Title	**CareerAi: An AI Powered Resume tool**
Course	CS/QTM/LING-329: Computational Linguistics
Authors	Zeshan Zahid, BA in Computer Science, zeshan.zahid@emory.edu
	Eric Ahn, BS in Computer Science, eric.ahn@emory.edu
	Daniel Ambrose, BS in Computer Science, danny.ambrose@emory.edu
	Nicholas Correa-Perez, BS in Computer Science, nicholas.correa-perez@emory.edu

Abstract

When it comes to the job market, nothing is more important than the resume. It is the end all be all to standing out as a candidate to potential employers[1]. Our project wants to improve your chances, and to address this we aim to use the power of Natural Language Processing, WebScraping, and the capabilities of ChatGPT-4 API. We want to enhance users' resumes through automating the parsing and categorization of resume content with regular expressions and text processing techniques. Then, we will utilize AI to optimize the user's resume structure and enhance its presentation. The core of our system lies in its ability for us to extract and section off detailed resume information (e.g. skills, work experiences, and educational backgrounds) with text processing, and transform this data into a professionally formatted resume using the ChatGPT API while also allowing user input to the completed resume. Through thorough testing with ChatGPT and Claude alone (i.e. Latex output for Overleaf), we find that they fail at this task when given more detailed resumes and tend to have many inaccuracies overall. Additionally, by leveraging web scraping technologies our system will identify job opportunities that align closely with the user's qualifications and career goals. By providing a tailored list of job recommendations, we will enable users to apply with resumes that are not only refined but also specifically aligned with potential employer expectations. Our approach not only saves users time but also enhances their marketability within the job market. The impact our project can have on the field is set up to be profound by offering the dual advantage of enhanced resumes as well as targeted job matching. Through our use of advanced language processing and web technologies, our project can create a user-focused experience that enhances their job search with the evolving workforce.

1 Introduction

1.1 Objectives

- Resume Parsing: We will use Flask within Python to develop a robust program that handles the extraction and categorization of detailed resume sections and other information. This system will serve as the foundation for data handling and processing within our project.
- Resume Enhancement: Integrate the ChatGPT-4 API to refine resume formatting and content presentation. By feeding parsed and organized data into the API, we aim to produce resumes that are not only professionally formatted but also optimized for clarity and impact.
- Job Matching: Implement web scraping techniques to collect job listings that align with the user's skills and experiences. This will involve scraping job boards and company websites to gather current and relevant employment opportunities.
- Intelligent Matching Algorithm: Develop a unique algorithm that uses the extracted resume data in conjunction with the scraped job listings to identify the best job matches. This algorithm will compare user qualifications against job requirements to ensure high relevance and fit. This goal of ours does come with some barriers specifically due to time but we believe we could also utilize AI to help us with some of the algorithms to get faster results.
- Overall Overview: Our projects in its entirety should return back a perfected resume along with relevant jobs and positions that target the user specifically.

1.2 Motivation

In recent years, the resume has remained one of the most important aspects of the professional world. Just a few pages of paper is what holds the door for potential students and job seekers looking for employment. By streamlining the process of resume editing as well as giving users a base of job listings, we give users more time on their hands to use effectively towards the hunt in the job market rather than drafting resumes. Our approach in categorization of the resume to then use the ChatGPT API makes the AI less likely to make mistakes within editing and formatting as well as ensures each resume accurately reflects the user and makes their resume more appealing. While our project mainly applies to college degree job seekers we think it's most beneficial for first-time applicants, especially those who just graduated as they generally have the least amount of experience in crafting a resume. Our project not only helps the technical aspects of job searching but also empowers users by providing them with tools that enhance their marketability. This support is crucial in helping individuals navigate the competitive job landscape more successfully.

