Online ISSN: 3104-9338
Print ISSN: 3104-932X

Research on the Double-Edged Sword Effects of Artificial Intelligence on Enterprise Human Resource Management

Dahao Li*

Doctoral candidate in Business Administration, Al-Farabi Kazakh National University International Business School Almaty, Kazakhstan

Received: March 26, 2026

Revised: March 27, 2026

Accepted: March 28, 2026

Published online: April 11, 2026

To appear in: *International Journal of Advanced AI Applications,* Vol. 2, No. 5 (May 2026)

* Corresponding Author: Dahao Li (li_dakhao2@live.kaznu.kz)

Abstract. Artificial intelligence has been increasingly embedded into the full-cycle practices of enterprise Human Resource Management (HRM), forming a typical technological empowerment scenario with coexisting opportunities and predicaments. Based on the Technology Acceptance Model, Sociotechnical Systems Theory, and Organizational Justice Theory, this paper constructs an integrated theoretical framework to systematically dissect the double-edged sword effects of AI application in HRM. The positive dimensions are refined into four core paths: efficiency improvement in recruitment and selection, objectivity enhancement in performance management, personalized empowerment of employee development, and data-driven upgrading of strategic decision-making. The negative dimensions are summarized as four prominent risks: algorithmic bias and discriminatory infringement, employee privacy leakage and supervision anxiety, dehumanization of employment relations, and multi-level organizational resistance. Furthermore, this study explores the differential impact mechanisms of AI outcomes from four dimensions: AI system design, implementation procedures, organizational contextual factors, and individual employee differences. On this basis, targeted coping strategies are proposed from ethical design, transparent governance, employee participation, and human-AI hybrid collaboration, aiming to provide theoretical references and practical paths for enterprises to maximize the positive effects of AI and control potential risks in HRM practices.

Keywords: *Artificial Intelligence; Human Resource Management; Double-edged Sword Effects; Algorithmic Management; Ethical AI; Human-AI Collaboration*

1. Introduction

Driven by digital transformation and technological iteration, artificial intelligence has broken through the technical boundary of enterprise management and penetrated into the whole process of human resource management, covering recruitment screening, performance evaluation, training and development, turnover prediction, labor relations coordination and other core modules [5] [10]. Global technology giants such as IBM, Oracle, and SAP have launched integrated AI-HR platforms, which have realized the intelligent transformation from traditional manual operation to data-driven decision-making, and gradually changed the management logic and operation mode of organizational human resources. According to industry statistics, more than 68% of large multinational enterprises have adopted AI-based HR tools, and the market size of global AI HR applications is expected to exceed 35 billion US dollars by 2030, showing a rapid growth trend.

However, the integration of AI and HRM presents a complex and contradictory picture in practical application. Optimists believe that AI can liberate HR professionals from heavy administrative work, improve management efficiency and decision-making accuracy, and realize the strategic transformation of human resource management [2]. Pessimists point out that the widespread use of algorithmic decision-making has triggered a series of ethical and management crises, such as algorithmic discrimination, excessive supervision, humanistic indifference, and employee resistance, which seriously impact the fairness and sustainability of employment relations [9]. This dual effect determines that AI in the field of human resource management is not a simple technological tool, but a double-edged sword that can both empower management and induce risks.

The double-edged sword effect refers to the phenomenon that a single factor can produce both beneficial and harmful results, and its final performance depends on the combination of contextual factors and operation methods. In the context of digital HRM, clarifying the positive value, negative risks, influencing mechanisms and governance strategies of AI application is not only a theoretical frontier issue, but also a practical problem that enterprises must solve in digital transformation. At present, academic research has carried out exploratory discussions on AI applications in single HR modules such as recruitment [11] [13], performance management [8], and talent retention, and has also paid attention to the ethical risks and organizational responses of algorithmic decision-making [2] [5]. However, the existing literature lacks a systematic theoretical framework that integrates positive effects, negative effects, influencing mechanisms and governance countermeasures, and cannot fully explain the formation logic and

regulation path of the double-edged sword effect of AI in HRM.

To fill this research gap, this study takes the double-edged sword effect of AI in enterprise HRM as the core research theme, and constructs a comprehensive analysis framework based on three classical theories. The research objectives of this paper are as follows: first, to clarify the specific manifestations and formation paths of the positive effects of AI on HRM; second, to identify the types and potential hazards of negative effects; third, to reveal the key factors and action mechanisms that determine the differential outcomes of AI application; fourth, to put forward targeted governance strategies to amplify advantages and avoid disadvantages. The theoretical contributions of this study are reflected in three aspects: integrating multi-theoretical perspectives to enrich the research system of digital HRM; systematically deconstructing the double-edged sword effect to clarify the dual logic of AI empowerment and risk; proposing a multi-dimensional governance framework to provide theoretical guidance for practical application. In practice, this study can help enterprises rationally view the role of AI, optimize the design and implementation of AI-HR systems, balance technological efficiency and humanistic care, and promote the sustainable development of human resource management in the digital era.

