A Study on AI-Driven Recruitment and Its Ability to Identify Candidates Beyond Resume-Based Screening

Dr. Shashikala S.R.
Associate Professor, Department of Studies in Business Administration, Pooja Bhagavat Memorial Mahajana P.G. Centre, Mysuru, Karnataka, India

Abstract—**Organizations increasingly rely on artificial intelligence (AI) to screen, rank and interview candidates, and a growing share of practitioner commentary, including HR trend outlooks for 2026, frames the central question as whether automated systems can recognize genuine ability rather than merely well-packaged credentials. This article examines whether AI-supported, skills-based hiring can surface talent that résumé centered processes overlook, and under what conditions it does so without reproducing or amplifying bias.**

Using a structured narrative review of peer-reviewed and policy literature, supplemented by four analytical cases (a large-scale résumé-screening failure, algorithmic video interviewing, the regulatory response in New York City and the European Union, and skills-inference platforms), the study compares conventional and AI-enabled hiring on four dimensions: signal quality, fairness, transparency and accountability. The analysis suggests that AI broadens the talent pool chiefly when it is designed to evaluate demonstrated skills rather than to imitate past hiring decisions, and that its fairness benefits depend on data governance, auditing and meaningful human oversight.

Where models are trained on historical outcomes, they tend to encode the very patterns they were expected to correct. The article proposes a three-layer framework, comprising skills evidence, algorithmic support and accountable human judgment, and derives implications for HR leaders, regulators and educators. The study contributes a conceptual integration of the AI-recruitment, skills-based hiring and decision-bias literatures and sets out propositions for subsequent empirical testing.

Index Terms—**AI recruitment; skills-based hiring; algorithmic bias; hiring discrimination; human judgment; talent acquisition; human–AI collaboration; HR analytics**

I. INTRODUCTION

The résumé has served as the default gateway to employment for decades, yet it is an imperfect instrument. It records where a person has studied, who has employed them and how they describe themselves, all of which are proxies for ability rather than direct evidence of it. Such proxies disadvantage people whose capabilities were built through non-linear careers, informal learning, caregiving breaks, military service or work in under-recognized sectors. Fuller et al. (2021) describe this population as "hidden workers": motivated candidates filtered out by rigid screening rules, often for lacking a degree or for having an employment gap that says little about job performance. AI promises a remedy. Machine-learning tools can parse unstructured text, infer skills from project descriptions, match candidates to roles at scale and administer standardized assessments. Vendors argue that such tools reduce reliance on pedigree and individual recruiter whim.

Critics respond that algorithms trained on past decisions inherit past prejudice, that opaque scoring undermines due process, and that efficiency gains may be bought at the cost of fairness (Barocas & Selbst, 2016; Raghavan et al., 2020). The debate has moved from the margins of HR technology into mainstream management agendas, and practitioner outlooks for 2026 identify the ability of automated hiring to find real capability, rather than keyword matches, as a pressing organizational challenge. This tension motivates the present study. The question is not whether AI is already used in hiring, since it plainly is, but whether it can identify talent that traditional processes miss, and whether human judgment remains necessary to make those identifications legitimate and trustworthy.

1.1 Research objectives

The study pursues five objectives:

1. To clarify how traditional, résumé-centered hiring and AI-enabled hiring differ in the signals they use to infer ability.
2. To assess the evidence on whether skills-based, AI-supported approaches widen access to overlooked candidates.
3. To analyze the mechanisms through which AI can reduce or amplify hiring bias.
4. To define the role of human judgment in AI-supported selection.
5. To propose a framework and practical guidelines for responsible adoption.

1.2 Research questions

- RQ1: Under what conditions can AI-enabled hiring identify capable candidates that résumé-based screening overlooks?
- RQ2: How do design choices (training data, features, validation, auditing) shape bias outcomes?
- RQ3: What forms of human oversight add genuine value rather than symbolic assurance?

1.3 Contribution and structure

The article integrates three literatures that are usually discussed separately: algorithmic fairness, personnel-selection science and human–AI collaboration. Section 2 reviews the literature. Section 3 describes the methodology. Section 4 presents the analysis, cases and findings. Section 5 sets out recommendations and implications, and Section 6 concludes.

