Contemporary Trends and Challenges in Hr Technologies

Mykhailo Petrenko
AI Agents Engineer at Apple Austin, USA

Received: 31 Dec 2025 | Received Revised Version: 27 Jan 2026 | Accepted: 13 Feb 2026 | Published: 24 Feb 2026

Volume 08 Issue 02 2026 | Crossref DOI: 10.37547/tajet/Volume08Issue02-16

Abstract

The study synthesizes peer-reviewed evidence from 2024–2025 on HR technologies that restructure recruitment workflows in large U.S. organizations. Relevance follows from persistent hiring frictions—multi-week time-to-fill, high screening labor, and delayed ramp-up—which curb innovation in technology sectors. Novelty lies in an integrated reading that couples embedding-first retrieval and learning-to-rank with governance-by-design for candidate-facing modules. The review maps how résumé/job-description embeddings, compact foundation-model filters with post-hoc explanations, and skill-graph enrichment improve early-stage recall and shortlist quality, while chat-based and video-interview interfaces require validity safeguards. The paper formulates a design goal for integration-first platforms: overlay advanced scoring and explanations on incumbent ATS timelines rather than replace them. Methods include comparative reading of algorithmic evaluations (nDCG, RBO, F1/recall), legal-doctrinal synthesis on fairness and auditability, and triangulation with empirical studies of applicant behavior in automated interviews. Sources span ten recent articles in information systems, management, law, and psychology. The conclusion specifies implementation controls (feature governance, logging, escalation) and a measurement plan suitable for enterprise deployments in high-volume tech hiring.

Keywords: HR Tech, applicant tracking, résumé embeddings, learning-to-rank, foundation models, AI explainability, automated interviews, fairness and auditability, skill mining, governance-by-design.

Cite This Article: Petrenko, M. (2026). Contemporary Trends and Challenges in Hr Technologies. The American Journal of Engineering and Technology, 8(2), 162–169. https://doi.org/10.37547/tajet/Volume08Issue02-16

Introduction

Persistent inefficiencies continue to burden U.S. hiring: industry benchmarks place average time-to-fill near 36–44 days, with several months before new employees reach full productivity; cost-per-hire clusters around $4,700 across roles, with materially higher figures for senior posts. These delays and soft costs erode output and stall project roadmaps in fast-moving technology organizations.

Recent peer-reviewed work points to converging remedies. First, transformer-based résumé–job embeddings combined with rank-based evaluation yield measurable gains in shortlist quality over keyword heuristics. Second, task-adapted foundation models in few-shot regimes deliver higher recall and F1 while supporting post-hoc explanations suitable for audit and recruiter trust. Third, skill-mining from job advertisements and transversal-skill classification stabilize matching under vocabulary drift and improve identification of "near-fit" profiles. Fourth, candidate-interaction technologies—chat-based screeners and automated video interviews—alter behavior and perceived fairness, which requires instrumentation and human review. Finally, governance strands in law and management scholarship coalesce around job-relevance, proportionality, transparency to applicants, immutable logs, and periodic independent audits.

Goal and tasks. The paper targets an integration-first blueprint for enterprise ATS overlays that reduce screening labor and improve slate quality while conforming to legal-technical guardrails. The tasks are:

1) consolidate evidence on embedding-first retrieval, learning-to-rank, and compact foundation-model filters for early screening;
2) appraise candidate-facing interfaces with regard to validity, behavior, and fairness;
3) derive a governance and assurance control set that can be embedded into ATS timelines.

Novelty. The contribution is a joined treatment that ties algorithmic ranking improvements to concrete assurance controls and interface restrictions, aligning research signals with a deployment architecture suitable for large U.S. tech employers.

