SPEAKEZ

AI Language Translator

Kaali Lovell | Pavan Polireddy | Rakesh Tirumala

Steve St Fleur | Tanvi Prakash Gavali | Venkat Sai Pallapu

AGENDA

1- Sprint 0 Recap	2- Project Overview	3- Backlogs	4- Team Metrics	5- Product
• Team Roles & Responsibilities • Improvements • Project Description • Team Working Agreement • Personas	• MVP • Languages & Tools • Algorithms • Diagrams	• User Stories • Acceptance Criteria • Sprint 1 Stories • Test Cases • Story Completion	• Team Velocity • Burndown Charts • Retrospective	• Sprint 2 Stories • Application Screenshots • API • Live Application Demo

TEAM ROLE & RESPONSIBILITIES

Rakesh   Tirumala
Team Lead/ Developer

Tanvi   Prakash   Gavali
Scrum Master/ Developer

Kaali   Lovell
Tester/ Product Designer

TEAM ROLE & RESPONSIBILITIES

Steve St Fleur
Developer

Venkat Sai Pallapu
Developer

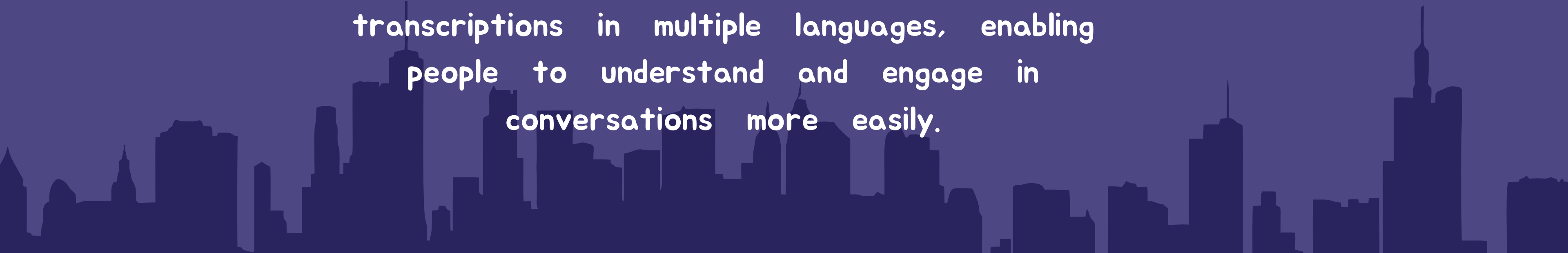**Pavan Polireddy**
Developer

IMPROVEMENTS

- Team Roles and Responsibilities:
  - updated the Developers' roles.
- Project Description:
  - the project description chart was added.
- Team Working Agreement:
  - expanded on the DoD (Release DoD, MVP DoD).
- Sprint Schedule:
  - add small milestones for the project to the deadline dates.
- Technologies:
  - specified the names under the technologies we plan to use.
- WikiPage:
  - added email link for team members.
  - added mp4 download of recording to the WikiPage.
  - added the files to the repository and then link to the WikiPage.

PROBLEM STATEMENT

Understanding foreign languages is often a challenge, as each language has its own unique structure, nuances, and meanings. Effective communication with native speakers becomes difficult without proper understanding, yet mastering every language is nearly impossible. To address this issue, we aim to create an application that helps bridge the communication gap by converting audio into transcriptions in multiple languages, enabling people to understand and engage in conversations more easily.

PROJECT DESCRIPTION

Project Name:	SpeakEZ
Team:	Bytes
Project Description:	For users who need to communicate across different languages, the AI-powered transcription and translation web application is a seamless solution that converts audio into multilingual text that enables users to understand conversations effortlessly, regardless of language barriers. Unlike current tools that may require manual translation or lack real-time accuracy, our application integrates advanced AI model to provide precise and real-time transcription and translation, improving cross-language communication efficiency across cultures and regions.
Benefit Outcomes:	Better Communication, Easy Accessibility
Github Link:	https://github.com/htmw/2024F-Bytes/wiki

TEAM WORKING AGREEMENT

Communication:

- The team will communicate with each other through Zoom and WhatsApp.
- Google Docs will be utilized for collaborative document editing, sharing resources, and taking notes during team discussions.
- Team members should alert the team in advance if they are unable to join a Zoom meeting.
- Team members are expected to provide regular updates on task progress and promptly communicate any obstacles or delays to the group.
- Team members are encouraged to actively participate in discussions, providing feedback and support as needed.
- During all discussions, team members should listen carefully, speak clearly, and stay on topic to avoid confusion
- Each team member should complete all assigned work before the deadline. If they are struggling with a task, they should contact the team as soon as possible to resolve the issue. Failure to complete work before the deadline can negatively affect the team's project timeline.

TEAM WORKING AGREEMENT

Work Division & Participation:

- Jira, and GitHub will be used to track and divide all tasks.
- The entire project should be divided into equal parts, and all team members should be given equal responsibility.
- If a team member is not participating to the level the other members are, the team leader is permitted to assign necessary tasks to that member.
- Each team member is responsible for their assigned tasks and their contribution to the overall success of the project.

TEAM WORKING AGREEMENT

Meetings:

- Regular team meetings will be scheduled twice a week on Sundays at 2:00 pm and Thursdays at 5:30 pm via zoom to review individual and collective progress. Failing to join these meetings will negatively affect the group.
- If unable to attend a scheduled meeting, team members should inform the group in advance.
- In urgent situations, additional meetings may be arranged with at least 24 hours' notice.
- Meeting agendas will be sent at least 24 hours prior to the meeting.
- Meetings will commence and conclude on time, with members expected to join 2 minutes prior to the start.
- All team members are required to have their Zoom cameras open during the meetings. They are required to show respect, actively listen, and participate by sharing feedback or suggestions.
- Team agrees to follow these agile practices throughout the project:
  - Meet at the beginning of each sprint to discuss tasks for the upcoming sprint.
  - Meet at the end of each sprint to present the completed work and gather feedback.

TEAM WORKING AGREEMENT

Respect and Professionalism:

- All communication and interaction within the team will be conducted with respect; diverse viewpoints and inputs should be appreciated.
- Feedback should be constructive and centered on the project and not on personal traits.
- If a member disagrees with a decision, they can request a review at the next team meeting.
- The team aims to create a welcoming space where everyone feels comfortable sharing their ideas
- Any disagreements should be addressed openly and respectfully within the team. If an internal resolution is not possible, the team will seek advice from the professor.

TEAM WORKING AGREEMENT

Definition of Done (Release DoD) for the Language Translator App:

- The translation and transcription feature works as intended and meets all acceptance criteria and requirements.
- The app successfully translates text between the specified languages without errors.
- The app successfully transcribes speech to text.
- Codes have been reviewed and approved by all team members.
- The app has been manually tested for usability and performance.
- Codes are documented with comments where necessary.
- User guide and technical documentation have been updated accordingly.
- The app is deployable and runs without issues.
- All user stories and tasks for the current sprint have been completed.
- The app contains all necessary features required for the MVP.

TEAM WORKING AGREEMENT

Definition of Done (MVP DoD) for the Language Translator App:

- The application can transcribe the given audio into text without error.
- The application can successfully translate the transcribed audio to the specified language without errors.
- A user is able to upload or delete an audio file to the application without errors.
- Codes have been reviewed and approved by all team members.
- The app has been manually tested for usability and performance.
- Codes are documented with comments where necessary.

TEAM WORKING AGREEMENT

Team Member	Email
Kaali Lovell	kl18422n@pace.edu
Tanvi Prakash Gavali	tg17504n@pace.edu
Steve St Fleur	ss79968n@pace.edu
Pavan Polireddy	jp04368n@pace.edu
Rakesh Tirumala	rt33583n@pace.edu
Venkat Sai Pallapu	vp36147n@pace.edu

PERSONAS

JOSÉ RODRIGUEZ, DOCKS WORKER

Demographics:
- Age: 53
- Gender: Male
- Location: Louisiana, USA
- Education: High school diploma
- Experience: 15+ years

Skills and Qualifications:
- Forklift License
- Management in handling and securing cargo using forklifts
- Excellent team worker, organized, and physically capable

Goals and Needs:
- José is originally from Mexico but has relocated to USA for better working opportunities for his children. He can understand very little English as most of his coworkers speak Spanish as well. During his old age, he started having frequent visits to the doctors where they speak in English with him, which he has difficulty understanding. He does not want to burden his children with the task of translating, and wants to retain confidentiality between himself and his doctor. It is also expensive to hire professional translators – beyond his budget. Jose is looking for an application where he can converse with the doctors without any communication hindrances, and keep transcripts as part of his medical record – both before and after translation.

