

## *1. Statement regarding the goal of the thesis or portfolio*

*What do you intend to discover or create during your Honors Thesis/Portfolio experience?*

For the Honors portfolio, I intend to build a web platform to help people apply to jobs in a more easy and effective manner. In Spring 2021, my team and I in CS 490S: Software Entrepreneurship developed a proposal and business plan for our resume-building platform called Resumeble. We discovered a few important key points. On average, candidates will apply to 276 job applications before receiving an offer. The main reason is that around 75% of all resumes get rejected before they reach the hiring manager through an application tracking system, which collects, sorts, and ranks the resumes it receives for open positions. Most resumes are eliminated because they don't match the job description to lack of keywords, formatting, and impact. Additionally, 54% of potential employers get rejected due to their social media accounts such as their LinkedIn Profile. Also, on average, it takes a candidate at least four to six hours to create a resume alone.

From these key points from above, we looked into current competitors in this specific area in the market. We found that competitors such as ResumePuppy, VMock, and JobScan focus on one or two areas. They are mostly focused on creating well formatted resumes and can possibly analyze the resume based on the job description or a set of key criteria through Artificial Intelligence and Natural Language Processing. However, there is a clear lack and need for helping candidates create and tailor their resumes quickly and effectively. It's vital that we capitalize on this issue as it will help benefit both the job applicant and recruiter. For the job applicant, our platform looks to reduce the time for people to create resumes based on the job description. We want to help candidates spend more time on applying to more job opportunities while also giving them the best chance to overcome the applicant tracking system and get a recruiter to take a look at their resume. For the recruiters and the company they represent, we look to reduce the number of applications that get tossed for formatting mistakes or missing key details that could have rejected a qualified candidate from being considered for an open position in the company. This way, we want to help recruiters identify candidates that are more qualified based on the job description and bring in better talent to the company.

For the thesis I will be working on the project with my classmate Herlin Rijo to develop the product, however we will be focusing on two separate sets of tasks for each of our proposals. I will focus on building three main aspects: developing services for our full-stack application related to authentication and user management; NLP and AI development for keyword analysis and resume ranking; as well as development and cloud management through AWS, Docker, and Nginx. More specifically, I'll be responsible for creating an AI system that extracts specific information from a user's profile and creates a resume based on a job description inputted by the user.

Herlin will be focusing on learning full-stack development and implementing both the visual aesthetics of the product and backend database counterpart where user information will be stored and the AI draws data from. His focus will be to develop the profile builders users will utilize by the AI system to create a general resume. Users will be able to input their data such as past jobs, education, skills and experiences. Then using this information, the AI system will be able to pull key characteristics of the user and give recommendations and analytics for areas of improvement in their profile and the resume created.

Through the Fall 2021 to Spring 2022 year, I intend to explore how we can help people apply to jobs in a more easy and effective manner and build a platform designed to help a fairer job application process which helps the job applicant and recruiters.

## 2. Key readings

*List the items that you plan to read in preparation for your own work. A literature review in a format appropriate for your discipline is required as part of the proposal for the 499T/P semester.*

### **ACM citation**

1. Vedant Bhatia, Prateek Rawat, Ajit Kumar, and Rajiv Ratn Shah. 2019. End-to-End Resume Parsing and Finding Candidates for a Job Description using BERT. (October 2019). Retrieved May 14, 2021 from <https://arxiv.org/abs/1910.03089>
2. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. (May 2019). Retrieved July 1, 2021 from <https://arxiv.org/abs/1810.04805>
3. Shiqiang Guo, Folami Alamudun, and Tracy Hammond. 2016. RésumMatcher: A personalized résumé-job matching system. (April 2016). Retrieved May 19, 2021 from <https://www.sciencedirect.com/science/article/pii/S0957417416301798>
4. Jing Li, Aixin Sun, Jianglei Han, and Chenliang Li. 2020. A Survey on Deep Learning for Named Entity Recognition. (March 2020). Retrieved June 30, 2021 from <https://arxiv.org/abs/1812.09449>
5. Sam Newman. 2018. Building Microservices: Designing Fine-Grained Systems, O'Reilly Media, Inc, USA.
6. Tim Zimmermann, Leo Kotschenreuther, and Karsten Schmidt. 2016. Data-driven HR - Résumé Analysis Based on Natural Language Processing and Machine Learning. (June 2016). Retrieved May 19, 2021 from <https://arxiv.org/abs/1606.05611>

## 3. Communication

- *How often will you meet with your Committee Chair?*

I'm looking to communicate with my committee chair at least one hour once per week from September to December.

- *What are your Committee Chair's expectations of such meetings?*

Meetings will be conducted remotely through Zoom or in person. I expect to go over submitted assignments and the progress done on the overall project. In certain weeks, we'll be displaying the working prototype to the professor for feedback.

#### **4. Specialized training**

*Does your proposed activity involve live human beings? If so, your committee chair may need to contact the UMass Office of Human Research Protection to establish whether your proposed work falls under the federal definition of human subject research. Please indicate the status of your proposed work in this regard*

We won't have any human subjects in the project and the proposed work does not fall under the federal definition of human subject research.