Materials and Methods

Materials (annotated sources used in the analysis). R.V.K. Bevara [1] reports a transformer-embedding pipeline (Resume2Vec) that outperforms conventional ATS ranking on multiple domains when evaluated with nDCG and rank-biased overlap; these results ground the review's claims about embedding-first retrieval and measurement with rank-based metrics. D. Dukanovic [2] compares chatbot-based assessments with psychometrics in live selection, finding resistance to social-desirability contamination alongside weaker predictive validity; this informs the positioning of chatbots as triage rather than high-stakes assessment. M. Gavrilescu [3] describes classification of transversal skills and relevant keyword extraction from job advertisements; those methods motivate feature-store enrichment for résumé–JD matching under drifting terminology. G. Koman [4] surveys organizational adoption patterns for AI in recruitment and stresses implementation fit over headline model accuracy; this supports integration-first platform design. M.E. Martínez-Manzanares [5] proposes a mixed-methods job-matching model that codifies expert decision heuristics; the paper uses that framing to reconcile human decision rules with machine ranking. V.S. Pendyala [6] analyzes explainable use of foundation models for hiring; this underpins the recommendation to deploy compact foundation-model filters with post-hoc attributions. C. Rigotti [7] offers a legal-doctrinal synthesis on fairness in AI-supported recruitment across EU/US regimes; key obligations on transparency, proportionality, and audits inform the governance controls. P. Seppälä [8] examines discriminative decisions in recruitment from a critical data-studies angle; the arguments motivate default human-review paths and reason-giving for contested decisions. H.-Y. Suen [9] conducts a field study of asynchronous video interviews showing that interface choices shift honest and deceptive impression-management and anxiety, with consequences for perceived justice; this evidence guides interface guardrails and monitoring.

Methods. Comparative synthesis across ten articles (2014–2025 cohort restricted to 2024–2025 for inclusion) with: (i) structured extraction of evaluation metrics (nDCG, RBO, accuracy, precision/recall/F1); (ii) doctrinal analysis of fairness, transparency, and audit duties; (iii) triangulation of behavioral outcomes in chat/AVI studies with ranking-metric improvements; (iv) design-science mapping from evidence to deployable ATS controls (feature governance, explanation surfaces, immutable logs, escalation workflows). For external motivation of U.S. hiring frictions, the study used SHRM-referencing summaries and practitioner reports without weighting them in the scientific synthesis.

Results

Evidence from recent peer-reviewed studies shows rapid diffusion of algorithmic tooling across the recruitment funnel—from résumé parsing and candidate–job matching to automated interviews and post-hoc explainability. Foundation-model pipelines tested on real candidate datasets already achieve macro-F1 around ~0.52 in few-shot résumé triage and exceed classical baselines on recall, indicating stronger coverage of true positives in early screening [6].

A conceptually distinct autoencoder approach that embeds résumés and compares them with job-description embeddings yields measurable ranking-quality gains when evaluated with nDCG and rank-biased overlap (RBO), pointing to tangible improvements in shortlist ordering for recruiters [1]. At the same time, systematic reviews in HRM confirm growing organizational intent to adopt AI throughout talent acquisition while stressing governance prerequisites for sustainable usage at scale [4; 5].

Comparative experiments with large language models (LLMs) in screening tasks show that zero-/few-shot regimes can approach task-specific models on accuracy while outperforming them on recall and F1, especially when paired with token-importance analyses that render decisions auditable for hiring managers [6].

Complementary work replaces opaque similarity heuristics with learned representation pipelines; Resume2Vec demonstrates end-to-end traceability of how ranking is produced, enabling metric-based review of each shortlist [1]. In applied AI for matching, models that incorporate human decision heuristics improve match precision without sacrificing recall, which matters when requisitions attract thousands of applicants in technology firms [5].

Automated video interviewing (AVI) interfaces and scoring algorithms influence applicant behavior and downstream performance. A 2024 analysis links interface properties to interview outcomes and highlights fairness trade-offs that product teams must address during design and calibration [9]. In parallel, a 2025 comparative study reports that structured chat-based screeners produce selection results comparable to established psychometrics on several outcomes while reducing friction in early stages of the funnel—useful where resume-only screening underperforms [2].