1.3 Problem Statement

Our project addresses a specific NLP challenge by accurately extracting and interpreting complex resumes that standard AI tools like ChatGPT and Claude struggle with. These tools often misinterpret detailed resumes, sometimes adding or omitting critical information and failing to deliver professionally appealing formats[2]. Our approach involves breaking down resumes into categorized sections for more precise processing. We then use our model to format the data into a highly professional layout which overcomes the limitations of general AI tools and ensures that resumes are both accurate and effective in the job market.

1.4 Innovation Component

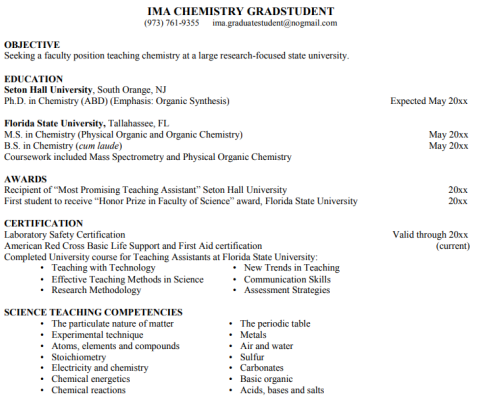
Our project introduces a unique approach to resume creation and extraction that differentiates itself from other resumes services and decreases inaccuracies caused from just relying on AI tools like ChatGPT or Claude. By parsing through a resume and forming it into categories based on contextual partitions like skills, relevant coursework, and experiences our system allows us to break down the task into smaller bits, enhancing each one with AI. Then with the enhanced data from each category we parse through it again to format it into a visually compelling and appealing resume. Additionally, our system integrates real job market data through web scraping different employment opportunities which allows users to have a good starting point to view their available opportunities.

2 Background

2.1 Related work and limitations

In today's day and age with advanced resume tools and AI chatbots like Claude there are many options in creating and editing resumes. Tools like MyPerfectResume and others similar do a great job integrating a layout to allow resume input and then return templates with certain edits and format changes. These sorts of services lack user input allowing users to make edits to smaller pieces of the resume which we will correct in our approach. Other chatbot services are a jack of all trades but can be very time consuming and can make errors in detailed or unique formats. These errors can be phantom changes caused by the LLMs not following the prompt as the resume gets longer[4]. This makes it very time intensive playing around with the chatbot with different prompts in hopes that you get a good output even though they still lack optimal format and will need major user input and tinkering. Additionally, neither service allows users to see real time job postings that are within the job market and match the user's skills and experience. AI chatbots can show general positions but due to the way they pull data cant show real time positions[5]. Our approach is different because it bridges the gap between these tools by using NLP to split up the prompts and text extraction and then retrieval to then form the professional resumes and at the end allowing user input to make slight corrections to sections while also with the unique approach of using the data we collect to find

real time job positions that are tailored to the users skill sets and educational background. Though their are some limitations especially with the way we will have to categorize the data in categories to get the best results it will be hard to do that manually but with the use of regular expression I believe we can cover most cases. To demonstrate the limitations of ChatGPT and Claude, we input a two-page example resume into



IMA CHEMISTRY GRADSTUDENT
(973) 761-9355 ima.graduatestudent@nogmail.com

OBJECTIVE
Seeking a faculty position teaching chemistry at a large research-focused state university.

EDUCATION
Seton Hall University, South Orange, NJ
Ph.D. in Chemistry (ABD) (Emphasis: Organic Synthesis) Expected May 20xx

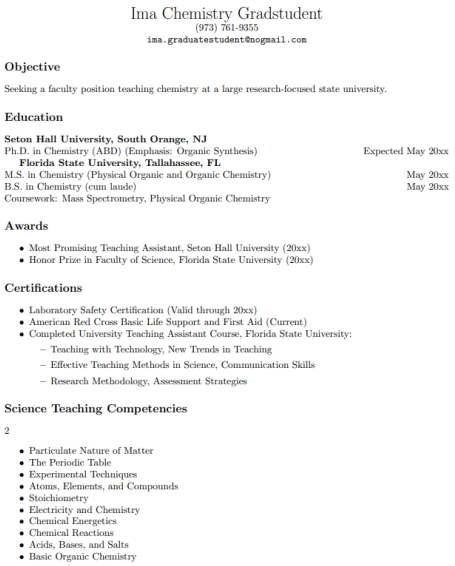
Florida State University, Tallahassee, FL
M.S. in Chemistry (Physical Organic and Organic Chemistry) May 20xx
B.S. in Chemistry (*cum laude*) May 20xx
Coursework included Mass Spectrometry and Physical Organic Chemistry