ANJALI CHITRE, LAWYER

Demographics:
- Age: 27
- Gender: Female
- Location: Edinburgh, UK
- Education: LLM in International Law
- Experience: 4+ years

Skills and Qualifications:
- Skilled in conducting in-depth research regarding international legal issues
- Proficient in explaining legal concepts and mediating in legal cases
- Professional, showcasing confidence and problem solving skills.

Goals and Needs:
- Anjali's job is to provide legal advice to multinational corporations who need to comply with different regulations for both their U.K and Swedish subsidiaries. To not disclose business information and legal strategies to a 3rd party, Anjali would like an application which can transcribe and translate discussion during meetings without breaching confidentiality or having to create further confidentiality agreements with translators.

NICOLE SCHULER, EXECUTIVE ADMINISTRATIVE ASSISTANT

Demographics:
- Age: 27
- Gender: Female
- Location: Toronto, CA
- Education: Bachelors of Business Administration (BBA)
- Experience: 1 years

Skills and Qualifications:
- Proficient in Excel, Microsoft Office, Outlook, Google Workspace
- Certified Administrative Professional (CAP)
- Excellent organizational, administrative, and interpersonal skills

Goals and Needs:
- Nicole is originally from NYC but has relocated to Canada for work. Nicole has to attend daily meetings and transcribe the meetings for her higher-ups. Some of her co-workers switch between French and English in the meetings which frustrates Nicole since her native language is English and she has no experience with French. Nicole is looking for a tool that will help her efficiently transcribe the meetings and also assist with translating the French terms.
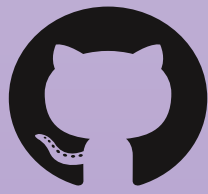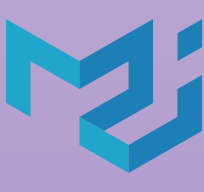
MVP DESIGN

ALGORITHMS

The application should be able to take audio as input and convert it into the desired language. To achieve this, we need an AI model for automatic speech recognition and speech translation. So, the most suitable neural network to use in order to perform this task is a Transformer.
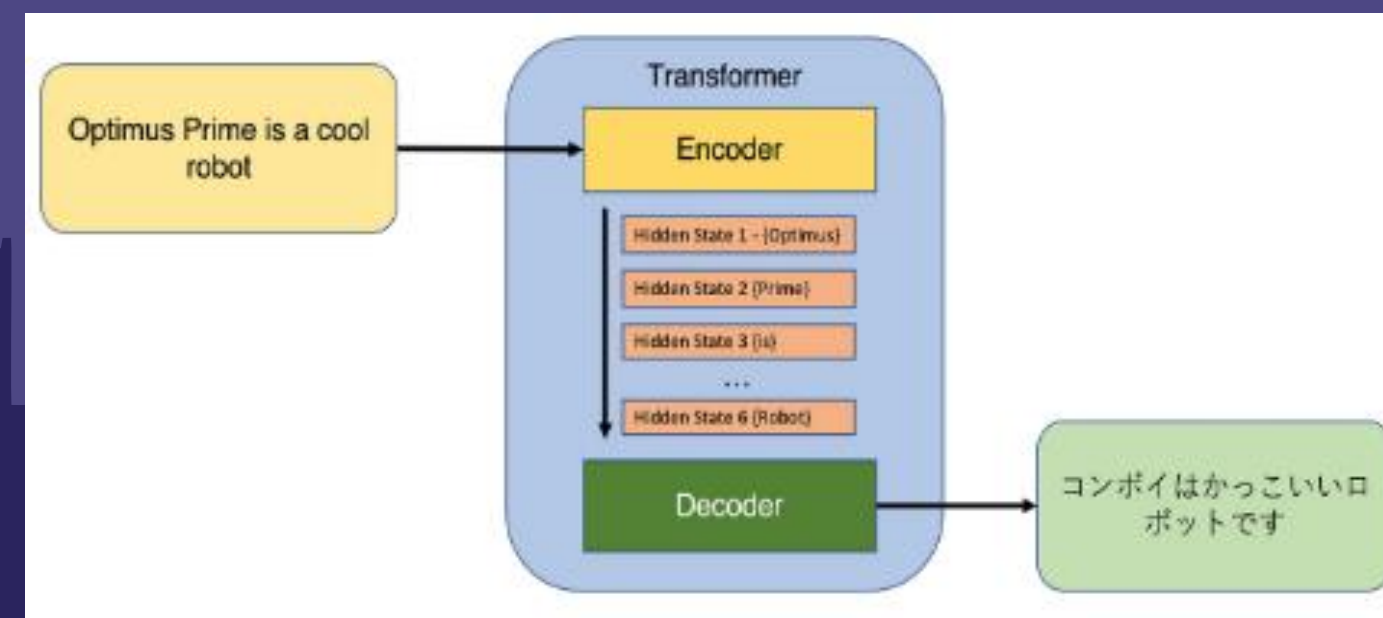
ALGORITHMS

A transformer is a neural network architecture especially designed for processing sequential data like natural language processing (NLP). They usually use an encoder-decoder structure but make use of attention mechanism in order to capture relationships between elements in a sequence. Unlike RNNs, transformers process data in parallel which makes them faster and more scalable.
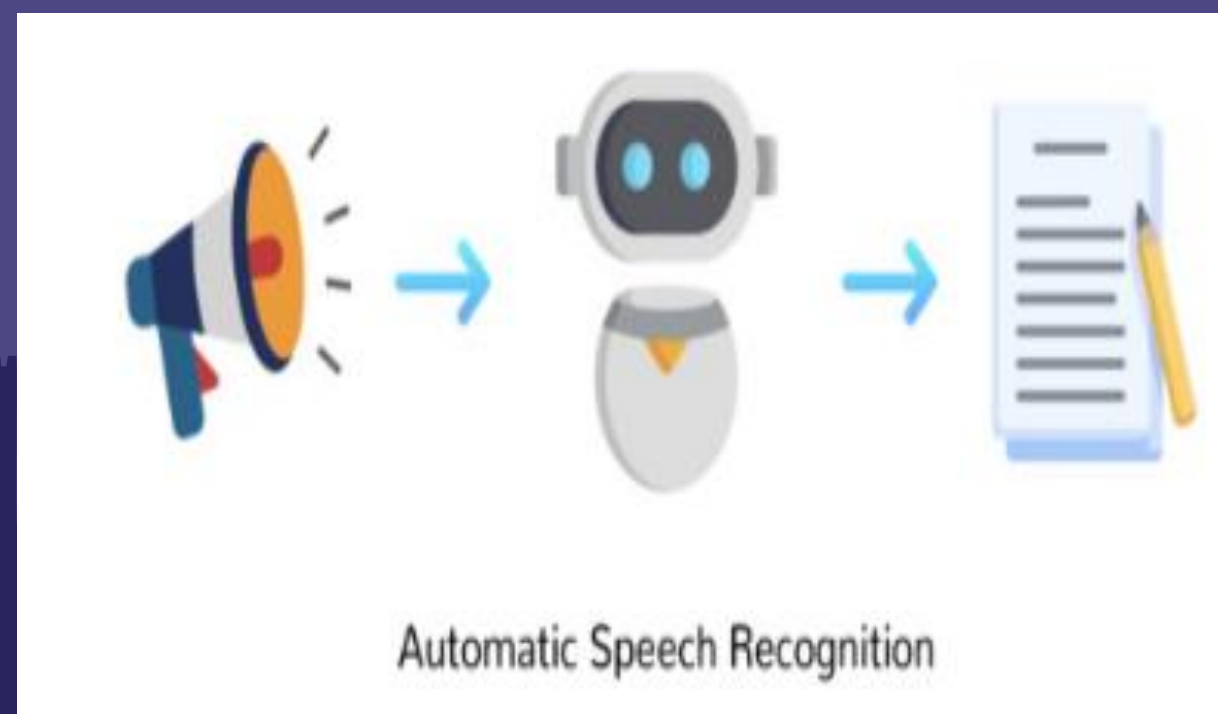
ALGORITHMS

