✍️ Product requirements

🎯 Objective

The project encompasses several key phases, starting with data collection and preparation through web scraping. This initial stage aims to build a solid foundation of knowledge for the Q&A agent. Following this, the development of a domain-specific Language Model (LLM) will take place, which is essential for equipping the Furhat robot with the ability to generate relevant and accurate responses. The subsequent integration of this agent with the Furhat robot marks a significant milestone, culminating in a platform where users can interact and receive information through a conversational interface. Additionally, the project will enhance the platform with receptionist functionalities, making it a versatile tool for users.

Find more details in Project Overview 🟦 Project Overview

📊 Success metricse

Goal	Metric
Simplify user's experience when browsing website.	Customer Satisfication increases

🤔 Assumptions

1. **Advanced Technical Infrastructure**: It is assumed that there is access to advanced technical infrastructure and resources necessary for web scraping, developing a domain-specific Language Model (LLM), and integrating these technologies with the Furhat robot. This includes hardware, software, and networking capabilities that can support the complexities of AI and robotics development.
2. **Availability of Data**: The project assumes that the targeted websites, such as Melbourne Connect or CIS, have a wealth of information available for scraping and that this data is accessible and can be legally and ethically used to train the Language Model. It also assumes that the data is structured or can be structured in a way that is conducive to effective web scraping and processing.
3. **Expertise in AI and Robotics**: There is an assumption that the project team, including the Product Manager, Scrum Master, Architecture Lead, Quality Assurance Lead, Development Environment Lead, and Deployment Lead, possesses the necessary expertise in AI, machine learning, natural language processing, and robotics. This expertise is crucial for developing a domain-specific LLM and integrating it with the Furhat robot to achieve the desired outcomes.
4. **User Engagement and Interaction**: The project presupposes that users will be willing to engage with and accept information from a robotic platform. It assumes that the interaction between users and the Furhat robot will be seamless, natural, and effective in providing the desired information and assistance, thereby enhancing the user experience.
5. **Scalability and Adaptability**: There is an underlying assumption that the system developed will be scalable and adaptable to future needs and technologies. This includes the ability to update the knowledge base with new information, refine the LLM for better accuracy and relevance, and integrate additional functionalities into the Furhat robot as user needs evolve.

🛠️ Functional scopes

1. **UI interactive interface:** Users can interact through the UI interface, which will present the main information and functions of the Q&A platform. The user enters or expresses the information to be queried, and the UI interface presents the result in a natural language

interaction.

2. **Furhat Robot interaction:** Furhat Robot will be integrated into the Q&A platform. The robot can receive user information and establish communication with data, models, etc. Finally, the conclusions generated by the Language Model can be passed to the UI interactive interface.
3. **Domain-Specific Language Model function:** Create a specialized LLM tailored to the needs and context of the Q&A system. The model will be designed to understand and process queries in a specific domain, enabling it to generate accurate and context-relevant responses.

🗒️ Requirements

Requirement	User Story	Estimation	Priority	Justification	Jira Issue
UI interactive interface Support text interaction in the robot interface.	U1.1: As a user, I want to be able to support text interaction in the robot interface, so I can type and express what I need to query.	SMALL	MUST H...	This is a must have because it is the basic function of user interaction. It may involve only the front end and is expected to be small.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-14
UI interactive interface Generate an livechat box when asking the question.	U1.2: As a user, I want the robot to generate an livechat box when I ask my question, so I can gain the information and answer I need directly.	SMALL	MUST H...	This is a must have because it is the basic function of user interaction. It may involve only the front end and is expected to be small.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-15
Furhat Robot interaction Must summary the information quickly.	U2.1: As a user, I want to interact with the Furhat robot in a conversational manner to obtain information directly, so that I can save time by not having to search and filter information on the web myself.	LARGE	MUST H...	This is a must have because the robot needs to feedback the information needed by the user, and involves the model, UI and robot interaction, which is expected to be a large project.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-2
Domain-Specific Language Model function Must provide accurate information of Melbourne Connect to user when asked.	U3.1: As a user interested in services offered at Melbourne Connect, I want the robot to provide detailed information (such as room's information, location and provided services) from websites of	LARGE	MUST H...	This is a must have because the language model needs to analyze user problems and obtain relevant information of Melbourne Connect, the project is expected to be large	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-3