II. LITERATURE REVIEW

2.1 The limits of traditional hiring

Personnel-selection research has long shown that predictive validity varies sharply across methods. Schmidt and Hunter (1998) reported that work-sample tests and structured interviews predict performance far better than unstructured interviews or years of experience, a conclusion refined but broadly supported by Sackett et al. (2022), who found that structured interviews rank especially high once range-restriction corrections are revised. Despite this, many organizations still lean on informal impressions and credential screens. Kahneman, Sibony and Sunstein (2021) show that such judgments carry substantial "noise": different evaluators reach different conclusions about the same candidate, and the same evaluator varies across occasions.

Bias compounds noise. Field experiments in which identical résumés carry different names have repeatedly shown unequal callback rates (Bertrand & Mullainathan, 2004). Human gatekeepers therefore cannot be treated as a neutral benchmark against which algorithms are judged. The relevant comparison is between two imperfect systems.

2.2 AI in recruitment

Tambe, Cappelli and Yakubovich (2019) observe that HR is a difficult domain for machine learning because outcomes such as "good performance" are ambiguous, datasets are small relative to other fields, and decisions raise fairness and legitimacy concerns. Black and van Esch (2020) distinguish AI applications across the hiring funnel, from sourcing and chatbots to résumé parsing, gamified assessment and video analysis, and warn that managers often adopt tools without understanding what is being measured. Köchling and Wehner (2020), in a systematic review, conclude that research on discrimination by algorithmic hiring is fragmented and that many deployed tools are not accompanied by transparent validation.

Evidence on benefits is mixed but suggestive. Li, Raymond and Bergman (2020) show that a hiring algorithm designed to value exploration, that is, to give weight to candidates with uncertain but potentially high upside, can both increase the share of hires from underrepresented groups and raise the quality of those selected, compared with a model that simply exploits historical patterns. Their finding implies that the objective function matters as much as the technology.

2.3 Skills-based hiring

Skills-based hiring replaces degrees and job-title histories with direct evidence of competence, such as assessments, work samples, portfolios and verified credentials. It rests on the personnel-selection finding that demonstrated performance outperforms proxies.

Fuller et al. (2021) report that employers who remove unnecessary degree requirements can reach talent pools that are larger and more diverse, although many firms continue to demand credentials in practice, creating a gap between stated policy and screening behavior. AI is a natural complement because skills

inference at scale is laborious for people but tractable for models, which can map a candidate's described work to a skills taxonomy even when job titles differ.
Skills taxonomies bring their own risks. Skill labels may be defined by incumbents' profiles, may favor skills that are easy to observe over those that are not, and may drift as jobs change. Inferred skills are probabilistic estimates, not verified facts.

2.4 Algorithmic bias and fairness

Barocas and Selbst (2016) explain how data mining can produce disparate impact even without discriminatory intent, through biased labels, unrepresentative samples and proxy variables. Mehrabi et al. (2021) catalogue the sources of bias across the data-to-deployment pipeline. Raghavan et al. (2020) examine vendor claims and find that practices for assessing bias are inconsistent and rarely subject to independent scrutiny. Sánchez-Monedero, Dencik and Edwards (2020) add that legal and technical definitions of non-discrimination do not align across jurisdictions, so a tool "fair" under one standard may fail another. Bogen and Rieke (2018) argue that bias can enter at each stage of the hiring funnel, including who sees a job advertisement, so auditing only the final ranking is insufficient.
Importantly, the literature does not support a simple verdict that algorithms are biased and humans are not, or the reverse. Kleinberg et al. (2018) show in the context of judicial decisions that algorithmic predictions can outperform human judgments and, if carefully built, reduce disparities, while stressing that results depend on the target outcome and data quality.

2.5 Human judgment and human–AI collaboration

Meehl's (1954) classic work established that statistical rules often outperform clinical intuition, yet people frequently distrust algorithms after seeing them err, a tendency termed algorithm aversion (Dietvorst et al., 2015). The opposite risk is automation bias, where reviewers defer to a score. Raisch and Krakowski (2021) frame this as an automation augmentation paradox: organizations that use AI to replace judgment may lose the contextual understanding they need, whereas those that use it to augment judgment must manage the tensions that arise. Langer, König and Papathanasiou (2019) find that candidates perceive highly automated interviews less favorably, a reminder that legitimacy in the eyes of applicants is part of effective hiring.

