GuideSense

Your Trusted Navigation Companion

Team: Power Angers

Sprint 3

Agenda

Power Anger

Afziya Waknis

Project Manager/
Developer
Aw14091n@pace.edu

Kaiyin Chen

Database Administrator
Kaiyin.chen@pace.edu

Power Anger

Ritesh Singh

Devops Engineer
Ritesh.singh@pace.edu

Dinesh Gopi Sunkara

Developer
Ds46669n@pace.edu

Power Anger

Rushabh Makwana

Backend Developer
Rm36294n@pace.edu

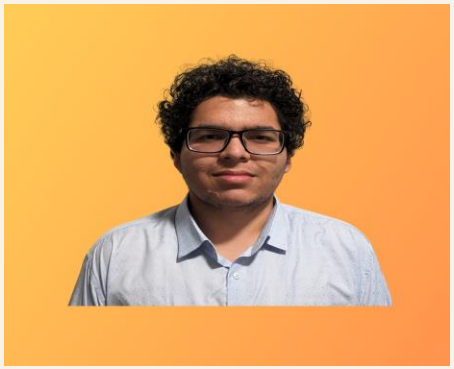
Vaibhav Thapliyal

Developer Lead
Frontend Developer
Vaibhav.thapliyal6@gmail.com

Power Anger

Min Jung

Quality Analyst
Mj34564n@pace.edu



Hrishikesh Shah

Frontend Developer
Hs75142n@pace.edu

Improvements

- User stories (US1, US7)
- Burndown chat (Slide 39)
- Code Demo Video
- Retrospective Video
- API code explanation(Slides 52-54)

Problem Statement

- Navigating through everyday environments presents significant challenges for blind and visually impaired individuals., moving vehicles, and the lack of real-time guidance.
- Traditional mobility aids like canes and guide dogs, while helpful, often fall short in providing comprehensive and real-time information about obstacles, directions, and surroundings.
- This lack of real-time situational awareness can lead to increased risks of accidents and restrict the independence of visually impaired individuals.
- Therefore, there is a pressing need for an innovative solution that leverages modern technology to enhance mobility and safety for the visually impaired community.
- This project aims to create a web app that uses object detection and voice commands to help blind people navigate safely by warning them about obstacles, giving directions, and describing their surroundings in real time.

Project Description (improved)

Project Name:	GuideSense
Team:	Power Angers
Project Description:	For visually impaired individuals who need assistance navigating independently in urban and outdoor environments, the real-time object detection web app is a computer vision and AI-powered solution that identifies objects, obstacles and crosswalks while providing real-time audio guidance and GPS-based navigation support. Unlike traditional mobility aids or existing navigation apps that lack real-time object detection, our application offers a seamless, intelligent, and accessible way for visually impaired users to navigate confidently.
Benefit Outcomes:	Increased Independence, Enhanced Safety, Improved Mobility, Greater Confidence, Accessibility, and Inclusivity.
GitHub Link:	https://github.com/htmw/2025S-Power-Anger/wiki

Team Working Agreement

CS-691 SPRING 2025
TEAM WORKING AGREEMENT
TEAM-POWER ANGERS

Communication

- Team will Communicate with each other through Email and WhatsApp
- There is going to be a team meeting where all 8 members are required to join on every Tuesday after 9pm.
- Technical Team meeting where developers would join the call for brief about the tasks and the updates it will be on tuesday and thursday at 9pm
- Team members are expected to update beforehand if the they are going to be absent for the meeting and asked to be updated till the next meeting
- Each team member should complete the given task before the deadline.In some one case was not able to do so then they should inform it to the rest of the team so they could divide the task

Work Division and Participation

- The entire project work should be divided into equal parts and equal responsibility should be given to all team members.Mebers are expected to select and contribute to the task in which their skill are best fit.
- Jira, Github will be used to track and divide all our work
- Every team member should update about their task 2 times in a week.

- Each team member should complete their part of work before the deadline.If one fails to do so immediately report to the rest of the teammates and take assistance.
- In case a team member is absent in the team meeting,members must support the decision taken in the meeting.

Team Working Agreement

Respect

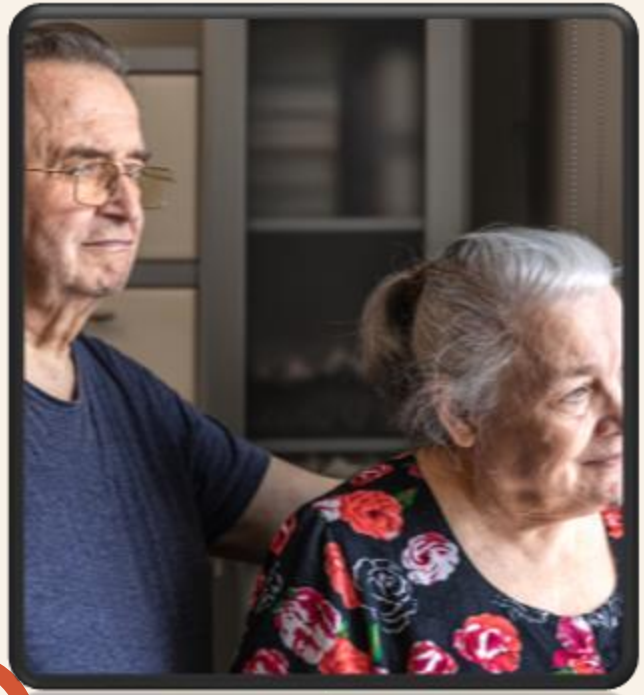
- It is essential that all team members have a chance to share their opinion and make any suggestion without judgement.The team project is team effort,taking advantage of collective knowledge to come up with solutions.
- All members agree to respect each other's personal schedules and listen to each other's perspective.

TEAM MEMBERS	EMAIL
Afziya Waknis	aw14091n@pace.edu
Kaiyin Chen	kaiyin.chen@pace.edu
Ritesh Singh	rs98576n@pace.edu
Dinesh Gopi Sunkara	ds46669n@pace.edu
Rushabh Makwana	rm36294n@pace.edu
Vaibhav Thapliyal	vt18517n@pace.edu
Min Jung	mj34564n@pace.edu
Hrishikesh Shah	hs75142n@pace.edu

Persona One: The Independent Blind User

- **Name:** Aisha Khan
- **Age:** 32
- **Occupation:** Software Developer (works remotely)
- **Background:** Aisha has been blind since birth. She is highly tech-savvy and relies on assistive technology. She is motivated to use tools that enhance her independence and streamline daily activities.
- **App Usage Scenario:** Aisha wants to use the app for navigation while walking, especially in unfamiliar areas, and identify objects which could be a barrier for seamless tasks.
- **Needs:** Seamless integration with existing assistive technologies (screen readers, voice control), accurate location services, reliable information about accessible routes and environments, and robust privacy features.
- **Goals:** Increased independence, improved access to information, enhanced safety while navigating, and streamlined daily task management.

Persona Two: The Concerned Family Member



- **Name:** Robert Chen
- **Age:** 65
- **Relationship:** Son of a visually impaired senior
  "I worry about my mother's safety. I hope this app can help me stay connected and ensure she's doing okay."
- **Background:** Robert's mother is losing her vision due to macular degeneration. He lives in a different city and wants to stay connected and provide support remotely.
- **App Usage Scenario:** Robert wants to use the app to track his mother's location (with her consent), receive alerts if she deviates from her usual routes, remotely assist her with tasks like medication reminders, and communicate with her easily through the app's accessible interface.
- **Needs:** User-friendly interface, reliable location tracking (with privacy safeguards), remote assistance features, accessible communication tools, and clear instructions for setup and use.
- **Goals:** Increased peace of mind, improved communication with his mother, ability to provide remote support, and enhanced safety for his visually impaired parent

Persona Three : Partially sighted Teenager

- **Name:** David Miller
- **Age:** 16
- **Background:** David has low vision due to a genetic condition. He can see some things with the aid of glasses or magnifiers.

"I want to be able to do the same things my friends do. I hope this app can help me navigate more easily and access information."

