ECE 445

SENIOR DESIGN LABORATORY

FINAL REPORT

Campus Tour Guide by AI-Powered Autonomous System

Team #21

XUANBO JIN
(xuanboj2@illinois.edu)
HAO REN
(haor2@illinois.edu)
YUNTONG GU (yuntong7@illinois.edu)
WEIANG WANG
(weiangw2@illinois.edu)

TA: Xinlong Huang

March, 18, 2024

Contents

1 Introduction

1.1 Problem and Solution Overview

Anyone entering a place for the first time, like an university, can be quite challenging. Knowing where you are, how to get to your destination, how to optimize your routes, knowing factors that will influence your routes can be complicated. Having a real-time interactive system that guides people through this process is needed. It has been possible yet not able to scale because it's not open-sourced, and its hardware isn't standardized, and is expensive. The interaction isn't versatile enough to adapt well under the ever-changing applications. A cheap and versatile solution is needed.

1.2 Motivation

The most traditional paradigm is having human tour guide guiding a group of visitors. When the era of electrical and electronics engineering came, engineers started automating this process. They designed specific pipelines to emulate this process. However, this automation has intrinsic problem with cost, scalability, as well as generality. Currently, when artificial intelligence prevails, simply inserting AI as a component in the pipeline has significant downsides. Apart from previous problems, most of the components in the pipeline is not AI-powered, wasting computational resources and efficiency. In light of these problems, we completely shift the pipeline design to a Multi-agent AI operation network design, with each of its component AI powered. It can scale easily as all agents have simple and general interfaces. The average cost is also amortised as the number of agents increase. Most importantly, the paradigm of all the downstream applications, including campus tour guide, is shifted from 1 guide leading X visitors to X guides leading 1 visitors, letting users to these applications fully exploit the power of AI.

1.3 Visual Aid

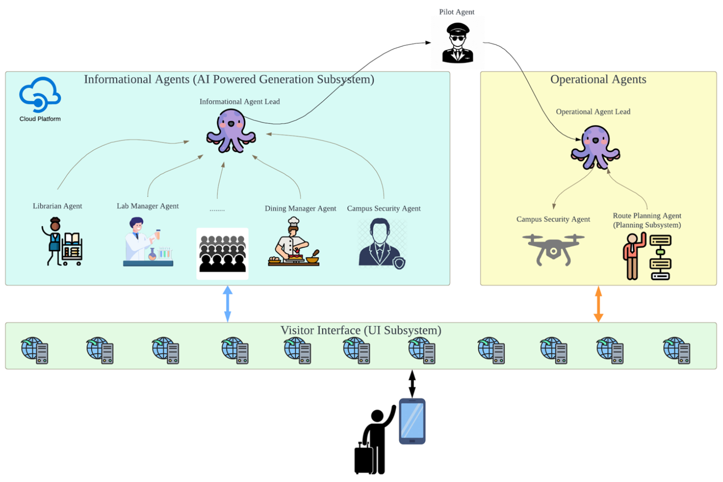


Figure 1: Multi-agent AI operation Visual Illustration

1.4 High-level requirements list

- The AI agents network system should be **responsive**. It should respond to user's request appropriately. It should give appropriate guidance to user in both **informational and operational** ways.
- The **interface** must be clean and useful. It should gives user easy access to the service.
- The system must distribute and **map** the request to the correct agent who is most useful in a certain service and is able to merge and **reduce** the response from many agents into one organized response. The choice of agent must be optimized to ensure the best holistic accuracy.

1.5 External Subsystem: Motion Control Subsystem (UAV)

The UAV model used in our project is the MFP450, a medium-sized drone platform with a 410mm wheelbase. It is equipped with a Pixhawk 6C open-source flight controller, M8N-GPS, brushless motors, custom hard-shell batteries, Minihomer telemetry, an integrated optical flow ranging module, camera, and other devices. This UAV meets the requirements for stable flight both indoors and outdoors, and it is suitable for various applications including teaching and development. This is an off-the-shelf open source UAV, so we won't go in depth here.

2 Design

2.1 Block Diagram

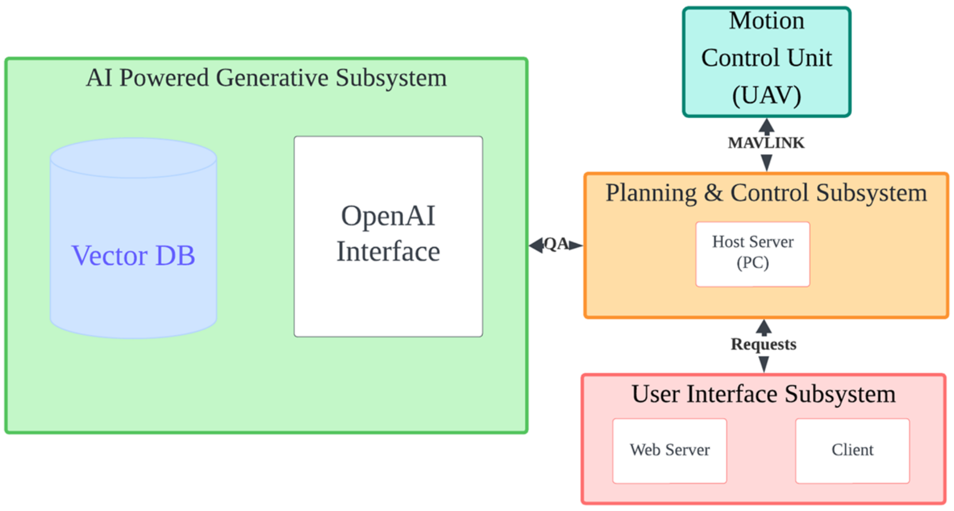
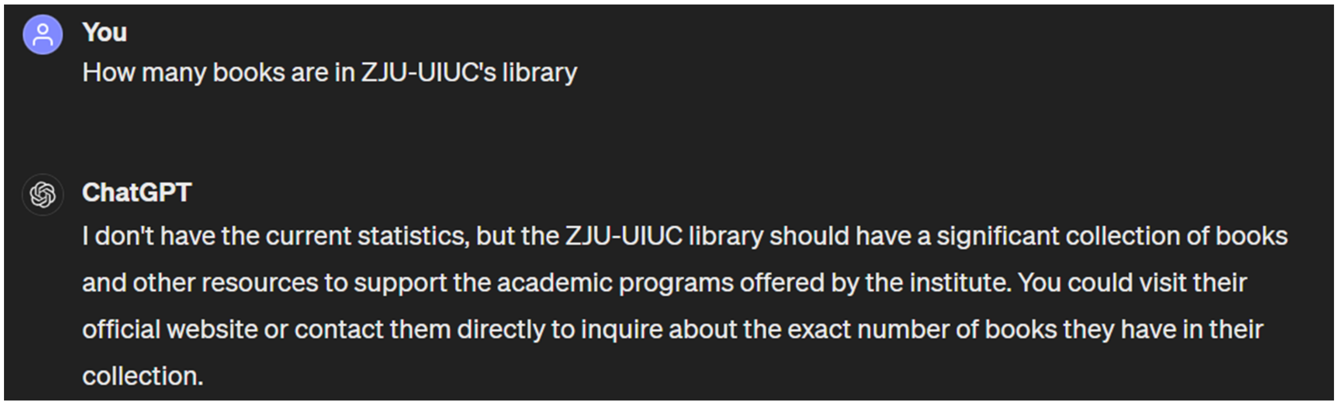


Figure 2: Block Diagram

2.2 Design Motivation

Information is unavailable and unorganized

Data are ubiquitous and exist in various formats, including multimedia. Efficiently collecting data for an entire campus is labor-intensive. However, the greater challenge lies in effectively utilizing this multimedia data and retrieving the most relevant information from an enormous dataset.



You
How many books are in ZJU-UIUC's library

ChatGPT
I don't have the current statistics, but the ZJU-UIUC library should have a significant collection of books and other resources to support the academic programs offered by the institute. You could visit their official website or contact them directly to inquire about the exact number of books they have in their collection.

Figure 3: Problems emerged when queries are associated campus related data

Figure 3 illustrates the problem that may arise when a user queries campus-related information. Due to the relatively small amount of relevant data compared to the total dataset size, the agent struggles to retrieve the corresponding data based on the user's intent.

Data path is very long for each query

Unlike applications hosted on a single device, our system is distributed across multiple devices, including two personal computers (PCs) acting as host servers, one web server, a mobile phone, a UAV, and several cloud services hosted by OpenAI. This system spans different countries and continents.