AWARDS
Recipient of "Most Promising Teaching Assistant" Seton Hall University 20xx
First student to receive "Honor Prize in Faculty of Science" award, Florida State University 20xx

CERTIFICATION
Laboratory Safety Certification Valid through 20xx
American Red Cross Basic Life Support and First Aid certification (current)
Completed University course for Teaching Assistants at Florida State University:
- Teaching with Technology
- Effective Teaching Methods in Science
- Research Methodology
- New Trends in Teaching
- Communication Skills
- Assessment Strategies

SCIENCE TEACHING COMPETENCIES
- The particulate nature of matter
- Experimental technique
- Atoms, elements and compounds
- Stoichiometry
- Electricity and chemistry
- Chemical energetics
- Chemical reactions
- The periodic table
- Metals
- Air and water
- Sulfur
- Carbonates
- Basic organic
- Acids, bases and salts

(a) Figure 1



Ima Chemistry Gradstudent
(973) 761-9355
ima.graduatestudent@nogmail.com

Objective
Seeking a faculty position teaching chemistry at a large research-focused state university.

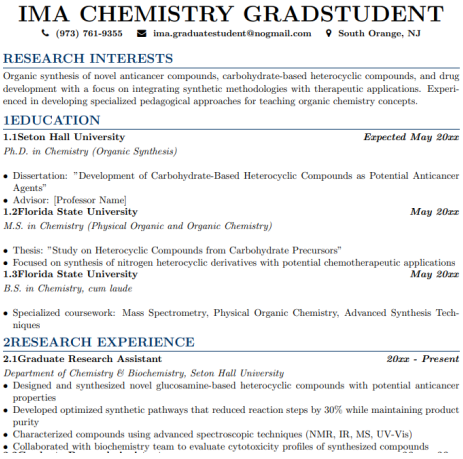
Education
Seton Hall University, South Orange, NJ
Ph.D. in Chemistry (ABD) (Emphasis: Organic Synthesis) Expected May 20xx
Florida State University, Tallahassee, FL
M.S. in Chemistry (Physical Organic and Organic Chemistry) May 20xx
B.S. in Chemistry (cum laude) May 20xx
Coursework: Mass Spectrometry, Physical Organic Chemistry

Awards
- Most Promising Teaching Assistant, Seton Hall University (20xx)
- Honor Prize in Faculty of Science, Florida State University (20xx)

Certifications
- Laboratory Safety Certification (Valid through 20xx)
- American Red Cross Basic Life Support and First Aid (Current)
- Completed University Teaching Assistant Course, Florida State University:
  - Teaching with Technology, New Trends in Teaching
  - Effective Teaching Methods in Science, Communication Skills
  - Research Methodology, Assessment Strategies

Science Teaching Competencies
2
- Particulate Nature of Matter
- The Periodic Table
- Experimental Techniques
- Atoms, Elements, and Compounds
- Stoichiometry
- Electricity and Chemistry
- Chemical Energetics
- Chemical Reactions
- Acids, Bases, and Salts
- Basic Organic Chemistry

(b) Figure 2



IMA CHEMISTRY GRADSTUDENT
(973) 761-9355 ima.graduatestudent@nogmail.com South Orange, NJ

RESEARCH INTERESTS
Organic synthesis of novel anticancer compounds, carbohydrate-based heterocyclic compounds, and drug development with a focus on integrating synthetic methodologies with therapeutic applications. Experienced in developing specialized pedagogical approaches for teaching organic chemistry concepts.

