AICA *(AI CAREER ASSISTANT)*: A JOB MATCHING PLATFORM USING LARGE LANGUAGE MODEL WITH RETRIEVAL-AUGMENTED GENERATION

A Thesis Proposal Presented to
The Faculty of the College of Engineering and Computing Studies
University of St. La Salle Bacolod

In Partial Fulfillment of
the Requirements for the Degree
Bachelor of Science in Computer Science

GAMBOA, APRIL FAITH J.
HIDALGO, JANPOL S.
MAHANDOG, HEIDINE MARIE J.
SANTIA, NATHANIA ELOUISE A.

May 26, 2025

APPROVAL SHEET

The thesis entitled **"AICA *(AI CAREER ASSISTANT)*: A JOB MATCHING PLATFORM USING LARGE LANGUAGE MODEL WITH RETRIVALl-AUGMENTED GENERATION "** presented by **APRIL FAITH GAMBOA, JANPOL HIDALGO, HEIDINE MARIE MAHANDOG, and NATHANIA ELOUISE SANTIA**, in partial fulfilment of the requirements for the degree of **Bachelor of Science in Computer Science** of the University of St. La Salle Undergraduate Programs has been evaluated and approved by the panel of evaluators.

PANEL OF EVALUATORS

JIM JONATHAN DECRIPITO, PhD.
Chair

EDDIE DE PAULA, PhD
Member

EISCHIED ARCENAL, PhD.
Member

ANTHONY J. GALLEGO, PhD.
Adviser

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Background

The widening gap between graduates' skills and industry demands has emerged as a critical issue, especially in the technology sector. Despite the increasing number of tech graduates, many struggle to secure employment that aligns with their training and capabilities (Abramov et al., 2023; Wenjing, 2023). As technology continues to evolve, it reshaps job requirements in ways that often leave fresh graduates without the necessary technical expertise and soft skills required to thrive in modern workplaces. In the Philippines, the implementation of the K–12 curriculum aimed to equip students with more relevant competencies, yet companies still report that many new hires remain underprepared (PIDS, 2023). This disconnect contributes to rising underemployment, slows career progression, and impedes national economic development.

Traditional job portals and recruitment systems have proven inadequate in addressing these challenges, primarily relying on keyword-based matching that fails to capture the semantic relationships between candidate skills and job requirements. Research by Kelley, Ksoll, and Magruder (2022) demonstrated that conventional job portals can even lead to lower employment rates in the short term, as job seekers develop unrealistic expectations while waiting for better offers that never materialize. These platforms often overlook qualified candidates due to differences in terminology while recommending unqualified applicants based on superficial keyword overlap. Furthermore, the disconnect between job location preferences and actual opportunities, combined with reservation wages that exceed market rates, reveals fundamental flaws in how traditional systems facilitate job matching.

The emergence of artificial intelligence particularly Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) presents new opportunities to improve these systems. Recent studies have shown promising results: Magron et al. (2024) demonstrated that LLMs trained on synthetic job data outperformed traditional supervised methods in skill extraction, while Kavas et al. (2024) developed interactive LLM-based systems that integrate recruiter input for multilingual job-candidate matching. However, most of these innovations focus on improving employer-side recruitment workflows rather than supporting job seekers in articulating their capabilities more clearly.

Therefore, this study aims to develop and evaluate AI Career Assistant (AICA), a job matching platform that uses Large Language Models and Retrieval-Augmented Generation. AICA is designed to bridge the critical gap between graduate skills and industry demands by enabling Filipino technology students and recent graduates to articulate their competencies effectively and discover relevant career opportunities.

Statement of the problem

A persistent skills mismatch continues to affect new graduates entering the job market, particularly in the technology sector. The disconnect between what students learn in school and what employers actually need continues to affect employability and career alignment. The system aims to serve as a career alignment tool for Filipino technology students and graduates, helping them better present their capabilities and connect with jobs that match their true potential.

This study addresses the following key problems:

1. What features do users consider important in a job matching platform that helps them find opportunities aligned with their skills?
2. How can a Large Language Model with Retrieval-Augmented Generation be applied to accurately rank job opportunities based on a user's declared skills?
3. What is the level of users acceptance of the AICA platform in terms of :
    a. Usability
    b. Job Matching
    c. Resume Builder Support

Conceptual Framework

The conceptual framework was guided by our goal to assess and improve the career readiness of tech students and graduates by evaluating their hard and soft skills to match them to relevant jobs.

This research focused on addressing the challenges faced by tech students and graduates to align their existing skills with relevant professional roles. The AICA platform was developed in response to the growing need for an accurate and personalized tool that facilitated this alignment by matching users to appropriate career opportunities. The framework was structured around four core components: Input, Process, Output, and Outcome.

The Input stage represented the foundational elements and data sources that initiated the AICA system's operations. This included User Skills + Profile, which encompassed the explicit information provided by job seekers, such as their skills and work experiences.

Complementing this was the Resume Builder, a tool designed to assist users in effectively articulating their technical and soft skills, thereby creating structured and comprehensive professional profiles. The technological backbone of the input stage was the "LLM + RAG Engine" (Large Language Model with Retrieval-Augmented Generation Engine). This signified the core AI technology that was employed for advanced natural language understanding, information retrieval from external knowledge bases (such as up-to-date job market data or skill taxonomies), and coherent text generation for summarization or recommendation. Finally, the Skill Matching Algorithm was identified as an input, representing the pre-defined logic and methodologies that, in conjunction with the LLM + RAG engine, were used to compare and align user skills with job requirements.

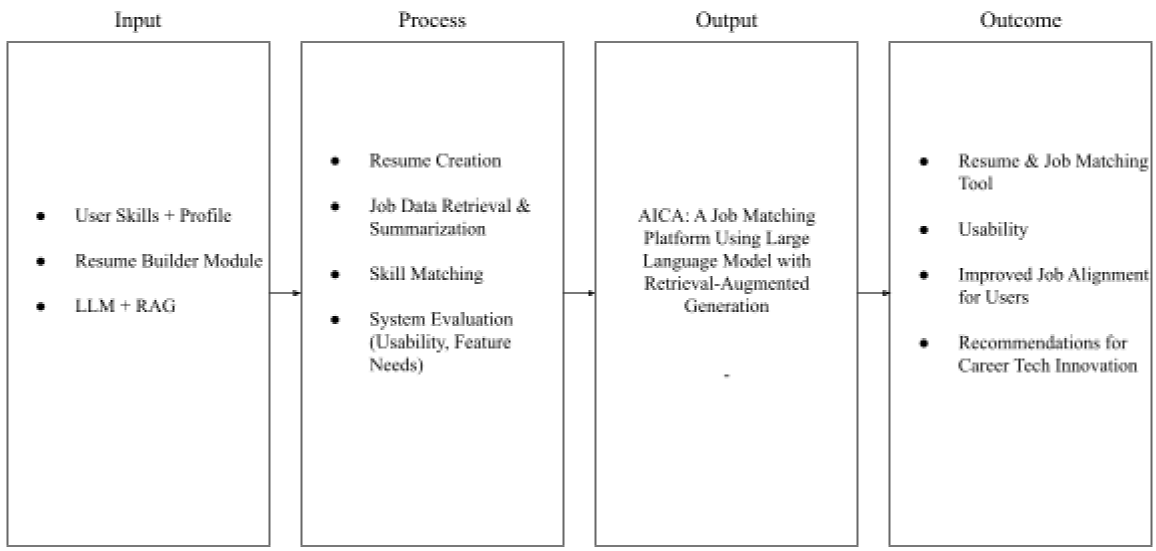
Following the input, the Process stage detailed the core operations and transformations performed by the AICA platform. The first process was Resume Creation, facilitated by the Resume Builder Module and driven by user inputs, resulting in a refined and AI-analyzed resume. Subsequently, Job Data Retrieval & Summarization occurred, where the system actively sourced job postings, potentially through methods like web scraping or API integrations as suggested by related literature (e.g., Budharapu et al., n.d.) and then utilized the LLM to summarize these descriptions, extracting key requirements and responsibilities. The central process was Skill Matching, where the LLM + RAG engine and the Skill Matching Algorithm analyzed the processed user resumes and summarized job descriptions to identify semantic similarities and contextual fits, moving beyond simple keyword matching, as highlighted by research on BERT-based systems (Almalki, 2025). This stage also incorporated System Evaluation (Usability, Feature Needs), an iterative feedback loop crucial for refining the platform based on user interactions, performance metrics, and identified needs for new functionalities, ensuring the system remained user-centered and effective.

The Output of this development and operational process was the tangible deliverable: "AICA: A Job Matching Platform Using Large Language Model with Retrieval-Augmented Generation." This was the fully functional platform itself, embodying all the preceding inputs and processes. It represented a tool designed to provide a superior job matching experience by leveraging advanced AI capabilities as detailed in this study.

