ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

AUTOMATED PROFESSIONAL DOCUMENT GENERATION USING LARGE LANGUAGE MODELS FOR CAREER APPLICATIONS

C. Sreerag
Department of Master of Computer Applications Progressive Education Society's Modern College of Engineering Pune, India

Article DOI: https://doi.org/10.36713/epra27875
DOI No: 10.36713/epra27875

ABSTRACT

The widespread adoption of Applicant Tracking Systems (ATS) in modern recruitment has created a structural disadvantage for job seekers who lack professional writing expertise or familiarity with algorithmic screening criteria. While automated hiring tools have significantly improved the effi-ciency of candidate evaluation from the employer's perspective, applicants continue to depend on static resume builders that provide formatting assistance without engaging with content quality. This paper proposes a conceptual framework for an AI-assisted resume generation platform that leverages the generative capabilities of Large Language Models (LLMs) to transform unstructured career data into structured, ATS-compatible pro-fessional documents. The proposed architecture integrates a structured user input layer, a role-specific prompt engineering pipeline, and a semantic document rendering engine. Unlike the prevailing focus of recruitment AI research – which centers on automated candidate screening – this work explicitly addresses the applicant's document preparation challenge. The framework aims to reduce the cognitive effort associated with resume writing while improving structural consistency, semantic alignment with job descriptions, and overall document quality. Empirical evalu-ation through prototype implementation constitutes the primary direction of future work.

INDEX TERMS – *Large Language Models, AI-Driven Resume Generation, Applicant Tracking Systems (ATS), Prompt Engi-neering, Natural Language Generation, Career Automation.*

I. INTRODUCTION

The digitization of recruitment processes over the past decade has fundamentally altered how organizations iden-tify and evaluate potential candidates. As the volume of applications for individual job postings continues to grow, employers have increasingly turned to Applicant Tracking Systems to manage initial screening at scale. These systems parse and rank incoming resumes based on structural con-ventions, keyword density, and semantic relevance to job descriptions—often before any human reviewer examines a submission. As a consequence, a candidate's first hurdle is no longer a recruiter but an algorithm.

This shift has placed an undue burden on applicants. A technically qualified individual may be systematically elimi-nated not because of insufficient experience, but because their resume does not conform to the formatting expectations or vocabulary patterns that a particular ATS is configured to recognize. Weak achievement statements, non-standard section headings, and absent role-specific terminology are among the most common and avoidable causes of automatic rejection.

Despite the availability of numerous online resume building platforms, these tools address only the presentational dimen-sion of the problem. They offer visual templates and layout controls, but the responsibility for generating substantive, targeted content remains entirely with the user. The cognitive effort required to translate raw professional experience into compelling, ATS-optimized prose is precisely what these tools leave unaddressed.

Recent developments in Large Language Models have demonstrated that generative AI is capable of nuanced content creation tasks, including contextual paraphrasing, structured text synthesis, and style-adapted writing. These capabilities present a compelling opportunity to reimagine resume creation as an AI-assisted process rather than a manual one. This paper introduces a conceptual framework for such a system—one that accepts structured career data as input and produces pro-fessionally written, ATS-compatible resume content as output.

The primary objectives of this research are:

- To propose an architecture for the end-to-end automated generation of professional, role-tailored resumes using generative AI.
- To design a prompt engineering methodology that ensures high-quality, factually grounded resume content while minimizing the risk of AI hallucination.
- To specify a document rendering pipeline that produces semantically structured, machine-readable PDF outputs compatible with standard ATS parsers.
- To outline a privacy-conscious data model that limits the exposure of personally identifiable information during AI inference.

II. RELATED WORK

The application of AI to human resources and career ser-vices has attracted considerable research interest. A systematic review of the existing literature reveals, however, that the vast majority of contributions are oriented toward employer-side automation, leaving applicant-facing tooling as a compara-tively underexplored domain.

A. Automated Screening and Scoring

A substantial portion of recent research investigates the use of language models to automate candidate evaluation work-flows. Gan et al. [1] proposed a multi-agent LLM architecture in which discrete agents handle document summarization, qualification extraction, and candidate ranking in parallel, enabling scalable processing of large applicant pools. Comple-menting this, Lo et al. [2] developed an explainable screening framework that generates structured justifications alongside ranking decisions, specifically to address concerns around transparency and fairness in algorithmic hiring. Both systems are designed to serve the recruiter's evaluation workflow and do not offer any mechanism through which applicants can improve or adapt their submitted documents.