- **App Usage Scenario:** David wants to use the app to magnify text and images, identify colors, navigate public transportation, and access audio descriptions of videos and other media. He needs an app that is easily customizable to his specific visual needs.
- **Needs:** Customizable display settings (font size, contrast, color schemes), reliable object and text recognition features, seamless integration with magnification tools, and accessible interface that can be used with limited vision.
- **Goals:** Increased independence, improved access to information, enhanced social participation, and greater confidence in navigating the world.

MVP

- A client (web browser) connects to the server and sends a WebRTC offer
- The server sets up a peer connection, processes the offer, and sends back an answer
- When video starts streaming from the client, each frame is:
  - Received by the server
  - Processed through YOLO (object detection)
  - Enhanced with visual indicators (boxes around objects)
  - Sent back to the client in real-time

Technologies & Tools (improved)

NodeJs: RunTime 

React: Frontend 

Python: Object Detection 

Clerk: Authentication 

MongoDB: Database 

AWS: Deployment 

Jira: Project Management 

GitHub: Code Versioning 

OneDrive: File Sharing 

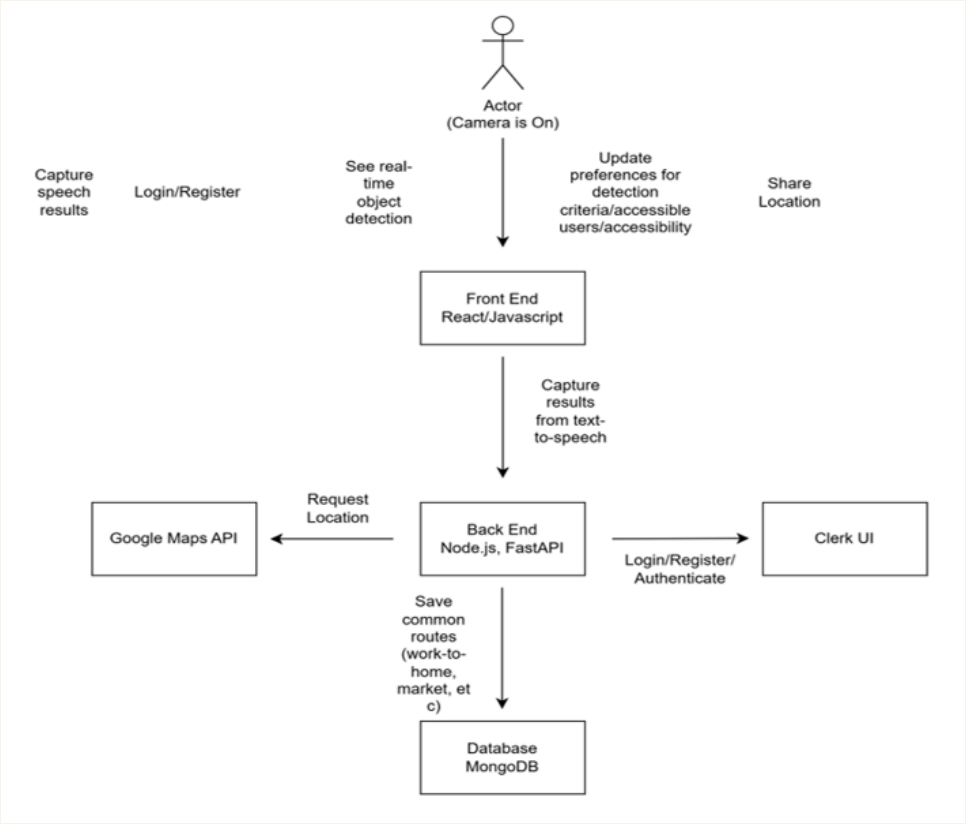
IdeaBoardz: Sprint Retrospective 

AI Model

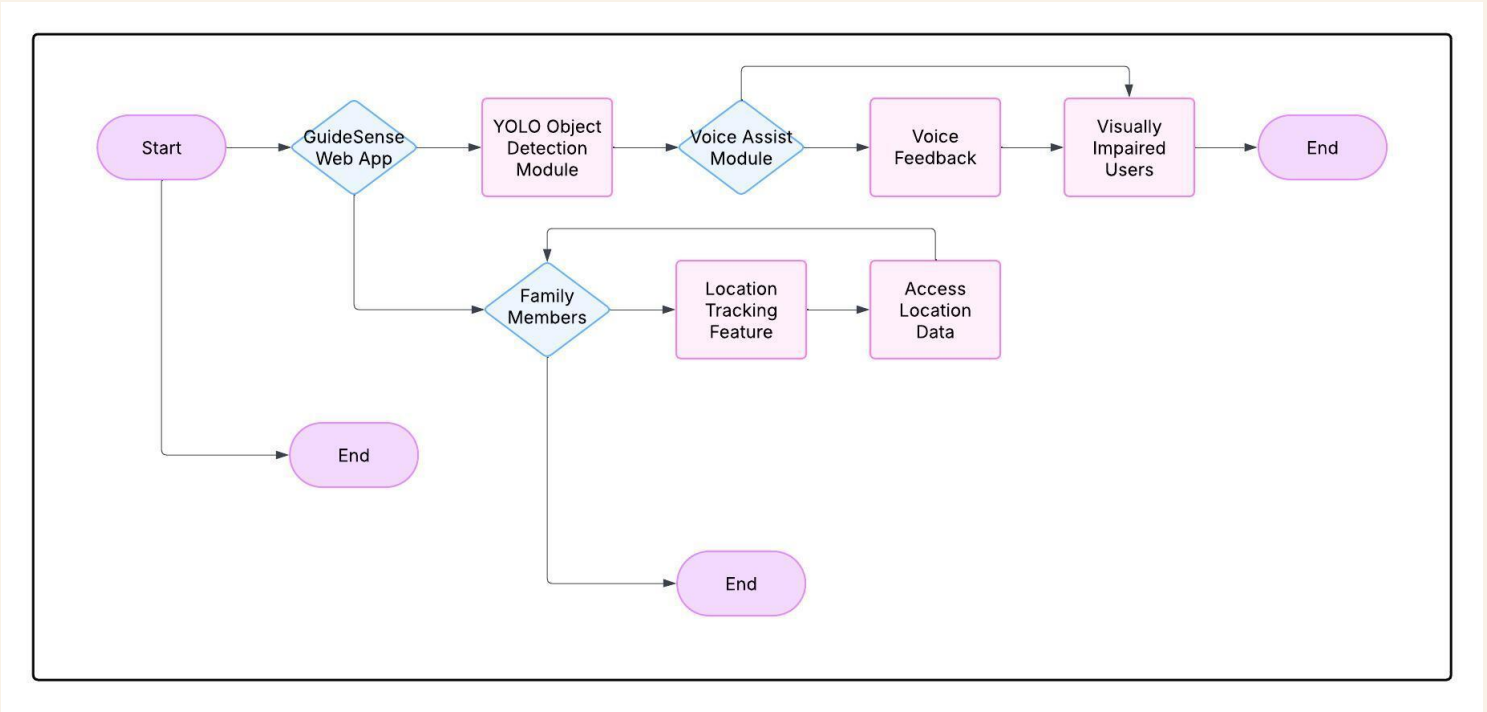
YOLO, which stands for ‘You Only Look Once,’ is a state-of-the-art object detection model known for its real-time speed and accuracy.

- Real-time Speed: Processes the entire image in a single pass, making it one of the fastest models for object detection.
- Grid-Based Detection: Divides the image into a grid, with each cell responsible for detecting objects within it.
- Accurate Bounding Boxes: Predicts precise bounding boxes and confidence scores for each detected object.
- Pre-trained on COCO: Leverages the massive COCO dataset (80 object categories, 200,000+ images) for robust and accurate general object detection
- Refined Results: Uses Non-Maximum Suppression (NMS) to eliminate overlapping detections and ensure only the most accurate results are shown.