Finally, the Outcome stage described the anticipated impacts and benefits resulting from the deployment and utilization of the AICA platform. The primary outcome was a Resume and AI Matching Tool, providing users crafting resumes with an intelligent assistant for finding highly relevant job opportunities. This, in turn, was expected to lead to improved Job Alignment for Users, directly addressing the critical gap between graduate skills and industry demands by facilitating better-quality matches. Beyond individual user benefits, a broader outcome was the generation of Recommendations for Career Tech Innovation. Insights derived from the platform's operation, user interaction patterns, and successful matching criteria could inform future developments in career technology, potentially influencing educational curricula alignment or identifying emerging skill demands within the tech sector.

Figure 1.

Conceptual Framework of the Study



Scope and Limitations

This study presented the design, development and evaluation of AICA, an intelligent job matching platform that integrated Large Language Models (LLMs) with Retrieval-Augmented Generation (RAG). The platform was intended to assist tech students and recent graduates in proving their employability by helping them improve how they saw their skills and by aligning with their qualifications and ideals.

The scope of this research leveraged the capabilities of LLMs to understand semantic nuances in user inputs and applied RAG to retrieve real-time job market insights. The system presented a statistical percentage of skills the users were inclined to which they could use to improve their resume. This platform was particularly intended for graduating students and recent graduates in the technology field. However, the platform was not limited to these groups but could also be utilized by anyone seeking employment within the technology industry. Designed with a specific focus on the technology job market, AICA used publicly accessible data sources particularly in Indeed to deliver relevant and timely job recommendations for students and early-career job seekers.

Designed with a specific focus on the technology job market, AICA used publicly accessible job data to deliver relevant and timely job recommendations for students and early-career job seekers. In particular, this study considered job postings extracted from four major online employment platforms in the Philippines: JobStreet, Indeed, Kalibrr, and OnlineJobs.ph. These platforms were selected based on their reliability, accessibility, and coverage. Other job platforms could also be considered if they offered valuable or specialized listings relevant to the target users.

Despite the platform's innovation, several limitations were recognized. Firstly, the system could not guarantee that all relevant and time-sensitive job opportunities would be captured and delivered to users in real time. Second, the platform currently supported only the English language, which could limit accessibility for non-English-speaking users. This could pose challenges for non-English-speaking users, potentially affecting accessibility, user experience, and overall inclusivity for a global audience. Third, the platform did not support embedding documents, images, or other non-textual files. This also could help encourage users to enhance their own resume in the platform. Although the system implemented AI, it lacked an AI-powered resume generator that provided real-time writing support or duplicated every feature of advanced commercial resume-building products. While AICA utilized data from major platforms like Indeed, it could not guarantee the capture of every relevant job posting across all possible job sites or company career pages, especially given the periodic nature of the updates. Consequently, AICA was designed solely to support job seekers and did not offer end-to-end solutions for job postings or employer-side functionalities. Moreover, the comprehensiveness of job listings was constrained by the data sources. For instance the system would only update. Finally, the job matching scope did not currently include capabilities for employers or an end-to-end job advertising and application management system; instead, it was largely focused on assisting job seekers.

Significance of the study

The development of AICA was highly relevant in addressing the persistent gap between the skills of technology graduates and the competencies demanded by today's evolving job market. This study contributed a practical and innovative solution by providing an AI-powered resume builder and job-matching platform that used Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) to align user-declared skills with real-time job opportunities. The research benefited multiple groups:

Job Seekers and Technology Graduates. AICA provided a structured and accessible platform for job seekers and graduates to present both technical and soft skills through an AI-supported resume builder. It offered personalized insights into how their skill sets aligned with current job opportunities, enabling more informed decisions regarding career paths and job applications.

University of St. La Salle Career Development Center. The Career Development Center (CDC) could integrate AICA into its career preparation initiatives, enabling students to articulate their technical and soft skills more effectively through the platform's structured resume builder. The AI-generated job recommendations provided students with real-world insight into how their skill sets aligned with current industry demands. Additionally, the data collected from user profiles and matching results could support the CDC in offering more targeted career guidance, ensuring that graduates were more confident and competitive in the job market.

Employers and recruitment professionals. The platform gave better job matches by understanding the details of both the job and the applicant, not just matching keywords. It also gave personalized job suggestions based on the user's skills, experience, and goals. In addition, it helped users improve their resumes by giving easy-to-understand feedback on what could be made better.

Academic Researchers. This study brought together ideas from different fields like language processing, finding information, and job skill alignment, showing how they could work together in one system. It could also serve as a starting point for future research, helping others test how well this kind of technology worked in other areas like course recommendations or mentor matching.

Future Researchers. Future researchers could take inspiration from this work by improving accuracy, expanding features, or adapting it to specific industries or local job markets. This study served as a useful reference for creating accessible and skill-based career support tools.

Definition of Terms

Large Language Model (LLM). A large language model (LLM) was an artificial intelligence system trained on massive datasets to understand and generate human-like text based on patterns in language (Brown et al., 2020). These models used deep learning, particularly transformer architectures, to perform a wide range of tasks such as translation, summarization, and question answering (Vaswani et al., 2017).

Operationally, LLM was utilized primarily in two key post-resume-creation stages. First, it processed and analyzed raw job description texts fetched from external sources, extracting key requirements and generating semantic embeddings for the RAG system's knowledge base. Second, after the RAG component retrieved potentially relevant job listings, the LLM performed a detailed semantic comparison between the user's structured profile (from the in-app resume builder) and each retrieved job description, assigning a match score and generating a justification for the alignment, thereby powering the core job matching decision.

Retrieval Augmented Generation (RAG). A method that combined large language models with an external knowledge retrieval system, enabling the model to access relevant documents during generation to improve factual accuracy and context-awareness (Lewis et al., 2020).

Operationally, Retrieval-Augmented Generation (RAG) within AICA functioned as the intelligent information retrieval backbone for the job matching process. When a user's profile (from the in-app resume builder) was processed into a query embedding, the

RAG system's retriever component efficiently searched this database to identify and fetch a set of the most semantically similar job listings.

Review of Related Literature

This chapter presents a review of related literature essential to the study, including concepts and prior research relevant to the development of AICA. It covers literature on the skills gap in the tech sector, the significance of soft and hard skills.

Skills Gap and Employment Challenges in the Tech Sector

Advancements in technology have influenced educational and academic standards, requiring curriculum designers to ensure that students acquire skills aligned with market demands (Kureková, Beblavý & Thum-Thysen, 2015; Vitale, Bowyer & Bayerlein, 2020). The Covid-19 pandemic worsened the volatility of labor markets, causing job scarcity as industries were severely impacted by the global economic stagnation. As a result, graduates' employability faced significant challenges, necessitating the development of new attitudes and skills to adapt to the post-Covid-19 environment (Muhammad Iqmal, Azuan, Mohd Asri, & Saffa Nasuha, 2020). As a result, the standard for graduate employability has been raised, requiring graduates to be more competitive to succeed in the labor market.

In the Philippine context, a survey conducted by the People Management Association of the Philippines (PMAP) revealed that 40% of fresh graduates struggle to secure employment due to skill deficiencies. According to Melba Tutor, a Philippine Institute for Development Studies (PIDS) consultant, many graduates perceive themselves as lacking crucial job performance skills—particularly communication, critical thinking, and problem-solving abilities. Notably, 25% of graduates attribute their employment difficulties to outdated college curricula that fail to align with contemporary industry requirements.

These findings reflect a broader global trend. The rapid evolution of the technology sector has highlighted a growing mismatch between the skills possessed by job seekers and those demanded by employers. Despite the increasing number of tech graduates, many still lack the specific competencies required by the industry, leading to a persistent skills gap. For instance, Essama-Nssah, Bahar, and Chen (2020) found that the shift toward export-intensive and high-skill industries in the United States has resulted in significant job displacement for low-skilled workers, even as opportunities expand for those with advanced skill sets. Their findings emphasize the need for upskilling initiatives to ensure that workers can transition into emerging sectors and adapt to changing job requirements. Although their study is focused on the U.S. context, the implications are equally relevant in the global tech landscape, where similar disparities in workforce readiness persist.

It is crucial to provide all tech graduates with opportunities to develop the skills necessary to fully engage in the future workforce, thereby fostering more inclusive and sustainable innovation to enhance their competencies.

Soft and Hard Skills in Tech Employment

A study from Harvard University highlights that a person's success is influenced not only by technical expertise (hard skills) but also by emotional regulation and interpersonal abilities (soft skills) (Cahyadiana, 2020). Researchers have emphasized that effective workplace performance depends on a blend of hard and soft skills (van der

Vleuten et al., 2019; Lyu & Liu, 2021). This insight offers educators and trainers a valuable chance to adopt integrated approaches in skill development, combining both technical and interpersonal aspects (Valderrama, 2025).