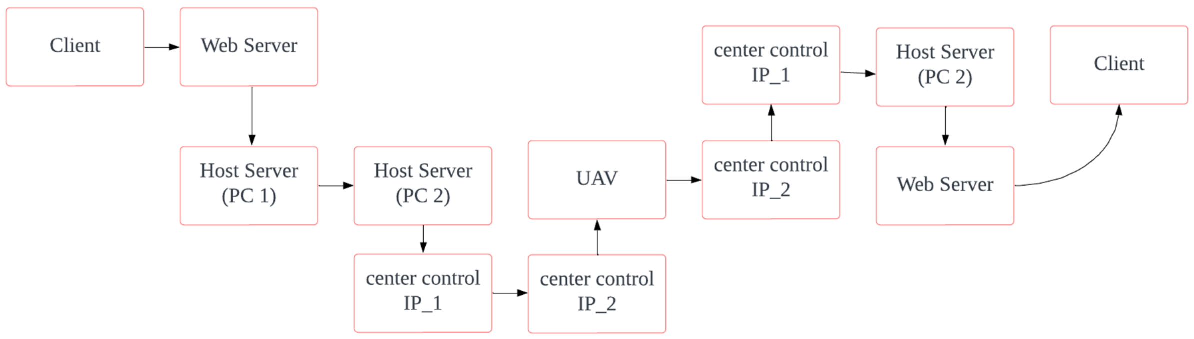
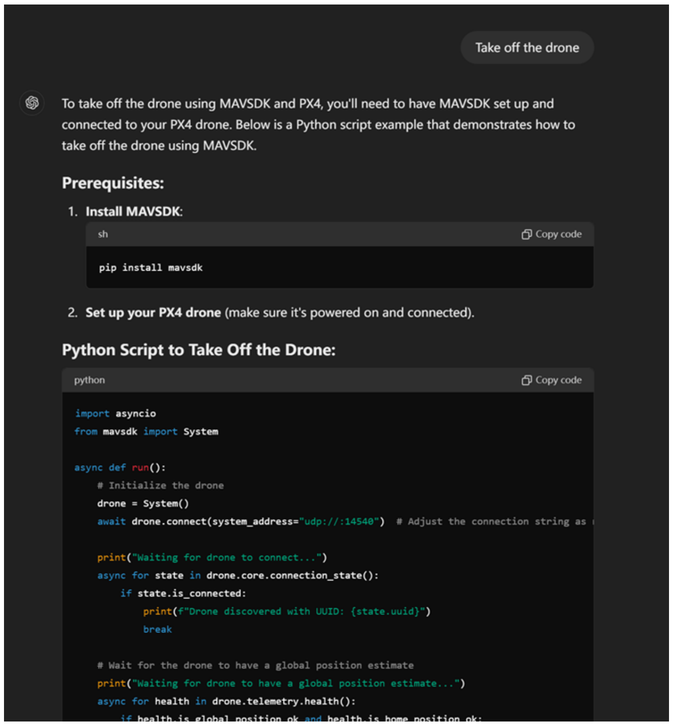


Figure 4: Long datapath for each user's query

The long datapath shown in Figure 4 suggests that it is beneficial to group the devices and abstract the services behind them as subsystems. Based on this idea, we grouped the service managing the web server and mobile phone as the user interface subsystem. We grouped the service managing the UAV and the host server that communicates with it as the planning and control unit. We also grouped the service hosted on OpenAI and the host server that communicates with it as the AI-powered response generation subsystem.

Huge gap between informational and operational

Even with proper grounding, where the OpenAI agent is provided with relevant information and its intended domain of expertise, the generative responses produced by the LLM can be highly unstable and unpredictable.



Take off the drone

To take off the drone using MAVSDK and PX4, you'll need to have MAVSDK set up and connected to your PX4 drone. Below is a Python script example that demonstrates how to take off the drone using MAVSDK.

Prerequisites:

1. **Install MAVSDK:**

```sh
pip install mavsdk
```

2. **Set up your PX4 drone** (make sure it's powered on and connected).

Python Script to Take Off the Drone:

```python
import asyncio
from mavsdk import System

async def run():
    # Initialize the drone
    drone = System()
    await drone.connect(system_address="udp://:14540")  # Adjust the connection string as

    print("Waiting for drone to connect...")
    async for state in drone.core.connection_state():
        if state.is_connected:
            print(f"Drone discovered with UUID: {state.uuid}")
            break

    # Wait for the drone to have a global position estimate
    print("Waiting for drone to have a global position estimate...")
    async for health in drone.telemetry.health():
        if health.is_global_position_ok and health.is_home_position_ok:
```

Figure 5: Problems when user types in the command

Another significant observation from our initial testing is that there is a 2 out of 25 chance that GPT could generate a malicious response potentially harmful to the UAV. This requires careful grounding, meaning we must establish a set of restrictions on communication between AI agents and the UAV.

Complete paradigm change needed

Our group proposes a multi-agent network, focusing on generalization while supporting automation and visualization. Imagine a system built for a single-CPU game: with more resources, a person can add another CPU, but the system must be entirely rebuilt because it cannot distribute work across multiple CPUs. The same logic applies to our system. Imagine a system using a UAV to guide users and answer questions about 10 locations. This system would need a complete overhaul if expanded to 10 UAVs or required to answer questions about 1000 locations, as it is designed for 1 UAV and 10 locations. It cannot manage large datasets or multiple UAVs. Our system, however, can handle arbitrarily large amounts of knowledge with a sufficiently large database and can expand to support any number of operational agents, given the resources to build them.

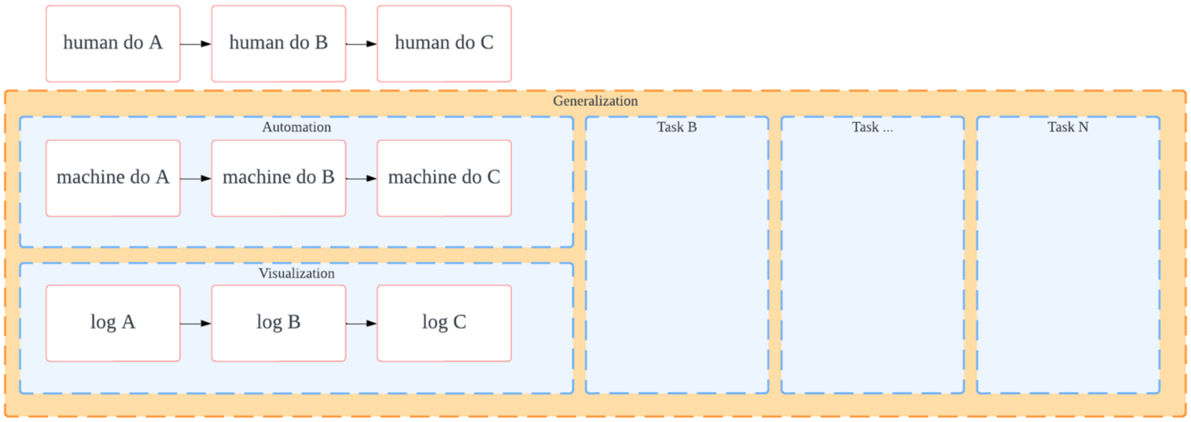


Figure 6: Paradigm change

2.3 AI-powered response generation subsystem

The AI-powered response generation subsystem focuses on building an assistant to respond to user queries about the ZJU-UIUC campus. The user's query can be a question about the campus or a command to the drone. The text generation and embedding modules are powered by OpenAI [1]. The subsystem acts as the brain of the campus tour guide. It consists of different AI-powered **agents** that help the system generate responses and interact with users.

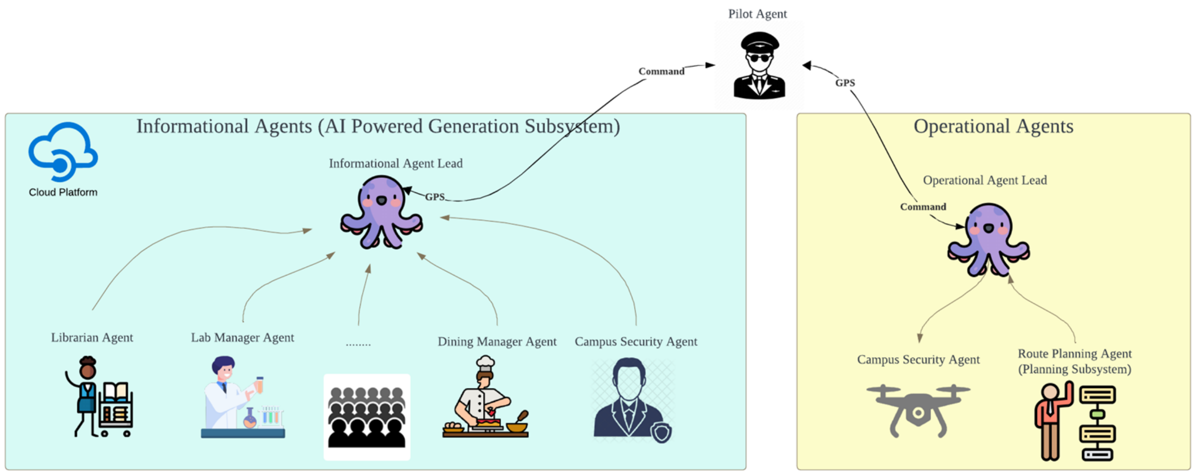
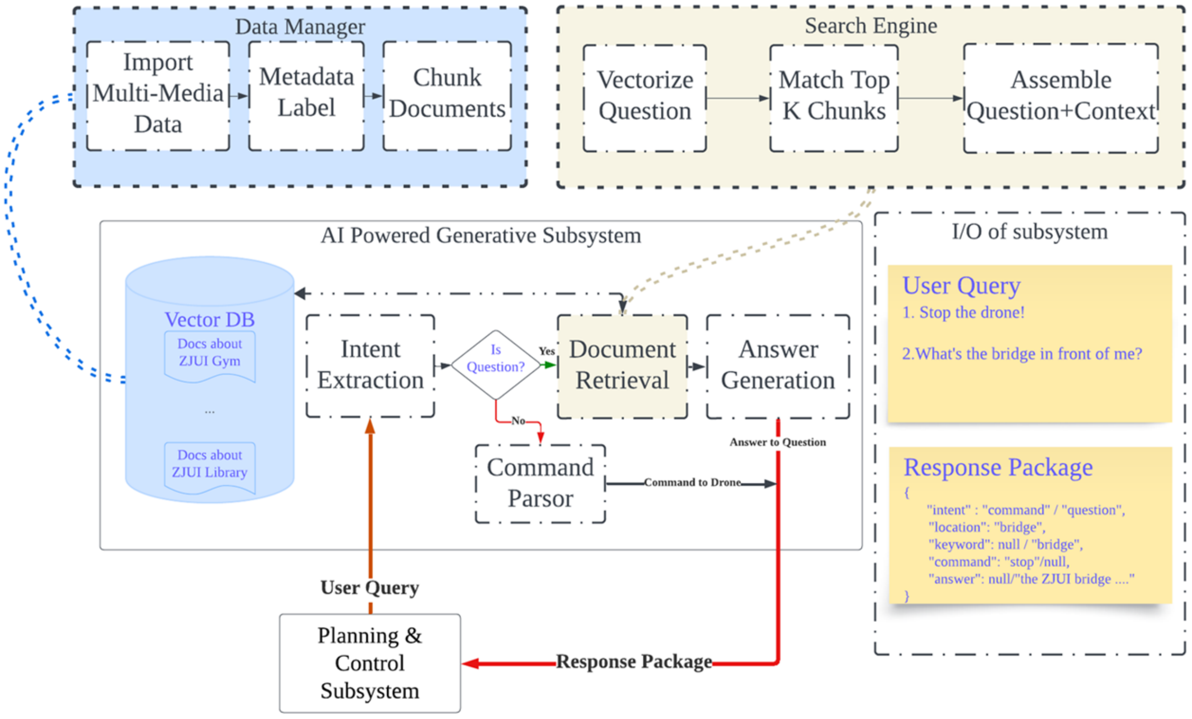