In the encoder-decoder structure, the encoder processes the input data, which can be text, audio, etc., and generates encodings. The decoder takes these encodings, uses the derived context, and generates the output sequence.
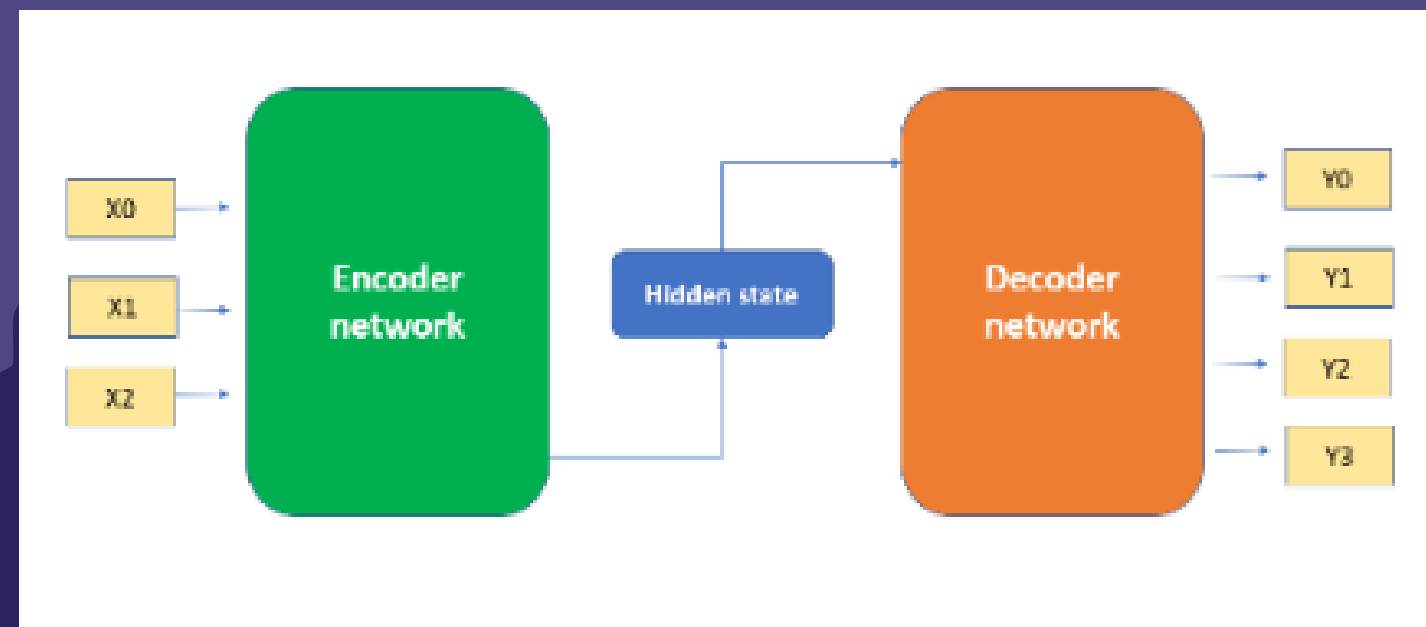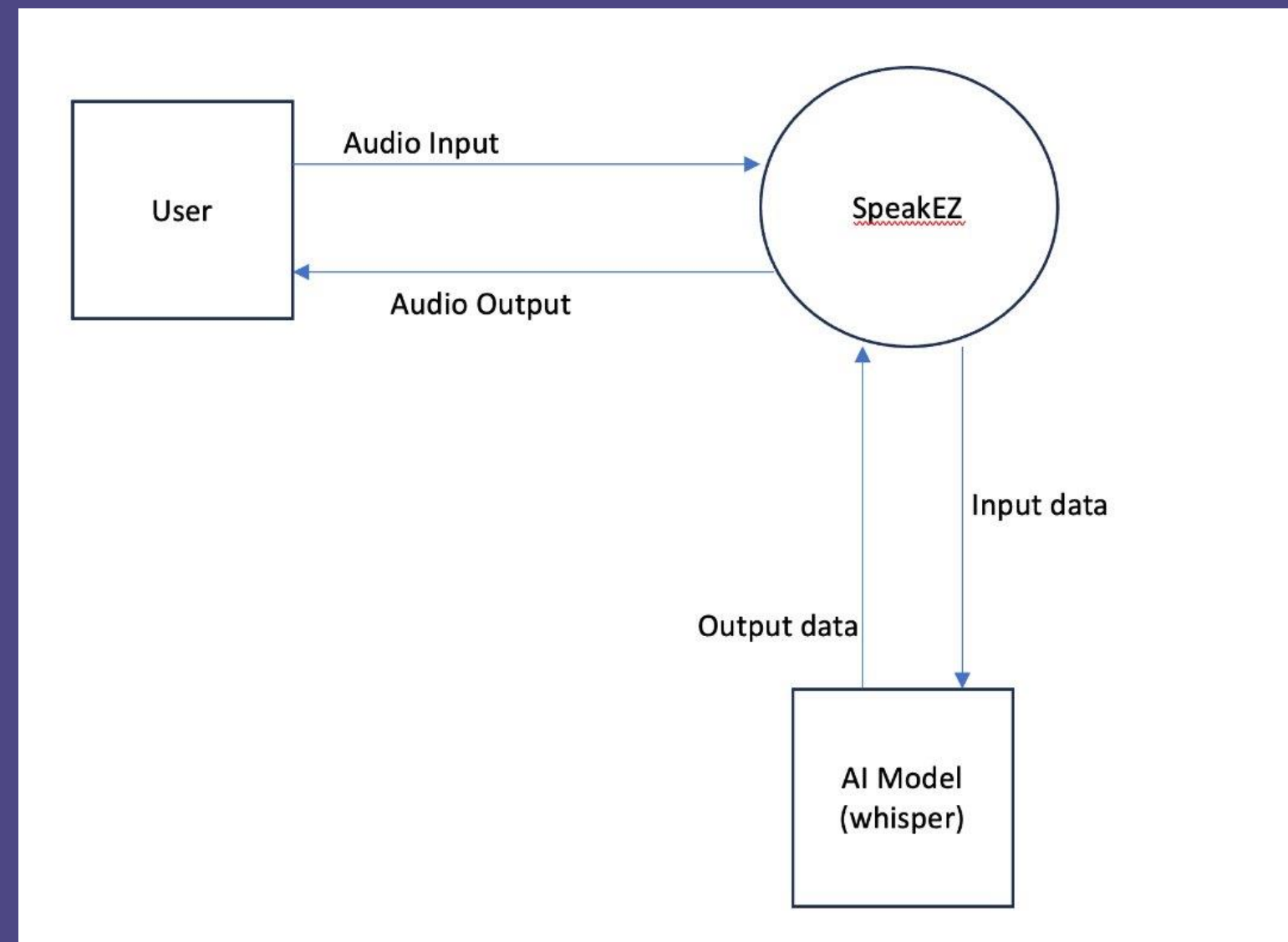
ALGORITHMS