One skill that employers are seeking is the basic skills or we call it the soft skills. When we examine the definition of soft skills, we can see that they are linked to emotional intelligence, or the capacity to recognize ourselves in the context of others and how our actions influence others (A. K. Touloumakos 2020, B. X. F. Da Silva, V. Carolina Neto, and N. H. S. Gritti 2020). Soft talents may be used in a variety of jobs and sectors. As a consequence, even if you don't fit the precise profile in a job description, you may discover that you possess many of the needed attributes. In addition to this, soft skills are essential for personal development and job market success (Volkov et al., 2022). They fall into two groups—those that help manage oneself, like creative thinking and stress management, and those that facilitate effective interpersonal interactions, such as leading and negotiating (Gejdoš et al., 2021).

On the other hand, hard skills are defined as technical, measurable abilities required for specific tasks, such as programming, driving, or welding (Lyu & Liu, 2021). While soft and hard skills serve different purposes, possessing both enhances an individual's employability, as the combination provides a more comprehensive and competitive skillset (Rainsbury et al., 2002; Succi & Canovi, 2020). Many roles now consider a mix of both skill types essential.

The expectations for knowledge, skills, and abilities in higher education have notably risen, highlighting the growing importance of soft skills in reaching professional excellence (Vasylenko et al., 2022). While both hard and soft skills are essential for career success, soft skills are often viewed as more critical due to their broad applicability across various roles and their rising relevance in the future (Pranic et al., 2021).

Job Portals, Expectations, and Employment Outcomes

The rise of online job portals has introduced scalable infrastructure for matching job seekers with employment opportunities. However, their effectiveness depends not only on the quality of job listings but also on users' expectations and behaviors. c(2022) conducted a randomized control trial in India to evaluate how access to an online job portal affected job-seeking behavior and employment outcomes among vocational training graduates. Surprisingly, the study found that portal access led to lower employment rates in the short term, suggesting the presence of voluntary unemployment. Job seekers raised their reservation wages and reduced job acceptance while waiting for better offers that never materialized. These findings underscore a critical challenge in digital job matching platforms: unrealistic expectations can reduce job uptake rather than enhance it.

The study highlights how information asymmetry and biased beliefs about job prospects can lead job seekers to hold out for roles that are either unavailable or unsustainable. Interestingly, the group that received more frequent job-related messages from the portal ("priority treatment") exhibited smaller drops in employment over time. This implies that more consistent exposure to real job listings helped some job seekers recalibrate their expectations and return to work. The employment rebound was especially evident among older and less affluent job seekers, who faced higher opportunity costs from prolonged unemployment and were quicker to adjust their beliefs when offers fell short of expectations.

These findings raise broader concerns about how platform design and user perception interact in employment technologies. The portal studied primarily catered to low-skilled workers through SMS alerts — a method suited to India's vocational workforce. However, the misalignment between job location (mostly in Delhi) and job seekers' geographic preferences contributed to the mismatch. Additionally, many users' reservation wages exceeded market rates, revealing a disconnect between aspiration and opportunity.

The implications for systems like AICA are significant. While AI-driven platforms using LLMs and RAG can enhance semantic understanding of skills and job descriptions, expectation management and user education remain essential components. Users must be guided to develop realistic views of their employment options and skill fit, especially in environments where job quality, location, and wage variation are major constraints.

Kelley et al.'s (2022) work reinforces that the success of AI-augmented matching platforms does not solely depend on technical sophistication. Rather, it hinges on whether these systems can effectively inform, educate, and engage job seekers—preventing inflated expectations and promoting informed decision-making. Integrating behavioral design principles into platform development could help close the loop between job seeker behavior and AI-driven matching logic.

Large Language Models in Resume Matching and Job Alignment

The growing use of artificial intelligence in recruitment processes has introduced new opportunities and challenges in how candidates are evaluated for employment. One of the most significant developments in this area is the application of **Large Language Models (LLMs) such as OpenAI's G**PT-4, which are capable of interpreting unstructured text data, summarizing content, and performing semantic reasoning (Brown et al., 2020). These models have been increasingly adopted in recruitment contexts for tasks such as resume screening, job description summarization, and skill-job alignment.

In a recent observational study, Vaishampayan et al. (2025) investigated the effectiveness of GPT-4 in evaluating resumes across various dimensions such as work experience, education, and certifications. The study involved 736 real-world resumes and assessed how well GPT-4 could rate them compared to human recruiters. Results showed that while LLMs like GPT-4 demonstrated a reasonable ability to approximate human scoring, there were notable discrepancies in specific areas—particularly in evaluating nuanced qualifications and soft skills. For instance, the model was more lenient in interpreting skill matches but stricter in assigning scores to certifications. These inconsistencies suggest that while LLMs are promising tools for resume evaluation, they still lack the contextual understanding and judgment of human reviewers.

The performance of LLMs in resume matching is highly influenced by **prompt design and contextual framing**. The same study found that the use of *chain-of-thought (CoT) prompting* and *example-based prompts* significantly improved alignment with human assessments, particularly in categories involving structured information like education history. However, even with prompt engineering strategies in place, the model's maximum agreement with human reviewers reached only 61.6%, indicating that LLMs are not yet fully reliable for independent resume evaluation.

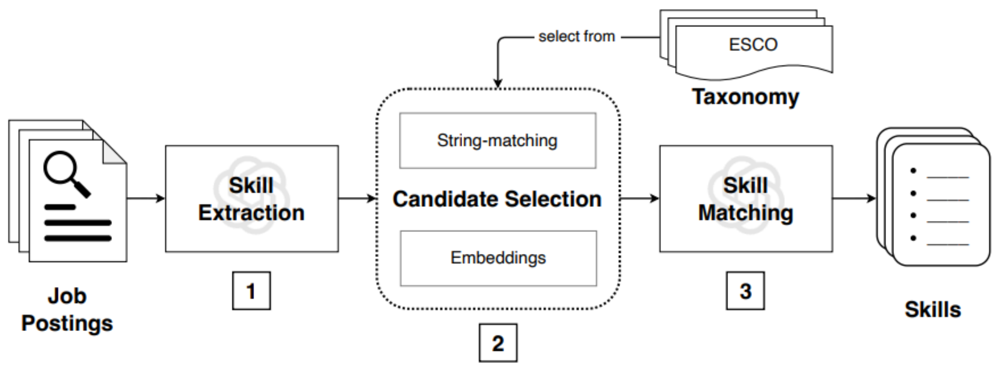
Ethical considerations also play a critical role in evaluating the use of LLMs in hiring. Vaishampayan et al. (2025) examined how LLMs rated candidates across demographic groups and found both encouraging and concerning patterns. While GPT-4 sometimes demonstrated **less bias than human reviewers**, inconsistencies in scoring across racial and gender lines were still present. This raises concerns about the fairness and transparency of AI-powered hiring tools, and reinforces the need for **Human-Centered AI (HCAI)** approaches, where LLMs serve as support tools rather than autonomous decision-makers (Amershi et al., 2019).

The implications of these findings are highly relevant for platforms like SkillMatch.AI, which aim to bridge the gap between graduate skills and industry requirements through AI-enhanced resume matching. While LLMs offer scalability and consistency, their limitations in interpreting resume nuances and ensuring fairness underscore the importance of integrating **user feedback** and **content validation** in the system's design. Moreover, current literature lacks a strong focus on how users themselves can best present their technical and soft skills in ways that LLMs can effectively interpret, particularly through resume-building interfaces. This represents a critical research gap that SkillMatch.AI aims to address.

Magron et al. (2024) introduced JOBSKAPE, a framework designed to generate synthetic job postings to enhance skill matching. By creating a comprehensive synthetic dataset, SkillSkape, the study demonstrated that LLMs trained on this data outperformed traditional supervised methods in skill extraction tasks, highlighting the potential of synthetic data in improving LLM applications. JOBSKAPE uses generative large language models (LLMs) to curate meaningful skill combinations and generate appropriate job descriptions containing these combinations. Moreover, for each job posting, the LLM identifies key skills and tasks within the job ad while omitting irrelevant information.

Figure 2.

Three-step Skill Extraction and Matching Pipeline.



Note. JOBSKAPE is a framework designed to generate synthetic job postings to improve skill matching in labor markets. From *JOBSKAPE: A Framework for Generating Synthetic Job Postings to Enhance Skill Matching* (p. __), by A. Magron, A. Dai, S. Montariol, M. Zhang, & A. Bosselut, n.d., arXiv.

