

Ethical Risks and Mitigation Strategies Arising from AI-Enabled Human Resource Management in the Public Sector

Aoran Wang^a, Dan Jiang^{*}

School of Humanities and Social Sciences, Northeast Petroleum University,
Daqing, 163000, China

^{*} Corresponding Author Email: jiangdan1009@126.com, ^a18845382178@163.com

Abstract. Artificial intelligence technologies are reshaping human resource management (HRM) processes in the public sector through intelligent algorithms and big data analytics. Their applications in three key scenarios—talent recruitment, performance evaluation, and training—play increasingly critical roles. This study identifies major ethical risks: challenges to decision-making fairness caused by algorithmic black boxes, ambiguity in accountability attribution within human-machine collaboration, and power imbalances resulting from data monopolization. To address these, we propose a "Tripartite Governance Framework for Technology": establishing algorithmic auditing mechanisms, refining ethical standards for human-AI collaboration, and fostering a digital governance culture. This framework provides actionable pathways for the digital transformation of public sector HRM in the AI era.

Keywords: Artificial Intelligence; Public Sector; Human Resource Management; Algorithmic Governance; Digital Ethics.

1. Application Scenarios of AI-Enabled Practices in the Public Sector

1.1. Era Context and Trends

With breakthroughs in general AI technologies like DeepSeek, China's public sector is undergoing a wave of digital transformation. The 14th Five-Year Plan for Digital Government Construction explicitly proposes building "Smart HR Systems". This dual imperative—balancing efficiency gains with equity and justice—has elevated AI governance as a defining challenge in public administration.^[1]As shown in Fig 1.



Fig 1 Modern Society Entering the AI Era (Source: Internet)

According to 2023 data from the U.S. federal government, AI-driven resume screening systems have significantly improved the efficiency of civil servant recruitment. However, this advancement has simultaneously ignited controversy over "technological unemployment." For instance, amid

existing employment challenges for contemporary college graduates, AI not only replaces certain public sector positions but also introduces new risks—highlighting the dual effects of AI technology in public human resource management.

1.2. AI Reshapes the Structural Framework of the Public Sector

1.2.1 Intelligent Recruitment Systems

Exemplified by the Hangzhou Municipal Government, the adoption of algorithm-optimized job-matching models and multi-dimensional competency mapping technologies has reduced the hiring cycle for critical positions by 65%. Cities like Guangzhou and Shenzhen have similarly transformed their civil servant recruitment: shifting from error-prone, labor-intensive manual screening to AI-driven processes with fixed parameters. By issuing simple AI commands, they now efficiently identify qualified candidates—significantly enhancing precision, employee quality, and operational efficiency.^[2]

While traditional bureaucracy emphasizes rigid rules, AI breaks hierarchical constraints. This injects fluidity into public sector structures, balancing efficiency with standardized rationality. Blockchain further enables distributed sharing of employee competency data (e.g., Xiong'an New Area's "Blockchain + Cadre Archives" initiative). As shown in Fig 2.

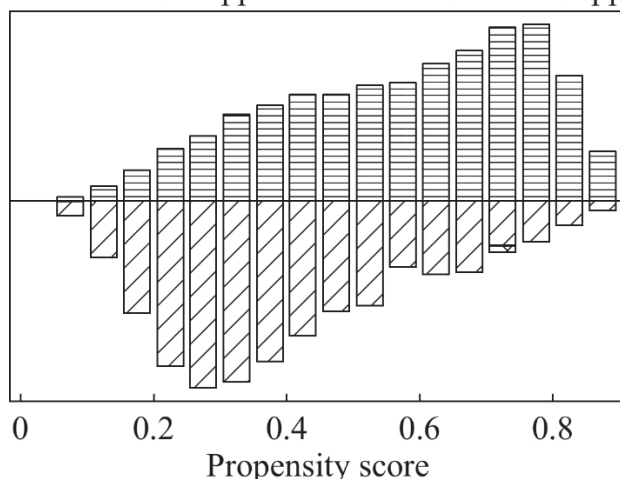


Fig 2 Human capital, urban social inclusiveness and the career development of migrant population(The picture is from the Internet)

1.2.2 Dynamic Performance Systems

With technological advancements, big data-driven 360-degree evaluations now analyze public sector performance through 22 quantitative indicators—including work hours, collaboration frequency, and innovation outputs—enabling more objective, accurate, and granular assessments. This empowers employees to optimize their contributions.