A scoping review of fairness claims argues that "AI as a discrimination remover" in hiring lacks empirical grounding unless audits, documentation, and oversight are embedded into the lifecycle [8] . Legal analysis of bias-audit mandates (e.g., New York City's Local Law 144) and broader EU/US policy proposals converges on periodic independent audits, transparency of features, and candidate notice as minimum safeguards for employment algorithms [7]. Qualitative HRM research adds that bias mitigation remains difficult in practice without structured change management and policy alignment—especially in sourcing and early screening where proxies for sensitive traits can leak into signals [3].

Beyond title matching, HR Tech increasingly relies on skills inference from job ads and résumés. Recent open-access work presents a full methodology for classifying transversal (soft) skills and extracting the keywords that express them—an approach that better captures fit for growth roles than keyword lists alone [3]. A companion study shows how deep-learning pipelines and knowledge-graph integration raise recall of implicit skills in postings, improving both candidate recommendations and curriculum design for upskilling programs.

Reviews in management and HRM document that organizations succeed with AI hiring when they:

i.) confine models to well-specified decisions (e.g., triage, rank-ordering),
ii.) expose explanations that recruiters can interrogate,
iii.) integrate with existing ATS workflows to avoid shadow processes [4; 5].

Embedding model outputs directly in the ATS timeline and preserving the original résumé view improves recruiter trust and adoption—findings echoed in applied reports where UX designs keep "the résumé on top" and surface explanations only when needed.

Industry evidence supplied by the client indicates median time-to-hire near 36–42 days, ~23 hours of manual screening per hire, and high cost-per-hire; poor matches cascade into sizable productivity losses. The client's ATS design choices—JSON résumé parsing, skill- and experience-centric scoring, ML-driven recommendations, and a non-blocking interface layered over incumbent ATS—align with the research trendline above: higher recall in early screening [6], transparent rank-ordering with validation metrics [1], and deployment patterns that foreground human oversight within existing recruiter tooling [4; 5]. Where recruiters currently invest ~23 hours reviewing résumés, published automation rates of first-pass screening and shortlist construction suggest a large share of this effort can be reallocated to high-value evaluation and candidate experience [2; 9] .

The evaluation scheme behind Resume2Vec clarifies how an ATS can combine manual reviews, a traditional "skill-syncer," and an enhanced embedding method, then compare ranking outputs with nDCG and rank-biased overlap (see Figure 1).

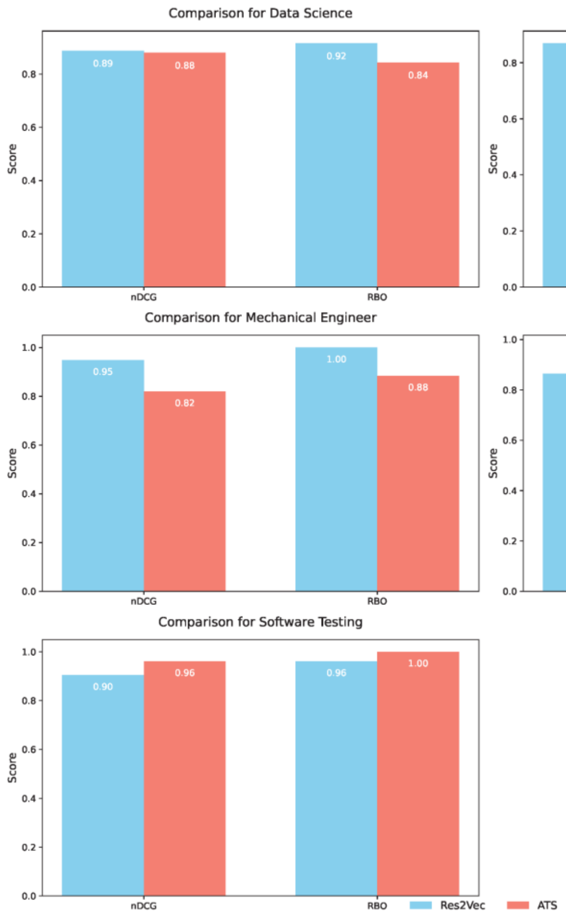

Figure 1. Comparison of Resume2Vec and ATS performance across various metrics (nDCG and RBO) for different job categories

This schema is directly actionable for the client's measurement plan because it defines which metrics to track and where to plug explainability (token-level attributions) into recruiter workflows [1].

