IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

# CareerCraft AI: An Intelligent Employment Support System Integrating Resume Building, Applicant Tracking System (ATS) Analysis, and GeoBased Job Matching

**Samidha S. Shende[1], Srushti S. Khate[2], Srushti M. Deshmukh[3], Anuj G. Gondane[4], Prof. N. K. Chede[5]**

Department of Artificial Intelligence and Data Science, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra, India[1]

Department of Artificial Intelligence and Data Science, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra, India[2]

Department of Artificial Intelligence and Data Science, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra, India[3]

Department of Artificial Intelligence and Data Science, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra, India[4]

Department of Artificial Intelligence and Data Science, Prof. Ram Meghe Institute of Technology and Research, Amravati, Maharashtra, India[5]

**Abstract:** Job seekers often rely on separate tools for resume analysis, resume enhancement, job-description matching, interview practice, and job search. This paper presents CareerCraft AI, an integrated, candidate-facing web prototype that combines these functions in one workflow. The system uses rule-based extraction for PDF and DOCX resumes and a hosted large language model (LLM), with rule-based fallbacks for selected tasks. The paper describes the architecture, data model, processing workflow, and exploratory observations from a deployed instance. Illustrative outputs demonstrate the available interfaces, while consistency checks identify unresolved differences between displayed scores and simple count-based calculations. Exploratory testing also identified parsing limitations, variation in generated content, and unsupported achievement metrics in some resume rewrites. These observations motivate factual review, transparent score provenance, privacy safeguards, and systematic evaluation. The study documents an integrated prototype; its accuracy, reliability, fairness, and effectiveness have not been established.

Keywords: CareerCraft AI; Applicant Tracking System; Large Language Models; Resume Analysis; Job Matching; Interview Preparation.

## I. INTRODUCTION

Recruitment increasingly relies on digital platforms and automated screening, making it important for job seekers to present their qualifications clearly and align resumes with job requirements. Students and first-time applicants may struggle to assess resume quality, identify skill gaps, prepare for interviews and find suitable vacancies across separate tools.

Career-support tools often handle resume analysis, editing, job-description matching, interview preparation and job search independently. Integrating these functions may reduce the need to switch between services, although integration alone does not establish analytical accuracy or improved employment outcomes.

CareerCraft AI was developed as a candidate-facing web prototype that combines resume analysis, AI-assisted resume enhancement, job-description matching, self-introduction generation, multiple-choice mock-interview practice, and external job search. Its processing approach combines rule-based extraction of information from PDF and DOCX resumes with a hosted large language model (LLM) for selected analysis and content-generation tasks. Rule-based fallbacks are also documented for selected functions. The prototype is intended to provide job seekers with a consolidated environment for preparing and reviewing application materials.

 This work is licensed under a Creative Commons Attribution 4.0 International License 

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

CareerCraft AI also raises reliability concerns: parsing errors can affect downstream analysis, while LLM-generated scores and text may vary or include unsupported claims. Evaluation must therefore consider output consistency, transparency and factual reliability, not merely feature availability.

This paper documents CareerCraft AI's architecture and workflow, reports limited exploratory observations, and proposes an evaluation approach for future work. It addresses three questions: (1) How are rule-based resume processing and LLM-based tasks integrated in a candidate-facing workflow? (2) What output behaviours and internal-consistency issues were observed in exploratory use? (3) What evaluation is needed before making claims about accuracy, reliability, fairness, or effectiveness?

The contribution is a structured account of an integrated prototype, its documented processing and data flow, selected observations from use, and a proposed evaluation plan. The paper introduces no new scoring algorithm and does not claim superiority over existing career-support tools.

The remainder of this paper reviews related work, describes the system methodology and data model, presents the observations and their implications, discusses limitations and future evaluation, and concludes with the principal findings.

## II. Literature Review

This section is a selective, thematic review of sources whose bibliographic details could be verified. It is not a systematic review. Peer-reviewed articles, preprints and one industry report are distinguished in the text and in the reference list, and statements about each source are limited to what its abstract or accessible full text supports.

### A. Resume Parsing and Information Extraction

Nguyen et al. [2] found that LLM-based skill extraction did not outperform supervised baselines overall, although it handled syntactically complex skill mentions better. They also noted that hosted-model updates limit reproducibility. Khelkhal and Lanasri [3] describe a preprint combining CV text extraction, rule-based and lightweight machine-learning methods, and embedding-based skill matching. Their expert evaluation of approximately 1,000 French and English CVs identified difficulties with complex layouts. In contrast, CareerCraft AI uses rule-based parsing before passing both raw text and parsed fields to the LLM (Section III-B). Since the parser has not been evaluated against annotated resumes, its errors may propagate to downstream analysis.

### B. ATS-Oriented Resume Assessment and LLM-Based Judgment

Applicant tracking systems (ATS) are widely used for initial screening [1]. Although commercial candidate-side ATS checkers exist, the sources reviewed here provide no peer-reviewed evaluation of their scoring. Vaishampayan et al. [4] compared zero-shot GPT-4 scores with human ratings for 736 resumes submitted to real job openings. They found only minor correlation, indicating that the scores were not interchangeable; chain-of-thought prompting improved ratings, while LLM scores did not show larger group differences than human ratings. Zheng et al. [5] found that strong LLM judges agreed with human preferences in over 80% of chat-assistant comparisons, but also identified position, verbosity, and self-enhancement biases and limitations in reasoning. These findings cannot be directly generalized to resume scoring. Together, they highlight the need to validate LLM-generated resume scores against human ratings and test their repeatability.

Amir Hussin et al. [6] describe a web-based system that screens resumes for recruiters and reports matched and missing skills to job seekers, while also generating ATS-friendly resumes. It combines named-entity recognition, regular expressions, and Sentence-BERT similarity, with TF-IDF as a fallback. The study reports functional and extraction tests, but their details were not examined here.

### C. Job-Description Matching

Ajjam and Al-Raweshidy [7] proposed a semantic-similarity framework for job-description matching and reported higher similarity scores than a keyword-based baseline in simulated tests. CareerCraft AI instead uses an LLM to extract job requirements, compare them with resume content, and return a match score and keyword lists. Although this approach can account for differently worded skills, its outputs depend on the model and may not be repeatable.

### D. Job Recommendation and Location-Aware Search

Zhao et al. [8] describe an industrial-scale job-to-candidate recommender using fused embeddings and location signals, while Bied et al. [9] present a large-scale job-recommendation system with fairness analysis. Al Awad and Omar [10] combine semantic retrieval with explicit constraints such as location, work mode, and seniority. Vyas et al. [11] describe an explainable matching system using Transformer embeddings, skill knowledge graphs, and reranking based on factors including skills, experience, and location.

Unlike these systems, CareerCraft AI uses an external job-listing API to retrieve listings by role or domain and optional city. Location acts as a search constraint, but resume-based ranking, skill-based filtering, and relevance evaluation are not documented.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

E. AI-Assisted Interview Preparation

Research on AI in interviews has mostly examined employer-side assessment. In a field study with 152 real applicants, Suen and Hung [12] found that AI-based asynchronous video interviews increased cognitive trust relative to a non-AI condition, that tangibility and transparency features raised cognitive and affective trust, and that immediacy had no statistically significant effect. This setting differs from the text-based, multiple-choice practice in CareerCraft AI, so the study is used only as context for how explanations might be presented. Among the sources reviewed here, none evaluates candidate-side practice tools.

