Human by Design: The Influence of Artificial Intelligence and Human Stewardship in Recruitment

Christine Marquis, DM
Center for Organizational Wellness and Belonging (COWEB)
4035 South Riverpoint Parkway, Phoenix, AZ 85040-0723
USA
cvgmarquis@gmail.com

Abstract: The rapid adoption of artificial intelligence (AI) in recruitment has outpaced ethical oversight, creating risks related to transparency, bias, and candidate dignity. This qualitative descriptive study examined how organizations implement and govern AI-enabled recruitment processes, emphasizing the role of human stewardship in ethical decision-making. Data was collected through semi-structured interviews with HR, diversity, and IT professionals, and analyzed using Socio-Technical Systems theory (STS) and the Ethics of Care (EoC) framework. Findings revealed the perceived value of AI depends heavily on disciplined human oversight, formal governance routines, and clear accountability mechanisms. Participants stressed that transparency, escalation protocols, and documented human review are essential to mitigate ethical risks. The study suggests that AI-enabled systems must be intentionally designed as human-centered socio-technical systems to ensure equitable hiring and institutional accountability.

Introduction

Integrating artificial intelligence (AI) into recruitment processes is reshaping how organizations manage hiring and evaluate candidates, with implications for education-to-work pathways. As AI-enabled systems mediate screening, assessment, and selection, they influence how individuals experience institutional judgment, transparency, and access to opportunity (Drage & Mackereth, 2022). Research indicates the accelerated adoption of AI has outpaced ethical governance, leaving gaps in oversight and explainability that undermine trust and accountability (Gabriel, 2020; Hagendorff, 2020). These risks are amplified when AI systems rely on historical training data that may encode prior organizational and societal biases, thereby automating exclusionary practices under the appearance of objectivity (Burrell & McAndrew, 2023).

From a socio-educational perspective, recruitment increasingly functions as a formative encounter in which institutional values are learned and reinforced. AI-enabled recruitment, therefore, operates as a learning environment within the education-to-work pipeline, shaping how candidates, practitioners, and institutions interpret transparency, accountability, and access to opportunity.

The purpose of this qualitative descriptive study was to examine how organizations implement and govern AI-enabled recruitment processes, with particular attention to the role of human stewardship in ethical decision-making. Data was collected through semi-structured interviews with HR leaders, diversity professionals, and information technology specialists, supplemented by recruitment-related documentation to support triangulation.

The study was guided by Socio-Technical Systems theory (STS), which emphasizes the joint optimization of technical systems and social structures, and the Ethics of Care (EoC) framework, which foregrounds relational responsibility and the preservation of human dignity in decision-making contexts (Oruc & Sarikaya, 2011; Zhang et al., 2023). Findings indicate the perceived value of AI is inseparable from disciplined human oversight, formal governance routines, and transparent accountability mechanisms, reinforcing the importance of designing AI-enabled recruitment as a human-centered socio-technical system.

Methods

This study used a qualitative descriptive design to examine how AI is implemented and governed within organizational decision-making. Qualitative descriptive approaches provide comprehensive, low-inference accounts of real-world practices that remain close to participants' language and experiences (Sandelowski, 2000). This design enabled the examination of AI-enabled systems as socio-technical environments in which human judgment, organizational routines, and ethical considerations intersect. The approach aligned with the study's focus on describing operational governance practices rather than evaluating system performance.

Data was collected through semi-structured interviews with human resource leaders, diversity professionals, and information technology specialists, supplemented by document analysis. Interviews enabled participants to reflect on governance routines, ethical decision points, and human oversight practices associated with AI-enabled systems (Anthony & Jack, 2009). Document analysis included organizational policies and publicly available regulatory materials to support triangulation and contextual interpretation of interview data (Yin, 2018).

Analysis followed Braun and Clarke's (2022) six-phase reflexive thematic analysis to identify patterns across participants' perspectives and documentation. The analytic process was informed by STS theory and the EoC framework, which guided attention toward relational accountability and human responsibility within technology-mediated environments.

