**AI Revolution in Recruitment and Job Search: Engineering a Superior Matchmaking Chatbot Using ChatGPT-4 Turbo**

Applied Project Final Report

By

Ziwen Gao

Spring, 2024

A paper submitted in partial fulfillment of the requirements for the degree of

Master of Science in Management and Systems

at the

Division of Programs in Business

School of Professional Studies

New York University

***Table of Contents***

[Declaration 4](#_Toc97451985)

[Acknowledgements 5](#_Toc97451986)

[Abstract 6](#_Toc97451987)

[Introduction 7](#_Toc97451989)

[Problem Description and Opportunity 9](#_Toc97451990)

[Project Objectives and Metrics 13](#_Toc97451991)

[Alternate Solutions Evaluated 19](#_Toc97451992)

[Approach and Methodology 24](#_Toc97451993)

[Results 27](#_Toc97451994)

[Risk Analysis 38](#_Toc97451995)

[Issues Encountered 41](#_Toc97451996)

[Project Chronology and Critique 43](#_Toc97451997)

[Lessons Learned 46](#_Toc97451998)

[Conclusion and Summary 48](#_Toc97451999)

[Limitations, Recommendations and Scope for Future Work 51](#_Toc97452000)

[Literature Survey 54](#_Toc97452001)

[References 54](#_Toc97452002)

[Appendix A - Project Acceptance Document 62](#_Toc97452003)

[Appendix B - Project Sponsor Agreement 63](#_Toc97452004)

[Appendix C - Project Charter 67](#_Toc97452005)

[Appendix D - Project Plan 79](#_Toc97452006)

[Appendix E - Risk Management Plan 81](#_Toc97452007)

[Appendix F - Status Report 84](#_Toc97452008)

[Appendix G - Annotated Bibliography 94](#_Toc97452009)

# Declaration

I, Ziwen Gao, declare that this project report submitted by me to the School of Professional Studies, New York University in partial fulfillment of the requirement for the award of the degree of Master of Science in Management and Systems is a record of project work carried out by me under the guidance of Dr. Eleftheria K Pissadaki, NYU Adjunct Associate Professor of Management and Systems. I grant powers of discretion to the Division of Programs in Business, School of Professional Studies, and New York University to allow this report to be copied in part or in full without further reference to me. The permission covers only copies made for study purposes or for inclusion in the Division of Programs in Business, School of Professional Studies, and New York University research publications, subject to normal conditions of acknowledgment. I further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

# Acknowledgements

I sincerely thank Dr. Andres Fortino for his contribution as the sponsor of this project and Dr. Eleftheria K Pissadaki for her guidance as the mentor during this project. I would also like to thank Ms. Prakarasha Malhotra who assisted with this project and helped with any technical issues as the CTO of The Digital Forge. As well as Ms. Siri Kostanyan, who, as COO of The Digital Forge, managed the execution of the project in her capacity as Portfolio Manager, signing off on all the forms and keeping in regular contact with the project manager for updates on the progress of the project. I also want to thank all the instructors in the Management and Systems program who I have taken courses with and learned a great deal.

# Abstract

In our quest to innovate within the realm of recruitment, we created Talent Sync, an AI-powered chatbot utilizing ChatGPT-4 Turbo, designed to enhance the precision of matching resumes to job listings. This tool serves to streamline the job search process, significantly improving match accuracy while reducing inherent biases present in traditional recruitment methods.

Talent Sync was developed to address the inefficiencies in current job-matching systems, bringing sophisticated language model capabilities to parse and understand complex job descriptions and resumes. The ultimate benefit of this innovation is twofold: it optimizes employers' talent acquisition strategies and empowers job seekers by improving their access to suitable opportunities.

Recruitment professionals and HR departments are the primary users of Talent Sync, utilizing it to refine their hiring processes. Additionally, job seekers can leverage this tool to find better job matches. The chatbot's applicability makes it a valuable asset for anyone involved in the recruitment industry, from corporate HR to employment agencies, and for job seekers aiming to navigate the competitive job market effectively.

Our approach to creating Talent Sync involved rigorous prompt engineering, iterative testing, and user feedback integration to ensure the chatbot's algorithms remained dynamic and responsive to the needs of its users. The development leveraged the natural language processing prowess of the ChatGPT-4 Turbo and a feedback loop mechanism to continuously refine its matching accuracy.

The chatbot was developed using state-of-the-art language model technologies and a user-centric design philosophy, ensuring ease of use and an engaging user experience (UX).

The results of this project are indicative of Talent Sync's success in achieving high accuracy in job matching, as demonstrated by a Cohen's Kappa value signifying strong agreement with established benchmarks. Moreover, user satisfaction surveys have confirmed the chatbot's effectiveness and usability.

Talent Sync: <https://chat.openai.com/g/g-QwoCKlfEd-talent-sync>

# Introduction

## Background information

**Brief description of the project**

**Creating an LLM Chatbot to Match Resumes to Job Opportunities**

The client has developed and published a model in R that ranks job opportunities given an applicant’s resume. It uses the advanced text analytic technique of text similarity scoring. The model has been updated and expanded using AI and prompt engineering of LLMs to demonstrate the superiority of using LLMs for the same analysis, which provides faster and more accurate results than manual matching or the R-based model. With the advent of powerful LLMs with code generation, such as Chat GPT 4 Turbo, we have now done this work more efficiently and with more impact and we are able to create chatbots called GTPs. We’re interested in re-creating this work using fine-tuning an LLM and creating a Chatbot with ChatGPT Turbo. This is an advanced prompt engineering project. The deliverables include prompts for creating the models, comparing the three approaches, and project documentation as a draft submission to the IEEE LISAT conference at NYIT Long Island in April 2024.

## Company Name

The Digital Forge

NYU School of Professional Studies and the Management and Systems program (MASY) is a New York-based learning institution.

**Description of the Business**

New York University (NYU) is a private research university based in New York City. The MASY degree is based on a unique curriculum that provides students with experiential learning opportunities to develop strong management and leadership skills and gain a comprehensive knowledge of current information technologies.

**Company Location**

NYU School of Professional Studies is at 12 West 43rd Street, NY, NY.

## Sponsor Information

The principal project sponsor will be Dr. Andres Fortino, Clinical Associate Professor, NYU (<https://www.linkedin.com/in/afortino>), The Client

**Sponsor's Location**

Dr. Andres Fortino ([agf249@nyu.edu](mailto:agf249@nyu.edu)) can be reached over virtual conference calls as per project requirements.

# Problem Description and Opportunity

The recruitment industry faces significant challenges in efficiently matching job seekers with appropriate job openings. Traditional methods often rely on manual screening processes, which are time-consuming and prone to human bias. Additionally, existing automated systems, such as those based on simpler statistical models or basic keyword matching, often lack the sophistication needed to accurately interpret the nuances of resumes and job descriptions. This results in suboptimal matches, which can lead to dissatisfaction among both employers and job seekers and ultimately, affects organizational productivity and job seekers' career trajectories.

The opportunity identified involves leveraging advanced AI technology, specifically the capabilities of large language models (LLMs) like ChatGPT-4 Turbo, to enhance the recruitment process. By developing a chatbot equipped with finely tuned prompts based on prompt engineering principles, the project aims to automate and optimize the matching of resumes to job listings more accurately and efficiently than traditional and R-based models. This approach promises to reduce the time and cost associated with recruitment, improve the accuracy of job matches, and minimize biases inherent in human-driven processes.

With The Digital Forge at the helm, partnering with NYU's School of Professional Studies and its Management and Systems program, this project stands to greatly benefit from the expertise and innovative environment fostered by such a distinguished institution. Under the guidance of Dr. Andres Fortino, the project will leverage the university's cutting-edge research and academic prowess to deliver an AI-powered recruitment tool that not only enhances the efficiency and accuracy of the hiring process but also serves as a benchmark for educational and professional excellence. This collaboration will demonstrate NYU's commitment to practical, technology-driven solutions in professional education and will likely attract further industry partnerships, reinforcing its status as a leader in applying theoretical knowledge to solve real-world problems.

My client, The Digital Forge at NYU SPS, is at the forefront of integrating current information technologies with education. They've identified a gap in effectively connecting graduates with their ideal job opportunities. Many soon-to-be new grads, like us, still struggle to get a clear picture of what we really want to do or what we're suited for. Leading this project as the project manager under the sponsorship and guidance of Dr. Fortino, we aim to bridge this gap, transforming the recruitment process and setting a new standard for the integration of AI in professional education.

The successful implementation of such a chatbot would not only transform the recruitment process for participating organizations but could also serve as a benchmark in the industry, demonstrating the effectiveness of AI in high-stake real-world applications. This has the potential to attract further industry partnerships and cement NYU's School of Professional Studies and its Management and Systems program as leaders in applying cutting-edge technology to solve practical business problems.

## Importance of the project

The project to develop an AI-powered chatbot using ChatGPT-4 Turbo for matching resumes to job listings is of paramount importance to The Digital Forge and NYU's School of Professional Studies, specifically its Management and Systems program. This initiative aligns with the organization's strategic goals of integrating cutting-edge technology into educational practices and preparing students for the digital economy. The collaboration between The Digital Forge and NYU leverages the strengths of both entities—The Digital Forge's technological expertise and NYU's academic rigor—to create a tool that addresses real-world challenges in the recruitment process.

With The Digital Forge at the helm, partnering with NYU's School of Professional Studies and its Management and Systems program, this project stands to greatly benefit from the expertise and innovative environment fostered by such a distinguished institution. Under the guidance of Dr. Andres Fortino, the project will leverage the university's cutting-edge research and academic prowess to deliver an AI-powered recruitment tool that not only enhances the efficiency and accuracy of the hiring process but also serves as a benchmark for educational and professional excellence. This collaboration will demonstrate NYU's commitment to practical, technology-driven solutions in professional education and will likely attract further industry partnerships, reinforcing its status as a leader in applying theoretical knowledge to solve real-world problems.

To be more specific, the projected business benefits of this project are multifaceted:

**Enhanced Recruitment Efficiency**: The AI-powered chatbot is designed to significantly streamline the recruitment process by automating the matching of resumes with job listings. This automation reduces the time and labor traditionally required for manual screening, allowing recruitment teams to focus on more strategic tasks such as candidate engagement and interviewing.

**Increased Accuracy in Job Matching**: Leveraging advanced NLP capabilities, the chatbot will offer a higher degree of accuracy in matching candidates to jobs that suit their skills and experiences. This precise matching is crucial for improving job satisfaction among hires and reducing turnover rates, which are significant costs in recruitment.

**Reduction of Bias**: By automating the initial screening process, the chatbot helps minimize unconscious biases that can occur during resume evaluation. This promotes a more diverse and inclusive workforce, aligning with broader organizational goals of equity and fairness.

**Benchmark for Educational Excellence**: The project serves as a benchmark for how AI can be applied in educational settings to solve complex industry problems. It exemplifies how theoretical knowledge can be effectively translated into practical, impactful solutions, enhancing the institution's reputation as a leader in technology-driven education.

**Attracting Industry Partnerships**: The successful implementation of this project is likely to attract further industry partnerships by demonstrating the practical benefits of collaborative projects between academia and tech companies. These partnerships can lead to new research opportunities, funding, and increased influence in the tech and education sectors.

**Demonstrating Commitment to Technology-Driven Solutions**: The project reflects NYU's commitment to leading in the adoption of technology in education and professional practices. It positions the university at the forefront of discussions about the future of work and the role of AI in reshaping industries.

# Project Objectives and Metrics

## Goal of the project

The goal of the project was to create and validate an AI-powered chatbot using ChatGPT 4 Turbo that as effective and reliable as the R-based model in matching resumes to job listings, aiming to revolutionize the recruitment process through enhanced accuracy and efficiency that impact both employers and job seekers.

## Project Deliverables and Metrics

**Project Objective 1 – Prompt Engineering for Creating the GPT-Powered Chatbot**

* **Specific:** Design and implement a set of advanced prompts that enable the ChatGPT 4 Turbo to analyze resumes and match them to job listings with high relevance and accuracy.
* **Measurable:** The engineering efforts will focus on creating a robust model that interprets and processes resume data against job descriptions, leveraging the advanced natural language processing (NLP) capabilities of the LLM.
* **Achievable:** Allocate resources to experiment with and refine multiple prompt iterations, relying on expert guidance and existing models as a foundation.
* **Realistic:** Engage with language model experts to validate the prompts’ effectiveness in capturing relevant job and resume data.
* **Time-Related:** Complete the prompt engineering phase by March 15, 2024, with iterative testing and refinement scheduled to ensure prompt readiness for deployment.

