COSC 4P02 Progress Report I

Chat Bot for Niagara on the Lake (NOTL) Museum: MuseumMate

Prepared by Maheen Samad, Fahad Arain, Robert Morabito

05/03/2023

Table of Contents

1. Team

1.1 Members

Below is a list of all the members a part of this team and their respective roles:

Maheen Samad: ***ScrumMaster and Product Owner***
6452270
Ms17ay@brocku.ca

Fahad Arain: ***Developer***
6770127
Fa19ra@brocku.ca

David Bailey: ***Developer***
6675482
Db18fv@brocku.ca

Robert Morabito: ***Developer***
7093230
Rm20mg@brocku.ca

Jordan Bharati: ***Front-end Developer***
6591556
Jb18vs@brocku.ca

Dana Dobrosavljevic: ***Developer***
7066855
Dd20eq@brocku.ca

Sarah Howcroft: ***Developer***
6381800
Sh17kq@brocku.ca

1.2 Scrum Meetings

Our team thoroughly understood that scrum meetings are an essential component of the scrum methodology; thus, we prioritized carrying out scrum meetings at least twice a week. Specifically, we have met every Tuesday and Thursday since January 12th, 2023 where we discuss our plans for the following working days, until the next meeting. We also discuss any issues we were currently facing, both as a group or individually. Scrum meetings have been an essential building block for our teams success thus far as it's robustness has enabled better communication, allowed us to stay updated on one another's progress, and sanctioned team cohesion and collaboration.
Below is a list of each scrum meeting date so far, as well as notes taken by the scrum master overviewing the purpose of each specific meeting:

Scrum Meeting #	Sprint	Date	Points of Discussion
1	0	Thurs Jan 12	• Getting to know the group. • Brainstorming ideas for the project. • Discussing scrum meeting schedule. • Finalizing topic for the project. • Deciding roles of each team member. • Planning schedule dates for each sprint. • Setting up GitHub page.
2	0	Tues Jan 17	• Editing and submitting progress report. • Choosing Jira as platform for storing progress. • Finalizing scope of sprint 1 using requirements and user stories. • Creating product backlog for sprint 1. • Prioritizing of each feature and user story. • Creating sprint backlog for sprint 1. • Dividing tasks for sprint 1.
3	0	Thurs Jan 19	• Goals for week discussed. • Discussed progress of preliminary research pertaining to GPT and museum itself. • Planned steps to overcome challenges and clarify individual understanding of the project. • Conveyed any resources or aid needed from scrum master (ex. floor plans, research, etc.).
4	1	Tues Jan 24	• Note progress over last couple days. • Continue GPT research and natural language processing. • Designed idea of simple UI for viewing outputs. • Discussed creation of database.
5	1	Thurs Jan 26	• Demonstrated for basic Input and output functions on the chatbot. • Showed latest pull for homepage, and admin login page. • Created a portion of the product backlog document.
6	1	Tues Jan 31	• Demonstrated admin login authentication. • Plan for linking database and chatbot. • Discussed research and plan for database population.
7	1	Thurs Feb 2	• Discussed general progress in last couple days. • Conveyed any resources or aid needed from scrum master (ex. Artifact research, etc.).
8	1	Tues Feb 7	• Finalized discussion for any

			remaining tasks.
9	1	Thurs Feb 9	• Time slot allocated for sprint review. • Planned for sprint 2, requirements specification for sprint 2 and create product backlog. • Created sprint 2 backlog. • Allocated tasks to team members.
10	2	Tues Feb 14	• Online meeting went over print progress so far. • GPT integration into current working system demonstrated. • Conveyed any resources or aid needed from scrum master (ex. GPT funding, meeting with Naser, etc.).
11	2	Thurs Feb 16	• Discussed any addition features system would need. • Front end recommendations discussed. • Backend integration issues addressed.
12	2	Thurs Feb 23	• Discussed plans for direction integration and path finding. • Used wipe boards to map out.
13	2	Tues Feb 28	• Planned for meeting with Professor Naser, plan demonstration and list of topics that want to be said to him. • Practiced demo.
14	2	Thurs Mar 2	• Scrum meeting with Professor Naser, notes taken on future suggestions and changes that will be implemented in upcoming sprints. • Sprint review. • Had Sprint retrospective and discussed plans to implement suggestions made by Professor.

2. Progress Thus Far

2.1 Completed Sprints

After each sprint was a sprint review where the team discussed methods that worked well from the previous sprint and would continue to be used in future sprints, as well as the omittance of certain practices or ideas that did not contribute to productivity. Section 2.2 outlines the specifics of what was completed in each sprint.

2.2 Implemented Features

2.2.1 SPRINT 1: Jan 24 – Feb 14

Below is a list of features and tasks successfully carried out in Sprint 1:

1. Created home page displaying option to use as a user or admin.
2. For admin, redirects to admin login page.
3. For admin, authenticates login.
4. For user, redirects to chatbot page.
5. For users, able to select which category they have inquiries for.
6. For admin, able to select which category they have updated information for.
7. Allow users to input a question about events and have an answer outputted.
8. Allow users to input a question about research and have answers outputted.
9. Allow users to input questions about museum and have answer outputted.
10. Create database for each category (events, research, museum info, directions).
11. Layered search to capture exhibit names from user input using GPT-3 as a classifier.
12. Client page for museum admin to update database for each category.
13. Conducted fine tuning development and research (changed to a prompt-based method).
14. Established GPT research and funding.
15. Defined map and museum layout.
16. Gathered holder information for population.
17. Inputted holder information for each category into database.

Specifically, Sprint 1 progress and features implemented, any issues encountered is discussed more in depth below:

1. Creating a home page for user chatbot and admin login was a simple and fundamental step to progressing with this project. The simplicity and importance of this task made it quite trivial and straight forward resulting in no challenges faced.
2. Creating a user-facing page to deploy the chatbot upon selection another simple and fundamental task for the project. The task was completed in time for other tasks to be worked on during the same sprint, as well, this task was associated with redirecting the user which created some friction for developers that had not coded with JavaScript and/or worked with React.js, resulting in the need for developer research.
3. For each page that was created, navigational tools were required, which lead to UI elements such as buttons being required in this sprint. Although the task was also quite trivial, an understanding was required in order to integrate the buttons to the redirect tool and the pages that connects to them.
4. Allowing the user to input questions into the Chatbot and receive responses related to their questions, required us to use this sprint time for establishing the framework for the future task of implementing GPT-3. This task involved having a chat box with a simple UI that showed the user input displayed in on the web page with a simple, temporary reply setup. There were little to no challenges with this task when it came to implementation of the chat-bot but rather just the UI design which had which would change according to the response of the group members.
5. The next task was creating a page for login, which was also quite simple and finished in time. The final product included using a Google authentication rather than implementing a sign in/login, after some research was conducted and it was found that using a Google authentication was a simple solution that would save time on implementation.

6. To authenticate the user login using the new idea of Google login was completed with relative ease. There were no challenges the API for Google login handled authentication and authorization.
7. The database was created with sample data that would later be implemented with the chatbot. Some researcher was required to better understand the API Key and the functionality of a .env file working with gitignore.
8. Able to input information into the database, this information represented potential answers to the user's question. For simplicity and testing purposes, fake data was inputted, in later sprints this fake data will be replaced with accurate data from the museum.
9. String search was implemented in order to confirm the pulling of data from the database but not used as GPT was not yet integrated.
10. Integration of all previous steps with each other. The task that worked together were the redirecting to pages, google authentication, and the updates forms working. These tasks were integrated with minimal challenges.
11. Fine tuning development and research was a task that included research about GPT, and its integration for human interacts. As research was being done, the fine tuning was also taking place. The challenges that came along with this was the ability for the GPT to response in a certain way with such details such as profanities and toxic replies being suppressed.
12. GPT research for integration with working web-app alongside funding. A more research based tasked included the integration of GPT into working setup for chat box. Upon learning that training the GPT on a set of data requires certain tokens which cost $0.06 per token that are approximately 1000 words.
13. Map and museum layout defined initially the understanding of implementing a directional component would require a layout of the museum. Due to the smaller size and difficulty of acquiring such blueprints, the task was not completed rather was scraped.
14. Gather information for population much like the task of the directional component was not followed through but rather a set of demo data was used to work with the GPT component.
15. Restyling front end overtime was a long task that was completed with certain UI related challenges. This task took place throughout the whole sprint with each step things were changing and being added.
16. The creation of the homepage displaying option of user or admin was quite easy and completed on time.
17. Using the React-router-dom the direct to admin login page was simple.
18. Much like the admin redirecting using the React-router-dom the direct to chatbot page was simple.
19. The selection of inquiring about specific categories was not completed. The reasoning for the incompletion of this task was due to GPT's ability to understand what inquiries the users seeks.
20. The ability to select which category was updated for the admin was not completed, yet it was changed so that there is a single form which included a title field with form that came with subfields and so a category could be created and changed upon filing out the form found in the admin page.
21. Much like the previous task in which a category specific question was asked, the group had formed a new strategy that including GPT and eliminated fine tuning after research was adopted which revolved around the idea of GPT recognizing the user input without specification of a category.

