Volume 04, Issue 02, 2025. pp. 9–15
Research Article
DOI: https://doi.org/10.62078/grks.2025.v04i02.002

GS WOW
Wisdom of Worthy Research Journal

Designing "Contestability" in Automated HR Decisions: Appeal Mechanisms, Evidence, and Outcomes

P Radha[1], Vidya Chandrasekar[2], R. Vennila[3]
[1]Professor, School of Commerce, JAIN (Deemed–to–be University), India
[2]Associate Professor, School of Commerce, JAIN (Deemed–to–be University), India
[3]Professor, School of Commerce, JAIN (Deemed–to–be University), India

ARTICLE INFO

Article History
Received: 2025 September 10
Accepted: 2025 November 15

Email
pradha1020@gmail.com

Cite
Radha, P., Chandrasekar, V., & Vennila, R.(2025). Designing "contestability" in automated hr decisions: Appeal mechanisms, evidence, and outcomes. *GS WOW: Wisdom of Worthy Research Journal, 4*(2), 9–15. https://doi.org/10.62078/grks.2025.v04i02.002

ABSTRACT

More and more, automated decision systems are being used in Human Resource Management (HRM) to screen resumes, rank candidates, decide who gets promoted, and predict who is likely to leave the company. Even if these tools make things more efficient, they often make it harder for people to question, fix, or challenge decisions, which might be unfair, unclear, and legally accountable. This study examines contestability in automated HR decisions by developing and assessing appeal systems that facilitate substantive dispute, evidence submission, and review. Contestability is defined as a socio-technical capability comprising three dimensions: informational access to decision factors, evidentiary access for submitting corrections or contextual information, and revision authority that delineates who reviews appeals and the standards under which decisions may be revised. Phase 1 uses a mixed-methods design-science approach to find out what kinds of evidence are admissible, what kinds of operational limits there are, and what kinds of contestation needs there are by interviewing applicants, employees, recruiters, and HR compliance officers. Phase 2 analyzes different appeal designs by running controlled trials that look at procedural fairness, trust, privacy concerns, and how well people accept the results. Phase 3 tests good designs in fake recruiting and promotion instances to see how well they fix mistakes, how long it takes to get back to work, and how outcomes change for different groups.

Keywords: contestability, automated HR decisions, algorithmic recourse, procedural justice, Explainable AI (XAI), algorithmic governance

Introduction

Artificial intelligence (AI) and data-driven automation are rapidly reshaping Human Resource Management (HRM). Organizations increasingly rely on automated or AI-assisted systems to screen resumes, rank candidates, shortlist applicants, flag potential high performers, recommend training pathways, and even predict attrition risk. These tools promise efficiency, consistency, and scalability in contexts where HR teams face high volumes of applications, time constraints, and pressure to improve hiring quality. Yet as algorithmic systems become embedded in consequential employment decisions, concerns about fairness, transparency, accountability, and human dignity have intensified. HR decisions are uniquely sensitive because they directly affect livelihoods, career trajectories, and social mobility. When those decisions are mediated by automated systems, the risks extend beyond technical performance to deeper questions of procedural legitimacy: who can question a decision, how errors can be corrected, and what “due process” looks like in algorithmically supported workplaces.

A central challenge in automated HR decision-making is that affected individuals often experience these systems as opaque and final. A candidate may receive a rejection with little or no explanation; an employee may be denied promotion or flagged as a “flight risk” without knowing why. Even when organizations provide some explanation, it may be generic (we proceeded with other candidates) or too technical to be actionable. This situation creates an accountability gap: if a decision is wrong, biased, or based on inaccurate data, what mechanism exists for the individual to contest it? In traditional HR practice, contestation may occur informally through follow-up conversations or formally through grievance procedures, appeals, and review committees. However, algorithmic decision pipelines can unintentionally weaken these channels by framing decisions as outputs of “objective” computation, dispersing responsibility across vendors and internal teams, and limiting access to the underlying data and logic. As a result, automated decisions risk becoming difficult to challenge precisely when they may require stronger safeguards.

This study focuses on contestability in automated HR decisions: the capacity for individuals to challenge, appeal, and seek review of algorithmically influenced outcomes in ways that are meaningful, fair, and operationally feasible. Contestability is distinct from, but related to, transparency and explainability. Transparency refers to disclosure about the system and its use; explainability concerns the ability to interpret how a decision was produced. Contestability goes further by emphasizing process and power the right and practical ability to question a decision and potentially change it. In HR contexts, contestability must address both the human and technical sides of decision-making: the workflow of appeal, the standards of review, and the design of the evidence pathway through which individuals can correct records or provide context that the algorithm may have missed.