B. Semantic Parsing and Skills Standardization

A parallel research thread concerns the normalization of resume content into machine-interpretable schemas. Sinha et al. [5] conducted a comprehensive review of foundational NLP approaches to resume parsing, identifying the limitations of rule-based methods in handling the structural heterogeneity common across real-world documents. More recently, Walid et al. [4] demonstrated that fine-tuned LLM ensembles substan-tially outperform single-model baselines when mapping ex-tracted skills to standardized occupational taxonomies. These parsing systems, while technically sophisticated, function as components within employer-side data pipelines rather than tools accessible to individual applicants.

C. Applicant-Facing Resume Generation

Research specifically targeting the applicant's document creation process remains comparatively limited. Zinjad et al. [3] presented a pipeline that employs LLMs to tailor existing resume content to specific job descriptions. While this work is the closest in spirit to the proposed framework, it addresses only the personalization of pre-existing resumes rather than the generation of structured professional content from raw, unformatted career inputs. It also does not consider the rendering requirements necessary for ATS compatibility.

D. Research Gap

The reviewed literature collectively demonstrates a pro-nounced asymmetry: substantial investment has been directed at helping employers evaluate candidates more efficiently, while the tools available to applicants have remained largely unchanged. Current resume building platforms provide struc-tural scaffolding and visual formatting, but they do not engage in content generation. The proposed framework addresses this gap by deploying LLM-driven natural language generation on the applicant's behalf—transforming raw career information into structured, high-quality, ATS-optimized professional doc-uments.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

The proposed system is conceptualized as a decoupled, layered web application in which each functional concern is handled by a dedicated architectural component. This separa-tion allows individual layers to be independently maintained, scaled, or replaced without disrupting the overall pipeline.

A. High-Level Architecture

The system comprises four primary layers, illustrated in Fig. 1.

1) **Client Layer (Frontend):** A structured input interface through which applicants submit career information in discrete, typed fields—job titles, organizational affili-ations, employment dates, and raw responsibility de-scriptions. A real-time document preview panel allows applicants to observe formatting and content changes before export.
2) **Service Layer (Backend):** A RESTful API server that manages authentication, session state, and relational data storage. This layer acts as a controlled gateway between the frontend interface and the AI inference engine, enforcing access controls, applying privacy filters, and assembling structured prompts from validated user in-puts before any data is forwarded for AI processing.
3) **Inference Layer (AI Engine):** The core genera-tive component of the system. This layer interfaces with a foundational LLM API, submitting structured prompts that incorporate the applicant's career data alongside role-specific engineering directives. The en-gine returns professionally formatted resume con-tent—including achievement-oriented bullet points and contextual professional summaries—tailored to the ap-plicant's target role.
4) **Rendering Engine:** A dedicated microservice that com-piles the finalized resume content into semantic HTML and converts it to a downloadable PDF via a headless browser-based rendering pipeline. This component is specifically designed to produce DOM-based output rather than canvas-captured images, ensuring that the exported document retains a fully machine-readable text layer compatible with standard ATS parsers.

B. Prompt Engineering Strategy

The quality and reliability of LLM-generated resume con-tent is highly sensitive to the design of the input prompts. The proposed inference layer employs a structured, multi-strategy prompting methodology informed by established practices in the literature [6], [7]:

- **Role Prompting:** Each prompt initializes the model with an expert persona—that of a senior professional resume writer with domain knowledge relevant to the applicant's target industry. This persona conditioning consistently produces output that is more formally structured and professionally appropriate than zero-shot generation.
- **Few-Shot Exemplar Injection:** Prompts incorporate curated examples demonstrating the transformation of vague, passive responsibility statements into quantifiable, action-oriented achievement descriptions. These exem-plars follow the Action Verb – Task Scope – Measurable Outcome structure widely recommended in professional career development practice.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

- **Constraint Directives:** To mitigate the risk of AI hallucination—a well-documented limitation of generative language models [8]—prompts explicitly instruct the model to base all generated content exclusively on the information supplied by the applicant, to avoid fabricating metrics or credentials, and to adhere to specified length constraints for each resume element.