In-context learning pipeline for end-to-end skill matching using an LLM to extract skills from job ads, then do candidate selection using heuristics, and last, do skill matching with a constrained taxonomy

A self refinement step using LLMs (Madaan et al., 2023) ensures label quality in the refined SKILLSKAPE dataset, assessed through offline metrics In addition, they leverage an LLM to match skills in synthetic job posting sentences to the ESCO taxonomy. This 47 pipeline has three steps, visualized in Figure 2: 1) skill extraction from the sentence, 2) candidate selection from the taxonomy, and 3) skill matching to the list of candidates

Meanwhile, Seedat and van der Schaar (2024) presented Matchmaker, a compositional LLM program aimed at table schema matching across heterogeneous data sources. The implementation of large language models (LLMs) represents a significant advancement in how LLMs can be utilized for structured data tasks, particularly in table schema matching. Unlike conventional LLM applications that rely on end-to-end prompting or fine-tuning, Matchmaker leverages a compositional LLM program structured into discrete steps: candidate generation, refinement, and scoring. Each component is executed through strategically designed prompts, capitalizing on the reasoning abilities of LLMs to decompose complex schema matching into manageable subtasks. This modular use of LLMs allows for more transparent and controllable operations, aligning well with tasks that require logical consistency and domain-specific reasoning.

Large Language Models (LLMs) form the foundation of Matchmaker, serving as key components within a compositional program composed of multiple language model calls. Specifically, LLMs exhibit several appealing properties and capabilities for schema matching:

- Contextual understanding: LLMs have been pre trained on a vast corpora of information, equipping them with extensive prior knowledge spanning different contexts and settings (Chowdhery et al., 2022). This contextual understanding enables LLMs to effectively reason about schema hierarchies and identify potential matches.

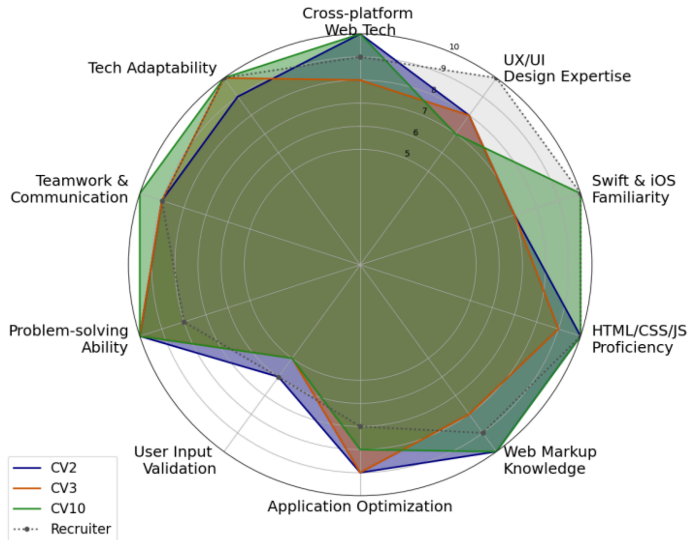
- Hypothesis proposers: LLMs have been shown to be “phenomenal hypothesis proposers” (Qiu et. al., 2024) , making them particularly useful for candidate generation tasks.
- Capable rankers: LLMs have been shown to be highly capable at relevance ranking; assessing the suitability of candidates given a query and a set of options (Zhuang et. al., 2023), especially “when ranking candidates retrieved by multiple candidate generators” (Hou Y et. al., 2024).

In addition, Kavas et al. (2024) developed an interactive system combining LLMs with recruiter expertise for optimized multilingual job offer and applicant CV matching. The system utilizes a fine-tuned classification model aligned with the European Skills, Competences, Qualifications, and Occupations (ESCO) taxonomy, facilitating accurate matching and relevance scoring through sequential LLM-driven interactions.

Their system utilizes the Mixtral 8x7B LLM (Jiang et al., 2024) from mistral.ai4 for its demonstrated ability to outperform previous models including GPT-3.5 Turbo, Claude2.1, Gemini Pro, and Llama 2 70B - chat model on human benchmarks. The Mistral model is deployed using an Ollama Docker image 5 , facilitating local operation and easy deployment across various environments.

Figure 3.

Example scoring results from our interface, demonstrating the evaluation of candidates based on skills and recruiter feedback.



Note. This paper presents an approach for improving multilingual job offer and CV matching using large language models and recruiter expertise. From *Using Large Language Models and Recruiter Expertise for Optimized Multilingual Job Offer–Applicant CV Matching* (p. __), by H. Kavas, M. Serra-Vidal, & L. Wanner, n.d. Retrieved May 24, 2025

To evaluate and score candidate CVs based on job-specific qualifications, they implement a four-step sequential LLM chain. First, following the extraction-style technique from Nguyen et al. (2024), they apply in-context learning with Mixtral LLM using five examples to identify relevant skills and competencies from job descriptions, particularly accommodating longer sentence structures. Second, the extracted skills are shown to recruiters through an interactive interface, where they can mark items as irrelevant or add comments. Third, another Mixtral LLM processes this recruiter feedback to evaluate each candidate's fit, factoring in their background and annotated skill relevance. Finally, the same model assigns a score from 1 to 10, based on skill alignment and recruiter input, averaging results across three iterations to minimize bias and improve reliability.

In summary, LLMs have shown significant potential in transforming recruitment systems, particularly in tasks like resume summarization and semantic skill matching. However, their limitations in contextual understanding, susceptibility to prompt variability, and potential bias make it essential to pair these technologies with user-centered design, transparent validation mechanisms, and human oversight. The integration of LLMs into career tools must be guided by both **technical performance metrics** and **user experience research**, ensuring that AI-enhanced systems empower users rather than marginalize them.

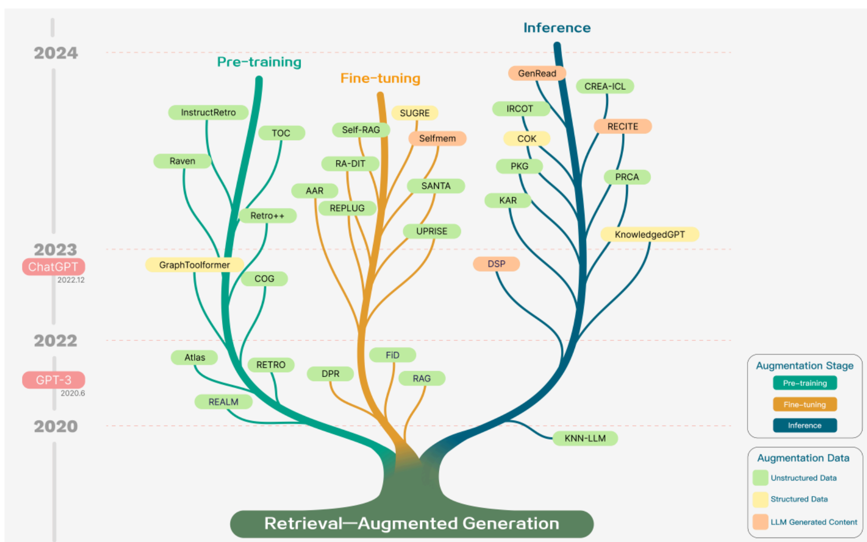
Use and Impact of Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG) has emerged as a significant breakthrough in the landscape of natural language processing (NLP), bridging the gap between static pre-trained models and the dynamic, evolving nature of human knowledge. Traditional large language models (LLMs) such as GPT and BERT, while powerful, are limited by the information encoded at the time of training, leading to hallucinations and outdated responses. RAG overcomes this constraint by integrating a retriever component with a generator. According to Gao et al. (2024), this hybrid architecture significantly improves factual accuracy and responsiveness in generation tasks by combining the strengths of retrieval-based methods and generative capabilities.

The structure of a RAG system generally consists of three main stages: retrieval, augmentation, and generation. First, a retriever model selects the most relevant documents or context from a knowledge base based on the input query. Then, the augmenter incorporates this retrieved content to provide additional context. Finally, the generator uses this context-enriched input to produce more informed and accurate outputs. The development of Retrieval-Augmented Generation (RAG) models is shown in Figure 4 of Mei et al. (2025). Most studies about RAG started coming out after 2020. A major turning point happened in December 2022 when ChatGPT was released, marking the start of a new era in natural language processing focused on large language models (LLMs).

Figure 4.

A timeline of existing RAG research. The timeline was established mainly according to the release date.



Note. This survey explores Retrieval-Augmented Generation (RAG) techniques that enhance large language models by incorporating external knowledge sources. From *Retrieval-Augmented Generation for Large Language Models: A Survey* (p. __), by Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang, & H. Wang, n.d., Shanghai Research Institute for Intelligent Autonomous Systems.

In practice, RAG has shown its value in various real-world applications. One notable example is its application in AI-driven resume analysis and job matching, as discussed by Kharde (2024). In this domain, RAG enables HR systems to extract relevant experiences from a resume and align them with job descriptions in real-time, thereby improving candidate screening efficiency. The study highlights how the dynamic retrieval of job-relevant information allows RAG-based systems to adapt to various industries and job roles with minimal retraining, offering scalable and explainable AI solutions.