2.2.2 SPRINT 2

Below is a list of features and tasks successfully carried out in sprint 2:

1. Query ChatBot integration
2. Integration of chatbot to front end
3. Speech to text option
4. Text to speech option
5. Front end integration
6. Creating database information for each category
7. Connecting database and chatbot
8. Create floor plans
9. Create method to complete directional calculations in next sprint
10. Front end updates
11. Directional chatbot integration
12. Populating database

Specifically, sprint 2 progress and features implemented, any issues encountered is discussed more in depth below:

1. Query ChatBot integration. This task was completed on time and the ChatBot is fully functional with accurate that are pull from the web and specific information provided by the development team on the NOTL museum. The challenges that were faced during this task including the conversion of python code to JavaScript while using the OpenAI API, and the integration of the premade front end of the ChatBot as the code was written by a different developer and was created before the completion of the OpenAI API.
2. Natural language understanding tasks. These tasks involved extracting relevant information from the user input and funneling it through a series of GPT-3 queries to understand the context of the input and the information being asked. Challenges faced with this were creating accurate GPT-3 queries with high success rates.
3. Integration of Chatbot. This task included refining the UI component of the ChatBot with in the web-app. The challenges that were faced during this task were not too drastic, but the method of trial and error resulted in the best user experience of the ChatBot
4. Speech to text option. This task was not completed, as there were last minute integration issues and minor errors within the code.
5. Text to speech. Alongside the speech to text option this task was also not completed. Created with the speech to text option, the text to speech also had similar complications.
6. Front end integration. This task was a test and refinement of the significant amount of changes made during this sprint in order to give the best user experience, this task was completed in time. The challenges included the method of trial and error in order to achieve the best results.
7. Creating database information for each category. This task is a subset of populating the database with categories such as exhibits, artifacts, and much more. The task was

not implemented to completion but rather a demo data set was used with one category filled which was used to integrate with the system.

8. Connecting database and ChatBot: The task included ChatBot pulling information from the database in order to know specific information on the museum, the task was completed but not tested thoroughly as there isn't a large set of data.

9. Create floor plans. A task that will set the future foundation for the implementation of directions within the museum. A demo floor was implemented for the time being for the lack of information provided by the museum. The floor plans will be created in a way that any future floor plans can implemented on top of the current. This task was completed with minimal challenges.

10. Front end updates. A mid-way point to make certain changes that go along with the user experience of the front end. There was no challenges and the task was completed with ease.

11. Create method to complete directional calculations in the next sprint. A research-based task that had the developers look into different path finding algorithms best suited for the museum. Upon more research it was deemed that the A-Star algorithm would be best suited for such a task. The idea of future implementation had to be kept in mind while doing this research, following this there were not many other challenges that hindered this task.

12. Directional ChatBot integration. To integrate the directions of the path finding algorithm the developers initially had thought, the implementation of a prompt asking for directional ChatBot rather than a query based ChatBot would be most viable. Ultimately, the developers completed the task with a prompt from the query-based chat asked for the starting and end location of the location the user wanted to go to. This was not quite as challenging but rather a decision for future UI use need to be made.

13. Populating the ChatBot. The failure of this task was due to the changing of minds for the developers to rather make a demo data and shift the population task to the next sprint where 500 plus sets of data are going to be used to populate the database by using a scraper.