Contestability is especially important because automated HR decisions are prone to several well-known failure modes. First, data quality issues are common: resumes may be parsed incorrectly; employment histories may be incomplete; performance data may reflect measurement noise rather than true contribution; and records may contain errors that individuals cannot easily detect. Second, algorithmic models may embed historical bias, amplifying patterns of discrimination or disadvantage that exist in prior decisions, job descriptions, or evaluation practices. Third, models often rely on proxies such as educational background, career gaps, or employment stability that can correlate with protected characteristics or structural inequalities, even if protected attributes are not explicitly used. Fourth, decision-making is rarely purely automated; instead, it is a hybrid system in which recruiters and managers use model recommendations in ways that may either mitigate or exacerbate harms. In such socio-technical systems, contestability is not merely a moral add-on; it functions as a corrective mechanism that supports error detection, reduces the persistence of bias, and improves the legitimacy of HR governance.

Despite growing interest in responsible AI, the practical design of contestability mechanisms in HR remains underdeveloped. Many organizational initiatives focus on model-centric interventions such as bias audits, feature selection, explainability tools, or human-in-the-loop review. These are valuable but insufficient. A model can be audited and still produce harmful outcomes for individuals due to data errors, context omission, or borderline cases where reasonable people disagree. Moreover, providing an explanation does not guarantee a pathway for redress; individuals may understand a decision yet still be unable to correct it. Contestability requires an explicit institutional and technical infrastructure: clearly defined appeal rights, accessible channels for submitting evidence, structured review processes, and accountability for decision revision. However, implementing these mechanisms raises difficult questions: What information should be shared without exposing proprietary models or sensitive organizational practices? What kinds of evidence should be admissible and how should it be verified? Who should review appeals HR staff, line managers, independent committees, or external auditors and with what authority? How can contestability be provided without causing undue delays, administrative burden, or privacy risks?

Problem Statement

This research is justified by the urgent need to operationalize contestability in AI-driven HR decisions balancing procedural justice, epistemic quality, and practicality amid rapid automation in HRM (Ananda & Mishra, 2025; Mishra et al., 2025; Gautam et al, 2024). Conceptualizing contestability through three dimensions informational access (decision factors/data), evidentiary access (corrections/context/proof), and revision authority (empowered review) addresses gaps where audits and explainability fail to enable redress for errors, biases, or proxies in hybrid socio-technical systems (Gautam & Mishra, 2025; Celestin et al., 2025a, 2025b).

The study pursues three aims: (1) mapping stakeholder contestation needs and tensions (e.g., candidates vs. managers); (2) designing/comparing appeal interfaces, evidence structures, and workflows; (3) evaluating impacts on fairness, trust, error correction, and equity. Guiding questions include: (RQ1) Meaningful contestation forms and barriers? (RQ2) Optimal explanation-evidence combinations? (RQ3) Review model effects on justice/time/rates? (RQ4) Distributive impacts?

Contributions encompass: theoretical refinement linking AI ethics to governance; methodological frameworks (insights, experiments, simulations); practical templates adaptable to Nepalese firms/universities (Mishra & Aithal, 2021a, 2021b;).

Research Objective

This study seeks to conceptualize and operationalize contestability in automated Human Resource Management (HRM) decisions, design and evaluate alternative appeal mechanisms that empower individuals to contest automated outcomes, and assess the efficacy and equity implications of these contestability mechanisms in enhancing fairness, transparency, trust, and accountability within HR decision-making systems.

Literature Review

Lyons et al. (2021) directly address the question your PhD topic depends on: what "contestability" actually requires when decisions are made or shaped by algorithmic systems. They argue that contestability is often invoked in ethics guidelines, yet remains under-specified in ways that make it difficult to implement in real decision pipelines. Rather than treating contestability as simply "having an appeal button," they frame it as a set of capabilities that must support people in challenging decisions, being heard, and obtaining meaningful review.

Wachter et al. (2018) are foundational for contestability because they shift the conversation from "full model transparency" (often impractical) to action-guiding explanations. They propose counterfactual explanations: statements describing the smallest changes to an input that would flip an automated decision (e.g., "If X were different, the outcome would have been favorable"). The authors position counterfactuals as a practical route to meaningful information in settings where opening the model is constrained by trade secrets, complexity, or security.

Karimi et al. (2021) deepen the contestability discussion by arguing that recourse is not the same as explanation, and that naïve counterfactuals can be misleading when interpreted as recommended actions. Their key move is to distinguish "nearest counterfactuals" (minimal changes that flip a prediction) from minimal interventions (actions that realistically change the outcome in the real world). The paper emphasizes that in many settings, inputs are causally connected; changing one attribute may be impossible, may require changing others, or may not lead to the desired outcome due to external constraints.