F. Fairness, Explainability and Privacy

Raghavan et al. [13] examined what vendors of algorithmic pre-employment assessments disclose about development, validation and bias mitigation, from technical and legal perspectives. Kumar et al. [14] survey fairness definitions and mitigation approaches in recruitment recommender systems and contrast them with EU and US employment law, and Mujtaba and Mahapatra [15], in a preprint, review biases, fairness metrics, mitigation methods and auditing approaches for AI-driven recruitment. CareerCraft AI supports candidates and makes no selection decisions, but the questions these works raise about how scores are produced, validated and communicated apply equally to any automated resume score. An empirical comparison of LLM and human resume ratings [4] found no larger group differences for the LLM, but it concerns a different system and setting and does not show that CareerCraft AI is free of bias. Resumes also contain personal data, and the documented design transmits resume text to an external LLM provider (Section III-H).

G. Comparison and Research Gap

TABLE I VERIFIED RELATED WORK AND ITS RELEVANCE TO CAREERCRAFT AI

| Ref. | Source type and setting | Main focus | Relevance to, and limits for, CareerCraft AI |
|---|---|---|---|
| [1] | Industry report; employer survey | Prevalence of automated screening | Motivates candidate-side ATS awareness; not peer reviewed |
| [2] | Workshop paper; six skill datasets | LLM in-context learning versus supervised extraction | Motivates validating LLM judgments; CareerCraft AI parses by rules |
| [3] | Preprint; employer-side pipeline | Hybrid CV extraction, embeddings and weighted, explainable matching; expert evaluation | Parallel to local parsing plus semantic scoring; reports layout difficulties |
| [4] | Conference paper (Findings); 736 resumes | Zero-shot GPT-4 versus human resume ratings | Most direct evidence on LLM resume rating; scores not interchangeable with human ratings |
| [5] | Conference benchmark paper | Agreement and biases of LLM judges | Context for LLM-produced scores; different task |
| [6] | Journal article; web system | Recruiter and job-seeker modules; NER and Sentence-BERT; ATS-friendly resume generation | Integrated workflow exists; evaluation details not examined here |
| [7] | Journal article; simulated tests | Semantic similarity versus keyword matching | Alternative to LLM-judged matching; different method |
| [8] | Preprint; industrial system | Embedding retrieval with a location signal | Scale and method differ from the API proxy used here |
| [9] | Conference paper; public employment service | Scalable recommendation with fairness analysis | Context on scale and fairness |
| [10] | Preprint; offline evaluation | Hybrid retrieval, filters, explanations | Supports keeping location as an explicit constraint |
| [11] | System-demonstration paper | Embeddings, knowledge graph, explainable reranking | Different method and scale |
| [12] | Journal article; 152 applicants | Trust in AI video interviews | Context for explanation design; setting differs |
| [13] | Conference paper; vendor review | Disclosure and validation of assessment tools | Framing for score transparency |

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

| Ref. | Source type and setting | Main focus | Relevance to, and limits for, CareerCraft AI |
|---|---|---|---|
| [14] | Journal survey | Fairness and law in recruitment recommenders | Fairness framing; recommender focus |
| [15] | Preprint; survey | Bias, metrics, mitigation, auditing | Fairness framing for any resume scoring |

The reviewed literature addresses individual stages of career support, including resume parsing and skill extraction [2], [3], LLM-based resume assessment [4], [5], job matching and recommendation [7]–[11], interview-related trust [12], and fairness and transparency in automated recruitment [13]–[15]. Existing systems also combine multiple recruitment functions, including resume screening, skill-gap feedback and resume enhancement [6]. Therefore, integration alone is not claimed as a novel contribution of CareerCraft AI. Instead, this paper documents a candidate-facing prototype that combines resume analysis, enhancement, job-description matching, interview practice and job search, and examines the internal consistency and limitations of its observed outputs. The available evidence is exploratory and does not establish the system's accuracy, reliability or effectiveness. This research gap is limited to the scope of the literature reviewed and does not imply that no comparable platforms exist.

## III. METHODOLOGY

The system description below is drawn from the authors' implementation notes and technical documentation; screenshots of a deployed instance (Section V) are used where they confirm a feature. Section V labels each statement by its level of evidence.

The "seminar report" cited below is the authors' earlier departmental project report; it is not a published source.

### A. System Architecture

CareerCraft AI follows a client-server architecture (Fig. 1). The front end is a React 18 single-page application using React Router, Axios, and Tailwind CSS. The back end is a Node.js/Express REST API, with PostgreSQL accessed through Sequelize. Authentication uses JSON Web Tokens (JWTs), while Google identity and a hosted LLM and job-listing API provide external services. The implementation notes identify Groq as the LLM provider, Adzuna for job listings, Netlify for front-end hosting, and Render for PostgreSQL hosting. These provider details are documented in the notes but are not independently confirmed by the screenshots. Inconsistencies in the documented charting approach remain unresolved.

 This work is licensed under a Creative Commons Attribution 4.0 International License 

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

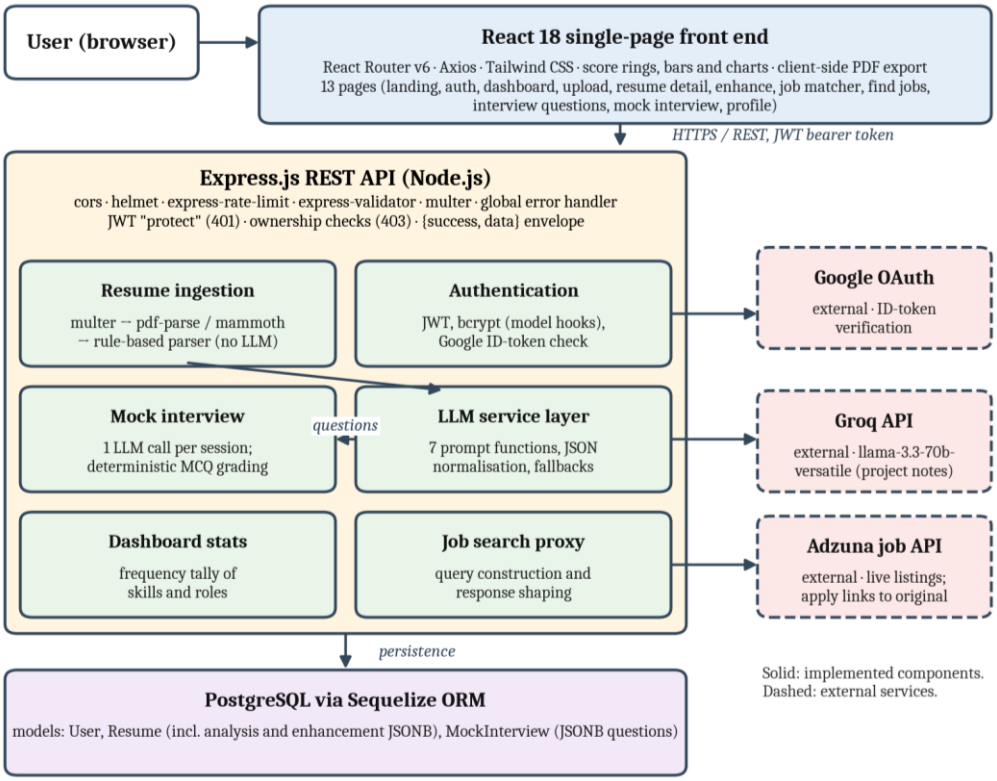



Fig. 1 Architecture of CareerCraft AI, drawn from the implementation notes. Dashed boxes are external services. An earlier project report also showed a resume-builder module, which is omitted because no from-scratch builder appears in the implemented pages. External service names are as given in the implementation notes and were not confirmed from screenshots (LLM provider, job-listing provider).

