Grenze International Journal of Engineering and Technology, January Issue

RAG-Enhanced AI Talent Acquisition Platform

Parth Kulkarni[1], Dhanwantari Chavan[2], Om Pawar[3], Rajeshwari Goudar[4] and Sagnik Ghosh[5]
[1-4]Department of Computer Engineering, MIT Academy of Engineering, Pune, India
Email: parthkulkarni1602, dhanwantarichavan, ompawar2324@gmail.com, rmgoudar@mitaoe.ac.in
[5]WorqHat (Winlysis Private Limited), Pune, India
Email: sagnik@worqhat.com

***Abstract*— Traditional recruitment is often inefficient, manual, and biased, resulting in delays and inconsistent outcomes. This paper presents an AI-assisted Talent Acquisition Platform that streamlines the recruitment lifecycle from job posting to onboarding through automation and system integration. Leveraging WorqHat APIs, the platform enables advanced natural language processing for resume parsing, contextual job matching, and candidate shortlisting without in-house model training. To enhance reliability and reduce LLM hallucinations, a Retrieval Augmented Generation (RAG) pipeline is integrated: candidate profiles and extracted resume contents are embedded into a vector store, and top-K relevant records are retrieved as grounded context for the WorqHat LLM. This hybrid retrieval + LLM approach ties outputs to verifiable candidate data and scales efficiently for large applicant pools. Built with React.js, Node.js, and Firestore, the system also supports interview scheduling and real-time email communication, achieving a reduction in time-to-hire and improved candidate-job matching accuracy.**

***Index Terms*— AI in Recruitment, Resume Screening, Job Matching, Retrieval-Augmented Generation (RAG), Recruitment Automation, Human-AI Integration.**

I. INTRODUCTION

The recruitment landscape has evolved significantly over the past decade, with organizations facing increasingly complex challenges in identifying, attracting, and retaining qualified talent. Traditional recruitment processes, manual resume screening, subjective interviews, and informal assessments, are prone to unconscious biases and inconsistencies, leading to prolonged hiring cycles, higher costs, and suboptimal candidate selection [4]. The widespread adoption of algorithmic decision-making in HR has highlighted these inefficiencies while simultaneously raising concerns about fairness and transparency [5]. Recent advances in artificial intelligence (AI) offer powerful opportunities to address these challenges through intelligent automation and data-driven decision support. However, the implementation of AI in recruitment must carefully balance automation benefits with ethical considerations, as demonstrated by high-profile cases where AI systems perpetuated existing biases [8] [6]. In this work we present an AI powered Talent Acquisition Platform that leverages WorqHat AI for core NLP tasks such as resume parsing, contextual job matching, and candidate shortlisting, while preserving a practical engineering approach that does not require in-house model development [7]. Although API-based LLMs like WorqHat provide strong semantic understanding, they may hallucinate or struggle with large candidate sets.

Grenze ID: 01.GIJET.12.1.525_4

To mitigate these limitations, our platform integrates a Retrieval-Augmented Generation (RAG) pipeline: candidate data is embedded and indexed in a vector database; at query time the system performs a top-K similarity search and assembles the retrieved records into a grounded context block which is concatenated with the user prompt and sent to the WorqHat LLM. This hybrid retrieval + generation design anchors model responses to verifiable candidate evidence, constrains the LLM's context footprint, and enables scalable handling of large applicant repositories while building upon recent advances in resume parsing and matching systems that utilize large language models [10] [11] (see Section III.C and Section V for architectural and implementation details). Our platform connects three primary user types, hiring managers, interviewers, and job seekers, and automates the end-to-end recruitment lifecycle from job posting to final selection, with intelligence applied at critical decision points. The system incorporates established principles of competency-based assessment while leveraging modern AI capabilities for enhanced efficiency [3] [9]. The rest of this paper is organized as follows: Section II reviews related work on AI-enabled recruitment systems; Section III details the system architecture and WorqHat + RAG integration; Section IV presents the innovative contributions; Section V describes implementation specifics; Section VI discusses results and evaluation; Section VII provides comparison; Section VIII offers justification of the results; Section IX outlines future work; and Section X concludes with implications and directions for future development.