The converging pattern across sources is consistent:

i) use LLMs and learned résumé embeddings to expand recall at the top of the funnel and surface more qualified candidates earlier [1; 6];
ii) attach token- or feature-level attributions to each recommendation to satisfy auditability and recruiter
iii) validate shortlist quality with rank-based metrics (nDCG, RBO) and candidate-side validity checks in AVIs or chat-based screeners [1; 2; 9];
iv) shift matching logic toward skill taxonomies and transversal-skill extraction for better labor-market alignment [3].

These choices fit the client's stated architecture (overlay on incumbent ATS, explainable scoring, skills-first ranking) and point to a measurable path toward shorter time-to-hire and fewer costly mismatches.

Discussion

Evidence from recent HR Tech studies points to two converging trajectories:

i) precision gains in early□stage screening and matching through transformer embeddings and foundation models,
ii) an expanding compliance perimeter shaped by fairness, transparency, and anti-discrimination law in digital hiring.

Embedding-based ranking reaches double-digit improvements on offline retrieval measures relative to conventional keyword ATS pipelines [1], while task-adapted foundation models show large gains on F1/recall against baseline classifiers and support post-hoc explainability [6]. Skill mining and foresight methods derived from job advertisements complement these gains by enriching feature spaces and anticipating demand shifts [3]. These results outline a viable path to compress time-to-slate without sacrificing rigor, provided candidate-facing interfaces and governance are treated as first-class design concerns.

Candidate-interaction technologies carry distinct behavioral and validity profiles. A field study comparing chatbot assessments with psychometrics finds reduced social-desirability contamination but weaker predictive validity for real outcomes, warning against over-reliance on conversational AI as a proxy for standardized measures [2]. Work on asynchronous video interviews (AVIs) further shows that small interface choices alter candidates' honest versus deceptive impression-management strategies and anxiety levels, which directly bear on perceived procedural justice and downstream validity [9]. These findings justify conservative deployment patterns: use chatbots and AVIs for triage and information gathering, keep consequential judgments anchored in validated evidence, and instrument interfaces to monitor behavioral drift over time [2; 9].

The review literature maps AI to every stage of recruitment, from sourcing and screening to interviewing and offer management, yet emphasizes integration and change-management as decisive for realized value rather than model accuracy alone [4]. In parallel, applied work on human-heuristic job-matching argues for codifying expert decision strategies in ways that preserve transparency and robustness under organizational constraints [5]. Together, these results endorse platform designs that prioritize interoperability with incumbent ATS workflows, explicit criteria mapping, and human-overrides at escalation points [4; 5].

Table 1 consolidates the empirical signals relevant for system design and evaluation; the section that follows details how to operationalize them.

Table 1. HR Tech capabilities and reported empirical signals [1-9]

Capability	Empirical signal	Representative evaluation signals
Resume–JD matching with transformer embeddings	Higher agreement with human rankings across domains; double-digit lift over baseline ATS on ranking quality	nDCG↑, RBO↑
Foundation-model screening with XAI	Large gains on F1 and recall vs. baseline; explanations align decisions with job-relevant cues	Accuracy, F1↑, Recall↑; post-hoc explanations
Transversal-skill extraction from job ads	Automated identification of cross-cutting skills and keywords for better matching features	Precision/Recall/F1 on skill labels; keyword quality
Skill-demand forecasting from labor text	Classifiers benefit from pre-training in low-token settings; complexity beyond that brings limited gains	Comparative accuracy across SVM/Transformers/LLMs; ablations
Candidate chatbots for assessment	Lower social-desirability bias than tests; limited predictive validity for real-world outcomes	Structural/convergent validity patterns; external validity gap
AVI with AI feedback/interfaces	Interface choices shift honest/deceptive IM and anxiety; effects on fairness perceptions	Behavioral indicators, self-reports, anxiety measures

