STM JOURNALS

Current Trends in
Information Technology
https://journals.stmjournals.com/ctit

ISSN: 2249-4707 (Online)
ISSN: 2348-7895 (Print)
Volume 16, Issue 1, 2026
January–April
DOI (Journal): 10.37591/CTIT

Review CTIT

Human-in-the-Loop AI in HR Decision-Making: Insights from Big 4 AI Governance Reports

Laxmee Vachher[1,*], Shyamali Dubey[2]

Abstract

The integration of artificial intelligence (AI) in human resource (HR) decision-making has transformed recruitment, performance evaluations, and talent management. However, biases embedded in AI-driven HR systems present significant ethical and operational challenges. Human-in-the-Loop (HITL) AI offers a hybrid approach that combines AI efficiency with human oversight to enhance fairness and accountability. This paper systematically analyses HITL AI in HR decision-making using qualitative analysis of AI governance reports from Big 4 consulting firms. Findings indicate that HITL AI enhances fairness and compliance in HR decisions but faces challenges in scalability, workforce acceptance, and regulatory complexity. The study provides insights into governance frameworks and best practices for balancing AI efficiency with human oversight. Additionally, this paper builds upon prior research on AI bias in HR analytics by extending the discussion towards practical implementations of HITL frameworks and assessing their effectiveness through industry insights.

Keywords: AI governance, human-in-the-loop AI, HR analytics, Big 4 consulting, algorithmic bias, ethical AI, explainable AI (XAI)

INTRODUCTION

The adoption of artificial intelligence (AI) in human resource (HR) decision-making is increasing, yet concerns about fairness, compliance, and bias persist. Human-in-the-loop (HITL) AI offers a hybrid approach in which human oversight mitigates the algorithmic bias. However, challenges such as human bias in oversight, cost implications, and gaps in AI literacy must be addressed. This study explores these complexities while presenting governance strategies outlined by the Big 4 consulting firms.

Increasing reliance on AI in HR analytics has led to significant advancements in efficiency and decision-making accuracy. AI-driven tools are widely used for talent acquisition, performance management, workforce planning, and employee engagement. These technologies promise objectivity and data-driven insights. However, AI models often inherit biases from historical datasets and algorithmic frameworks, leading to potential discrimination in HR decisions. The absence of human intervention in AI-driven HR processes can result in unfair hiring practices, biased promotions, and flawed performance evaluations.

HITL AI has emerged as a crucial approach to addressing these challenges by integrating human judgment at key decision-making points. Unlike fully automated systems, HITL AI allows human experts to validate, modify, or override AI-generated recommendations, ensuring fairness and ethical decision-making. Regulators and industry leaders have increasingly emphasized the role of human oversight in AI to ensure transparency, accountability, and compliance with emerging policies.

***Author for Correspondence**
Laxmee Vachher
E-mail: drlv1447@bbdu.ac.in

[1]Associate Professor, School of Management, Babu Banarasi Das University, Lucknow, Uttar Pradesh, India
[2]Associate Professor, School of Management, Babu Banarasi Das University, Lucknow, Uttar Pradesh, India

Received Date: September 15, 2025
Accepted Date: September 20, 2025
Published Date: February 07, 2026

Citation: Laxmee Vachher, Shyamali Dubey. Human-in-the-Loop AI in HR Decision-Making: Insights from Big 4 AI Governance Reports. Current Trends in Information Technology. 2026; 16(1): 43–50p.

Moreover, the HITL approach is gaining traction owing to its ability to bridge the gap between AI efficiency and human ethical reasoning. While AI can process large volumes of data rapidly, it often lacks the contextual awareness and adaptability provided by human oversight. Organizations implementing HITL AI benefit from a hybrid system in which AI handles data processing and pattern recognition, while humans intervene to correct biases and ensure that decisions align with ethical and legal standards. The incorporation of human feedback loops further refines the AI models over time, making them more robust and adaptive to evolving HR practices.

As AI regulations continue to evolve, industries are recognizing the necessity of integrating HITL models to comply with global standards such as the EU AI Act and Equal Employment Opportunity Commission (EEOC) guidelines. Many enterprises are actively shifting toward AI governance frameworks that emphasize transparency, fairness, and accountability, with consulting firms such as Deloitte, PwC, EY, and KPMG leading the way in providing best practices for HITL AI adoption. These firms stress the importance of explaining ability, bias audits, and human validation checkpoints to ensure that AI-driven HR decisions align with corporate social responsibility and legal compliance.