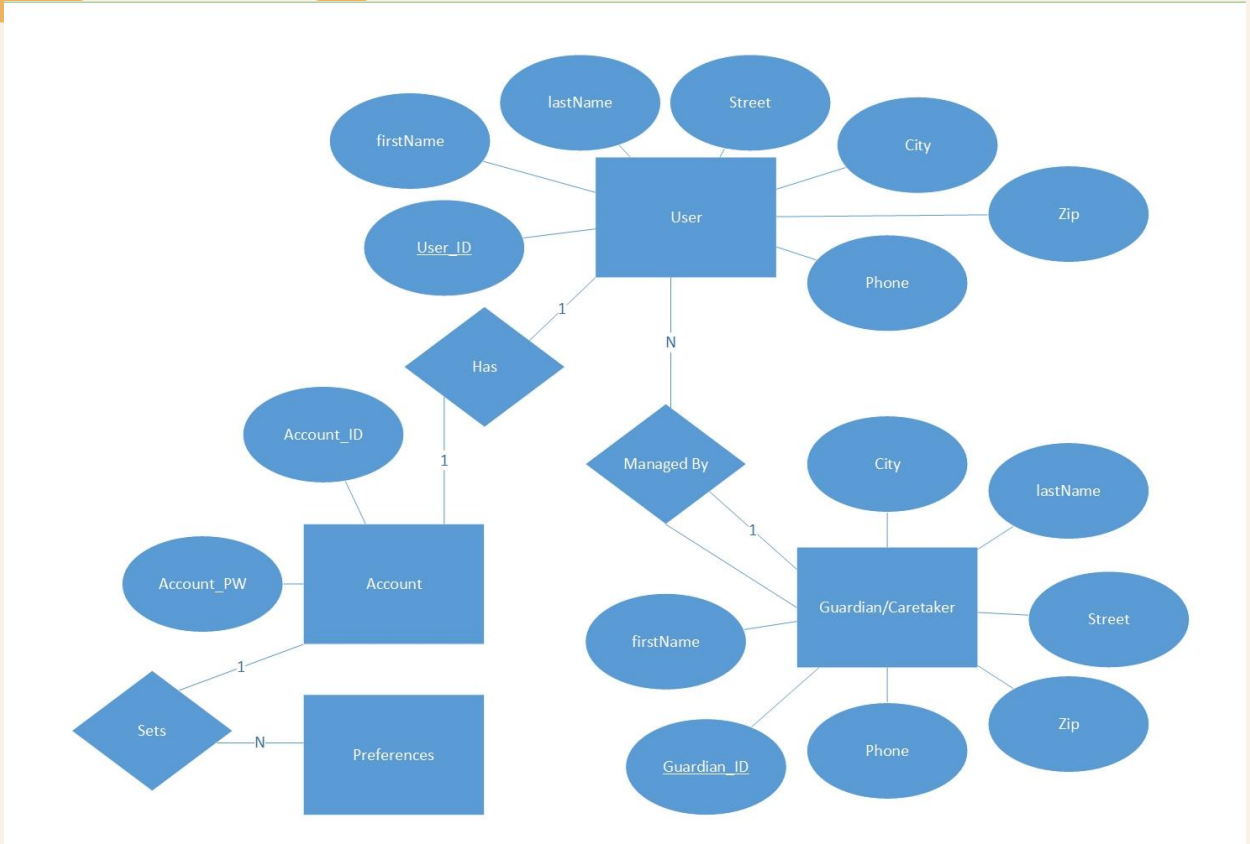
Architecture Diagram



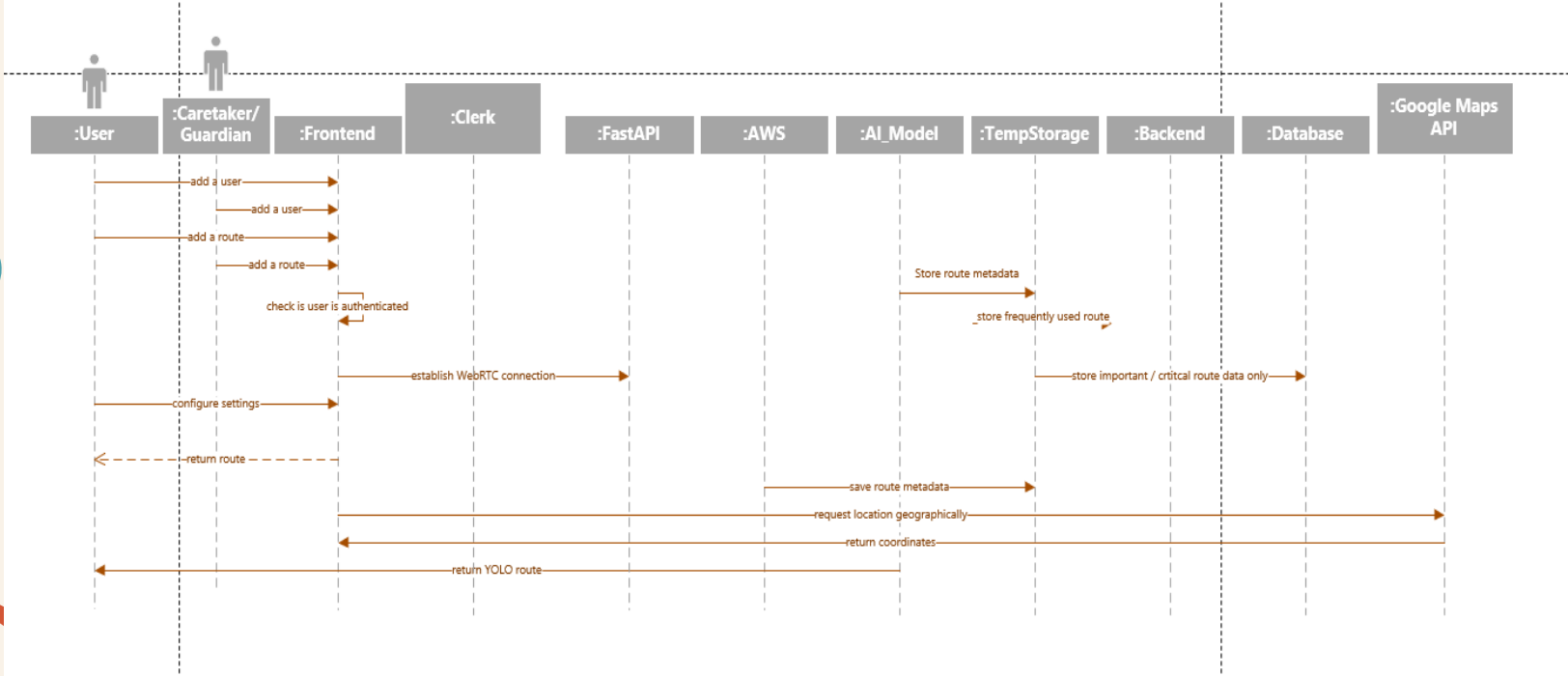
Context Diagram



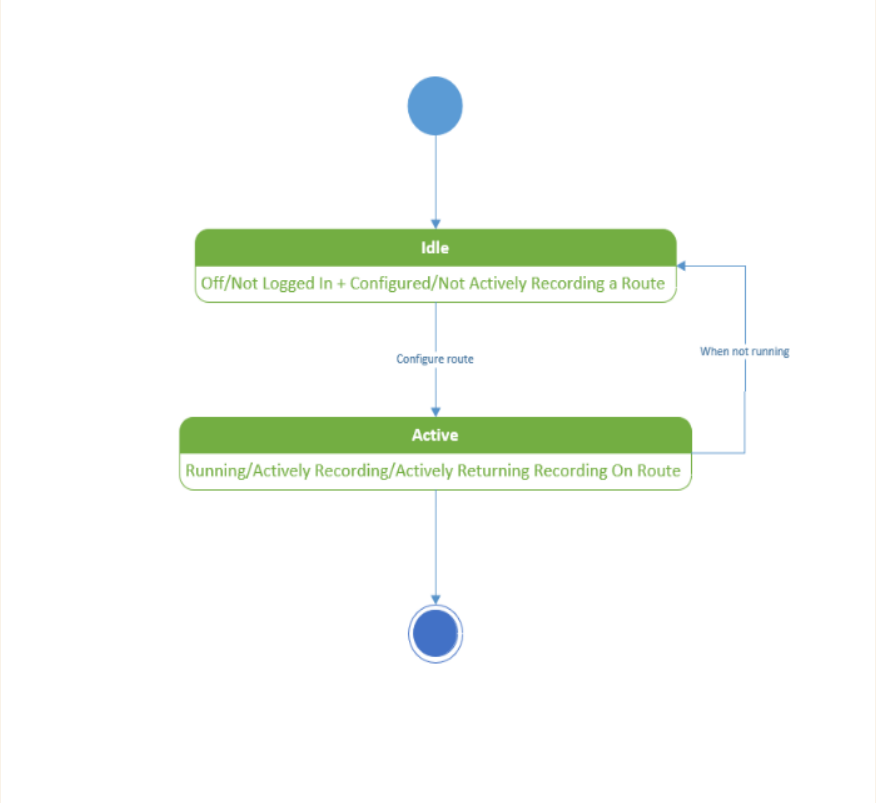
ER Diagram(improved)