2.6 Research gap

Existing work tends to assess either the predictive performance of algorithms, their fairness, or candidate reactions in isolation. Few studies connect skills-based hiring design, bias mechanisms and the structure of human oversight in a single analytical framework. This article addresses that gap.

III. RESEARCH METHODOLOGY

3.1 Research design

The study adopts an interpretive, conceptual design combining a structured narrative literature review with comparative case analysis. This approach suits an emerging, fast-moving topic in which controlled empirical evidence is limited and the aim is theory-building and proposition development rather than statistical generalization.

3.2 Sources and search strategy

Peer-reviewed articles, working papers from recognized research institutions, regulatory documents and reputable investigative reporting were identified through major databases (e.g., Scopus, Web of Science, Google Scholar, SSRN) using combinations of the terms algorithmic hiring, AI recruitment, skills-based hiring, hiring discrimination, algorithmic fairness, human oversight and algorithm aversion. Sources were included if they (a) addressed selection or recruitment decisions, (b) offered empirical evidence, a conceptual contribution or regulatory analysis, and (c) were written in English. Practitioner reports were used to contextualize trends but not as primary evidence of effectiveness.

3.3 Analytical framework

Sources and cases were coded against four dimensions derived from the selection and fairness literatures:

- Signal quality: what evidence of ability is used, and is it validated against job performance?
- Fairness: are outcomes monitored for disparate impact, and at which funnel stages?
- Transparency: can candidates and decision-makers understand why a decision was reached?
- Accountability: who is responsible for the decision, and can it be contested?

Cases were chosen purposively because they are well documented in public sources and illustrate contrasting design philosophies. Findings are therefore analytic generalizations, not claims about prevalence.

3.4 Limitations

The study relies on secondary evidence, and publicly reported cases may over-represent failures. Vendor performance claims are difficult to verify independently. The framework is conceptual and requires empirical testing.

IV. ANALYSIS, EXAMPLES AND FINDINGS

4.1 Case A: Learning from past decisions can replicate past bias

In 2018, Reuters reported that Amazon had abandoned an experimental résumé-ranking system after discovering that it penalized résumés associated with women, because it had been trained on a decade of submissions dominated by male applicants (Dastin, 2018). Engineers reportedly tried to neutralize specific terms, but could not guarantee that other proxies would not emerge. The lesson is structural: a model optimized to reproduce historical hiring outcomes treats past preferences, including discriminatory ones, as the definition of quality. It cannot discover overlooked talent because overlooked talent is, by construction, absent from the positive examples. This case illustrates the risk Barocas and Selbst (2016) describe as biased labels combined with proxy variables.

4.2 Case B: Algorithmic interviews and the validity question

Automated video-interview tools score recorded responses using language, and in some earlier products facial and vocal features. Concerns about scientific validity and opacity prompted public scrutiny, and at least one major vendor discontinued facial analysis in its assessments. Langer et al. (2019) show candidate reactions are less favorable to highly automated formats, raising questions about both employer brand and the willingness of strong candidates to participate. The analytic point is that a tool may measure something consistently without measuring what predicts job success. Reliability is not validity, and without a documented link to performance the tool may merely substitute a new proxy for the old ones.

4.3 Case C: Regulation as a design constraint

Regulators have begun to translate fairness concerns into obligations. New York City's Local Law 144 requires employers using automated employment decision tools to commission independent bias audits and to notify candidates. The European Union's AI Act places AI systems used in recruitment and worker management in its high-risk category, which brings requirements for risk management, data governance, documentation and human oversight. Litigation in the United States, such as Mobley v. Workday, has also tested whether vendors can be held accountable as agents in hiring decisions. These developments show that accountability is shifting from a voluntary ethic to a compliance requirement, and that organizations cannot assume responsibility lies with the software supplier.

4.4 Case D: Skills inference and exploration-oriented design

Skills-inference platforms map candidates' experience, assessments and credentials to a skills taxonomy and match them to roles, including internal mobility. Used well, they enable a delivery driver with logistics-planning experience or a caregiver who managed budgets and schedules to be recognized for transferable capabilities that no job title announces. The exploration-based model studied by Li et al. (2020) provides evidence that algorithmic design can deliberately broaden the pool. However, benefits depend on three conditions: skills data must be validated rather than self-asserted; taxonomies must be reviewed for embedded bias; and results must be tested against actual job performance.

