

Capstone Project Proposal – Virtual Technical Leader

Bhavya Nileshbhai Vadher (C0894977), Jesus Andres Correal Ortiz (C0872634),
Mary Luz Gomez (C0891136), Rehan Khokhar (C0896278), Viki Kanubhai Patel
(C0906295)

Artificial Intelligence and Machine Learning, Lambton College

AML 2404 2: AI and ML Lab

Dr. Simrandeep Kaur

September 28th, 2023

Virtual Technical Leader (VTL)

Handling information inside big corporations is very challenging; every department generates information daily, such as company policies, procedures to run a machine, processes to place purchase orders, or more advanced knowledge, such as research results and experimental analysis. All that information is generally placed in different systems, fragmenting employee access to information. Additionally, and in most cases, the applications used for search are built on basic algorithms. This leads to immense hours wasted for an individual reviewing many files or trying to find the right person when trying to answer a particular question.

VTL aims to solve this issue by creating an Artificial Intelligence entity in the organizations. This role would act as a knowledge discovery system that every employee can contact when any question is related to the organization. Having the power of NLP and LLM will allow employees to have conversations related to company topics, which will reduce the time for employees to get the answers required for the job, make the onboarding process more straightforward, eliminate work duplications by having an easy understanding of what has already been done, and finally, will release time for the most specialized talent who will be able to focus more on strategic projects rather than answering routine questions.

VTL will transform the information and knowledge generated in companies into gold by enabling all individuals in the organization with easy access and comprehension.

Statement of Need

Companies generate vast amounts of information and knowledge in different forms, primarily distributed in diverse systems and handled by mostly obsolete applications using simple search systems, making it difficult to get the exact needed answers. This situation creates several areas for inefficiency in an organization.

Firstly, unreachable information or knowledge means a considerable investment in producing and storing it is worthless. Secondly, the quantity and quality of decisions can be seriously impacted if the right know-how is not used. Thirdly, big corporations can have expensive teams working on the same thing. Finally, employees are less productive if they spend considerable time getting the necessary information.

In this last regard, according to a 2012 McKinsey report, employees spent an average of 1.8 per day or 9.8 per week looking for information. Therefore, it could be said that for every five employees, one only searches for information in the company without adding real value to the organization.

We did a simple analysis based on the above statistics to understand how much it could cost an organization to have the wrong mechanisms to manage the company's knowledge and information. Assuming an average salary of CAD 80,000 annually per employee, we determined that the cost for a company would be around CAD 20M per 1,000 employees.

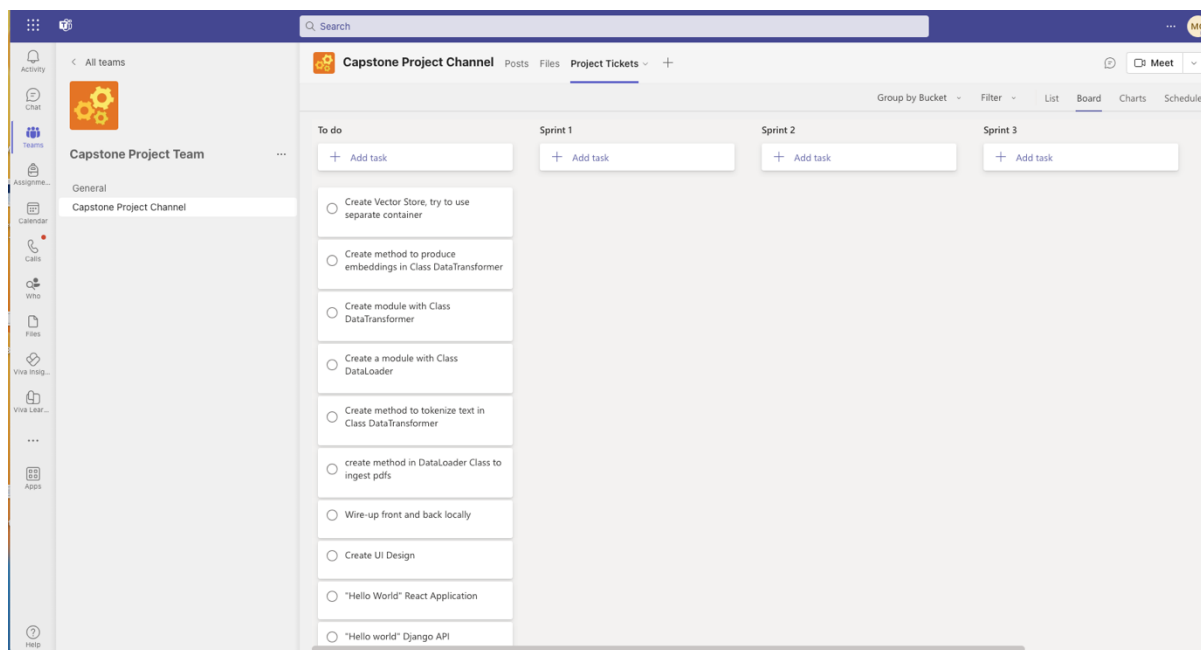
In conclusion, the information and knowledge discoverability in an organization is critical, and having a lousy management system can create severe losses for an organization. Additionally, we firmly believe that using Artificial Intelligence, specifically NLP with a Large Language Model combined with the semantic search concept, we can effectively solve each organization's knowledge and information discoverability problem. However, as proof of concept in VTL, we will tailor/simulate one specific company to demonstrate the system's effectiveness.