Whisper is a transformer based on the encoder-decoder structure. It was trained on 680,000 hours of labelled speech data. The model can perform both speech recognition and speech translation.
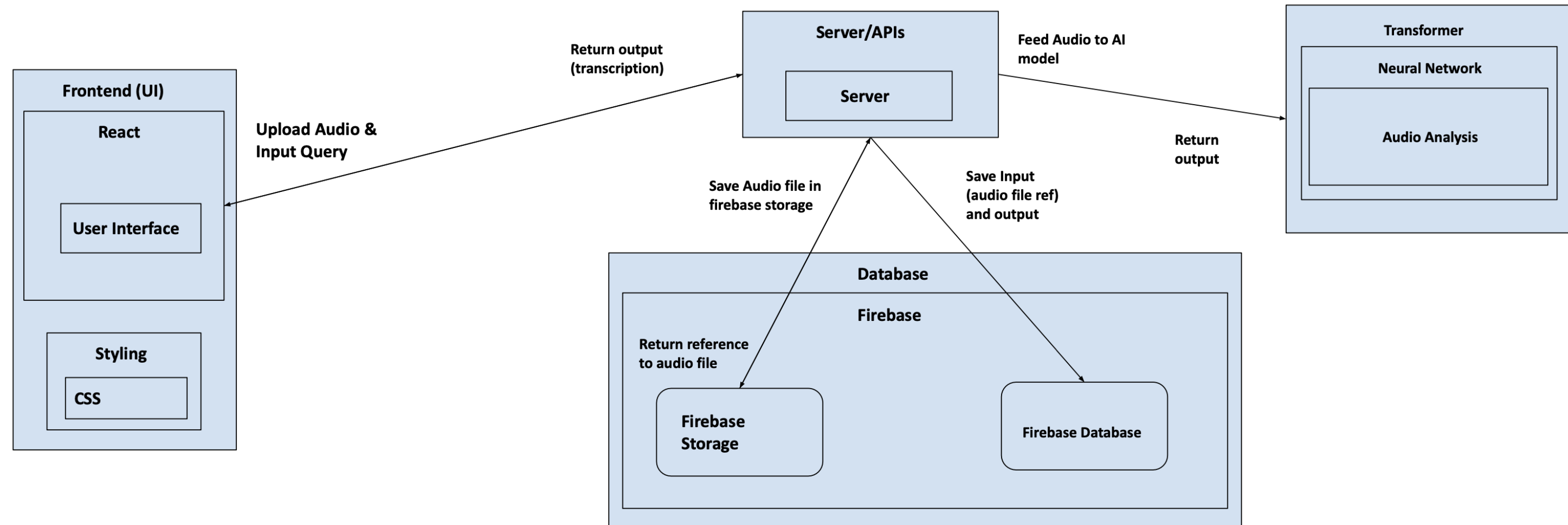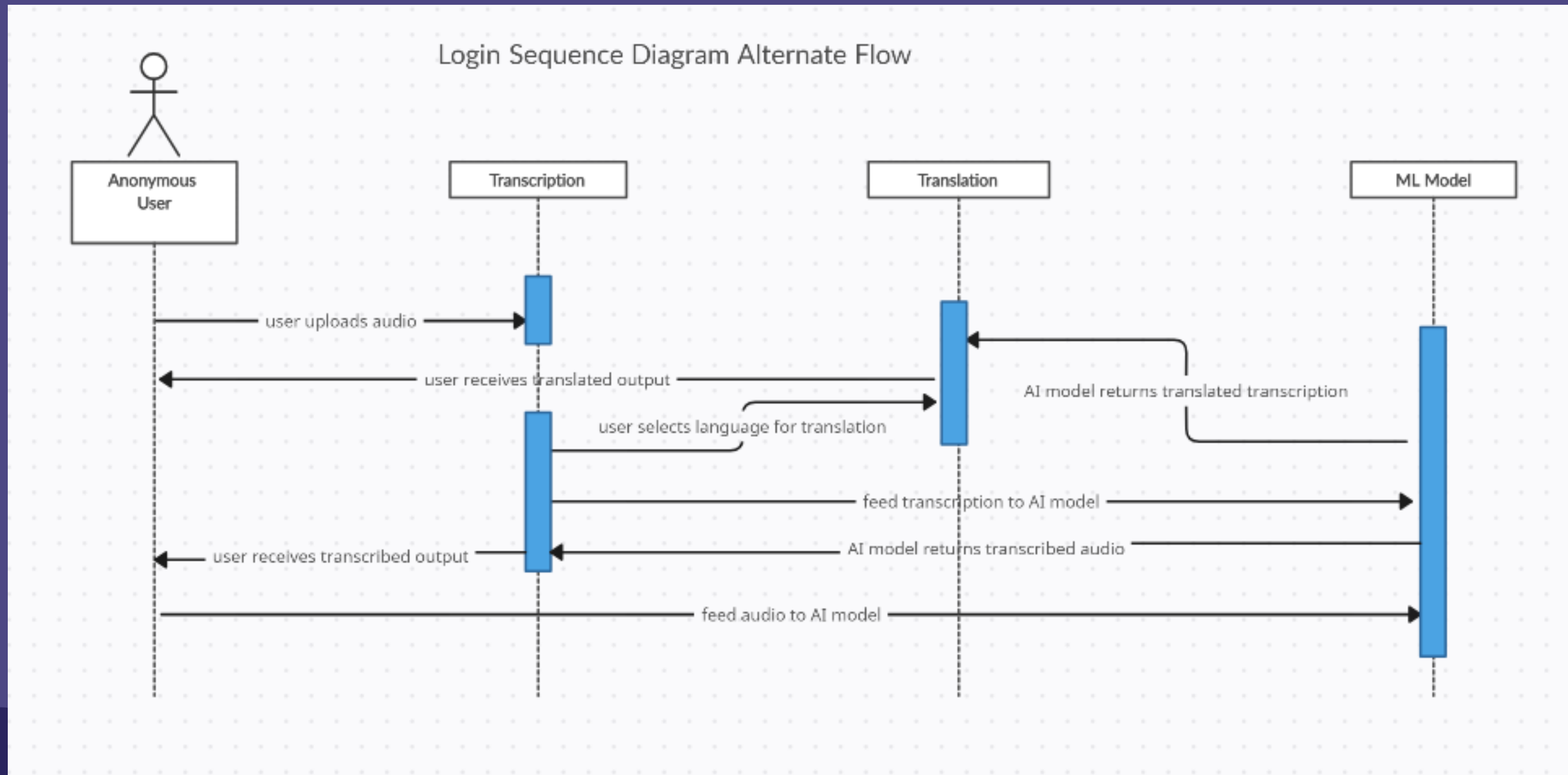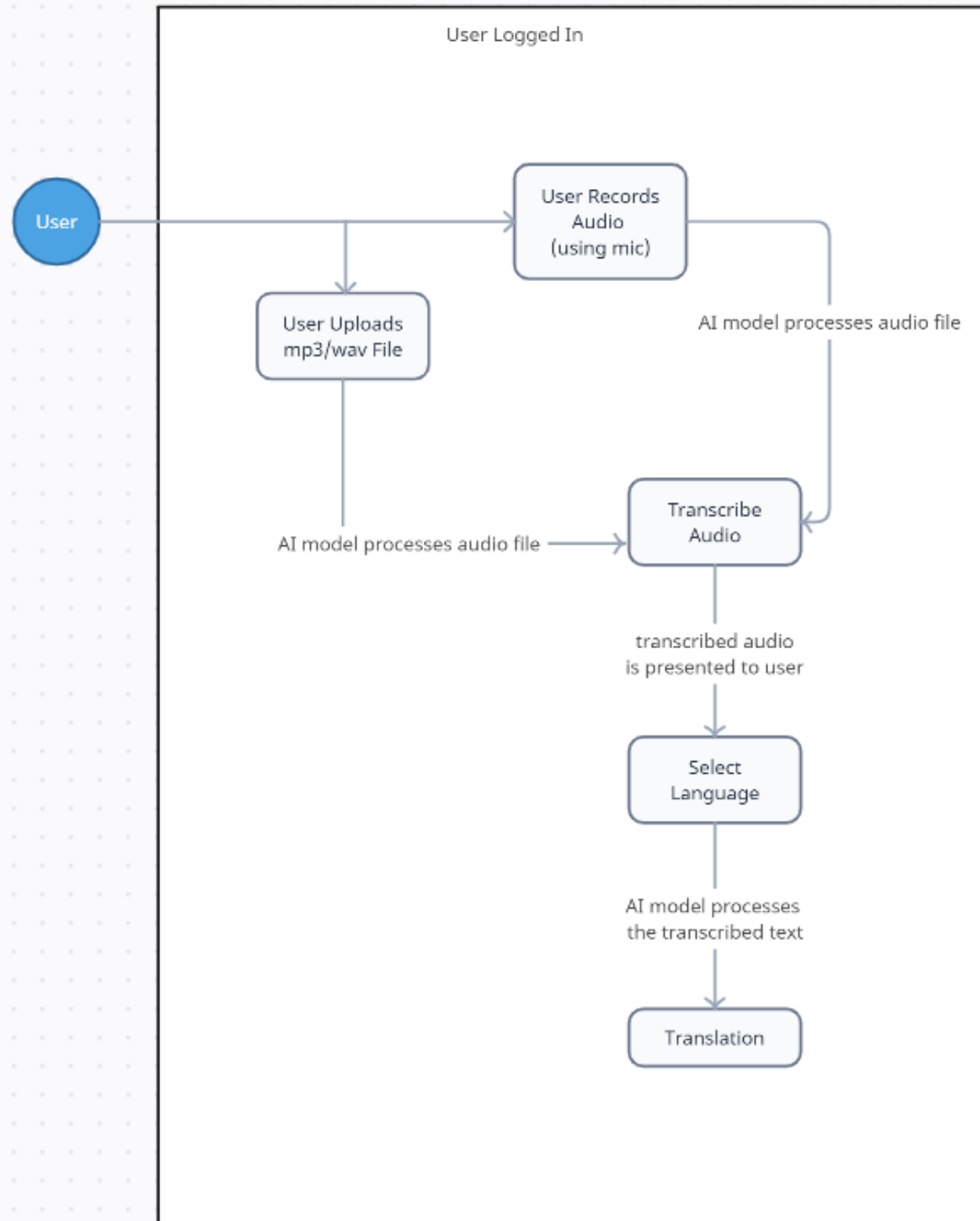