2. Theoretical Foundations

2.1. Technology Acceptance Model

The Technology Acceptance Model (TAM) was proposed by Davis (1989) to explain and predict users' behavioral intention to adopt new information technologies. The core logic of the model is that users' acceptance behavior is determined by two key perceptions: perceived usefulness and perceived ease of use. Perceived usefulness refers to the degree to which individuals believe that using a certain technology can improve their work performance; perceived ease of use refers to the degree to which individuals believe that using a certain technology does not require excessive effort. On this basis, Venkatesh et al. (2003) expanded the model into a unified theory of acceptance and use of technology, incorporating social influence and facilitating conditions into the analysis framework, which enhanced the explanatory power of the model in organizational scenarios.

In the scenario of AI-enabled HRM, the Technology Acceptance Model provides a micro-level analytical perspective for understanding employees' and managers' responses to AI systems. When HR practitioners and employees perceive that AI systems can effectively improve work efficiency, reduce operational burden, and enhance decision-making quality

(high perceived usefulness), and the system operation is simple and easy to understand (high perceived ease of use), they will take a positive attitude towards AI application and actively participate in system use [14]. Conversely, if users believe that AI systems threaten their work autonomy, replace their core responsibilities, or have complex operation and opaque logic (low perceived usefulness and ease of use), they will produce resistance and rejection, which will lead to the failure of AI implementation. This theoretical perspective emphasizes that the effectiveness of AI application does not solely depend on the technical level of the system, but is closely related to users' subjective perception and psychological acceptance, which lays a foundation for analyzing the micro-foundation of the double-edged sword effect.

2.2. Sociotechnical Systems Theory

Sociotechnical Systems Theory was founded by Trist and Bamforth (1951) in the study of coal mining production systems. The core view of the theory is that any organization is a complex system composed of interactive social subsystems and technical subsystems; the optimal operation of the organization requires the coordinated adaptation of the two subsystems, and the introduction of technical systems must fully consider the matching of social factors such as organizational structure, employee relations, and work design. Bostrom and Heinen (1977) further applied this theory to the field of information system management, pointing out that the failure of most management information systems is not due to technical defects, but due to the neglect of social subsystem factors, resulting in the disconnection between technology and people.

In the context of AI integration into HRM, Sociotechnical Systems Theory provides a holistic analytical framework for balancing technological efficiency and humanistic care. AI, as a technical subsystem, has the advantages of high efficiency, objectivity, and data processing, but it must be matched with the social subsystem including organizational culture, employee attitudes, power relations, and work design [1]. When AI systems are designed to assist human work rather than replace human judgment, and fully consider the social attributes and emotional needs of employees, the technical and social subsystems can form a synergistic effect, and the positive effects of AI will be highlighted. If enterprises only pursue technological efficiency and ignore the social impact of AI implementation, such as damaging interpersonal trust, destroying employment relations, and weakening organizational cohesion, it will lead to system imbalance and trigger a series of negative effects. This theory reveals the systemic logic of AI's double-edged sword effect and provides a theoretical basis for constructing a human-AI hybrid management model.

2.3. Organizational Justice Theory

Organizational Justice Theory focuses on employees' perception of fairness in organizational decision-making and its impact on individual attitudes and behaviors. Greenberg (1987) divided organizational justice into distributive justice and procedural justice; distributive justice refers to the fairness of the distribution of organizational resources and outcomes, and procedural justice refers to the fairness of the decision-making process. Colquitt (2001) further integrated the theoretical framework and added interactional justice, which refers to the fairness of interpersonal treatment received by individuals in the process of organizational management. The three dimensions of justice jointly affect employees' trust, satisfaction and organizational commitment.

In the application of AI in HRM, Organizational Justice Theory is a key theoretical tool for explaining employees' psychological perception and behavioral response to algorithmic decision-making. AI systems participate in key HR decisions such as recruitment, promotion, performance appraisal and termination, and employees' perception of the fairness of these algorithmic decisions directly determines their acceptance of AI [2]. If employees believe that AI decisions are based on fair data and transparent processes (procedural justice), the results are reasonable and unbiased (distributive justice), and the system fully respects individual dignity (interactional justice), they will recognize the application of AI and form positive organizational behavior. On the contrary, if the algorithm is a "black box", the decision results are discriminatory, and the management process lacks humanistic care, employees will perceive injustice, resulting in resistance, distrust and even departure, which amplifies the negative effects of AI. This theory clarifies the psychological mechanism of employees' response to AI and provides a theoretical basis for the ethical design and transparent governance of AI-HR systems.