Moreover, research from BAAI (2024) explores the integration of RAG in multimodal systems and highlights its scalability in handling large-scale knowledge-intensive tasks. The paper investigates how RAG can be optimized for long-form question answering, where grounding responses in multiple, diverse knowledge sources is crucial. The authors also address computational challenges, including latency and memory usage, and propose architectural enhancements such as batch retrieval and knowledge distillation to improve performance.

Despite its strengths, RAG faces several limitations. One of the critical issues is the quality of the retriever. If irrelevant or noisy documents are retrieved, the generator may produce inaccurate outputs. Moreover, integrating RAG into end-to-end pipelines often requires complex orchestration between the retriever and generator, making real-time deployment computationally expensive. Researchers continue to explore how to refine retriever models, enhance indexing mechanisms, and improve document filtering to ensure higher-quality generations.

To conclude, Retrieval-Augmented Generation represents a crucial shift toward more grounded, adaptable, and context-aware language models. By combining retrieval-based reasoning with generative fluency, RAG addresses many of the pitfalls seen in traditional LLMs. Its applications across domains, from resume screening to academic question answering, demonstrate its practical relevance and potential. However, challenges in retriever quality, efficiency, and system complexity suggest that while RAG is promising, it remains an evolving field with ample room for improvement and innovation.

Job Data Collection

Web scraping is a way of extracting vast volumes of data from websites in an automated manner. A significant portion of this data comes in the form of unstructured HTML, which needs to be converted into structured formats such as spreadsheets or databases before it can be used in various applications. Web scraping is a common method for extracting data from websites and can be carried out in several ways, including using online tools, dedicated APIs, or writing custom scraping scripts. Many major websites, such as Google, Twitter, Facebook, LinkedIn, and Stack Overflow, provide APIs that allow access to their data in a standardized and developer-friendly format. This is the best choice, but there are other places that either don't allow users to access vast volumes of data in a standardised format or aren't technologically advanced enough (Pandey, 2021).

In that case, scraping the website for data using Web Scraping is the best option. The crawler and the scraper are needed for web scraping. The crawler is an artificial intelligence algorithm that searches the internet for specific data by following links across the network. A scraper, on the other hand, is a tool designed to remove information from a website. The scraper's design can vary greatly depending on the project's complexity and scope, but it must be able to extract data quickly and accurately. One might wonder how this works. Web scrapers will retrieve all of the data on a platform or only the data that a user is looking for. It's best if you define the details you're looking for so that the web scraper just extracts the information quickly (Pandey, 2021).

Web scraping, essential in AI, data science, and cybersecurity, involves the automated extraction of data from websites, benefiting e-commerce, data analytics, and research (Khder, 2021). Web scraping is an essential part of data science and AI. Web scraping, an essential component of data science and AI, extends beyond data extraction to involve intelligent parsing and organization for real-world applications (Khder, 2021). Moreover, natural language processing (NLP) excels in extracting significant insights from text-heavy pages, integrating technologies for both quantitative and qualitative data analysis (Cai, 2021). Intelligent web scraping is key for market and customer sentiment analysis, competitor monitoring, and price tracking in business (Khder, 2021). In scholarly studies, intelligent web scraping enhances large dataset aggregation. Automated resume evaluation and web scraping revolutionize recruitment by minimizing errors and biases (Mishra, et al., 2023).

Top Reliable Job Sites in the Philippines

The rise of digital labor platforms in the Philippines has significantly transformed how job seekers and employers connect. These online platforms act as digital labor market intermediaries, streamlining the employment process through user-friendly features and advanced matching technologies (Agaton et al., 2024).

JobStreet

According to SimilarWeb, a global digital analytics platform, JobStreet is the most visited employment platform in the Philippines, reflecting its dominance in the digital labor market. It provides an extensive range of job listings and features such as salary match tools and industry-specific filters. According to Guild Solutions Inc. , an Australia-based recruitment and workforce solutions company, JobStreet tops their ranking of job sites due to its expansive network and employer partnerships. Furthermore, according to Moneymax, a Philippine-based personal finance comparison website, JobStreet is highly trusted by Filipino job seekers for its regular updates and comprehensive postings.

Indeed

According to Guild Solutions Inc., Indeed is the second most reliable job site, known for aggregating listings from various sources. It is especially popular for its minimalist design and powerful search engine. SimilarWeb ranks Indeed among the top job platforms in the country, based on web traffic. In addition, Indeed's employer reviews and resume upload options offer added value for job hunters.

OnlineJobs.ph

According to Moneymax, OnlineJobs.ph is ideal for Filipinos seeking home-based or freelance roles, with many employers hiring virtual assistants and remote workers. Unlike broader job boards, OnlineJobs.ph focuses on digital and remote work, making it a key player in the online employment space.

Kalibrr

According to Guild Solutions Inc., Kalibrr is a fast-growing platform recognized for its AI-enhanced matching system and professional profile creation. It is especially prominent among white-collar professionals. The platform also provides assessment tools and branding services, enhancing employer-employee alignment. As Agaton et al. (2024) note, Kalibrr serves both as a job board and an end-to-end hiring service provider.

Bossjob

For Guild Solutions Inc., Bossjob is distinguished by its chat-based application system, allowing real-time interaction with recruiters. It is particularly useful for professionals looking for mid-level and executive roles. According to Moneymax, Bossjob's AI-driven notifications and streamlined dashboard contribute to its effectiveness.

Online job platforms have made it much easier for Filipinos to find work and for companies to find the right people. Websites like JobStreet, Indeed, OnlineJobs.ph, Kalibrr, and Bossjob each offer different features that help job seekers connect with employers more quickly and easily. These platforms are now a big part of how people look for jobs in the Philippines, especially with the help of technology like AI, chat systems, and job matching tools.

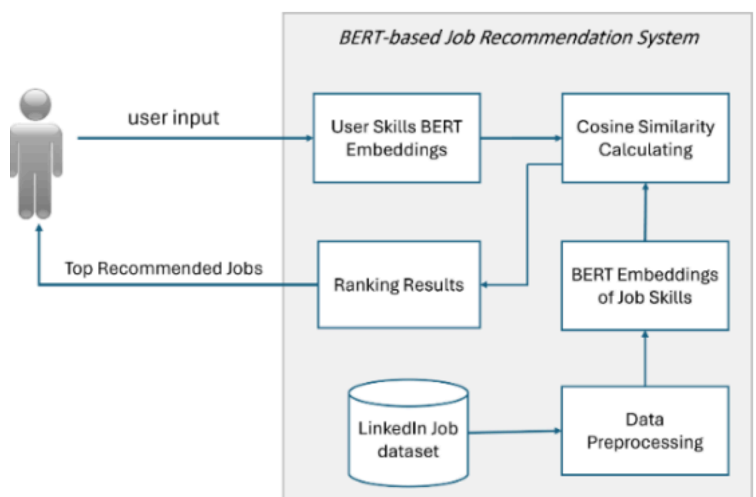
Leveraging Transformer-Based Language Models for Enhanced Job Matching

Building on the capacity of AI to understand job market data, the application of advanced transformer-based language models specifically for job recommendation and matching represents a significant leap forward. These models excel at capturing contextual nuances in language, which is crucial for bridging the semantic gap between how job seekers describe their skills and how employers articulate their needs. The pursuit of more accurate and personalized job matching has increasingly turned towards these models to overcome the limitations of earlier systems.

Addressing these limitations, Almalki (2025) presents a job recommendation system that utilizes the Bidirectional Encoder Representations from Transformers (BERT) model, a powerful NLP framework, to analyze data from a LinkedIn dataset. The core of this system lies in its ability to discern deep semantic relationships within textual data, thereby enabling precise and personalized recommendations through contextual matching of skills and preferences. By harnessing BERT's bidirectional contextual capabilities, Almalki's research offers an alternative approach to the job-matching process, aiming to significantly improve recommendation accuracy over conventional content-based methods.

Figure 5.

The proposed system architecture



Note. This study proposes a BERT-based job recommendation system leveraging LinkedIn data to improve personalized job matching. From *BERT-based Job*

Recommendation System Using LinkedIn Dataset (p. __), by K. M. Rabby, M. S. Islam, & S. M. Masum, 2024, ResearchGate.

Almalki's work (2025) highlights how BERT's ability to process information bidirectionally allows for a more profound contextual understanding of both job descriptions and user profiles, moving beyond superficial keyword matching to true semantic similarity. The reported effectiveness in recommendation accuracy demonstrates the practical benefits of such an approach. This is directly relevant to the AICA platform, which proposes to use Large Language Models (LLMs) – an evolution of transformer architectures – coupled with Retrieval-Augmented Generation (RAG). While BERT focuses on encoding and understanding, LLMs with RAG offer enhanced generative capabilities and the ability to incorporate external, up-to-date knowledge. This combination, as envisioned in AICA. The insights from Almalki (2025) regarding the efficacy of transformer models in semantic matching bolster the foundational premise of AICA, while AICA seeks to extend these capabilities by incorporating RAG for more dynamic, context-aware, and empowering job matching experiences.