This study explores how Big 4 consulting firms advise organizations to implement HITL AI in Human Resource Management (HRM). Their insights offer valuable strategies for mitigating AI bias, ensuring compliance with regulatory frameworks, and fostering the adoption of ethical AI in HR practices. Furthermore, this study builds on prior research on AI bias in HR analytics by shifting the focus from identifying biases to evaluating strategies for incorporating human oversight to enhance fairness and accountability in AI-driven HR decision-making. Additionally, this paper examines the challenges of integrating HITL AI, including workforce resistance, scalability, cost implications, and the need for AI literacy training within HR teams. These findings will contribute to a deeper understanding of how HITL AI can shape the future of fair and transparent HR analytics.

OBJECTIVES

- Examine the role of HITL AI in mitigating bias in HR decision-making.
- Analyze Big 4 consulting firms' strategies for AI governance and fairness audits in HR analytics.
- Evaluate challenges and future considerations for implementing HITL AI at scale.

LITERATURE REVIEW

The concept of HITL AI in HR has gained prominence owing to rising concerns about algorithmic bias in HR analytics. AI-driven tools analyze large datasets to predict candidate suitability, assess employee performance, and optimize workforce planning. However, multiple studies have indicated that these AI models may inadvertently amplify existing biases in historical data [1]. The bias in AI-driven HR systems often stems from imbalanced training datasets, inadequate data representation, and inherent algorithmic limitations. Consequently, there is growing emphasis on integrating human intervention in AI-driven HR decision-making to ensure fairness, transparency, and accountability.

AI bias mitigation, regulatory compliance, and governance strategies have been studied extensively. Recent research highlights the role of fairness-aware retraining of AI models in HR decisions [1]. Additionally, human-AI interaction loops have been shown to increase the predictive accuracy in hiring and performance evaluations [2]. However, empirical studies quantifying the effectiveness of the HITL AI remain limited.

Defining Human-in-the-Loop AI in HR

HITL AI refers to systems that incorporate human oversight at key decision points in automated processes. Unlike fully automated AI models, HITL AI ensures that human reviewers validate, modify, or override AI-generated decisions to prevent errors and bias. This hybrid model is particularly valuable in HR analytics, in which fairness, ethical considerations, and compliance with regulatory standards are critical.

Human-in-the-Loop AI for Bias Mitigation

HITL AI ensures human oversight at critical decision points, allowing experts to validate, adjust, or override AI-generated recommendations. Deloitte (2024) [3] stressed that AI fairness audits and bias detection tools are crucial in identifying and mitigating bias. Their research highlighted the need for continuous human monitoring and interventions to counteract model drift and unintended biases. Furthermore, PwC (2023) [4] underscores the role of explainable AI (XAI) in HR analytics, advocating periodic human reviews to effectively interpret and validate AI outputs. Studies also indicate that organizations integrating human feedback loops into AI-driven hiring processes experience improved fairness and transparency [2].

Governance Frameworks for HITL AI in HR

The implementation of the HITL AI requires well-defined governance frameworks. EY (2023) and KPMG (2024) [5, 6] emphasize the importance of ethical AI governance, risk assessment methodologies, and compliance mechanisms. EY focuses on integrating fairness-aware algorithms to ensure that the AI models used in HR adhere to ethical and legal standards. The KPMG highlights the necessity of risk assessment frameworks that enable organizations to identify potential ethical dilemmas and biases before deploying AI-driven HR analytics. Case studies have demonstrated that organizations adopting structured governance models have greater AI accountability and improved HR decision-making outcomes.

Regulatory Compliance and HITL AI in HR

Compliance with AI regulations, such as the EU AI Act, EEOC guidelines, and ISO AI standards, necessitates the adoption of HITL models in Human Resource Management (HRM). Organizations that integrate human oversight into AI decision-making processes report higher legal adherence and fewer discrimination claims [2]. Deloitte (2024) [3] suggested that organizations must develop AI governance strategies aligned with these regulatory frameworks to avoid legal and reputational risks. A key emerging trend is the push for AI to explainability, ensuring that HR-related AI decisions remain interpretable and justifiable in legal and ethical contexts.

Challenges in Implementing HITL AI

Despite these advantages, the implementation of HITL AI in HR is challenging. One key issue is ensuring unbiased human oversight. Studies have shown that even human reviewers can exhibit cognitive biases, potentially reinforcing discriminatory outcomes rather than mitigating them [1]. Structured bias training programs and audit mechanisms are being increasingly explored as solutions to address these concerns.