2.4. Theoretical Integration Framework

Based on the above three theories, this study constructs an integrated theoretical framework for the double-edged sword effect of AI on HRM. The Technology Acceptance Model explains the micro-psychological mechanism of users' acceptance of AI, which is the micro-foundation of AI effect realization; Sociotechnical Systems Theory emphasizes the synergistic matching of technical and social subsystems, which is the systemic guarantee of AI effect; Organizational Justice Theory reveals the fairness perception mechanism of employees to algorithmic decisions, which is the psychological premise of AI effect. The three theories complement each other: TAM focuses on individual perception, Sociotechnical Systems Theory focuses on

organizational system matching, and Organizational Justice Theory focuses on fairness perception. Together, they constitute a multi-level, multi-dimensional theoretical system to explain why AI produces both positive and negative effects in HRM, and provide a logical basis for analyzing influencing factors and proposing governance strategies.

3. Positive Effects of AI on HRM

3.1. Enhanced Efficiency in Recruitment and Selection

AI has realized the intelligent transformation of recruitment and selection, and greatly improved the operational efficiency of the recruitment link. Traditional recruitment relies on manual resume screening, which has problems such as long time-consuming, high labor cost and large subjective error. AI-powered automated resume screening tools can use natural language processing and machine learning technologies to quickly identify keywords, match job requirements, and screen qualified candidates from massive resumes. Relevant data shows that AI resume screening can shorten the average recruitment cycle by 55%-60%, and process thousands of resumes in a few minutes, which is dozens of times more efficient than manual screening.

Intelligent interview assistants and chatbots further optimize the recruitment process. AI chatbots can conduct preliminary communication with candidates 24 hours a day, answer frequently asked questions about positions, enterprises and welfare, schedule interviews automatically, and feed back information to HR in real time [11]. Video interview AI systems can analyze candidates' expressions, language logic and behavioral characteristics through computer vision and speech recognition technology, providing auxiliary evaluation references for recruiters. These applications reduce the repetitive work of HR professionals, enabling them to focus on high-value activities such as candidate relationship maintenance and cultural fit assessment.

In terms of cost control, AI recruitment tools have significant advantages. Compared with traditional recruitment methods, AI-enabled recruitment can reduce recruitment costs by 20%-30%, mainly reflected in the reduction of intermediary agency fees, administrative labor costs and time costs [13]. For enterprises, shortened recruitment cycles mean that vacant positions are filled faster, reducing the production and operational losses caused by talent gaps; for candidates, AI recruitment provides faster response speed and more standardized communication, improving the candidate experience. The efficiency improvement of recruitment and selection is the most direct positive effect of AI on HRM, laying a foundation

for enterprises to quickly obtain high-quality talents.

3.2. Improved Objectivity in Performance Management

AI applications effectively alleviate the subjective bias in traditional performance management and improve the objectivity and fairness of evaluation. Traditional performance appraisal is highly dependent on managers' subjective judgment, which is prone to recency bias, halo effect, central tendency and interpersonal favoritism, leading to the deviation of evaluation results from actual performance [8]. AI performance management systems collect objective data from multiple sources, including project completion rate, work quality indicators, customer satisfaction, team collaboration efficiency and other behavioral and result data, and conduct comprehensive evaluation through algorithm analysis, reducing the interference of human subjective factors.

The real-time feedback function of AI optimizes the dynamic management of performance. Traditional performance appraisal mostly adopts annual or semi-annual periodic evaluation, with delayed feedback and poor timeliness. AI systems can monitor employees' work data in real time, provide instant feedback on work defects and improvement directions, help employees adjust work strategies in a timely manner and achieve continuous performance improvement [5]. For example, sales employees can obtain real-time data on customer conversion and performance completion through AI systems, and adjust sales strategies according to feedback; R&D personnel can track project progress and task completion through AI tools to ensure the smooth progress of projects.

In addition, AI performance analytics supports fair and transparent evaluation results. The system generates quantitative evaluation reports based on objective data, and managers can refer to algorithmic suggestions to make evaluation decisions, reducing the impact of personal preferences. Some enterprises have applied AI performance systems to achieve multi-dimensional and full-cycle evaluation, making performance appraisal more standardized and credible, enhancing employees' recognition of performance management, and promoting the formation of a fair competition organizational atmosphere.

3.3. Personalized Employee Development

AI empowers employee training and career development to achieve personalized and precise management at scale. Traditional training modes adopt a one-size-fits-all approach, which cannot match the individual differences in employees' skill gaps, learning abilities and career plans, resulting in low training efficiency and poor effectiveness. AI-powered learning

platforms can analyze employees' skill attributes, work performance, learning habits and career aspirations through big data technology, and recommend personalized training courses, learning paths and development plans [8]. For example, for new employees, the system can provide targeted induction training; for senior employees, it can recommend advanced management or technical courses to meet their career promotion needs.

AI career pathing tools provide personalized development guidance for employees. Based on the matching analysis of employees' personal abilities and organizational talent needs, the system can predict employees' career development potential, recommend suitable positions and promotion paths, and formulate skill improvement plans [10]. This personalized development service makes employees feel the organization's attention and investment in their growth, enhances employees' sense of belonging and work engagement, and reduces the turnover risk of high-potential talents. For enterprises, personalized employee development helps to accurately fill the skill gap, optimize the talent structure, and realize the coordinated development of individual growth and organizational goals.