*If work on your thesis or portfolio should require specialized training (e.g., human/animal subjects testing, lab-safety instruction, survey technique), please specify the type of training required and when and from whom that training will be obtained.*

For the thesis, the specialized training that would be required would be through understanding AI/ML techniques. I'll be educating myself by reading some of the literature to understand how resumes get ranked by applicant tracking systems. Most of the articles involve how researchers use NLP models such as BERT and IR methods to rank resumes. I hope to take some of this information to reverse engineer the process and develop a rule-based method to identify how to create resumes that tailor to the job description. Beyond researching these methods, I'll be auditing into DeepLearning.AI's online course for Natural Language Processing. The course would give me good practice using Tensorflow and developing NLP models for common business and real-world use cases.

I will also plan to gain experience in using Docker, an open platform used to separate applications from infrastructure so that we can deploy our software quickly and efficiently. I plan to gain experience and insight by reading the books Sam Newman. 2018. Building Microservices: Designing Fine-Grained Systems by Martin L. Abbott and Michael T. Fisher. 2015. The art of scalability: scalable web architecture, processes, and organizations for the modern enterprise.

## 5. Methods or agenda

*What are you going to do to complete the portfolio/thesis? In language appropriate for your disciplinary area, describe how you will decide on your methods or agenda.*

*If you have already decided on your methods or agenda and have begun or will start your work in the 499Y semester, then describe, again in language appropriate for your disciplinary area, what you will do.*

*If you are not in a position to describe your methods or agenda at this stage, note that you are required to do so in your proposal for the 499T/499P semester.*

In the fall 2021 semester, I intend to build out some of the core features necessary to this application. Herlin Rijo and I plan to divide the project into separate features and work collaboratively to integrate the features together.

One of my main focuses on the project will be the full-stack development of the product. I'll be working with our proposed tech stack to develop the authentication & user management of our application. Users will be able to login, register, reset their password, and manage their user settings within our application. Additionally, I will be researching and developing the technology to turn a user's profile page into an optimized resume. This will involve learning techniques in Natural Language Processing and Information Retrieval to rank a user's job experience and projects from the profile to a resume.

To develop these features, I will be adopting the Agile software development methodology which involves discovering requirements and developing solutions through the collaborative effort of my partner. In addition, throughout development I will conduct sprints in which I will showcase specific work that has to be completed and made ready for review. My work will be reviewed by the product owner, Professor Neena Thota. Before implementing any of the code I will be developing user cases, flow diagrams, and mock UI diagrams regarding the platform. After implementation, I plan to perform tests of usability at different stages of platform development to ensure everything is working and on track.

## 6. Timeline and Grading

### 6.1. Describe a graded assignment due by the W (withdrawal) date.

- The draft of the introduction and background section of the 499P proposal shall be submitted before the withdrawal deadline on October 16th, 2021. The introduction will discuss the motivation for the project, the methodology, and significance of the work in

itself. The background section of the 499P proposal will provide an overview of the readings done and how these readings play a role in our project.

## *6.2. Describe and give dates for other first-semester assignments.*

9/3/2021 - Written Summary of key readings 5  
9/10/2021 - Written Summary of key readings #1  
9/17/21 - Written Summary of key readings #2 + #3  
9/24/21 - Written Summary of key readings #4 + #6  
10/1/21 - Initial Development of Project User Stories and Requirements  
10/8/21 - Completed, detailed, development of Data flow diagrams & API research  
10/15/21 - Draft of Introduction & Background Section of 499P Proposal  
10/16/21 - Withdrawal Date  
10/22/21 - Completed, detailed, development of UI mockups and sequence diagrams  
10/29/21 - Draft Of Methodology of 499P due  
11/5/21 - Draft of Communication + Timeline Due  
11/12/21- Evaluation Sections + References  
11/19/21 - First Draft of 499T/P Proposal  
11/26/21- Thanksgiving Week  
12/3/21 - Final Submission of 499T/P Proposal

## *6.3. Give the date by which the second committee member will have been selected. Give the name, if known.*

- The second committee member will be chosen at latest by October 10th, 2021.

## *6.4. Give the date by which the first draft of your literature review will be submitted to your advisor.*

- The date by which the first draft of the literature review will be submitted is by October 29th

## *6.5. Give the dates by which (a) a draft of your 499T/P proposal will be submitted to your advisor, and (b) your 499T/P proposal will be uploaded to PATHS (this deadline should be before the end of the 499Y semester).*

- The first draft of the 499T/P proposal will be due on November 19th, 2021.
- It should be uploaded to PATHS by December 3rd, 2021.

## *7. COVID-19 Impact Statement*

***As explained in a recent campus-wide email from Chancellor Subbaswamy and on the Get Ready for Fall 2021 webpage, UMass Amherst expects campus life to return to normal operations in Fall 2021. In an effort to be prepared given the ongoing COVID-19 pandemic, after consulting with your committee chair, please enter a statement that describes alternatives to your Face-to-Face research plans in the 2021-2022 academic calendar and how you will address those impacts to ensure the successful completion of your work. This statement should address: (a) whether your planned activities require in-person access to on-campus or other resources (labs, studios, archives, performance spaces, field research sites, etc.) and if so, whether your committee chair has obtained all required permissions for you to access these resources in person through their Research and Library Operating Plan (if required); (b) what your contingency plans are in the event that health concerns warrant a return to fully remote operations significantly before the last of classes on December 8th.***

In the 2021-2022 academic year, we will not require in-person access to campus resources such as labs, studios, and archives. The planned activities can be done synchronously. In the event that health concerns warrant a return to fully remote operations significantly before the last day of classes on December 8th, we shall operate as normal. Meetings will still be conducted through Zoom and assignments will still be sent through email, or some other form of assignment submission platform.