Automatic Speech Recognition

CONTEXT DIAGRAM

ARCHITECTURE DIAGRAM

Frontend (UI)

React

User Interface

Styling

CSS

Upload Audio & Input Query

Return output (transcription)

Server/APIs

Server

Feed Audio to AI model

Return output

Transformer

Neural Network

Audio Analysis

Save Audio file in firebase storage

Save Input (audio file ref) and output

Database

Firebase

Return reference to audio file

Firebase Storage

Firebase Database

SEQUENCE DIAGRAM

Login Sequence Diagram Alternate Flow

STATE DIAGRAM

User Logged In

User

User Records
Audio
(using mic)

User Uploads
mp3/wav File

AI model processes audio file

AI model processes audio file → Transcribe
Audio

transcribed audio
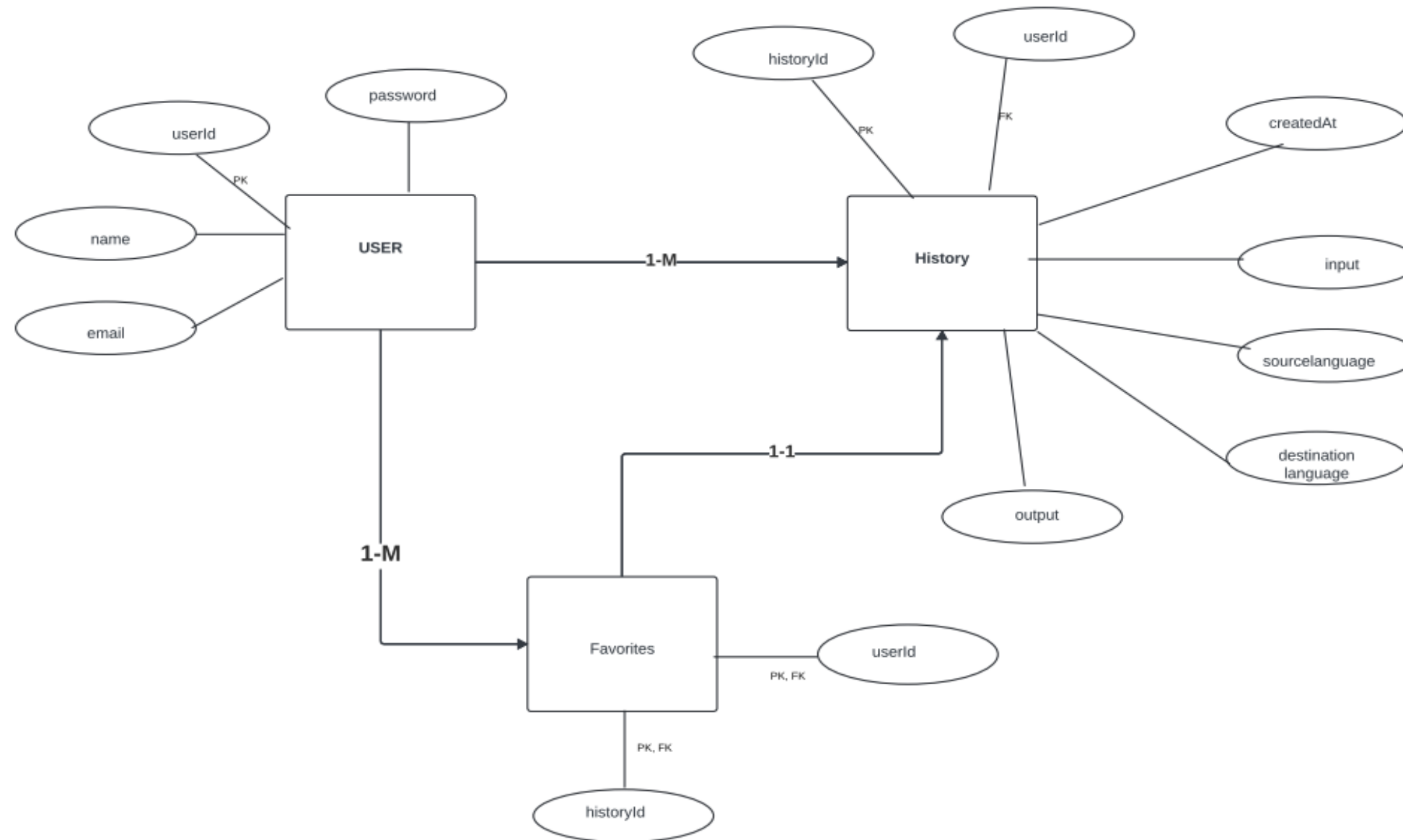is presented to user

Select
Language

AI model processes
the transcribed text

Translation

CLASS DIAGRAM

ER DIAGRAM

USER STORIES

🔖	LTA-118	File Size Limitation
🔖	LTA-116	Error Message for Unsupported File Types
🔖	LTA-119	Microphone Button for Voice Input
🔖	LTA-120	Accurate Translation of Audio Input
🔖	LTA-122	Audio Recording Feedback
🔖	LTA-124	Audio Input Time Limitation
🔖	LTA-126	Dropdown for Output Language Selection
🔖	LTA-127	Copy Translated Text
🔖	LTA-128	Download Button for Translated Text
🔖	LTA-129	Supported Formats for Downloaded File
🔖	LTA-130	Ability to Reuse Past Translations
🔖	LTA-131	Ability to Save All Past Translations
🔖	LTA-132	Toggle Button for Light/Dark Mode
🔖	LTA-133	Persisting Theme Preference Across Sessions

USER STORIES

🔖 LTA-126	Dropdown for Output Language Selection	TO DO ⌄
🔖 LTA-127	Copy Translated Text	TO DO ⌄
🔖 LTA-128	Download Button for Translated Text	TO DO ⌄
🔖 LTA-129	Supported Formats for Downloaded File	TO DO ⌄
🔖 LTA-130	Ability to Reuse Past Translations	TO DO ⌄
🔖 LTA-131	Ability to Save All Past Translations	TO DO ⌄
🔖 LTA-132	Toggle Button for Light/Dark Mode	TO DO ⌄
🔖 LTA-133	Persisting Theme Preference Across Sessions	TO DO ⌄
🔖 LTA-134	Marking and Unmarking a Translation as a Favourite	TO DO ⌄
🔖 LTA-135	Accessing Favourite Translations	TO DO ⌄
🔖 LTA-137	Dropdown Menus for Profile and Settings	TO DO ⌄
🔖 LTA-138	Configure Firebase Database Based on UML Diagram	TO DO ⌄

USER STORIES

LTA-115	Upload Button for File Selection
LTA-117	File Upload Progress
LTA-121	Option to Confirm or Cancel the Input
LTA-125	Dropdown for Input Language Selection
LTA-140	Implement Worker Functionality for Handling Incoming Messages
LTA-142	Send Audio Data for Transcription to Whisper Model
LTA-139	Integrate AI Model for Audio and File Input Translation
LTA-146	show the Progress of the audio file transcription
LTA-136	Header Display with Logo
LTA-141	Convert Audio File into Audio Data
LTA-145	Enable Copying Transcription from Text Box
LTA-123	Language Selection for Translation
LTA-144	Display Transcribed Text in a Text Box
LTA-143	Ensure Proper Handling and Order of Messages in Worker