B. Resume Ingestion and Deterministic Parsing

Resume uploads are validated on both the client and server, including file type and a 5 MB size limit. The server uses Multer to enforce an extension whitelist and assign a unique filename. Text is extracted using pdf-parse for PDF files and Mammoth for DOCX files; subsequent parsing is deterministic and does not use an LLM (Fig. 2). Regular expressions extract contact details, links, GPA and durations, while a section-based parser organises education, skills, projects and experience into structured fields. The parsed data is stored with the resume and passed to the LLM alongside the raw text. The parser is intended for English-language, text-based resumes; image-only documents and OCR are unsupported. The exact heading patterns and short-line threshold are not consistently documented.

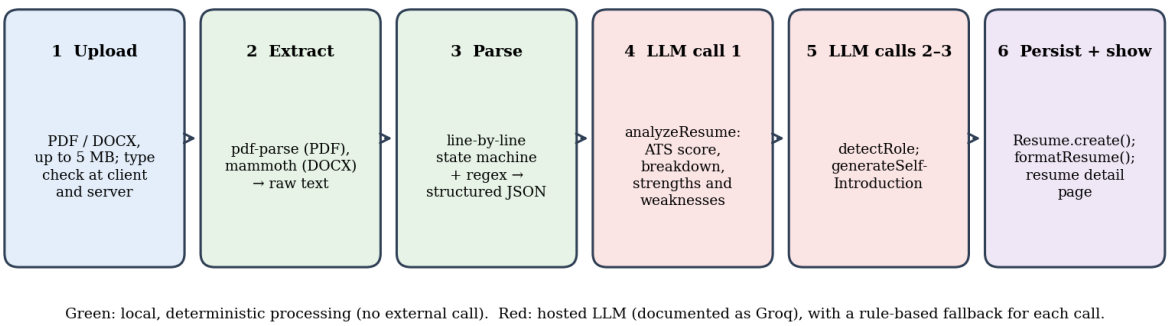



Fig. 2 Upload-to-display pipeline. Steps 1–3 are local and deterministic; steps 4–5 are three sequential LLM calls, each with a rule-based fallback.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

**Algorithm 1. Section-parsing state machine (as described in the implementation notes)**

1. Split the normalised text into lines; set inSection ← false.
2. For each line, test it against the section-header patterns (education, skills, projects, experience). On a match, close the current entry and set the current section.
3. For a line inside a section, compute isBullet (starts with •, -, *, | or >), looksLikeSentence (contains verbs such as "built", "developed", "using", "with") and isShort (length below a threshold that the notes do not specify).
4. isTitle ← ¬isBullet ∧ isShort ∧ ¬looksLikeSentence ∧ (length > 3).
5. If isTitle, close the previous entry and open a new entry titled with this line; otherwise append the line to the current description and merge any technology keywords detected in it.
6. At the end of a section, push the final entry. Return arrays of {title, description, technologies[]}.

C. LLM Service Layer

CareerCraft AI uses a shared LLM service containing seven prompt functions for resume analysis, role detection, self-introduction generation, interview-question generation, resume enhancement, job matching and answer evaluation. The documented model is Llama 3.3 70B Versatile, accessed through the Groq API, with requests configured for JSON responses. Returned fields are normalised before results are stored or displayed. Five functions have independent rule-based fallbacks to maintain functionality when LLM requests fail; the answer-evaluation function is documented but is not called by any route. Although the interface documentation mentions GPT-4o, the implementation notes identify Groq. This provider discrepancy remains unresolved, and the actual model used by the deployed application has not been independently verified.

D. ATS Scoring and Resume Enhancement

CareerCraft AI uses an LLM-based ATS rubric comprising ten positive criteria totalling 100 points and four deductions (Table II). The model evaluates the criteria and deductions rather than application code calculating the score. The relationship between the rubric and the four category scores displayed in the interface—Keyword Match, Format Compliance, Content Quality and Section Completeness—is undocumented.

TABLE II ATS SCORING RUBRIC EMBEDDED IN THE PROMPT, AS LISTED IN THE IMPLEMENTATION NOTES

| Positive criterion | Points | Deduction | Points |
|---|---|---|---|
| Professional summary present | +15 | Passive language | −10 |
| Quantifiable achievements | +15 | No metrics | −10 |
| Strong action verbs | +15 | Missing summary | −5 |
| Five or more skills | +10 | Grammar issues | −5 |
| Experience present | +10 | | |
| Project with technology stack | +10 | | |
| Education complete | +10 | | |
| Certifications | +5 | | |
| Contact information | +5 | | |
| Consistent formatting | +5 | | |

Resume enhancement uses a single LLM call, with a general or job-description-tailored prompt depending on whether the supplied job description exceeds 50 characters. Results include severity-rated issues, an enhanced summary and rewritten content. They are stored as JSONB with a timestamp, allowing users to revisit results without triggering another call. Users can edit and preview the enhanced resume and export it to PDF using jsPDF.

E. Job-Description Matching

The Job Matcher accepts pasted job-description text or an uploaded file, using the same PDF and DOCX extraction libraries described in Section III-B. Inputs shorter than 50 characters are rejected by both the interface and server. A single LLM call compares the resume, job description and parsed skills to identify requirements, matched and missing keywords, skill-gap categories, a match score and improvement suggestions. Although the seminar report defines the match percentage as the proportion of extracted job keywords found in the resume (Eq. 4, Section IV-C), the implementation notes indicate that the score is returned by the model. The deployed calculation remains unresolved. Results are not persisted, so revisiting the page triggers a new LLM call.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

F. Self-Introduction and Mock Interview

CareerCraft AI generates a self-introduction from the parsed resume and detected role using an LLM call. Its mock-interview module generates 20 multiple-choice questions per session across seven categories: skill-technical, skill-scenario, skill-concept, project-technical, project-scenario, HR-skill and HR-project. Each question includes four options, a stored answer key and an explanation. Responses are evaluated deterministically by comparing the selected answer with the stored key, and progress is saved after each answer to support session resumption. The final score is the percentage of correct answers. Although the seminar report describes typed responses evaluated by an LLM, the documented implementation uses multiple-choice questions; the free-text evaluation function is not called by any route.

G. Job Search

The Find Jobs page takes a role or domain and an optional city. The server validates that the query is non-empty, builds a request, according to the implementation notes, to the Adzuna search endpoint for India with the application credentials, the query and the location, and maps each result to a compact record (title, company, location, description, salary range, posting date and apply URL). The apply URL is the aggregator's redirect to the original posting, and the page opens it in a new tab. The notes describe no ranking against resume skills and no logging of searches, which the seminar report had proposed.

H. Data Model, Security and Privacy

The documented schema contains seven entities, while the implementation notes identify three Sequelize models: User, Resume and MockInterview. Passwords are hashed using bcrypt, and authentication uses signed JSON Web Tokens (JWTs) valid for 30 days. Protected routes validate tokens and check resource ownership. Google sign-in verifies the ID token on the server, while environment variables, cross-origin restrictions and validation middleware support application security. No penetration test or security audit is documented.

Resume text, parsed fields and job descriptions are sent to an external LLM provider, while uploaded files are stored on the server's disk. The available documentation does not establish whether users receive a consent notice, what retention terms apply at the provider, or how long files are retained. Although a resume-deletion control is visible (Fig. 3), deletion of associated files and records has not been verified. These privacy and data-protection issues require assessment before wider deployment.