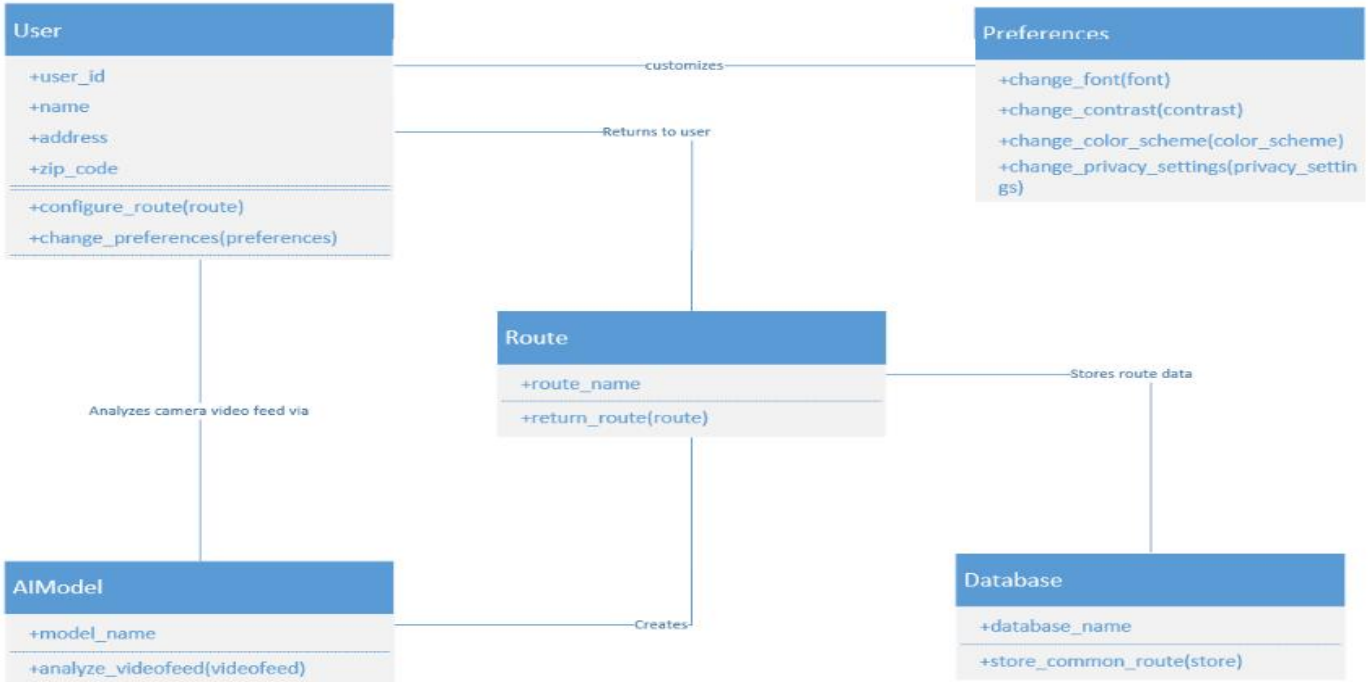
Sequence Diagram



State Diagram



Class Diagram



Sprint 2 Recap

1. Improved detection model
2. Implemented Text-to-Speech functions
3. Completed technical paper (documentation)
4. Integrated Front-end and Back-end

Product Backlog (improved)

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
US1	2	As a visually impaired user, I want to receive audio alerts when potential obstacles are detected through my device camera, so that I can navigate safely in unfamiliar environments.	The FastAPI backend should process frames using the YOLO model and return a JSON object with detected obstacles or important objects (e.g., door, chair, or pedestrian), ensuring minimal delay for real-time assistance.	5
US2	2	As a visually impaired user, I want a single, easy-to-find button that allows me to capture video using my device's camera, so that I can receive audio descriptions of my surroundings to help me navigate safely.	The application must provide an accessible interface that allows visually impaired users to easily capture videos through the phone camera and receive real-time audio feedback on detected objects.	8
US3	2	As a visually impaired user, I want to hear descriptions of objects around me, so that I can understand my surroundings and move around safely.	The system should accurately identify objects in the user's surroundings and provide clear, real-time auditory descriptions through text-to-speech, ensuring the user can understand and navigate their environment safely	8

Product Backlog

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
US4	2	As a visually impaired user, I want to store my data in the app's frontend, so that I can interact with the application to track my information.	The frontend should allow users to store and display data in real-time, reflecting changes in the interface.	3
US5	1	As a visually impaired user, I want to securely sign up, log in, and manage my account, so that I can access personalized features within the app.	The user should be able to securely sign up, log in, manage their account, reset their password, and maintain their session, with proper error handling and redirection after authentication using Clerk	8
US6	1	As a visually impaired user, I want to enter address and receive navigation directions, so that I can easily find the route to my destination.	The user should be able to input the address, and the system should display the route using Google Maps, providing turn-by-turn navigation with estimated time of arrival and distance.	8

Product Backlog

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
TS1	1	As a developer, I need to write test cases for Sprint 0 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 1 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5
TS2	1	As a developer, I need to write test cases for Sprint 1 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 2 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5
TS3	2	As a developer, I need to write test cases for Sprint 2 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 3 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5
TS4	3	As a developer, I need to write test cases for Sprint 3 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 4 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5

Product Backlog (improved)

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
TS4	2	As a developer, I need to write a technical paper documenting the methodology, findings, and outcomes of our project, so that it can be shared with stakeholders, academic peers, or for publication purposes.	The technical paper should comprehensively document the project's problem, methodology, results, analysis, and conclusions, following a clear, structured format with proper citations and adhering to the required submission guidelines.	5
TS5	1	As a developer, I need to set up the MongoDB database according to the ER diagram, so that the data structure aligns with the application requirements, ensuring efficient storage and retrieval of user information.	The database should be set up according to the ER diagram with proper schema, indexing, and relationships	5
TS6	2	As a developer, I need to document the API endpoints and integrations for FastAPI, YOLO object detection, and Google Maps, so that the team can easily understand how to interact with the services and ensure smooth integration.	The API documentation should clearly describe all FastAPI endpoints, YOLO object detection integration, Google Maps functionality, include example requests and responses, and provide authentication details, error codes, and troubleshooting guidelines.	5

Product Backlog (improved)

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
TS7	1	As a developer, I need to research and evaluate different AI models for real-time object detection and explore WebRTC and other methods for connecting the backend and frontend, so that we can select the most suitable technologies for our project's requirements.	The research should evaluate various AI models for real-time object detection and communication methods like WebRTC, comparing performance, scalability, security, and deployment requirements, and provide a recommendation based on the project's needs.	5
TS8	1	As a developer, I need to create an diagrams(Architecture, Sequence, Class, ER, State, Context) for the entire product, which includes all core components, services, and interactions such as the frontend, backend, external APIs, database, and third-party integrations, so that the overall structure and flow of the system are clearly understood and can be implemented effectively.	The architecture diagram should clearly represent all major components (functional and non-functional) of the product, their interactions, data flow, and error handling mechanisms, using a standard tool for easy understanding and team review.	13

Product Backlog (improved)

ID	Sprint	User Story / Technical Story	Acceptance Criteria	SP
TS9	1	As the Project Manager, I want to ensure seamless coordination between technical components and stakeholders, So that the system meets both user needs and technical specifications	The system should facilitate regular communication between stakeholders and the development team, track progress against project goals, ensure alignment between user needs and technical specifications, and maintain clear documentation for all key technical components and requirements	1
TS10	2	As a Developer, I want to review and update the diagrams	Diagrams should be upto date and should be reviewed with the team	1