Figure 7: Multi-agent communication architecture

2.3.1 System Architecture and Design Overview

The system is architected to efficiently handle two distinct types of queries: **informational** queries and **operational** queries. The architecture of each agent is modular, with each module specializing in tasks such as intent detection, data retrieval, and response generation and validation. This modular design facilitates scalability and maintenance while ensuring the system can evolve to incorporate future enhancements or functionalities. The single agent system specializes in one specific task, and together they compose the AI-agents network. Many tasks require the cooperation and communication between different agents, as shown in Figure 7.



1. Stop the drone!

2.What's the bridge in front of me?

```
{
    "intent" : "command" / "question",
    "location": "bridge",
    "keyword": null / "bridge",
    "command": "stop"/null,
    "answer": null/"the ZJUI bridge ...."
}
```

Figure 8: Single-agent architecture

The single-agent architecture can be broken into the following parts:

- Vector Database(DB) storing related campus material
- Intent Extraction module extracting user's intent.
- Search Engine module search for related entries in Vector DB
- Answer Generation

To build each agent, we have two sides of tasks:

- Data side tasks: tasks related to multi-media data collection and processing.
- Agent side tasks: tasks related to response generation process.

The data side tasks consist of various work items. First, one must spend times to mine insights from data. This data mining process is crucial to determine the methods used to process data. For instance, If the data for each location is not too long, then we can categorize data by location and handle queries related to each location separately. Based on the insights gained from the multi-media data, we can handle different types of data with different methods. Once we finalize the methods, we can scale up the data size and systematically collect and process data. Hence, the third work item is multi-media data gathering, cleaning and tagging. Eventually when the data gathering is completed, we need to present statistics in clear and organized ways.

The agent side also has many components in its pipeline. The first module of the pipeline is intent identification. It identifies and categorizes user's intent. Then we have a search engine module that extracts the corresponding data which is processed by tremendous data-side effort. Then we have a answer generation module which references the data extracted from search engine, generating a ungrounded response. By ungrounded, we mean the response generated by answer generation module can be nonideal, meaning it can be too long, too dangerous, inaccurate. Hence, we need a protection-unit module to ground the output generated by answer generation module. These modules form a pipeline for the AI powered response generation agent.

The interface of the subsytem is listed as follows:

Table 1: Input and Output of the AI-Powered Response Generation Subsystem

Field Name	Type	Meaning
User Query	Input	Campus related query or a command to drone
User Location	Input	Current GPS location of user
Answer	Output	Answer to user's question
Command	Output	Parsed output Command to the drone

2.3.2 Data side effort: Multi-media Data Collection

The GPT model does not have the ability to answer questions related to ZJU-UIUC. It failed to answer the 100 testing questions completely. To supply pertinent information regarding to users' query, we need a set of data tailored for our application. The data we will use are in multiple forms as shown, it could be digital pdfs, or paper-based materials placed at many places at different locations within the campus. We utilize several methods to handle these different forms of information.

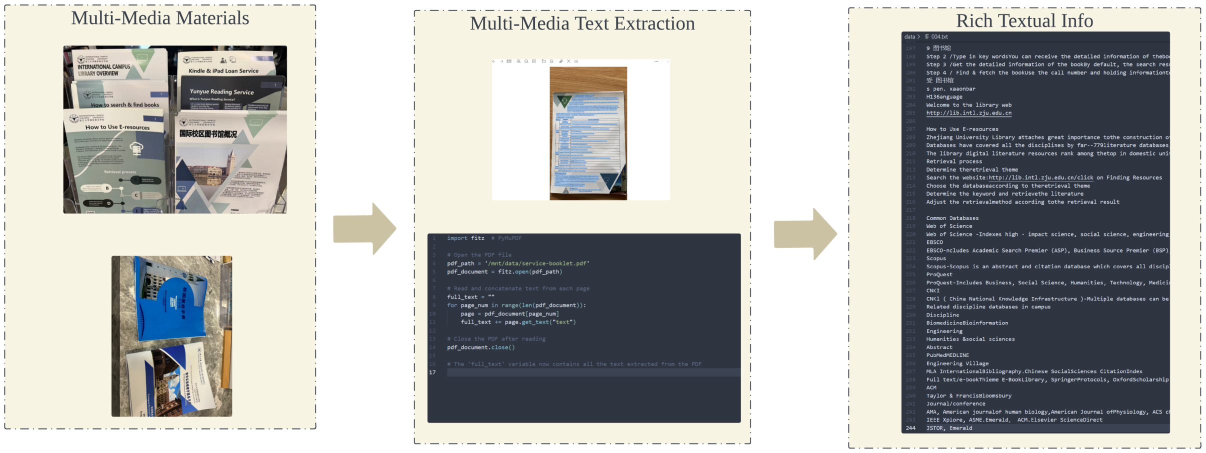
1. OCR tool backed by Wechat.
2. python tools to parse pdfs.

The code to parsing multi-media material is in this directory.

We convert these data-sources into uniform textual information so our agent can answer user's query with accurate and diverse information.

2.3.3 Agent side effort: Intent Identification and Retrieval

The intent identification module classifies users' input into three categories, as specified in the instruction section of the prompt. This module relies on a straightforward query to GPT



```python
import fitz  # PyMuPDF

# Open the PDF file
pdf_path = '/mnt/data/service-booklet.pdf'
pdf_document = fitz.open(pdf_path)

# Read and concatenate text from each page
full_text = ""
for page_num in range(len(pdf_document)):
    page = pdf_document[page_num]
    full_text += page.get_text("text")

# Close the PDF after reading
pdf_document.close()

# The 'full_text' variable now contains all the text extracted from the PDF

```

```
data > 004.txt
9 图书馆
Step 2 /Type in key wordsYou can receive the detailed information of thebo
Step 3 /Get the detailed information of the bookBy default, the search res
Step 4 / Find & fetch the bookUse the call number and holding informationt
受 图书馆
s pen. xaaonbar
H136anguage
Welcome to the library web
http://lib.intl.zju.edu.cn

How to Use E-resources
Zhejiang University Library attaches great importance tothe construction o
Databases have covered all the disciplines by far--779literature databases
The library digital literature resources rank among thetop in domestic uni
Retrieval process
Determine theretrieval theme
Search the website:http://lib.intl.zju.edu.cn/click on Finding Resources
Choose the databaseaccording to theretrieval theme
Determine the keyword and retrievethe literature
Adjust the retrievalmethod according tothe retrieval result

Common Databases
Web of Science
Web of Science -Indexes high - impact science, social science, engineering
EBSCO
EBSCO-ncludes Academic Search Premier (ASP), Business Source Premier (BSP)
Scopus
Scopus-Scopus is an abstract and citation database which covers all discip
ProQuest
ProQuest-Includes Business, Social Science, Humanities, Technology, Medici
CNKI
CNKl ( China National Knowledge Infrastructure )-Multiple databases can be
Related discipline databases in campus
Discipline
BiomedicineBioinformation
Engineering
Humanities &social sciences
Abstract
PubMedMEDLINE
Engineering Village
MLA InternationalBibliography.Chinese SocialSciences CitationIndex
Full text/e-bookThieme E-BookLibrary, SpringerProtocols, OxfordScholarship
ACM
Taylor & FrancisBloomsbury
Journal/conference
AMA, American journalof human biology,American Journal ofPhysiology, ACS c
IEEE Xplore, ASME.Emerald,  ACM.Elsevier ScienceDirect
JSTOR, Emerald
```

Figure 9: Workflow of Data Collection

3.5. However, given that GPT models typically achieve only about 70% accuracy on general multiple-choice tests [1], we must evaluate its performance on our specific task with caution. Enhancements might be achieved through refined prompt engineering [2].

The information retrieval module employs a Retrieval-Augmented-Generation approach. After extensive review [3], I opted for a sophisticated modular configuration. This agent integrates multimedia input processing, a search engine as the retriever, and a generation unit that processes and outputs the retrieved information.