I. Implementation Environment

The documented requirements are Node.js 18 or later, React 18 or later, PostgreSQL, and a current Chrome or Firefox browser; the documented minimum hardware is an Intel Core i5-class processor with 8 GB of RAM. Exact package versions were not supplied.

## IV. Data Model, Computations and Proposed Evaluation

This section states what the system stores and computes, separating deterministic computations that are documented from values produced by the LLM. No formula is given for LLM-produced scores because none is documented for the deployed code; where a formula is shown for such a score, it is labelled as a conceptual or intended definition.

A. Data Model

Table III compares the seven entities in the technical documentation with the three Sequelize models identified in the implementation notes. Parsed resume fields, ATS analysis and enhancement results are stored within the Resume record rather than separate tables. Job-match results are not persisted, and job-search logging is not documented. Foreign-key constraints and indexes are reported for Resume but not for Mock Interview; the foreign-key deletion rules remain unverified.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

TABLE III DOCUMENTED ENTITIES AND THEIR STATUS IN THE IMPLEMENTATION NOTES

| Documented entity | Documented attributes | Status in implementation notes |
|---|---|---|
| USER | id, name, email, password hash, Google id, created_at | User model; bcrypt hooks |
| RESUME | id, user_id, source_type, raw_text, file_path, created_at | Resume model; foreign key and indexes; extended with parsed data, analysis and enhancement fields |
| PARSED_FIELD | resume_id, section_name, content_json | No separate table; parsed data held in the resume record |
| ATS_RESULT | resume_id, overall_score, section_scores_json, target_role | No separate table; ATS analysis stored in the Resume record |
| JD_MATCH | resume_id, job_description, match_score, missing_keywords | Not persisted; recomputed at each visit |
| JOB_SEARCH_LOG | user_id, city, role, searched_at | No log described; search is a proxy |

B. Input and Output Structures

The parsed resume contains contact details and structured education, skills, projects and experience fields. ATS analysis returns an overall score, category scores, strengths and improvement areas, while resume enhancement provides rewritten content and severity-rated issues. Job matching returns a score, matched and missing keywords, skill-gap categories and suggestions. Mock-interview questions contain four options, an answer key, an explanation and response-related fields. These structures are based on implementation notes and screenshots; the exact JSON schemas were not supplied.

C. Computations

Rubric score (conceptual). The rubric in Table II can be written as a sum over criteria judged by the model, where $c_i$ and $d_j$ are binary judgments, $w_i$ are the positive points (summing to 100) and $p_j$ the deduction points:

$$S_{rubric} = \Sigma_i \, w_i \, c_i - \Sigma_j \, p_j \, d_j \quad (1)$$

The notes do not state whether the result is clipped to the range 0-100, and the model, not code, evaluates $c_i$ and $d_j$; the equation therefore describes the intended scheme and not a verified implementation.

Fallback score (documented, deterministic). When the LLM call fails, the analysis function computes a score from regular-expression tests, where q, a and s are 1 if quantified results, action verbs and a summary heading are found and 0 otherwise, and $[\cdot]$ is 1 if the condition holds:

$$S_{fb} = 40 + 10[|skills| \geq 5] + 10[|experience| > 0] + 10[|projects| > 0] + 10q + 10a + 5s \quad (2)$$

The formula produces values from 40 to 95, assuming each indicator is binary as defined above. Strengths and weaknesses in the fallback are generated from the parsed arrays and not from generic text.

Mock-interview score (documented, deterministic). With C correct answers among N = 20 questions:

$$S_{mock} = \text{round}(100 \cdot C / N) \quad (3)$$

Job-description match (documented intent). The seminar report defines the match percentage from the sets $K_{matched}$ of resume-matched keywords and $K_{JD}$ of keywords extracted from the job description:

$$M = 100 \cdot |K_{matched}| / |K_{JD}| \quad (4)$$

The notes instead describe the score as returned by the model, and the observations in Section V-B are not consistent with Eq. (4); the equation is therefore reported as the documented intent only. The dashboard aggregates skills and detected roles with a frequency tally over the user's resumes (time O(nk) for n resumes with k skills on average) and shows the ten most frequent skills.

D. Proposed Evaluation Methodology

Table IV defines measures for the aspects that remain untested. They are proposals and none has been applied. A test set of consented, anonymised resumes of varied layouts and formats, with paired job descriptions and expert annotations, should be assembled and its size and provenance reported. Repeatability would be assessed by running each resume-role pair r times (for example r = 5) with the generation settings recorded, and reporting the mean, standard deviation and range of scores. Agreement with expert ratings would be summarised by the Spearman rank correlation

$$\rho = 1 - 6\Sigma / (n \, (n^2 - 1)) \quad (5)$$

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

where ($d_k$) is the difference between the two ranks for resume (k), and (n) is the number of resumes. The displayed formula applies when there are no tied ranks; if ties occur, a tie-corrected calculation of Spearman's rank correlation should be used. For job-description matching, an internal-consistency check compares the reported score with (100m/(m+u)), where (m) and (u) are the displayed matched and missing keyword counts, respectively, alongside precision and recall against expert-annotated lists. For fairness, paired resumes differing only in name, gender-coded wording, or institution would be scored, and the score differences reported. Each run should log whether the result came from the LLM or a fallback, together with the model identifier and prompt version, to preserve provenance.

TABLE IV PROPOSED EVALUATION MEASURES (NOT YET APPLIED)

| Module or aspect | Proposed measure | Procedure |
|---|---|---|
| Resume parsing | Section-segmentation precision and recall; contact-field accuracy | Compare parser output with manual annotation on resumes of varied layouts |
| ATS scoring | Repeatability (SD and range over r runs); Spearman ρ with expert ratings; provenance (LLM or fallback) | Re-run identical inputs; collect independent expert ratings; log source of each score |
| Resume enhancement | Factual invariance: share of rewrites adding tools, skills, metrics or credentials absent from the source | Manual audit of rewrites against the original resume |
| Job-description matching | Precision and recall of matched and missing keywords; reproducibility; score-count consistency | Annotated resume-job pairs; repeated runs; check reported score against displayed counts |
| Job search | Result availability; precision at k judged manually; latency; failure rate | Fixed set of roles and cities; record outcomes and request times |
| Mock interview | Question relevance; answer-key correctness; explanation accuracy | Expert review of generated sessions |
| Fairness | Score difference between counterfactual resume pairs | Paired resumes differing only in protected-attribute proxies |
| Latency and robustness | End-to-end response time; rejection of invalid tokens; validator and rate-limit behaviour | Repeated timed requests; scripted requests to protected routes |
| Usability and trust | Task completion and perceived usefulness and clarity of explanations | Small user study with consent; instruments to be fixed beforehand |

The evaluations listed in Table IV are proposed procedures and have not been reported as completed experiments.