	Melbourne Connect, so that I can clearly understand the overview of Melbourne Connect without navigating through those websites.			due to the model involved.	
Furhat Robot interaction Must be able to filter and sort information based on user-defined criteria such as location, size, and function of the room.	U2.2: As a user, I want the robot to filter and sort information based on my provided criteria (e.g., location, size, function of the room), so that I can find what I'm looking for more efficiently.	MEDIUM	SHOULD...	This is a should have because filters and sorting functions can better present information. It involves robot interaction with UI, and the project is expected to be medium.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-4
UI interactive interface Must be able to guide user on how to using furhat robot.	U1.3: As a new user, I want the robot to offer an introduction on how to use it, so that I can quickly understand and start utilizing the platform services.	SMALL	SHOULD...	This is a should have because the introduction function allows users to better understand the role of QA robots. The front-end is involved, and the engineering quantity is expected to be small.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-5
Furhat Robot interaction Must be able to describe the key information based on the website summary to visually impaired user.	U2.3: As a visually impaired user, I want the robot to describe images and visual content from websites, so that I can comprehend visual information that I cannot see.	LARGE	COULD ...	This is a could have because this feature is an additional optimization feature to visually impaired users and does not affect the core functionality of the project. Due to the development of images, models, etc., a large amount of work is expected.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-6
UI interactive interface Must protect user's private information and provide clear privacy policies.	U1.4: As a user, I want the Q&A platform to protect my personal information and provide clear privacy policies, so that I can ensure the security and confidentiality of my data.	MEDIUM	MUST H...	This is a must have because user information and data privacy is very important, and we need to make sure that we are clear about it.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-7

Domain-Specific Language Model function Must be able to summary the information of the website clearly and quickly.	U3.2: As a user, I want the chatbot to automatically extract and summarize the main content of a website I provide, so that I can quickly grasp what the website is about without reading all the content.	LARGE	SHOULD...	This is a should have because it is based on the language model to complement the functionality. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-8
Domain-Specific Language Model function Must be able to recommend information based on user's interests and website summary.	U3.3: As a user looking for entertainment or news, I want the chatbot to recommend and summarize articles, blog posts, stories, or provide me with the latest headlines based on my interests, so that I can easily find engaging or informative content.	LARGE	MUST H...	This is a must have because it is the key function with language model to provide the correct and necessary answer. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-9
Domain-Specific Language Model function Must be able to extract information from various careers websites and provide a clear summary job list.	U3.4: As a job seeker, I want the chatbot to compile job listings from various websites, so that I can find opportunities that match my skills and preferences more easily.	LARGE	MUST H...	This is a must have because it is the key function with language model to provide the correct and necessary answer. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-16
Domain-Specific Language Model function Must be able to extract academic information from scientific websites.	U3.5: As a user doing research, I want to ask the chatbot specific questions about content found on academic or scientific websites, so that I can gather information efficiently for my studies or work.	LARGE	MUST H...	This is a must have because it is the key function with language model to provide the correct and necessary answer. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-11

Furhat Robot interaction Must be able to translate web content from various of languages.	U2.4: As a user learning a new language, I want the chatbot to translate content from websites in foreign languages(e.g. from Chinese to English), so that I can understand the content without being fluent in the language.	LARGE	COULD ...	This is a could have because it is based on the language model to complement the functionality. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-12
Furhat Robot interaction Must be able to avoid answering unrelated questions with the website.	U2.5: As a user, I want the robot to avoid answering questions that are unrelated to website navigation content or involve sensitive information, and to inform me when my queries are outside the scope of available information, to maintain professionalism and set clear expectations.	HIGH	MUST H...	This is a must have because it is the key function with language model to provide the correct and necessary answer. The project is expected to be large due to the model involved.	https://comp90082-2024-qa-koala.atlassian.net/browse/C2QK-13

📤 Out of Scopes

1. **Furhat Robot Hardware Customization and Modification**: Since we are only using Furhat's SDK, any work involving robot hardware is outside the scope of the project.
2. **Furhat Robot Maintenance**: We will use the SDK of Furhat Robotics for platform development, and there will be no maintenance of Furhat robot.
3. **Supports multiple language conversations:** At present, the robot mainly provides services for Melbourne Connect, and will communicate in English, the main local language. Implementing multilingual dialogue capabilities will significantly increase the complexity of the model and integration, and therefore may be outside the scope of the current project stage.
4. **Emotion recognition:** The current project only focuses on the question and answer content of the robot, and the emotion recognition part may require additional hardware and technology to be considered, which is not within the scope of the current project.
5. **Voice interaction and input：**We mainly consider that the robot question answering platform interacts with natural language, and voice input may have to consider more technical requirements, function integration and other issues, which are not within the scope of our discussion.