The integration of Theory X and Y elucidates AI's dual role: Theory X (assuming inherent laziness) leverages AI for rigorous oversight, reducing dereliction of duty by ensuring accountability; Theory Y (emphasizing motivation) uses AI to identify high-KPI performers, awarding merit-based incentives that foster a virtuous cycle of competition and service excellence. A pilot project in Shanghai's Xuhui District demonstrates this dynamism: AI performance systems reduced evaluation appeals by 37%, reflecting successful institutionalization of AI-driven management.^[3]

Grassroots units (e.g., communities) now adopt similar AI metrics. For instance, during community internships, performance in tasks like tobacco surveys and housing inspections is automatically scored by AI systems based on data quality and volume, enabling dynamic assessment through digital workflows.

1.2.3 Personalized Training Systems

The widespread application of intelligent technologies such as artificial intelligence and natural language processing is driving profound digital transformation across industries, continuously reshaping the boundaries of professional competencies and the functional core of traditional editorial roles. Particularly within the current environment of "formulaic" and "scripted" interviews, selecting truly suitable candidates and designing their future capacity-building programs pose significant challenges.^[4]In response, multiple municipal governments are adopting AI training systems modeled after the UK's NHS, achieving nearly 90% accuracy in skills gap prediction, while preliminary occupational competency forecasting models are undergoing testing.^[5]As shown in Fig 4, Table 1.

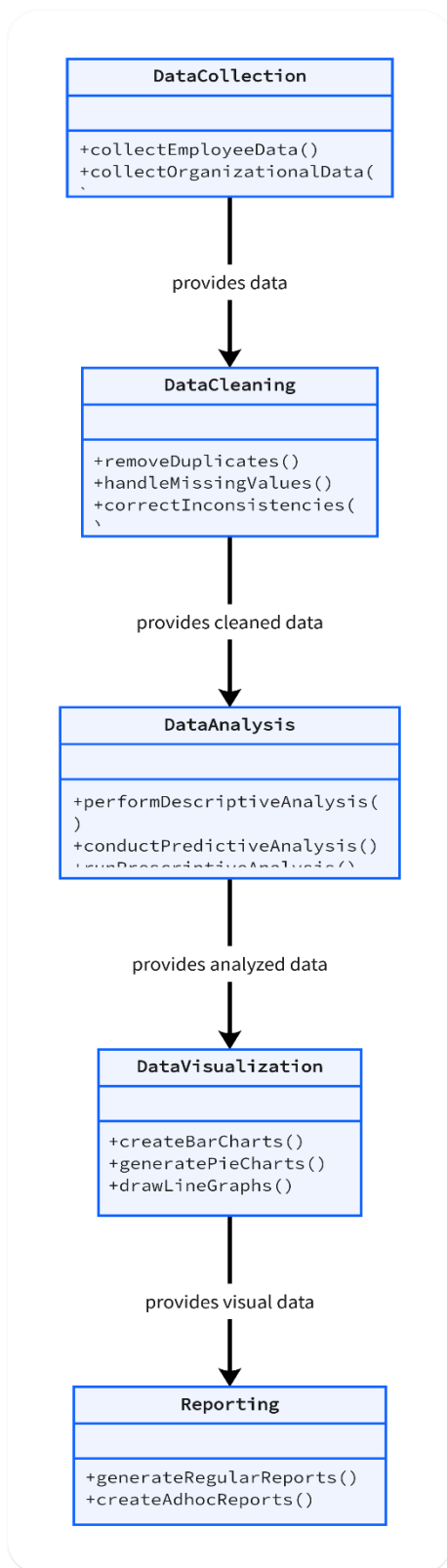


Fig3 Digital Application Scenarios in HRM (Source: Internet)

Table 1 Personalized Training Framework (Self-created)

Module	Technical Implementation	Application Scenario Example
	Technical Methods	
Intelligent Diagnosis	Decision Tree Algorithms	Auto-matching job competency models with individual skill gaps
Course Recommendation	Collaborative Filtering, Reinforcement Learning	Dynamically prioritizing learning content based on real-time needs
Effect Evaluation	Knowledge Graphs + Bayesian Networks	Tracking knowledge mastery and behavioral changes in real-time
Resource Allocation	Transfer Learning + Edge Computing	Local-server-based customized resources for remote civil servants
Integrated Application	Deep Learning Algorithms	Personalized training scheme generation