## V. RESULTS AND DISCUSSION

No systematic experiments have been conducted. This section presents the implementation status (Section V-A), single-run observations and consistency checks from a deployed instance (Section V-B), repeated attempts recorded in a further exploratory session (Section V-C), discussion (Section V-D) and limitations (Section V-E). Evidence is labelled as follows: D (documented) refers to information in the authors' notes or technical documentation; O (observed) refers to outputs visible in the supplied screenshots or recorded in the verification log; and U (unverified) indicates claims that are not demonstrated or conflict across sources.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

A. Implementation Status

TABLE V FUNCTIONS OF CAREERCRAFT AI AND THE EVIDENCE SUPPORTING THEM

| Function | Evidence | Status |
|---|---|---|
| Authentication (e-mail and Google sign-in) | D; O (signed-in session visible) | Session observed |
| Resume upload and rule-based parsing | D; O (parsed name and contact shown on analysis page) | Parsing output observed; discrepancy reported in Section V-C |
| ATS-style analysis with category breakdown | D; O (Fig. 3) | Output observed |
| Resume enhancement, editing and PDF export | D; O (Fig. 4; download control visible) | Enhancement observed; output varied between runs (Section V-C) |
| Job-description matching | D; O (Fig. 5) | Output observed; failed and completed results alternated (Section V-C) |
| Self-introduction generator | D (example in technical documentation); no screenshot supplied | Documented example only |
| Mock interview (20 multiple-choice questions) | D; O (Fig. 6) | Question interface observed |
| Job search with optional city filter | D; O (Fig. 7); provider named in notes only | Listings and links observed |

B. Single-Run Observations and Consistency Checks

Figs. 3–7 show selected outputs from a deployed instance of CareerCraft AI. Personal information was reviewed before publication, and the screenshots are presented for illustrative purposes. These examples are not validated performance results and do not establish the system's accuracy, repeatability, reliability, or effectiveness.

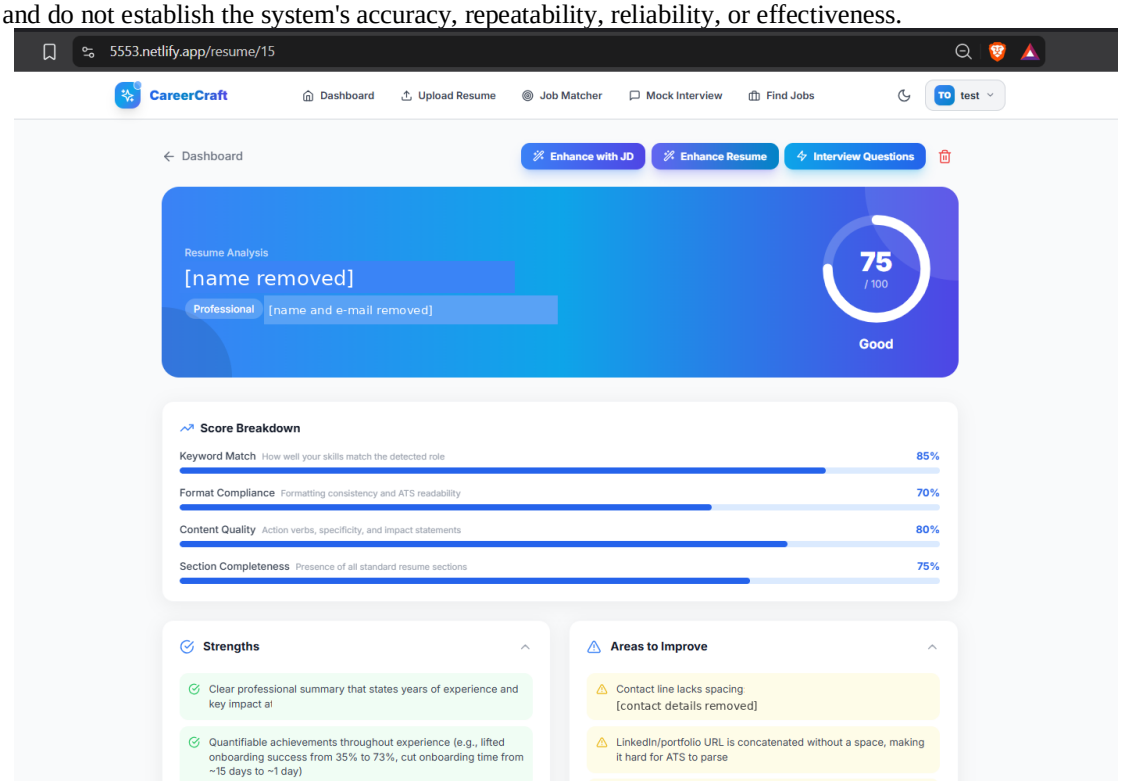



Fig. 3 Resume analysis screen with identifying details removed: overall score, four category bars, and the first strengths and areas to improve.

 This work is licensed under a Creative Commons Attribution 4.0 International License 

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

1) ATS Analysis: The page shows an overall score of 75/100, with category scores of 85, 70, 80 and 75 for Keyword Match, Format Compliance, Content Quality and Section Completeness, respectively (Fig. 3). The unweighted mean of these category scores is 77.5, which does not equal the displayed overall score of 75. The relationship between the overall score and category scores is therefore unresolved. The screenshot also displays strengths and areas for improvement, but it does not establish whether the score was produced by the LLM or by a rule-based fallback. The figure demonstrates the interface and does not establish scoring accuracy.

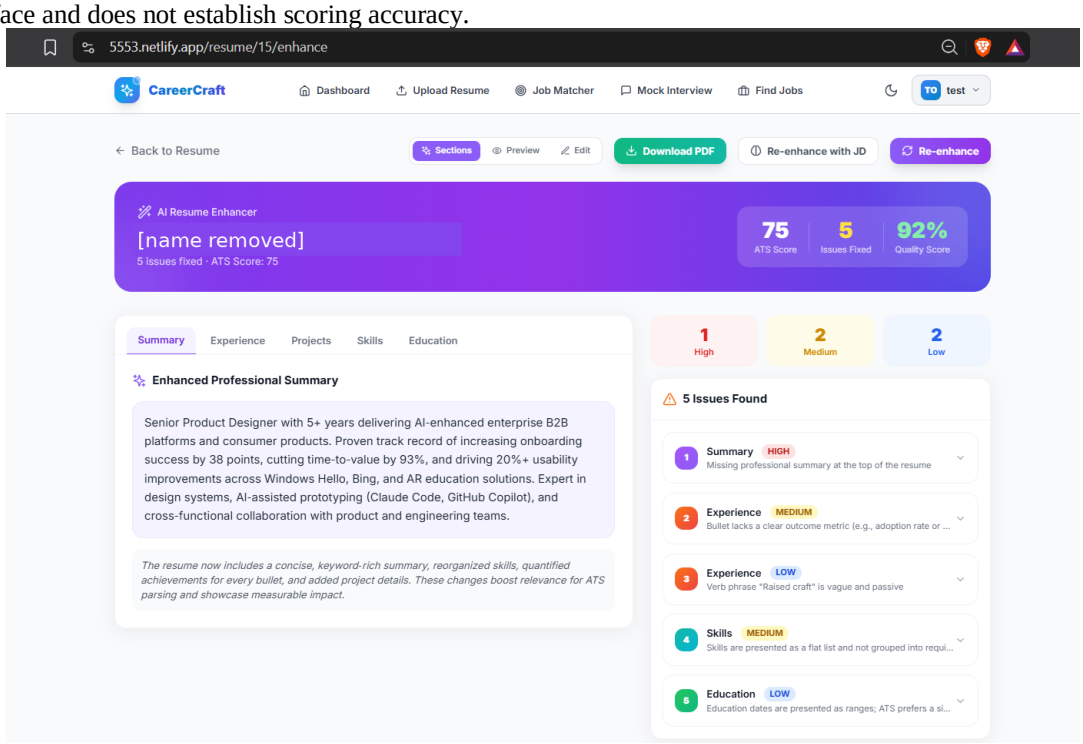



Fig. 4 Resume enhancement screen with identifying details removed: ATS score, number of issues fixed, a “Quality Score”, the enhanced professional summary and a severity-rated issue list.