TEST CASES

Feature to Test	Test Case	Test Steps	Test Data	Expected Result	Actual Result	Pass/Fail
"Upload" button	Ensure chosen file is uploaded	Click upload > file explorer opens > select audio file	.mp3 .wav	Selected audio file is presented on screen	Audio file was accepted	Pass
"Upload" button	Verify accepted audio type	Click upload > file explorer opens > searches through to ensure no other file types are shown to be accepted	.mp3 .wav .m4a .wma .flac	Only .mp3 and .wav files should be shown as accepted files in file explorer	The program allows for both .mp3 and .wav files to be selected	Pass
"Upload" button	User should be able to see the file name on screen once its selected	Click upload > file explorer opens > select audio file > file name shows on screen	.mp3 .wav	The audio file name is shown on screen after being selected	The file name was shown on screen after being selected	Pass
"x" Delete Button	Once user select the "x" icon, the file is deleted, and user is unable to see the name of file on screen	Click upload > file explorer opens > select audio file > file name shows on screen > click "x" icon next to file name > file deleted	.mp3 .wav	The file should be deleted and the file name should be removed from screen	User is able to click on the "x" icon, and the file was deleted. The file name was removed from display	Pass
Transcribe Button	Transcribe button should redirect page	Click "Transcribe" button > Page redirected to the transcription-translation page	.mp3 .wav	When a user click the "Transcribe" button, it redirect you to another page that shows the transcribed texts	User was redirected to the correct transcription-translation page	Pass
Transcribe Button	Transcribe button should initiate whisper model	Click upload > file explorer opens > select audio file > click "Transcribe" > inspect the page: console	.mp3 .wav	As the transcription process is occurring, and you're inspecting the page using the console, one should see the messages "DOWNLOADING", the AI model being called, and an object returned with the transcribed text.	The expected results occurred as expected, there were also some warnings shown.	Pass
Transcribe Button	Audio transcribed correctly	Click upload > file explorer opens > select audio file > click "Transcribe" > read transcribed text displayed	.mp3 .wav	What is said in audio should be shown as text in the text-box	The audio was transcribed correctly. Even when spoken in a foreign language such as Spanish, & German, it was transcribed correctly.	Pass
Page Functionality	Transcribe text presented in correct textbox	Click upload > file explorer opens > select audio file > click "Transcribe" > view where the transcribed text is displayed	.mp3 .wav	The transcribed text is shown in the text box specified to hold the transcribed text	The transcribed text was presented in the correct text-box	Pass
Page Functionality	Ability to edit transcribed text	Click upload > file explorer opens > select audio file > click "Transcribe" > add to the transcribed text & also delete some text from the text box	.mp3 .wav	Should be able to add or delete the transcribed text in the text box	User is able to add and remove from the transcribed text	Pass

SPRINT 1 STORIES

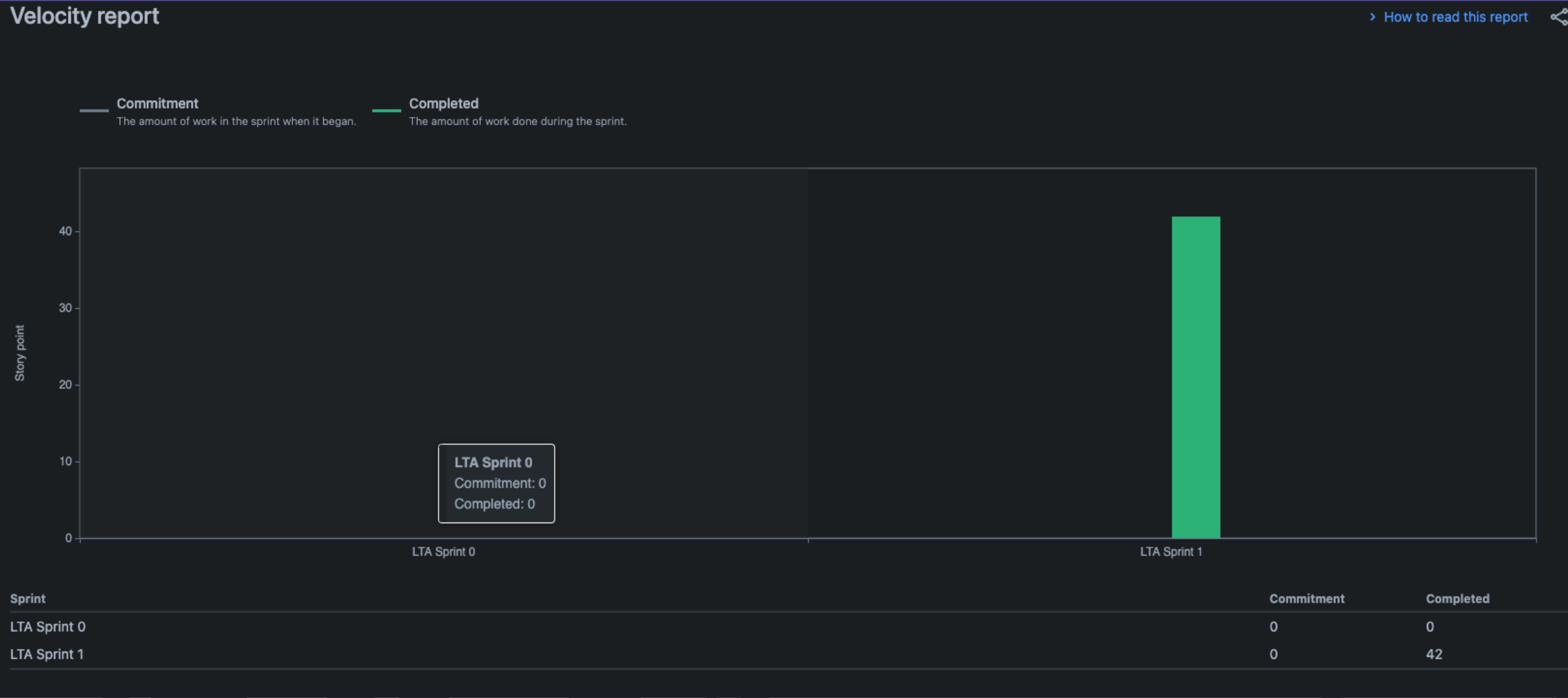|		
🔖	LTA-115	Upload Button for File Selection
🔖	LTA-117	File Upload Progress
🔖	LTA-121	Option to Confirm or Cancel the Input
🔖	LTA-125	Dropdown for Input Language Selection
🔖	LTA-140	Implement Worker Functionality for Handling Incoming Messages
🔖	LTA-142	Send Audio Data for Transcription to Whisper Model
🔖	LTA-139	Integrate AI Model for Audio and File Input Translation
🔖	LTA-146	show the Progress of the audio file transcription
🔖	LTA-136	Header Display with Logo
🔖	LTA-141	Convert Audio File into Audio Data
🔖	LTA-145	Enable Copying Transcription from Text Box
🔖	LTA-123	Language Selection for Translation
🔖	LTA-144	Display Transcribed Text in a Text Box
🔖	LTA-143	Ensure Proper Handling and Order of Messages in Worker

SPRINT 1 STORIES COMPLETED

Completed issues

Key	Summary	Issue type	Epic	Status	Assignee	Story points
LTA-115	Upload Button for File Selection	Story		DONE	VP	2
LTA-117	File Upload Progress	Story		DONE	SS	3
LTA-121	Option to Confirm or Cancel the Input	Story		DONE	VP	3
LTA-123	Language Selection for Translation	Story		DONE	RT	3
LTA-125	Dropdown for Input Language Selection	Story		DONE	PP	2
LTA-136	Header Display with Logo	Story		DONE	SS	1
LTA-139	Integrate AI Model for Audio and File Input Translation	Story		DONE	KL	5
LTA-140	Implement Worker Functionality for Handling Incoming Messages	Story		DONE	TG	5
LTA-143	Ensure Proper Handling and Order of Messages in Worker	Story		DONE	TG	3
LTA-141	Convert Audio File into Audio Data	Story		DONE	KL	3
LTA-142	Send Audio Data for Transcription to Whisper Model	Story		DONE	PP	3
LTA-144	Display Transcribed Text in a Text Box	Story		DONE	RT	5
LTA-145	Enable Copying Transcription from Text Box	Story		DONE	VP	2
LTA-146	show the Progress of the audio file transcription	Story		DONE	PP	2