AI skill gap analysis provides a basis for proactive talent development. The system can compare the current skill inventory of the workforce with the skill requirements required by the enterprise's future development strategy, identify the key skill gaps in the organization, and help the human resource department carry out targeted training and talent introduction in advance [5]. This predictive management mode changes the passive response of traditional HRM, realizes proactive talent layout, and provides talent guarantee for enterprises' long-term development.

3.4. Data-Driven Strategic Decision-Making

AI promotes the transformation of human resource management from administrative execution to strategic decision-making, and realizes data-driven scientific management. Traditional HRM relies on experience and qualitative judgment, lacking quantitative support for strategic decisions such as workforce planning, talent retention and organizational structure adjustment. AI-enabled people analytics platforms can mine and analyze massive workforce data, identify potential laws and trends, and provide quantitative decision-making support for strategic human resource management [8].

Predictive analytics is a core application of AI in strategic HR decision-making. For example, AI turnover prediction models can analyze the key factors affecting employees' departure intention by integrating demographic characteristics, work performance, salary level,

organizational atmosphere and other data, issue early warning signals for high-risk turnover employees, and help managers take intervention measures in advance [10]. Workforce planning AI tools can simulate the impact of different recruitment, training and mobility strategies on the future workforce scale, structure and cost, and optimize the workforce allocation plan to match the enterprise's business strategy.

AI improves the strategic value and organizational status of the human resource department. By providing accurate data insights and decision support, HR can participate in the enterprise's top-level strategy formulation, and transform from a functional support department into a strategic partner. For enterprise executives, AI-HR analytics provides a visual presentation of human capital value and risk, helping them accurately grasp the status of organizational talents and make scientific business decisions. The strategic upgrading of HRM driven by AI is an important positive effect, which enhances the core competitiveness of enterprises in the digital era.

4. Negative Effects of AI on HRM

4.1. Algorithmic Bias and Discrimination Risks

Algorithmic bias is the most prominent negative effect of AI in HRM, which may lead to discriminatory violations and damage the fairness of employment. Algorithmic bias in HRM scenarios essentially stems from two distinct and independent sources: data bias and model design bias, which can separately or jointly trigger discriminatory decision-making and form systemic unfairness in organizational HR practices [2].

Data bias derives from inherent flaws in the training data for AI models, the most common source of algorithmic bias in HRM practice. AI systems rely on historical organizational and industry data for iterative learning and decision-making; if the historical data contains implicit or explicit biases related to gender, age, race, region, or the dataset is unrepresentative, unbalanced, or incomplete for specific groups, the algorithm will amplify and solidify these pre-existing unfairness in the training process. For example, in the recruitment of technical positions, if the historical training data is dominated by male employees due to the traditional gender structure of the industry, the AI model may incorrectly associate gender with job competence, resulting in discriminatory screening against female candidates [13]. This type of bias is a passive reflection of the unfairness in the data source, with the AI model only reproducing and strengthening the original bias without independent judgment.

Model design bias originates from artificial technical defects in the development and design

stage of AI models, caused by inappropriate human intervention in the model construction process. Flawed algorithmic logic, unreasonable selection and weight setting of evaluation features, structural defects in model architecture, or the lack of effective debiasing algorithms in the design stage can all lead to model design bias. For instance, in the design of performance evaluation AI models, if developers set excessive weight on quantitative work output indicators and ignore qualitative indicators such as team collaboration and innovation contribution, the model will form a systemic bias against employees engaged in creative and collaborative work, leading to unfair and one-sided evaluation results.

Scholars have verified the existence of algorithmic discrimination in practical cases caused by the two types of biases above. Some AI recruitment systems have been found to downgrade candidates with female-typical names or experience in women's organizations (a typical result of data bias); some performance evaluation algorithms have shown racial bias by penalizing employees whose language and behavior characteristics are associated with specific ethnic groups (often a combination of data imbalance and unreasonable feature selection in model design) [2]. This algorithmic discrimination is concealed and procedural, making it difficult for individuals to defend their rights, and it is easier to form systemic unfairness in the organization.

Algorithmic bias brings serious legal and reputational risks to enterprises. In recent years, lawsuits and regulatory penalties caused by AI discriminatory decisions have increased in Europe and the United States; regulatory authorities have strengthened the review of algorithmic fairness in employment scenarios. For enterprises, algorithmic discrimination not only faces economic losses such as fines and compensation, but also damages the employer brand and social image, leading to the loss of talents and customers.

4.2. Employee Privacy Concerns and Surveillance Anxiety

The wide application of AI in HRM relies on the collection and analysis of a large amount of employee personal data, which triggers prominent privacy and security risks. AI monitoring systems can collect employees' work behavior data in an all-round way, including email content, chat records, attendance status, keystroke frequency, office track and even physiological characteristics [9]. Although these data are used for performance analysis and management decision-making, excessive collection and unauthorized use seriously infringe on employees' personal privacy.