2) Resume Enhancement: The page displays an ATS score of 75, five issues fixed and a “Quality Score” of 92% (Fig. 4). The issue counts comprise one high-, two medium- and two low-severity items, which sum to five. The basis of the Quality Score is not defined in the available documentation. The enhanced summary contains quantified claims, including a 38-point increase in onboarding success and a 93% reduction in time-to-value. Their support in the source resume must be checked before they are treated as factual improvements. The screenshot demonstrates the enhancement interface, but does not establish the factual accuracy or completeness of the generated content.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

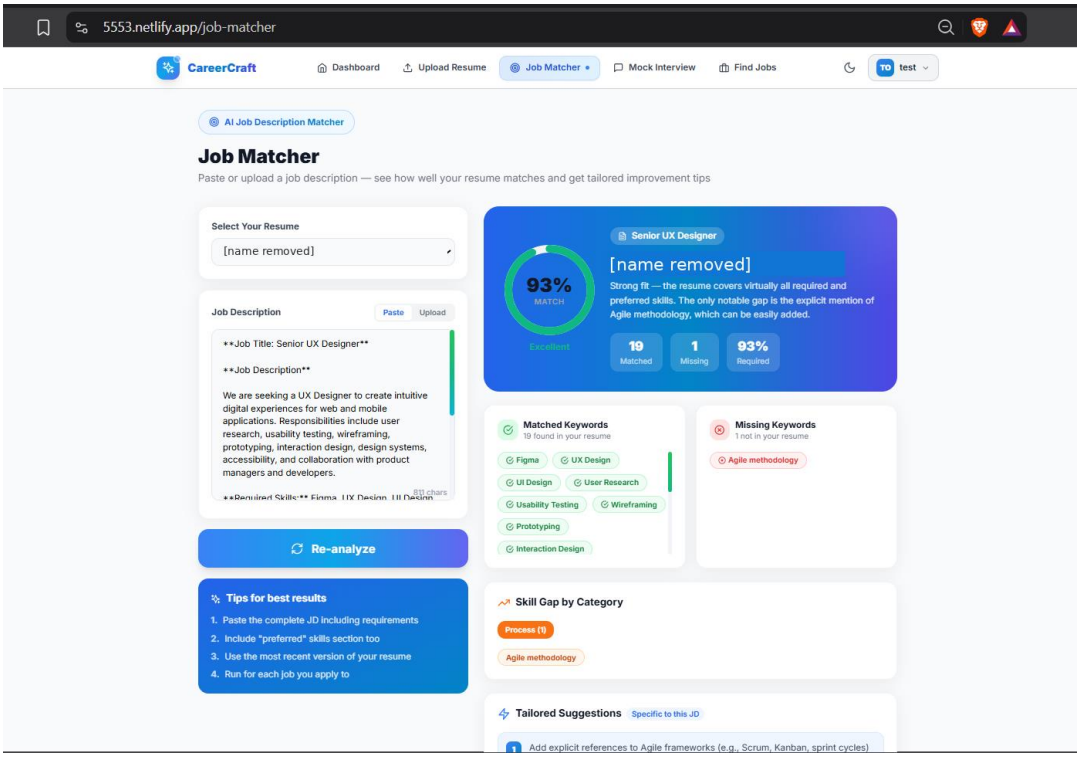



Fig. 5 Job-description matching interface (identifying details removed) showing a 93% match score, 19 matched keywords, one missing keyword and skill-gap recommendations for a Senior UX Designer role.

3) Job-Description Matching: The screenshot shows a 93% match for a Senior UX Designer role, with 19 matched keywords and one missing keyword (Fig. 5). Based on these counts, the matched-keyword proportion is 19/ (19 + 1) × 100 = 95%, which differs from the displayed score by two percentage points. The screenshot also shows a skill-gap category for the missing requirement, "Agile methodology," and tailored suggestions. The calculation used to produce the displayed score remains unverified; therefore, this result demonstrates the interface output, not validated matching accuracy.

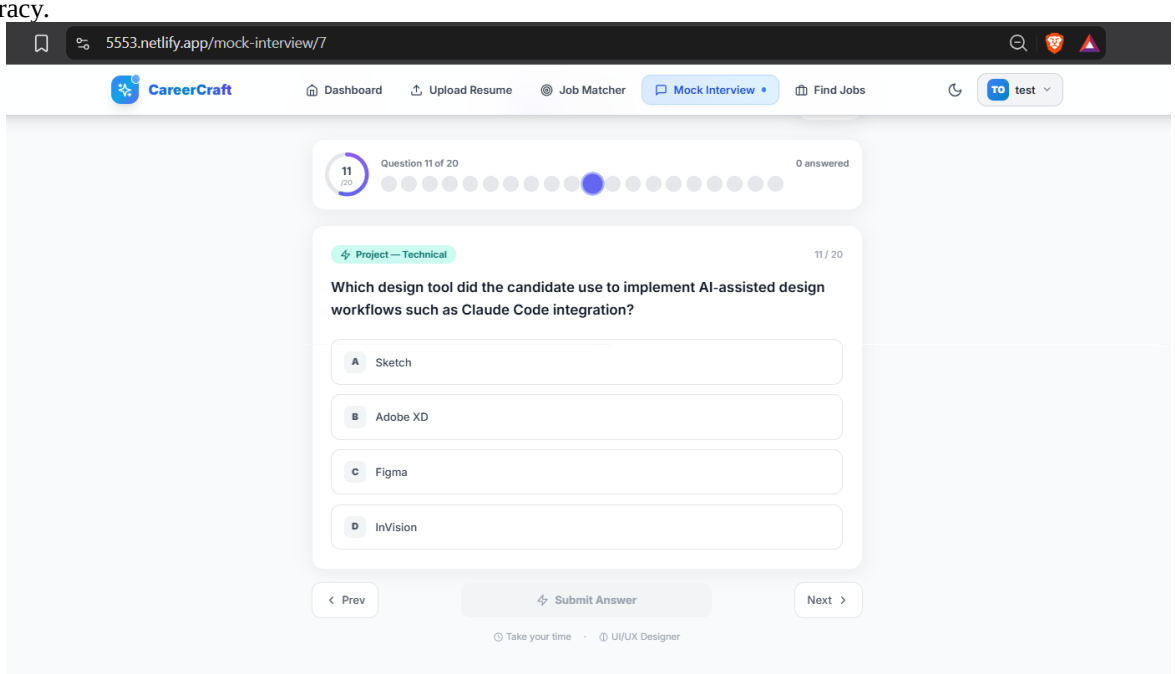



Fig. 6 Mock-interview interface showing question 11 of 20, the question category, four answer options and navigation controls for the UI/UX Designer role.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

4) Mock Interview: The screen shows question 11 of 20, a “Project-Technical” category badge, four answer options and the Prev, Submit Answer and Next controls (Fig. 6), consistent with the multiple-choice flow documented in the implementation notes and different from the seminar report's description of typed answers. The stored answer key is not visible in the screenshot. According to the implementation notes, the key is generated by the LLM when the session is created and the selected answer is later evaluated by comparing answer letters. The correctness of the generated answer key has not been independently verified.