SPRINT 1 STORIES NOT COMPLETED

Incomplete issues

Key ⇕	Summary ⇕	Issue type ⇕	Epic ⇕	Status ⇕	Assignee ⇕	Story points

All issues in sprint have been completed

TEAM VELOCITY- SPRINT 1

BURNDOWN CHART

Sprint burndown chart

> How to read this report

Sprint
LTA Sprint 1

Estimation field
Story points

Date - 26 September 2024 - 24 October 2024

Sprint goal - have an mvp

Remaining work
Number of story points left to complete this sprint

Guideline
Ideal burn rate
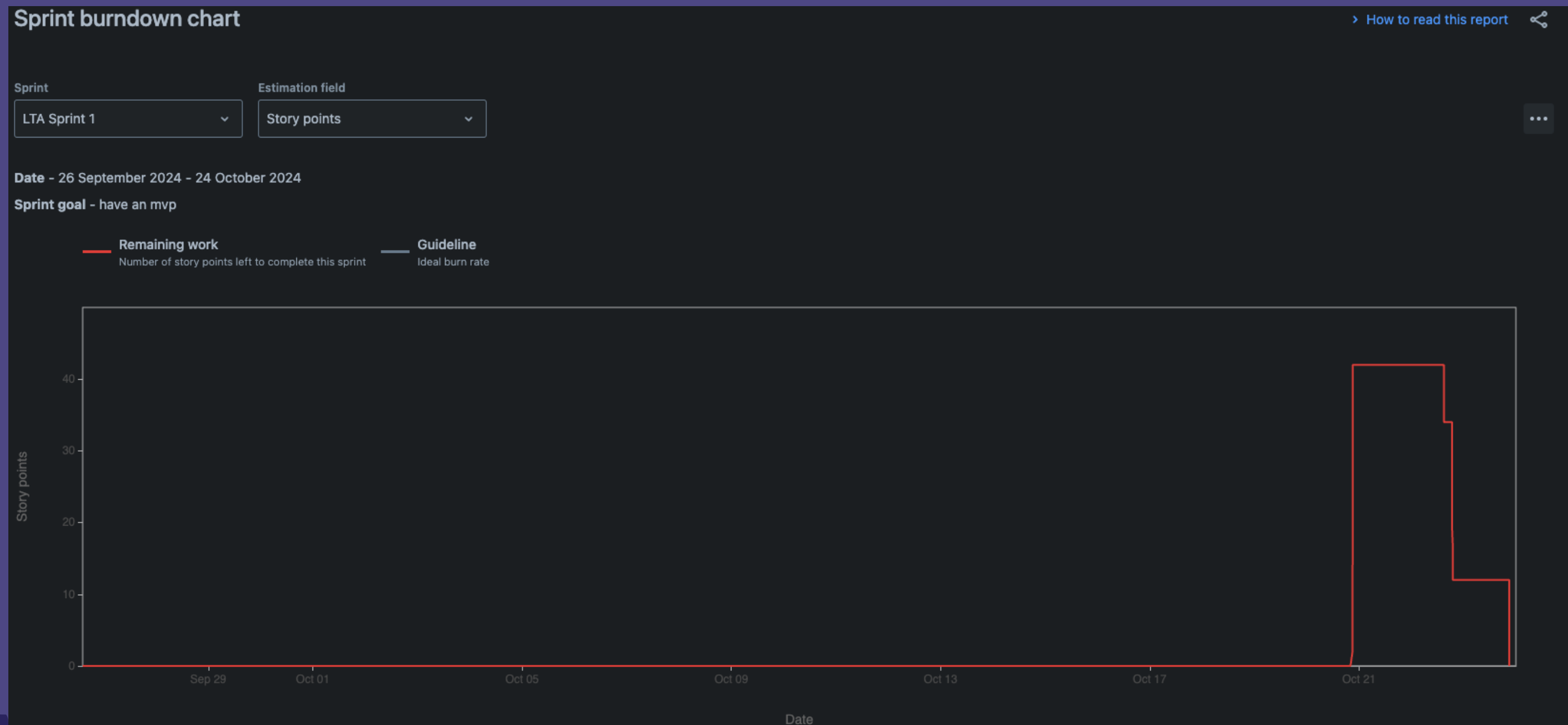
COMPLETED / COMMITTED RATIO

42 / 42

RETROSPECTIVE

What went well

What were you happy about this sprint

> Shout out to Kaali for continuing making agenda before every meeting and making polls for the meeting
>
> TANVI

> Shout out to Venkat and Steve for taking initiation to take tasks
>
> TANVI

What needs to be improved

What are some areas we could do better in

> I think the application needs a lot of improvement.
>
> ANONYMOUS

> improve communication
>
> ANONYMOUS

> There was a lot of stalling of time, like feedback wasn't given until next meeting, work wasn't shown until end, giving less time to review.
>
> TANVI

> Communication and collaboration. We can encourage teams members to regulary check in with each other about their tasks
>
> STEVE

> try not to miss too many meetings back to back. One should alert the team at least 30mins before so the team can reschedule the meeting so all can attend
>
> KAALI

> Demonstrate code as well instead of just output and also prefer explaining a bit of the code they worked on
>
> TANVI

> everyone should review all work that is posted and test out the code that has been pushed to the repo
>
> KAALI

> Code review session should be initiated
>
> PAVAN POLIREDDY

RETROSPECTIVE — ACTION ITEM

Application needs to be improved action item -
- after Friday presentation, we should do sprint planning on Sunday so that everyone will have their assigned tasks.

T

Improve Communication
-
- Text more if its about anything (issue, tasks, another session)
- Help team members (ask each other for help)

Solution -
- icebreaker
- screenshot in the group chat
- be more active in the group chat so that every team members knows about it instead of texting privately.

ATM

Work that has been done, review same day and give feedback same day

ATM

After all the agenda topics are completed, we can discuss the code that has been implemented since the last meeting.