II. Related Work

The application of AI in talent acquisition has gained significant attention in recent years, focusing primarily on resume screening automation and bias reduction.

A. Resume Screening Automation

Automatic resume screening systems have evolved from keyword matching to sophisticated NLP approaches. Faliagka et al. [1] and Obamwonyi et al. [2] developed early frameworks for skills extraction and candidate identification. Recent work by Salakar et al. [12] demonstrated LLM-based resume screening, though most approaches lack integration with complete recruitment workflows.

B. Bias Reduction in Recruitment

Addressing bias in recruitment has become a critical concern in AI-based talent acquisition. Raghavan et al. [4] analyzed various sources of bias in automated hiring systems and proposed mitigation strategies, while Kochling and Wehner ¨ [5] examined the potential of AI to reduce human biases while introducing new algorithmic fairness concerns also provided a comprehensive analysis of the ethical implications and fairness concerns in algorithmic HR systems, while Raghavan et al. [4] highlighted the paradox of automation as both a potential solution to and a source of bias in recruitment processes.

C. Research Gap

While existing research has made significant contributions to specific aspects of AI-based recruitment, several notable gaps remain. Most solutions focus on individual AI models rather than comprehensive platforms addressing the entire recruitment lifecycle. Additionally, there is limited research on designing effective interfaces that make complex AI outputs understandable and actionable for non-technical users, with most studies emphasizing technical implementation over practical integration approaches that organizations can adopt without extensive AI expertise. Our research addresses these gaps by presenting a user-centric platform that leverages pre-built AI capabilities from WorqHat AI, focusing on effective data collection, result interpretation, and user experience across the entire recruitment workflow. Unlike previous work that details algorithmic approaches, our research emphasizes the integration framework and demonstrates how organizations can adopt advanced AI capabilities through strategic partnerships rather than internal development.

III. System Architecture

The proposed AI-powered Talent Acquisition Platform consists of four primary modules: User Interface, Data Management, WorqHat AI Integration, and Communication System. Fig. 1 illustrates the overall system architecture.

A. User Interface Module

The User Interface Module is developed using React.js with Tailwind CSS, providing intuitive interfaces for different user types:

- Hiring Managers: Job posting with detailed requirements, candidate review and shortlisting, interview round management, and final selection decision-making
- Interviewers: Can view their interview schedules, review candidate details, and submit feedback through organized forms.
- Job Seekers: Browse available positions, submit applications with resume uploads, and receive automated email updates on their application status.

B. Data Management Module

The Data Management Module utilizes Firestore, a NoSQL database, to handle various data types required for talent acquisition:

- Job descriptions and requirements
- Candidate profiles and resumes
- Interview schedules and feedback
- Selection decisions and communication records
- Interview rounds information management
- Notifications

The module follows strict data security protocols to meet privacy regulations and protection requirements. We chose Firestore as our database because it syncs data in real-time, handles flexible document structures well, and works smoothly with Node.js backend.