The document retriever connects queries to relevant locations and retrieves the associated data. However, as data for each location increases, it may surpass GPT's token limit—the maximum number of tokens an LLM can process at one time. To manage this, the retriever breaks down data into manageable "chunks," summarizing each chunk and converting them into vectors using an OpenAI tool. The vector distance indicates the textual semantic similarity, which aids in identifying the most relevant chunks for a query.

2.3.4 Agent side effort: Protection Sub-Unit

AI agents suffer greatly from hallucination, which means the generation of plausible but incorrect or nonsensical information. This can be particularly problematic in scenarios requiring precise and accurate responses, such as UAV operations, where incorrect information can lead to dangerous or illegal actions.

To mitigate these risks, AI agents must be carefully grounded, especially on the operational side of the AI agent network. Grounding involves validating and verifying the generated outputs against predefined rules and contexts to ensure their accuracy and safety.

2.3.5 Goals and Verification

In order to verify the design, our system has the following 3 main target functionalities:

- Identify the intent of the input to the subsystem

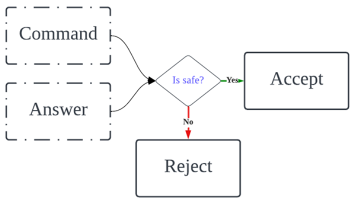


Figure 10: Simplified workflow of protection unit

- Fetch the correct external materials
- Output proper answer

In order to verify these functionalities, we design the following verification methods:

- The intent is evaluated automatically against the ground truth labeled by human on a testing dataset containing 100 sample questions.
- The retrieval accuracy will be evaluated automatically against ground truth labeled by human on a testing question set containing 30 human labeled questions on 4 testing locations.
- The end-to-end answer accuracy will be evaluated on the same dataset as retrieval accuracy. This evaluation will be conducted by human, giving a score of 5 and a comments pointing out potential issue. The comments are visualized and analyzed through word cloud.

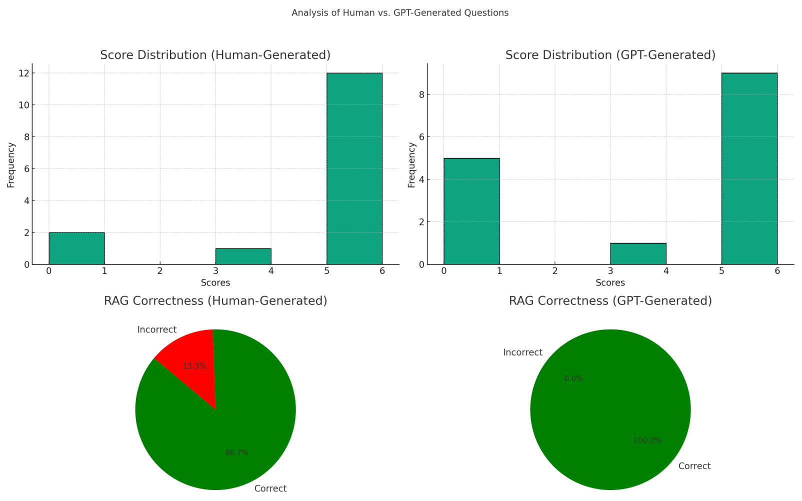


Figure 11: Analysis of End to End Accuracy as well as RAG accuracy

The code is available at project_repos

2.4 User Interface

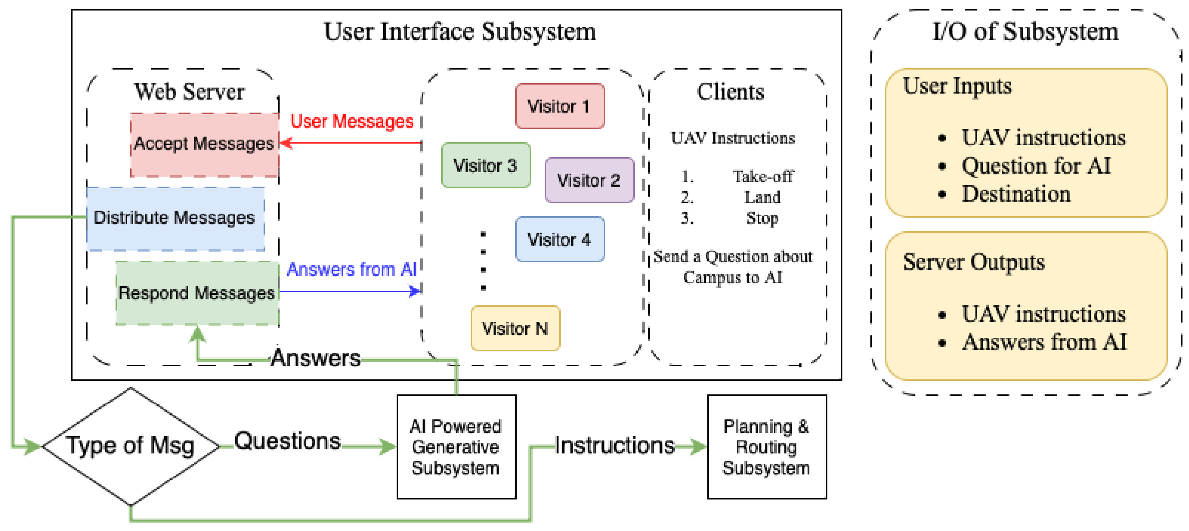


Figure 12: User Interface Diagram

The User Interface (UI) subsystem serves as the primary point of interaction between the users and the AI-powered response generation system. It is designed to be intuitive and user-friendly, enabling users to easily submit queries about the ZJU-UIUC campus and issue commands to the UAV.

2.4.1 Subsystem Architecture and Design

The User Interface (UI) subsystem is the principal conduit for user interactions within the AI-powered response generation system. It is meticulously crafted to be both intuitive and user-friendly, enabling seamless submission of queries and UAV command issuance.

The subsystem is comprised of several key components:

- A web server that accepts and distributes messages.
- Client interfaces for visitors to interact with the system.

The architecture delineates the following operational flow:

1. The web server receives messages from various visitors through the user interface.
2. Depending on the message type, identified as a question or a command, the server routes the message to the respective subsystem for processing.

The design leverages a set of defined APIs to manage the interactions between the UI and other subsystems, promoting real-time processing and ensuring data consistency and reliability. The modular nature of the design allows for scalable and maintainable enhancements, critical for future integration and functionality expansion.

The envisioned deliverables include:

- Hosting the web service on a remote server.
- Developing an easy-to-navigate, full-stack framework.

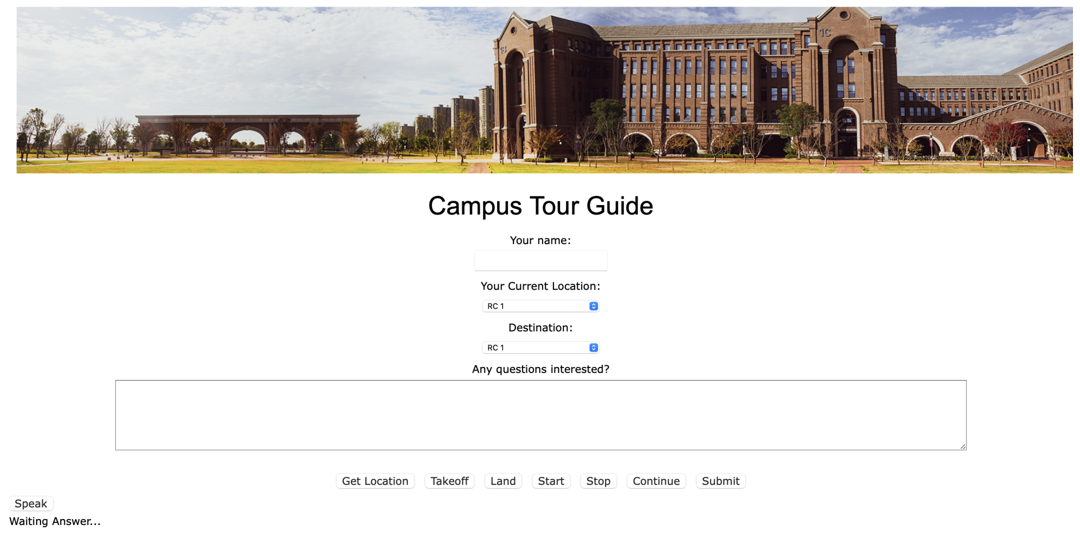


Figure 13: User Interface

- Establishing a robust connection between the web server and the AI and Planning subsystems.

This architecture is designed to provide a seamless, efficient, and secure user experience, whether it's for informational queries or operational control over the UAV.

2.4.2 Input and Output of the Subsystem

Table 2: Input and Output of the User Interface Subsystem

Field Name	Type	Meaning
UAV Instructions User2Server	Input	Take-off, Land, Stop instruction to UAV
User Questions	Input	Questions about ZJUI Campus
User Destination	Input	User's destination
AI Answer	Output	Answer to user's question
UAV Instructions Server2UAV	Output	Take-off, Land, Stop instruction to UAV

The input to the user interface subsystem is **the answer to the user's questions and the status of the drone**. The output of the subsystem is **questions by the user and commands to the drone**.

