

CST3140 Novel Interaction Technologies-VUI App

Joomun Muhammad Muddathir
Ibney Noorani
Middlesex University Mauritius Student
MJ806@live.mdx.ac.uk

INTRODUCTION

This study details the creation and testing of a voice-controlled interface designed to help Computer Science students find university-related information easily. Using voice commands, students can interact with the system to get both spoken and visual responses. The initial design was quickly sketched out to try out various ideas before building a working model using the Web Speech API. The prototype was then tested with users to spot any issues and to figure out how to make it better. The focus was on practical improvements rather than just getting positive feedback. The VUI name shall be MDXChatBot

DESIGN

In this section we will discuss about the conceptual framework and specifics of the VUI system

Purpose of the VUI System

The VUI system is purpose-built to streamline the information acquisition process for Computer Science students at the university. By leveraging voice commands, the system provides directional guidance across campus, appointment scheduling, updated course fees information, and detailed course descriptions, all while focusing on campus-centric queries.

Anticipated Features

The system's anticipated features include:

- **Directional Guidance:** Real-time navigation to campus locations such as the library, administration block, and lecture theaters.
- **Appointment Scheduling:** Allow users to set up meetings and academic consultations.
- **Course Information:** Access to current courses, their fees, and detailed information through the OpenAI API configured to prioritize campus-related data.
- **Courses Fees:** Based on updated data stored on the server.

Design Best Practices and Guidelines

1. **Mode vs. Wake Word:** A click-on-mic interface was chosen over wake-word detection for user control and privacy, supported by literature that suggests explicit user actions can reduce unintended activations and enhance user trust [1].
2. **VUI and UI Integration:** The system design considers the balance between voice interactions and visual feedback,

ensuring that information is accessible, and actions are confirmed visually, aligning with established UI principles for multimodal systems [2].

Anticipated Features & Slot Filling

The system features structured conversation flows, known as intents and utterances. For example, to get details about courses, users might say "Tell me about the computer science course," to which the system will respond with relevant information including content, fees, and duration. Slot filling is used to gather additional information from the user in a conversational manner. If a user wishes to schedule an appointment but only provides a date, the system will ask for the time, ensuring all necessary details are captured.

Sketches

1. Landing Page

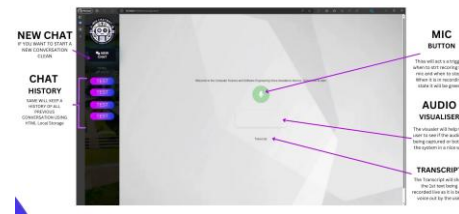


Figure 1 Landing Page

2. Chat based scenario.



Figure 2 Chat based scenario.

3. Map based scenario.

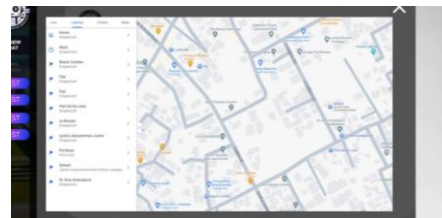


Figure 3 Map based scenario.

PROTOTYPE

Technical Architecture

The prototype was developed using Node.js with the Express framework to manage server-side operations, allowing for robust API integration and routing. The following APIs were integral to the prototype's functionality:

- **Google Maps API:** Facilitates real-time directional guidance within the campus, helping users navigate to various locations like the library, admin block, and lecture theaters.
- **Google Calendar API:** Supports the appointment scheduling feature, enabling students to book meetings directly through the VUI.
- **OpenAI API:** Utilized for answering complex queries related to course details, fees, and general information, ensuring the system remains contextually aware of the campus environment.
- **Web Speech API:** Utilized for enabling real-time voice recognition and speech synthesis, allowing users to interact with the system naturally.

Client-Side Implementation

On the client side, the Web Speech API was used for voice recognition and speech synthesis, allowing for a natural interaction mode where users can speak to the system and receive auditory feedback. JavaScript was used to manage client-side logic, including speech processing and UI interactions.