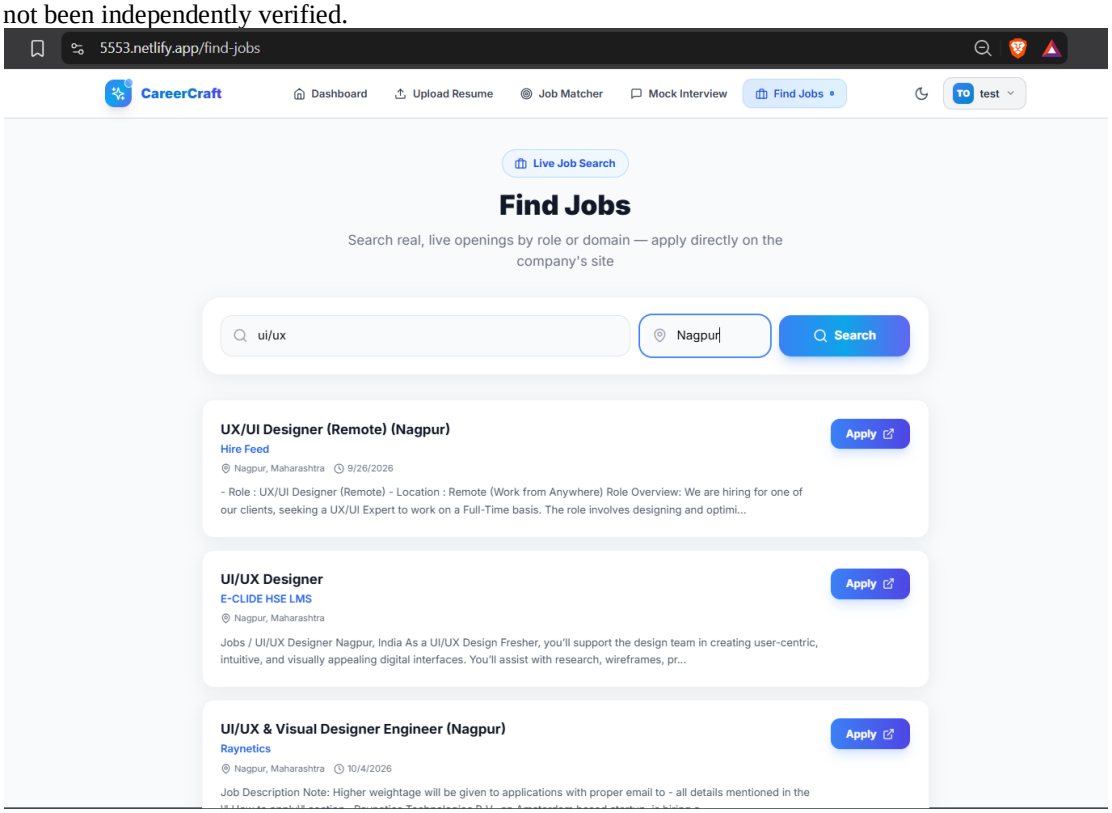



Fig. 7 Job-search interface displaying job listings for the queried role, including job titles, company names, locations, posting dates and application links.

5) Job Search: The page displays job listings with titles, company names, locations, posting dates and Apply controls that open external links (Fig. 7). The screenshot demonstrates the job-search and redirection workflow, but does not identify the listing provider or establish the authenticity, freshness or relevance of the results. The displayed listings illustrate that keyword-based searches can return roles whose technical specialisation may differ from the user's intended meaning. A single screenshot of one search cannot establish search precision or overall relevance, and no resume-based ranking is demonstrated.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.151099

TABLE VI REPORTED OUTPUTS AND INTERNAL-CONSISTENCY CHECKS (SINGLE RUNS; NOT VALIDATION)

| Output | Displayed or reported values | Check | Result |
|---|---|---|---|
| ATS score (Fig. 3) | Overall score: 75; category scores: 85, 70, 80, 75 | Mean of category scores = 77.5 | Overall score differs from the category mean by 2.5 points; scoring formula requires clarification. |
| Enhancement issues (Fig. 4) | 5 issues: 1 high, 2 medium, 2 low | 1 + 2 + 2 = 5 | Issue counts are consistent; the displayed Quality Score of 92% is undefined. |
| Job-description match (Fig. 5) | 93% match; 19 matched keywords; 1 missing keyword | 100*19/(19+1)=95% | Displayed score differs from the keyword proportion by 2 percentage points. |
| Mock interview (Fig. 6) | Question 11 of 20; Project-Technical category; four answer options | Visual inspection of the displayed interface | Multiple-choice question and navigation interface demonstrated. |
| Self-introduction (technical documentation) | Approximately 173 words | Word count reported in documentation | Output length is documented; latency and variability were not measured. |
| Job search (Fig. 7) | Job listings, company names, locations, posting dates and application links | Listing provider is not identifiable from the screenshot alone | Search interface and application-link workflow are demonstrated; listing freshness and relevance were not independently measured. |
| JD match, further session (Section V-C) | 78%; 21 matched; 8 missing | 100·21/(21 + 8) = 72.4% | Displayed score is 5.6 points higher; formula unknown. |
| JD match, failed attempts (Section V-C) | 50%; 1 matched; 0 missing; “analysis unavailable” | 100·1/(1 + 0) = 100% | Not a completed analysis; does not match the displayed 50%. |
| Mock-interview score (Section V-C) | 15 of 20 correct; 75% | Eq. (3): round(100·15/20) = 75 | Consistent with Eq. (3). |

C. Repeated Attempts in a Further Exploratory Session

A separate verification log dated 21 October 2026 records a further, uncontrolled exploratory session using one one-page test resume and exercise-specific job descriptions. It was not designed as a benchmark or repeatability study, and the raw response from the successful matching run was not captured. Table VI summarises the reported outputs and the internal-consistency checks; these observations should not be interpreted as validated performance results.

1) Resume Parsing: Although the test resume listed multiple skills, the Skills panel displayed only “git”, while Education and Experience contained fragmented text. Re-uploading the same PDF reproduced the display. The cause and downstream effects were not established.

2) Job-Description Matching: Of four attempts using the same resume and Junior UI/UX Designer description, three displayed “analysis unavailable” alongside a 50% score and one matched keyword; one attempt completed with 78%, 21 matched keywords and eight missing keywords. The completed count-based proportion is 72.4%, while the failed display's one matched and zero missing keywords would imply 100% under that simple formula. Because the deployed scoring formula and the raw successful response are unavailable, these are unresolved consistency questions, not proof of a specific implementation defect; failed displays are not valid match results.

3) Resume Enhancement: Across recorded runs, the ATS score and issue count remained at 75 and five, while the displayed quality score varied (85%, 92%, and 86%) and generated content changed. A tailored output included quantitative claims not traceable to the source resume, including 15,000 users within three months, a 25% reduction in message-retrieval time, and an 18% reduction in onboarding time. The exported PDF also omitted some source details and links and changed the displayed email address. These are unsupported additions or omissions in the inspected output; they were not established as false, and generated content requires source-by-source review before use.



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026
DOI: 10.17148/IJARCCE.2026.151099

4) Mock Interview and Job Search: One 20-question session produced 15 correct answers (75%), consistent with Eq. (3). Review of selected questions found equivalent answer options marked differently and questions whose assumptions or risks were insufficiently stated; the full question bank was not audited. Searches for Python Developer and Data Analyst roles in Nagpur returned listings and external application links, including a portal warning that one position might have closed. The observations demonstrate the search-and-redirect workflow only, not listing authenticity, freshness, or relevance.