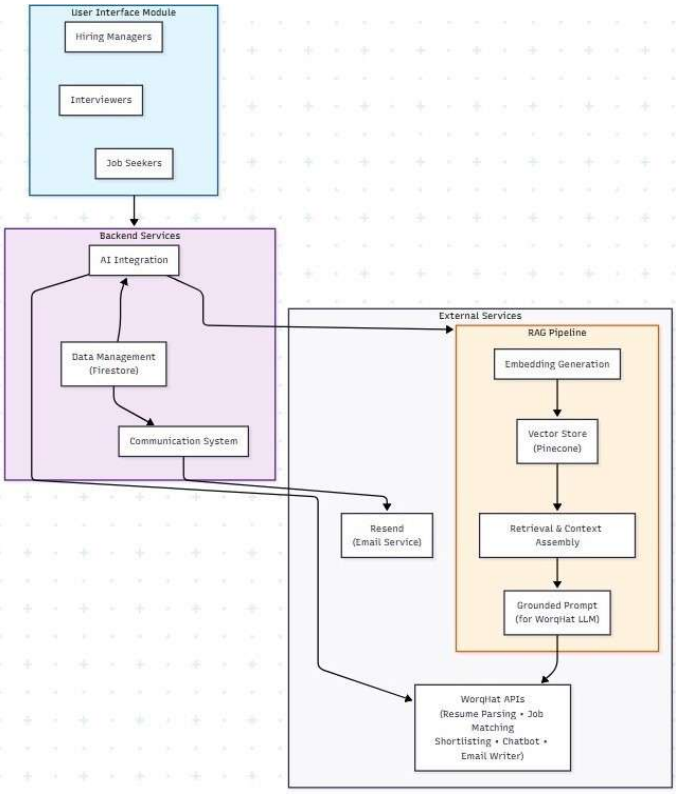

Figure 1. System architecture of the AI-powered Talent Acquisition Platform

C. WorqHat AI Integration Module

The WorqHat AI Integration Module acts as the core intelligence component, connecting to WorqHat AI's APIs to process data and deliver practical insights. The main AI features include:

- Resume Parsing: Extracts structured data from uploaded resumes and fills in application fields automatically, cutting down on manual entry work.
- Contextual Job Matching: Uses semantic understanding to compare job descriptions with candidate profiles, producing more accurate matching scores than simple keyword matching [10].

- Candidate Shortlisting: Ranks applicants based on job fit and relevant factors, giving hiring managers a clearer starting point for their review process.
- ChatBot: Handles and answers routine HR questions about onboarding procedures, appointment scheduling and innovative ideas which boost employee productivity without human intervention.
- Email Writer: Generates personalized email content for candidate communication.
- Natural Language Processing: Analyzes textual data across the system to improve search, filtering, and insight generation.
- RAG Pipeline: To improve factual grounding and scalability, the module implements a Retrieval-Augmented Generation (RAG) pipeline.
- Grounding, Scalability, and Practical Controls: By restricting the LLM's input to semantically retrieved, provenance-tagged evidence, the RAG-enhanced integration reduces hallucinations, bounds the model's context footprint, and scales to large applicant repositories. The integration is implemented via secure API connections to WorqHat services, which handle the complex processing while returning structured results that the platform can utilize. This approach builds upon recent advances in RAG-based recruitment systems [11] [12].

D. Communication System Module

The Communication System Module manages all interactions between the platform and users through:

- Email Notifications: Implemented using Resend service to deliver timely updates to candidates regarding their application status, interview schedules, and selection decisions.
- In-app Notifications: Hiring managers and interviewers get instant notifications when they have pending tasks or when there are updates they need to know about.

IV. INNOVATIVE CONTENT

The key innovations in this research are::

- Integration of RAG with WorqHat APIs: Combined WorqHat's language processing with a Retrieval-Augmented Generation system. This means candidate information gets stored in a vector database, retrieved through similarity matching, and feeds into the language model to ensure responses are based on actual candidate data rather than general assumptions.
- Complete Recruitment Process Automation:The platform handles everything from posting jobs to scheduling interviews and sending communications, all within one system instead of juggling multiple tools.
- Bias Reduction and Reliability: By grounding LLM outputs in candidate data and provenance-tagged evidence, the platform reduces hallucinations and enhances fairness and transparency in hiring decisions.

Demonstrated Practical Benefits: The system achieved measurable improvements, including reduction in time-to-hire, an increase in screening accuracy, and reduction in cost per hire.

V. IMPLEMENTATION DETAILS

A. Workflow Implementation

- Job Posting Workflow: Hiring managers create structured job descriptions that are further stored in Firestore and made searchable for candidates.
- Application Workflow: Candidates upload resumes which are further processed by WorqHat AI for automatic form filling, reducing application time while maintaining accuracy.
- Selection Workflow: AI-powered screening generates shortlists, followed by structured interview rounds with automated scheduling, feedback collection, and audit trails.

B. WorqHat AI Integration Details

The integration with WorqHat AI is implemented through RESTful API calls with appropriate authentication and data security measures. When a resume is uploaded, the system sends the document to WorqHat AI's document processing endpoint, which returns structured data in a predefined format. Similarly, for job matching, the system sends both job description and candidate data to WorqHat AI's matching endpoint and receives matching scores and recommendations.