2.4.3 Frontend Development

The frontend of the UI subsystem is developed using React and JavaScript, offering a dynamic and responsive web interface. The design features a minimalist layout to enhance user experience and facilitate ease of use. Key elements of the UI include:

- **Question Input Block**: A dedicated area where users can type in or voice their queries about the ZJU-UIUC campus.
- **Instruction Buttons**: Several interactive buttons designed to issue predefined commands to the UAV, such as "Take off," "Land," and "Stop". Besides, there is an additional button to send questions to AI-powered Generative System "Send".
- **Campus Image Display**: An image block that dynamically displays photographs of the ZJUI campus, which could be relevant to the user's queries or for showcasing UAV functionalities.

This design ensures that users have a straightforward and efficient way to interact with the system, whether seeking information or controlling the UAV.

2.4.4 Remote Server Setup

The subsystem utilizes Ali Cloud for hosting the remote server, establishing a robust and scalable infrastructure. The connection to the server is secured via SSH, ensuring encrypted communication channels. This server plays a critical role in:

- Managing connections between end-users and the Planning & Control subsystem.
- Facilitating data exchange between the UI and AI subsystems, ensuring seamless integration and real-time response capabilities.

This integration is designed to be highly efficient, minimizing latency and maximizing the accuracy and relevance of the information provided to the users.

2.4.5 Verification and Results

The verification of the User Interface Subsystem involved a series of rigorous tests and evaluations to ensure its functionality, usability, and performance met the project requirements. The following methodologies were employed:

1. **Usability Testing:** Usability testing sessions were conducted with a diverse group of users to evaluate the UI's ease of use, intuitiveness, and effectiveness in facilitating interaction with the AI-powered Campus Tour Guide UAV.
2. **Performance Testing:** Performance tests were carried out to assess the responsiveness and reliability of the UI subsystem under various load conditions, ensuring it could handle multiple concurrent user requests without degradation in performance.
3. **Integration Testing:** Integration tests were performed to validate the seamless communication between the UI subsystem and other project modules, including the AI subsystem and the Planning & Control subsystem.

The problem at hand pertains to the fluctuating relationship between querying or pulling frequency and packet loss/data error within a network environment. In one scenario, as the frequency of queries increases, the occurrence of packet loss decreases. This suggests a correlation where heightened querying aids in error correction or enhances packet delivery, thereby mitigating loss. Conversely, in another scenario, an inverse relationship emerges, where elevated pulling frequencies correspond with escalated packet loss. Here, the increased frequency potentially triggers network congestion or system overload, heightening the probability of errors occurring. These contrasting patterns highlight the complexity of managing network performance and the need for nuanced strategies to optimize data transmission and minimize loss.

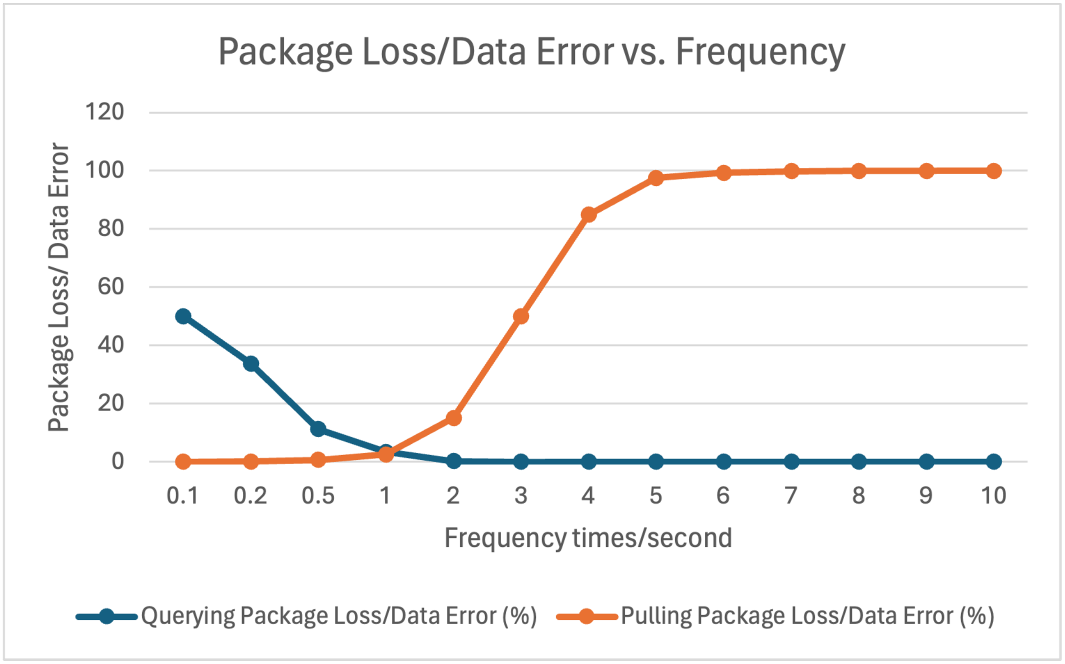


Figure 14: Package Loss & Data Error

The table illustrate the relationship between querying or pulling frequency and packet loss/data error. The first curve demonstrates a scenario where packet loss decreases as querying frequency increases. The second curve shows an inverse relationship where packet loss increases with higher pulling frequency.

Based on the observed relationships between querying or pulling frequency and packet loss/data error, we opt to set the frequency to once per second. This decision stems from a careful consideration of both scenarios outlined. In the first scenario, where packet loss decreases as querying frequency increases, a higher querying frequency undoubtedly enhances error correction mechanisms and ensures improved packet delivery, thus reducing loss. Conversely, in the second scenario, where higher pulling frequencies lead to increased packet loss, selecting a lower frequency mitigates the risk of network congestion and system overload, thereby decreasing the likelihood of errors.

2.5 Planning & Control Subsystem

The Planning and Control Subsystem is an integral component of our UAV operational framework, designed to process commands from the user interface, assess the current status of the drone, and issue precise navigational instructions based on the drone's specifications. This subsystem interfaces directly with PX4 APIs, an open-source flight control software, to monitor and control the UAV's flight parameters. The primary input to this subsystem is **the user's**

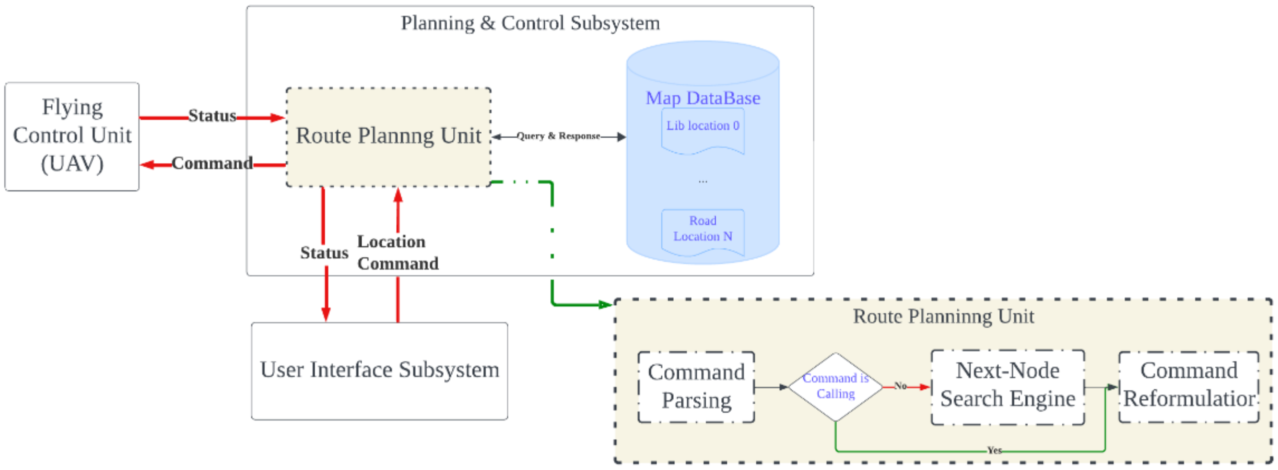


Figure 15: Planning and Control Block Diagram

command and the drone's status, while its output is the **command to the UAV**, ensuring that each operation is executed safely and efficiently.

2.5.1 Notation and Explanation

Table 3: Notations Used for Planning and Control Subsystem

Name	Meaning
Node	The map is continuous, we extract a sets of locations as nodes
Map Database	A data-store for the node
Next Node	The next node we plan to go to
Link	The minimum route unit linking 2 nodes
Command	An instruction to UAV or from user
Parsed Command	An instruction equivalent to a set of MAVSDK APIs
MAVSDK	a software dev kit consists of APIS instructing UAV
Search Engine	A module searching for next node given current location and command
Reformulation	A process formulating instruction to comply with MAVSDK APIs

2.5.2 Subsystem Architecture and Design

This subsystem obtains a command and it reformulates the command to the drone to execute. The reason why this module is essential and vital is this module helps visitor has a safe and comfortable trip. This subsystem objective is to find a short and comfortable route for user to take while taking the tour inside ZJU-UIUC campus.

This subsystem establishes a robust connection with a remote server to access vital flight data, including starting points, destinations, flight altitudes, and other navigational parameters. This connection is crucial for retrieving real-time information necessary for flight planning and control. The communication between the Planning & Control Subsystem and the remote server is facilitated through a secure, encrypted channel, ensuring the integrity and confidentiality of transmitted data. This setup allows the subsystem to:

- Obtain real-time updates on weather conditions, no-fly zones, and other environmental factors that may affect flight paths.
- Receive user-defined flight parameters such as starting location, destination, and preferred flight height.
- Update the UAV's mission parameters in response to changing conditions or user commands.