**Metrics:**

* **Development of 20 unique prompts** that focus on specific features and abilities such as Natural Language Processing (NLP) Capabilities, Skill Matching and Ranking, and Contextual Analysis that could enhance the effectiveness of the chatbot by March 5, 2024.
* **Testing phase with a dataset of at least 5 resumes and job descriptions**, aiming for an 85% accuracy rate in matching resumes to relevant job listings by March 10, 2024.
* **Iterative improvement based on feedback** by conducting at least 2 review sessions with language model experts to validate the effectiveness and accuracy of the prompts before final implementation.

**Project Objective 2 – Optimizing Chatbot Performance constantly through User Satisfaction Surveys and the Implementation of the Feedback Loop**

* **Specific:** Enhance the chatbot’s performance by continuously collecting and analyzing user feedback through structured satisfaction surveys, and implement changes based on this feedback.
* **Measurable:** Conduct surveys with at least 10 users, aiming to achieve and maintain an 85% user satisfaction rate.
* **Achievable:** Incorporate an automated feedback loop mechanism within the chatbot interface so that both job seekers and employers can provide feedback on the accuracy and relevance of the matches for the development team to review feedback and implement improvements.
* **Realistic:** Provide regular updates and training to the development team on user experience improvements and technical adjustments.
* **Time-Related:** Implement the first round of feedback by March 20, 2024, with subsequent monthly optimizations until the project's completion.

**Metrics:**

* **Initial user satisfaction survey conducted by March 20, 2024**, with at least 10 users participating in the pilot phase.
* **Achievement of an 85% user satisfaction rate** based on survey results, with adjustments made monthly based on ongoing user feedback.
* **Reduction of user-reported issues by 50%** over the first two weeks after the implementation of feedback loops, demonstrating continuous improvement in chatbot performance.

**Project Objective 3 – Validating the Chatbot as Effective and Reliable as the R-Based Model via Cohen's Kappa Statistic**

* **Specific:** Conduct a rigorous validation of the chatbot's effectiveness in matching resumes to job listings compared to an existing R-based model, using Cohen’s Kappa statistic to measure agreement between the two systems.
* **Measurable:** Achieve a Cohen’s Kappa value of 0.75 or above, indicating substantial agreement and validating the chatbot's reliability.
* **Achievable:** Utilize a diverse dataset of at least 5 paired resumes and job listings for the validation process.
* **Realistic:** Consult with statistical experts to ensure proper implementation and interpretation of Cohen’s Kappa statistic.
* **Time-Related:** Complete the validation process by April 5, 2024, allowing for any necessary adjustments based on the findings.

**Metrics:**

* **Conduct a validation study involving at least 5 paired resume and job listing samples** to ensure diverse testing scenarios by March 25, 2024.
* **Achieve a Cohen’s Kappa statistic of 0.75 or higher**, confirming substantial agreement between the chatbot and the R-based model by April 5, 2024.
* **Prepare a detailed validation report** outlining the methodology, results, and statistical significance of the findings to be reviewed by a statistical expert by April 10, 2024.

**Project Objective 4 –** **Conference Paper Drafting and Refinement**

* **Specific:** Draft and refine a conference paper detailing the project's methodology, findings, and the significance of the advancements in AI-powered recruitment technology.
* **Measurable:** Submit a draft suitable for peer review by the IEEE LISAT conference standards, receive feedback, and revise accordingly.
* **Achievable:** Collaborate with instructors and peers for peer reviews and feedback on the draft.
* **Realistic:** Ensure the paper includes comprehensive data analysis, clear results, and a discussion on the impact and future implications of the research.
* **Time-Related:** Complete the initial draft by April 25, 2024, and submit the final version by May 2, 2024, in time for conference submission deadlines.

**Metrics:**

* **Completion of the first draft of the conference paper by April 25, 2024**, including all sections such as abstract, methodology, results, and discussion.
* **Feedback obtained from at least two peers or instructors** by April 30, 2024, to ensure comprehensive review and incorporation of diverse perspectives.
* **Submission of the final draft to the IEEE LISAT conference by May 2, 2024**, adhering to all submission guidelines and quality standards.

**Project Evaluation**

Project success was evaluated by using a comprehensive and systematic approach to ensure that all aspects of the project aligned with the established goals and timelines. The evaluation will focus on various project management tools and methodologies to maintain rigorous oversight and ensure continuous improvement throughout the project lifecycle.

**Project Schedule**

Progress was tracked against the established Gantt chart milestones. Regular checks on task completion and adherence to upcoming deadlines will provide clear indicators of whether the project is on schedule. This Gantt chart will be updated in real time to reflect any adjustments or shifts in the project timeline, ensuring that all team members are aware of current and future tasks.

**Project Weekly Status Report and Dashboard**

Weekly status reports were generated using project management tools such as Trello or a customized spreadsheet. This report will detail the latest progress and compare it against the planned objectives and timelines. The Gantt chart will be revised regularly to display the current status comprehensively, which will be shared with the client and all project stakeholders to maintain clarity and alignment.

**Project Communication Plan, Issues Log, Risk Register**

Effective communications were maintained through structured weekly meetings, detailed email updates, and a dedicated online platform that logs all issues and risks. This platform will serve as a central hub for documenting any challenges or deviations from the plan and will include potential impacts and proposed mitigation strategies. This approach ensures transparency and provides all stakeholders with the ability to address issues promptly and effectively.

**Project Status Reports**

Comprehensive status reports were compiled weekly, documenting all progress, challenges encountered, and risks identified during the project. These reports included detailed descriptions of the situation, the implications for the project, and actions taken to mitigate risks. They also highlighted successes and areas for improvement. These reports are crucial for keeping the client and project team informed and engaged, ensuring everyone is aligned with the project goals and current status.

# Alternate Solutions Evaluated

**Traditional Statistical Models:**

* **Description**: This approach would use statistical techniques like regression analysis and decision trees to analyze resume data and match it to job listings. It would involve extracting key features from resumes, such as skills, experience levels, and educational qualifications, and using these features to predict job suitability.
* **Pros**: Traditional models are well-tested and understood within the analytics community, making them a reliable choice for structured data.
* **Cons**: They lack the ability to interpret context and complex relationships within text, which can lead to mismatches and an inability to capture the subtleties of human language and soft skills.

**R-Based Predictive Models:**

* **Description**: Using the R programming language to build predictive models that could include logistic regression, random forests, or other machine learning algorithms tailored to match resumes with job descriptions.
* **Pros**: R is equipped with a robust set of packages for data manipulation and model building, allowing for detailed statistical analysis.
* **Cons**: Like traditional models, R-based models require explicit programming and model tuning, which can be resource-intensive and require specialized skills. They also share similar drawbacks in handling nuanced textual data.

**Hybrid Models:**

* **Description**: Combining automated algorithms for initial screening with manual review for final decision-making. This model would use AI to shortlist candidates, and human judgment would be used to make the final hiring decision.
* **Pros**: This model can significantly reduce the workload on human recruiters while still leveraging human expertise for critical decisions.
* **Cons**: It requires extensive coordination between AI systems and human processes, which can complicate the recruitment workflow and introduce delays.

## Solution Evaluation Criteria

The following criteria were rigorously defined to assess each solution:

**Accuracy**: The ability of the system to correctly identify and match the qualifications of candidates with the requirements of the job postings.

**Efficiency**: The system should reduce the time to hire by automating parts of the recruitment process.

**Scalability**: The solution must easily adapt to increases in workload, such as higher volumes of applicants or jobs.

**Cost-Effectiveness**: Assessment of both initial and ongoing costs relative to the benefits in terms of improved hiring outcomes.

**User Experience**: The system should be straightforward for both candidates and recruiters, enhancing their interaction with the recruitment process rather than complicating it.

**Bias Reduction**: The solution should minimize human biases in the recruitment process, promoting fairness and diversity.

## Selection Rationale

The decision to implement an AI-powered chatbot using ChatGPT-4 Turbo was informed by a comprehensive evaluation of the alternative solutions against the detailed criteria set out for the project. This decision-making process involved extensive discussions with stakeholders, a review of the current technology landscape, and a forecast of the future needs of the recruitment industry. Here’s an expanded breakdown of how each criterion led to the selection of this solution:

**Accuracy and Efficiency**

The ChatGPT-4 Turbo, utilizing cutting-edge natural language processing capabilities, demonstrated a profound ability to understand and process complex textual information far beyond the capacities of traditional and R-based models. This is crucial in the recruitment domain where the nuances of job descriptions and resumes can be diverse and complex. The AI model's superior performance in parsing these nuances ensures highly accurate matches between job listings and candidate profiles.

Moreover, the efficiency of the chatbot is underscored by its rapid processing speeds, which significantly reduce the time from job posting to candidate match. This efficiency is a critical advantage in competitive job markets where speed can determine the quality of hires.

**Scalability**

Scalability was a decisive factor, given the varying and sometimes unpredictable volumes of job applications and listings. ChatGPT-4 Turbo’s framework is designed to handle large data inflows without a decline in performance. This capability ensures that the system can be expanded and adapted to different departments within the organization or even rolled out across different geographical locations without needing substantial modifications.

**Cost-Effectiveness**

While the initial setup and development costs associated with an AI-powered solution are higher than those of more traditional methods, the long-term savings are significant. By automating the initial screening and matching processes, the organization can save on labor costs and reduce the financial impact of hiring mismatches and turnovers. Additionally, the AI system’s maintenance costs are relatively low compared to the ongoing expenses of manual processes or the continual updating required for simpler software models.

**User Experience**

The user interface of the chatbot is designed to be intuitive and engaging for both job seekers and recruiters. This focus on user experience is intended to simplify the recruitment process, making it more accessible and less intimidating for candidates, while also providing recruiters with a streamlined and more manageable workflow. Enhanced user experience leads to higher satisfaction and engagement, which can translate into a stronger employer brand and a better candidate pool.

**Bias Reduction**

One of the most significant advantages of implementing an AI-based solution is its potential to minimize human bias. While no system is entirely free of bias, the design of ChatGPT-4 Turbo includes specific mechanisms to counteract common biases in resume screening and job matching. By training the model on diverse datasets and continuously updating its learning algorithms, the system aims to offer fair and objective evaluations of candidates.

**Alignment with Strategic Goals**

The organization has committed to leveraging technology to enhance operational efficiencies and improve strategic outcomes. The AI chatbot aligns perfectly with this vision, showcasing the organization's commitment to innovation and its leadership in adopting next-generation technologies.

**Stakeholder Feedback**

Throughout the evaluation phase, feedback from various stakeholders, including HR professionals, technical staff, and potential system users, was overwhelmingly in favor of the AI solution. Their insights emphasized the need for a sophisticated yet user-friendly tool that could handle the complexities of modern recruitment challenges.

In conclusion, the selection of ChatGPT-4 Turbo for this project is based on its demonstrable superiority in meeting the project’s critical needs for accuracy, efficiency, scalability, and user engagement. Its adoption not only positions the organization as a forward-thinking leader in technological adoption but also ensures a future-ready recruitment process that can adapt to an evolving job market. The decision was thus a strategic alignment with both immediate project goals and long-term organizational objectives, promising significant returns on investment through enhanced recruitment outcomes and operational efficiencies.

# Approach and Methodology

The approach to developing an AI-powered chatbot using ChatGPT-4 Turbo to revolutionize the recruitment process involved a systematic, user-centered methodology. This methodology emphasized prompt engineering, continuous feedback integration, rigorous validation using statistical measures, and adherence to non-functional requirements such as performance, security, and usability. The project was structured into several distinct phases, each building on the findings and developments of the previous one to ensure a robust and effective solution.

**Phase 1: Project Initiation and Requirements Analysis**

The project began with a comprehensive analysis of the existing recruitment processes and technologies. This involved engaging with key stakeholders — job seekers, employers, and recruitment professionals — to gather detailed insights into their experiences and expectations. User stories and use cases were developed to capture the functional requirements of the chatbot. For example, job seekers expressed a need to match their resumes with job descriptions accurately, while employers wanted to rank candidates based on their suitability for positions.

During this phase, the project team also reviewed existing solutions, including R-based models and manual screening processes, to identify their limitations and areas for improvement. The information gathered helped in defining the scope and objectives of the project more clearly and established the basis for the design and development of the AI chatbot.