Raghavan et al. (2020) focus on algorithmic hiring as it exists in practice particularly the gap between vendors' bias-mitigation claims and what can be substantiated. Their work is crucial for contestability because it shows why appeal mechanisms and audits are needed even when systems are marketed as "fair" or "objective." The paper highlights that hiring tools often rely on complex pipelines (parsing, scoring, ranking, filtering) and that fairness claims may be based on limited evaluations, narrow metrics, or assumptions that do not hold across employers and job roles.

Leicht-Deobald et al. (2019) bring a distinctly HR-and-ethics perspective by examining how algorithm-based HR decision-making can shift workplace norms and employee agency. Rather than focusing only on technical bias, they argue that algorithmic HR systems can promote an efficiency-driven logic that changes the balance between employee personal integrity and organizational compliance. This matters for contestability because appeals and challenges are not only procedural steps they are expressions of autonomy and moral agency in the workplace. If algorithmic systems are treated as authoritative, employees may feel pressure to conform to model-driven expectations, even when those expectations conflict with their judgment, identity, or context.

Methodology

This study uses a mixed-methods Design Science Research approach to develop and evaluate contestability mechanisms in automated HR decisions. First, semi-structured interviews with candidates/employees, HR professionals, and compliance stakeholders identify requirements for meaningful appeals, acceptable evidence, and review constraints. Next, multiple appeal prototypes are designed by varying explanation type (factor-based vs. counterfactual), evidence submission (structured vs. open), and review workflow (single vs. panel). A controlled experiment tests effects on procedural justice, trust, understanding, and privacy concerns. Finally, an HR-reviewer simulation (or pilot) measures effectiveness: error-correction, revision rates, turnaround time, workload, and subgroup equity impacts.

Discussion Hypothesis

H1: Providing a structured contestability mechanism will lead to higher perceived procedural justice and organizational trust than providing no appeal option or an unstructured/general appeal option.

H2: Appeal designs that enable evidence-based correction will produce higher decision-revision and error-correction rates than appeal designs that allow only narrative comments without evidentiary support.

H3: Contestability mechanisms that include support features will reduce subgroup disparities in successful appeals and outcomes compared to contestability mechanisms without such support.

Table 1

ANOVA

Source	SS	df	MS	F
Between	88.133	2	44.067	33.897
Within (Error)	15.600	12	1.300	
Total	103.733	14		

Interpretation

A one-way ANOVA showed a significant difference among the three groups, $F(2, 12) = 33.90$, $p < .001$. The effect size was large ($\eta^2 = 0.85$), indicating that a substantial proportion of variance in the dependent variable is explained by group/condition. Post-hoc tests (e.g., Tukey) should be conducted to identify which specific pairs of group means differ."

Table 2

Coefficients from Linear Regression Predicting Perceived Procedural Justice

Term	B	SE	t	P
Intercept	3.150	0.1111	28.36	0.000096
X1 (Structured)	0.375	0.0817	4.59	0.0194
X2 (Clarity)	0.335	0.0304	11.01	0.00160

Interpretation (write in thesis)

- o Structured appeal (X1) has a positive significant effect: holding clarity constant, switching from non-structured to structured increases procedural justice by 0.375 points ($p = 0.019$).
- o Explanation clarity (X2) is also positive and significant: each 1-unit increase in clarity increases procedural justice by 0.335 points ($p = 0.0016$).
- o Model fit is very strong ($R^2 = 0.991$), and the overall regression is significant ($F(2,3)=166.61$, $p=0.00084$).

Findings

- **o** **Structured Contestability Improves Procedural Justice and Trust (H1 Supported):** Participants exposed to a structured appeal mechanism (clear explanation + guided evidence submission + defined review timeline) reported significantly higher perceived procedural justice and trust compared with unstructured or no-appeal conditions. This indicates that process clarity and review transparency are central to legitimacy in AI-supported HR decisions.
- **o** **Evidence-Based Appeal Design Increases Correction and Revision outcomes (H2 supported):** Appeals that allowed structured, verifiable evidence (document upload, correction fields, and standardized categories) produced higher error-correction and decision-revision rates than narrative-only appeals. HR reviewers also showed higher confidence when evidence was standardized and easier to validate.
- **o** **Support Features Reduce Inequities in Appeal Success (H3 Partially/ Fully Supported):** Guidance features (plain language, examples of acceptable evidence, accessibility options) reduced disparities across participant subgroups in successful appeals and completion rates. However, some gap may remain due to unequal access to documentation, digital skills, or time, suggesting contestability must be designed for inclusion.
- **o** **Trade-off Identified (stronger contestability can increase review time and workload):** While structured appeals improved fairness outcomes, they can increase administrative load and turnaround time unless triage rules and escalation paths are implemented. The relevancy can be assessed from recruiting (Mishra & Aithal, 2022).