Transparent lack of data usage exacerbates employees' supervision anxiety. Many enterprises do not fully inform employees of the scope, purpose and usage rules of data collection when

applying AI monitoring systems, forming a "transparent employee" management model. Employees feel that their work behaviors are monitored and analyzed all the time, resulting in psychological pressure, anxiety and resistance [5]. Some employees even take perfunctory work, performative behaviors and other ways to "cope" with algorithmic supervision, which reduces work efficiency and creativity.

Data security risks further aggravate privacy threats. The centralized storage of massive sensitive employee data increases the risk of data leakage, theft and abuse. Once a data breach occurs, it will cause serious harm to employees' personal information and bring regulatory penalties and trust crises to enterprises [2]. Privacy and supervision issues have become important obstacles affecting employees' acceptance of AI and the sustainable operation of AI-HR systems.

4.3. Dehumanization of Employment Relationships

AI application leads to the weakening of human interaction and the dehumanization of employment relations, damaging the emotional connection between individuals and organizations. Traditional HRM relies on face-to-face communication between HR professionals, managers and employees, which conveys organizational care and humanistic warmth. When recruitment consultation, performance feedback, training guidance and other links are replaced by AI systems, employees mainly interact with machines and algorithms, and feel treated as data points rather than independent individuals with emotions and needs [9].

The dehumanization effect is more obvious in key employment decisions. When employees encounter dismissal, demotion, rejection of promotion and other results generated by algorithms, they often feel that the decision lacks human judgment, empathy and respect, even if the result is reasonable [2]. This sense of dehumanization reduces employees' trust in the organization, weakens organizational commitment and work engagement, and may even lead to negative behaviors such as absenteeism and resignation.

In addition, the increasing automation of HR processes significantly reduces the opportunities for meaningful interpersonal communication within organizations. As HR professionals become increasingly occupied with managing and operating AI systems, they often find themselves lacking direct communication with employees, which can hinder relationship-building. Moreover, managers tend to rely heavily on algorithmic reports generated by these systems and, as a result, may neglect crucial face-to-face interactions with their subordinates. This weakening of interpersonal interaction can have detrimental effects on the organizational culture and team cohesion, ultimately leading to the alienation of

employment relations. Such alienation not only impacts employee morale but also poses risks to the long-term sustainability and growth of the organization itself.

4.4. Multi-Level Organizational Resistance to AI Adoption

The application of AI in HRM encounters multi-level resistance from organizational stakeholders, affecting the implementation effect and sustainable operation of the system. First, HR professionals have resistance. They worry that AI will replace their core work tasks, leading to job redundancy and loss of professional autonomy, so they take negative attitudes such as perfunctory implementation and passive rejection [5]. Many HR professionals lack the digital skills to operate AI systems, and the fear of technical unemployment further intensifies resistance.

Second, line managers have resistance. AI systems participate in performance appraisal, talent evaluation and other decisions that originally belong to managerial authority, which makes managers feel that their decision-making power is weakened and their status is threatened [10]. Some managers believe that algorithms cannot grasp the complex context of work, and blindly relying on AI will lead to decision-making deviation, so they choose to ignore or bypass AI suggestions, resulting in low utilization of AI systems.

Third, ordinary employees have resistance. Employees resist AI systems that are perceived as unfair, intrusive and dehumanized; some employees take hidden resistance measures such as reducing work input, avoiding data submission and using alternative communication channels [2]. The multi-level organizational resistance increases the implementation cost of AI, reduces the operational efficiency of the system, and even leads to the complete failure of AI-HR projects.

5. Mechanisms Shaping AI Outcomes

5.1. System Design Factors

The design characteristics of AI systems directly determine whether the double-edged sword effect tends to be positive or negative, and are the core technical factors affecting AI outcomes. Transparency and explainability are the primary design principles. "Black box" algorithms with opaque decision-making logic make users unable to understand the basis of AI recommendations, prone to doubt and resistance; AI systems with explainable functions can clearly output decision-making factors, calculation processes and result basis, helping users understand and trust the system [2]. Enterprises should integrate explainable AI technology in

the design stage to improve the transparency of the system.

Fairness-embedded design is the key to avoiding algorithmic bias. AI systems should use diversified and representative training data to avoid data imbalance leading to bias; pre-deployment bias testing and post-operation real-time monitoring should be carried out to correct discriminatory tendencies in a timely manner [10]. The fairness design should cover all protected groups such as gender, race, age and disability to ensure that the algorithm does not produce differential infringement.

Human-in-the-loop design is an important guarantee for positive outcomes. The system should position AI as an auxiliary decision-making tool rather than an independent decision-maker, retaining human review and final decision-making power [5]. This design allows managers to consider complex contextual factors that cannot be captured by algorithms, ensuring that decisions are both efficient and humanistic, and balancing technical rationality and human judgment.