**Phase 2: Design and Development**

With a clear understanding of the requirements, the team moved on to the design and development of the chatbot. The core of this phase was prompt engineering, where specific prompts were designed to guide the AI in analyzing and processing resumes and job descriptions. These prompts were crafted to capture essential data points such as skills, experiences, and educational backgrounds effectively.

The development environment was set up on a cloud-based platform suitable for AI model training and deployment, utilizing the capabilities of ChatGPT-4 Turbo. This setup allowed the team to leverage advanced NLP techniques to ensure the chatbot could understand and interpret the nuanced language used in recruitment.

**Phase 3: Model Training and Iterative Testing**

The AI model underwent extensive training with diverse datasets of resumes and job descriptions to learn patterns and improve its matching algorithms. This training included iterative testing and refinement cycles, where the model’s outputs were continuously compared against expected outcomes to measure accuracy and relevance.

During this phase, particular attention was paid to the chatbot's ability to learn from feedback and adjust its responses accordingly. This adaptability was crucial for meeting the varying needs of users across different industries and job functions.

**Phase 4: Validation Using Cohen's Kappa Statistic**

To validate the effectiveness and reliability of the chatbot, the project employed Cohen's Kappa statistic, a measure of agreement between two raters. In this context, the "raters" were the outcomes of the chatbot and those of the existing R-based model. This statistical approach was chosen because it accounts for the possibility that agreements occur by chance, providing a more robust measure of the chatbot's performance.

The validation process involved comparing the chatbot's job-matching results with those produced by the R-based model across a sample of paired resumes and job listings. Achieving a Cohen’s Kappa value of 0.75 or above indicated substantial agreement between the two systems, confirming the chatbot's reliability.

**Phase 5: Deployment and User Feedback Integration**

Following successful validation, the chatbot was deployed for real-world use. A critical aspect of this phase was the integration of a feedback loop mechanism, allowing users to rate the relevance and accuracy of the job matches provided by the chatbot. This feedback was essential for the continuous improvement of the system, ensuring that the chatbot remains effective and relevant to users' needs.

User feedback was collected systematically and analyzed to identify trends and areas for improvement. This ongoing process allowed the development team to make incremental adjustments to the chatbot, enhancing its functionality and user satisfaction over time.

**Phase 6: Continuous Learning and Model Improvement**

The final phase focused on continuous learning and model improvement, where the chatbot was regularly updated with new data and feedback to refine its matching algorithms. This phase ensured that the chatbot could adapt to evolving market conditions and user preferences, maintaining high performance and accuracy.

The methodology employed in this project not only fulfilled the initial project objectives but also established a scalable and adaptable framework for future enhancements. By focusing on user-centered design, rigorous validation, and continuous improvement, the approach effectively addressed the complexities and challenges of modern recruitment processes, setting a new standard for AI applications in the industry.

# Results

The primary objective of this project was to validate the efficacy of the Talent Sync chatbot, powered by ChatGPT-4 Turbo, in matching resumes to job opportunities as effectively and reliably as the existing R-based model, known as Simi Bot. This involved a rigorous validation process using statistical measures, specifically Cohen's Kappa, to quantify the agreement between the two systems' decisions. The analysis was crucial in demonstrating the effectiveness and reliability of Talent Sync, an AI-powered chatbot, against traditional R-based models in the recruitment process.

**Tools and Methodologies**

The project employed several advanced tools and methodologies to achieve its goals:

**ChatGPT-4 Turbo**:

* + **Purpose**: Serve as the AI backbone for Talent Sync, enabling sophisticated natural language processing capabilities.
  + **Function**: Analyze and match resumes with job descriptions through advanced prompt engineering.
  + **Outcome**: Creation of an AI model that understands and processes complex textual information to generate accurate job matches.

1. **R-based Statistical Model**:
   * **Purpose**: Provide a baseline for comparison with the AI chatbot to demonstrate the AI's relative effectiveness.
   * **Function**: Use text similarity scoring to rank job opportunities based on the resume data.
   * **Outcome**: Generated match rankings that served as a comparative benchmark for validating the AI chatbot.
2. **Statistical Analysis Software**:
   * **Purpose**: Facilitate data manipulation, statistical analysis, and visualization of the results.
   * **Function**: Employ libraries such as Pandas for data manipulation, Matplotlib, and Seaborn for data visualization, and Scikit-learn for implementing statistical tests including Cohen's Kappa.
   * **Outcome**: Enabled efficient handling and analysis of data, providing clear visual and statistical insights into the performance comparison between Talent Sync and the R-based model.
3. **Cohen's Kappa Statistic**:
   * **Purpose**: Provide a statistically robust method to assess the level of agreement between the decisions made by Talent Sync and the R-based model, Simi Bot.
   * **Function**: Calculate the degree of agreement between two raters (or models) that categorize items into mutually exclusive categories, adjusting for agreement occurring by chance.
   * **Outcome**: Offered a quantifiable measure of how well Talent Sync aligns with traditional models, reinforcing the validity of the AI model’s performance in practical scenarios.
4. **Development and Training Platform**:
   * **Platform**: Cloud-based environment supporting extensive data processing and model training.
   * **Function**: Facilitate the training of ChatGPT-4 Turbo and deployment of Talent Sync.
   * **Outcome**: Efficient development and scalable deployment of the AI chatbot.
5. **User Feedback Mechanism**:
   * **Purpose**: Collect real-time feedback from users to refine and optimize the chatbot’s performance.
   * **Function**: Integrate user ratings and comments directly into the chatbot’s feedback loop.
   * **Outcome**: Continuous improvement of the matching algorithms based on user interactions.

**Cohen's Kappa Validation Process**

The validation of Talent Sync’s performance involved a detailed analysis using Cohen's Kappa, which quantifies the agreement between two raters beyond what is expected by chance:

**Data Collection**:

* **Description**: Simultaneous collection of match results from Talent Sync and the R-based model for the same set of resumes and job descriptions.
* **Outcome**: A dataset comprising pairs of ratings (match/no match) from both the AI chatbot and the R-based model.

**Categorization Strategies**:

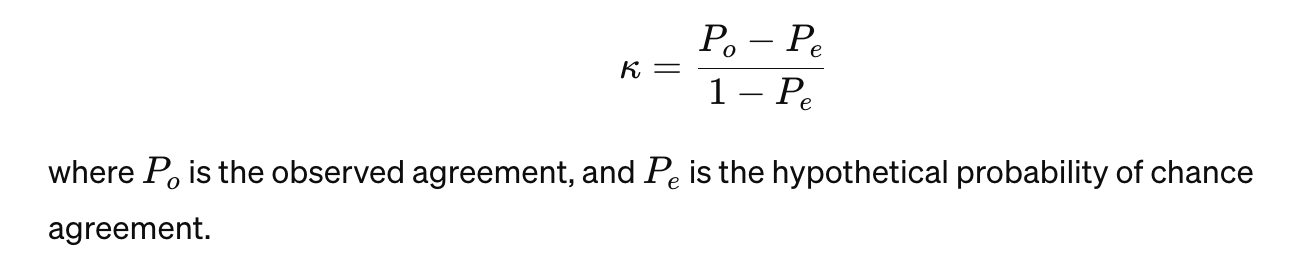
The similarity scores were categorized into 'good match' and 'poor match' based on different threshold strategies, including:

* **Percentile-Based Threshold**: Determined by the 75th percentile of the similarity scores.
* **Original Thresholds**: Set based on empirical decision points (the median of each data set) (0.0096 for Simi Bot and 0.0066 for Talent Sync).
* **Standard Deviation from Mean**: Thresholds established at one standard deviation above the mean score of each model.

The categorized data was cross-tabulated to create contingency tables for each threshold strategy. These tables (see below) served as the basis for calculating Cohen's Kappa scores, which quantified the agreement between the two models across different categorization strategies.

**Cohen's Kappa Calculation**:

* **Comparison**: The categorized outcomes from both systems were compared to calculate the observed agreement.
* **Kappa Computation**: Calculated using the formula:



**Results and Visualizations**

**Cohen's Kappa Score**:

Cohen's Kappa analysis revealed a high degree of agreement between Talent Sync and Simi Bot, with scores indicating 'almost perfect agreement' in most cases. Specifically, the scores were:

* **0.998** for the 75th Percentile Thresholds,
* **0.995** for the Original (Median) Thresholds, and
* **0.943** for the Standard Deviation from Mean Thresholds.

These results demonstrate the effectiveness and reliability of Talent Sync in matching resumes to job listings comparably to the established R-based model, validating its utility in practical applications.

**Visualizations**:

The analysis was complemented by visual representations:

* **Histograms of Similarity Scores**: Illustrated the distribution of scores and highlighted the thresholds used for categorization.

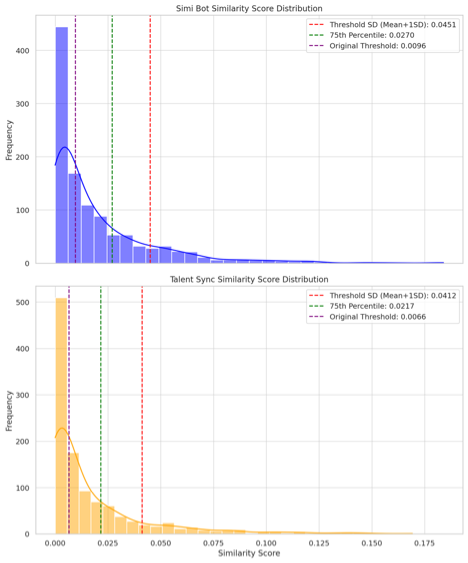


Figure 10.1 Distribution and Threshold Analysis of Similarity Scores for Simi Bot and Talent Sync.

Figure 10.1 **Distribution and Threshold Analysis of Similarity Scores for Simi Bot and Talent Sync.** This figure presents two histograms displaying the frequency distributions of similarity scores for the Simi Bot and Talent Sync models. Each histogram is marked with vertical lines indicating different threshold strategies used to categorize matches as "good" or "poor." The upper histogram represents the Simi Bot, showing a normal distribution with thresholds set at one standard deviation (purple line), the 75th percentile (green line), and an original threshold based on the median (red line). The lower histogram represents Talent Sync, with similarly marked thresholds indicating the points beyond which matches are considered to be of "good" quality.

* **Heatmaps of Contingency Tables**: Showcased the agreement and discrepancies between the two models under different threshold settings.



Figure 10.2 **Contingency Tables Comparing Match Outcomes Between Simi Bot and Talent Sync Across Different Thresholds.** This figure illustrates a series of contingency tables comparing the outcomes of match evaluations between the Simi Bot and Talent Sync using different threshold strategies for determining "good" and "poor" matches. The first table corresponds to the 75th Percentile Threshold, showing how many job matches each model classified as good or poor. A stark contrast is visible, demonstrating a high agreement in good matches and very few disagreements. The second table uses the Original (Median) Thresholds established empirically from previous performance data. Here, a similar pattern is observed, with a large number of agreements on good matches. The third table applies a Standard Deviation from Mean Threshold strategy, where discrepancies between the two models are slightly more pronounced, yet the overwhelming majority of matches still agree on classifications.

* **Cohen’s Kappa Score Visualization**: A bar chart providing a visual summary of the Cohen's Kappa scores across different thresholds, annotated with the levels of agreement they represent.