Sprint Summary

Sprint 1 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	status
US5	As a vision impacted user, I want to securely sign up, log in, and manage my account so that I can access personalized features.	The user should be able to securely sign up, log in, manage their account, reset their password, and maintain their session, with proper error handling and redirection after authentication using Clerk	8	Completed
US6	As a vision impacted user, I want to enter a source and destination and receive navigation directions so that I can easily find the best route to my destination.	The user should be able to input a source and destination address, and the system should display the best route using Google Maps, providing turn-by-turn navigation with estimated time of arrival and distance.	8	Completed
TS1	As a developer, I need to write test cases for Sprint 0 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 1 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5	Completed

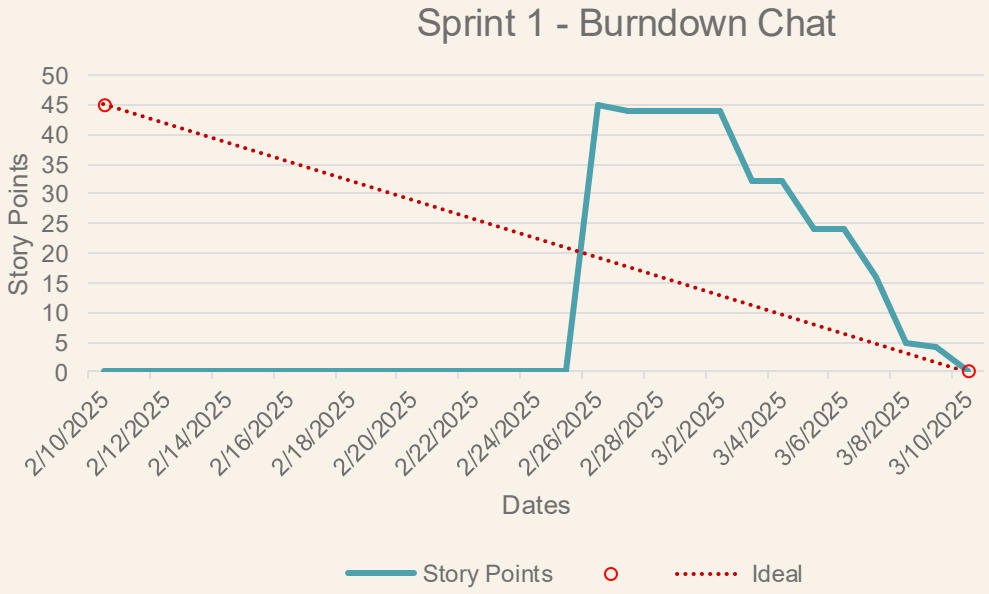
Sprint 1 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	Status
TS7	As a developer, I need to research and evaluate different AI models for real-time object detection and explore WebRTC and other methods for connecting the backend and frontend, so that we can select the most suitable technologies for our project's requirements.	The research should evaluate various AI models for real-time object detection and communication methods like WebRTC, comparing performance, scalability, security, and deployment requirements, and provide a recommendation based on the project's needs.	5	Completed
TS8	As a developer, I need to create an diagrams(Architecture, Sequence, Class, ER, State, Context) for the entire product, which includes all core components, services, and interactions such as the frontend, backend, external APIs, database, and third-party integrations, so that the overall structure and flow of the system are clearly understood and can be implemented effectively.	The architecture diagram should clearly represent all major components (functional and non-functional) of the product, their interactions, data flow, and error handling mechanisms, using a standard tool for easy understanding and team review.	13	Completed

Sprint 1 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	Status
TS7	As a developer, I need to research and evaluate different AI models for real-time object detection and explore WebRTC and other methods for connecting the backend and frontend, so that we can select the most suitable technologies for our project's requirements.	The research should evaluate various AI models for real-time object detection and communication methods like WebRTC, comparing performance, scalability, security, and deployment requirements, and provide a recommendation based on the project's needs.	5	Completed
TS9	As the Project Manager, I want to ensure seamless coordination between technical components and stakeholders, So that the system meets both user needs and technical specifications.	The system should facilitate regular communication between stakeholders and the development team, track progress against project goals, ensure alignment between user needs and technical specifications, and maintain clear documentation for all key technical components and requirements	1	Completed

Total Committed Story Points on Sprint 1: 45



Sprint 1 committed 45 story points and completed 45 story points.

Sprint 2 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	status
US1	As a visually impaired user, I want to receive real-time alerts when objects are detected through my phone camera	The FastAPI backend should process frames using the YOLO model and return a JSON object with detected obstacles or important objects (e.g., door, chair, or pedestrian), ensuring minimal delay for real-time assistance.	5	Completed
US2	As a visually impaired user, I want an intuitive interface that allows me to easily capture images using my phone camera, so that the system can analyze the surroundings and help me navigate safely.	The application must provide an accessible interface that allows visually impaired users to easily capture images through the phone camera and receive real-time audio feedback on detected objects.	8	Completed
US3	As a visually impaired user, I want the system to describe detected objects in my surroundings through text-to-speech, so that I can receive real-time auditory feedback and safely navigate my environment.	The system should accurately identify objects in the user's surroundings and provide clear, real-time auditory descriptions through text-to-speech, ensuring the user can understand and navigate their environment safely	8	Completed

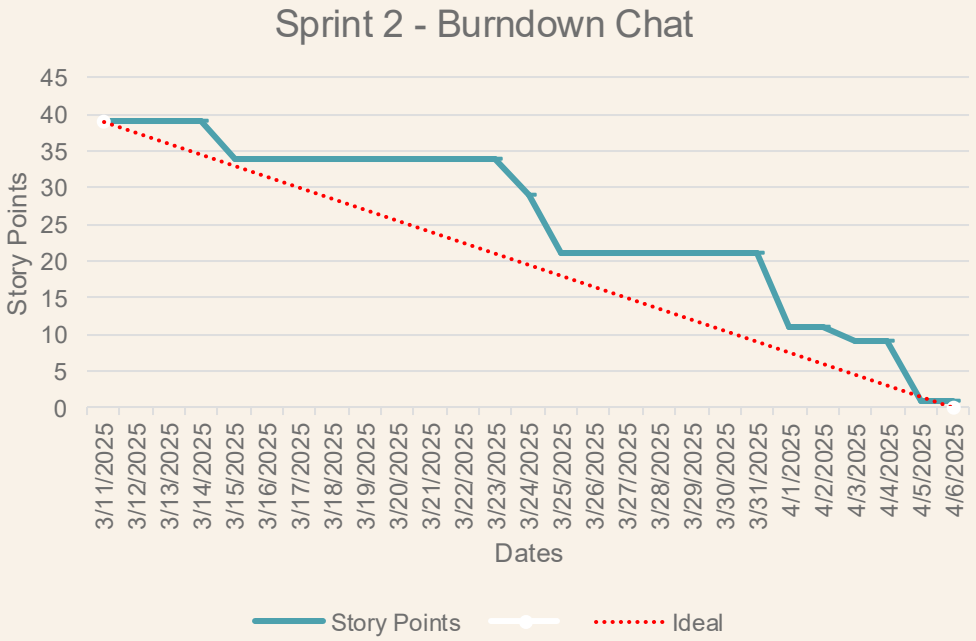
Sprint 2 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	status
TS6	As a developer, I need to document the API endpoints and integrations for FastAPI, YOLO object detection, and Google Maps so that the team can easily understand how to interact with the services and ensure smooth integration	The API documentation should clearly describe all FastAPI endpoints, YOLO object detection integration, Google Maps functionality, include example requests and responses, and provide authentication details, error codes, and troubleshooting guidelines.	5	Completed
TS10	As a Developer, I want to review and update the diagrams	Diagrams should be upto date and should be reviewed with the team	1	Inccomplete
TS4	As a developer, I need to write a technical paper documenting the methodology, findings, and outcomes of our project so that it can be shared with stakeholders, academic peers, or for publication purposes.	The technical paper should comprehensively document the project's problem, methodology, results, analysis, and conclusions, following a clear, structured format with proper citations and adhering to the required submission guidelines.	5	Completed

Sprint 2 User Stories

ID	User Story / Technical Story	Acceptance Criteria	SP	status
TS3	As a developer, I need to write test cases for Sprint 2 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 3 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5	Completed
US4	As a user, I want to store my data in the app's frontend and see the results reflected immediately so that I can interact with the application in real-time and track my information.	the frontend should allow users to store and display data in real-time, reflecting changes immediately in the app interface.	2	Completed

Total Committed Story Points on Sprint 2: 39



Sprint 2 committed 39 story points and completed 38 story points.