R

https://youtu.be/RANnmD3Wrd8

TIMELINE

Design

Design prototype in Figma

Front-end

Basic front-end functionalites
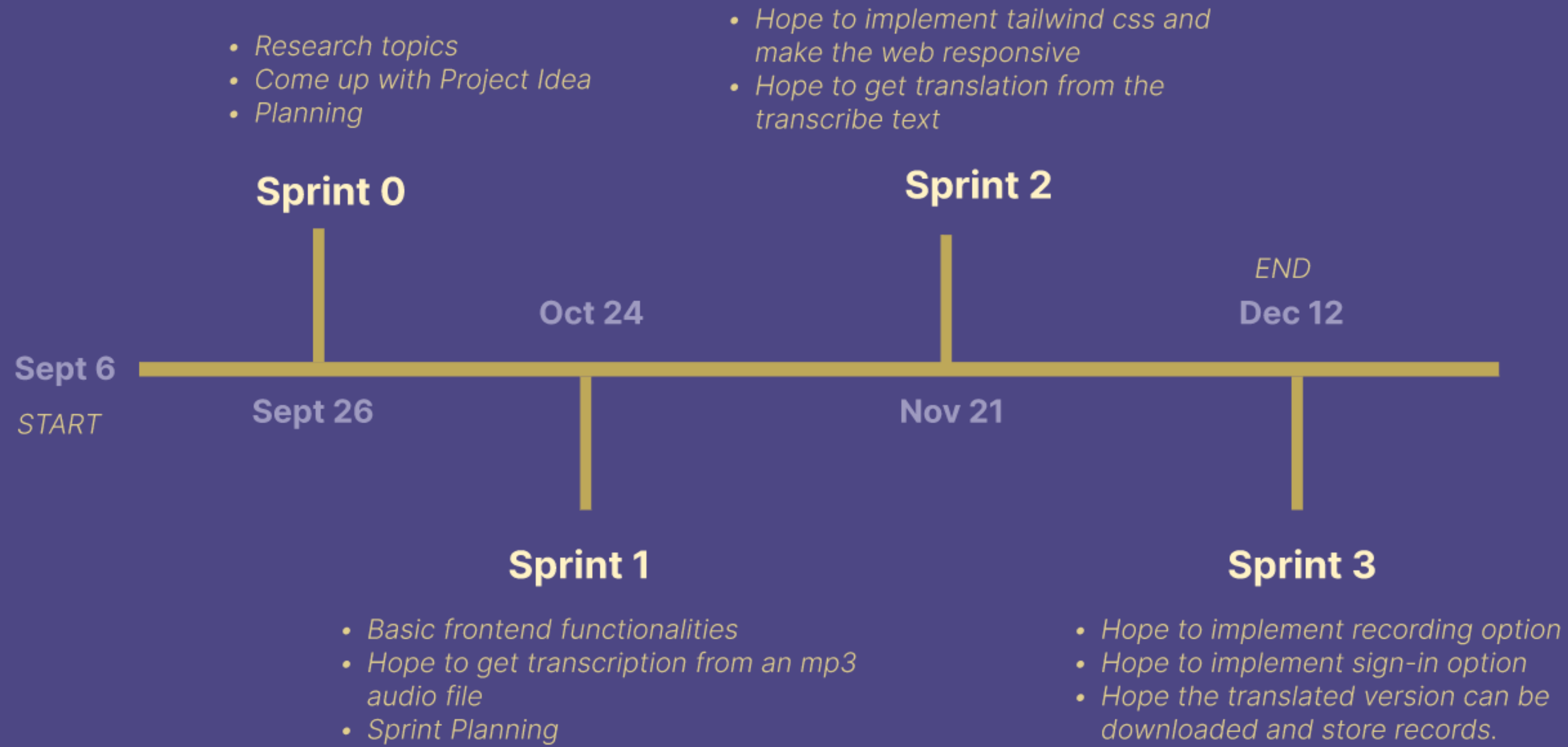
Back-end

Transcribe text, Test, Debug

Deliverable

MVP, Sprint Planning, Presentation, recording, wikipage, retrospective and p2p review

PROJECT SCHEDULE

- *Research topics*
- *Come up with Project Idea*
- *Planning*

- *Hope to implement tailwind css and make the web responsive*
- *Hope to get translation from the transcribe text*

Sprint 0

Sprint 2

END
Dec 12

Oct 24

Sept 6

Sept 26

Nov 21

START

Sprint 1

Sprint 3

- *Basic frontend functionalities*
- *Hope to get transcription from an mp3 audio file*
- *Sprint Planning*

- *Hope to implement recording option*
- *Hope to implement sign-in option*
- *Hope the translated version can be downloaded and store records.*

SPRINT 2 STORIES

🔖	LTA-118	File Size Limitation
🔖	LTA-116	Error Message for Unsupported File Types
🔖	LTA-119	Microphone Button for Voice Input
🔖	LTA-120	Accurate Translation of Audio Input
🔖	LTA-122	Audio Recording Feedback
🔖	LTA-124	Audio Input Time Limitation
🔖	LTA-126	Dropdown for Output Language Selection
🔖	LTA-127	Copy Translated Text
🔖	LTA-128	Download Button for Translated Text
🔖	LTA-129	Supported Formats for Downloaded File

SPEAKEZ SCREENSHOT

Sign In

Drag and drop
or **upload**
a mp3 or wave file
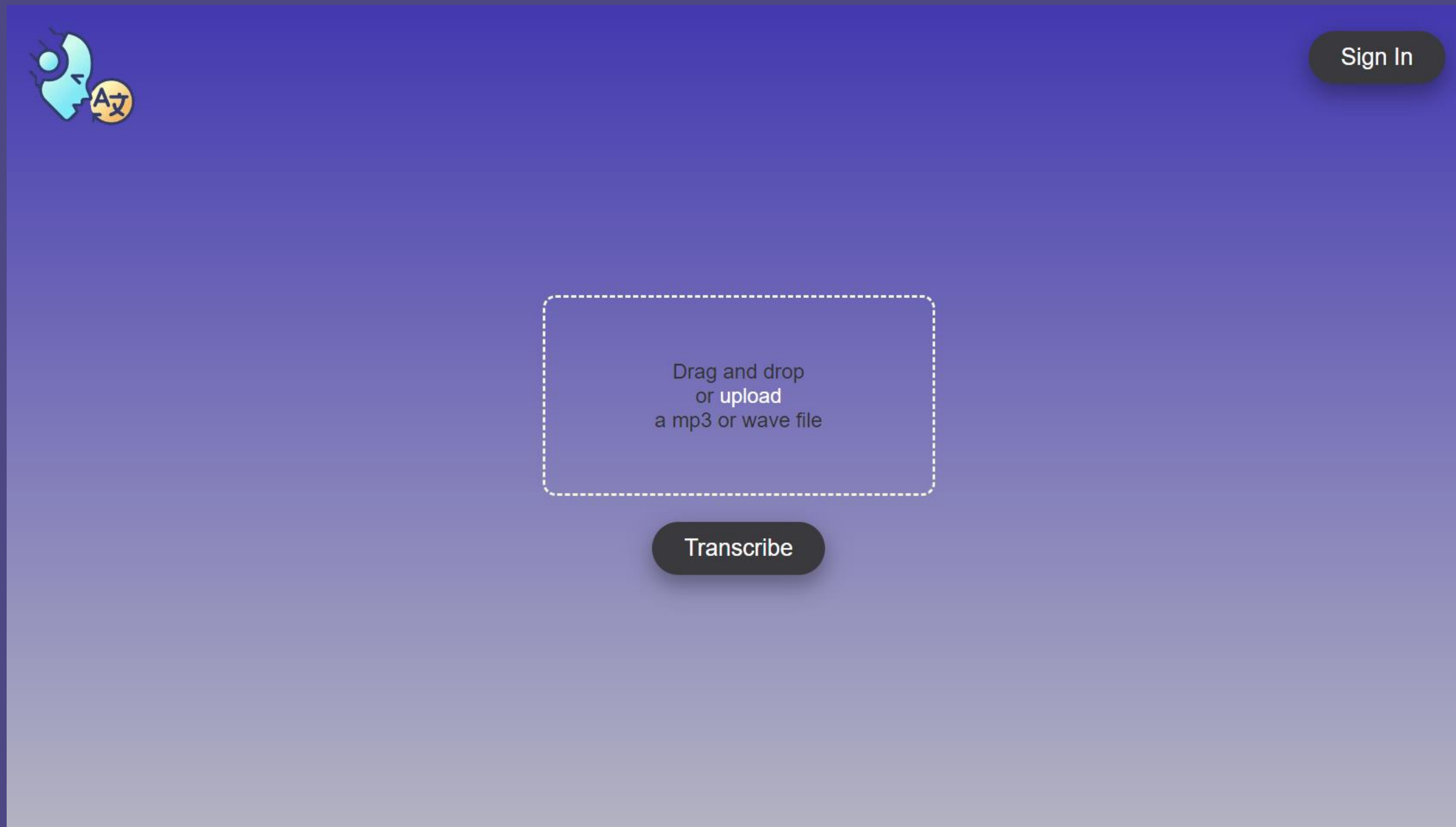
Transcribe

SPEAKEZ SCREENSHOT

Sign In

Translate

Select langu ⌄

Testing testing one, two, three.

Select Language

Copy transcribed text: 🗐   Copy translated text: 🗐

SPEAKEZ PIPELINE
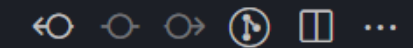
```js
import { env, pipeline } from "@xenova/transformers";
import { MessageTypes } from "./reference";
env.allowLocalModels = false;

class MyTranscriptionPipeline {
  static task = "automatic-speech-recognition";
  static model = "openai/whisper-tiny.en";
  static instance = null; // null to hold the instance of the pipeline

  // commnicate all the status with our main function // returns an instance of the pipeline
  static async getInstance(progress_callback = null) {
    if (this.instance === null) {
      // creates an instance by calling the pipeline function passing the task, model and optional callback to track loadin
      this.instance = await pipeline(this.task, null, { progress_callback });
    }

    return this.instance;
  }
}

// we add the eventlistener to listen for messages that we send from the main file and processes the audio data by calling
self.addEventListener("message", async (event) => {
  const { type, audio } = event.data; //takes the type and audio from the incoming message data
  if (type === MessageTypes.INFERENCE_REQUEST) {
    await transcribe(audio);
  }
});
```

SPEAKEZ DEMO

WIKIPAGE LINK

https://github.com/htmw/2024F-Bytes/wiki