Voice Interaction Flow

Below is the diagram for voice interface module and the 4 steps describe how it was implemented in our prototype. We based our flow on research conducted on voice driven life assistant system for visually impaired people where Voice interaction was crucial for accessibility and design guidelines were followed.[3]

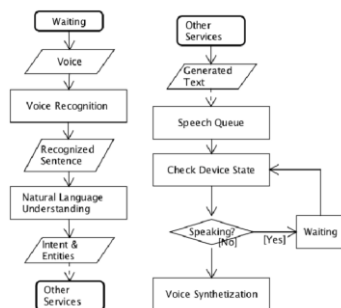


Figure 4 Flowchart diagram of voice interface module [3]

1. **Activation:** Users initiate voice interaction by clicking on the microphone icon, which starts the voice recognition process.
2. **Speech Processing:** Spoken commands are captured and sent to the server where they are processed to identify user

intent, whether it's asking for directions, inquiring about courses, or scheduling appointments.

3. **API Integration and Response:** The server handles the logic of interacting with the appropriate APIs based on the recognized intent and then sends back the information or confirmation to the client.
4. **Feedback:** Users receive spoken feedback through the Web Speech API's synthesis capabilities or visual indicators on the client interface, ensuring multimodal communication.

Visual and Interaction Design

The interface design includes a waveform visualization of the user's speech input, enhancing user engagement and providing visual feedback on the voice recognition process. Interactive elements like chat bubbles and modal dialogs for map displays and appointment scheduling contribute to an intuitive user experience.

Demonstration and Testing

A video demonstration of the prototype showcases its functionality, highlighting how different voice commands trigger specific system responses. The video includes examples of navigating the campus, scheduling appointments, inquiring about course fees, and asking general questions to illustrate the system's versatility and responsiveness. Please find the link

Video Demo - <https://youtu.be/K9YGMakxNKE>

Screenshots

1. General Chat

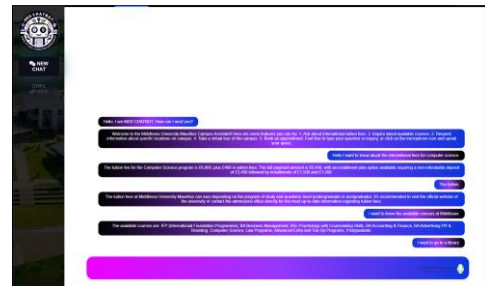


Figure 5 General Chat

2. Map Interaction

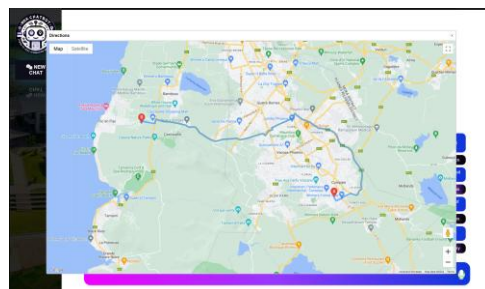


Figure 6 Maps interaction

3. Appointment Booking

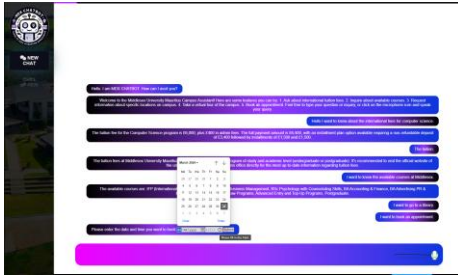


Figure 7 Appointment Booking

4.Appointment Confirmation

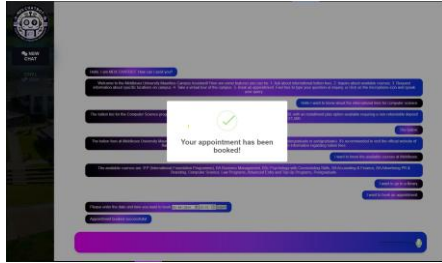


Figure 8 Appointment Confirmation

EVALUATION

The evaluation aims to address key questions to assess the usability and effectiveness of the VUI system:

- 1.How intuitively can users interact with the system to complete specific tasks?
2. What is the level of user satisfaction with the multimodal feedback provided?
- 3.Are there any notable navigational or interactional obstacles encountered by the users?

Evaluation Methods

A mixed-methods approach will be used, combining observational techniques, the System Usability Scale (SUS), and the User Experience Questionnaire (UEQ). These methods are chosen for their proven effectiveness in previous VUI studies, offering a comprehensive understanding of both quantitative usability metrics and qualitative user feedback. [7]

- 1.**Observational Techniques:** Real-time monitoring of users interacting with the system to identify any difficulties or bottlenecks in the user journey. [6]
- 2.**System Usability Scale (SUS):** A reliable, quick, and industry-standard tool to measure the usability of the system, providing a quantitative measure of user satisfaction. [4]
- 3.**User Experience Questionnaire (UEQ):** Assesses the broader aspects of user experience, including attractiveness,perspicuity,efficiency, dependability, stimulation, and novelty. [5]

Target Users

The evaluation will involve a diverse group of 5-10 Computer Science students to ensure a representative sampling of the user base. Participants will be selected randomly, ensuring anonymity and a broad range of user types, including undergraduates and international students.