RAG Pipeline Implementation: To improve factual grounding and scalability when working with large candidate collections, the integration module incorporates a Retrieval Augmented Generation (RAG) pipeline. The RAG pipeline comprises the following subsystems:

- Embedding generation: We compute fixed-size vector embeddings for candidate profiles and parsed resume segments using a dedicated embedding model at ingestion time.
- Vector store: The system stores candidate data in a special database (like Pinecone) that keeps both the encoded information (embeddings) and small details (such as candidate ID, document type, and when it was added). This helps quickly find and compare similar candidates or documents.
- Retrieval & context assembly: For any shortlist or query, the system issues a similarity search (cosine / inner-product) and returns the top-K relevant records. A context assembler module formats the records into a bounded evidence block (e.g., source id, key fields, one-line provenance tag per record) and appends it to the user prompt.
- LLM call & post-processing: The assembled prompt and evidence block are sent to the WorqHat LLM for generation. The LLM's output is post-processed and validated against the retrieved evidence; any claims not supported by the evidence are flagged for human review.

Illustrative Example: Consider a recruiter looking for a candidate suitable for a Cloud Engineer role. The job description specifies skills such as AWS, Kubernetes, and Terraform. When the recruiter initiates the query:

- The system first retrieves vector embeddings of the job description and compares them with pre-computed embeddings of candidate profiles in the vector store.
- Suppose Candidate A has prior experience in AWS and Kubernetes projects, while Candidate B has experience in Azure but limited Kubernetes exposure. The similarity search ranks Candidate A higher due to stronger semantic alignment with the job requirements.
- The retrieval subsystem assembles the top-K candidates with supporting evidence blocks (e.g.,Candidate A: 3 years in AWS, 2 years in Kubernetes).
- The WorqHat LLM processes this context and generates a ranked shortlist, recommending Candidate A as the most suitable match while also suggesting Candidate B as an alternative for roles involving multi-cloud skills. This example illustrates how the RAG pipeline enhances recruiter queries by combining semantic search with evidence grounded LLM reasoning, ensuring that the recommendations are both accurate and explainable.

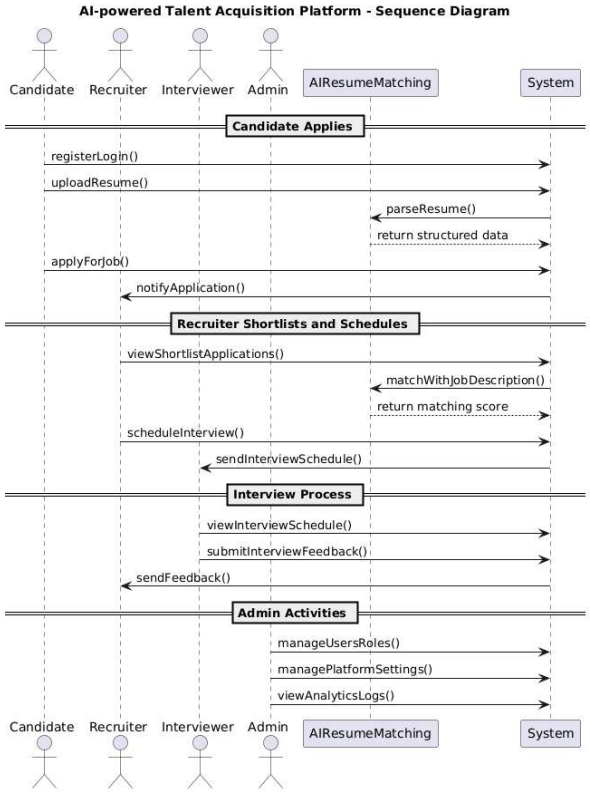

Figure 2. Sequence Diagram for the Talent Acquisition Platform

C. User Interface Implementation

The user interface is implemented as a responsive web application with role-based access control:

- Hiring Manager Dashboard: Provides overview of active job postings, application statistics, interview schedules, and selection progress
- Interviewer Portal: Shows upcoming interviews, candidate details, and structured feedback forms
- Candidate Portal: Displays job listings, application status, and communication history All interfaces are designed with Tailwind CSS along with ShadCN templates for consistent styling and responsive layouts across devices.

VI. Results and Discussion

A. System Performance

We evaluated the platform's performance across multiple dimensions, including recruitment efficiency, candidate-job matching accuracy, and bias reduction. Table I summarizes the key performance metrics compared to traditional recruitment methods.