Sprint 3 – Sprint Backlog

ID	User Story / Technical Story	Acceptance Criteria	SP	Status
TS12	As a developer, I need to create an AWS account for the project, so that we can deploy and host backend services, machine learning models, and other cloud-based components in a secure and scalable environment.	An AWS account should be successfully created, with access credentials securely stored and shared with authorized team members.	2	Completed
TS13	As a developer, I need to deploy the backend and model code on an AWS EC2 instance, so that the application can run continuously on the cloud and handle real-time requests from users.	The backend and model code should be successfully deployed and running on the EC2 instance, accessible via the instance's public IP or domain, and verified through test requests.	8	Completed
TS14	As a developer, I need to create a comprehensive deployment and installation manual for the project so that other team members and future developers can easily set up the environment, deploy the application, and maintain the system.	The manual should include clear, step-by-step instructions for environment setup, dependencies installation, code deployment, and running the application, and should be accessible to the entire team.	8	Completed

Sprint 3 – Sprint Backlog

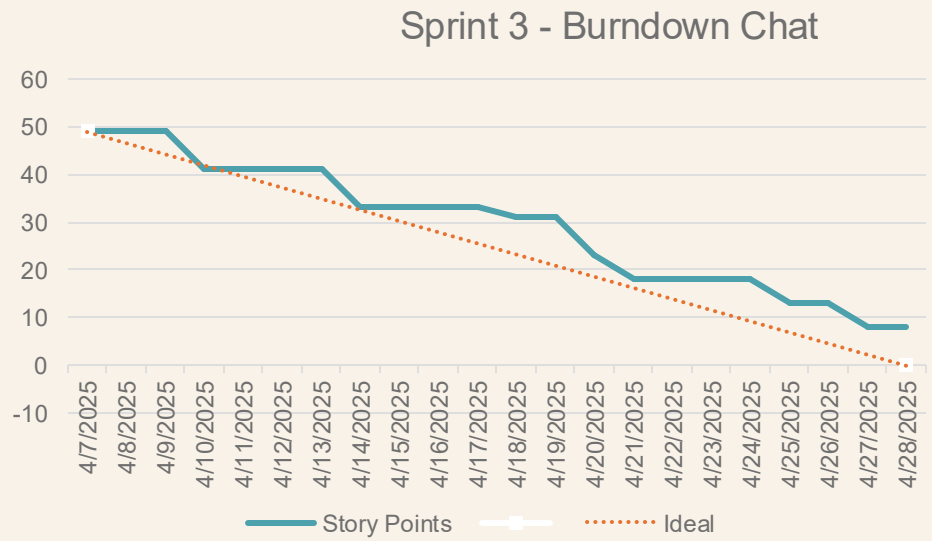
ID	User Story / Technical Story	Acceptance Criteria	SP	Status
TS15	As a developer, I need to deploy the project's database on an AWS RDS server so that we can ensure secure, scalable, and managed database access for our application across all environments.	The database should be successfully created and deployed on the RDS server, with appropriate security groups, credentials configured, and connectivity verified from the application backend.	3	Incomplete
US7	As a visually impaired user, I want to receive real-time audio updates about my surroundings so that I can navigate safely and independently.	The frontend and backend should work together to provide real-time object detection, user-specific responses, and data storage, with all core features functioning smoothly and verified end-to-end through user interaction.	8	Completed
US8	As a visually impaired user, I want a simple and accessible UI to input my source and destination so that I can easily receive navigation directions using Google Maps and move safely and independently.	The UI should allow users to enter source and destination locations, and upon submission, display route information and initiate navigation using Google Maps integration in an accessible format.	5	Completed

Sprint 3 – Sprint Backlog

ID	User Story / Technical Story	Acceptance Criteria	SP	Status
TS4	As a developer, I need to write test cases for Sprint 3 to ensure the implemented features are working as expected and meet the acceptance criteria.	All user stories and technical requirements from Sprint 4 should have corresponding test cases that cover positive, negative, and edge case scenarios.	5	Completed
TS16	As a developer, I need to draft a technical paper that documents the design, architecture, implementation details, challenges faced, and results of the project so that we can clearly present our work for academic or professional review.	The technical paper should include sections on problem statement, methodology, system architecture, technology stack, implementation details, challenges, results, and future work, and it should be well-structured and ready for submission or presentation.	5	Completed
US9	As a visually impaired user, I want to be able to connect with other users and grant them access to control my app temporarily, so that they can assist me in navigating or using certain features remotely when I need help.	The system should allow a user to securely share access with a trusted user, enabling them to view and interact with key controls in real-time, with the ability to revoke access at any time.	5	Incomplete

Total Committed Story Points on Sprint 3: **49**

Sprint 3 – Burndown Chart(improved)



Sprint 3 committed 49 story points and completed 41 story points.

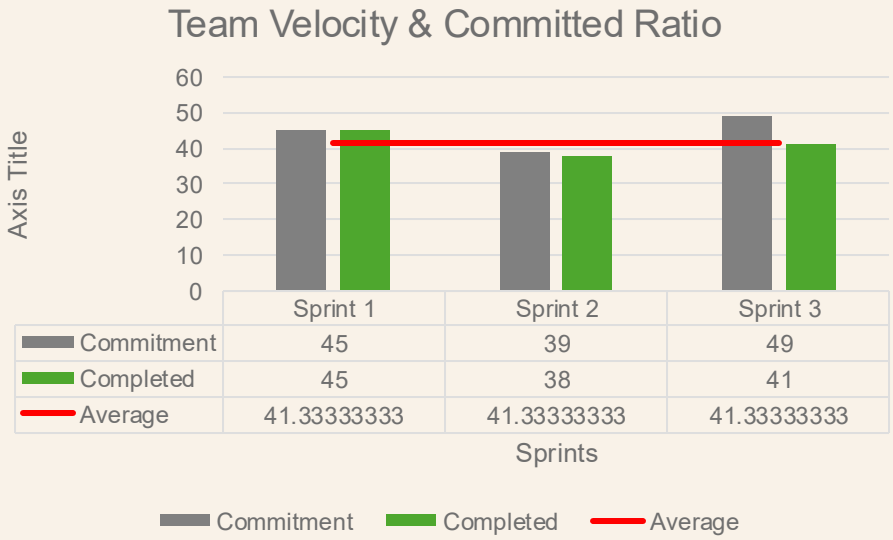
Sprint 3 Test Cases (improved)

ID	Story ID	Test Case	Expected Result	Status
1	TS12	Verify that a developer can successfully create a new AWS account with root user credentials	A new AWS account is created with a confirmed email address, secure password, and multi-factor authentication (MFA) enabled for the root user. The developer receives confirmation of successful account creation and can access the AWS Management Console.	Passed
2	TS13	Verify that the backend and model code can be successfully deployed on an appropriately configured EC2 instance	EC2 instance launches with proper configuration (correct type, security groups, IAM role). Backend and model code deploy successfully with all dependencies. Application initializes without errors, with logs confirming proper startup of all components	Passed
3	TS14	Verify that the deployment and installation manual contains all necessary steps and information	Manual includes complete setup instructions for environment configuration, application deployment, and system maintenance with accurate commands, dependencies, and configuration settings.	Passed

Sprint 3 Test Cases (improved)

ID	Story ID	Test Case	Expected Result	Status
4	US7	Verify that the application correctly identifies and communicates relevant objects and obstacles in the user's surroundings	Application accurately detects obstacles, landmarks, and potential hazards in the user's environment and provides clear, timely audio descriptions that match what's actually present.	Passed
5	US7	Verify that audio updates are delivered in a clear, timely, and non-disruptive manner	Audio notifications are delivered with appropriate volume, clarity, and frequency. Critical alerts are prioritized, and the user can easily understand the information without becoming overwhelmed or confused by excessive updates.	Passed
6	US8	Verify that visually impaired users can easily input their source and destination using accessibility features	UI elements are properly labeled for screen readers, voice input works correctly, and all input fields are navigable using keyboard shortcuts. Users can successfully enter locations without visual cues.	Passed
7	US8	Verify that navigation directions from Google Maps are delivered in an accessible format	Navigation directions are converted to clear audio instructions that provide timely guidance at appropriate intervals. Turn-by-turn directions include distance information and warning for upcoming turns.	Passed

Sprint 3 – Team Velocity & Committed Ratio



Out of a total 49 committed story points, 41 story points were completed. The committed ratio is 84**%.**

Retrospective – Sprint 3

What went well?