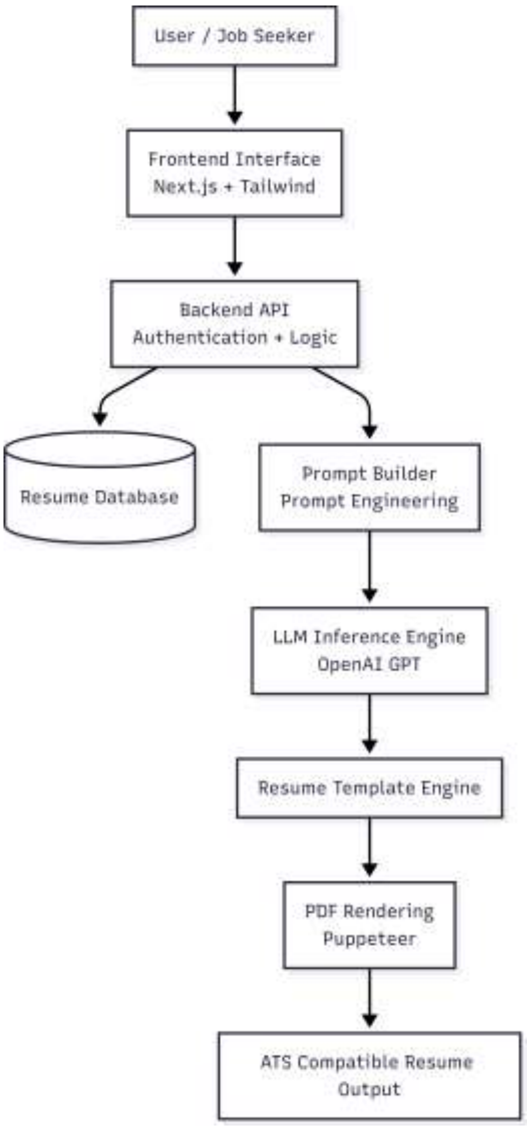

Fig. 1. Layered Architecture of the Proposed AI-Assisted Resume Generation Framework.

C. Data Schema and Privacy Design

Privacy is a foundational design consideration given that applicants must share sensitive career information to use the system effectively. The proposed data schema addresses this through several mechanisms:

- User authentication credentials are stored independently from resume content records, ensuring that career data is not directly linked to account identifiers in the primary database.

All data fields containing personally identifiable informa-tion (PII) are encrypted at rest using industry-standard symmetric encryption.

- Prior to transmission to the LLM inference endpoint, the service layer applies a sanitization pass that removes or pseudonymizes direct identifiers—including full names, contact details, and institutional references—replacing them with structured placeholders that are reinserted into the final document post-generation.

The complete resume generation workflow is illustrated in Fig. 2.

IV. Implementation Challenges and Solutions

Realizing the proposed architecture as a deployable system introduces several non-trivial engineering considerations. Two of the most significant are addressed below.