2.5.3 Input and Output of the Subsystem

Table 4: **Inputs and Outputs**

Inputs	**Outputs**
From the Unmanned Aerial Vehicle (UAV): • Latitude and longitude position of the UAV. • Current velocity, acceleration, and attitude (orientation) of the UAV.	**To the Unmanned Aerial Vehicle (UAV):** • Attitude control commands for the UAV. • Velocity and acceleration control commands for the UAV.
From the AI-Powered Generative System: • Responses generated by the AI system based on user queries.	**To the AI-Powered Generative System:** • User queries directed to the AI system.
From the User Interface System: • User commands and requests submitted through the interface.	**To the User Interface System:** • Responses provided to users based on their queries.

2.5.4 Data Collection and Tagging

We used an n*n adjacency matrix to represent the entire map, where n is the number of locations in the map. A link is a pair of nodes. A link is a viable path. Any connections between nodes that are not linked are not a valid path. For instance, if the connection between 2 nodes cross a lake which visitor is impossible to follow, it won't be our data store.

2.5.5 Algorithms for Subsystem

Let's define the mathematical model for the planning and control subsystem:

- **Nodes** N: A set of extracted locations from the continuous map, which serve as possible waypoints for the UAV.
- **Map Database** D: A datastore that contains the nodes and links information.
- **Links** L: Directed edges between nodes representing the minimum navigable path for the UAV. Each link connects two nodes and has associated costs (like distance, time, or energy consumption).
- **Commands** C: Instructions from the user or system that need to be executed by the UAV.
- **Parsed Commands** P: Translated commands into MAVSDK API calls.
- **MAVSDK** M: The software development kit used to control the UAV.

Given:

- **Current Node** $n_{\textbf{current}}$: The UAV's current position represented as a node.
- **Destination Node** $n_{\textbf{destination}}$: The target position the UAV needs to reach.

Objective

Find the optimal path Π from n_{current} to $n_{\text{destination}}$ minimizing the cost function F, which could include factors like distance, time, energy, etc.

Constraints

- The UAV can only travel along links in L from one node to another.
- The path must start near n_{current}

BFS Algorithm for Pathfinding

BFS algorithm can be used to find the path from n_{current} to $n_{\text{destination}}$ in a graph represented by nodes and links.

1. Initialize a queue Q and enqueue the starting node n_{start}.
2. Create a set S to track visited nodes and add n_{start} to S.
3. While Q is not empty:
   - Dequeue the front node n from Q.
   - If n is the destination node $n_{\text{destination}}$, terminate the search and process the result as needed.
   - For each neighbor n_{next} of n:
     - If n_{next} has not been visited (not in S):
       * Enqueue n_{next} into Q.
       * Add n_{next} to S.

The result of BFS algorithm is the path Π that optimizes the cost function F, providing an efficient route for the UAV from the starting point to the destination.

Implementation with MAVLink Python Package: The MAVLink Python package provides a comprehensive set of tools for communicating with the UAV, offering functionalities such as:

- Sending navigational commands and mission updates to the UAV.
- Receiving real-time status information, including location, battery level, and flight mode.
- Managing telemetry data to monitor and adjust flight parameters as needed.

This package is instrumental in bridging the gap between high-level operational commands and the low-level directives understood by the UAV, ensuring that the subsystem can effectively translate user intentions into actionable flight paths.

2.5.6 Verification and Results

The verification of the Planning & Control Subsystem involved a comprehensive evaluation of its capabilities in interpreting flight missions, generating optimal routes, and controlling the UAV during flight operations.

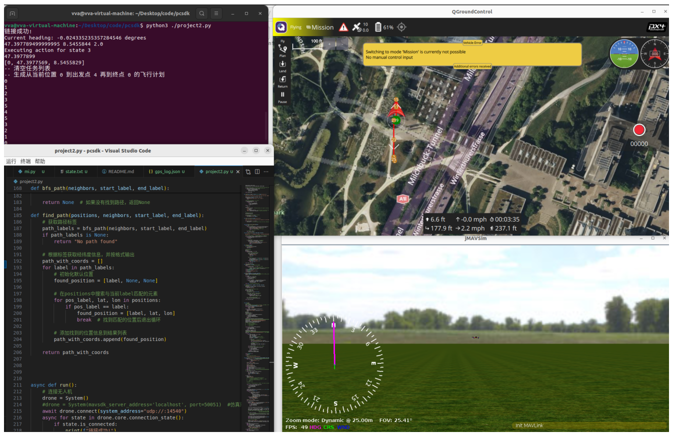
The methods and results of the verification process for the Planning & Control Subsystem are summarized as follows:

1. **Mission Interpretation Testing:** The subsystem successfully interpreted various flight missions specified in text files, accurately translating waypoints and other parameters into executable commands for the UAV.
2. **Route Generation Testing:** BFS algorithm consistently generated optimal flight routes, avoiding obstacles and adhering to operational constraints. The subsystem demonstrated reliability in finding efficient paths even in complex environments.
3. **Flight Control Testing:** Flight control tests confirmed the subsystem's capability to execute flight missions autonomously, including takeoff, landing, and waypoint navigation. MAVSDK scripts facilitated smooth operation and precise control of the UAV.
4. **Integration Testing:** Integration tests revealed seamless communication between the Planning & Control Subsystem and other project modules. The subsystem effectively received user-defined destinations from the UI subsystem, integrated AI-generated insights, and utilized real-time environmental data from the Sensing subsystem for adaptive flight planning.

Current Location	**Start Location**	**Destination**	**Route**
0	4	0	0-1-2-3-5-4-2-1-0
2	7	1	2-3-5-6-7-6-5-3-2-1
7	4	1	7-6-5-4-5-3-2-1

Table 5: Table of Locations and Routes

After inspection, the output of the path is the same as the output we expected, which matches the map we designed



```
vva@vva-virtual-machine:~/Desktop/code/pcsdk$ python3 ./project2.py
链接成功!
Current heading: -0.024335235357284546 degrees
47.397789499999995 8.5455844 2.0
Executing action for state 3
47.3977899
[0, 47.3977569, 8.5455829]
-- 清空任务列表
-- 生成从当前位置 0 到出发点 4 再到终点 0 的飞行计划
0
1
2
3
5
4
5
3
2
1
0
```

```python
def bfs_path(neighbors, start_label, end_label):

    return None  # 如果没有找到路径，返回None

def find_path(positions, neighbors, start_label, end_label):
    # 获取路径标签
    path_labels = bfs_path(neighbors, start_label, end_label)
    if path_labels is None:
        return "No path found"

    # 根据标签获取经纬度信息，并按格式输出
    path_with_coords = []
    for label in path_labels:
        # 初始化默认位置
        found_position = [label, None, None]

        # 在positions中搜索与当前label匹配的元素
        for pos_label, lat, lon in positions:
            if pos_label == label:
                found_position = [label, lat, lon]
                break  # 找到匹配的位置后退出循环

        # 添加找到的位置信息到结果列表
        path_with_coords.append(found_position)

    return path_with_coords



async def run():
    # 连接无人机
    drone = System()
    #drone = System(mavsdk_server_address='localhost', port=50051)  #仿真
    await drone.connect(system_address="udp://:14540")
    async for state in drone.core.connection_state():
        if state.is_connected:
            print(f"链接成功!")
```

Figure 16: UAV Flight mission by Planning & Routing Subsystem

2.5.7 Verification Table For Planning and Control Subsystem

The following is a table of validation items and results that we set for the subsystem:

Verification Item	**Simulation Results**	**Field Test Results**
Terminal control for takeoff and landing	Pass	Pass
Directional and distance flight	Pass	Pass
GPS coordinate-based flight	Pass	Pass
Route planning flight	Pass	Pass
Control flight using state.txt file	Pass	Pass
UI controlled flight	Pass	Pass
Asynchronous monitoring and alerting	Pass	Pass
Asynchronous monitoring for hovering at starting point	Pass	Pass

Table 6: Verification Items and Results

2.6 Sensor System

This section provides an overview of the sensor subsystem, detailing the specific sensors used for measuring humidity, temperature, and angular rate. Each sensor's type, model, and primary functionality are outlined to give a clear understanding of how they contribute to the overall system.

2.6.1 Introduction

We want to connect a temperature and humidity sensor to a PCB board, powered by a drone, with the sensor's temperature and humidity measurements displayed on an external LCD screen. This way, as the drone follows the user, the user can view real-time weather conditions.

Here is the schematic diagram of our overall sensor section. From the diagram, we can see that the entire circuit is controlled by the STM chip. Firstly, it converts the 5V regulated power supply from the aircraft input to a 3.3V voltage. Then, it connects to the temperature and humidity sensors as well as the speed sensor to obtain data. Subsequently, the acquired data is transmitted to the LCD display screen by the microcontroller for displaying the results.