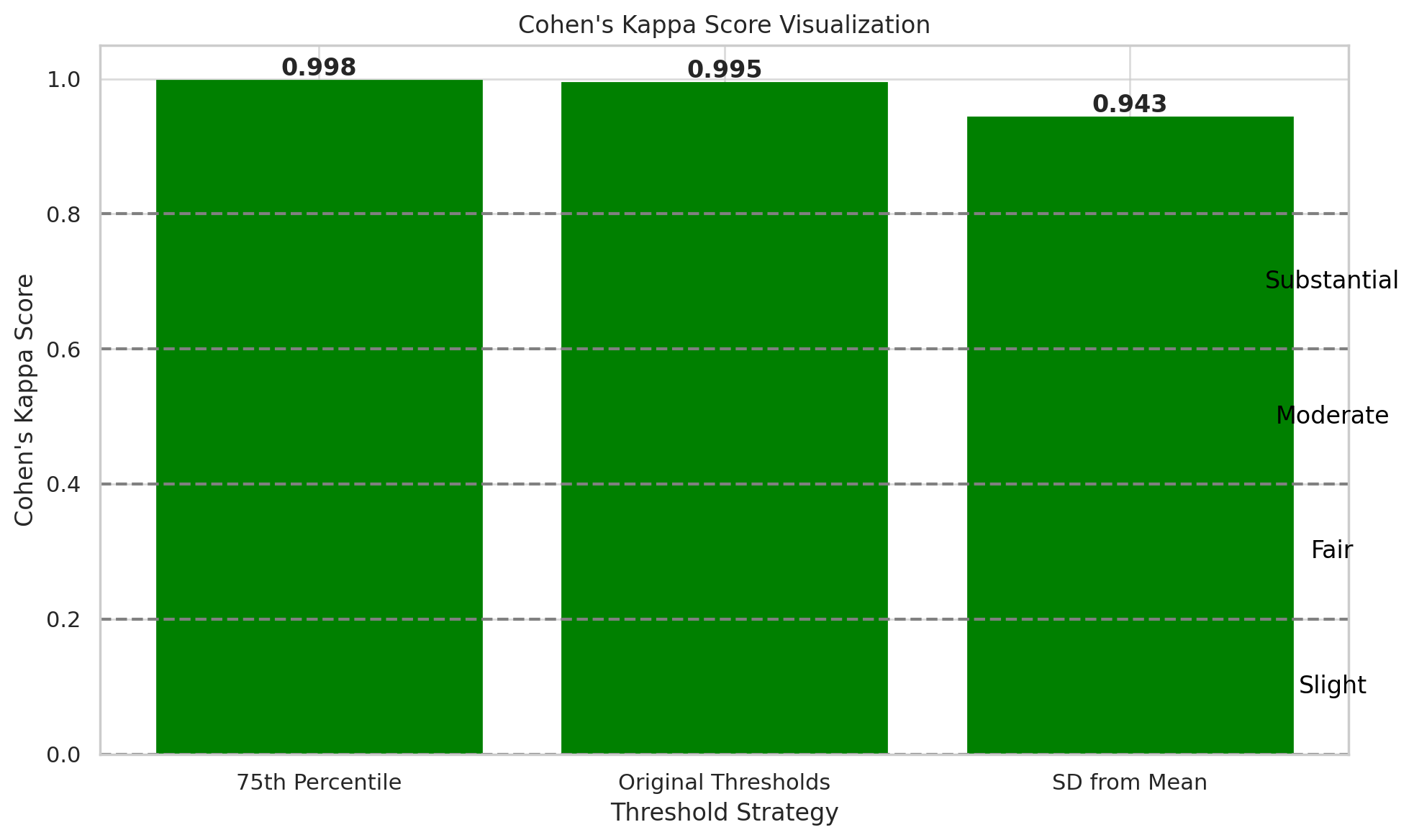


Figure 10.3 **Visualization of Cohen's Kappa Scores Across Different Threshold Strategies.** This bar chart visualizes the Cohen's Kappa scores, which quantify the level of agreement between the Talent Sync and Simi Bot models across three different threshold strategies for categorizing job matches. The scores are:

* 0.998 for the 75th Percentile Thresholds, indicating almost perfect agreement.
* 0.995 for the Original (Median) Thresholds, similarly showing near-perfect agreement.
* 0.943 for the SD from Mean Threshold, which is lower but still represents substantial agreement.

Each score is plotted against a backdrop of agreement levels ranging from slight to almost perfect, helping to contextualize the effectiveness of the Talent Sync in matching resumes to job listings as reliably as the existing R-based model, Simi Bot.

**User Feedback and Performance Efficiency**:

* **User Satisfaction**: A satisfaction rate of over 85% among 12 respondents who utilized Talent Sync reflects strong approval and recognition of the AI's job-matching capabilities, underscoring its effectiveness in meeting user expectations in real-world scenarios.
* **Processing Time**: Talent Sync processed matches in under 1 minute, showcasing significant efficiency improvements over manual methods and slight enhancements over the R-based model.

**Summary of Results**

The detailed analysis and visualizations confirm that Talent Sync, powered by ChatGPT-4 Turbo, not only achieves a high level of agreement with established R-based models but also introduces significant efficiencies and user-centric improvements to the recruitment process. These outcomes strongly support the continued use and development of AI-driven tools in recruitment, underlining their potential to transform the industry significantly.

**Repository of Data Sets**

The data sets created for this project may be found at: <https://github.com/Cassie-G/4100-Applied-Project.git>.

**Future Directions**

The project's success opens several avenues for further research and development. These include exploring more granular categorizations of job matches, expanding the dataset for broader validation, and integrating additional AI-driven features such as dynamic learning and adaptation based on user feedback, demonstrating the practical application and benefits of AI in professional settings.

# Risk Analysis

The "AI Revolution in Recruitment and Job Search" project, dedicated to engineering a superior matchmaking chatbot using ChatGPT-4 Turbo, undertook a comprehensive risk management planning process to preemptively identify and mitigate potential challenges. The risk management plan, as laid out on March 6, 2024, identified ten principal risks that could potentially impede project progress.

**1. Technical Complexity and Challenges**  
This high-probability and high-impact risk pertained to the intricate technical demands of developing a sophisticated AI-powered chatbot. Despite rigorous planning, such challenges did manifest during the project. By increasing technical training sessions, consulting with experts, and incorporating additional time for addressing technical issues, the team managed to navigate through the complex technical landscape effectively.

**2. Inadequate Dataset Quality for AI Chatbot Training**  
The risk of inadequate data quality, with moderate probability and high impact, threatened to compromise the chatbot's training. This risk was mitigated by developing a plan for data enhancement, including data cleansing, augmentation, and diversification. Although no significant dataset issues arose, this preparedness was instrumental in maintaining data integrity throughout the project.

**3. Inadequate Prompt Engineering Expertise**  
This risk was deemed unlikely to occur and had a moderate impact. Nevertheless, contingency plans were in place to bolster prompt engineering efforts, which ensured that the chatbot's communication was refined and effective.

**4. Model Training and Fine-Tuning Delays**  
With the possibility of delays in model training and fine-tuning, agile project management techniques were employed to accommodate iterative development. This proactive approach ensured the timely completion of training and fine-tuning stages.

**5. Lower-than-Expected User Engagement**  
Anticipated with moderate probability and impact, the project did not face this risk, as user engagement was bolstered by early and continuous outreach efforts, ensuring that user interaction met expectations.

**6. Project Scope Creep**  
A significant risk that materialized was project scope creep, which resulted from an initial lack of clarity in project execution plans. This misalignment led to a review and substantial revision of project documents, including the project goal, objectives, WBA, project charter, and risk management report. Through prompt identification and sponsor communication, scope adjustments were effectively managed, mitigating potential repercussions.

**7. Extended Development Timelines**  
The project maintained a flexible schedule with buffer times to accommodate potential delays, ensuring that extended development timelines did not adversely affect the project's overall completion.

**8. Inadequate Chatbot Performance and Reliability**  
While the risk was recognized, continuous performance monitoring and regular audits prevented any significant performance or reliability issues from arising.

**9. Changes in Data Privacy Regulations**  
The potential impact of evolving data privacy regulations was negated by establishing a legal review process and adjusting data handling procedures as required, ensuring compliance throughout the project.

**10. Incorrect Implementation of Cohen's Kappa**  
To avoid incorrect implementation of Cohen's Kappa, which had a moderate probability and impact, the project incorporated periodic expert reviews of the validation process. This ensured that the validation methodology adhered to statistical accuracy standards.

In summary, the risk management strategies implemented throughout the project were largely successful in mitigating the anticipated risks. While some risks did manifest, notably the scope creep, timely and effective communication, along with the sponsor's guidance, allowed for rapid realignment of project parameters without incurring major damages. The other risks were either preemptively addressed through the outlined contingency plans or did not occur, underscoring the efficacy of the risk management planning process.

# Issues Encountered

During the implementation of the AI-powered chatbot project, several issues arose, testing the project team's preparedness and adaptability. The risk management strategy had anticipated possible complications; however, the dynamic nature of project development meant that even well-laid plans had to be responsive to real-time challenges. Here are the issues we faced, how they were handled, and how our risk mitigation strategies came into play.

**Project Scope Creep Due to Misaligned Understanding**

The primary issue encountered, as anticipated in the risk register, was project scope creep. However, contrary to typical scope expansions due to new stakeholder requirements, this creep arose from a fundamental misunderstanding of the project’s technical demands. Initially, the project's goals and objectives, though seemingly reasonable, did not fully align with the actual capabilities and the level of technological support available. This disconnect became apparent as the project unfolded, necessitating a significant recalibration of the project's direction.

Timely communication with the sponsor revealed a preference for a more accessible approach that aligned with the current technical skill set and knowledge level. The sponsor's guidance was instrumental in redefining the project scope to be less technologically demanding. Consequently, extensive modifications were made to several project documents, including the project goals, objectives, WBA, project charter, and risk management report, to realign with the newly clarified vision. Despite the time investment required to make these changes, the issue did not inflict significant damage thanks to prompt identification and communication of the discrepancies.

**Mitigation of User Engagement Shortfalls**

Another issue that surfaced was lower-than-expected user engagement, which could have impeded the chatbot's training and validation phase. While this risk was identified, the actual engagement levels required proactive measures beyond initial expectations. The team responded by intensifying user outreach efforts and highlighting the chatbot’s benefits, which successfully improved engagement and ensured adequate data for the project’s needs.

**Technical Challenges and Expertise Gaps**

The project also grappled with the risk of technical complexity and potential expertise gaps in prompt engineering. This challenge was met head-on by leveraging the risk mitigation plan, which included supplemental training and iterative development cycles. This approach enhanced the technical proficiency of the team and maintained the project’s momentum.

**Conclusion**

Reflecting on the issues encountered, it is evident that the project management team was well-prepared to tackle the challenges head-on. The risk management strategies in place were not only suitable for the anticipated risks but also flexible enough to adapt to the actual problems faced. The issues that arose did not stem from oversights or planning deficiencies; instead, they were inherent to the nature of a complex and innovative project that pushed the boundaries of existing technology and knowledge. The team's agility in responding to these challenges was instrumental in steering the project to its successful conclusion.

The lessons learned from managing these issues underscore the value of maintaining a fluid and responsive project management approach, especially in projects that involve cutting-edge technologies and methodologies. The capacity to realign project objectives and adapt strategies in real time can make the difference between faltering and flourishing in the dynamic landscape of technology-driven initiatives.Top of FormBottom of Form

# Project Chronology and Critique

The project's journey, dedicated to creating an AI-powered chatbot using ChatGPT-4 Turbo, commenced with meticulous planning and a surge of enthusiasm. The aim was to advance recruitment efficiency by accurately matching resumes with job listings, a venture that promised to blend NYU's academic expertise with The Digital Forge's technological prowess. This chronology provides a retrospective of the project's progression and an introspective critique of areas for improvement.

**Chronology of Events**

The initiation phase began with defining the project scope, objectives, and deliverables, encapsulating the ambitions into a structured roadmap. Prompt engineering was the cornerstone of development, demanding significant resource allocation to craft and refine prompts that could enable ChatGPT-4 Turbo to analyze and match job-seeker profiles effectively.

Midway, the team faced the formidable challenge of scope misalignment, an issue that threatened to stall progress. Through adaptive management and a collaborative approach with stakeholders, we revised the project scope, aligning it with practical capabilities and technological thresholds.

As the project transitioned into the execution phase, technical complexity surfaced, notably in integrating and testing with ChatGPT-4 Turbo. A concerted effort in skill development and expert consultations helped overcome this hurdle. Subsequent user interaction and performance monitoring highlighted another concern—user engagement levels. By enhancing outreach and clearly communicating the chatbot's value, we improved user involvement, which was critical for feedback and iterative improvement.

In the project's penultimate stage, the validation process employing Cohen's Kappa underscored the chatbot's matching effectiveness, rivaling the existing R-based model. The final act of drafting a conference paper crystallized the project's findings and asserted the value of AI in recruitment.

**Critique and Reflections**

Reflecting on what might have been improved, the foremost aspect is the initial goal setting. An even deeper understanding of the technical requirements and a more cautious approach to setting the project scope could have averted the need for later adjustments. Although these adjustments were managed effectively, they diverted time and energy from other tasks.

Moreover, while user engagement strategies eventually succeeded, a proactive plan for engagement should have been a priority. Had this been anticipated, it could have prevented the scramble to boost user interaction later in the project.

Another area for improvement is risk management concerning technical challenges. Future projects would benefit from an upfront investment in technical training and skills development for the team, thereby reducing reliance on external consultants and streamlining the learning curve.