4.5 Comparative synthesis

Dimension	Traditional résumé-centered hiring	AI trained on past hires	AI-supported skills-based hiring
Signal quality	Proxies (pedigree, titles, tenure)	Proxies learned from history	Demonstrated skills and work samples
Fairness risk	Individual bias, noise	Systematic, scalable bias	Reduced if audited; taxonomy bias remains

Transparency	Low but explainable by people	Often opaque	Higher if skills evidence is shown
Accountability	Hiring manager	Diffuse (vendor vs. employer)	Explicit human owner required
Capacity to find overlooked talent	Low	Very low	Moderate to high

4.6 Findings

Finding 1: The objective determines the outcome. AI discovers overlooked talent only when it is aimed at demonstrated capability. Systems trained to imitate previous hires inherit prior exclusions and may harden them at scale (Case A; Li et al., 2020).

Finding 2: Skills evidence is the active ingredient. The strongest gains arise from replacing proxies with work samples, structured assessments and verified skills, which the selection literature already supports. AI contributes mainly by making this evidence cheaper to collect and compare, not by supplying new insight about people (Schmidt & Hunter, 1998; Sackett et al., 2022).

Finding 3: Bias is a pipeline problem. Discrimination can enter through job advertising, sourcing, parsing, scoring and final choice. Audits limited to the last step miss upstream exclusion (Bogen & Rieke, 2018).

Finding 4: Measurable is not the same as valid. Consistent scores from an interview or game tool offer no assurance of job relevance unless validated (Case B).

Finding 5: Human judgment matters most at specific points. Humans add the most value when defining what success means, validating the tool, reviewing borderline and contested cases, and hearing appeals. They add little when asked to rubber-stamp scores, where automation bias undermines oversight (Raisch & Krakowski, 2021; Dietvorst et al., 2015).

Finding 6: Legitimacy is part of effectiveness. Candidate trust, legal compliance and brand reputation depend on explainability and recourse, not only on predictive accuracy (Langer et al., 2019; Case C).

4.7 Propositions for future research

- P1: AI systems trained on skills evidence will reach a more diverse qualified applicant pool than those trained on historical hiring outcomes.
- P2: The effect of AI on hiring fairness is moderated by the frequency and independence of bias audits.
- P3: Human review improves selection quality when reviewers have decision authority and structured criteria, but not when they only confirm algorithmic ranks.
- P4: Transparent explanations of skill-based decisions increase candidate perceptions of procedural justice.

V. RECOMMENDATIONS AND IMPLICATIONS

5.1 A three-layer framework

The analysis suggests a framework in which each layer has a distinct function. The evidence layer collects validated indicators of capability, including work samples, structured interviews and verified credentials. The algorithmic layer organizes this evidence, matches skills to requirements and flags patterns for review, with models validated against job outcomes and tested for disparate impact. The judgment layer keeps accountable people in charge of defining success criteria, interpreting edge cases, deciding and handling appeals.

5.2 Recommendations for organizations

1. Specify the job before the model. Define skills and performance criteria through job analysis, and drop credential requirements that lack a demonstrated link to performance.
2. Avoid training solely on past hires. Where historical data must be used, test for proxy variables and consider exploration-oriented designs.
3. Audit across the funnel. Commission independent audits covering advertising, screening and selection, and repeat them after model or data changes.
4. Require vendor evidence. Ask for validation studies, documentation of training data and the right to test the tool in your own context.
5. Design oversight that can say no. Give reviewers authority, time and structured criteria, and track how often humans override the system and with what results.
6. Provide transparency and recourse. Inform candidates when automated tools are used, explain the evidence evaluated and offer a human review route.

7. Build capability. Train recruiters and managers in basic data literacy so they can question model outputs.

5.3 Implications for policy

Regulators should favor requirements for independent auditing, documentation and contestability over bans on particular technologies, and should work toward harmonized definitions of discrimination so that organizations operating across borders can comply meaningfully. Public bodies might also encourage skills-based hiring in their own practices, which would signal demand for credible skills evidence.