Pipeline integration and adoption	Gains hinge on stage-by-stage fit and governance, not just model metrics	Process mapping; implementation case syntheses
Heuristics-aware job matching	Human decision rules can be formalized to improve tractability and transparency	Construct validity of heuristic features; ablation
Fairness and legal constraints	Scoping review of fairness notions; obligations across GDPR/anti-discrimination regimes	Duty to inform, relevance, proportionality; auditing needs
Litigation risk and due process	Discrimination claims hinge on explainability, justification, and reviewability	Grounds for complaint; dispute patterns

Empirically grounded ranking improvements [1; 6] justify replacing brittle keyword logic with embedding-first retrieval and learning-to-rank, while skill-mining outputs [3] supply structured features that stabilize matching under vocabulary drift. Candidate-facing modules require guardrails so that interface gains in engagement do not erode criterion validity or fairness expectations [2; 9].

Fairness and compliance emerge as product requirements rather than ex-post audits. A legal-technical synthesis shows divergent fairness notions among applicants and practitioners, with GDPR's fairness, transparency, and proportionality duties interacting with anti-discrimination law across jurisdictions [7]. Scholarship on discrimination disputes underscores the need for reasons, audit trails, and human review in contested decisions [8]. Table 2 distills these demands into actionable design controls that can be embedded into an ATS or a companion "decision-support" layer.

Table 2. Governance and assurance controls mapped to evidence and system components [2; 3; 7-9]

Governance need	Concrete control in an ATS/decision-support layer	Technical anchor
Job-relevance and proportionality	Feature whitelists tied to JD requirements; automatic exclusion of non-job-relevant signals	Feature governance; JD–feature linkage reports
Transparency to applicants	Plain-language notices and optional model-output summaries at screening points	Explanation APIs; templated notices
Auditability of screening	Immutable decision logs with feature attributions; replay datasets	Model registry; signed inference records
Contestability and review	One-click escalation to human review with full case bundle	Human-in-the-loop override UI
Interface-induced bias control	A/B-guardrails on chatbot/AVI prompts; IM/anxiety monitoring with thresholds	Interface experimentation framework
Feature drift and labor-market shift	Periodic refresh of skill taxonomies; alignment with foresight classifiers	Skill-graph sync; drift detectors

For the sponsor's ATS concept (integration-first "layer" rather than a monolithic suite), the literature supports three design commitments. First, center the retrieval stack on resume/job embeddings and learning-to-rank with human-aligned objectives, since those correlate with assessor preferences and lift slate quality [1]. Second, add a compact foundation-model filter with post-hoc explanations only where structured signals are sparse, because few-shot LLMs close much of the gap with supervised baselines on F1/recall and produce inspectable rationales [6]. Third, harvest transversal-skill and demand-forecast signals from job-text corpora into the feature store to stabilize ranking under term drift and to surface "near-fit" candidates on emerging technologies [3]. This combination fits large-volume U.S. tech requisitions where hundreds of resumes arrive per opening and internal teams need a defensible short-list fast, without replacing the incumbent ATS workflow [4].

Risk posture improves when candidate-facing modules inherit the governance controls in Table 2. Chatbot screeners are best positioned as guided information-collection tools rather than as high-stakes assessors given the predictive-validity gap [2]. AVI interfaces get instrumented for impression-management and anxiety signals, with thresholds triggering a human-review path to protect procedural justice expectations [9]. Fairness reviews focus on job-relevance, proportionality, and reason-giving rather than demographic parity targets alone, reflecting the scoping analyses in the legal scholarship [7; 8].