AI-Powered Analysis of Job Postings for Labor Market Insights

The increasing complexity and dynamism of the modern labor market necessitate advanced tools for understanding job requirements, skill demands, and emerging trends. Artificial Intelligence (AI), especially Natural Language Processing (NLP), offers powerful capabilities to extract meaningful insights from the vast amounts of textual data available in online job postings. Exemplifying this, Müller and Düsing (2024) recently developed and validated a measurement instrument based on deep learning and NLP techniques to forecast the technical feasibility of AI performing specific job tasks, using unstructured job postings as input. Their research, which showcased high predictive power and strong correlation with traditional metrics derived from sources like O*NET, demonstrates the efficacy of AI in dissecting job descriptions to understand underlying task characteristics and their amenability to automation.

While the primary focus of Müller and Düsing (2024) was on assessing AI's impact, their methodological success in extracting nuanced task characteristics from unstructured job postings in real-time is highly pertinent. Their use of explainable AI to identify specific words and phrases indicative of AI feasibility further underscores the depth of analysis achievable. This capability to deeply parse and interpret job requirements is a foundational element for systems like AICA, which must accurately understand employer needs to facilitate effective job matching, even though AICA's end goal is candidate-job alignment rather than AI impact forecasting. The validation of AI-driven analysis of dynamic job posting data lends strong support to AICA's approach of leveraging real-time job opportunities and empowering users to articulate their competencies effectively in this evolving landscape.

Acknowledging LLM Limitations and Exploring Hybrid AI Architectures

The recognized challenges with Large Language Models (LLMs)—spanning reliability, cost, speed, and their noted struggles with complex logical reasoning tasks such as constraint satisfaction, as critically analyzed by Chu-Carroll et al. (2024) in "Beyond LLMs: Advancing the Landscape of Complex Reasoning"—drive the exploration of enhanced AI architectures. The research of Chu-Carroll et al. (2024), featuring the neuro-symbolic EC AI platform that pairs a robust logical engine with LLMs for ancillary functions, illustrates the significant performance gains achievable over unassisted LLMs in these areas, reinforcing the value of LLM augmentation. AICA's architectural design responds to these insights by incorporating Retrieval-Augmented Generation (RAG). This isn't a formal symbolic reasoner, but a vital mechanism for LLM enhancement, specifically enabling AICA to counter 'hallucinations,' tap into dynamic, real-time job market information (e.g., current job postings), and elevate the factual accuracy and contextual fit of its job recommendations, thereby striving for a superior and more dependable user experience in career navigation.

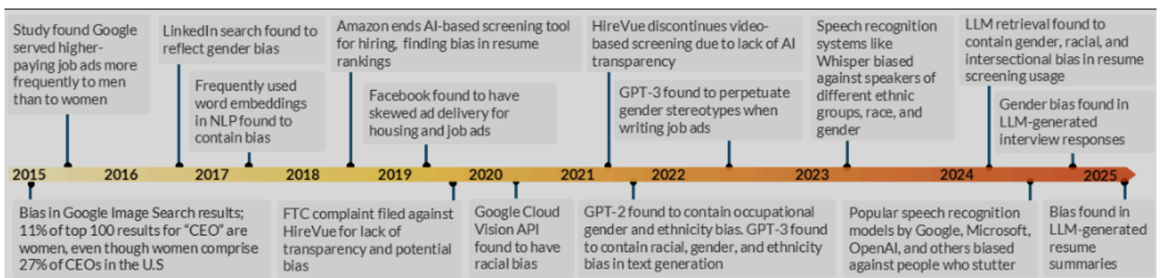
Ethics and Bias in AI Hiring Systems

A major theme in recent literature is the risk of bias and unfairness in AI-driven recruitment. Studies consistently find that if training data or algorithms reflect historical inequities, automated screening can amplify discrimination. For example, Brookings research found that open LLM-based screening showed significant gender and racial bias (particularly against Black men). Likewise, Chen (2023) reviews evidence that AI recruiting can inadvertently perpetuate biases (e.g. on gender or ethnicity) due to skewed data or assumptions. Unchecked, these systems may filter out qualified applicants based on protected traits. To mitigate these issues, the literature emphasizes both technical and governance remedies. Technical measures include curating unbiased training datasets, applying fairness-aware ML techniques, and maintaining algorithmic transparency (so decisions can be audited). Administrative measures include internal ethical guidelines and external oversight (e.g. independent audits or regulatory standards). Recent surveys advocate regular bias testing and multi-step review of AI decisions: once bias is detected, the root cause must be addressed and fairness must be re-evaluated continuously. Overall, the field highlights that AI hiring tools must be carefully designed and monitored –

ensuring compliance with fairness standards and legal frameworks (such as the EU's proposed AI Act for high-risk hiring AI) – to prevent unjust outcomes. These insights inform AICA's development by underlining the need to build in bias-mitigation (e.g. blind screening options) and explainability mechanisms from the outset.

Figure 6.

Timeline of notable incidents highlighting biases identified in AI systems relevant to or influencing recruitment applications



Note. From *Aligning Job Taxonomies: A Global Perspective on Matching Frameworks* (p. __), by L. Wolfram, N. Gutiérrez, M. Campero, & M. Fana, 2024, arXiv.

Soft and Hard Skills: Alignment and AI Assessment

Both technical (hard) and interpersonal (soft) skills are crucial for job fit. Classic research suggests well-developed soft skills account for ~85% of career success, far outweighing purely technical knowledge. Employers likewise emphasize communication, teamwork, adaptability and problem-solving as top requirements for new hires. AI can assist in evaluating both kinds of skills. For hard skills, LLMs and NLP pipelines excel at parsing resumes or cover letters to extract keywords (e.g. programming languages, certificates, degrees) and match them to job requirements. For soft skills, emerging work shows AI can help quantify these once-subjective traits. For instance, Ciaschi & Barone (2024) demonstrate that AI-based systems can "effectively gauge" attributes like communication style and emotional intelligence – tasks traditionally done by human judges – by analyzing text and audio cues. By modeling patterns in language use or simulating interview interactions, AI can offer a more consistent, scalable soft-skill assessment. Reviews note that integrating AI for soft-skill evaluation "holds immense promise" for talent assessment (enabling companies to screen interpersonal abilities across many candidates), even while they caution about bias and transparency. In summary, the literature highlights that soft skills are key determinants of employability and that AI tools can play a role in measuring them alongside hard skills, thereby improving the match between a candidate's profile and job demands.

PSSUQ Framework for a Qualitative Research

Recent studies have continued to validate and expand the use of the Post-Study System Usability Questionnaire (PSSUQ) as a tool for evaluating user satisfaction and system usability, particularly in the context of digital platforms and emerging technologies. Although the PSSUQ is primarily a quantitative tool, its framework can effectively inform qualitative research design. The three subscales—system usefulness, information quality, and interface quality—can be used as thematic categories to guide the development of interview protocols, focus group questions, and qualitative coding schemes

A growing body of literature highlights the value of integrating the PSSUQ framework into qualitative research design. For instance, Kizilcec et al. (2021) demonstrated that using PSSUQ subscales as thematic guides in interviews and focus groups allowed for a more nuanced exploration of user experiences with online learning platforms. Their qualitative analysis, structured around the PSSUQ's core dimensions, provided deeper insights into specific usability challenges and areas for system improvement.

Similarly, recent research by Alshamari and Mayhew (2019) emphasized that adapting PSSUQ items into open-ended qualitative prompts enabled researchers to capture detailed user narratives about satisfaction, perceived usefulness, and interface issues in healthcare applications. This approach facilitated the identification of usability barriers that might not be evident through quantitative scores alone.

Furthermore, the reliability and validity of the PSSUQ have been reaffirmed in contemporary usability studies. Lewis (2018) reported that the instrument maintains high internal consistency (Cronbach's alpha > .90) across diverse digital environments, supporting its continued use as a foundational framework for both quantitative and qualitative usability research.

In summary, literature confirms that the PSSUQ framework is not only a reliable quantitative instrument but also an effective guide for qualitative research. By structuring qualitative data collection and analysis around the PSSUQ's validated dimensions, researchers can achieve a holistic understanding of user experience and system usability.

METHODS

This section describes the research methodology used to develop and evaluate the AICA platform. It presents the research design, participant selection, data gathering procedures, research instruments, and analytical approaches used to address the study's research questions. Additionally, it discusses the statistical treatments applied, as well as the ethical considerations observed throughout the research process.

