71–79. <https://images.template.net/wp-content/uploads/2016/04/11122659/balanced-scorecard-key-performance-indicators.pdf>
8. Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705–717. <https://doi.org/10.1037/0003-066X.57.9.705>
9. Cappelli, P., & Tavis, A. (2016). The performance management revolution. *Harvard Business Review*, 94(10), 58–67.
10. Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: Why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1–11. <https://doi.org/10.1111/1748-8583.12090>
11. Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly*, 44(2), 350–383. <https://doi.org/10.2307/2666999>
12. Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., & Vracheva, V. (2017). Psychological safety: A meta-analytic review and extension. *Personnel Psychology*, 70(1), 113–165. <https://doi.org/10.1111/peps.12183>

13. De Neve, J.-E., Kaats, M., & Ward, G. (2024). *Workplace wellbeing and firm performance* (Wellbeing Research Centre Working Paper 2304). University of Oxford. <https://doi.org/10.5287/ora-bpkbjayvk>
14. Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85(6), 956–970. <https://doi.org/10.1037/0021-9010.85.6.956>
15. Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Sage Publications.
16. Parent-Rochelleau, 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>
17. Doerr, J. (2018). *Measure What Matters: OKRs: The Simple Idea that Drives 10x Growth*. Portfolio. <https://sunbytes.io/app/uploads/2022/04/Measure-What-Matter-John-Doerr.pdf>
18. Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. *The International Journal of Human Resource Management*, 28(1), 3–26. <https://doi.org/10.1080/09585192.2016.1244699>
19. Pulakos, E. D., Hanson, R. M., Arad, S., & Moye, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavior change. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8(1), 51–76. <https://doi.org/10.1017/iop.2014.2>
20. Latham, G. P., & Locke, E. A. (2007). New developments in and directions for goal-setting research. *European Psychologist*, 12(4), 290–300. <https://doi.org/10.1027/1016-9040.12.4.290>
21. Bankins, S., Formosa, P., Griep, Y. et al. (2022). AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context. *Inf Syst Front* 24, 857–875 <https://doi.org/10.1007/s10796-021-10223-8>
22. Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, 45(2), 159–182. <https://doi.org/10.1002/job.2735>
23. Chowdhury, S., Dey, P., 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 literacy: An integrative framework and research agenda. *International Journal of Human Resource Management*, 36(2), 267–314. DOI: [10.1016/j.hrmr.2022.100899](https://doi.org/10.1016/j.hrmr.2022.100899)
24. Rigamonti E, Gastaldi L, Corso M (2024), Measuring HR analytics maturity: supporting the development of a roadmap for data-driven human resources management. *Management Decision*, Vol. 62 No. 13 pp. 243–282, doi: <https://doi.org/10.1108/MD-11-2023-2087>
25. Madad, S., Maumy-Bertrand, M., Bertrand, F., & Valero, Y. (2024). *NLP approach to ground employee performance evaluation*. In *Proceedings of the International Conference on Connected Innovation and Technology (IC-CITX.O) & The Smart Healthcare International Conference (SHeIC)*, Danang, Vietnam. <https://hal.science/hal-05555901v1/document>
26. Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. P., Bakker, A. B., Bamberger, P., Bapuji, H., Bhawe, D. P., Choi, V. K., Creary, S. J., Demerouti, E., Flynn, F. J., Gelfand, M. J., Greer, L. L., Johns, G., Lemoine, G. J., Probst, T. M., Putnam, L. L., van Vugt, M. (2021). COVID-19 and the workplace: Implications, issues, and insights for future

research and action. *American Psychologist*, 76(1), 63–77.
<https://doi.org/10.1037/amp0000716>

27. Dewe, P., & Cooper, C. L. (2021). *Work and stress: A research overview*. Routledge/Taylor & Francis Group.

28. Rousseau, D.M. and Barends, E.G.R. (2011), Becoming an evidence-based HR practitioner. *Human Resource Management Journal*, 21: 221-

235. <https://doi.org/10.1111/j.1748-8583.2011.00173.x>

Стаття надійшла до редакції 19.04.2026

Стаття прийнята до друку після рецензування 26.04.2026