Limitations temper generalization. Resume2Vec's ranking lifts vary by domain, which implies heterogeneous returns in niche roles and necessitates domain-specific evaluation sets [1]. Foundation-model filters achieve strong F1/recall against baseline classifiers but remain sensitive to prompt/shot selection and may concentrate errors on long-tail profiles [6]. Skill-mining pipelines derived from job ads reflect employer signaling rather than the full distribution of productive skills in the labor force; foresight classifiers benefit from pre-training under short-text constraints but show diminishing returns from added model complexity [3]. Chatbot personality inference trades off lower social-desirability bias against weaker criterion validity [2], while AVI interface effects complicate comparisons across implementations and vendors [9].

For an integration-first product like HireSight, the most defensible near-term roadmap is: embed-first retrieval and ranking [1; 6]; skill-graph enrichment with transversal skill signals and near-term demand markers [3]; conservative chatbot/AVI usage with monitoring and human-escalation [2; 9]; and default activation of legal-technical controls for job-relevance, transparency, logging, and review [7; 8]. This trajectory aligns with enterprise adoption patterns mapped in the review literature and addresses the documented hiring inefficiencies through measurable slate-quality gains and governance by design.

Conclusion

Evidence supports replacing brittle keyword filters with résumé–JD embeddings and learning-to-rank, measured by nDCG/RBO; compact foundation-model filters improve recall/F1 in sparse-signal cases and produce post-hoc attributions suitable for audit and recruiter inspection.

Candidate-facing chatbots and automated video interviews are positioned as low-stakes triage and information-collection modules; interface designs require anxiety/IM monitoring and human escalation to preserve procedural justice.

Governance-by-design enters the platform baseline: whitelist features to JD-relevance, present concise notices to applicants, retain signed inference logs, and offer one-click human review.

This integration-first blueprint fits large-volume tech requisitions and aligns with an enterprise deployment that overlays advanced retrieval, explainability, and skill-graph signals onto incumbent ATS timelines, addressing screening labor and slate quality without displacing recruiter judgment.

Reference

1. Bevara, R. V. K., Mannuru, N. R., Karedla, S. P., Lund, B., Xiao, T., Pasem, H., Dronavalli, S. C., & Rupeshkumar, S. (2025). Resume2Vec: Transforming applicant tracking systems with intelligent resume embeddings for precise candidate matching. Electronics, 14(4), 794. https://doi.org/10.3390/electronics14040794
2. Dukanovc, D., & Krpan, D. (2025). Comparing chatbots to psychometric tests in hiring: Reduced social desirability bias, but lower predictive validity. Frontiers in Psychology, 16, 1564979. https://doi.org/10.3389/fpsyg.2025.1564979

3. Gavrilescu, M., Leon, F., & Minea, A.-A. (2025). Techniques for transversal skill classification and relevant keyword extraction from job advertisements. Information, 16(3), 167. https://doi.org/10.3390/info16030167
4. Koman, G., Boršoš, P., & Kubina, M. (2024). The possibilities of using artificial intelligence as a key technology in the current employee recruitment process. Administrative Sciences, 14(7), 157. https://doi.org/10.3390/admsci14070157
5. Martínez-Manzanares, M. E., Urias-Paramo, J. J., Waissman-Vilanova, J., & Figueroa-Preciado, G. (2024). An empirical job matching model based on expert human knowledge: A mixed-methods approach. Applied Artificial Intelligence, 38(1), Article 2364158. https://doi.org/10.1080/08839514.2024.2364158
6. Pendyala, V. S., Thakur, N. B., & Agarwal, R. (2025). Explainable use of foundation models for job hiring. Electronics, 14(14), 2787. https://doi.org/10.3390/electronics14142787
7. Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. Computer Law & Security Review, 53, 105966. https://doi.org/10.1016/j.clsr.2024.105966
8. Seppälä, P., & Małecka, M. (2024). AI and discriminative decisions in recruitment: Challenging the core assumptions. Big Data & Society, 11(1). https://doi.org/10.1177/20539517241235872
9. Suen, H.-Y., & Hung, K.-E. (2024). Revealing the influence of AI and its interfaces on job candidates' honest and deceptive impression management in asynchronous video interviews. Technological Forecasting and Social Change, 198, 122978. https://doi.org/10.1016/j.techfore.2023.122978