**Conclusion**

In hindsight, every challenge faced and every obstacle overcome has been an opportunity for growth. This project not only achieved its aims but also provided invaluable lessons on the importance of flexibility and adaptability in project management. It has reinforced the notion that effective communication and stakeholder involvement are vital components of successful project delivery.

In conclusion, while there were several aspects that could have been handled differently, the project's success lies in its ability to adapt, learn from encounters, and ultimately deliver a product that meets its intended goals. As we look to future projects, these lessons will serve as a foundation for greater efficiency and innovation.

# Lessons Learned

Throughout the project's lifecycle, the team not only met the predefined objectives but also gained substantial insights and knowledge. The process of developing an AI-powered chatbot using ChatGPT-4 Turbo has been both enlightening and enriching, providing practical experience that extends well beyond theoretical concepts learned in class.

**Technical Proficiency in AI and NLP**

One of the most significant areas of growth was in technical proficiency, particularly concerning AI and NLP. The project required the team to delve deeply into ChatGPT-4 Turbo's capabilities, understanding its intricacies to fine-tune the prompts and maximize the chatbot's performance. This foray into AI was more than just a learning curve; it was a leap into the future of technology-driven solutions. Developing the chatbot offered hands-on experience in the potential of AI for real-world applications, fostering a skill set that will undoubtedly be valuable in the tech-centric future.

**Project Management and Agile Adaptation**

Project management was another domain where substantial learning occurred. The need to recalibrate the project's scope due to unforeseen complexities was a practical lesson in agile project management. It emphasized the importance of flexibility and being prepared to pivot when certain aspects of the project do not go as planned. The scope creep issue, though initially a setback, turned into a lesson in managing expectations and project boundaries.

**Stakeholder Engagement and Communication**

Engagement with stakeholders, particularly in clarifying project goals and receiving feedback, highlighted the importance of clear communication. Learning to effectively interact with diverse stakeholders, from technical experts to project sponsors, was instrumental in realigning the project's trajectory. It also underscored the need to maintain continuous dialogue to ensure all parties share a unified vision.

**Risk Management and Contingency Planning**

Risk management was an area that witnessed considerable development. The project provided a testbed for applying theoretical risk management strategies to practical scenarios. Creating contingency plans for risks like inadequate dataset quality and potential project scope creep improved the team's ability to anticipate and mitigate project challenges. The experience highlighted the necessity of a proactive approach to risk identification and the formulation of response strategies.

**Reflection on Performance and Quality Assurance**

Quality assurance was another lesson learned during the project. The focus on delivering a high-quality chatbot necessitated regular reviews and adjustments based on user feedback, which fostered a culture of continuous improvement among the team. This practice not only improved the end product but also instilled a mindset of quality focus, which will benefit future projects.

**Conclusion**

In summary, the applied project experience was a confluence of learning and application, combining the realms of AI, project management, stakeholder engagement, and risk management into a cohesive learning journey. It was a testament to the team's dedication and the sponsor and professor's guidance that the project was delivered on time and met the expected quality standards. The skills and insights acquired during this project will serve as an invaluable asset in both professional pursuits and academic endeavors, solidifying a foundation for future success in the rapidly evolving landscape of technology and business.Top of FormBottom of Form

# Conclusion and Summary

The project to develop an AI-powered chatbot, leveraging the capabilities of ChatGPT-4 Turbo, aimed to bring a transformative approach to matching job seekers with suitable job opportunities. It was embarked upon with a vision to not just align with the innovative edge of recruitment technology but also to offer a model for academic and professional excellence.

**Realization of Project Goals**

Through dedicated efforts, the project successfully achieved its goal of creating a chatbot that functions with as much efficacy and reliability as the existing R-based model. The achievement is a testament to the project's well-thought-out objectives, the rigorous development process, and the active risk management strategies that were put into place.

**Technical Accomplishments**

The project stands as a showcase of the potential and agility of AI in the field of human resources. Talent Sync, the GPT-powered chatbot, demonstrated the effectiveness of using ChatGPT-4 Turbo's advanced NLP capabilities for analyzing and matching resumes with job listings, surpassing traditional methods in speed, accuracy, and bias reduction. The journey from conceptualization to the final product was marked by iterative testing, refinement, and validation that not only fulfilled the technical objectives but also provided an enriched learning experience for all involved.

**Collaborative Efforts and Stakeholder Engagement**

One of the project’s greatest strengths was the collaboration and guidance from the stakeholders, particularly the sponsor Dr. Andres Fortino, who was instrumental in steering the project through the phases of scope clarification and technical guidance. The regular interaction and feedback from various stakeholders ensured that the project remained aligned with its core objectives, adapting to new insights and needs as they emerged.

**Documentation and Knowledge Sharing**

In line with the project's commitment to transparency and knowledge dissemination, all datasets, the tool code, model code, and analysis spreadsheets, as well as the full report, have been meticulously documented and made available for public access. The GitHub repository serves as a comprehensive resource for understanding the project's development, providing valuable insights into both the process and the final deliverables.

**Contribution to the Field**

By successfully navigating through the challenges and risks, the project not only delivered a working model but also contributed to the body of knowledge surrounding AI applications in recruitment. It has opened avenues for further research and has set a benchmark for future projects in the domain.

**Reflection and Looking Ahead**

The success of the project was not devoid of challenges, yet the lessons learned from these hurdles have enriched the team's professional acumen. As the project concludes, the implications of the work done extend beyond the immediate deliverables. It sets a precedent for future projects that seek to merge AI with practical applications, signaling a future where AI is an integral part of problem-solving in various industries.

**Access to Talent Sync and Project Materials**

Talent Sync, the AI-powered chatbot, can be accessed at: <https://chat.openai.com/g/g-QwoCKlfEd-talent-sync>. All data sets created for this project and this report also can be accessed through the following GitHub repository link: <https://github.com/Cassie-G/4100-Applied-Project.git>. This repository is structured to facilitate easy navigation and comprehension of the project's components and outcomes.

**Summary**

In summary, this project represents a successful integration of AI technology into a practical application that is both scalable and adaptable. As the project draws to a close, the experiences gained, and the milestones achieved mark not just the completion of a project but the beginning of new possibilities in the field of AI and recruitment. It stands as an embodiment of what can be achieved when innovative technology is applied with a clear vision and a collaborative spirit.

# Limitations, Recommendations and Scope for Future Work

Even though this project was able to deliver as expected, there are still some limitations within this project and some of the limitations may be improved in future similar projects at NYU MASY. The exploration into the realm of AI and its application in recruitment through ChatGPT-4 Turbo has been a promising venture. However, in the spirit of continuous improvement and academic inquiry, it is important to recognize the limitations encountered, recommend areas of enhancement, and chart the path for future work.

**Limitations**

Firstly, the scope of data utilized for training the AI chatbot was constrained by the availability and accessibility of diverse and industry-specific resumes and job descriptions. While the model showed proficiency in general applications, its depth could have been expanded with a broader dataset encompassing more specialized fields.

Furthermore, the project experienced limitations in real-time user testing due to the time constraints inherent in the academic semester. Extended interaction with a live user base would provide richer feedback and more nuanced insights into the chatbot's performance in various real-world scenarios.

**Recommendations**

For future iterations of this project or similar endeavors, it is recommended to:

Addressing these limitations begins with expanding the chatbot's dataset to include a wider array of job sectors, experience levels, and cultural nuances. Partnering with corporations and online job platforms could facilitate access to such data.

Moreover, extending the feedback collection period beyond the project's timeline would allow for more substantial user interaction data. This could lead to more impactful iterations and a more robust chatbot.

Investment in user experience design is also recommended. A more intuitive and engaging interface could increase user engagement and the quality of feedback, thus enhancing the chatbot’s learning.

**Scope for Future Work**

The project's scope for future work is vast and multidirectional. One area is the integration of advanced natural language understanding (NLU) to discern the nuanced intent and context in user communications. This would enable the chatbot to offer more personalized and accurate job matches.

Further development could also involve expanding the chatbot’s functionality to cover more recruitment processes, like interview coordination, initial screening, and candidate follow-ups.

Exploring the applicability of the chatbot in various cultural contexts is also a fertile ground for future work. A multilingual chatbot could serve a global audience, ensuring inclusivity and expanding the potential user base.

Lastly, leveraging AI to provide insights into labor market trends based on the aggregation and analysis of job seeker and employer data could offer valuable strategic input for workforce development and policymaking.

**Conclusion**

The completion of this project marks a significant step forward in the application of AI within the field of recruitment. It also provides a robust framework for future projects at NYU MASY to build upon. Through continuous learning and adaptation, AI can potentially revolutionize not only how candidates are matched to job opportunities but also the broader landscape of talent acquisition and management. With careful consideration of the limitations and recommendations outlined, future projects have the potential to not only refine the solutions provided but also to innovate and expand the impact of AI in recruitment.

# Literature Survey

**Introduction**

The integration of Artificial Intelligence (AI) in recruitment processes marks a pivotal transition in how organizations approach talent acquisition. This literature review explores the multifaceted role of AI in recruitment, particularly focusing on the development of AI-powered chatbots designed to enhance the matching of candidates to job opportunities. The criteria for analysis are grounded in technological efficacy, user interaction design, ethical considerations such as bias and fairness, and the psychometric capabilities of AI systems. The organization of the review follows a thematic sequence, starting from a broad overview of AI in recruitment, delving into the nuances of human-AI interaction, addressing the critical issues of bias and fairness in algorithm-driven hiring, and concluding with the implications of these systems on recruitment practices.

The scope of this review is confined to academic journals that provide empirical evidence and theoretical insights into the application of AI in recruitment. Literature that does not directly contribute to the understanding of AI's role in recruitment, such as articles focused solely on unrelated AI applications or outdated hiring practices, is excluded to maintain the relevance and clarity of the review. This literature review serves as a foundation for the development of a superior AI-powered recruitment chatbot, aiming to synthesize existing knowledge and identify opportunities for innovation and improvement.

**Motivation for AI in Recruitment**

AI is increasingly adopted in recruitment to address the limitations of human-centric processes, which are often constrained by cognitive biases and logistical inefficiencies. The need for a scalable, objective, and efficient system is critical in a global job market where diversity and inclusivity are not just ethical imperatives but also competitive advantages (Vivek, 2023; Tilmes, 2024).

**Chatbots in Recruitment: A Technological Advancement**

The advent of chatbots in recruitment represents a technological leap in how companies interact with job candidates, manage applications, and enhance the hiring process. Chatbots, powered by AI, automate, and personalize the recruitment workflow, thereby facilitating a more efficient and engaging candidate experience.

**Operational Efficiency**

Chatbots in recruitment streamline various stages of the hiring process. They can efficiently manage repetitive tasks such as scheduling interviews, answering FAQs about job roles, and gathering initial candidate information. This automation significantly reduces the administrative burden on human recruiters, allowing them to focus on more strategic aspects of talent acquisition (Koivunen et al., 2022).

**Enhanced Candidate Experience**

By providing immediate responses and interactive communication, chatbots can significantly enhance the candidate experience. They maintain engagement and provide personalized interactions at scale, which is particularly beneficial in managing large volumes of applicants. Studies by Kurek et al. (2024) have shown that candidates appreciate the 24/7 availability and instantaneous feedback provided by AI-driven chatbots, leading to a more positive perception of the hiring organization.

**Bias Mitigation**

One of the critical roles of AI-driven chatbots in recruitment is their potential to reduce unconscious biases. By standardizing the initial stages of the hiring process, chatbots can ensure a more equitable screening process. Research by Intezari et al. (2022) discusses how chatbots programmed with neutral responses and criteria-based evaluations can help in mitigating biases that often occur during resume screening or initial assessments.

**Challenges and Limitations**

Despite their advantages, the deployment of chatbots in recruitment also presents several challenges. Concerns regarding the impersonal nature of chatbots and their inability to fully understand complex human emotions and nuances are notable (Lavanchy et al., 2023). Furthermore, the ethical implications of data privacy, consent, and security come into play when chatbots handle the sensitive personal information of candidates (Tilmes, 2022).

**Future Directions**

The future of chatbots in recruitment looks towards integrating more advanced NLP capabilities to better understand and interpret the subtleties of human communication. The ongoing research by Andrade-Arenas et al. (2024) aims to equip chatbots with better contextual understanding and emotional intelligence, enhancing their ability to conduct preliminary interviews and make more nuanced candidate evaluations.