An increasing number of municipal governments are adopting AI-driven personalized training systems, signaling an inevitable trend toward AI-integrated development. Public sectors leverage machine learning and big data analytics to overcome traditional "one-size-fits-all" training constraints. For example: Suzhou Industrial Park's intelligent platform uses NLP to analyze civil servant roles, generating personalized modules (policy/emergency management) that improved job-role alignment by 50% [6]; while Singapore's "Learning Companion" employs knowledge graphs to shift training from passive to proactive paradigms. These cases demonstrate AI's capacity to diagnose skill gaps and optimize learning paths, yet challenges persist: algorithm bias from data silos and suboptimal human-AI interaction in virtual training.[7]

2. Ethical Risks Arising from AI Applications

2.1.Challenges to Algorithmic Decision-Making Fairness

Algorithmic systems risk perpetuating data biases—e.g., historical male-dominated hiring patterns train AI to favor male candidates, exacerbating gender discrimination in employment. This not only triggers public backlash but undermines recruitment equity. One city's AI promotion system reduced female promotion rates by 12% due to biased historical data, sparking widespread criticism. Similarly, while AI resume screening improves efficiency, it may cause "degree discrimination" by prioritizing elite university graduates, marginalizing competent candidates from ordinary institutions and reinforcing meritocracy-based social inequities.

2.2.Accountability Dilemmas in Human-AI Collaboration

The non-interpretability of deep neural networks triples accountability complexity, hindering decision traceability. Legal gaps compound this: China's Civil Service Law lacks AI accountability provisions, enabling public agencies to misuse AI for power abuse without safeguarding employee rights. AI-integrated performance monitoring—deployed amid incomplete disciplinary frameworks—blurs responsibility lines for violations, frequently triggering disputes.

2.3.Power Alienation Through Data Elements

As pivotal digital-era assets, data elements drive public and private sectors under the "Data Factor X" paradigm.[8]

Excessive Personal Data Collection: A Shenzhen-based system storing >3 million employee behavior records faced lawsuits. Beyond storage challenges, such data leaks to malicious actors could cause irreversible harm.^[9]

Privacy and Security Breaches: Telecom fraud rings increasingly purchase or steal private data illegally. Espionage actors exploit technical vulnerabilities to pilfer scientific/military data—threatening individual privacy and homeland security.

Data power disparities (e.g., education levels, AI literacy) distort information access. AI trained on unverified or low-quality data propagates false outputs, disrupting research, work, and daily life.

3. Governance Pathway Construction

3.1. Technological Governance Dimension

Establish algorithmic auditing mechanisms to prevent AI misuse, guiding ethical deployment. Continuously expand audit dimensions (data cleansing compliance, model interpretability, decision transparency) and implement multi-phase oversight (development-deployment-maintenance) to ensure AI remains aligned with public interest, sustaining technological governance stability of the country.^[10]

3.2. Institutional Innovation Dimension

Governments should enact Public Sector AI Ethical Guidelines enforcing core principles—fairness priority, minimal harm, citizen empowerment—through dual-track mechanisms: ethics review committees + public hearings. Concurrently, refine relevant laws and regulations. Society must promote AI literacy—educating citizens on responsible usage and risks—to build an empowered "AI-for-good" ecosystem. Individuals should develop AI competency, embrace lifelong learning, enhance self-discipline, and utilize AI ethically rather than exploit it for malicious purposes. As shown in Table 2.

Table 2 Three-Dimensional Application Framework (Self-created)

Theoretical Tool	Technical Application Scenario	Effectiveness Metric
Scientific Management Theory	Intelligent Scheduling System	Productivity per Employee Increase
New Public Management Theory	Performance Data Visualization Dashboard	Decision Response Speed Improvement
Inclusive Management Theory	Disability Hiring Algorithm	Special Group Hiring Rate Growth

By integrating New Public Management theory with data-driven decision-making, we enable human-AI collaborative governance through a New Public Service lens. This empowers public servants to provide scientifically grounded, context-sensitive recommendations—delivering grassroots-level institutional services that concretely benefit every citizen. As shown in Fig 5.