1EDUCATION
1.1Seton Hall University ***Expected May 20xx***
Ph.D. in Chemistry (Organic Synthesis)
- Dissertation: "Development of Carbohydrate-Based Heterocyclic Compounds as Potential Anticancer Agents"
- Advisor: [Professor Name]

1.2Florida State University ***May 20xx***
M.S. in Chemistry (Physical Organic and Organic Chemistry)
- Thesis: "Study on Heterocyclic Compounds from Carbohydrate Precursors"
- Focused on synthesis of nitrogen heterocyclic derivatives with potential chemotherapeutic applications

1.3Florida State University ***May 20xx***
B.S. in Chemistry, cum laude
- Specialized coursework: Mass Spectrometry, Physical Organic Chemistry, Advanced Synthesis Techniques

2RESEARCH EXPERIENCE
2.1Graduate Research Assistant ***20xx - Present***
Department of Chemistry & Biochemistry, Seton Hall University
- Designed and synthesized novel glucosamine-based heterocyclic compounds with potential anticancer properties
- Developed optimized synthetic pathways that reduced reaction steps by 30% while maintaining product purity
- Characterized compounds using advanced spectroscopic techniques (NMR, IR, MS, UV-Vis)
- Collaborated with biochemistry team to evaluate cytotoxicity profiles of synthesized compounds

(c) Figure 3

Figure 1: Outcomes of AI

both LLMs (Figure 1). The prompt just asked for ChatGPT and Claude to improve the resume and produce the output in Latex so that we could create a pdf of the output in Overleaf (Figure 2 and Figure 3).

For Claude, the LLM you recommended, we immediately found that it made up information in Figure 3 with the dissertation because there is no mention of a dissertation in the original resume. It likely made up this information because it was viewing the resume as a whole instead in parts. We believe that this issue and other inaccuracies will be solved using our approach of categorization of resume content. Also more tinkering with this showed similar results where it would leave of chuncks of information or just make them up.

For ChatGPT, in Figure 2 you can immediately tell that the output is bland and does not stand out at all. The ChatGPT version looks arguably worse than the original resume such as with the Science Teaching Competencies section. Our approach will use a preset optimal resume structure that will inherently eliminate this issue.

You could argue that inputting sections into ChatGPT and Claude instead would produce the same result, but we want to streamline this process for users so that there is as little manual effort as possible. Additionally, the integration of best-fit job postings sets our project apart from any related works while providing an innovative, accurate, and streamlined user experience. Also at time when outputting a Latex file or other similar structure both LLMS would sometimes crash.

3 Proposed Approach

3.1 Methodology

To begin our technical approach, we will accurately parse through the resume given by the user and clean it up to prepare for edits. We then parse again and use various techniques, regular expression and others to then split the resume into categories like profile, experience, and skills. Once categorized we can prompt the GPT API to enhance the content grammatically and professionally and refining all aspects and then once we receive the JSON response back we then reimplement the data into the new chosen resume format. Our resume structure is chosen by a highly rated resume template that many professional programs and jobs recommend. Once the resume is created then we scrape job opportunities based on the users skills and education to get tailored jobs and we will return the top choices based on an algorithm we will write to compare both sets of skills needed within the job and those possessed by the user.