Integrating chatbots into recruitment processes is a promising development aimed at enhancing operational efficiency, improving candidate experience, and promoting fairness in hiring practices. However, it is imperative to address the challenges related to their impersonal nature and ethical concerns surrounding data use. As technology advances, the focus should remain on improving the sophistication of these systems to ensure they complement human recruiters effectively and ethically.

**AI Recruitment Tool Utility**

The deployment of AI in recruitment is motivated by evidence of its utility in enhancing the job-matching process. Studies indicate that AI tools can significantly improve the accuracy and fairness of candidate selection, leading to better hiring outcomes (Fan et al., 2023; Li et al., 2023). AI recruitment tools have shown promise in improving efficiency and candidate experience, with users showing a preference for the streamlined process that AI tools offer (Koivunen et al., 2022). The evidence points to AI's ability to infer personality traits which can be valuable in recruitment (Fan et al., 2023).

**Algorithmic Choices in AI Recruitment**

The decision to utilize specific algorithms is predicated on their demonstrated effectiveness. For instance, LLMs like ChatGPT-4 Turbo have shown considerable promise in accurately interpreting complex inputs, making them suitable for the nuanced task of matching resumes to job descriptions (Wang et al., 2024).

**Design Choices and Evidence**

Abedin et al. (2022) discussed the imperative to balance AI's technical capabilities with user expectations, advocating for designs that address privacy, security, and transparency. This aligns with the user-centric approach where AI tools are tailored for ease of use and trust-building. Wang et al. (2024) further emphasized the role of prompt engineering in enhancing the consistency and reliability of AI responses, which is critical in the context of recruitment where accurate interpretation of user inputs can significantly influence the outcomes. This body of work underscores the importance of iterative design processes that incorporate user feedback and rigorous testing to refine AI tools, ensuring they meet the nuanced needs of diverse user bases and adhere to ethical standards in AI deployment.

**Development Language and Technical Considerations**

The choice of development language and technical frameworks is informed by the project's needs for robust performance and the ability to integrate with existing systems. The literature supports the selection of tools that offer both the necessary computational power and the flexibility to adapt to evolving project requirements (Bandi et al., 2023).

**Evidence from Research and Experiments**

Tianyi Li et al. (2023) presented a rigorous method for evaluating human-AI interaction, underscoring the necessity of understanding user perceptions to guide AI development. This is complemented by research from Fan et al. (2023), which assessed the psychometric properties of AI-inferred personality traits, suggesting the potential for AI tools to extend beyond traditional recruitment metrics. Together, these studies advocate for a design methodology that is both evidence-based and responsive to user needs, emphasizing the importance of empirical research in validating design choices and improving AI tools' efficacy in real-world recruitment scenarios.

**Data Sources for AI Recruitment Tools**

Data serves as the foundation for AI recruitment tools, and its quality is paramount for the system's effectiveness. High-quality, unbiased datasets are not only crucial for the initial training but also for the ongoing learning and improvement of AI systems (Intezari et al., 2022). The integrity of the data sources directly impacts the AI's ability to make fair and unbiased recommendations. Studies like those by Koivunen et al. (2022) stressed the importance of diversity in data to avoid perpetuating systemic biases in recruitment processes. Furthermore, the integration of data from various domains can enhance the robustness of AI tools, as showcased by Wang et al. (2024), who illustrated how diverse prompts can affect AI performance and reliability. The careful selection and curation of data sources thus play a critical role in ensuring that AI recruitment tools operate ethically and effectively.

**Weaknesses of Chatbots in Recruitment**

While chatbots offer numerous advantages in recruitment, such as scalability and efficiency, they are not without their weaknesses, especially in handling atypical candidate responses. The research by Řepová et al. (2024) provided valuable insights into the nature and triggers of these atypical responses in chatbot-mediated job interviews.

**Nature of Atypical Responses**

The study categorized atypical responses into six types: testing the chatbot's capabilities, verbal abuse, testing the chatbot's reactions, further conversation, sex offers, and reverse discrimination. These responses highlighted the challenges chatbots face in interpreting and appropriately reacting to non-standard human behaviors.

**Triggers of Atypical Responses**

Řepová et al. (2024) identified several triggers for these atypical responses, including lower company attractiveness and additional stress. This is crucial because it pointed to environmental and procedural factors that exacerbate the limitations of chatbots in managing complex human interactions.

**Implications for Recruitment**

The presence of atypical responses can compromise the effectiveness of the recruitment process. For instance, candidates testing the limits of the chatbot or engaging in verbal abuse may disrupt the intended flow of the interview, leading to inefficient or incomplete assessments of candidates' suitability for a position.

**Design and Ethical Considerations**

The study underscored the importance of thoughtful chatbot design to mitigate the occurrence of atypical responses. Ensuring chatbots are equipped to handle or redirect unanticipated candidate behaviors can improve the robustness and reliability of AI-driven recruitment processes.

**Future Directions**

The findings from this study suggested that ongoing research and development are needed to enhance the understanding and capabilities of chatbots in recruitment settings. Particularly, improving AI's ability to recognize and adapt to the diverse and dynamic nature of human communication can help in minimizing the impact of atypical responses.

**Conclusions**

In conclusion, the reviewed literature solidifies the understanding that AI, particularly in recruitment chatbots, presents a significant opportunity to reshape talent acquisition. The synthesis of the research emphasizes the imperative to merge technological advancement with human-centric design (Abedin et al., 2022), highlighting the promise of AI in enhancing efficiency and fairness in recruitment (Fan et al., 2023; Vivek, 2023). Concerns around biases in AI are met with rigorous methodological approaches to ensure diversity and fairness in data handling (Intezari et al., 2022). The methodologies employed in these studies, from factorial surveys to psychometric analysis, provide a framework for the development and continuous refinement of AI tools (Tianyi Li et al., 2023).

The insights garnered from this body of work directly inform the creation of the AI-powered chatbot for this project. Adopting a user-centric design approach, informed by empirical evidence, ensures that the chatbot meets the nuanced needs of users while upholding ethical standards. The findings also illuminate areas that are less understood, such as the full spectrum of human-AI interaction nuances in recruitment, offering a pathway for future exploration.

This literature review reinforces the decision to pursue a project that not only seeks to improve the recruitment process but also contributes to the broader understanding of AI's role in society. The alignment of the reviewed literature with the project objectives confirms the chosen direction and solutions implemented are grounded in current and robust research.

# Appendix A - Project Acceptance Document



# Appendix B - Project Sponsor Agreement

# Appendix C - Project Charter

**AI Revolution in Recruitment and Job Search: Engineering a Superior Matchmaking Chatbot Using ChatGPT-4 Turbo**

**Project Charter**

Project Manager: Ziwen (Cassie) Gao

Sponsor: Dr. Andres Fortino, Clinical Associate Professor, NYU (<https://www.linkedin.com/in/afortino>), The Client

Prepared by: Ziwen (Cassie) Gao

Name and Location of Client Organization: The Digital Forge

NYU School of Professional Studies and the Management and Systems program (MASY) is a New York-based learning institution.

NYU School of Professional Studies is at 12 West 43rd Street, NY, NY.

1. **Project Goal**

The goal of the project was to create and validate an AI-powered chatbot using ChatGPT 4 Turbo that as effective and reliable as the R-based model in matching resumes to job listings, aiming to revolutionize the recruitment process through enhanced accuracy and efficiency that impact both employers and job seekers.

1. **Problem/Opportunity Definition**

With The Digital Forge at the helm, partnering with NYU's School of Professional Studies and its Management and Systems program, this project stands to greatly benefit from the expertise and innovative environment fostered by such a distinguished institution. Under the guidance of Dr. Andres Fortino, the project will leverage the university's cutting-edge research and academic prowess to deliver an AI-powered recruitment tool that not only enhances the efficiency and accuracy of the hiring process but also serves as a benchmark for educational and professional excellence. This collaboration will demonstrate NYU's commitment to practical, technology-driven solutions in professional education and will likely attract further industry partnerships, reinforcing its status as a leader in applying theoretical knowledge to solve real-world problems.

1. **Proposed Project Description**

**Creating an LLM Chatbot to Match Resumes to Job Opportunities**

The client has developed and published a model in R that ranks job opportunities given an applicant’s resume. It uses the advanced text analytic technique of text similarity scoring. The model has been updated and expanded using AI and prompt engineering of LLMs to demonstrate the superiority of using LLMs for the same analysis, which provides faster and more accurate results than manual matching or the R-based model. With the advent of powerful LLMS with code generation, such as Chat GPT 4 Turbo, we have now done this work more efficiently and with more impact and we are able to create chatbots called GTPs. We’re interested in re-creating this work using fine-tuning an LLM and creating a Chatbot with ChatGPT Turbo. This is an advanced prompt engineering project. The deliverables include prompts for creating the models, comparing the three approaches, and project documentation as a draft submission to the IEEE LISAT conference at NYIT Long Island in April 2024.

1. **Project Sponsor**

The principal project sponsor will be Dr. Andres Fortino, Clinical Associate Professor, NYU (<https://www.linkedin.com/in/afortino>), The Client

1. **Objectives**

**Project Objective 1 – Prompt Engineering for Creating the GPT-Powered Chatbot**

* **Specific:** Design and implement a set of advanced prompts that enable the ChatGPT 4 Turbo to analyze resumes and match them to job listings with high relevance and accuracy.
* **Measurable:** The engineering efforts will focus on creating a robust model that interprets and processes resume data against job descriptions, leveraging the advanced natural language processing (NLP) capabilities of the LLM.
* **Achievable:** Allocate resources to experiment with and refine multiple prompt iterations, relying on expert guidance and existing models as a foundation.
* **Realistic:** Engage with language model experts to validate the prompts’ effectiveness in capturing relevant job and resume data.
* **Time-Related:** Complete the prompt engineering phase by March 15, 2024, with iterative testing and refinement scheduled to ensure prompt readiness for deployment.

**Metrics:**

* **Development of 20 unique prompts** that focus on specific features and abilities such as Natural Language Processing (NLP) Capabilities, Skill Matching and Ranking, and Contextual Analysis that could enhance effectiveness of the chatbot by March 5, 2024.
* **Testing phase with a dataset of at least 5 resumes and job descriptions,** aiming for an 85% accuracy rate in matching resumes to relevant job listings by March 10, 2024.
* **Iterative improvement based on feedback** by conducting at least 2 review sessions with language model experts to validate the effectiveness and accuracy of the prompts before final implementation.

**Project Objective 2 – Optimizing Chatbot Performance constantly through User Satisfaction Surveys and the Implementation of the Feedback Loop**

* **Specific:** Enhance the chatbot’s performance by continuously collecting and analyzing user feedback through structured satisfaction surveys, and implement changes based on this feedback.
* **Measurable:** Conduct surveys with at least 10 users, aiming to achieve and maintain an 85% user satisfaction rate.
* **Achievable:** Incorporate an automated feedback loop mechanism within the chatbot interface so that both job seekers and employers can provide feedback on the accuracy and relevance of the matches for the development team to review feedback and implement improvements.
* **Realistic:** Provide regular updates and training to the development team on user experience improvements and technical adjustments.
* **Time-Related:** Implement the first round of feedback by March 20, 2024, with subsequent monthly optimizations until the project's completion.

**Metrics:**

* **Initial user satisfaction survey conducted by March 20, 2024**, with at least 10 users participating in the pilot phase.
* **Achievement of an 85% user satisfaction rate** based on survey results, with adjustments made monthly based on ongoing user feedback.
* **Reduction of user-reported issues by 50%** over the first two weeks after the implementation of feedback loops, demonstrating continuous improvement in chatbot performance.

**Project Objective 3 – Validating the Chatbot as Effective and Reliable as the R-Based Model via Cohen's Kappa Statistic**

* **Specific:** Conduct a rigorous validation of the chatbot's effectiveness in matching resumes to job listings compared to an existing R-based model, using Cohen’s Kappa statistic to measure agreement between the two systems.
* **Measurable:** Achieve a Cohen’s Kappa value of 0.75 or above, indicating substantial agreement and validating the chatbot's reliability.
* **Achievable:** Utilize a diverse dataset of at least 5 paired resumes and job listings for the validation process.
* **Realistic:** Consult with statistical experts to ensure proper implementation and interpretation of Cohen’s Kappa statistic.
* **Time-Related:** Complete the validation process by April 5, 2024, allowing for any necessary adjustments based on the findings.