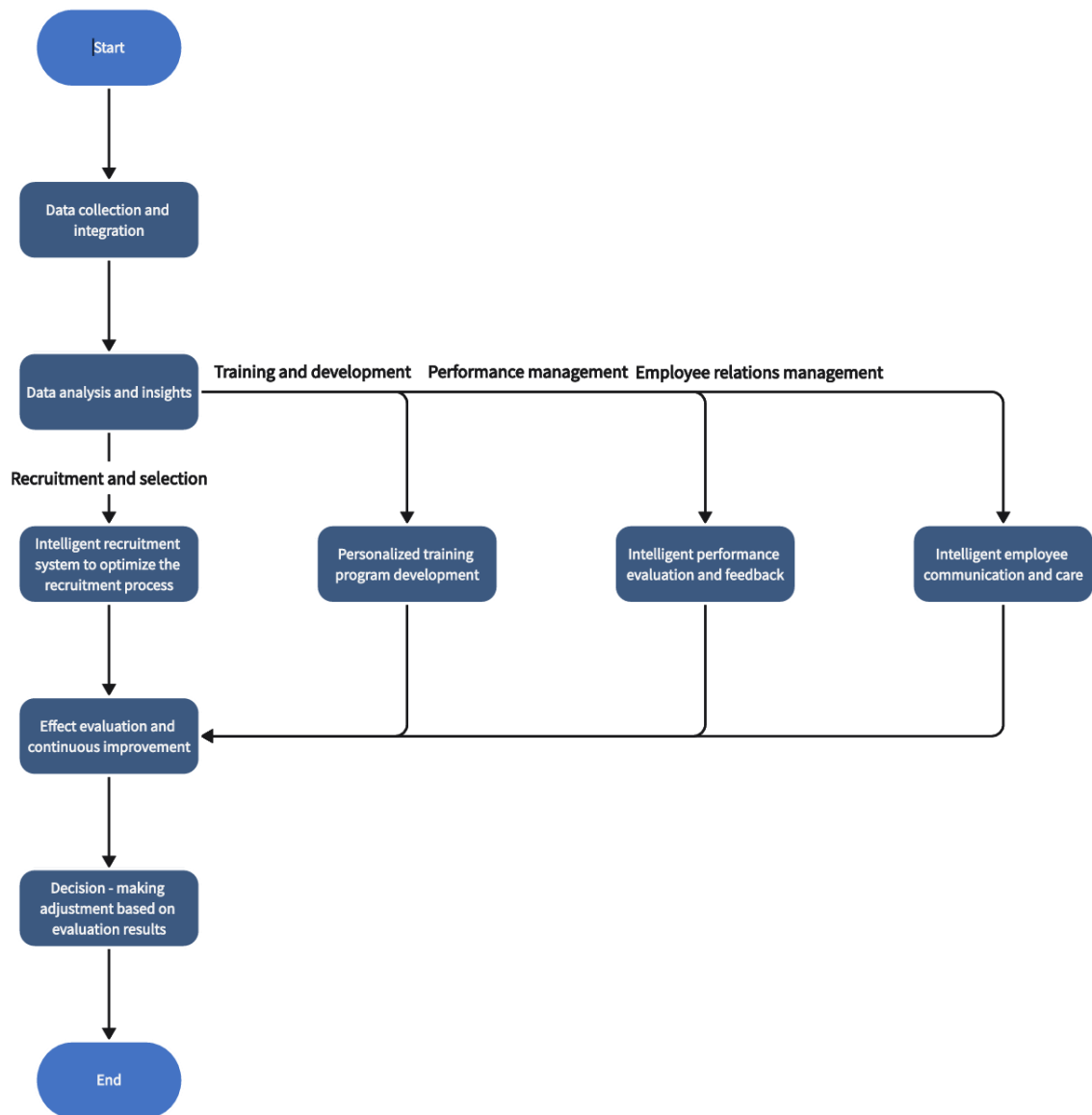


Fig 4 Governance Pathway Flowchart (Self-created)

3.3.Capacity Building Dimension

3.3.1 Developing Digital Governance Literacy Curricula

Integrate algorithmic thinking, data literacy, and human-AI ethics with Taylor’s Scientific Management Principles through standardized, specialized modules to enhance efficiency. Establish a tiered certification system (basic/intermediate/advanced) and develop risk matrices using two-dimensional assessment models (data breach probability vs. impact severity). Leverage IoT devices (e.g., smart ID badges) for real-time data collection, minimizing manual errors, and deploy VR simulations to train civil servants in routine and emergency scenarios.^[11]