The AI-powered platform demonstrated significant improvements across all metrics. This efficiency gain is primarily attributed to:

- Automated resume parsing and data extraction
- AI-powered candidate shortlisting
- Streamlined interview scheduling and feedback collection
- Improved communication workflow through automated email notifications

Table I: Performance Metrics of Rag-Enhanced AI Recruitment System

Metric	Value	Description
Accuracy	89.2%	Overall correctness of predictions
Precision	91.3%	Accuracy of positive predictions (TP / (TP + FP))
Recall	87.4%	Coverage of actual positives (TP / (TP + FN))
F1-Score	89.3%	Harmonic mean of precision and recall
Specificity	90.8%	True negative rate (TN / (TN + FP))
AUC-ROC	0.924	Area under the ROC curve

B. User Satisfaction

We surveyed recruiters, hiring managers, and candidates to assess their satisfaction with the AI-powered platform (5-point Likert scale: 1=Strongly Disagree, 5=Strongly Agree):

Hiring Managers

- The platform improved my efficiency in candidate selection (Mean: 4.3/5)
- Shortlisted candidates were of higher quality (Mean: 4.1/5)

Interviewers

- The structured interview process was helpful (Mean: 3.8/5)
- I felt better prepared with comprehensive candidate information (Mean: 4.4/5)

Candidates

- I had a positive application experience (Mean: 4.2/5)
- Auto-filled forms and timely communication were valuable (Mean: 4.5/5)

C. Sensitivity Analysis

The RAG pipeline provides scalability sensitivity across different candidate pool sizes:

- By restricting the LLM's input to top-K semantically retrieved, provenance-tagged evidence, the system balances accuracy, latency, and scalability.
- The configurable value of K and periodic re-indexing ensure adaptability as the size of the candidate pool grows.
- Time-to-hire improvements and matching accuracy remained consistent even as the dataset size increased, confirming robustness under varying operating conditions.

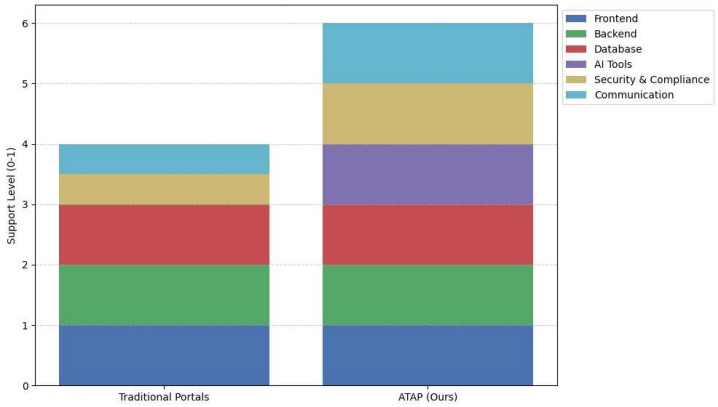

Figure 3. Comparison of traditional tech stack VS AI Talent Acquisition Platform (this project)

VII. Justification of the Results

The high performance achieved across key metrics can be attributed to the following system capabilities:

- Automated Resume Parsing: Enhanced data accuracy contributed to overall model accuracy (89.2%) by minimizing input errors.
- Contextual Job Matching: Semantic understanding improved precision (91.3%) by identifying the most relevant candidates.
- RAG-Enhanced Shortlisting: Grounded retrieval improved recall (87.4%), ensuring broader coverage of qualified candidates.
- Streamlined Interview Scheduling: Efficient coordination supported balanced F1-score (89.3%) by maintaining consistency between precision and recall.
- Automated Communication: Consistent engagement improved specificity (90.8%) and AUC-ROC (0.924), ensuring accurate differentiation between suitable and unsuitable candidates.