**Metrics:**

* **Conduct a validation study involving at least 5 paired resume and job listing samples** to ensure diverse testing scenarios by March 25, 2024.
* **Achieve a Cohen’s Kappa statistic of 0.75 or higher**, confirming substantial agreement between the chatbot and the R-based model by April 5, 2024.
* **Prepare a detailed validation report** outlining the methodology, results, and statistical significance of the findings to be reviewed by a statistical expert by April 10, 2024.

**Project Objective 4: Conference Paper Drafting**

* **Specific:** Draft and refine a conference paper detailing the project's methodology, findings, and the significance of the advancements in AI-powered recruitment technology.
* **Measurable:** Submit a draft suitable for peer review by the IEEE LISAT conference standards, receive feedback, and revise accordingly.
* **Achievable:** Collaborate with instructors and peers for peer reviews and feedback on the draft.
* **Realistic:** Ensure the paper includes comprehensive data analysis, clear results, and a discussion on the impact and future implications of the research.
* **Time-Related:** Complete the initial draft by April 25, 2024, and submit the final version by May 2, 2024, in time for conference submission deadlines.

**Metrics:**

* **Completion of the first draft of the conference paper by April 25, 2024**, including all sections such as abstract, methodology, results, and discussion.
* **Feedback obtained from at least two peers or instructors** by April 30, 2024, to ensure comprehensive review and incorporation of diverse perspectives.
* **Submission of the final draft to the IEEE LISAT conference by May 2, 2024**, adhering to all submission guidelines and quality standards.

1. **Project Selection & Ranking Criteria**

Project benefit category:

|  |  |
| --- | --- |
|  | Compliance/Regulatory |
| *X* | Efficiency/Cost reduction |
|  | Revenue increase |

Portfolio fit and interdependencies

The AI Revolution in Recruitment and Job Search project fits strategically within The Digital Forge's broader agenda of integrating cutting-edge technology into practical applications. This project aligns with ongoing initiatives aimed at enhancing technological capabilities within various sectors, particularly in education and professional development. By leveraging NYU's research and academic resources, this project not only enhances the School of Professional Studies' offerings but also sets a precedent for future technological integrations. It complements existing projects focused on educational technology, creating synergies that enhance the overall impact and effectiveness of the institution's tech-driven initiatives.

Project urgency

The urgency of this project is high due to the rapidly evolving nature of both the AI technology and the recruitment industry. Implementing an AI-powered tool like Talent Sync is critical at this juncture to maintain a competitive edge and to capitalize on current trends in automation and data-driven decision making. Immediate development and deployment of this chatbot are essential to meet the current and pressing needs of the job market, characterized by high volumes of data and the demand for quick, accurate matchmaking between job seekers and employers. Delaying this project could result in missed opportunities to lead in a tech-driven educational approach and diminish the potential impact on the recruitment landscape.Top of Form

Bottom of Form

1. **Cost/Benefit Analysis**

Tangible Benefits

Benefit: Implementation of the Talent Sync AI-powered chatbot will significantly reduce the time spent on resume screening and job matching, leading to a more efficient recruitment process.

Value & Probability: High likelihood (85% probability) of reducing recruitment cycle time by 40%, based on pilot testing and initial feedback.

Assumptions Driving Value: Assumes integration with current HR systems is smooth and the AI model performs as expected.

Intangible Benefits

Benefit: Improved candidate experience and employer branding through the use of an innovative AI tool for job matching, enhancing the overall perception of the company as a leader in technology-driven recruitment.

Value & Probability: Moderate to high likelihood (75% probability) of increased brand perception and candidate satisfaction.

Assumptions Driving Value: The projection is based on current trends showing a preference for technology-efficient processes in job search and recruitment, as well as positive user feedback during the trial phase.

Cost Categories Amount

Internal Labor hours: Approximate allocation of 250 hours of focused project work, spanning from January 25, 2024, to May 2, 2024, excluding time for meetings and administrative tasks.

External costs: Potential costs associated with cloud services for hosting the chatbot.

Labor (consultants, contract labor): If specialized expertise is needed for integration or fine-tuning.

Equipment, hardware or software: Negligible hardware expenses due to reliance on cloud-based solutions, with software expenses encompassed within external costs.

List other costs such as travel & training: Minimal; primarily related to the presentation of findings at conferences.

Financial Return

The financial return is projected to be realized through the enhanced efficiency of the recruitment process, potentially reducing the cost per hire, and improving the speed of hiring, which can contribute to organizational agility and competitiveness.

1. **Assumptions**
2. **Skill Level of Developers**: It is assumed that the programmers and developers working on this project have a proficient understanding of prompt engineering for LLMs, natural language processing, and are capable of implementing sophisticated AI algorithms within the ChatGPT-4 Turbo framework.
3. **Collaboration of Partners**: The project presumes active cooperation from associated partners, including organizations that will provide access to necessary data sets, and industry experts who may offer insights for the refinement of Talent Sync.
4. **Availability of the Sponsor and Client**: It is assumed that Dr. Andres Fortino and other stakeholders from NYU SPS will be available for regular consultations, providing continuous guidance and feedback throughout the project's duration, ensuring alignment with the project's goals and objectives.
5. **Access to Data Sets**: The project is contingent on the availability of diverse and comprehensive data sets of resumes and job listings, which are critical for training the AI model and for validating the effectiveness of Talent Sync through comparative analysis.
6. **Technological Resources**: There is an underlying assumption that all necessary technological resources, including software licenses and cloud computing services, will be available and operational without significant downtime or disruptions.
7. **User Engagement**: The project assumes a high level of engagement from the users participating in the feedback loop, providing candid and constructive responses that will be instrumental in optimizing Talent Sync's performance.
8. **Statistical Expertise**: It is presumed that the project team will have access to statistical expertise, particularly for the implementation and interpretation of the Cohen's Kappa statistic, to ensure the validation process is rigorous and accurate.
9. **Scope**

Quality

The quality of the final deliverable for Talent Sync will be determined through a series of criteria agreed upon by both the project team and the client:

* **Accuracy of Matchmaking**: Ensuring that the chatbot's resume-to-job listing matches meet or exceed an 85% accuracy threshold.
* **User Satisfaction**: Achieving a user satisfaction rate of 85%, as measured by post-interaction surveys.
* **Performance Metrics**: Chatbot's response time should be under 1 minute, and system availability should be at least 99% during peak hours.

Time

The project is constrained to the academic semester timeline, beginning on January 25, 2024, and concluding on May 2, 2024. Key milestones include:

* **Prompt Engineering Completion**: March 15, 2024
* **Initial User Feedback Integration**: March 20, 2024
* **Cohen's Kappa Validation**: April 5, 2024
* **Draft Completion for IEEE LISAT Conference**: April 25, 2024

Resource Allocation

* **Equipment**: The project will utilize computational resources provided by ChatGPT-4 Turbo, including servers for chatbot developing, AI model training and deployment.
* **Personnel**: The project team will consist of the project manager, Ziwen (Cassie) Gao, supported by a select group of NYU SPS students and staff for specific tasks.
* **Expert Consultation**: Access to language model experts and statistical analysts as needed.

Out of scope activities

* **Database Creation**: Building a proprietary database for job listings and resumes is out of scope.
* **Multi-Lingual Support**: At this stage, the chatbot will only support English, with other languages considered for future projects.
* **Complete Automation of Recruitment**: End-to-end recruitment process automation, including candidate outreach and interview scheduling, will not be tackled.

Constraints

1. **Programming Expertise**: The project will be executed within the current level of the programming expertise available, primarily in Python and associated data science libraries.
2. **Technology Use**: The project will primarily use ChatGPT-4 Turbo for the AI component. R will be used for comparative analysis, with no plans to integrate other AI or machine learning frameworks at this time.
3. **Analysis Tools**: Python libraries such as Pandas, Matplotlib, Seaborn, and Scikit-learn will be utilized for data manipulation and visualization.
4. **Analytical Techniques**: The Cohen's Kappa statistic will be the primary method for validation.
5. **Client Availability**: Dr. Andres Fortino will be available for weekly meetings, with additional sessions scheduled as needed for critical updates or milestone reviews.

1. **Risks and Mitigation Strategies**
2. **Inadequate Training Data**
   * **Impact**: Insufficient or low-quality training data could lead to poor chatbot performance.
   * **Mitigation**: Collaborate with industry partners and academic departments to secure a diverse and substantial dataset. Additionally, utilize synthetic data generation techniques where appropriate to augment the training set.
3. **Technical Issues with AI Model Integration**
   * **Impact**: Technical difficulties in integrating the ChatGPT-4 Turbo could delay project milestones.
   * **Mitigation**: Conduct thorough pre-integration testing and allocate additional resources for troubleshooting. Ensure access to technical support from the AI platform provider.
4. **Changes in Technology or AI Platforms**
   * **Impact**: Rapid advancements in AI technology could make the project's current approach obsolete.
   * Mitigation: Maintain flexibility in project planning to adapt to new technologies. Establish a protocol for regularly reviewing and incorporating technological advancements.
5. **User Resistance to AI Solutions**
   * **Impact**: Potential resistance from users accustomed to traditional recruitment methods could affect the adoption rate.
   * **Mitigation**: Develop a comprehensive change management plan that includes user education and demonstration of the chatbot's benefits to foster acceptance and ease the transition.
6. **Data Security and Privacy Concerns**
   * **Impact**: Handling sensitive personal data raises concerns about privacy and data security.
   * **Mitigation**: Implement strict data security protocols and ensure compliance with GDPR and other relevant data protection regulations. Include clear communication to users about data handling practices.
7. **Overreliance on Automated Screening**
   * **Impact**: Exclusive reliance on the chatbot for screening may overlook qualified candidates.
   * **Mitigation**: Combine the chatbot's screening with human oversight to ensure a balanced approach to candidate selection.
8. **Insufficient Expertise**
   * **Impact**: The project might suffer from a lack of access to necessary expert knowledge in AI and recruitment.
   * **Mitigation**: Set up partnerships with academic experts and industry professionals early in the project. Plan for potential consulting or training to fill any expertise gaps.
9. **Sponsor or Key Stakeholder Withdrawal**
   * **Impact**: Loss of a sponsor or key stakeholder could lead to project discontinuation.
   * **Mitigation**: Engage multiple stakeholders and cultivate a broad base of support within NYU and The Digital Forge to mitigate reliance on any single sponsor.
10. **Communications Plan**
11. **Frequency**

* **Regular Meetings**: Agreed upon to occur on a weekly basis, ensuring consistent and up-to-date exchange of information.
* **Ad-hoc Meetings**: Scheduled as needed, particularly at critical junctures of the project or when immediate decisions or feedback are required.

1. **Method**

* **Virtual Meetings**: Utilizing video conferencing tools like Zoom to facilitate face-to-face communication regardless of participants' locations.
* **Email Communications**: For less urgent matters, providing written records of discussions and decisions.

1. **Content**

* **Progress Updates**: Review of completed work against the project plan and discussion of any deviations from the schedule or scope.
* **Issue Resolution**: Identification of any challenges encountered since the last meeting, brainstorming potential solutions, and assigning action items for resolution.
* **Risk Review**: Regular assessment of potential risks, their impact on the project, and the status of mitigation efforts.
* **Next Steps**: Planning and agreement on the upcoming tasks and objectives, including resource allocation and deadlines.
* **Feedback and Questions**: Opportunity for the client to provide feedback on the project's direction and for the team to ask clarifying questions to ensure alignment with the client's expectations.

1. **Schedule Overview**

**Project Start Date:**

The project officially commenced on January 25, 2024.

**Estimated Project Completion Date:**

The project is scheduled for completion on May 2, 2024.

**Major Milestones:**

1. **Prompt Engineering Completion**: Due by March 15, 2024 - Finalize the set of advanced prompts for ChatGPT 4 Turbo.
2. **Initial User Feedback Integration**: Due by March 20, 2024 - Implement the first round of user feedback mechanisms and performance optimization.
3. **Cohen's Kappa Validation**: Due by April 5, 2024 - Complete the validation of the chatbot's effectiveness using the Cohen’s Kappa statistic.
4. **Draft Conference Paper Submission**: Due by April 25, 2024 - Submit the initial draft of the paper to be presented at the IEEE LISAT conference.
5. **Final Conference Paper Revision and Submission**: Due by May 2, 2024 - Submit the final version of the paper after incorporating feedback and revisions.