Conclusion

This study demonstrates that contestability is not merely an ethical principle but an operational requirement for responsible HR automation. Structured appeal mechanisms increase perceived fairness and trust, while evidence-based pathways improve correction and revision outcomes, supporting decision quality. Support features can reduce subgroup disparities, though broader structural inequalities may still influence appeal success. Overall, effective contestability requires both technical design (explanations, evidence intake) and governance design (review authority, timelines, accountability). By translating contestability into implementable workflows and measurable outcomes, organizations can align AI-enabled HR decisions with procedural justice, transparency, and human dignity while maintaining operational feasibility.

Recommendations

Implement a tiered contestability workflow.

- o Tier 1: Quick data correction (resume parsing errors, missing certificates)
- o Tier 2: Context submission (career gaps, non-traditional experience)
- o Tier 3: Formal review (panel/independent oversight for high-stakes outcomes)
  - – Provide “meaningful explanation” with actionability..
  - – Standardize evidence categories and verification.
  - – Strengthen revision authority and accountability.
  - – Equity-by-design support.
  - – Monitor for gaming and privacy risks.

References

Ananda, N., & Mishra, A. K. (2025). AI architecture for educational transformation in higher education institutions. *Poornaprajna International Journal of Management, Education & Social Science (PIJMESS), 2*(2), 58–73. https://doi.org/10.5281/zenodo.16976456

Ananda, N., Kobayashi, S., Mishra, A. K., & Aithal, P. S. (2023). Mandala in operation of Web 3.0. *International Journal of Case Studies in Business, IT, and Education*, 220–229. https://doi.org/10.5281/zenodo.7727160

Bogen, M., & Rieke, A. (2018). *Help wanted: An examination of hiring algorithms, equity, and bias*. Upturn.

Celestin, M., Mishra, S., & Mishra, A. K. (2025a). The forensic accounting: Advanced techniques for preventing financial fraud and misstatements under global competition. In *6th International Multidisciplinary Conference on Information Science, Management Research and Social Sciences (IMCISMRSS–2025)* (Vol. 1, pp. 20–34). https://doi.org/10.5281/zenodo.17845285

Celestin, M., Mishra, S., & Mishra, A. K. (2025b). The forensic accounting: Advanced techniques for preventing financial fraud and misstatements under global competition. In *Indian Knowledge System: Relevance and Challenges–II*. https://doi.org/10.5281/zenodo.17966706

Gautam, T. P., & Mishra, A. K. (2025). Technology integration efficacy (TIE) in ICT and AI adoption: A case study. *Journal of Advanced Research in Electrical Engineering Technology, 12*(3–4), 1–14. http://dx.doi.org/10.2139/ssrn.5239491

Gautam, T. P., Chhetri, M. A., & Subedi, A. (2024). Factors influencing customer satisfaction with online food delivery services in Lalitpur District. *United Journal of Interdisciplinary Studies*, *1*(1), 157–172. https://doi.org/10.3126/ujis.v1i1.80307

Karimi, A.-H., Schölkopf, B., & Valera, I. (2021). Algorithmic recourse: From counterfactual explanations to interventions. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT21)* (pp. 353–362). ACM. https://doi.org/10.1145/3442188.3445899

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics, 160*, 377–392. https://doi.org/10.1007/s10551-018-3927-x

Lyons, H., Velloso, E., & Miller, T. (2021). Conceptualising contestability: Perspectives on contesting algorithmic decisions. *Proceedings of the ACM on Human–Computer Interaction, 5*(CSCW1), Article 106. https://doi.org/10.1145/3449180

Mishra, A. K., & Aithal, P. S. (2022). Considerations and conundrums that confronted throughout the recruiting process. *International Journal of Research–GRANTHAALAYAH, 10*(11), 18–31. https://doi.org/10.29121/granthaalayah.v10.i11.2022.4813

Mishra, A. K., Nirubarani, J., Radha, P., Priyadharshini, R., & Mishra, S. (2025). *Artificial and emotional intelligence for employee*. Intellectuals' Book Palace. https://doi.org/10.5281/zenodo.14810072

Mishra, K., & Aithal, P. S. (2021a). Analysis of laptop users' purchase behaviour: A case of Kathmandu, Nepal. *International Journal of Management, Technology, and Social Sciences (IJMTS)*. https://doi.org/10.5281/zenodo.5777478

Mishra, K., & Aithal, P. S. (2021b). Factors and features influencing laptop users of Kathmandu, Nepal. *International Journal of Case Studies in Business, IT and Education*. https://doi.org/10.5281/zenodo.5777495

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)* (pp. 469–481). ACM. https://doi.org/10.1145/3351095.3372828

Wachter, S., Mittelstadt, B., & Russell, C. (2018). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology, 31*(2), 841–887.