5.2. Implementation Processes

The implementation process of AI systems is equally as important as their technical design; therefore, a standardized and human-centered approach to implementation can significantly mitigate potential negative effects. Effective communication with all stakeholders is a vital prerequisite for ensuring a smooth implementation process. Prior to launching AI systems, enterprises should clearly articulate the purpose, functionalities, usage guidelines, and rights protection measures associated with these systems to HR professionals, managers, and employees. Addressing common concerns related to job security, privacy protection, and fairness is crucial in order to alleviate fears and uncertainties, thereby reducing resistance to change and fostering a more positive reception of the technology among all parties involved [2].

Phased and pilot implementation is a practical strategy. Instead of large-scale full deployment at one time, enterprises should carry out small-scale pilot tests in a single department or a single module, collect user feedback, optimize system design and operation processes, and then gradually promote them to the whole organization [5]. The phased implementation helps to accumulate experience, resolve risks and improve user acceptance.

Comprehensive user training and support are essential. Enterprises should provide targeted training for different groups: for HR professionals, focus on AI system operation, data analysis and ethical governance skills; for managers, focus on how to combine AI suggestions with

practical management; for ordinary employees, focus on system use methods and rights protection channels [10]. Adequate training can improve users' technical literacy and operational confidence, promoting the effective use of AI systems.

5.3. Organizational Context

Organizational contextual factors regulate the relationship between AI application and HRM effects, and determine the adaptation degree of AI in the organization. Organizational culture is a key contextual factor. Enterprises with an innovative, open and trustworthy culture are more tolerant of new technologies, and employees are more willing to accept and try AI systems; enterprises with a conservative, rigid and distrustful culture are more prone to resistance and conflict, amplifying the negative effects of AI [5].

Leadership commitment and support determine the implementation intensity. Senior leaders' positive attitude towards AI, investment of resources and demonstration of use can form a guiding effect within the organization and promote the smooth promotion of AI systems [10]. If leaders lack attention, insufficient investment or inconsistent signals, the implementation of AI will lack institutional guarantee and resource support, leading to poor results.

Resource matching is a basic guarantee. The successful application of AI requires matching technical infrastructure, data management capabilities, professional teams and organizational systems [8]. Enterprises with sufficient digital resources can optimize system operation and user experience; enterprises with insufficient resource investment will face problems such as system lag, data errors and insufficient support, leading to negative outcomes.

5.4. Individual Differences

Individual differences of employees lead to heterogeneous perceptions and responses to AI systems, affecting the final effect of AI application. Technical self-efficacy is a key individual factor. Employees with high technical self-efficacy have strong confidence in using new technologies, can quickly adapt to AI systems, and perceive more positive effects; employees with low technical self-efficacy are afraid of operating difficulties and technical substitution, and are prone to negative perceptions and resistance [14].

Trust in organization and management regulates individual responses. Employees who trust the enterprise and managers believe that the organization will use AI fairly and protect their rights and interests, and have higher acceptance of AI systems; employees with low trust are skeptical of AI application, worried about being treated unfairly and invaded privacy, and amplify negative perceptions [2].

Demographic and occupational characteristics also have an impact. Studies show that young employees, highly educated employees and employees in technical positions have higher acceptance of AI; older employees, less educated employees and employees in traditional positions are more cautious about AI [5]. Enterprises should formulate differentiated communication and training strategies based on individual differences to improve the overall acceptance of AI systems.

6. Countermeasures for Enterprises

6.1. Ethical AI System Design: Foundation of Risk Prevention

Enterprises should take ethical design as the core of AI-HR system construction to prevent risks from the source. First, establish a clear AI ethical governance framework, formulate ethical principles covering fairness, transparency, accountability, privacy protection and human-centeredness, and run through the whole life cycle of AI system development, deployment and operation [2]. The ethical principles should be combined with labor laws and regulations to ensure compliance.

Second, implement full-cycle algorithmic bias governance. Conduct bias audit on training data before system deployment to eliminate discriminatory information; test the algorithm for disparate impact on different groups to ensure fair decision-making; establish a real-time monitoring mechanism after deployment, track the decision results of the system, and correct bias in a timely manner [10]. Record the whole process of bias testing and governance to form a compliant evidence chain.

Third, strengthen the explainability and human oversight of the system. Adopt explainable AI technology to enable users to clearly understand the decision basis of the system; set up a multi-level human review mechanism, and AI can only provide recommendations for key decisions such as recruitment, promotion and dismissal, and the final decision shall be made by managers to ensure that humanistic care and contextual judgment are not missing [5].

6.2. Transparent Implementation Practices: Path to Trust Building

Transparent and standardized implementation practices can enhance stakeholder trust and reduce organizational resistance. First, carry out full-process transparent communication. Disclose the objectives, functions, data usage scope, supervision mechanisms and appeal channels of AI systems to all employees through meetings, manuals, online platforms and other forms; regularly publish the operation status and effect of AI systems to ensure employees' right

to know [2].