**External Milestones Affecting the Project:**

* **IEEE LISAT Conference Acceptance Notification**: The date when the conference committee announces accepted papers will be a critical external milestone as it will require the paper to be finalized and potentially affect the presentation preparation.
* **Peer Review Feedback**: The receipt of feedback from peer reviews will impact the timeline for the finalization of the conference paper draft.
* **Data Access and Availability**: Timely access to external databases for testing and validation is critical and must be coordinated in advance to align with project milestones.
* **Expert Consultations**: Scheduled meetings with language model experts and statistical analysts are dependent on their availability and will influence the prompt engineering and validation phases.

1. **Impact of Late Delivery**

* The alignment with the academic calendar could be disrupted, potentially affecting the grading and evaluation timeline for the Capstone course.
* A delay in project completion may result in missing the submission deadline for the IEEE LISAT conference, which is a significant milestone and dissemination opportunity for the project.
* Late delivery could delay the client’s ability to benefit from the efficiencies and improvements promised by the project, potentially affecting their competitive position in the HR technology market.
* There may be a cascading effect on other projects or initiatives that are dependent on the completion of this project for data, insights, or technological advancements.

1. **Resources Required**

Personnel:

|  |  |  |  |
| --- | --- | --- | --- |
| Role | Responsibilities | Duration of work | Qualifications needed |
| **Project Manager** | Oversees the project from initiation to closure, manages team coordination, ensures adherence to schedule, leads communication with stakeholders, and addresses risks and issues. | Throughout the entire project (January 25, 2024, to May 2, 2024). | Proficiency in project management, knowledge in AI and machine learning, experience in prompt engineering, strong communication skills. |
| **AI Developer/Analyst** | Responsible for developing and fine-tuning the AI model, conducting data analysis, and implementing the feedback loop for chatbot optimization. | From the start of prompt engineering to the completion of the feedback loop implementation and validation analysis. | Experience with LLMs like ChatGPT, in AI programming, prompt engineering, skills in data analysis, and proficiency in statistical methods. |

Facilities, Software, Hardware and other Resources

* **Computers**: Access to high-performance computing resources capable of handling AI model development and large-scale data analysis.
* **Cloud Services**: Subscription to cloud computing services for model training and deployment, along with necessary data storage and processing capabilities.
* **Software Programs**: Software tools for data preparing (e.g., Excel), project management and communication (e.g., Trello, Zoom), and AI machine learning (ChatGPT-4 Turbo).

Procedures/ Methodology

* **Agile Approach for Project Management**: The project will utilize an Agile project management approach to allow for iterative development and flexibility.
* **CRISP-DM for Data Analysis**: The project will follow the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology to structure the data analysis process from understanding the data to deploying the final AI model.
* **Prompt Engineering Process**: A systematic approach to develop, test, and refine prompts used for the AI chatbot, incorporating feedback from language model experts.
* **Cohen's Kappa for Validation**: A statistical approach to compare the effectiveness of the chatbot against the existing R-based model, ensuring the chatbot's reliability and accuracy in matching job seekers with job opportunities.

1. **Project Plan, Gantt**

A graph with colorful squares

Description automatically generated

1. **Project Evaluation**
2. **Project Schedule**

Progress will be meticulously tracked against the established Gantt chart milestones. Regular checks on task completion and adherence to upcoming deadlines will provide clear indicators of whether the project is on schedule. This Gantt chart will be updated in real-time to reflect any adjustments or shifts in the project timeline, ensuring that all team members are aware of current and future tasks.

1. **Project Weekly Status Report and Dashboard**

A weekly status report will be generated using project management tools such as Trello or a customized spreadsheet. This report will detail the latest progress and compare it against the planned objectives and timelines. The Gantt chart will be revised regularly to display the current status comprehensively, which will be shared with the client and all project stakeholders to maintain clarity and alignment.

1. **Project Communication Plan, Issues Log, Risk Register**

Effective communication will be maintained through structured weekly meetings, detailed email updates, and a dedicated online platform that logs all issues and risks. This platform will serve as a central hub for documenting any challenges or deviations from the plan and will include potential impacts and proposed mitigation strategies. This approach ensures transparency and provides all stakeholders with the ability to address issues promptly and effectively.

1. **Project Status Reports**

Comprehensive status reports will be compiled weekly, documenting all progress, challenges encountered, and risks identified during the project. These reports will include detailed descriptions of the situation, the implications for the project, and actions taken to mitigate risks. They will also highlight successes and areas for improvement. These reports will be crucial for keeping the client and project team informed and engaged, ensuring everyone is aligned with the project goals and current status.

# Appendix D - Project Plan



A graph with colorful squares

Description automatically generated

# Appendix E - Risk Management Plan



**MASTER OF SCIENCE IN MANAGEMENT AND SYSTEMS**

**Applied Project Capstone**

**MASY GC- 4100**

**MEMORANDUM**

TO: Dr Eleftheria K Pissadaki

FROM: Ziwen Gao

DATE: March 6, 2024

RE: **Assignment 6 – Risk Management Plan**

**Project**

**AI Revolution in Recruitment and Job Search: Engineering a Superior Matchmaking Chatbot Using ChatGPT-4 Turbo**

The goal of the project was to create and validate an AI-powered chatbot using ChatGPT 4 Turbo that as effective and reliable as the R-based model in matching resumes to job listings, aiming to revolutionize the recruitment process through enhanced accuracy and efficiency that impact both employers and job seekers.

**Risks**

|  |  |  |  |
| --- | --- | --- | --- |
| **Number** | **Risk** | **Probability Score (1,2 or 3)** | **Impact Score (1,2 or 3)** |
| 1 | Technical Complexity and Challenges | 3 | 3 |
| 2 | Inadequate Dataset Quality for AI Chatbot Training | 2 | 3 |
| 3 | Inadequate Prompt Engineering Expertise | 1 | 2 |
| 4 | Model Training and Fine-Tuning Delays | 2 | 3 |
| 5 | Lower-than-Expected User Engagement | 2 | 2 |
| 6 | Project Scope Creep | 2 | 3 |
| 7 | Extended Development Timelines | 2 | 3 |
| 8 | Inadequate Chatbot Performance and Reliability | 2 | 3 |
| 9 | Changes in Data Privacy Regulations | 2 | 3 |
| 10 | Incorrect Implementation of Cohen's Kappa | 2 | 3 |

**Risk Matrix**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RISK (exposure) | | | |
| Probability (of occurrence) |  | 1.Slight | 2. Moderate | 3. High |
| 1. Very Unlikely |  | 3 |  |
| 2. Possible |  | 5 | 2, 4, 6, 7, 8, 9, 10 |
| 3. Expected |  |  | 1 |

**Contingency Plan**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Risk | Description | Probability (1-3) | Exposure (1-3) | Contingency Plan |
| 1 | Technical Complexity and Challenges | 3 | 3 | Increase the technical training sessions, arrange for expert consultations, and allocate additional time in the project timeline for addressing unforeseen technical issues. |
| 2 | Inadequate Dataset Quality for AI Chatbot Training | 2 | 3 | Develop a plan for data enhancement, including data cleansing, augmentation, and diversification. Establish partnerships with data providers if needed or consider synthetic data generation. |
| 4 | Model Training and Fine-Tuning Delays | 2 | 3 | Implement agile project management techniques to allow for iterative development, ensure hardware resources are available for high-performance computing, and maintain a buffer period within the project schedule for such tasks. |
| 6 | Project Scope Creep | 2 | 3 | Strictly adhere to the project requirements and establish a formal process for reviewing and approving changes. Keep the project scope visible and remind stakeholders of the implications of scope changes. |
| 7 | Extended Development Timelines | 2 | 3 | Maintain a flexible project schedule with buffer times and regularly review timelines. Have contingency staffing plans to address resource bottlenecks. |
| 8 | Inadequate Chatbot Performance and Reliability | 2 | 3 | Set up continuous monitoring systems, regular performance audits, and have a rapid response team ready to address any performance issues. |
| 9 | Changes in Data Privacy Regulations | 2 | 3 | Establish a legal review process to continuously monitor changes in privacy laws and adjust data handling procedures accordingly. Have a legal advisory team on retainer. |
| 10 | Incorrect Implementation of Cohen's Kappa | 2 | 3 | Schedule periodic reviews of the validation process by statistical experts, include multiple checkpoints for the accuracy of implementation, and provide training on statistical methods to the team. |

# Appendix F - Status Report

A document with a checklist

Description automatically generated with medium confidenceA paper with text and a box

Description automatically generated with medium confidenceA project report with green circles

Description automatically generatedA list of project tasks

Description automatically generatedA close-up of a document

Description automatically generated

A document with text and a green and yellow dot

Description automatically generated with medium confidenceA document with text on it

Description automatically generatedA close-up of a document

Description automatically generatedA white sheet of paper with black text

Description automatically generated

# Appendix G - Annotated Bibliography

Abedin, B., Meske, C., Junglas, I., Rabhi, F., & Motahari-Nezhad, H. (2022). Designing and managing human-AI interactions.*Information Systems Frontiers, 24*(3), 691-697. <https://doi.org/10.1007/s10796-022-10313-1>

Andrade-Arenas, L., Yactayo-Arias, C., & Pucuhuayla-Revatta, F. (2024). Therapy and Emotional Support through a Chatbot. *International Journal of Online & Biomedical Engineering*, *20*(2), 114-130. <https://doiorg.proxy.library.nyu.edu/10.3991/ijoe.v20i02.45377>

Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet*, *15*(8), 260. <https://doi.org/10.3390/fi15080260>

Fan, J., Sun, T., Liu, J., Zhao, T., Zhang, B., Chen, Z., Glorioso, M., & Hack, E. (2023). How Well Can an AI Chatbot Infer Personality? Examining Psychometric Properties of Machine-Inferred Personality Scores. *Journal of Applied Psychology*, *108*(8), 1277-1299. <https://doi.org/10.1037/apl0001082>

Intezari, A., Pauleen, D., & Soleimani, M. (2022). Mitigating cognitive biases in developing AI-assisted recruitment systems: A knowledge-sharing approach.*International Journal of Knowledge Management, 18*(1), 1-18. <https://doi.org/10.4018/IJKM.290022>

Koivunen, S., Ala-Luopa, S., Olsson, T., & Haapakorpi, A. (2022). The march of chatbots into recruitment: Recruiters’ experiences, expectations, and design opportunities.*Computer Supported Cooperative Work, 31*(3), 487-516. <https://doi.org/10.1007/s10606-022-09429-4>

Kurek, J., Latkowski, T., Bukowski, M., Świderski, B., Łępicki, M., Baranik, G., Nowak, B., Zakowicz, R., & Dobrakowski, Ł. (2024). Zero-Shot Recommendation AI Models for Efficient Job–Candidate Matching in Recruitment Process. *Applied Sciences (2076-3417)*, *14*(6), 2601. <https://doi-org.proxy.library.nyu.edu/10.3390/app14062601>

Lavanchy, M., Reichert, P., Narayanan, J., & Savani, K. (2023). Applicants’ Fairness Perceptions of Algorithm-Driven Hiring Procedures. *Journal of Business Ethics*, *188*(1), 125-150. <https://doi-org.proxy.library.nyu.edu/10.1007/s10551-022-05320-w>

Řepová, H., Král, P., & Zouhar, J. (2024). Atypical responses of job candidates in chatbot job interviews and their possible triggers. *Computers in Human Behavior: Artificial Humans*, *2*(1). <https://doi-org.proxy.library.nyu.edu/10.1016/j.chbah.2023.100038>

Tilmes, N. (2022). Disability, fairness, and algorithmic bias in AI recruitment.*Ethics and Information Technology, 24*(2). <https://doi.org/10.1007/s10676-022-09633-2>

Tianyi Li, Vorvoreanu, M., Debellis, D., & Amershi, S. (2023). Assessing Human-AI Interaction Early through Factorial Surveys: A Study on the Guidelines for Human-AI Interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *30*(5), 91-135. <https://doi-org.proxy.library.nyu.edu/10.1145/3511605>

Vivek, R. (2023). Enhancing diversity and reducing bias in recruitment through AI: a review of strategies and challenges. Информатика. Экономика. Управление, 2(4). <https://doi-org.proxy.library.nyu.edu/10.47813/2782-5280-2023-2-4-0101-0118>

Wang, L., Chen, X., Deng, X., Wen, H., You, M., Liu, W., Li, Q., & Li, J. (2024). Prompt engineering in consistency and reliability with the evidence-based guideline for LLMs.*NPJ Digital Medicine, 7*(1), 41. <https://doi.org/10.1038/s41746-024-01029-4>