Project Methodology, Activity, and Outcomes

Methodology

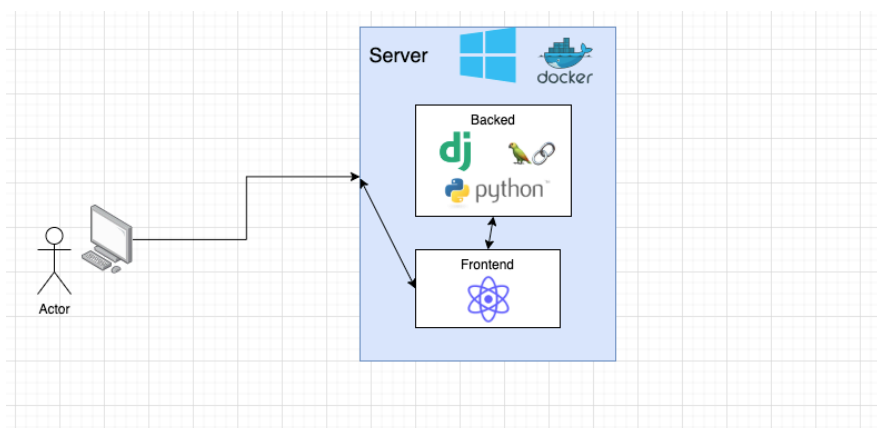
We will use the Scrum framework as a project management tool. For this purpose, we created a Microsoft Teams channel, which will be used to conduct our daily async updates, define the tasks for the sprint, which will be every week, track the project progress,

and develop and follow up on the tasks for each sprint. Below is a picture of our project board, where several tasks have already been created and waiting to be moved to the first Sprint.



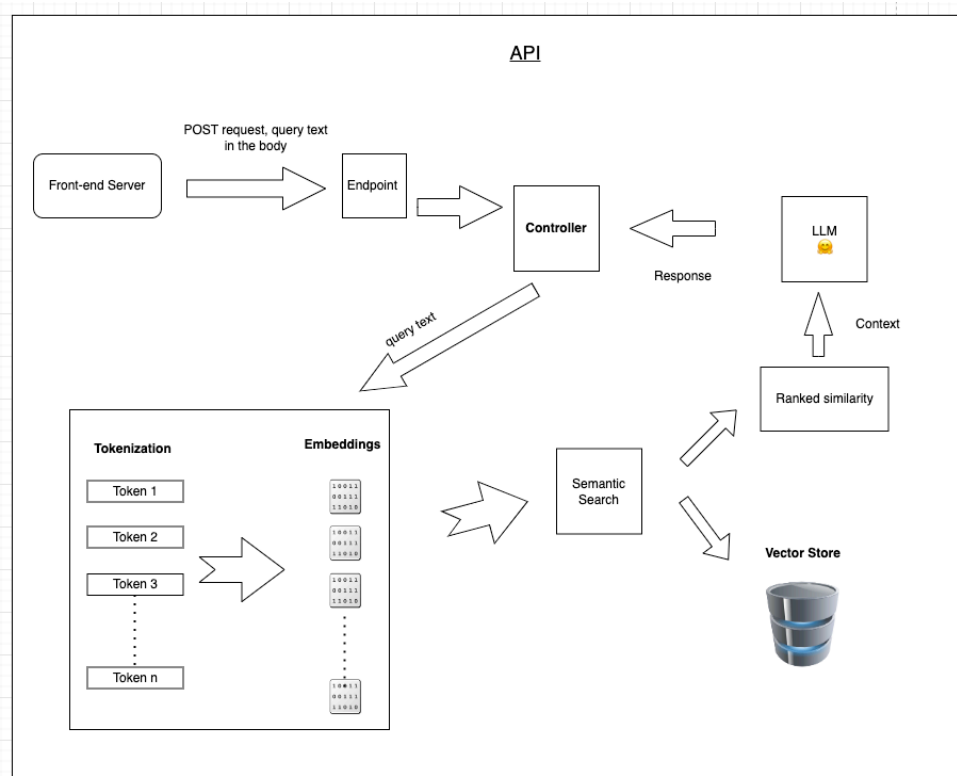
Activity

Application Architecture



We will implement a simple architecture with a Docker container holding the backend and frontend servers. The frontend will be a React application that will capture the user text