- Integrated API on time
- Team meetings were short but effective
- Clear prioritization of tasks and features
- Comprehensive documentation covering all use cases
- Successfully implemented professor's feedback
- Completed core development features on schedule
- Effective code synchronization and review process before merging code

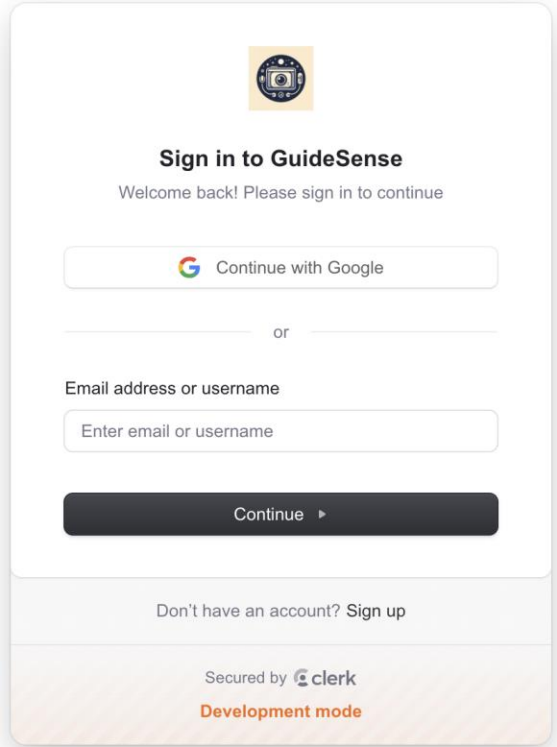
What need to improve?

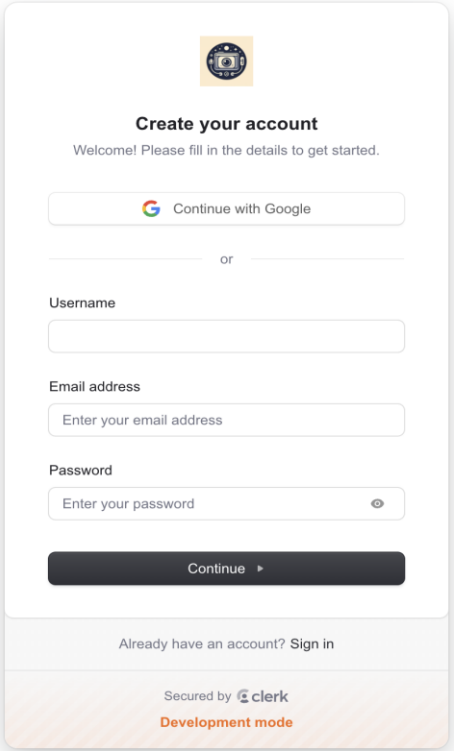
- Better planning and timeline management
- Ensuring everyone joins sprint meetings regularly
- Avoiding last-minute merging of features
- Setting up initial API infrastructure earlier in the process
- Standardizing the installation manual to make it more simplistic with clearer steps

Action Items

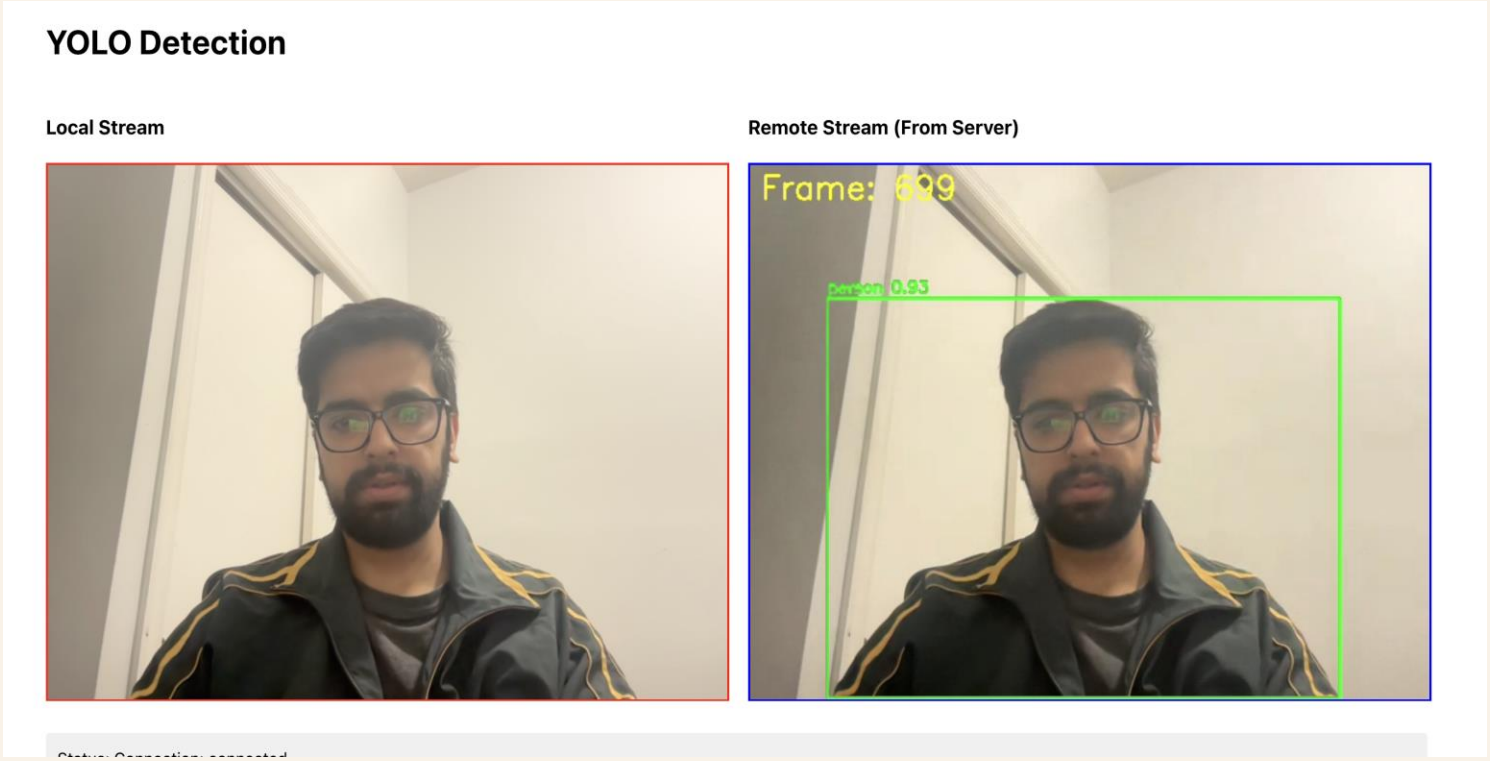
- Implement weekly timeline check-ins to assess progress against milestones
- Send calendar invites with automatic reminders 1 hour before meetings
- Establish a code freeze policy 24-48 hours before sprint end
- Test installation instructions with a team member who wasn't involved in writing them