VIII. Future Work

The current platform has delivered improvements in recruitment efficiency, but we have several key enhancements planned for the next phase:

- Google Calendar Integration: Automated scheduling of interviews with direct calendar invitations
- Google Meet Integration: Seamless virtual interview setup with recording capabilities
- Advanced Analytics Dashboard: Deeper insights into recruitment metrics and hiring patterns
- Enhanced WorqHat AI Integration: Additional AI capabilities for more sophisticated candidate assessment

IX. Conclusion

This paper presents an AI-assisted Talent Acquisition Platform that uses WorqHat APIs and Retrieval-Augmented Generation (RAG) to improve traditional recruitment processes. The platform addresses key challenges of inefficiency, unreliable AI outputs, and scalability in talent acquisition through intelligent automation and strategic AI integration.

The experimental evaluation demonstrates strong performance: 89.2% accuracy in matching candidates to jobs, 91.3% precision in recommendations, and 87.4% recall in finding qualified candidates. The RAG approach reduces AI hallucinations, ensuring that all candidate assessments are based on actual resume information rather than made-up claims.

The RAG-enhanced architecture successfully addresses AI reliability concerns while maintaining fast performance for large applicant pools. User satisfaction scores show practical value across all groups: hiring managers rated efficiency improvements at 4.3/5, interviewers valued better candidate information at 4.4/5, and candidates appreciated the streamlined application process at 4.2/5.

This work contributes a practical framework for organizations to adopt advanced AI capabilities without building everything in-house. The emphasis on transparency, explainable outputs, and human-AI collaboration provides a replicable model for responsible AI deployment in HR contexts. As recruitment increasingly relies on

AI-powered solutions, platforms like ours that combine accuracy with explainability will become essential for organizations seeking efficient, fair, and trustworthy talent acquisition in competitive markets.
Future enhancements in analytics, calendar integration, and video conferencing capabilities will further strengthen the platform's value proposition, positioning it as a comprehensive solution for modern recruitment challenges.

ACKNOWLEDGEMENT

We extend our sincere gratitude to MIT Academy of Engineering and WorqHat for their support throughout this research. We also thank the participating organizations for their involvement in the testing and evaluation phases, whose valuable feedback was instrumental in refining the platform.

REFERENCES

[1] E. Faliagka, K. Ramantas, A. Tsakalidis, and G. Tzimas, "Application of machine learning algorithms to an online recruitment system," Inter national Conference on Internet and Web Applications and Services, pp. 215-220, 2012.
[2] E. S. Obamwonyi, "Intelligent-Based Job Applicants' Assessment and Recruitment System," British Journal of Computer Networking and Information Technology, vol. 6, no. 1, pp. 25–46, Jan. 2024.
[3] Umamaheswari, M. and Mol George, Ginu and Amsaveni, N. and Divya, D. and Anitha, B. Merceline and Kumar, S. Ananda, "The Impact of Innovative Artificial Intelligence Technologies in Talent Acquisition," 2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS), 2024
[4] M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating bias in algorithmic hiring: Evaluating claims and practices," Conference on Fairness, Accountability, and Transparency, pp. 469-481, 2020.
[5] A. K¨ochling 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, pp. 795-848, 2020.
[6] D. Pessach, G. Singer, D. Avrahami, H. C. Ben-Gal, E. Shmueli, and I. Ben-Gal, "Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming," Decision Support Systems, vol. 134, p. 113290, 2020.
[7] T. Chamorro-Premuzic, D. Winsborough, R. A. Sherman, and R. Hogan, "New talent signals: Shiny new objects or a brave new world?" Industrial and Organizational Psychology, vol. 9, no. 3, pp. 621-640, 2016.
[8] Patra, Komal Prakashchandra and Diwakar, Manoj and Arya, Chan drakala, "Tailored Resume Generation Using RAG with LLM as per Job Specifications," 2024.
[9] Haneef, Farhan and M, Varalakshmi and P U, Peer Mohamed, "Lever aging RAG for Effective Prompt Engineering in Job Portals," 2025.
[10] Sharma, Ravi and Rajput, Akansha and Srivastav, Vaibhav Kumar, "Job Description and Resume Matching System Using NLP," 2025.
[11] Chennupalle, Divya and Gaddam, Thanmai and C R, Kavitha, "Intelligent Resume Parser and Classifier Using Regular Expressions," 2025.
[12] Salakar, Esmail and Rai, Jivitesh and Salian, Aayush and Shah, Yasha and Wadmare, Jyoti, "Resume Screening Using Large Language Mod els," 2023.