2.3 Successful and Effective Methods

The following is a list of practices that have been effective and significant contributors to our progress thus far:
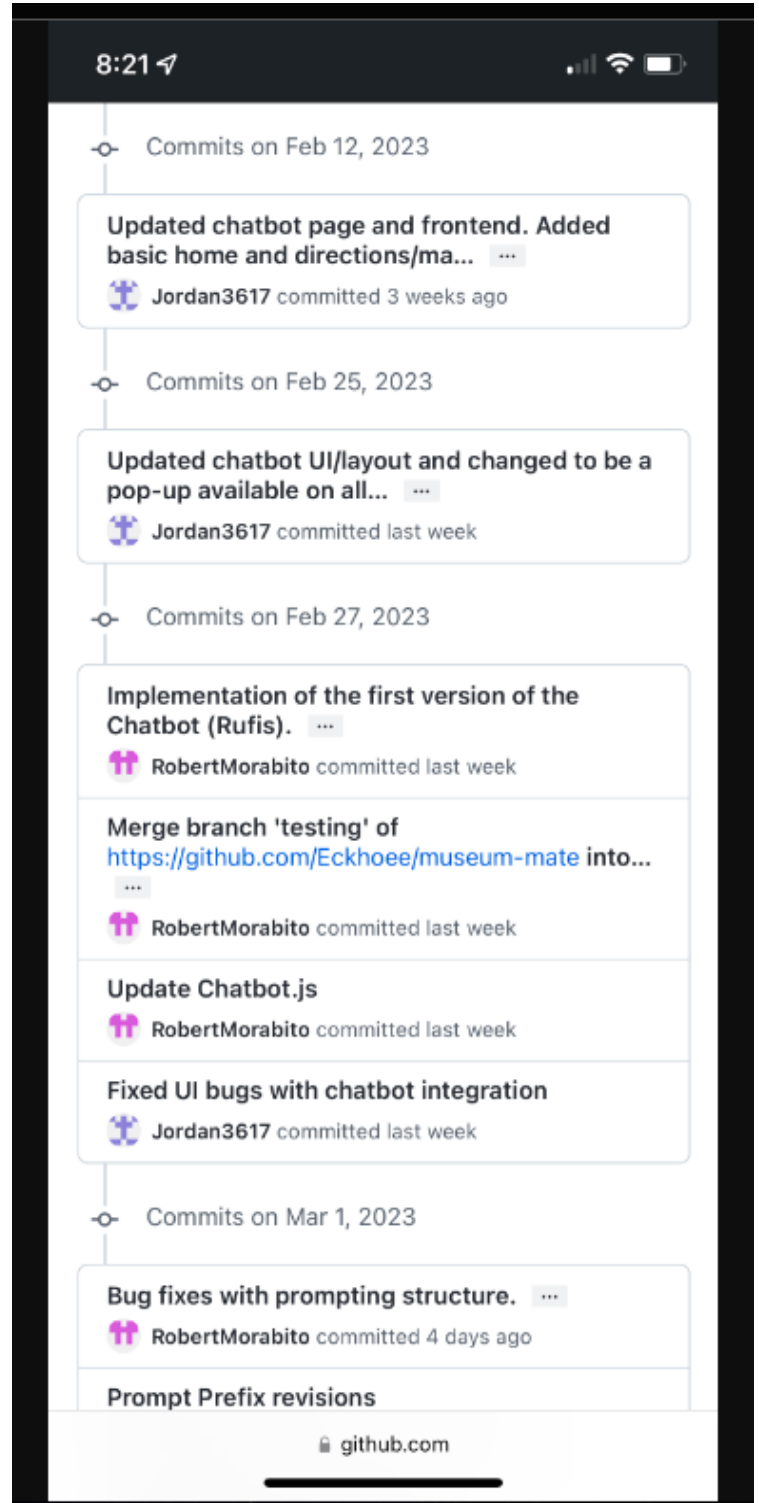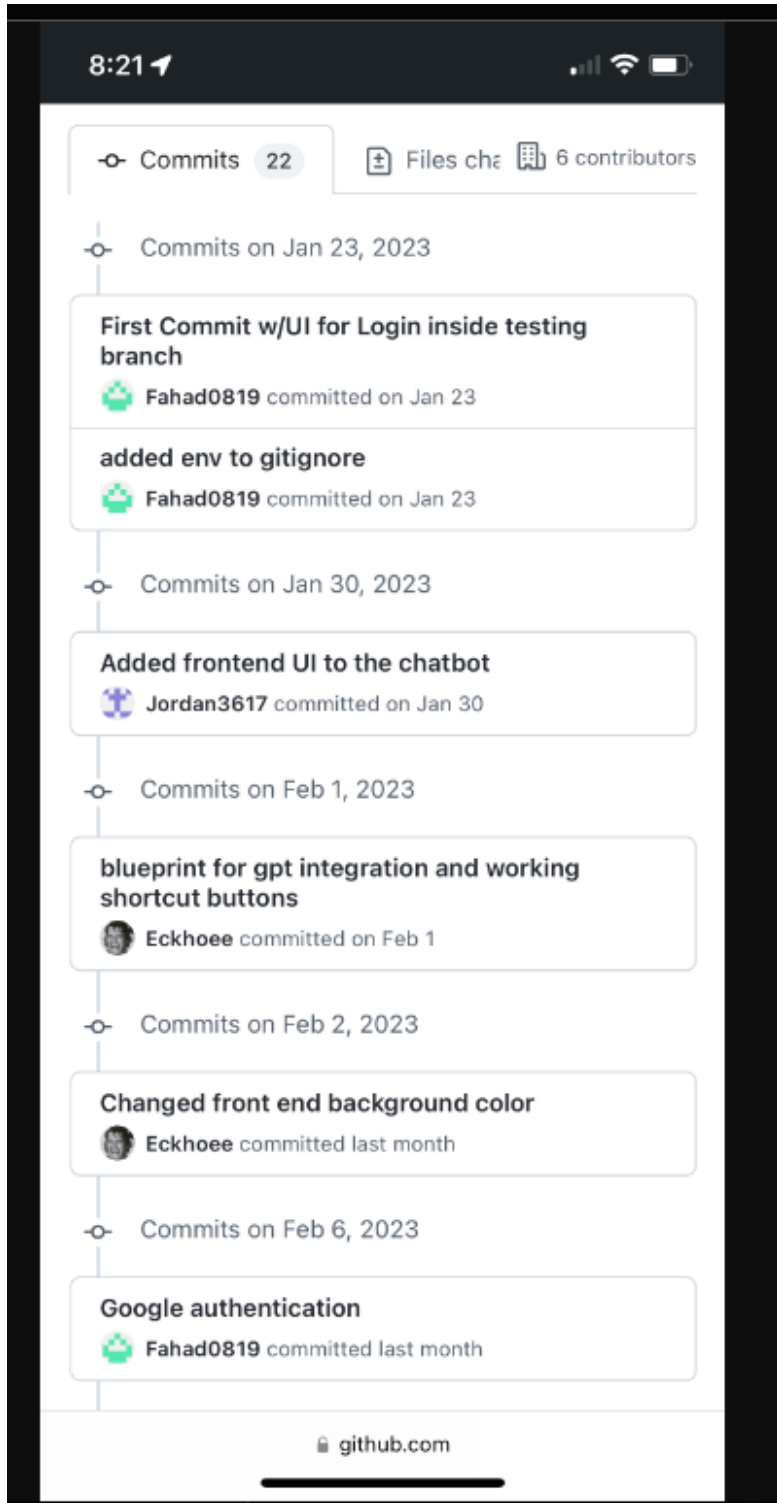
- Using Jira, recording tasks to do and what's been completed
- Regular scrum meetings twice a week
- Daily check ins in Discord channel
- Pair programming – assigning two developers to complete a task together, was proved to be more efficient and faster (For example, Fahad and Dana completed database requirements together within two days, whereas, working separately for a week before had no progress)
- Refactoring and commenting code wherever possible (improved others understanding code so its easier to build off of)
- Communicating difficulties and issues on a daily basis
- Resorting to scrum master for external resources such as floor plans, data, etc
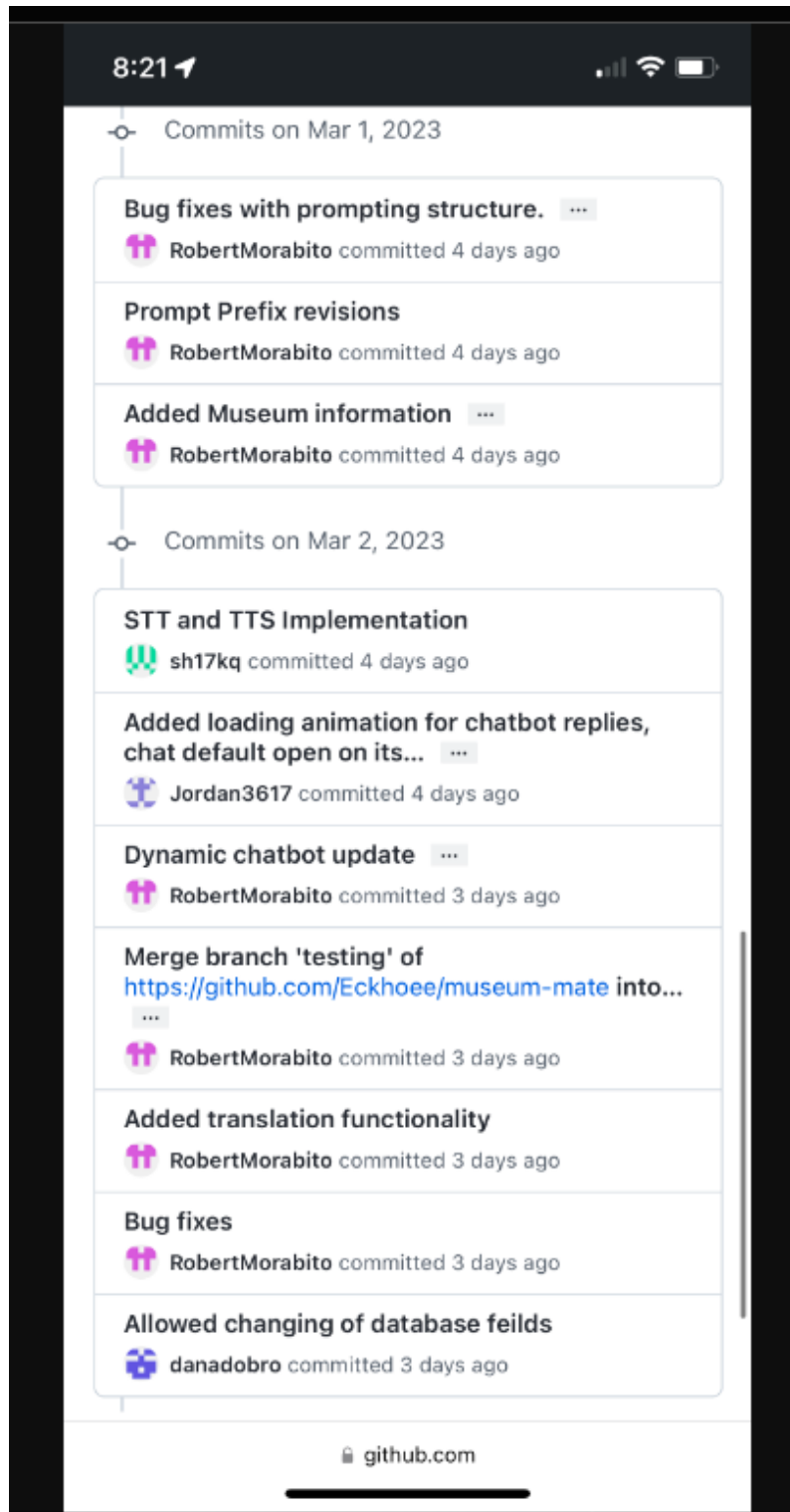
2.4 Team Contributions & Issues Encountered

Above, in section 2.2 issues encountered by the whole group were discussed, this section outlines issues encountered by each specific group member. This table represents each team member, tissues they encountered with the project so far, and the area in which they contributed to:

Member Name	Role	Contributions	Issues Encountered
Fahad Arain	Developer	• Database creation and administration • Admin authentication and google sign in • Integration of all components together	• Staying aware of changes made by other developers in order to integrate code into system
Maheen Samad	ScrumMaster & Product Owner	• All documentation • Proposal Brief • Product Backlogs • Requirements Documentation • Hosted all meetings, documented tasks done	• Workload balance – how much is too much or too little in a task? • Finding GPT funding
Robert Morabito	Developer	• GPT-3 integration • Chat-bot logic system (all the layers data goes through) • Training data creation/ prompt design • Output translation into several languages	• Integrating all individual work done with other's work • Learning JavaScript to integrate original python code
Jordan Bharati	Developer	• Front end development and implementation • Front end design and layout • Chatbot UI	• Learning JavaScript as a first-time user
David Bailey	Developer	• Sprint 1 front end development with Jordan • Research for directional/ mapping component • Aided in integration of components together • Github utilization	• Learning JavaScript as a beginner • Understanding how many tasks are too much or too less in a sprint – workload balance
Dana Dobrosavljevic	Developer	• Database creation and administration • Populating database • Allowing admin to update database information • Connected project to firebase	• Connecting to firebase • Enabling user to enter data which shows up on the database
Sarah Howcroft	Developer	• Text to speech, speech to text planning	• Learning JavaScript on a time crunch

Specific team member contributions are documented from
our GitHub page below:

8:21 ⬈

Commits on Mar 1, 2023

Bug fixes with prompting structure. ⋯
🧑 RobertMorabito committed 4 days ago

Prompt Prefix revisions
🧑 RobertMorabito committed 4 days ago

Added Museum information ⋯
🧑 RobertMorabito committed 4 days ago

Commits on Mar 2, 2023

STT and TTS Implementation
🎣 sh17kq committed 4 days ago

Added loading animation for chatbot replies, chat default open on its... ⋯
🤖 Jordan3617 committed 4 days ago

Dynamic chatbot update ⋯
🧑 RobertMorabito committed 3 days ago

Merge branch 'testing' of https://github.com/Eckhoee/museum-mate into... ⋯
🧑 RobertMorabito committed 3 days ago

Added translation functionality
🧑 RobertMorabito committed 3 days ago

Bug fixes
🧑 RobertMorabito committed 3 days ago

Allowed changing of database feilds
🧩 danadobro committed 3 days ago

🔒 github.com

3. Future Sprints

3.1 Sprint 3

This section will go into the details which outline features which will be implemented in Sprint 3. Sprint 3 has commenced March 1 and will complete March 14. As agile is adaptable to stakeholder feedback, additional feature recommendations done by sprint reviews, sprint retrospectives, and meetings with stakeholders allowed developers and the scrum master to plan for the following features to be implemented by the end of Sprint 3:

- Populate database from legitimate data from museum rather than temporary and fake data (Ex. The War of 1812)
- Create web scraper in order to retrieve information from museum database
- Ensure chatbot responds to museum related data (such as the War of 1812, etc.) that is retrievable from database
- Training of GPT on data by category rather the set of all data
- Implement text to speech and speech to text feature for variation in input for users
- Implementation of a-star algorithm for directional component, this component will provide users with directions to different exhibits, aiding in navigation around the museum
- Refining chatbot to better accommodate directional component
- Update user interface to meet product owners' requirements

3.2 Sprint 4 and 5

This section will outline the upcoming plans for sprint 4, scheduled to start on March 15th, and complete March 29th. It is important to note that this section is susceptible to change as a sprint review and sprint retrospective, as well as our first presentation will determine any sort of changes or addition that will be needed in our current plan. The most integral portion of sprint 4 will be to implement a visual component into the working system. Specifically, after a successful sprint, the system will be able to output visuals such as pictures, diagrams and videos to the user based on their question. For example, rather than physical commands to get from Exhibit A to Exhibit B, the chatbot will output a small map which provides directions similar to small scale maps such as mall maps. This feature needs to be adaptable to change such as when exhibits are changed or deleted through the admin side. Such requirement is crucial in the event where an exhibit no longer exists, or a new exhibit is created. Additionally, the chatbot will output resources such as videos and pictures when scraping the internet for research purposes in cases where the user enters research related questions about artifacts and history. Another plan for this sprint is to finalize a user-friendly interface that results in an optimal user experience. Ultimately, any sort of setbacks or reevaluation will also be incorporated into this sprint.
Essentially, sprint 5, which will start on April 1st and complete April 11th. The purpose of this sprint is to have an intensive testing phase. Although, testing has and will be done throughout each sprint, after each feature is implemented, the team plans on going through a strict testing phase to ensure each feature is tested and is accurate in the response or feedback it provides the user with. Another key part of this sprint is to select a batch of potential real-life users of the system and receive feedback on their user experience, and then implement any improvements. This sprint will also include the planning for the final report, and the final presentation.