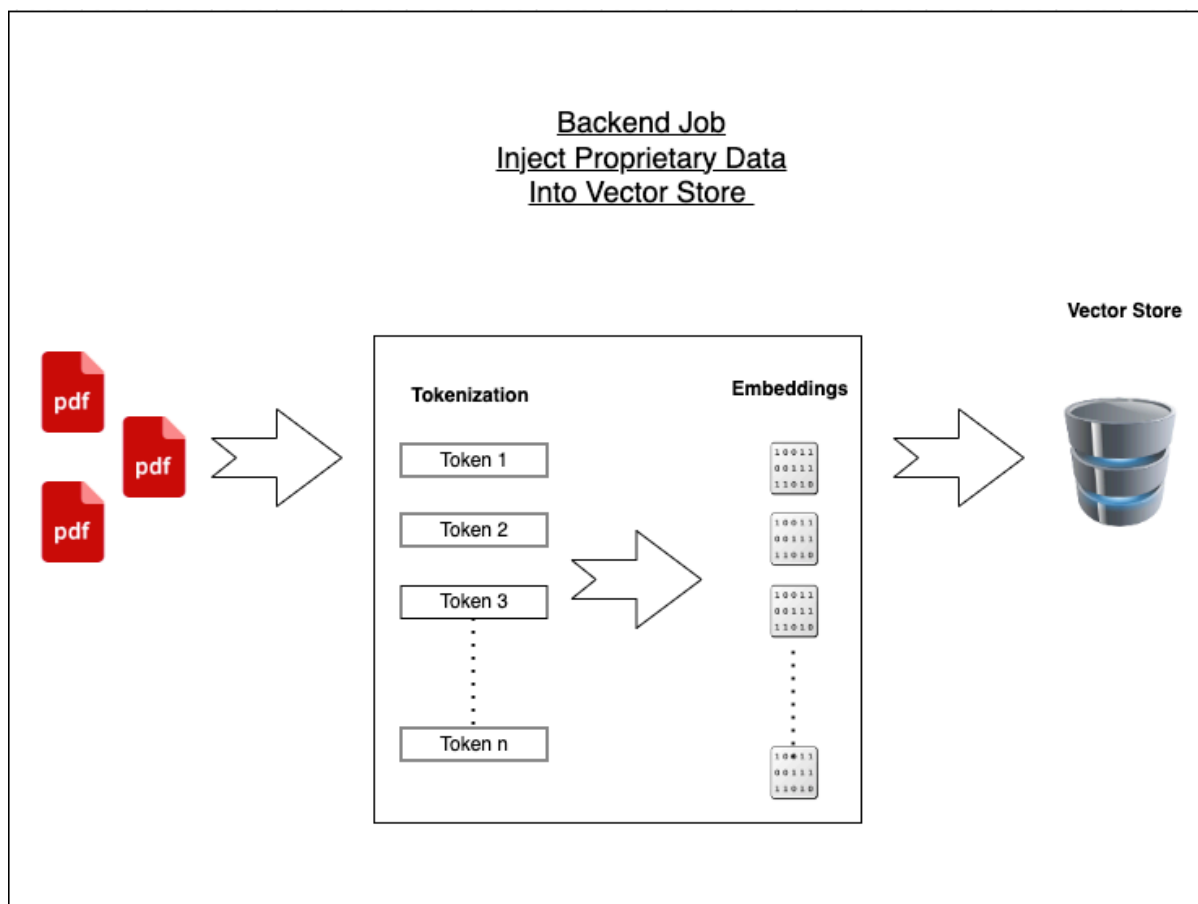
input and then send a request to a Django API. The following graph summarizes the processes in the backend.



The graph above shows the flow of work in the application's backend. First, the text query is passed to a unit that will transform the data into embeddings, allowing the semantic search to locate the most similar information in the vector store. Finally, the information will be passed to a trained Large Language Model using this context to respond to the frontend application appropriately.

Backend Job

The following graph shows a job we will build to transform the company information and store it in the vector store.



Outcomes

VTL will be, in this initial state, a fully functional application, including a UI that will aim for an intuitive user interaction with the application. Furthermore, we will develop an API that will handle the backend tasks along with a job that will load and transform the company information to be stored in a vector store.

We expect the app will work satisfactorily, being able to answer questions about a selected company in a natural and human-like way, delivering quick responses to the queries and a pleasant overall experience.

Evaluation

We will use the backend job to create a vector store with what we call “Company Information.” This information will be new for the model and could be domain-specific data

from the internet. Thereafter, we use the application to ask questions related to the data loaded into the vector store ("Company Information"), and based on the responses, we can conclude if the responses are appropriate.

References

- ProProfts (2019, July 19th). Does Your Workforce Spend Too Much Time Searching for Information? Retrieved from <https://www.proprofskb.com/blog/workforce-spend-much-time-searching-information/#:~:text=According%20to%20a%20McKinsey%20report,average%E2%80%9494searching%20and%20gathering%20information.>
- Xenit (2012, July). Do workers still waste time searching for information? Retrieved from <https://xenit.eu/do-workers-still-waste-time-searching-for-information/>
- McKinsey (2019, January). The social economy: Unlocking value and productivity through social technologies. Retrieved from <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-social-economy>
- Alejandro AO (2023, May 29th). Chat with Multiple PDFs | LangChain App Tutorial in Python (Free LLMs and Embeddings). Retrieved from <https://www.youtube.com/watch?v=dXxQ0LR-3Hg>