Project Demo

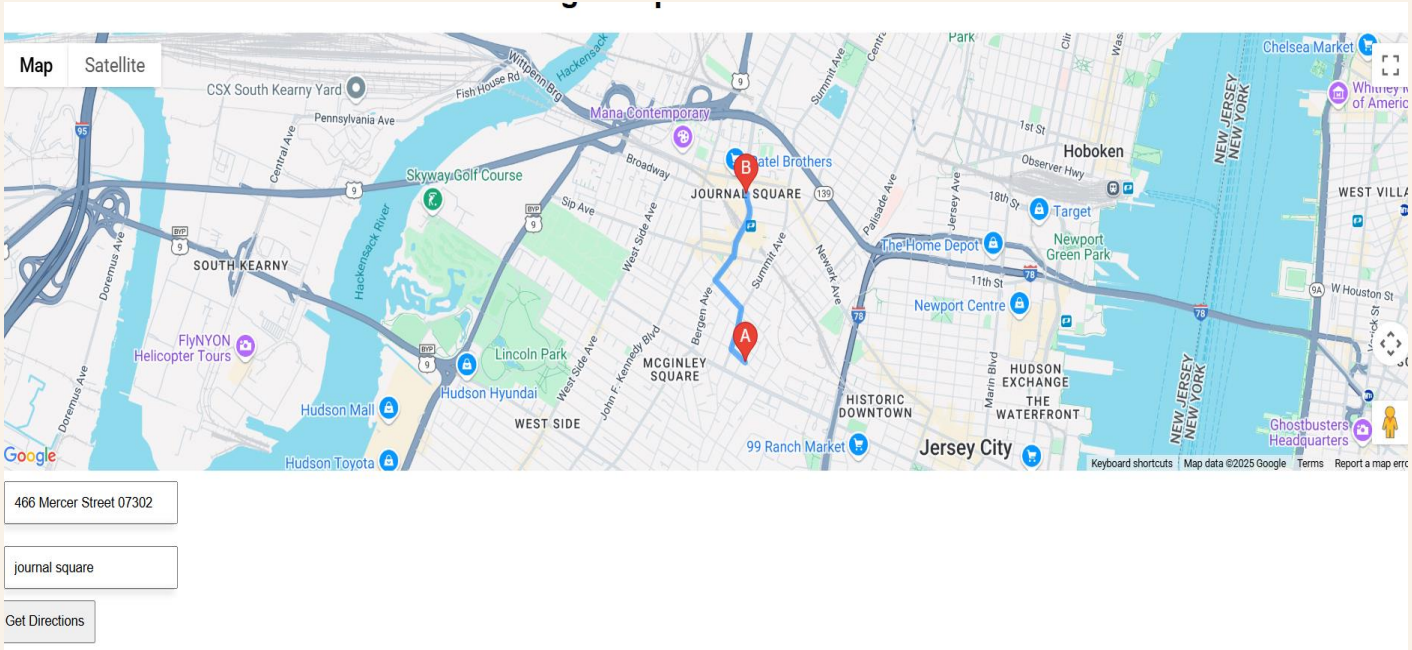




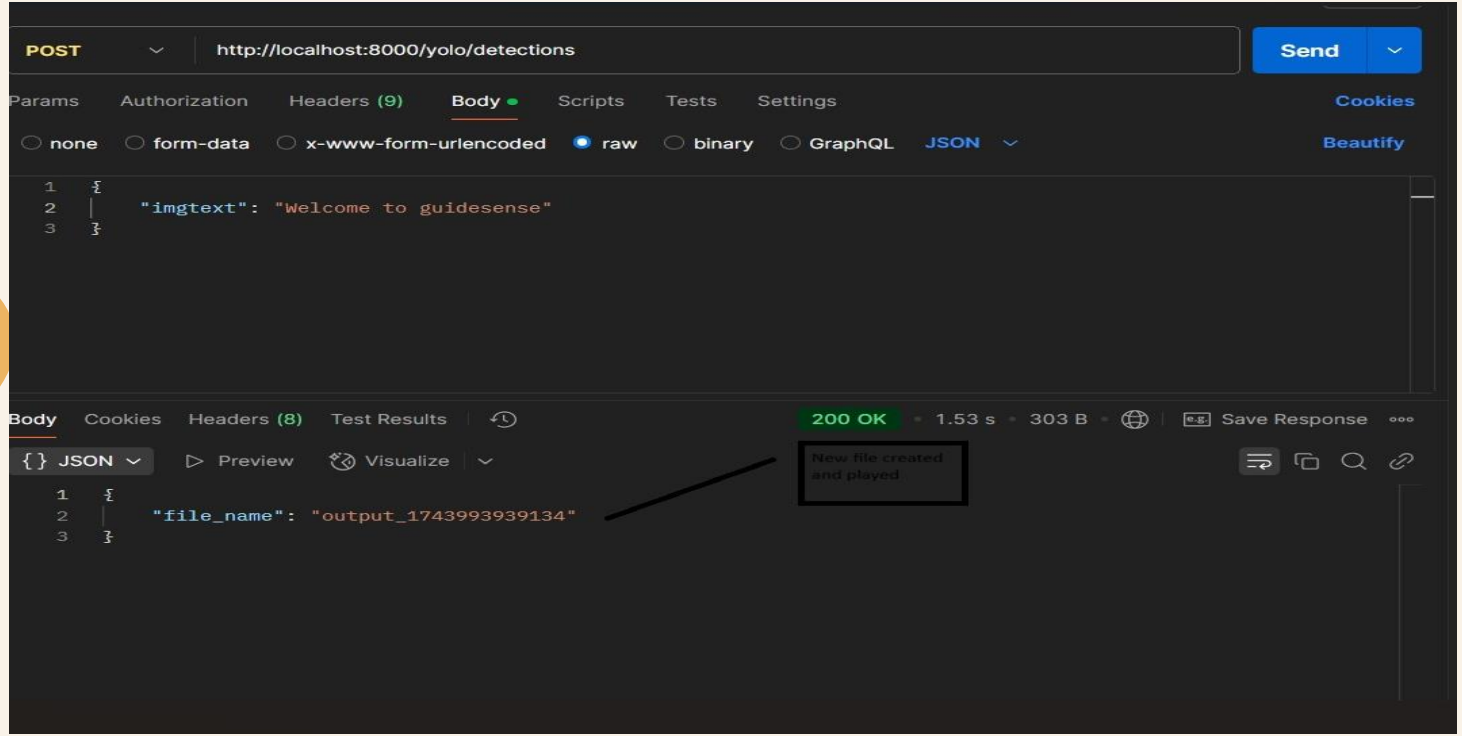
Project Demo



Project Demo



Project Demo



Backend API

Detections API (to detect messages from frontend and pass it to text to speech API)

```
detections: async (req, res) => {
    try {
      console.log(req.body);

      // Extract detected objects and add to accumulated set
      if (req.body.detections && req.body.detections.length > 0) {
        req.body.detections.forEach(det => {
          accumulatedDetections.add(det.class_name);
        });
      }

      const currentTime = Date.now();
      // Check if cooldown period has passed
      if (currentTime - lastSpeechTime >= speechCooldown && accumulatedDetections.size > 0) {
        // Find new objects that weren't in the last announcement
        const newObjects = [...accumulatedDetections].filter(obj => !lastAnnouncedDetections.has(obj));
        // Find objects that disappeared
        const removedObjects = [...lastAnnouncedDetections].filter(obj => !accumulatedDetections.has(obj));
```

Backend API

```
// Only generate speech if there are changes
        if (newObjects.length > 0 || removedObjects.length > 0) {
          let detectionText = "";

          if (newObjects.length > 0) {
            detectionText += ` I now see ${newObjects.join(', ')}. `;
          }

          if (removedObjects.length > 0) {
            detectionText += ` I no longer see ${removedObjects.join(', ')}.`;
          }

          // Generate speech
          const result = await yoloService.detections(JSON.parse('{"imgtext":"' + detectionText + '"}'));

          // Update last announced set to current set
          lastAnnouncedDetections = new Set(accumulatedDetections);

          // Reset the timer and accumulated detections
          lastSpeechTime = currentTime;

          return res.status(200).json({file_name: result});
```

Backend API

```
} else {
          // If we have the same objects as before, don't generate speech
          lastSpeechTime = currentTime; // Still reset the timer
          return res.status(200).json({message: "No change in detected objects"});
        }
      } else {
        // If we're in cooldown period, just acknowledge
        return res.status(200).json({message: "Detection added to queue"});
      }
    } catch (error) {
      return res
        .status(500)
        .json({ message: "An error occurred" });
    }
  },
```

Project Wikipage Link

https://github.com/htmw/2025S-Power-Anger/wiki

Live Application Demo