Continue reading to see a current working version of the Museum Mate system below.
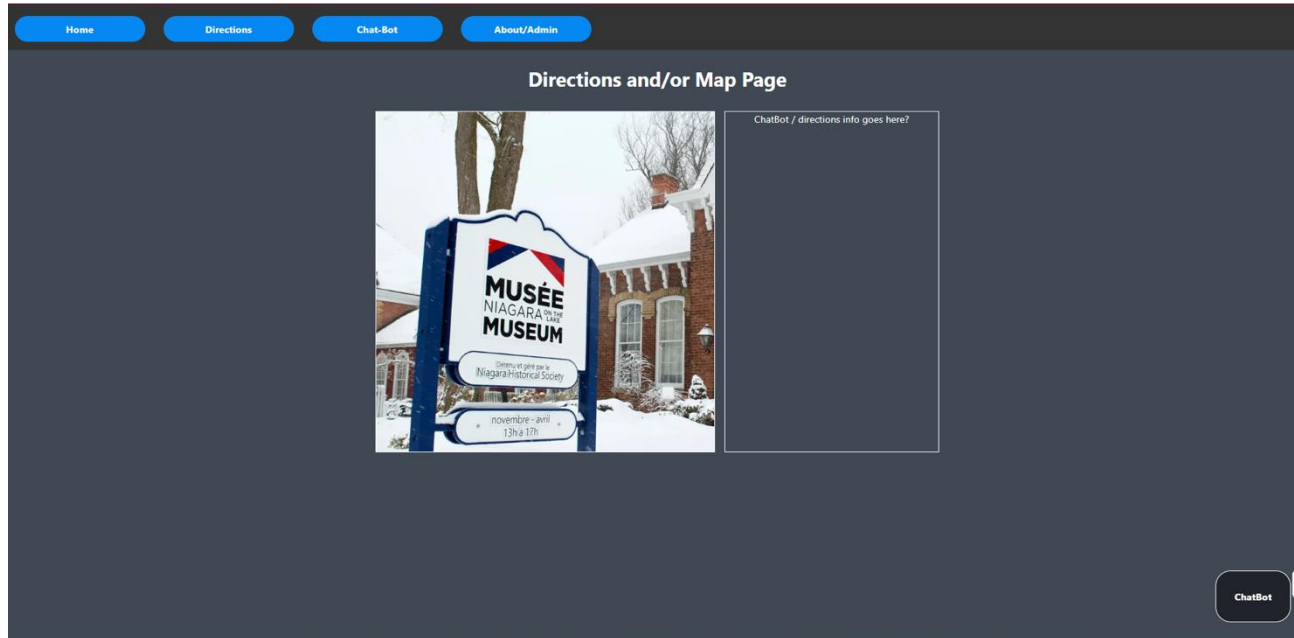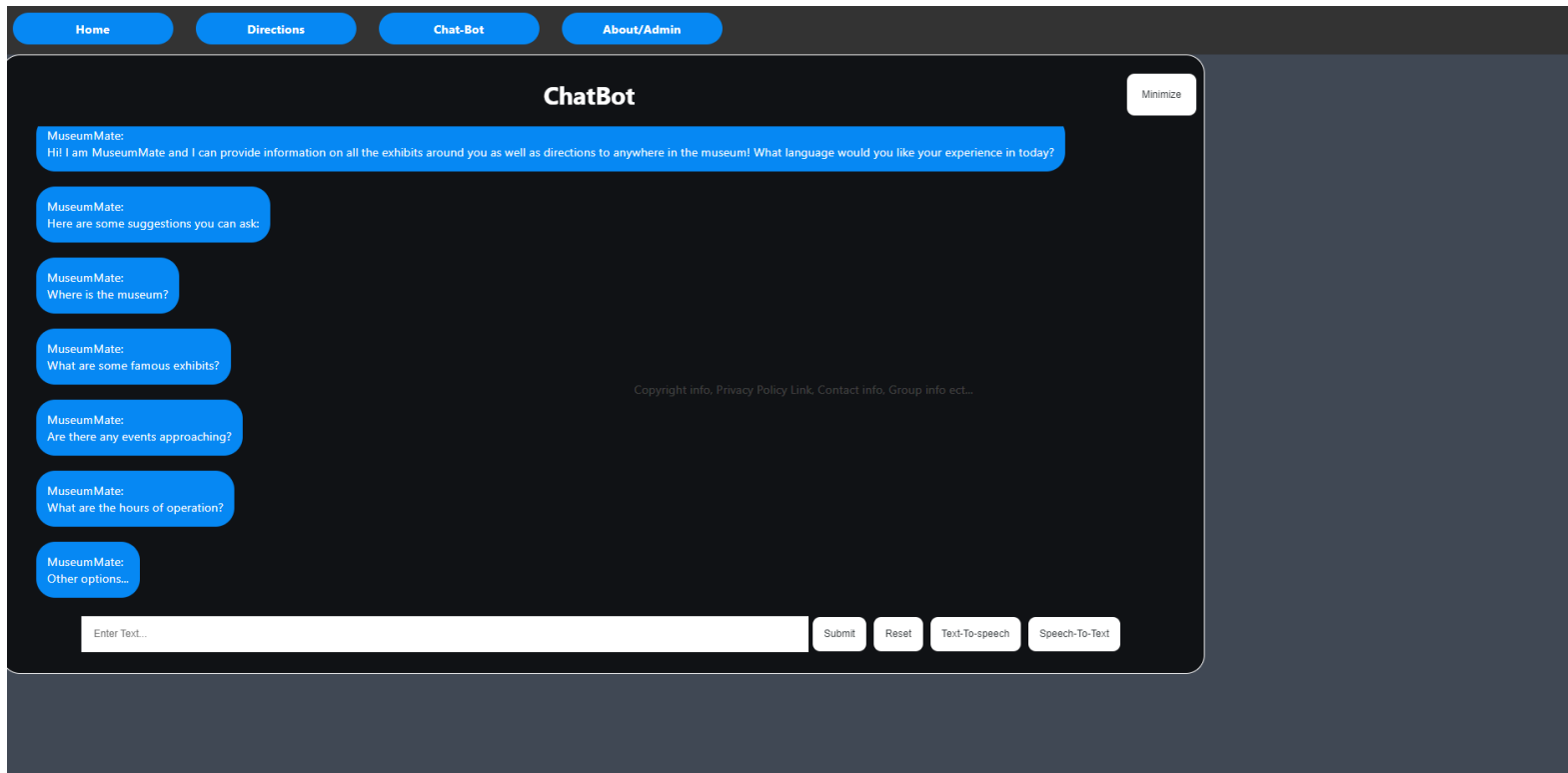
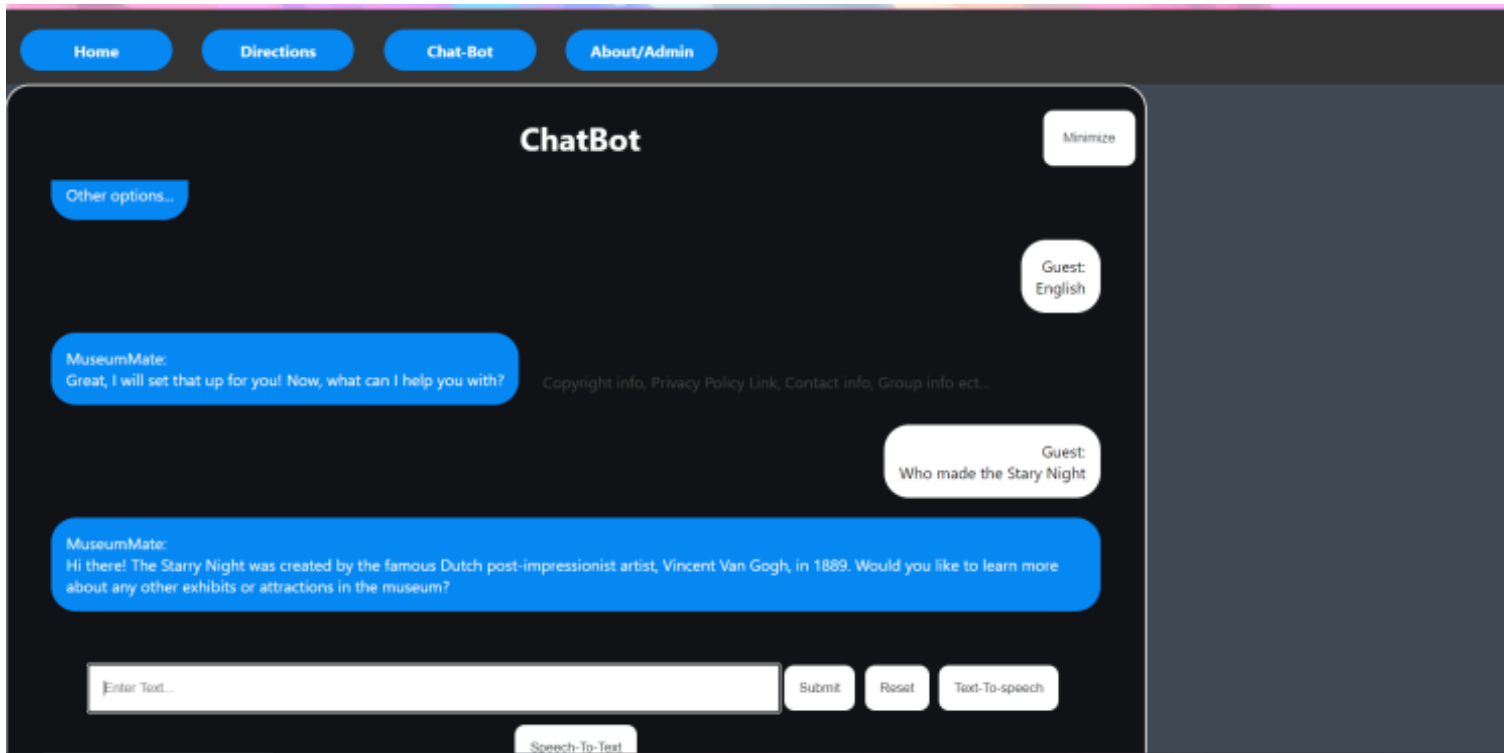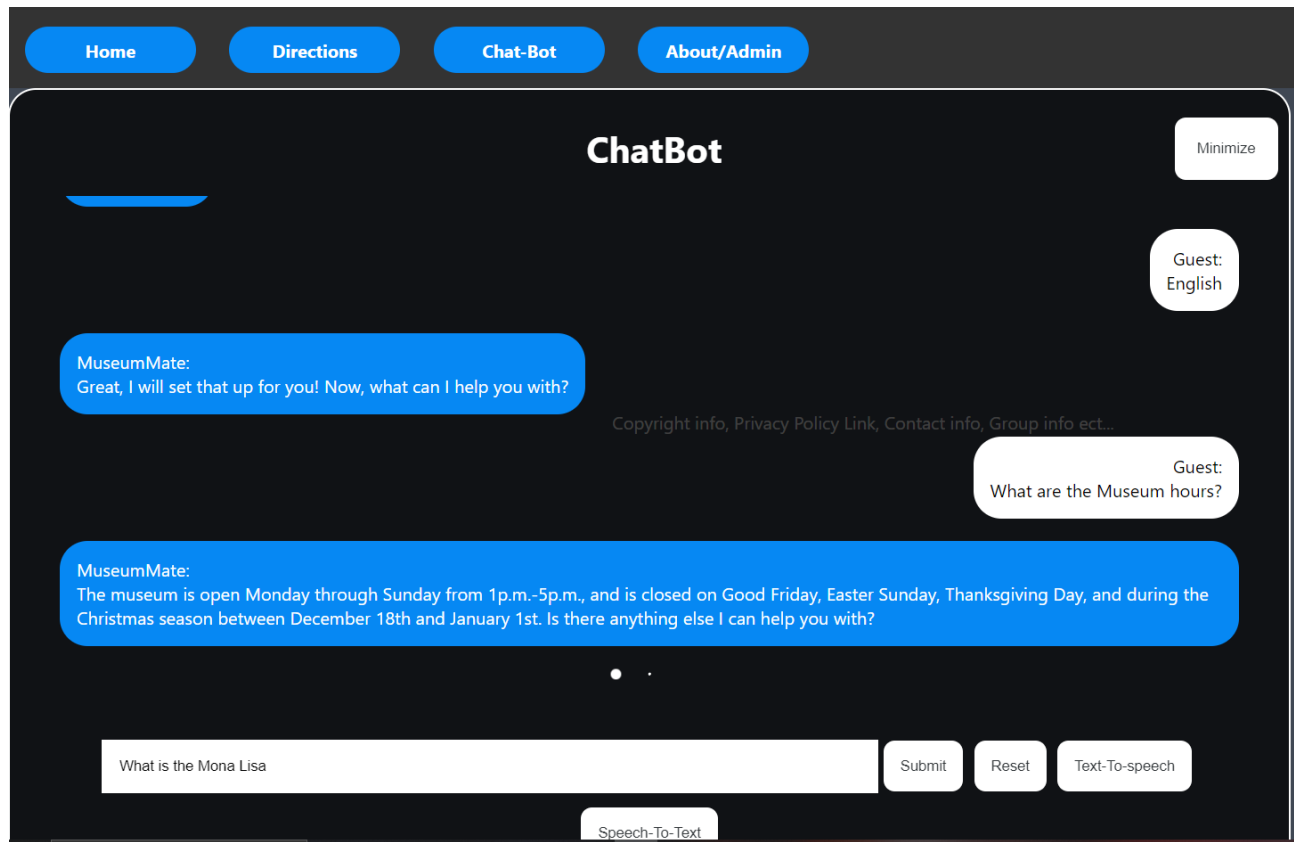# 4. Current Working System