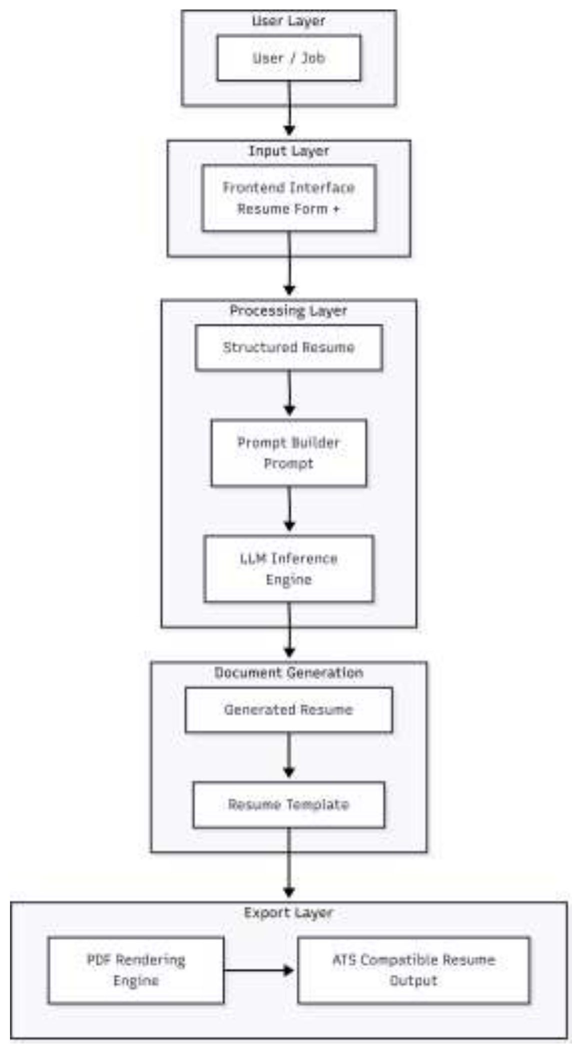

Fig. 2. LLM-Based Resume Generation Pipeline.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 12| Issue: 5| May 2026|| Journal DOI: 10.36713/epra2013 ||SJIF Impact Factor 2026: 8.753 || ISI Value: 1.188

Generation Latency: LLM inference introduces variable response latency that, if unmanaged, would result in a poor user experience. Rather than blocking the interface until the full response is available, the proposed system employs a server-sent events (SSE) streaming approach, delivering gen-erated tokens to the client interface incrementally as they are produced. This significantly reduces perceived wait time and provides the applicant with immediate, progressive feedback during content generation.

ATS-Compatible PDF Rendering: A common failure mode in existing resume tools is the generation of PDF files by capturing rendered HTML on a canvas element. While visually accurate, such outputs are image-based and contain no selectable or machine-readable text, rendering them effectively invisible to ATS parsers. The proposed rendering engine avoids this by constructing resume content exclusively from semantic HTML elements—standard paragraph, heading, and list tags—and converting the resulting document object model to PDF through a headless browser pipeline. This approach ensures that the exported file retains a fully accessible, structured text layer parseable by any standards-compliant ATS.

V. FUTURE EVALUATION METHODOLOGY

As the proposed system represents a conceptual architecture, empirical validation is deferred to a subsequent implementa-tion phase. The planned evaluation protocol will assess the system across five dimensions, summarized in Table I.

TABLE I
PROPOSED EVALUATION METRICS FOR THE AI-ASSISTED RESUME GENERATION SYSTEM

Metric	Description	Measurement Method
ATS Compat-ibility	Resume parse success rate across multiple screening engines	Open-source ATS parser and resume screening li-braries
Content Qual-ity	Grammatical correctness, professional tone, and achievement clarity	Blind human evaluation panel
User Effort	Time-on-task and num-ber of manual edits re-quired post-generation	Controlled user study
Generation Time	End-to-end resume generation latency per section	Instrumented server-side logging
User Satisfac-tion	Perceived usefulness and overall experience	Post-session structured survey

Evaluation participants will be recruited from three popu-lations: recent graduates with limited resume writing expe-rience, mid-career professionals seeking role transitions, and career services practitioners capable of providing expert-level assessments of document quality. Comparative baselines will include manually authored resumes and outputs from estab-lished template-based resume builders, providing a meaningful reference point against which the system's contribution can be measured.

VI. LIMITATIONS AND ETHICAL CONSIDERATIONS

Deploying generative AI in the context of professional career documentation introduces specific risks that must be acknowledged and incorporated into the system design.

Hallucination and Factual Accuracy: Despite the con-straint directives embedded in the prompting strategy, large language models retain a non-trivial probability of generating plausible but factually inaccurate content [8]. In a resume context, fabricated metrics or embellished credentials could have serious professional and reputational consequences for the applicant. The system interface must therefore present all AI-generated content as a draft subject to mandatory user review, rather than a final deliverable.

Bias in Generated Content: LLMs are trained on large corpora that reflect historical patterns in professional writing. These patterns are known to encode biases related to gender, ethnicity, and occupational domain, which may manifest subtly in the vocabulary or framing used to describe certain roles or professional backgrounds. Addressing this will require systematic bias auditing during the evaluation phase, with find-ings informing both model selection and any post-generation filtering strategies.