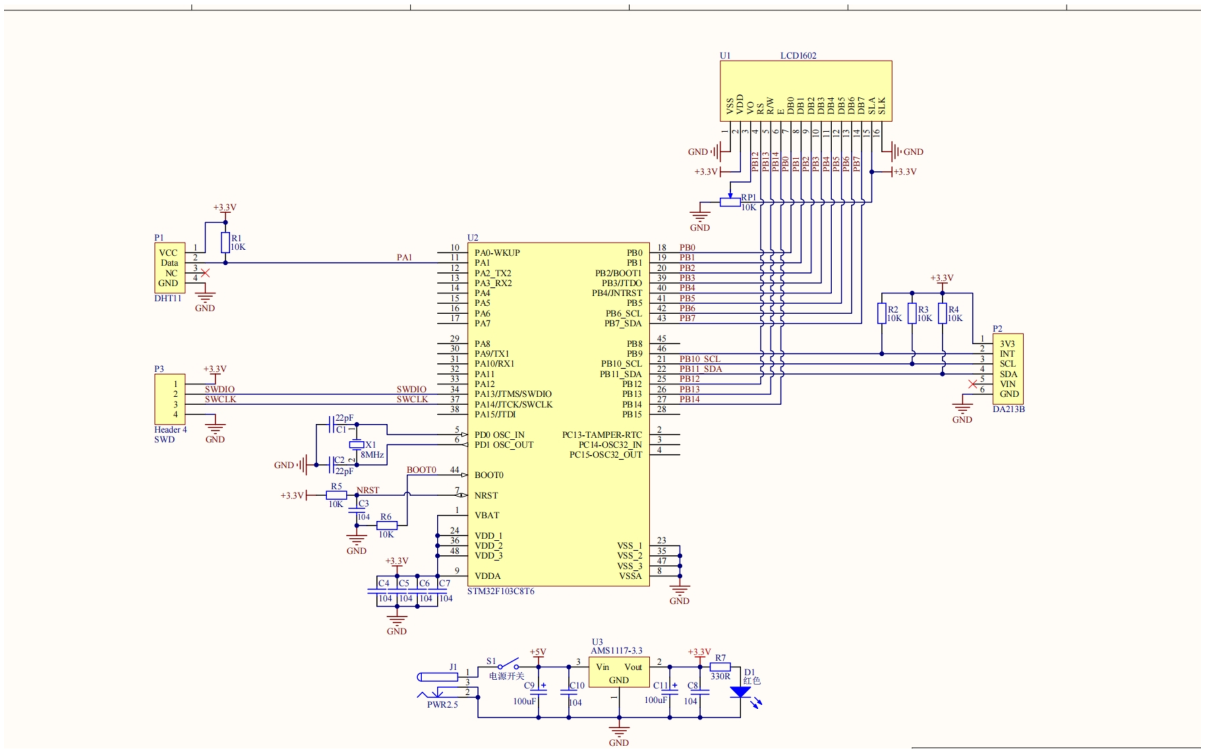


Figure 17: Circuit Design

2.6.2 Temperature and Humidity Sensor

- **Type & Model:** DHT-11

- **Functionality:** We selected the DHT11 sensor mainly because of its simplicity, ease of use, and low cost, as well as its ability to accurately measure environmental temperature and humidity. Its digital output makes data reading and processing more convenient, and it offers high accuracy and stability, providing reliable temperature and humidity measurements in various environmental conditions to ensure flight safety and stability for our project.[4]

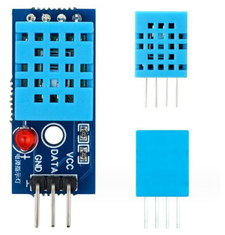


Figure 18: DHT 11 Outside view

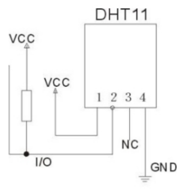


Figure 19: DHT11 pin diagram

2.6.3 Accelerometer Sensor

- **Type & Model:** DA213B

- **Functionality:** The DA213B accelerometer sensor was chosen primarily for its high precision, sensitivity, and stability. Capable of accurately measuring acceleration in three axial directions with digital output, it offers convenient data reading and processing. Additionally, its low power consumption and small footprint make it suitable for embedded systems and mobile device applications. In our project, the DA213B sensor provides precise acceleration data for attitude control and motion tracking, enhancing flight stability and accuracy.[5]

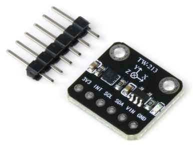


Figure 20: DA213B Outside view

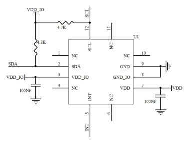


Figure 21: DA213B pin diagram

2.6.4 LCD Display

- **Type & Model:** LCD1602

- **Functionality:** The LCD1602 screen was selected primarily for its simplicity and affordability. It can display 16 columns and 2 rows of characters, providing basic text display functionality. Utilizing the standard HD44780 controller, it communicates easily with microcontrollers through a simple interface. Moreover, the LCD1602 screen has low power consumption, making it suitable for embedded systems and mobile devices. In our project, it provides a convenient way to display sensor data, system status, and user interface elements, simplifying user interaction and making system operation more intuitive.

Figure 22: LCD1602 Outside view

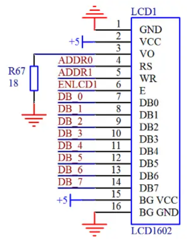


Figure 23: LCD1602 pin diagram

2.6.5 Verification and Requirement

There are several functionalities we want to achieve for the sensor unit:

1. The stability of converting 1.5V voltage to 3V voltage.
2. Whether the two sensors can operate properly, achieve data measurement, and transmit data.
3. Whether the LCD screen can display normally and show the data we need in real-time.

We verified these goals using the following verification methods.

1. Utilize a stable power supply to provide a 1.5V input voltage, connecting it to the voltage conversion circuit. Measure the output voltage using an oscilloscope or digital multimeter and record its value. Continuously observe the stability of the output voltage over a period of time (e.g., 30 minutes to 1 hour), noting any fluctuation range. Repeat the test steps to check the stability of the circuit under different temperature and load conditions.

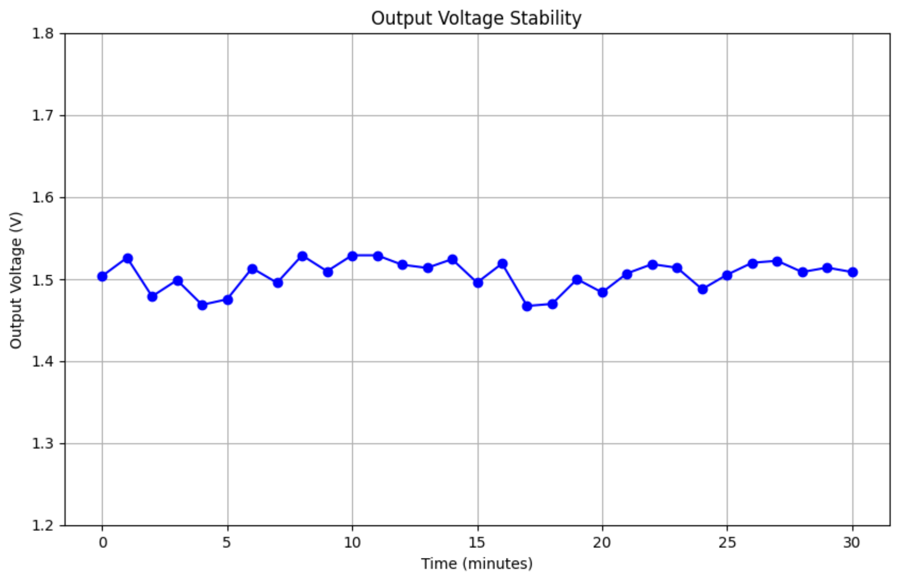


Figure 24: Voltage Test

2. Connect the sensors to the test board and use an oscilloscope or microcontroller to read the sensor's output data. Simulate sensor operation under different environmental conditions, such as varying temperature, humidity, and light levels, and observe the sensor's response. Verify whether the sensor output data matches the expected values and record any anomalies. Test the data transmission functionality to ensure that sensor data can be successfully transmitted to the microcontroller or other devices and correctly interpreted.

Environmental Condition	**Temp. (°C)**	**Humidity (%)**	**Output T(°C)**	**Output Humidity (%)**
Indoor Environment	25	50	24.8	51
Indoor Environment	25	50	24.9	50
High Temp.and Humidity	40	80	41.2	79
High Temp.and Humidity	40	80	40.9	81

Table 7: Test Data for DHT11 Sensor

Speed (m/s)	Result (m/s)
0.1	0.09
0.5	0.48
1.0	1.05
2.0	2.02

Table 8: Accelerometer Performance Test

3. Connect the LCD screen to the test board and provide appropriate power. Send test data to the LCD screen and observe whether the screen displays correctly. Check whether the LCD screen can display various types of data, including text, numbers, and graphics. Under simulated working conditions such as temperature changes and vibration, verify whether the LCD screen can stably display data. If possible, conduct long-term testing to confirm the stability and durability of the LCD screen.

Test Scenario	Temperature (°C)	Display Content	Display Result
Outdoor Environment	28	Text: ″Hello World″	Displayed Correctly
Outdoor Environment	30	Numbers: ″12345″	Displayed Correctly
Outdoor Environment	25	Graphics	Displayed Correctly

Table 9: LCD Screen Test Data

2.7 Tolerance Analysis

Ensuring the robustness of the UAV system involves conducting a thorough tolerance analysis. This analysis focuses on identifying potential vulnerabilities within the system, assessing risks, and devising strategies to mitigate these risks.

2.7.1 Route Planning Stability

Route planning is susceptible to various environmental factors that can affect the UAV's ability to navigate effectively. These factors include wind, physical obstacles, and the presence of visitors within the campus.

Risks:

- Wind can significantly alter the UAV's course, leading to deviations from the planned route and potentially unsafe conditions.
- Obstacles such as trees and buildings may not only hinder the UAV's path but also pose a risk of collision.
- Visitors moving unpredictably through the campus can introduce dynamic variables, complicating the UAV's navigation and safety protocols.