4.1 Current Working System

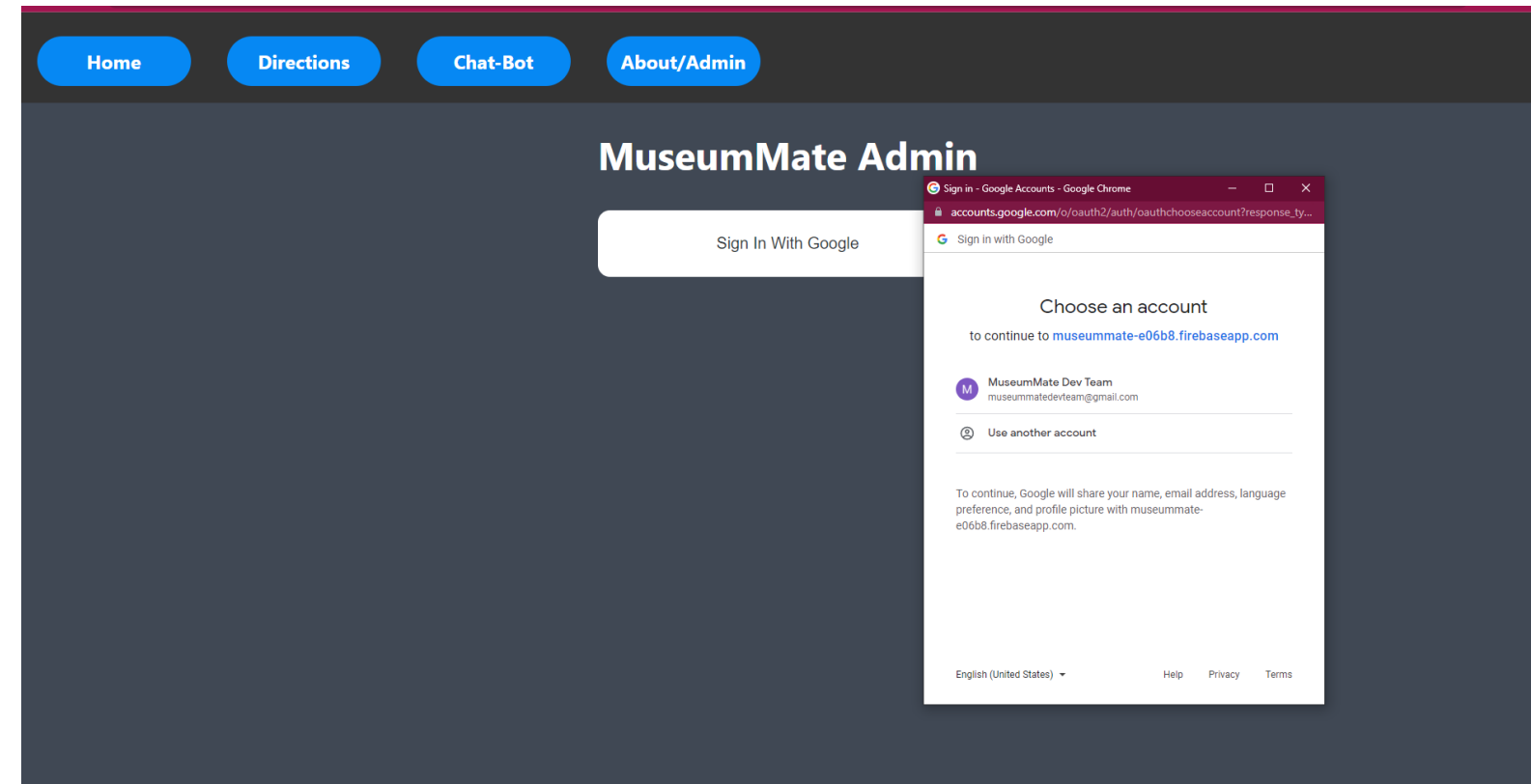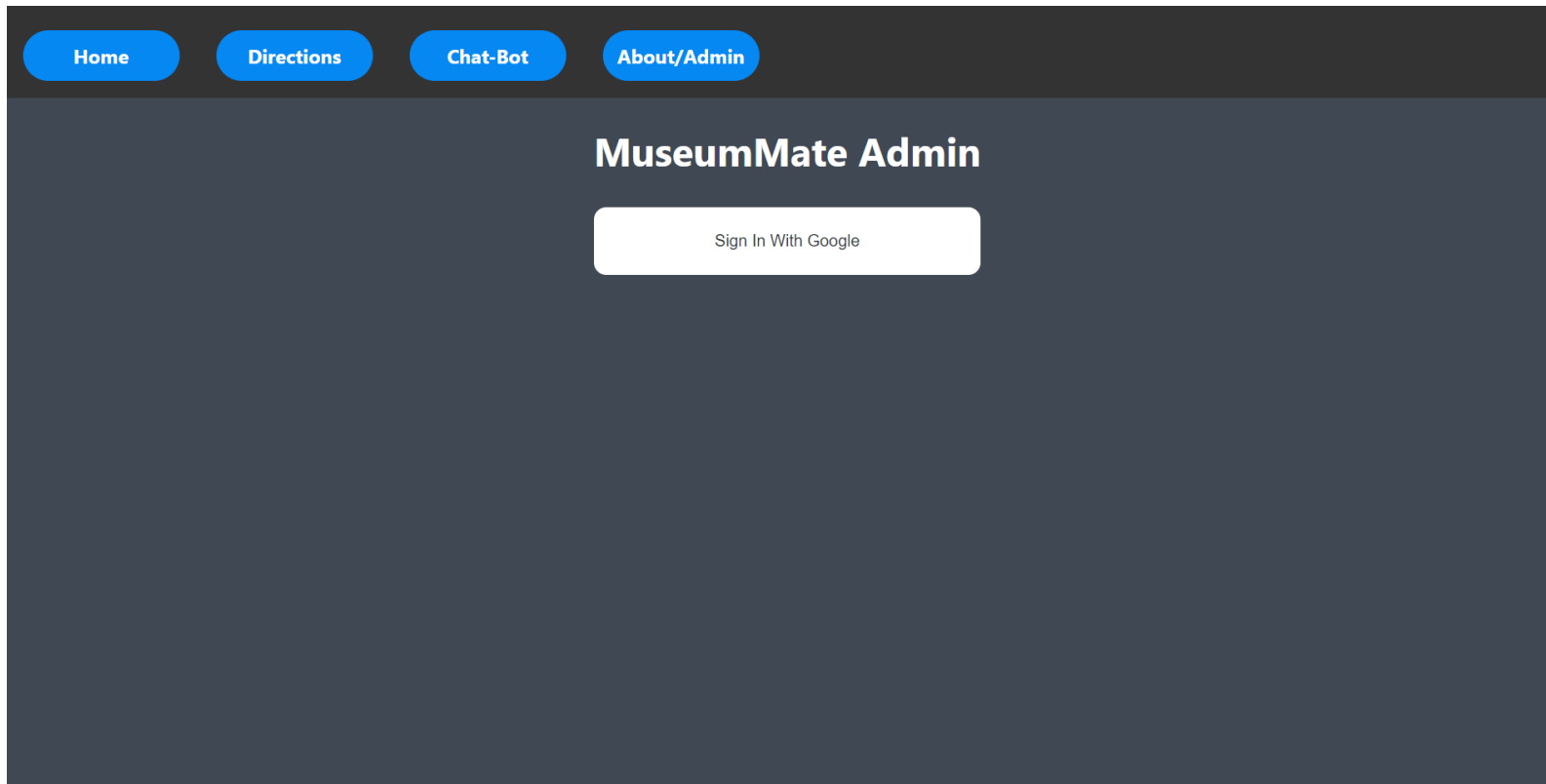Homepage
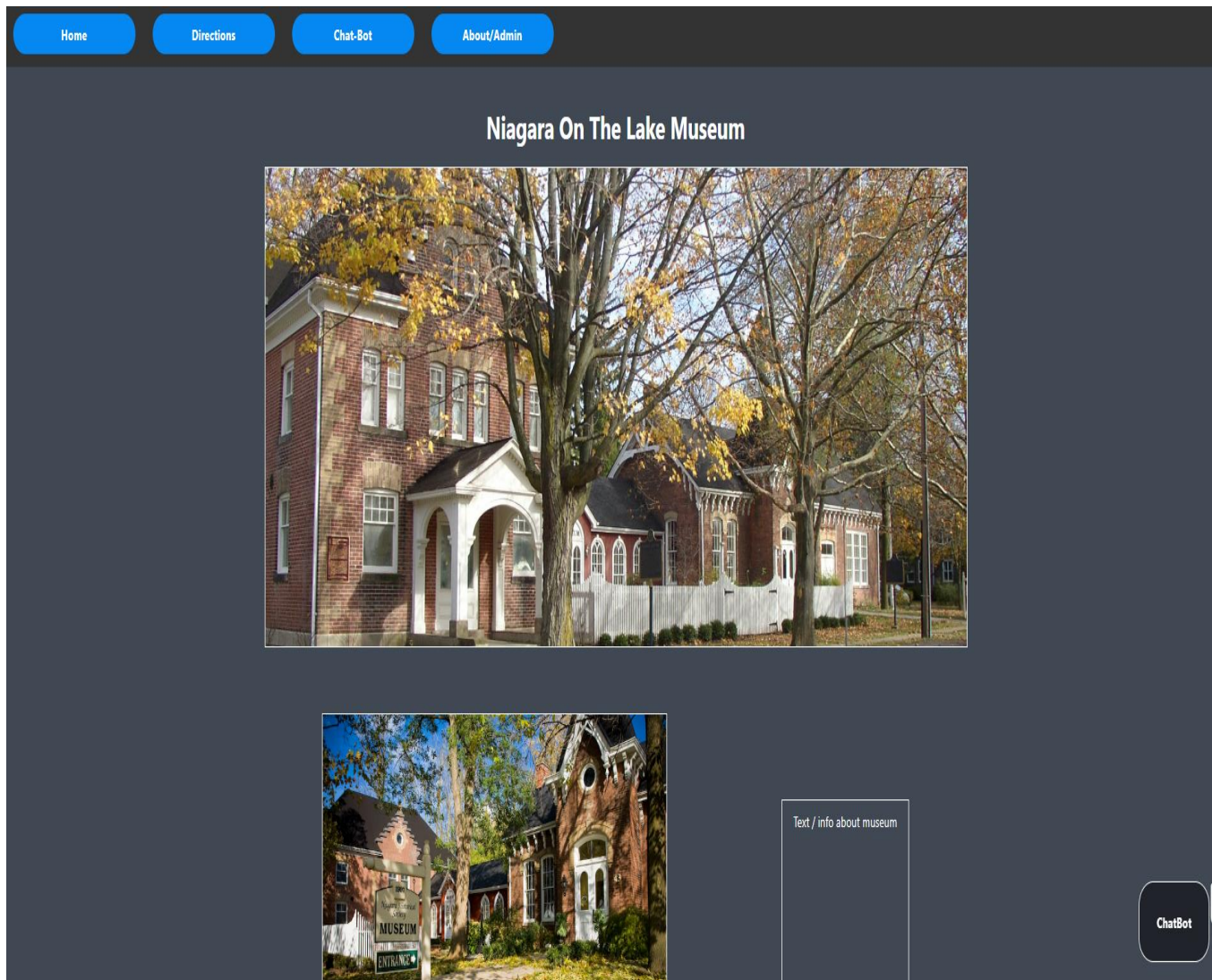
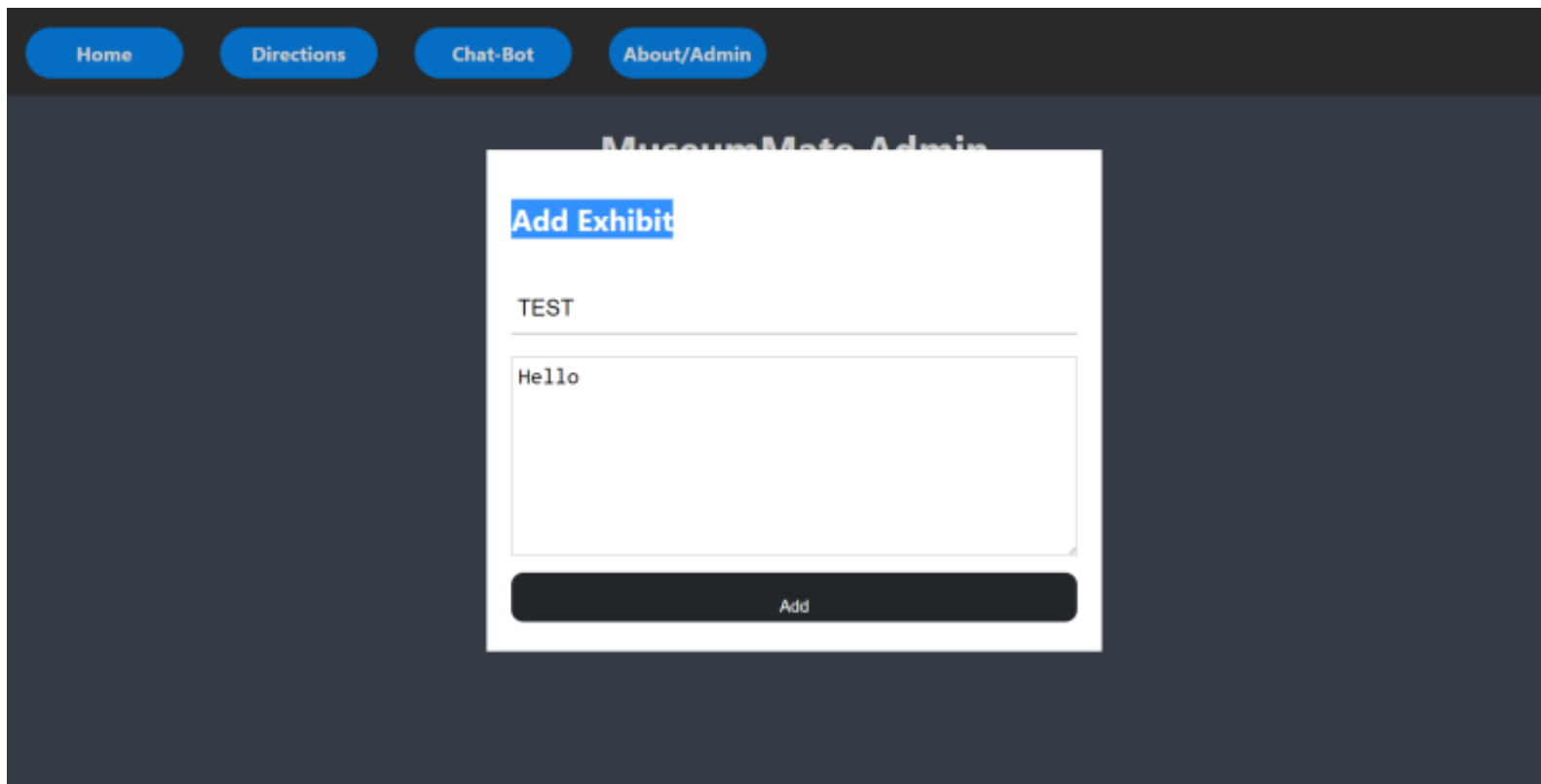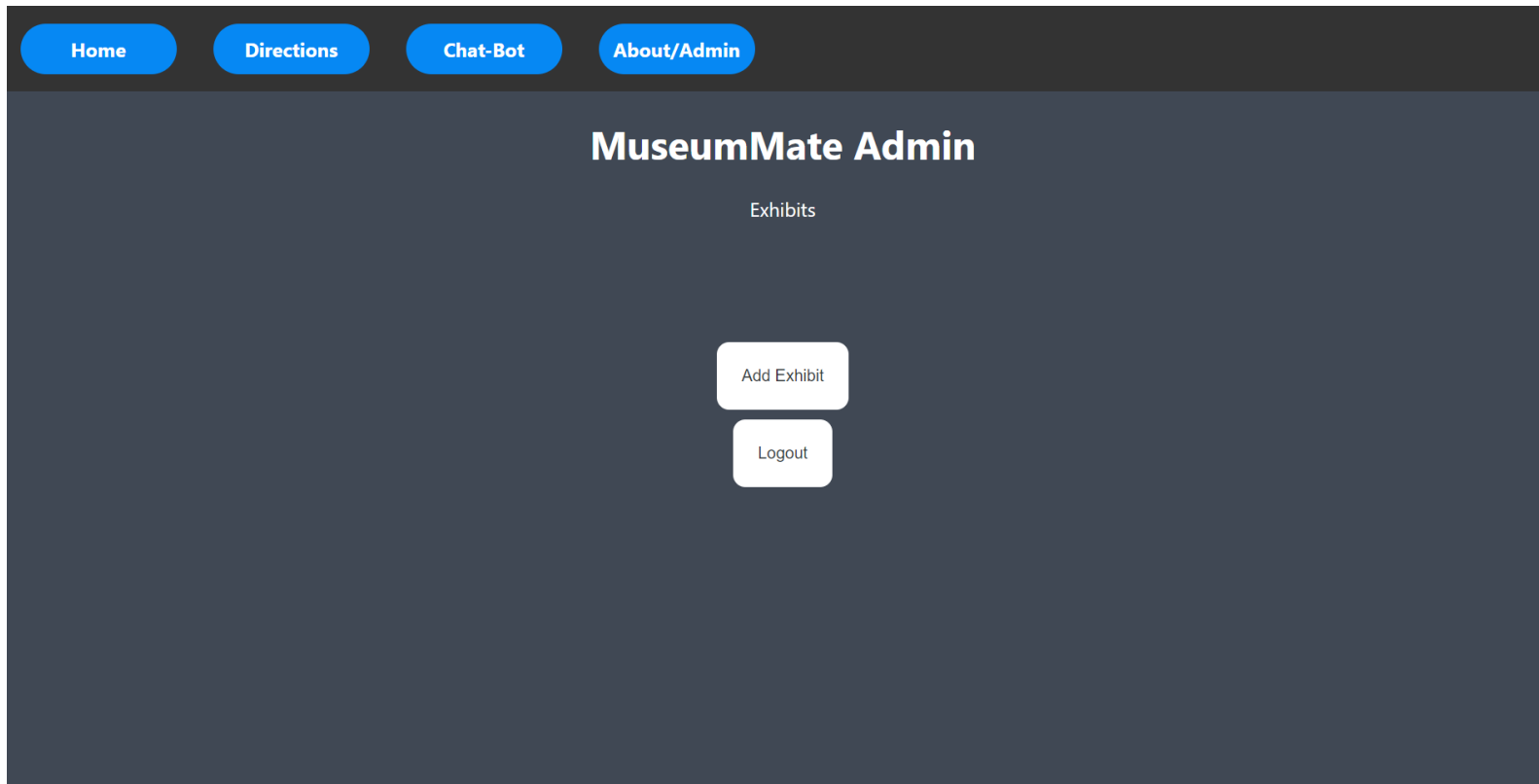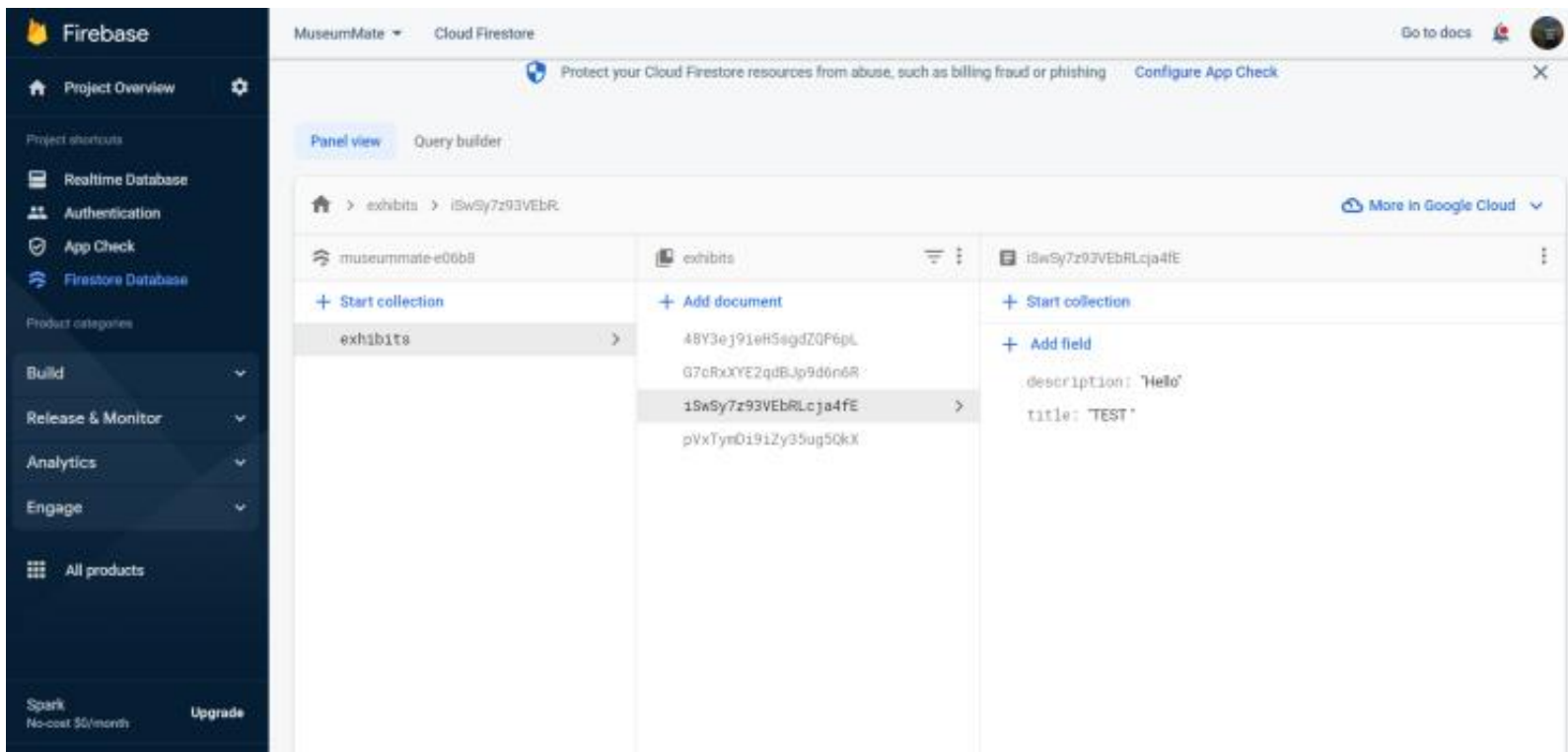Directional and Map Page (Sprint 3 will be completed)

Chatbot Responses:

Admin Authentication and database manipulation:

Translation Feature:

| Home | Directions | Chat-Bot | About/Admin |

ChatBot

Minimize

Other options...

Guest:
Hindi

MuseumMate:
Great, I will set that up for you! Now, what can I help you with?

Copyright info, Privacy Policy Link, Contact info, Group info ect...

Guest:
what are museum hours

MuseumMate:
म्यूज़ियम सोमवार से रविवार तक, 1 बजे से 5 बजे तक खुला है। हम गुड फ्राइडे, ईस्टर संग्रह दिवस, थैंकसगिविंग और दिसंबर 18 से जनवरी 1 तक के दौरान क्रिसमस सीजन में बंद हैं। आपके लिए अभी कुछ और कर सकते हैं

Enter Text...   Submit   Reset   Text-To-speech

Speech-To-Text