Second, promote employee participation in governance. Invite employee representatives to participate in AI system pilot testing, demand design and effect evaluation; collect employees' opinions and suggestions on AI application, and optimize the system and management methods based on feedback [5]. Participatory governance makes employees feel respected and enhances their sense of ownership.

Third, formulate complete AI management policies. Clarify the norms of data collection, storage, use and destruction, the scope of application of algorithmic decisions, the rights and obligations of employees, and the appeal process for unfair decisions; publicize the policies to all employees and strictly implement them to provide institutional guarantee for standardized AI application [10].

6.3. Employee Involvement Mechanisms: Guarantee of Rights Protection

Improving employee participation and rights protection mechanisms can alleviate employees' concerns and enhance their positive perception of AI. First, establish a smooth feedback and appeal channel. Set up a special mailbox, hotline or online platform for employees to report problems such as algorithmic bias, privacy infringement and unfair decisions; respond to and handle feedback in a timely manner to ensure that employees' voices are heard [5].

Second, improve the human review appeal mechanism. Employees who disagree with AI-based decisions can apply for human review; the organization should form an independent review team to re-evaluate the decision, fully consider the employee's situation and contextual factors, and give a reasonable explanation and result [10]. The appeal mechanism ensures that employees have a way to protect their rights and enhances the fairness of management.

Third, carry out employee AI literacy training. Popularize AI basic knowledge, ethical norms and rights protection knowledge to employees, improve employees' ability to understand and use AI systems, and enhance their awareness of self-protection and participation in governance [2]. Literacy training helps eliminate employees' fear of AI and promote positive interaction between people and AI.

6.4. Hybrid Human-AI Work Models: Path to Synergistic Empowerment

Constructing a hybrid human-AI work model is the core strategy to give full play to the advantages of both sides and achieve synergistic empowerment. Figure 1 visually presents the hierarchical division of labor, interactive mechanism and application scenarios of the human-AI hybrid collaboration model in the full cycle of enterprise HRM, which intuitively reflects

the complementary operation logic of AI and humans in different HR core modules.

First, clarify the division of labor between humans and AI. AI is responsible for repetitive, standardized, data-intensive tasks such as resume screening, data statistics and real-time monitoring; humans are responsible for creative, emotional, complex judgment tasks such as cultural fit evaluation, interpersonal communication, conflict resolution and key decision-making [5].

Second, clarify the division of labor between humans and AI. AI is responsible for repetitive, standardized, data-intensive tasks such as resume screening, data statistics and real-time monitoring; humans are responsible for creative, emotional, complex judgment tasks such as cultural fit evaluation, interpersonal communication, conflict resolution and key decision-making [5].

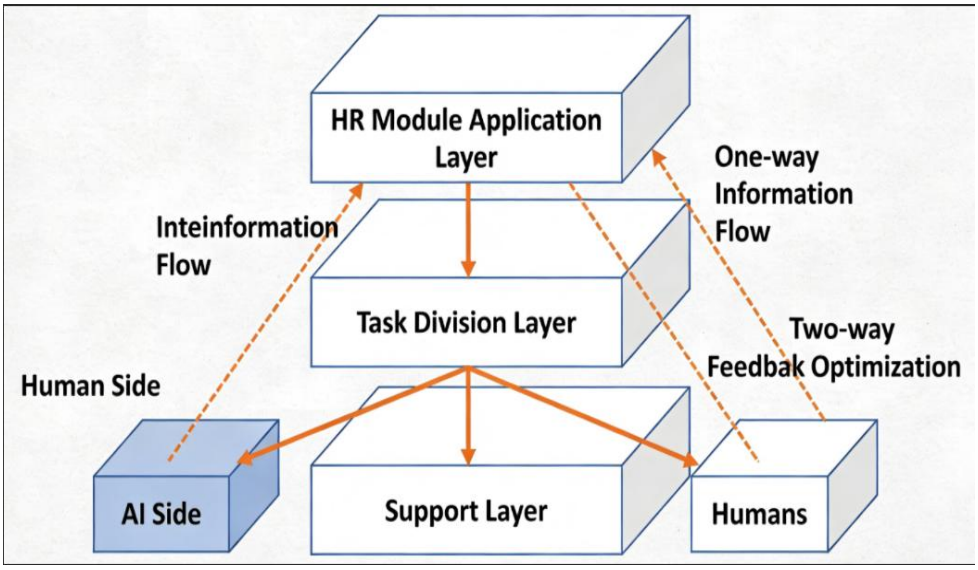

Figure 1.

Third, optimize the organizational process and team structure matching the hybrid model. Adjust the workflow of HRM to adapt to the collaborative operation of humans and AI; train cross-disciplinary teams with HR expertise and AI technology to promote the deep integration of technology and management [10]. The hybrid model realizes the complementary advantages of efficiency and humanization, and maximizes the positive effects of AI while controlling risks.