5.4 Implications for education and research

Business schools and professional bodies can embed AI governance and selection science in HR and MBA curricula. Researchers should test the propositions above using field experiments, audit studies and longitudinal comparisons of hire performance, and should examine contexts outside North America and Europe, where labor-market structures and legal standards differ.

VI. CONCLUSION

AI can help organizations look beyond the résumé, but it does not do so automatically. The technology is a means of collecting and comparing evidence at scale; whether that evidence reveals hidden talent or reproduces old exclusions depends on what the system is told to value, what data it learns from and who is accountable for its use. The analysis indicates that skills-based design, ongoing auditing and genuinely empowered human judgment are mutually reinforcing, and that removing any one weakens the others. Organizations should therefore treat AI as an instrument for improving the quality of evidence used in hiring rather than as a substitute for managerial responsibility. Future empirical work should test whether the proposed framework improves both the diversity and the performance of hires.

REFERENCES

[1]. S. Barocas and A. D. Selbst, "Big data's disparate impact," *California Law Review*, vol. 104, no. 3, pp. 671–732, 2016.

[2]. M. Bertrand and S. Mullainathan, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American Economic Review*, vol. 94, no. 4, pp. 991–1013, 2004.

[3]. J. S. Black and P. van Esch, "AI-enabled recruiting: What is it and how should a manager use it?" *Business Horizons*, vol. 63, no. 2, pp. 215–226, 2020.

[4]. M. Bogen and A. Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*. Upturn, 2018.

[5]. J. Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters*, Oct. 10, 2018.

[6]. B. J. Dietvorst, J. P. Simmons, and C. Massey, "Algorithm aversion: People erroneously avoid algorithms after seeing them err," *Journal of Experimental Psychology: General*, vol. 144, no. 1, pp. 114–126, 2015.

[7]. J. B. Fuller, M. Raman, E. Sage-Gavin, and K. Hines, *Hidden Workers: Untapped Talent*. Harvard Business School and Accenture, 2021.

[8]. D. Kahneman, O. Sibony, and C. R. Sunstein, *Noise: A Flaw in Human Judgment*. Little, Brown Spark, 2021.

[9]. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, "Human decisions and machine predictions," *The Quarterly Journal of Economics*, vol. 133, no. 1, pp. 237–293, 2018.

[10]. A. Köchling and M. C. Wehner, "Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development," *Business Research*, vol. 13, no. 3, pp. 795–848, 2020.

[11]. M. Langer, C. J. König, and M. Papathanasiou, "Highly automated job interviews: Acceptance under the influence of stakes," *International Journal of Selection and Assessment*, vol. 27, no. 3, pp. 217–234, 2019.

[12]. D. Li, L. R. Raymond, and P. Bergman, "Hiring as exploration," NBER Working Paper No. 27736, National Bureau of Economic Research, 2020.

[13]. P. E. Meehl, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. University of Minnesota Press, 1954.

[14]. N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," *ACM Computing Surveys*, vol. 54, no. 6, Art. no. 115, 2021.
[15]. M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating bias in algorithmic hiring: Evaluating claims and practices," in *Proc. 2020 Conf. on Fairness, Accountability, and Transparency*, 2020, pp. 469–481.
[16]. S. Raisch and S. Krakowski, "Artificial intelligence and management: The automation–augmentation paradox," *Academy of Management Review*, vol. 46, no. 1, pp. 192–210, 2021.
[17]. P. R. Sackett, C. Zhang, C. M. Berry, and F. Lievens, "Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range," *Journal of Applied Psychology*, vol. 107, no. 11, pp. 2040–2068, 2022.
[18]. J. Sánchez-Monedero, L. Dencik, and L. Edwards, "What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems," in *Proc. 2020 Conf. on Fairness, Accountability, and Transparency*, 2020, pp. 458–468.
[19]. F. L. Schmidt and J. E. Hunter, "The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings," *Psychological Bulletin*, vol. 124, no. 2, pp. 262–274, 1998.
[20]. P. Tambe, P. Cappelli, and V. Yakubovich, "Artificial intelligence in human resources management: Challenges and a path forward," *California Management Review*, vol. 61, no. 4, pp. 15–42, 2019.
[21]. Society for Human Resource Management, 2026.