Wind's impact on UAV navigation can be analyzed using vector mathematics, specifically the wind triangle theory. The ground speed vector ($\vec{V_g}$) of the UAV is the vector sum of its airspeed vector ($\vec{V_a}$), which is the speed and direction relative to the air, and the wind speed vector ($\vec{V_w}$), which represents the speed and direction of the wind. This relationship is given by:

$$\vec{V_g} = \vec{V_a} + \vec{V_w} \tag{1}$$

2.7.2 GPS Locating Error

Global Positioning System (GPS) technology is crucial for UAV navigation, offering real-time location data that guides the UAV's flight path. However, GPS signals can be subject to interference from environmental factors, such as atmospheric conditions, buildings, and signal jamming, leading to potential errors in location accuracy.

Risks:

- **Mission Failure:** Critical missions requiring precise location data, such as aerial photography or targeted delivery, could fail due to inaccurate positioning.

GPS locating error (E_{gps}) can be influenced by several factors, including signal propagation delay, atmospheric conditions, and multipath errors. The total error can be modeled as a combination of these factors:

$$E_{gps} = E_{propagation} + E_{atmospheric} + E_{receiver} \tag{2}$$

3 Cost and schedule

3.1 Cost Analysis

3.1.1 Labor

The labor cost is calculated based on the working hours and wage pricing of each team member. We set the hourly wage at 100 RMB based on market research and the skill levels of team members. Considering the total project duration of 8 weeks with 40 hours of work per week, the total working hours per team member are: 320 hours. Therefore, the labor cost per team member is:

$$100 \text{ RMB/hour} \times 320 \text{ hours} = 32000 \text{ RMB}$$

We chose an hourly wage of 100 RMB, which is based on market wage levels and the skill and experience levels of team members. According to survey data from the Institute of Electrical and Electronics Engineers (IEEE) [6], the average salary for graduates in Electrical and Computer Engineering (ECE) is around 200,000 RMB per year. Calculated on a full-time basis, the average hourly wage is approximately 100 RMB.

3.1.2 Parts

The table below provides a breakdown of the parts and their estimated costs:

Description	Manufacturer	Part #	Quantity	Cost (RMB)
Drone	PixHawk	MFP450	1	5174
Mavlink Module	Amovlab	-	1	680
Temperature Sensor	Aosong	DHT11	2	10
Acceleration Sensor	MiraMEMS	DA213B	2	15
LCD Screen	Touglesy	LCD1602	1	20
PCB Board	Custom	-	1	50
Simple Application Server	Alibaba Cloud	-	1	49 per month
ChatGPT4 API	OpenAI	-	1	240 per month

Table 10: Parts List and Estimated Costs

3.1.3 Grand Total

The grand total cost of the project can be calculated by summing up the labor cost and the cost of parts:

- Labor: 32,000 RMB
- Parts: 6,819 RMB

Grand Total: 32000 + 6819 = 38,819 RMB

4 Conclusion

4.1 Achievement

In this project, we implemented an AI-based campus tour guide assistant. We utilized a drone as the guiding platform to facilitate a seamless campus tour experience while providing informative responses to users' inquiries. The AI agent network system we developed prioritizes responsiveness, ensuring prompt and relevant answers to users' requests. This system not only offers guidance on campus exploration but also addresses user queries comprehensively.

Moreover, we emphasized the importance of a clean and user-friendly interface, enabling easy access to the service. A well-designed interface enhances user engagement and satisfaction, enhancing the overall tour experience.

Additionally, our system incorporates a sophisticated agent distribution and mapping mechanism. Requests are intelligently routed to the most suitable agent capable of addressing specific services. Furthermore, the system optimizes agent selection to ensure the highest accuracy in response aggregation. By merging and organizing responses from multiple agents, we deliver a cohesive and comprehensive user experience, enhancing the effectiveness of our AI-based campus tour guide assistant.

4.2 Uncertainties

Despite the well-structured design, there are still uncertainties that are left to explore for this project. For instance, we still doesn't fully tested the cases where the UAV is far away to the central control unit. We also didn't detect whether human follows the drone. This can be supported by a fusion with computer vision.

4.3 Future Work

Due to the assumptions we made in this project, and the uncertainties left to investigate, there are a series of future work to accomplish:

- Fuse with computer vision to detect the human presence, and even gesture to boost the human-UAV interaction
- Extend the distance UAV can receive signal by setting more than 1 central controls.

4.4 Ethical and Safety Considerations

The development and deployment of the AI-guided tour guide drone raise important ethical considerations that must be addressed. [7] Safety is important in the development and operation of the AI-guided tour guide drone. Several safety measures will be implemented to mitigate risks and ensure the well-being of users and developers. Firstly, the drone's hardware and software systems will undergo rigorous testing and validation to ensure their reliability and stability. Emergency shutdown protocols will be in place to address malfunctions or emergencies promptly. Secondly, strict battery safety protocols will be enforced to prevent accidents related to lithium polymer (LiPo) batteries. This includes regular inspection, proper storage, and careful handling to minimize the risk of fire or explosion.

References

[1] OpenAI, J. Achiam, S. Adler, *et al.*, *Gpt-4 technical report*, 2024. arXiv: 2303.08774 `[cs.CL]`.

[2] L. Weng, “Prompt engineering,” *lilianweng.github.io*, Mar. 2023. [Online]. Available: https://lilianweng.github.io/posts/2023-03-15-prompt-engineering/.

[3] P. Zhao, H. Zhang, Q. Yu, *et al.*, *Retrieval-augmented generation for ai-generated content: A survey*, 2024. arXiv: 2402.19473 `[cs.CV]`.

[4] Aosong Electronics Co., Ltd. “Aosong Product 21.” (2024), [Online]. Available: http://www.aosong.com/products-21.html (visited on 04/19/2024).

[5] MiraMEMS. “MiraMEMS Product 1.” (2024), [Online]. Available: http://www.miramems.com/product-1.html (visited on 04/19/2024).

[6] “IEEE (Institute of Electrical and Electronics Engineers) Salary Survey,” 2022.

[7] IEEE. “”IEEE Code of Ethics”.” (2016), [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html (visited on 02/08/2020).

Appendix A Requirement & Verification Table

Table 11: Subsytem Index Table

Subsystem Name	Subsystem Index(S-Index)
AI Powered Response Generation Subsystem	1
User Interface Subsystem	2
Planning and Control Subsystem	3
Sensor Unit Subsystem	4

Table 12: Requirement & Verification Table

S-Index	Requirement	Verification	Points
1	Detect user intent with at least 25% accuracy	Test with a diverse set of input queries and verify the accuracy of intent detection.	2
1	Detect user intent with at least 50% accuracy	Continue testing and refining to achieve higher accuracy.	2
1	Generate response within 60 seconds	Measure retrieval time with various queries to ensure performance within the initial time limit.	2
1	Generate response within 40 seconds	Measure retrieval time with various queries to ensure performance within the initial time limit.	2
1	Generate response within 30 seconds	Optimize system to improve performance and meet the final time requirement.	1
1	Can fetch correct external material with 20% accuracy	Testing the agent's RAG accuracy by using manually labeled dataset containing 30 questions for 4 testing locations.	1
1	Can fetch correct external material with 50% accuracy	Testing the agent's RAG accuracy by using manually labeled dataset containing 30 questions for 4 testing locations.	2

Continued on next page

Table 12 continued from previous page

S-Index	Requirement	Verification	Points
1	Can fetch correct external material with 70% accuracy	Testing the agent's RAG accuracy by using manually labeled dataset containing 30 questions for 4 testing locations.	1
1	Generated answers must match the user's intent with an accuracy of at least 25% in given dataset	Compare generated answers from the dataset containing 30 questions with human labeled answer to calculate initial accuracy.	3
1	Generated answers must match the user's intent with an accuracy of at least 50% in given dataset	Compare generated answers from the dataset containing 30 questions with human labeled answer to calculate initial accuracy.	3
2	The web server must handle and route messages with less than 2 seconds latency	Test message routing on the server under load and measure latency	2
2	The client interface must provide intuitive access for users to submit queries and control the UAV	Conduct usability testing with participants to assess ease of use and intuitiveness	1
2	UAV command buttons must send correct instructions to the UAV subsystem with 100% accuracy	Test each button and verify that the correct command is sent to the UAV subsystem	1
2	The web server must send instructions and questions separately to different hosts	Test two hosts if they receive correct messages	1
2	Obtain the user's GPS position as the starting position	Test if the two hosts receive the GPS signal	1
2	The host which process the questions can display answers correctly	Check if the UI can display reasonable answers	2
2	The host which process the instructions can send commands to UAV	UAV can take-off, Stop, Continue, Land correctly and in time	2
3	Must accurately process user commands and drone status within 10 second	Perform stress testing with simultaneous user commands and verify response time	3

Continued on next page

Table 12 continued from previous page

S-Index	Requirement	Verification	Points
3	Must accurately process user commands and drone status within 1 second	Perform stress testing with simultaneous user commands and verify response time	3
3	Must optimize the UAV route based on the current status	Test with different scenarios (no-fly zones, different areas) to verify route optimization	3
3	Should maintain a secure and encrypted connection to the remote server	Verify the encryption standards and conduct penetration testing to assess security	1
3	Must integrate seamlessly with the PX4 APIs for flight control	Execute a series of flight tests to ensure proper integration and control	1
4	Power supply successfully power the hardware unit	Power supply LED works correctly	1
4	Sensors can operate properly	Connect the sensors to the test board Use oscilloscope to read the output data.	2
4	LCD screen can display normally	Connect the LCD screen to the test board Provide appropriate power. Send test data to the LCD screen Check if the screen displays correctly.	2
1,2 3,4	End to End works correctly	Can complete a guide for appointed locations while iteracting with visitors	5