Evaluation Plan

Participants will be invited via email and social media channels with clear instructions on the evaluation process. The setup will include a brief introduction to the VUI system, followed by a guided interaction session where participants perform a series of tasks using the system.

Data Collection and Analysis

Both quantitative and qualitative data will be collected:

1. Quantitative data from the SUS will be analyzed to provide an overall usability score.
2. Qualitative feedback from the UEQ and observational notes will be used to identify patterns in user interaction, preferences, and areas for improvement.
- 3.Data analysis will include statistical methods for SUS scores and thematic analysis for qualitative feedback, aiming to draw comprehensive insights into the system's performance and user experience.

Ethical Considerations

Participants will be fully informed about the purpose of the study and the nature of their involvement. Consent forms will be provided, ensuring that participants understand their rights to withdraw at any time and the confidentiality of their responses.

EVALUATION FINDINGS

The evaluation of the VUI system was conducted with 5 Computer Science students. Below we shall analyze the evaluations

Table of Results

User	Task Completion Rate	SUS Score	Positive Feedback	Negative Feedback
1	100%	85	Intuitive commands	Occasional Misunderstanding
2	80%	75	Clear Responses	Slow response time
3	60%	70	Helpful Guidance	Confusing Navigation prompts
4	90%	80	Accurate information	Limited error handling
5	70%	65	Easy to initiate	Inconsistent Response when saturated

Figure 9 Table of Results

Analysis of Positive and Negative Aspects

- 1.**Positive Feedback:** Users found the voice commands intuitive and the system's responses generally clear and accurate. The ease of initiating interactions was praised.
- 2.**Negative Feedback:** Common issues included occasional misunderstandings of user commands, slow response times, confusing navigation prompts, limited error handling, and inconsistent feedback.

Common Issues Identified

Multiple users reported slow response times and occasional misunderstandings, indicating a need for optimization in speech recognition and system performance.

Overall Findings

The system demonstrated potential in providing a useful interface for accessing university-related information. However, the user experience varied, with some facing challenges in navigation and interaction speed.

Challenges Found

1. Ensuring consistent and rapid processing of voice commands was a primary challenge.
2. Balancing between comprehensive responses and user attention span, particularly in navigational guidance.

CONCLUSION

The evaluation of the VUI system for Computer Science students highlighted its potential in simplifying access to university services through voice commands. Despite positive feedback on user-friendly interactions, issues like slow response times and occasional misunderstandings were noted. Addressing these challenges through system optimization and improved speech processing will be key to enhancing usability. Ultimately, refining the VUI system based on user feedback is essential for developing a more effective and intuitive tool for educational support.

REFERENCES

1. [1] C. Pearl, Designing Voice User Interfaces. "O'Reilly Media, Inc.," 2016.
2. [2] P. Romanowski and M. Guardado, The Many Faces of Multilingualism. Walter de Gruyter GmbH & Co KG, 2020.
3. [3] R. Chen, Z. Tian, H. Liu, F. Zhao, S. Zhang, and H. Liu, "Construction of a voice driven life assistant system for visually impaired people," 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), May 2018, doi: <https://doi.org/10.1109/icaibd.2018.8396172>.
4. [4] "Analisis Website STIMATA Menggunakan System Usability Scale (SUS)," Jurnal Ilmiah Komputasi, vol. 20, no. 3, Sep. 2021, doi: <https://doi.org/10.32409/jikstik.20.3.2776>.
5. [5] A. Marcus, Wentao Wang, and Springerlink (Online Service, Design, User Experience, and Usability. Practice and Case Studies : 8th International Conference, DUXU 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part IV. Cham: Springer International Publishing, 2019.
6. [6] J. S. Dumas and J. Redish, A practical guide to usability testing. Exeter, Angleterre ; Portland, Or.: Intellect, 1999.

7. [7] J. W. Creswell and V. L. Plano Clark, Designing and conducting mixed methods research, 3rd ed. Los Angeles: Sage, 2018.