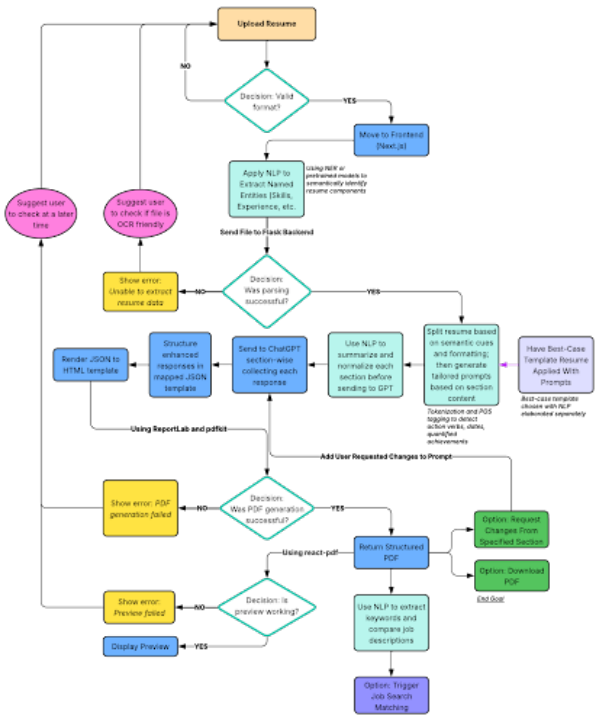


Figure 2: Structural Diagram

3.2 Framework

Our framework will consist of basic Flask to use for creating the programs within python. We will then incorporate that and connect it with Nextjs for the user interface or if needed could just be HTML. We don't see a need to implement a database as the user information will just be held and used during the session.

3.3 Implementation and Quality Assurance

Our implementation strategy relies on a split approach of labor. We have four members so we split each task into phases and all four members have a task during each phase. We have hard deadlines of one week per phase to make sure we implement on time and stay ahead of deadlines. To insure quality we have already created a github repository for the project that each member has access to then keep changes on a linear timeline and everybody on the same page. We plan to implement testing as we go through each phase as well as a larger test at the end of each phase to make sure are capabilities are met. We plan to also ask students in class to test certain aspects and demo each phase to get feedback for changes to keep quality at the peak.

4 Timeline

4.1 Weekly Schedule

- In Week 1 we set up the GitHub repository and deployed the initial project Zeshan and Danny worked in Flask to build the resume upload and parsing features and organized resumes into clean structured categories Nicholas and Eric started working on a job scraper based on education and skills To reduce risk we split into pairs to catch issues early and keep communication open

- For Week 2 Danny will set up ChatGPT API prompts to send resume categories and make sure the responses match our structure Zeshan will parse those responses into the improved resume format Eric will connect the scraper to the skills we have stored and Nicholas will begin building a job matching algorithm to rank results based on user skills

- In Week 3 Nicholas will create the UI for uploading resumes and making edits through a chat interface and generate a PDF with top job matches Zeshan and Danny will finalize the JSON format and Eric will add a function to let users edit their resume categories after reviewing them

- Week 4 will be focused on testing fixing bugs and making final adjustments Week 5 will include live demos and using feedback to make last updates before the final submission

- To manage risk we plan regular check ins and track all tasks in a shared Google Sheet Everyone has access to the GitHub repo and can switch tasks if needed We split work based on tasks instead of assigning frontend or backend roles so everyone stays involved in all parts of the project We use GitHub and group chat to keep communication constant and open

References

[1]Why Is a Resume Important? (Types and Why You Need One) — Indeed.Com, www.indeed.com/career-advice/resumes-cover-letters/why-is-a-resume-important. Accessed 24 Mar. 2025.

[2]Piper Center for Vocation and Career. "Ai and Your Resume: Advantages, Limitations, and Tips." AI and Your Resume: Advantages, Limitations, and Tips – Piper Center for Vocation and Career, wp.stolaf.edu/pipercenter/ai-and-your-resume/. Accessed 23 Mar. 2025.

[3]Break down Your Prompts for Better AI Results - Relevance AI, relevanceai.com/prompt-engineering/break-down-your-prompts-for-better-ai-results: :text=Decomposed

[4]Ibm. "What Are Ai Hallucinations?" IBM, 13 Feb. 2025, www.ibm.com/think/topics/ai-hallucinations.

[5]"How to Scrape Websites with Chatgpt in 2025." Crawlbase, 3 Apr. 2024, crawlbase.com/blog/chatgpt-web-scraping/: :text=No

A Contributions

We all contributed equally to the report

- Zeshan Zahid 25 percent
- Daniel Ambrose 25 percent
- Eric Ahn 25 percent
- Nicholas Correa-Perez 25 percent