### D. Discussion

The observations identify issues for further investigation rather than establish system-wide failure. The displayed ATS overall score (75) differs from the mean of its four category scores (77.5), and the "Quality Score" has no documented calculation. The job-matching score (93%) also differs from the 95% proportion implied by 19 matched and one missing keyword in the illustrated case. The scoring formulas and score provenance therefore require clarification.

The interface screenshots demonstrate selected workflows, but they do not establish parsing accuracy, the factual reliability of enhanced resumes, the correctness of generated interview answer keys, or the freshness and relevance of job listings. In the further exploratory session, matching failures were accompanied by default-looking score and keyword values, enhancement runs varied and omitted some source details, and selected quiz items had ambiguous or under-specified options.

These findings are illustrative and are based on a small number of examples without controlled experiments or systematic repeated testing. Future evaluation should assess parsing accuracy, score consistency, factual invariance, matching quality, answer-key correctness, job relevance, fairness, security, and usability before claims about effectiveness can be made.

Practical improvements suggested by these observations include withholding match scores when analysis fails, calculating count-based match percentages programmatically where that definition is intended, recording whether outputs came from the LLM or a fallback, comparing rewritten content with the source before export, and reviewing generated answer keys. These are recommendations for future work; their implementation and effects have not been tested.

### E. Limitations

This study is limited by the absence of systematic experiments and the use of a small number of illustrative screenshots. The accuracy of resume parsing, validity and repeatability of generated scores, factual reliability of enhanced content, job-matching performance, and correctness of mock-interview answer keys remain unverified. The study also does not establish fairness across user groups, security through independent auditing, or the freshness and relevance of job listings. Consequently, the findings should be interpreted as exploratory observations rather than evidence of validated system performance.

Implementation details are based on notes and technical documentation rather than source-code verification. The raw response for the successful matching run was not captured, and results from the further session did not reproduce the illustrated matching result. These constraints limit reproducibility and the strength of conclusions.

## VI. Conclusion and Future Scope

CareerCraft AI is an integrated, candidate-facing prototype combining rule-based resume parsing with LLM-based analysis, resume enhancement, job-description matching, self-introduction and interview-question generation, and job search through an external listing API. This paper documents the architecture, data model, processing workflow, and selected observations from a deployed instance. The examples demonstrate interface functionality, while internal-consistency checks and exploratory testing identify questions about score calculation, parsing, generated-content fidelity, and quiz quality. The study does not establish the system's accuracy, repeatability, fairness, or practical effectiveness.

Future work should first implement the evaluation protocol proposed in Table IV and resolve outstanding implementation questions, particularly the calculation and provenance of ATS and job-match scores. Score provenance should be logged and displayed, while matched and missing keyword counts and match percentages should be calculated programmatically to ensure reproducibility. Further improvements include embedding-based or hybrid lexical-semantic matching [7], [10], ranking job listings against resume skills, persisting job matches and search history, adding a guided resume builder, supporting additional languages and image-based resumes through OCR, and evaluating free-text interview answers. Usability and user trust should also be studied with candidates. Any recruiter-facing extension would introduce additional fairness and transparency concerns, requiring separate evaluation.

## References

[1] J. B. Fuller, M. Raman, E. Sage-Gavin and K. Hines, "Hidden Workers: Untapped Talent," Harvard Business School Project on Managing the Future of Work and Accenture, Sep. 2021 (industry report; not peer reviewed).

DOI: 10.17148/IJARCCE.2026.151099

[2] K. Nguyen, M. Zhang, S. Montariol and A. Bosselut, "Rethinking Skill Extraction in the Job Market Domain using Large Language Models," in Proc. First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024), St. Julian's, Malta, Mar. 2024, pp. 27–42, doi: 10.18653/v1/2024.nlp4hr-1.3.

[3] K. Khelkhal and D. Lanasri, "Smart-Hiring: An Explainable End-to-End Pipeline for CV Information Extraction and Job Matching," arXiv:2511.02537, Nov. 2025 (preprint).

[4] S. Vaishampayan, H. Leary, Y. B. Alebachew, L. Hickman, B. Stevenor, W. Beck and C. Brown, "Human and LLM-Based Resume Matching: An Observational Study," in Findings of the Association for Computational Linguistics: NAACL 2025, Albuquerque, NM, USA, Apr. 2025, pp. 4823–4838, doi: 10.18653/v1/2025.findings-naacl.270.

[5] L. Zheng, W.-L. Chiang, Y. Sheng, S. Zhuang, Z. Wu, Y. Zhuang, Z. Lin, Z. Li, D. Li, E. P. Xing, H. Zhang, J. E. Gonzalez and I. Stoica, "Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena," in Advances in Neural Information Processing Systems 36 (NeurIPS 2023), Datasets and Benchmarks Track, 2023, arXiv:2306.05685.

[6] A. A. Amir Hussin, S. A. Suparman and H. S. Rushdan, "An Intelligent NLP-Based System for Resume Screening and Applicant Tracking Systems (ATS)-Optimised Resume Generation," International Journal on Perceptive and Cognitive Computing, vol. 12, no. 2, pp. 129–134, 2026, doi: 10.31436/ijpcc.v12i2.719.

[7] M. H. Ajjam and H. S. Al-Raweshidy, "AI-Driven Semantic Similarity-Based Job Matching Framework for Recruitment Systems," Information Sciences, vol. 724, art. no. 122728, 2026 (available online 2025), doi: 10.1016/j.ins.2025.122728.

[8] J. Zhao, J. Wang, M. Sigdel, B. Zhang, P. Hoang, M. Liu and M. Korayem, "Embedding-Based Recommender System for Job to Candidate Matching on Scale," arXiv:2107.00221, 2021 (preprint).

[9] G. Bied, S. Nathan, E. Perennes, M. Hoffmann, P. Caillou, B. Crépon, C. Gaillac and M. Sebag, "Toward Job Recommendation for All," in Proc. 32nd Int. Joint Conf. on Artificial Intelligence (IJCAI-23), AI for Good Track, 2023, pp. 5906–5914, doi: 10.24963/ijcai.2023/655.

[10] H. A. Al Awad and K. F. Omar, "Developing an Intelligent Job Recommendation System Using Semantic Retrieval and Explainable AI Techniques," arXiv:2605.27656, May 2026 (preprint).

[11] M. Vyas, A. Chakraborty and V. Gupta, "JobMatchAI – An Intelligent Job Matching Platform Using Knowledge Graphs, Semantic Search and Explainable AI," in Proc. 64th Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations), San Diego, CA, USA, Jul. 2026, pp. 526–535, doi: 10.18653/v1/2026.acl-demo.52.

[12] H.-Y. Suen and K.-E. Hung, "Building Trust in Automatic Video Interviews Using Various AI Interfaces: Tangibility, Immediacy, and Transparency," Computers in Human Behavior, vol. 143, art. no. 107713, 2023, doi: 10.1016/j.chb.2023.107713.

[13] 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 (FAccT '20), 2020, pp. 469–481, doi: 10.1145/3351095.3372828.

[14] D. Kumar, T. Grosz, N. Rekabsaz, E. Greif and M. Schedl, "Fairness of Recommender Systems in the Recruitment Domain: An Analysis from Technical and Legal Perspectives," Frontiers in Big Data, vol. 6, art. no. 1245198, 2023, doi: 10.3389/fdata.2023.1245198.

[15] D. F. Mujtaba and N. R. Mahapatra, "Fairness in AI-Driven Recruitment: Challenges, Metrics, Methods, and Future Directions," arXiv:2405.19699, 2024 (preprint).