Scalability is another challenge, particularly for large organizations that process vast amounts of HR data daily. PwC (2023) [4] highlights the necessity of developing scalable governance models that allow organizations to maintain human oversight without significantly reducing AI efficiency. This involves leveraging semi-automated review systems where AI flags potentially biased decisions for human review rather than requiring human intervention in every instance. Additionally, integrating AI literacy training into HR teams is essential to ensure effective AI governance and decision-making.

Case Studies on HITL AI in HR

Several frameworks have been proposed to effectively integrate human oversight into AI systems. Deloitte (2024) [3] emphasized the role of fairness audits, bias detection tools, and explainable AI (XAI) in ensuring transparency. PwC (2023) [4] highlighted the importance of periodic human intervention in AI-driven HR processes to validate predictions and improve decision-making reliability. EY (2024) and KPMG (2025) [5, 6] focused on developing governance models that incorporate ethical AI practices, risk assessment methodologies, and compliance mechanisms. These reports collectively stress that HITL AI models should not replace human HR professionals but rather augment their capabilities by providing data-driven insights while maintaining ethical and fair decision-making.

In addition to industry reports, real-world applications of HITL AI have provided valuable insights into its impact. Case studies of leading multinational corporations demonstrate that integrating human oversight into AI-driven HR analytics improves hiring fairness, enhances diversity, and increases compliance with regulatory guidelines. However, challenges such as scalability, training requirements, and subjective biases introduced by human reviewers remain key concerns that require further study [7, 8].

Furthermore, HITL AI is increasingly being integrated with fairness-aware algorithms to enhance predictive accuracy. Research highlights that organizations employing a combination of HITL and algorithmic fairness mechanisms report improved equity in hiring outcomes. Additionally, the literature underscores the role of continuous feedback loops, where human reviewers provide real-time corrections to AI-driven decisions, thereby refining the model over time [9–11]. This dynamic adaptation mechanism ensures that AI models remain responsive to evolving workplace diversity and ethical considerations.

Future Considerations for HITL AI in HR

As organizations continue to adopt AI-driven HR analytics, the role of HITL AI will become increasingly critical [12, 13]. Future research should focus on developing advanced fairness-aware algorithms that reduce the need for human intervention while maintaining ethical decision-making. Additionally, policymakers and industry leaders must collaborate to establish standardized AI governance frameworks that ensure consistency in HITL AI implementation across organizations and industries [14]. The key areas for further exploration include the following.

- *Empirical evaluation of HITL AI effectiveness:* More longitudinal studies assessing the tangible impact of HITL AI on reducing bias in HR decisions.
- *Scalability models:* Identifying best practices for balancing automation and human intervention in large-scale HR operations.
- *HR workforce training:* Ensuring HR professionals receive adequate AI literacy training to effectively interact with and oversee AI-driven systems.
- *Standardized AI governance frameworks:* Establishing industry-wide benchmarks for measuring AI fairness and accountability.

Overall, HITL AI presents a viable solution for reducing algorithmic bias in HR decision-making. By leveraging insights from Big 4 consulting firms and industry case studies, organizations can implement best practices for integrating human oversight into AI-driven HR analytics, ensuring fairness, transparency, and compliance with evolving regulatory standards [15–17].

METHODOLOGY

This study systematically reviewed 12 AI governance reports from Deloitte, PwC, EY, and KPMG, using thematic content analysis. Key themes such as bias reduction, regulatory compliance, and best practices for AI implementation in HR were identified and analyzed [18].

This study employed secondary data analysis to assess the implementation of HITL AI in HR decision-making. The research relies on qualitative and quantitative secondary data from academic research papers, Big 4 consulting firm reports, industry white papers, regulatory documents, and corporate case studies. By conducting a systematic review of existing literature and consulting reports, this study aims to synthesize key insights into how organizations integrate human oversight into AI-driven HR practices [19, 20].

The methodology involves:

- *Comparative analysis:* Examining AI governance strategies from Deloitte, PwC, EY, and KPMG to identify common themes, best practices, and unique contributions to HITL AI adoption.
- *Case study review:* Documented case studies of organizations implementing HITL frameworks were analyzed to assess their effectiveness in balancing AI efficiency and human judgment.

- *Regulatory framework evaluation:* Review AI governance policies, such as the EU AI Act, EEOC guidelines, and ISO AI standards, to understand how compliance measures impact HITL AI adoption in HR.
- *Bias mitigation techniques:* Identifying fairness-aware algorithms, risk assessment methodologies, and transparency initiatives that enhance the ethical deployment of AI in HR decision-making.

This approach ensures a comprehensive understanding of how organizations leverage HITL AI, while addressing potential risks and limitations.