3.3.2 Public-Aligned Reform and Development

Surveys covering all age groups across multiple regions (n>240, n>30), conducted via community outreach and social media, reveal >50% public preference for digitized/virtualized cultural artifacts—key vehicles of urban cultural heritage. Word clouds highlight "digitalization" and "accessibility" as top priorities. Thus, public sectors must center reforms on citizen needs: internally optimizing AI

utilization while externally aligning with societal trends. The keywords confirmed in the questionnaire include: artificial intelligence, digitalization, and cultural creativity.

4. Conclusion

Fundamentally, upholding competency-based principles is paramount to prevent nepotism and cronyism in public appointments. To this end, HR experts in the public sector continuously develop rationalized management technologies—such as AI-driven job analysis, position evaluation, personnel assessment, performance systems, and training cost-benefit analytics—to establish objective HR mechanisms. These advance equitable incentives: equal pay for equal work, proportional rewards for labor input, and zero entitlement without contribution.

The shift from traditional personnel administration to intelligent public HR management demands governments transition from task-centric to human-centered models. A proactive, dutiful, and competent civil service remains the cornerstone of good governance and effective public administration.

AI is reshaping the foundational logic of public sector HR, embodying a tension between technorationality and public values. Our proposed governance framework—synergizing technological innovation, institutional restructuring, and capacity building—provides systemic solutions to the "efficiency paradox." Future efforts must prioritize large-model applications in public personnel management, refine human-centric intelligent governance ecosystems, and accelerate the public sector's AI-enabled transformation to advance China's rejuvenation in the intelligent era.

5. Acknowledgments

Funding: 2020 Research Startup Fund for Talent Introduction in Heilongjiang Provincial Universities, Project Title: "Research on Innovation Mechanisms of Urban Community Interactive Governance in Promoting Grassroots Governance Efficiency" (Project No.: 13051202022), Based on the theory of modernization of rural grassroots governance and the current situation of regional differences in modernization of rural governance, the study analyzes the influencing factors of modernization of rural governance using the CRITIC-entropy power method, the TOPSIS comprehensive evaluation method and the regression analysis method, The research on the theoretical system of Chinese modernization and its response to the current social development reality in China, funded by the Basic Scientific Research Operating Expenses of Provincial Universities, Project Number: 2023JCYJ-03., 2023.07.

Reference

- [1] Li Deyi. Introduction to Artificial Intelligence[M]. Beijing: China Science and Technology Press, 2021.74.
- [2] General Office . Guiding Opinions on Accelerating the Standardization, Regularization, and Facilitation of Government Services[Z]. 2022-02-07.
- [3] Weng Yiqin. New Vision of Intelligent Creative Editing Training[J]. Cultural Industry, 2025, (17): 22-24.
- [4] Wang Fang, Zhang Yu. Ethical dilemmas of digital transformation in public sector human resource management[J]. Chinese Administrative Management, 2023(5): 45-52.
- [5] World Economic Forum. The Future of Jobs Report 2023[R]. Geneva: WEF, 2023.
- [6] Suzhou Industrial Park Management Committee. Construction Plan for Suzhou Industrial Park Smart Training Platform[R]. Suzhou: Suzhou Industrial Park Management Committee, 2021.
- [7] Singapore Government Technology Agency. White Paper on the Civil Servant Learning Manager (SGLearn) System[R]. Singapore: Singapore Government Science and Technology Agency, 2022.
- [8] Russell S. Human Compatible: Artificial Intelligence and the Problem of Control[M]. London: Penguin Books, 2019: 89-102.

- [9] Yan Bing, Zheng Xia, Wang Bo. "Data Element ×" Scenario-Driven Citizen Green and Low-Carbon Action Operation Guide System V1.0[CP]. 2023SR1234567. Beijing: 2025.
- [10] European Commission. Ethics Guidelines for Trustworthy AI[EB/OL]. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, 2021-04-08.
- [11] Osborne D, Gabbler T. Ditching Bureaucracy: Five Strategies for Reinventing Government [M]. Translated by Zhou Dunren, et al. Shanghai: Shanghai Translation Publishing House, 2006: 112-135.