Research Design

This study employed a mixed-methods research design within a developmental framework to create and evaluate AICA, combining qualitative and quantitative assessment approaches to comprehensively address the research questions and objectives.

To guide the creation and refinement of AICA, the study adopted a developmental research design, which was particularly suited to the systematic development and evaluation of AI-driven platforms in real-world employment contexts. This approach facilitated iterative cycles of design, implementation, and assessment, ensuring that the platform effectively met user needs and technical requirements. The technical development and evaluation of the semantic job matching algorithm which leveraged LLM with RAG to extract skills from user inputs and match them with real-time job opportunities was systematically addressed through the development and implementation stages of this design process. This process involved designing and refining an algorithm that could accurately rank job opportunities based on the semantic alignment between extracted skills and job descriptions, ensuring relevant and personalized recommendations for each user.

The research adopted a sequential approach, beginning with a qualitative exploration of user requirements through focus group discussions and semi-structured interviews with third- and fourth-year technology students, as well as recent graduates. This qualitative phase aimed to capture nuanced insights into the challenges faced by job seekers and their expectations of AI-assisted career alignment tools, thereby informing the identification of essential features for AICA. This phase was essential for identifying requirements that could not emerge through structured surveys alone, informing evidence-based feature development for the AICA platform.

To evaluate user acceptance of AICA, the study applied a mixed-method approach. The quantitative component consisted of Likert-scale items adapted from the Post-Study System Usability Questionnaire (PSSUQ) to assess usability. Additional research-made items were included to measure user satisfaction with the relevance of job matches generated by the platform.

The qualitative component consisted of post-task interviews that explored how effectively the resume builder supported users in presenting their technical and soft skills. Interview responses underwent thematic analysis to identify recurring patterns related to usability, and perceived alignment with job opportunities. Together, this mixed-methods evaluation offered a comprehensive understanding of AICA's usability, functional effectiveness, and its potential to assist users in navigating the job search process.

Respondents

The respondents of the study are selected using a combination of purposive and convenience sampling techniques, tailored to suit the specific characteristics of each target group. According to Tongco (2007), purposive sampling, while traditionally aligned with qualitative methods, is also useful in quantitative research, especially in exploratory stages or when studying specific populations. This technique is effective when the research requires participants with particular characteristics. On the other hand, convenience sampling allows researchers to gather data from participants who are readily accessible, especially when time or access to a larger pool is limited.

- Recent graduates (within 1 to 2 years of graduation) from the same programs who are undergoing job applications.
- Third and Fourth Year Technology Students currently enrolled in degree programs such as Bachelor of Science in Computer Science (128 student population) or Bachelor of Science in Information Technology (226 student population) at the University of St. La Salle

For recent graduates, convenience sampling is employed due to the practical constraints of accessing this group. Unlike current students who are formally enrolled and easier to reach through institutional channels, recent graduates are dispersed and no longer bound to a central communication structure. Their participation is often subjected to availability, current employment status, and willingness to respond, which makes random or stratified sampling impractical. The findings from this group are not intended to be generalized to all graduates, but rather to inform the study through the experiences of individuals who are likely to have used or benefitted from platforms aimed at job matching,

Meanwhile, purposive sampling is used for current IT and CS students, as they represent a targeted group with specific academic backgrounds aligned with the study's goals. The total population of these 2 courses is 354. To ensure that the sample size is statistically valid, the study used Slovin's Formula to determine the minimum number of respondents required at a 95% confidence level with a 5% margin of error:

$$n = \frac{N}{1 + Ne^2}$$

Where:

- n = sample size
- N = population size (354)
- e = margin of error (0.05)

$$n = \frac{354}{1 + 354(0.05)^2} = \frac{354}{1 + 354(0.0025)} = \frac{354}{1.885} \approx 188$$

Thus, a sample size of approximately 188 respondents is required to ensure statistical validity. The study may collect responses from 181 to 283 participants to account for variability and data quality assurance, ensuring that the final dataset remains within the accepted confidence interval range and supports reliable analysis.

Research Instrument

The instruments to be used by the researchers in this study include a mix of qualitative and quantitative tools designed to gather comprehensive data from 4th year technology students and recent technology graduates. One of the primary instruments is a researcher-made interview guide, that is used to conduct semi-structured interviews. These interviews are held either face-to-face or through online platforms such as Zoom, depending on the availability and preference of the participants. The interview sessions aim to collect detailed feedback and insights about the participants' experiences, perspectives, and expectations regarding career platforms and how they navigate job-seeking processes within the technology sector.

For the quantitative component of the study, a structured survey will be administered through Google Forms. The survey will consist of Likert-scale items adapted from the Post-Study System Usability Questionnaire (PSSUQ) to measure core usability dimensions such as system usefulness, interface quality, and overall satisfaction. Additionally, researcher-made items will be included to specifically assess user satisfaction with the relevance and personalization of job recommendations generated by AICA.

The usability of the resume builder component, which falls under the User Acceptance, are assessed through post-task interviews conducted immediately after participants complete a guided resume creation task. The interview questions, while researcher-made, are designed to capture reflections on system usability, skill articulation, and perceived alignment of the generated resume with job requirements. The qualitative data will provide insights into how the platform supports users in presenting their technical and soft skills effectively.

To ensure content validity of the researcher-made instruments, the Good and Scates method are applied. For the quantitative instrument, Cronbach's alpha are used to assess internal consistency and reliability of both the PSSUQ-adapted items and the researcher-made subscales.

This combination of qualitative insight, technical testing, and quantitative measurement ensures that the instruments are well-aligned with the study objectives and are capable of accurately capturing user experiences and system performance

Data Gathering Procedure

The data gathering process for this study was conducted in multiple stages, following a systematic, ethically responsible, and sequential approach aligned with the study's developmental and mixed-methods research design. Data collection spans the platform's design, implementation, and evaluation phases to ensure a well-rounded understanding of AICA's effectiveness and usability.

The qualitative data for the needs assessment involved 10–15 participants selected purposively from the same pool of eligible respondents for the broader study. This targeted sample size is appropriate for qualitative research aimed at generating rich, contextual insights into user needs. In contrast, the mixed-methods approach for the broader user acceptance evaluation will involve administering a structured survey and post-task interviews to the full sample of approximately 188 participants, ensuring sufficient representation for statistical analysis and generalizability of findings.

In the initial phase, the researchers will conduct qualitative data collection through semi-structured interviews with purposely selected third- and fourth-year technology students and recent graduates. The purpose of these interviews is to explore job search challenges, and expectations for AI-assisted job-matching platforms. A structured interview guide aligned with the study's objectives is used to ensure consistency across participants. All participants were asked to sign an informed consent form prior to participation, which outlines the purpose of the study, voluntary participation, confidentiality, and their right to withdraw at any time. Interview responses were recorded (with consent), transcribed, and analyzed thematically to identify design-relevant user needs and recurring themes. Findings from this phase will directly inform the feature development of AICA.

Following the initial needs assessment and system development, the researchers proceeded to the user acceptance evaluation phase, which integrates both the quantitative and qualitative assessment of the AICA platform. Participants completed a guided resume-building task using AICA, after which post-task interviews are conducted to gather reflections on their experience. These interviews focused on how effectively the system supported them in presenting their technical and soft skills for job matching, and was transcribed and analyzed thematically.

Concurrently, participants completed a structured quantitative survey administered through Google Forms. This survey will include Likert-scale items adapted from the Post-Study System Usability Questionnaire (PSSUQ) to assess dimensions such as system usefulness, information quality, interface quality, and overall satisfaction. In addition, researcher-made items are incorporated to measure satisfaction with the relevance and personalization of the job recommendations generated by the platform. Data from the surveys are gathered through online forms or printed questionnaires and analyzed using descriptive statistics, including frequencies, means, and standard deviations.

All participants across phases are provided with informed consent forms detailing the purpose of the study, their rights as participants, and the measures taken to ensure the confidentiality of their responses. Data are securely stored and anonymized for analysis.

Quantitative data from the survey are processed using spreadsheet or statistical software, while qualitative data from interviews and usability testing are coded and analyzed thematically. This combination of methods allowed the researchers to triangulate findings and develop a comprehensive evaluation of the AICA platform from both technical and user-centered perspectives.

Statistical Treatment

Quantitative data collected through the structured survey, comprising Likert-scale items adapted from the Post-Study System Usability Questionnaire (PSSUQ) and researcher-made questions, are analyzed using descriptive statistics. Frequencies, means, and standard deviations are computed to summarize participant responses across key dimensions of usability, such as system usefulness, and overall satisfaction, as well as perceptions of the relevance and personalization of job recommendations. To ensure the reliability of the quantitative instrument, Cronbach's alpha is calculated for the overall scale and for each subscale to assess internal consistency.