FINDINGS AND DISCUSSION

HITL AI in HR Decision-Making: Key Insights

HITL AI has emerged as a pivotal framework for mitigating bias and enhancing fairness in HR decision-making. The integration of human oversight allows for ethical intervention in AI-generated recommendations, thereby reducing the risks associated with fully automated HR processes. Key findings from the analysis of Big 4 consulting firms' AI governance reports highlight the following aspects.

Effectiveness in Bias Mitigation

- HITL AI significantly reduces bias in recruitment, performance evaluations, and talent management by enabling human reviews at critical decision points.
- Organizations that implement fairness-aware AI models combined with human oversight report increased diversity in their hiring outcomes.
- Periodic AI fairness audits ensure that algorithmic biases are identified and addressed proactively.

Enhancing Explainability and Transparency

- Explainable AI (XAI) is a core focus of HITL AI governance, ensuring that HR professionals understand the AI-generated recommendations.
- Deloitte and PwC emphasized the need for interpretability in AI models, enabling HR teams to validate decision-making processes.
- Employee trust in AI-driven HR analytics improves when AI recommendations are explainable and are subject to human validation.

Regulatory Compliance and Legal Considerations

- HITL AI aligns with emerging regulatory frameworks such as the EU AI Act and EEOC guidelines by ensuring fairness and accountability.
- Human oversight helps organizations avoid the legal risks associated with biased AI-driven hiring practices.
- The AI governance strategies proposed by Big 4 firms emphasize the importance of compliance audits and risk assessment methodologies.

Big 4 Consulting Firms' Insights on HITL AI Governance

An analysis of AI governance reports from Deloitte, PwC, EY, and KPMG revealed strategic approaches to HITL AI adoption.

These firms advocate AI governance frameworks that integrate human oversight and ensure fairness, transparency, and compliance in HR decision-making Table 1.

Table 1. Focus area of Big 4 consulting firms.

Big 4 consulting firms	Key focus areas in HITL AI governance
Deloitte	Regulatory compliance, fairness audits, bias detection tools
PwC	Explainable AI (XAI), periodic human intervention, and ethical AI training
EY	Fairness-aware algorithms, structured risk assessment methodologies
KPMG	AI-human collaboration models, legal risk mitigation strategies

CHALLENGES IN HITL AI IMPLEMENTATION

Despite its benefits, HITL AI adoption presents several challenges that organizations must address for successful implementation.

Technical Challenges

- *Scalability issues:* Large enterprises face difficulties in scaling HITL AI because of the high volume of HR decisions that require human intervention.
- *Data processing constraints:* Ensuring real-time AI recommendations while allowing human validation can slow down decision-making.
- *Algorithm adaptation:* Continuous refinement of AI models based on human feedback is essential but resource intensive.

Organizational Challenges

- *Workforce resistance:* Employees may distrust AI-driven decision-making, perceiving it as intrusive or as a threat to job security.
- *HR skill gaps:* Many HR professionals lack AI literacy, making it difficult to effectively oversee AI-driven processes.
- *Cost of implementation:* Establishing HITL AI frameworks requires investment in training, software updates, and compliance audits.

Ethical and Governance Challenges

- *Human bias in oversight:* While HITL AI aims to mitigate algorithmic bias, human reviewers may introduce their own cognitive biases.
- *Legal accountability:* Determining the responsibility for AI-driven HR decisions remains complex, especially when human intervention modifies AI output.
- *Standardization issues:* The lack of uniform AI governance frameworks across industries creates inconsistencies in HITL AI implementations.

FUTURE CONSIDERATIONS FOR HITL AI IN HR

To optimize HITL AI implementation, organizations should focus on the following areas.

Standardizing AI Fairness Benchmarks

Establish industry-wide AI fairness evaluation criteria to consistently measure HITL AI effectiveness. Develop best practices for periodic bias audits and compliance monitoring.

Enhancing AI Literacy in HR Teams

Conduct targeted training programs to equip HR professionals with the skills required to oversee AI-driven decision-making. Foster collaboration between HR and data science teams to improve AI interpretability.

Improving AI-Human Collaboration Models

Implement semi-automated review systems where AI flags high-risk decisions for human validation rather than requiring full human intervention. Leverage feedback loops to refine AI models continuously, thereby improving fairness and accuracy over time.

Adapting to Evolving Regulations

Stay updated with global AI governance policies to ensure compliance with emerging legal standards. Collaborate with policymakers and industry experts to shape ethical AI adoption frameworks.