Results

Human Oversight and Safeguards

Human oversight emerged as a consistent practice shaping how artificial intelligence–enabled systems were used within organizational decision environments. Participants described routine involvement of human reviewers at key decision points, particularly when AI-generated outputs affected candidate screening, ranking, or exclusion. Rather than relying on automated recommendations as final determinations, leaders emphasized escalation protocols, exception handling, and manual review processes to address ambiguous or high-impact outcomes. These safeguards were embedded within formal workflows, including documented review checkpoints and role-based accountability, rather than applied informally or at individual discretion.

The study's findings indicated that participants viewed human oversight as essential to maintaining ethical responsibility and organizational credibility when using AI-enabled tools. Collectively, these findings indicate that AI use was framed as decision support rather than decision replacement, with human judgment retained as a governing mechanism.

Governance Routines and Accountability

Governance routines played a central role in shaping how AI–enabled systems were monitored and managed within organizational decision environments. Participants described formal policies, review procedures, and documentation requirements that guided the use of AI beyond the initial implementation. Accountability for AI-related decisions was typically distributed across defined roles rather than assigned to a single individual, with responsibility shared among human resources, information technology, and compliance functions. These routines included periodic reviews of system outputs, documentation of decision-making rationales, and mechanisms to address identified issues or discrepancies. Rather than treating governance as a static control, organizations described it as an ongoing process that evolved alongside system use. These practices illustrate how accountability was operationalized through routine organizational processes rather than relying solely on individual discretion.

Transparency and Explainability

Transparency and explainability were described as key features shaping the implementation and communication of AI-enabled decision processes within organizational settings. Participants reported variation in the extent to which AI use was disclosed and in the clarity with which decision logic and system outputs were explained to stakeholders. In some settings, limited visibility into how AI-generated recommendations were produced or applied constrained understanding of decision outcomes.

In contrast, participants described practices such as disclosing AI involvement, documenting decision stages, and providing opportunities for human clarification as supporting clearer interpretation of AI-assisted decisions. Explainability was not limited to technical detail but included how decisions were justified and communicated in accessible terms. These findings describe how transparency and explainability were

operationalized as procedural elements within AI-enabled systems rather than as abstract ethical principles.

Discussion

The findings indicate that the value of AI in organizational decision-making is inseparable from disciplined human stewardship. Human oversight, governance routines, and transparency mechanisms functioned as stabilizing structures that shaped how AI systems were interpreted and constrained in practice. From an STS perspective, AI does not operate independently but depends on the joint configuration of technical tools, human roles, and organizational processes to function responsibly (Oruc & Sarikaya, 2011). When these elements were aligned, AI-supported decisions were treated as provisional inputs subject to review, escalation, and correction. This framing challenges efficiency-driven narratives that position AI as a substitute for judgment. Human stewardship, therefore, emerges as a design requirement rather than a compensatory safeguard.

The study also highlights the importance of transparency and explainability as operational practices rather than abstract ethical commitments. Participants described disclosure, explanation, and access to human clarification as procedural elements that shaped how AI-enabled decisions were understood and managed. Viewed through an EoC lens, these practices reflect attentiveness, responsibility, and responsiveness to those affected by technology-mediated decisions (Parra Jounou & Tronto, 2024). Transparency served not only as a compliance mechanism but as a means of preserving dignity and agency within automated decision environments. For institutions operating at the education-to-work boundary, these findings underscore that AI literacy involves more than technical competence; it requires cultivating organizational capacities for ethical judgment and relational accountability. Designing AI-enabled systems with care and intentional oversight is, therefore, central to sustaining equitable and human-centered pathways.

Limitations

This study is bounded by limitations inherent in qualitative descriptive research. The interpretive nature of qualitative analysis introduces the possibility of subjectivity in how data are understood and represented. Although structured thematic procedures and triangulation supported analytic rigor, findings reflect participant perspectives rather than direct measurement of system performance. The descriptive design was not intended to support statistical generalization, but to provide contextually grounded insight into organizational practices. The sample was limited to experienced human resource, information technology, and diversity professionals within a defined geographic region, and did not include the perspectives of job seekers affected by AI-enabled recruitment. Additionally, the rapidly evolving landscape of artificial intelligence and related ethical guidance may influence how findings are interpreted over time.