Qualitative data from post-task interviews are subjected to thematic analysis to identify recurring themes, usability barriers, and user perceptions related to the AICA's effectiveness, including the resume builder's support for skill articulation and job alignment.. This qualitative analysis complements the quantitative results by providing deeper insights into user interactions and feedback.

Ethical Considerations

This research strictly adhered to the ethical principles of transparency, confidentiality, informed consent, and respect for persons.

First and foremost, data privacy and confidentiality were the number one concern. Users are required to input personal and professional information for the career alignment, which has to be securely stored and protected in accordance with data protection laws such as the Philippine Data Privacy Act of 2012. The system had to ensure that user data was not shared with third parties without explicit consent and that anonymization measures were in place when aggregating data for system evaluation or research purposes.

Additionally, informed consent has to be obtained from all participants who interacted with the AICA platform during its development and evaluation phases. Participants were clearly informed about how their data would be used, the scope and limitations of the AI technologies employed, and their right to withdraw from the study at any time without consequence.

Another ethical concern related to bias and fairness in AI-driven recommendations. Large Language Models (LLM) were also trained on wide datasets that could contain systematic biases. Moreover, algorithmic transparency was essential to build trust among users. While the underlying AI models were complex, the platform had to make efforts to explain, in user-friendly language, how job matches were made and what factors influenced these recommendations.

Moreover, algorithmic transparency is essential to build trust among users. While the underlying AI models were complex, the platform had to make efforts to explain, in user-friendly language, how job matches were made and what factors influenced these recommendations. This transparency empowered users to make informed decisions and encouraged accountability in automated decision-making.

REFERENCES

Alampay, E. A., & Soriano, C. R. R. (2024). Facilitating employment opportunities in the digital space: A study on the activities and services of digital labor market intermediaries in the Philippines.

Almalki, L., Al-Alsheikh, N., Bahareth, F., Abdullah, M., & Alsahafi, R. (2025). BERT-based job recommendation system using LinkedIn dataset. *Journal of Information Systems Engineering and Management, 10*(8s).

Alshamari, M., & Mayhew, P. (2019). Task-based usability evaluation of healthcare systems: A literature review.Journal of Biomedical Informatics, 94, 103188.

Ayers, S. (2020). Introduction to Web Scraping for Data Analysis. *Data Science Journal, 19*(1), 1–10.

Brown, T. B., et al. (2020). Language Models are Few-Shot Learners.

Caliskan, A. (2025, April 25). Gender, race, and intersectional bias in AI resume screening via language model retrieval. *Brookings*.

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications, 10*, Article 567.

Chu-Carroll, J., Beck, A., Burnham, G., Melville, D. O. S., Nachman, D., Özcan, A. E., & Ferrucci, D. (n.d.). Beyond LLMs: Advancing the landscape of complex reasoning. *Elemental Cognition*.

Ciaschi, M., & Barone, M. (2024). Exploring the role of Artificial Intelligence in assessing soft skills. In *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)* (Vol. 39, pp. 573–578). IEEE.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., & Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey.

Guild Solutions Inc. (n.d.). Top 10 job websites in the Philippines.

Kavas, H., Serra-Vidal, M., & Wanner, L. (n.d.). Using Large Language Models and Recruiter Expertise for Optimized Multilingual Job Offer -Applicant CV Matching. Retrieved May 24, 2025.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2021). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. Computers & Education, 104, 18-33.

Kelley, E., Ksoll, C., & Magruder, J. (2023). How do digital platforms affect employment and job search? Evidence from India. *Journal of Development Economics, 166*, 103176.

Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. International Journal of Human–Computer Interaction, 34(7), 577–590.

Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.

Moneymax. (2024, March 4). 16 online job sites in the Philippines.

Mujtaba, D. F., & Mahapatra, N. R. (2025, May 18). Fairness in AI-driven recruitment: Challenges, metrics, methods, and future directions (Version 3). *arXiv*.

Nabeel Seedat, & van. (2024). Matchmaker: Self-Improving Compositional LLM Programs for Table Schema Matching.

National Soft Skills Association. (2025, April 22). The soft skills disconnect.

Patil, P. A., & Kharde, A. (2025). Application of RAG (Retrieval-Augmented Generation) in AI-driven resume analysis and job matching.

Philippine Institute for Development Studies. (2023, July 24). *Skills mismatch hampers Philippine economic growth, experts say*.

Rojas-Galeano, S., Posada, J., & Ordoñez, E. (2022). A Bibliometric Perspective on AI Research for Job-Résumé Matching. *The Scientific World Journal, 2022*, e8002363.

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Salemi, A., & Zamani, H. (2024). Evaluating Retrieval Quality in Retrieval-Augmented Generation.

Saurabh Bhausaheb Zinjad, Bhattacharjee, A., Amey Bhilegaonkar, & Liu, H. (2024). ResumeFlow: An LLM-facilitated Pipeline for Personalized Resume Generation and Refinement. *ArXiv (Cornell University)*.

Sauro, J., & Lewis, J. R. (2023). Quantifying the user experience: Practical statistics for user research (3rd ed.). Morgan Kaufmann.

Similarweb. (2025). Top websites ranking - jobs and employment in Philippines.

Vaishampayan, S., Leary, H., Alebachew, Y. B., Hickman, L., Stevenor, B. A., Beck, W., & Brown, C. (2025). Human and LLM-based resume matching: An observational study. In *Findings of the Association for Computational Linguistics: NAACL 2025* (pp. 4808–4823). Association for Computational Linguistics.

Vaswani, A., et al. (2017). Attention is All You Need.

Wilson, K., & Caliskan, A. (2025, April 25). Gender, race, and intersectional bias in AI resume screening via language model retrieval. *Brookings Institution*.

APPENDICES

Appendix A: Informed Consent

Introduction

We are Nyx Arcana, a group of third-year Computer Science students from the University of St. La Salle – Bacolod. As part of our undergraduate thesis requirement, we are conducting a study titled *AICA (AI Career Assistant): A Job Matching Platform Using a Large Language Model with Retrieval-Augmented Generation.* The purpose of this study is to explore how an AI job alignment tool matches the skills of the users to the relevant careers in the tech industry.

Purpose of the Study

The purpose of this study is to develop and evaluate the AI Career Assistant (AICA), a job matching platform that integrates Large Language Models (LLMs) with Retrieval-Augmented Generation (RAG) to bridge the gap between the skills of tech graduates and the demands of the modern job market. By guiding users through a structured resume-building process and making use of AI to match the declared skills with relevant job opportunities, this tool aims to empower graduating students and recent graduates to better articulate their competencies and improve their career alignment. This study seeks to address the limitations of traditional job portals by introducing an intelligent, user-centered tool that supports an efficient job matching in the tech industry

Participant Selection

Participants in this study are selected based on specific inclusion criteria. The target participants are:

- Third- and fourth-year students currently enrolled in Computer Science or Information Technology programs, and
- Recent graduates (within the last 1–2 years) from related degree programs who are exploring job opportunities in the technology field.

This study uses purposive sampling for current students and convenience sampling for graduates. Only individuals who voluntarily agree and provide informed consent included in the research.

What Participation Involves

If you agree to participate, you will:

- Use the AICA web-based platform to build a resume and explore job matching features.
- Participate in a post-task interview where you will be asked to provide feedback on your experience with the platform's resume builder. The interview will be audio-recorded with your consent.
- Participate in a post-task interview where you will be asked to provide feedback on your experience with the platform's resume builder
- Complete a post-assessment survey, which includes Likert-scale questions adapted from the Post-Study System Usability Questionnaire (PSSUQ) and

additional researcher-made items designed to assess your perceptions of the platform’s usability and the relevance of the job recommendations.

The total participation time is approximately 20 to 30 minutes.

Confidentiality

The information to be collected in this research study are kept private and confidential. This means that we will do our best to not let anyone see or hear the information you give to us while you participate or after. Only the researchers will have access to the information that are collected from the respondents. Likewise, your responses will only be used for research purposes only

Voluntary Participation and Right to Withdraw

Your participation in this study is entirely voluntary. You have the right to refuse to answer any question or withdraw from the study at any time without any consequences. Your decision to participate or not participate will not affect your academic standing or your relationship with the university in any way.

Risks and Benefits

There are no known physical, emotional, or legal risks involved in participating in this study. While there is no monetary compensation, your participation may provide personal insight into your own skill strengths and areas for improvement. Additionally, your feedback will help us develop more effective and engaging tools for skill evaluation in the future.

Who to Contact

If you have questions or concerns about this study, you may contact:

Nathania Elouise A. Santia

Student Researcher

s2200756@usls.edu.ph

Consent Statement

I have read and understood the information provided above. I voluntarily agree to participate in this study. I am aware that my participation is confidential, and I may withdraw at any time without consequence.

[ ] I agree to participate in this study.

Name: ____________________________________

Signature: ____________________________________

Date: _____________________________________