External API Dependency: The proposed architecture's reliance on third-party LLM inference endpoints introduces operational risk. Disruptions to API availability, changes in pricing models, or policy updates from the provider could affect system reliability. Future iterations of the framework should evaluate the viability of self-hosted, open-source lan-guage models as an alternative inference backend—an ap-proach that would simultaneously address API dependency and strengthen the system's data privacy posture.

VII. CONCLUSION AND FUTURE WORK

This paper has presented a conceptual framework for au-tomated professional document generation using Large Lan-guage Models, designed specifically to support job applicants during the resume preparation stage. By integrating structured user input, a multi-strategy prompt engineering pipeline, and a semantically faithful PDF rendering engine, the proposed system aims to reduce the cognitive burden of resume writ-ing while producing ATS-compatible outputs consistent with professional industry standards.

Unlike the predominant focus of existing recruitment AI research—which has concentrated on employer-side candidate screening—this work is oriented toward addressing a gap on the applicant's side of the hiring process. The framework contributes a principled architecture that treats generative AI not as a formatting aid, but as an active collaborator in content creation.

Future work will proceed along three directions. First, a functional prototype will be implemented and subjected to the empirical evaluation protocol described in Section V, with particular attention to ATS compatibility and user effort reduction. Second, the framework will be extended to support multilingual resume generation, broadening its applicability to non-English-speaking professional contexts. Third, the feasibility of replacing the cloud-based inference layer with locally hosted open-source language models will be investigated, with the dual objective of eliminating external API dependency and further strengthening data privacy guarantees.

ACKNOWLEDGMENT

The author sincerely expresses gratitude to Dr. Shivani Budhkar, Department of Master of Computer Applications, Progressive Education Society's Modern College of Engineer-ing, Pune, for her valuable guidance, encouragement, and academic support throughout the development of this research work.

The author also acknowledges the support provided by the Department of Master of Computer Applications at Pro-gressive Education Society's Modern College of Engineering, Pune, for offering the academic environment and resources necessary for completing this research.

REFERENCES

1. *Gan, Q. Zhang, and T. Mori, "Application of LLM agents in recruitment: A novel framework for resume screening," arXiv preprint arXiv:2401.08315, 2024. [Online]. Available: https://arxiv.org/abs/2401. 08315*
2. *P. W. Lo et al., "AI hiring with LLMs: A context-aware and explainable multi-agent framework for resume screening," arXiv preprint arXiv:2504.02870, 2025. [Online]. Available: https://arxiv.org/abs/2504. 02870*
3. *S. B. Zinjad et al., "ResumeFlow: An LLM-facilitated pipeline for personalized resume generation and refinement," in Proc. ACM SIGIR, 2024. [Online]. Available: https://dl.acm.org/doi/10.1145/3626772*
4. *O. Walid et al., "MSLEF: Multi-segment LLM ensemble finetuning in recruitment," arXiv preprint arXiv:2509.06200, 2025. [Online]. Avail-able: https://arxiv.org/abs/2509.06200*
5. *Sinha, M. A. K. Akhtar, and A. Kumar, "Resume screening using natural language processing and machine learning: A systematic review," in Advances in Intelligent Systems and Computing, Springer, 2021, pp. 149–157. [Online]. Available: https://link.springer.com/chapter/10.1007/ 978-981-33-4859-2_21*
6. *T. Brown et al., "Language models are few-shot learners," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 1877–1901. [Online]. Available: https://arxiv.org/abs/2005.14165*
7. *L. Ouyang et al., "Training language models to follow instructions with human feedback," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 35, 2022, pp. 27730–27744. [Online]. Available: https://arxiv.org/abs/2203.02155*
8. *R. Bommasani et al., "On the opportunities and risks of foundation models," arXiv preprint arXiv:2108.07258, 2021. [Online]. Available: https://arxiv.org/abs/2108.07258*
9. *Vaswani et al., "Attention is all you need," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017. [Online]. Available: https://arxiv.org/abs/1706.03762*
10. *Radford et al.,"Language models are unsupervised multitask learners," Technical Report,2019. [Online]. Available: https://cdn.openai.com/better-language-models/language_models_are_ unsupervised_multitask_learners.pdf*