Implications for Practice

Findings from the study point to several actionable practices for organizations implementing AI-enabled decision systems. Human oversight should be intentionally designed into workflows, with clearly defined escalation points and documented authority for human reviewers to challenge or override automated outputs. Transparency and explainability must be operationalized through explicit disclosure of AI use, accessible explanations of decision stages, and structured opportunities for human clarification when outcomes are contested. Governance responsibilities should be distributed across roles rather than centralized, supported by routine review cycles and documentation practices that sustain accountability over time. Treating these elements as design requirements rather than discretionary safeguards helps prevent the normalization of unchecked automation.

For institutions situated at the education-to-work boundary, these practices represent essential learning competencies rather than solely technical requirements. Preparing practitioners to work effectively with AI-enabled systems requires developing skills in ethical judgment, critical interpretation of automated outputs, and responsible escalation when risks or ambiguities arise. Professional learning and leadership development should therefore emphasize how to question, contextualize, and govern the use of AI rather than merely how to operate tools. Embedding these competencies into organizational routines reinforces the idea that responsible AI adoption is learned through practice and reflection, positioning AI literacy as a human-centered capability grounded in accountability and care.

These findings also carry implications for the evolving role of academic institutions in preparing professionals to steward AI-enabled systems responsibly. As artificial intelligence becomes embedded in recruitment and other decision-making environments, the demand for roles integrating human resources, technology governance, compliance, and ethical oversight is likely to increase. Preparing practitioners for these interdisciplinary responsibilities requires educational programs that move beyond technical proficiency to include structured engagement with accountability, transparency, and relational ethics. Institutions situated at the education-to-work boundary are therefore positioned to cultivate stewardship competencies that support the integration of socially responsible AI across workforce pathways. Framing AI adoption as a socio-technical and ethical endeavor underscores the need for academic preparation that aligns professional formation with institutional responsibility and public trust.

Conclusion

This study demonstrates that the effective use of AI in organizational decision-making depends less on technical capability than on disciplined human stewardship. Findings indicated human oversight, governance routines, and transparency mechanisms functioned as integral components shaping how AI-enabled systems were implemented and constrained in practice. When these elements were embedded into routine processes, AI outputs were treated as provisional inputs rather than authoritative decisions. Framing AI-enabled systems as socio-technical environments underscores that ethical

responsibility is enacted through design choices, accountability structures, and human judgment rather than delegated to automation. For institutions operating at the education-to-work boundary, these findings reinforce the need to approach AI adoption as a human-centered endeavor grounded in care, responsibility, and intentional governance. Designing AI systems with humans at the center remains essential to sustaining equitable and accountable pathways.

References

Anthony, S., & Jack, S. (2009). Qualitative case study methodology in nursing research: An integrative review. *Journal of Advanced Nursing (Wiley-Blackwell)*, *65*(6), 1171–1181. https://doi.org/10.1111/j.1365-2648.2009.04998.x

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology, 9*(1), 3–26. https://doi.org/10.1037/qup0000196

Burrell, D. N., & McAndrew, I. (2023). Exploring the ethical dynamics of the use of artificial intelligence (AI) in hiring in healthcare organizations. *Bialystok Legal Studies / Bialostockie Studia Prawnicze*, *28*(4), 309–321. https://doi.org/10.2478/raft-2023-0037

Gabriel, I. (2020). Artificial intelligence, values, and alignment. *Minds & Machines, 30*(3), 411-437. https://doi.org/10.1007/s11023-020-09539-2

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds & Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Oruc, I., & Sarikaya, M. (2011). Normative stakeholder theory in relation to ethics of care. *Social Responsibility Journal, 7*(3), 381-392. https://doi.org/10.1108/17471111111154527

Parra Jounou, I., & Tronto, J. C. (2024). Care ethics in theory and practice: Joan C. Tronto in conversation with Iris Parra Jounou. *Contemporary Political Theory*, *23*(2), 269–283. https://doi.org/10.1057/s41296-024-00680-6

Sandelowski, M. (2000). Whatever happened to qualitative description? Research in *Nursing & Health, 23*(4), 334–340. https://doi.org/10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g

Zhang, X., Wang, X., Tian, F., Xu, D., & Fan, L. (2023). Anticipating the antecedents of feedback-seeking behavior in digital environments: a socio-technical system perspective. *Internet Research*, *33*(1), 388–409. https://doi.org/10.1108/INTR-09-2021-0672