7. Conclusion and Future Directions

7.1. Summary of Theoretical Contributions

This study systematically explores the double-edged sword effect of artificial intelligence on enterprise human resource management based on the Technology Acceptance Model,

Sociotechnical Systems Theory and Organizational Justice Theory, and forms a complete theoretical and practical framework. Theoretically, this study first integrates multi-dimensional theories to construct a comprehensive analytical framework for AI-HRM effects, clarifying the micro-psychological, organizational-systemic and fairness-perceptual mechanisms of AI's dual effects, and enriching the theoretical system of digital human resource management. Second, this study systematically deconstructs the positive effects of AI on HRM into four dimensions: efficiency, objectivity, personalization and strategicization, and the negative effects into four risks: algorithmic bias, privacy anxiety, dehumanization and organizational resistance, clarifying the dual logic and specific manifestations of AI empowerment and risk. Third, this study reveals the differential impact mechanisms from system design, implementation process, organizational context and individual differences, and proposes a four-in-one governance strategy of ethical design, transparent implementation, employee participation and hybrid collaboration, providing a theoretical basis for enterprises to regulate AI application.

7.2. Implications for Practice

The conclusions of this study have important practical guiding significance for enterprise managers, HR practitioners and technology providers. For enterprise executives, AI application in HRM cannot blindly pursue technological efficiency, but should balance technological advantages and ethical risks, attach importance to system design, organizational matching and employee acceptance, and take risk prevention and control as a necessary part of digital transformation. For HR leaders, they should take the initiative to master AI ethical governance and hybrid management capabilities, transform from traditional administrative executors to digital strategic managers, promote the coordinated development of AI technology and humanistic management, and enhance the strategic value of human resources. For AI technology providers, they should focus on the fairness, transparency and explainability of products in the development of HR-oriented AI systems, embed human oversight and ethical governance functions, and develop human-centered AI tools to meet the actual needs of enterprises.

7.3. Limitations and Future Research Directions

This study has certain limitations. As a theoretical research, this paper constructs a framework based on existing literature and practical cases, but lacks empirical testing such as questionnaire survey and empirical analysis to verify the causal relationship and impact effect of various variables. In addition, this study does not distinguish the differential effects of AI in

different industries, enterprise scales and institutional environments, and the universality of the conclusions needs to be further verified.

Future research can be carried out from four aspects. First, conduct empirical research to collect data through questionnaires, interviews and case studies, test the theoretical framework of this paper, verify the impact of various factors on the double-edged sword effect of AI, and clarify the boundary conditions of AI positive and negative effects. Second, explore the long-term dynamic impact of AI on employment relations, track the changes in employees' trust, engagement and organizational behavior after long-term use of AI systems, and reveal the evolution law of AI effects over time. Third, compare the effectiveness of different governance strategies, analyze the governance effects of ethical design, transparent implementation and other paths, and find the optimal combination strategy suitable for Chinese enterprises. Fourth, carry out cross-cultural and cross-institutional comparative research, explore the differences in AI-HRM effects and governance paths under different regulatory systems and cultural backgrounds, and provide a reference for global digital HRM practices.

References

[1] Bostrom, R. P., & Heinen, J. S. (1977). MIS problems and failures: A socio-technical perspective. *MIS Quarterly*, 1(3), 17-32.

[2] Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., & Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. *Human Resource Management Journal*, 33(3), 606-659.

[3] Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386-400.

[4] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

[5] Fernández-Vidal, J., Peral-Peral, B., & Gascó, J. L. (2025). Technology-driven change in human resource management: Reshaping talent management and organizational design. *Administrative Sciences*, 15(11), 452.

[6] Malyshev, V., Lipskyi, Y., Kovalenko, V., Gab, A., Shakhnin, D., & Orel, O. (2024). Assessment of the global artificial intelligence market in healthcare. *Technology audit and production reserves*, 6(4/80), 62-70.

[7] Greenberg, J. (1987). A taxonomy of organizational justice theories. *Academy of Management Review*, 12(1), 9-22.

[8] Ioniță, A., & Ștefan, S. C. (2025). Strategic human resource management in the digital era: Technology, transformation, and sustainable advantage. *Merits*, 5(4), 23.

[9] 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.

[10] Kim, S., Khoreva, V., & Vaiman, V. (2025). Strategic human resource management in the era of algorithmic technologies: Key insights and future research agenda. *Human Resource Management*, 64(2), 447-464.

[11] Marliyas, A., Ummah, M. A. C. S., & Gunapalan, S. (2026). The transformation of talent

acquisition through artificial intelligence in the context of Industry 4.0: A systematic literature review. *Journal of Business Research and Innovation*, 11(2).

[12] 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.

[13] Vaishnavi, D., Hada, V., Sharma, R., Singh, G., Singh, R. K., & Bhingardive, S. (2026). Integrating AI into human resource management: Implications for recruitment and retention. *Journal of Asia Entrepreneurship and Sustainability*, 22(1).

[14] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.