CONCLUSION

HITL AI presents a viable solution to AI bias in HR decision-making by integrating human expertise with algorithmic intelligence. Insights from the Big 4 AI governance reports highlight the best practices for balancing automation and human intervention to ensure fairness, transparency, and legal compliance. Organizations should implement AI literacy programs, periodic bias audits, and standardized governance frameworks. Future research should focus on the empirical testing of HITL AI in diverse HR settings to quantify its impact on bias mitigation and decision fairness. Establishing standardized AI governance benchmarks is crucial for ensuring consistency across industries.

REFERENCES

1. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM Comput Surv. 2021;54(6):1–35. doi:10.1145/3457607.
2. Wilson HJ, Daugherty P, Bianzino N. The jobs that artificial intelligence will create. MIT Sloan Manag Rev. 2017;58(4):14. Available from: https://sloanreview.mit.edu/article/will-ai-create-as-many-jobs-as-it-eliminates/
3. Ammanath B, Mitchell K. (2024). Preparing the workforce for ethical, responsible and trustworthy AI. [Online]. Deloitte US. Available from: https://www.deloitte.com/us/en/about/articles/ethical-technology-survey.html
4. PwC (2023). A Year of Solving Together. Global Annual Review 2023. [online] Available from: https://www.pwc.com/gx/en/global-annual-review/2023/pwc-global-annual-review-2023.pdf.
5. Serruys H. Navigating the AI revolution: the future of the HR role. Belgium: Ernst & Young Global Limited; 2024 Feb 15. Available from: https://www.ey.com/en_be/technical/tax/tax-alerts/2024/navigating-the-ai-revolution-the-future-of-the-hr-role
6. KPMG International Limited. Transparency Report 2025. London: KPMG International Limited; 2025. Available from: https://kpmg.com/xx/en/about/transparency-report.html
7. European Commission. Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Brussels: European Commission; 2021 Apr 21. Report No.: COM(2021) 206 final; 2021/0106(COD). 2021.
8. Karasik AW. EEOC's artificial intelligence evolution. Brief. 2024;53(2):12.
9. Benraouane SA. AI Management System Certification According to the ISO/IEC 42001 Standard: How to Audit, Certify, and Build Responsible AI Systems. New York: Productivity Press; 2024. 226 p. doi:10.4324/9781003463979.
10. Binns R. Fairness in machine learning: lessons from political philosophy. In: Friedler SA, Wilson C, editors. Proceedings of the 1st Conference on Fairness, Accountability and Transparency; 2018 Feb 23–24; New York, NY, USA. Proceedings of Machine Learning Research. Vol 81. Breckenridge, Colorado, USA: PMLR; 2018. p. 149–159. Available from: https://proceedings.mlr.press/v81/binns18a.html
11. Raji ID, Buolamwini J. Actionable auditing: investigating the impact of publicly naming biased performance results of commercial AI products. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society; 2019 Jan 27–28; Honolulu, HI, USA. New York (NY): Association for Computing Machinery; 2019. p. 429–435. doi:10.1145/3306618.3314244.
12. Binns R, Van Kleek M, Veale M, Lyngs U, Zhao J, Shadbolt N. "It's reducing a human being to a percentage": perceptions of justice in algorithmic decisions. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018 Apr 21–26; Montreal, QC, Canada. New York (NY): Association for Computing Machinery; 2018. p. 1–14. doi:10.1145/3173574.3173951.

13. Floridi L, Cowls J. A unified framework of five principles for AI in society. Harv Data Sci Rev. 2019 Jul 1;1(1):535–545. Available from: https://hdsr.mitpress.mit.edu/pub/l0jsh9d1
14. Binns R. On the apparent conflict between individual and group fairness. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20); 2020 Jan; Barcelona, Spain. New York (NY): Association for Computing Machinery; 2020. p. 514–524. doi:10.1145/3351095.3372864.
15. Barocas S, Hardt M, Narayanan A. Fairness and Machine Learning: Limitations and Opportunities. Cambridge (MA): MIT Press; 2023.
16. Crawford K. The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. New Haven (CT): Yale University Press; 2021.
17. Tolan S, Miron M, Gómez E, Castillo C. Why machine learning may lead to unfairness: evidence from risk assessment for juvenile justice in Catalonia. In: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law. New York: ACM; 2019. p. 83–92. doi:10.1145/3322640.3326705.
18. Zhang BH, Lemoine B, Mitchell M. Mitigating unwanted biases with adversarial learning. In: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. New York: ACM; 2018. p. 335–340. doi:10.1145/3278721.3278779.
19. Sandvig C, Hamilton K, Karahalios K, Langbort C. Auditing algorithms: research methods for detecting discrimination on internet platforms. Data Discrimination. 2014;22:4349–4357.
20. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–453. doi:10.1126/science.aax2342. PubMed: 31649194.

