

VocaFree

An Luu (301574874)

Parsa Ghaderi (301623337)

Jake Choi (301552446)

Caiden Merklin (301608302)

<https://github.com/CMPT-276-FALL-2024/project-07-hills>

Overview

Karaoke can be a fantastic way for friends and family to bond and have fun together, but it often comes with significant limitations that make it less accessible for many. Several barriers hinder participation, including busy schedules that leave little time for leisure activities, financial difficulties, as the costs associated with going to karaoke venues or purchasing home karaoke systems can be prohibitive, and the needs of individuals with physical or mental conditions, who may find it challenging to travel to venues. Additionally, teams in work settings may struggle to engage in team-bonding activities like karaoke due to tight schedules. Singing enthusiasts often feel limited by the narrow selection of songs available in traditional karaoke catalogs, while language learners also face challenges, as understanding and pronouncing foreign lyrics can be difficult without proper support.

Hence, our group came up with the idea of VocaFree, a free-to-use web application that uses YouTube song video links to create an authentic karaoke experience by separating vocals from the music of any song, and playing just the instrumental track, allowing users to sing any songs of their choice without the original song's vocals. VocaFree addresses challenges by offering a free, convenient, and accessible karaoke solution that can be enjoyed from home, regardless of time constraints, financial constraints, or travel limitations. Moreover, for teams and workplaces, VocaFree provides an engaging way to boost team spirit without taking up too much time or requiring employees to leave the office. For singing enthusiasts, the app opens up a wider range of song choices by allowing users to convert any YouTube song videos into karaoke tracks, overcoming the limitations of karaoke song catalogs that only include well-known songs. Language learners and cultural enthusiasts also benefit, as VocaFree offers translation and romanization/phonetic assistance for foreign language songs, helping users practice pronunciation of those languages while enjoying a fun activity.

VocaFree is versatile and accessible, making it an ideal solution for various potential users, including:

- **Individuals with Busy Schedules:** Those who may not have time to go out can easily enjoy karaoke at the convenience of their home.
- **Individuals Facing Financial Difficulties:** Karaoke can often be an expensive activity, with both the cost of going out to karaoke venues and the high price of home karaoke systems being barriers. VocaFree offers a cost-free solution, allowing users to enjoy a full karaoke experience from the comfort of their home without needing to invest in expensive equipment.
- **Individuals with Mental or Physical Conditions:** For those who struggle to leave their homes, VocaFree provides an inclusive activity that can be enjoyed in a safe environment.
- **Teams and Work Settings:** With tight schedules and heavy workloads, employees often struggle to find time for team-building activities. VocaFree allows teams to engage in

karaoke sessions from their offices or homes, enhancing team spirit without requiring significant time away from their responsibilities.

- **Singing Enthusiasts Who Want Access to a Wider Range of Songs:** For those who love singing and want to challenge themselves, VocaFree provides a solution when the songs they wish to sing aren't available in the catalogs of traditional karaoke places or apps. This allows users to sing along to any song by extracting the instrumental only from the music of YouTube song video links, making it ideal for both casual singers and those serious about honing their vocal skills.
- **Language Learners:** For songs in foreign languages, the app provides English pronunciations (romanization and phonetics) of the lyrics, enabling users to sing along even if they don't understand the original language, making it easier for language learners to practice speaking in that language.

Personas

Persona 1: Individuals with Busy Schedules

Name: Sarah Thompson

Age: 35

Occupation: Marketing Manager

Family: Single mother with 2 kids (aged 7 and 9)

Background: Sarah works long hours in a fast-paced job and has two children to take care of, and often struggles to find time for leisure activities.

Needs: Quick and easy access to entertainment that can fit into her busy schedule.

Goals: Enjoy activities at home without needing to travel, allowing her to unwind after work, while being able to look after her children.

How VocaFree Helps: She can easily use VocaFree for karaoke sessions during her downtime, providing relaxation without extra time commitment.

Persona 2: People Facing Financial Difficulties

Name: Stan Marshall

Age: 22

Occupation: College Student

Family: Late parents, 3 younger siblings

Background: Stan is a full-time student who is mindful of his expenses and often looks for free activities. He is on a very limited budget due to rent and needs to provide for his siblings, and is currently relying on his full-ride scholarship to subsidize his schooling.

Needs: Affordable entertainment options that don't strain his (extremely) limited budget.

Goals: To have fun with friends without spending money on outings.

How VocaFree Helps: VocaFree allows Stan to enjoy karaoke with friends for free, creating memorable experiences without financial stress.

Persona 3: Individuals with Mental or Physical Conditions

Name: Linda Garcia

Age: 67

Occupation: Retired

Family: None

Background: Linda has physical mobility challenges that limit her ability to travel. She enjoys singing but finds it difficult to participate in social activities outside her home, especially since there is no one to support or take care of her.

Needs: Accessible entertainment options that can be enjoyed without needing to leave the house, as well as opportunities for social interaction with friends.

Goals: To stay socially connected with her loved ones and enjoy shared experiences, while also having fun and expressing herself through music.

How VocaFree Helps: VocaFree allows Linda to invite friends to join her at her home for karaoke sessions, creating joyful memories and fostering connections from the comfort of her home.

Persona 4: Team Leaders Striving to Increase Team Spirit in Work Settings:

Name: Mike Patel

Age: 42

Occupation: Project manager at a Tech Company

Family: None, he thinks of his co-workers like family instead

Background: Mike manages a remote team and is looking for ways to strengthen team bonds.

Needs: Quick team bonding activities that don't require extensive travel/time commitments.

Goals: To enhance team spirit without disrupting work schedules.

How VocaFree Helps: Mike can organize fun karaoke sessions virtually or in the office, allowing his team to engage while saving time.

Persona 5: Singing Enthusiasts Who Want Access to a Wider Range of Songs

Name: Emma Rodriguez

Age: 24

Occupation: Music Teacher, Singer

Family: Parents who were both famous singers

Background: Emma loves singing and is always looking to improve her vocal skills. Some of the songs she wishes to sing do not have a karaoke version at the karaoke places she goes to.

Needs: A platform to practice singing, as well as access to karaoke versions of songs that are not available at most karaoke venues.

Goals: To challenge herself, hone her craft, and practice singing various songs.

How VocaFree Helps: VocaFree serves as a tool for Emma to practice her singing, while providing her with a vast library of songs, including those not typically found at karaoke venues.

The app allows her to create custom karaoke experiences and refine her vocal technique.

Persona 6: Language Learners

Name: Alex Smith

Age: 26

Occupation: Graduate Student in Linguistics

Family: Parents who both like to travel to different countries

Background: Alex is passionate about learning new languages. He is currently learning Korean and wants to speak it without having an accent.

Needs: Resources that are engaging and facilitate Korean speaking practice.

Goals: To improve Korean speaking skills while enjoying music and culture.

How VocaFree Helps: VocaFree's features like lyrics translation and romanization enable Alex to practice pronunciation and understand songs in different languages, enhancing his practice experience.

APIs/Technologies

The Central feature of VocaFree is the Ultimate Voice Remover (UVR), an open-source tool designed for separating vocals from instrumentals in audio tracks. We would be using this to separate vocals from instrumentals in songs, thus allowing for the karaoke features. The following are the APIs to be used and their features:

Genius API: The Genius API allows access to a database of song lyrics, artist information, and song metadata from the Genius platform. It allows you to search songs by title or lyrics, retrieve artist profiles, and get links to Genius pages where full lyrics can be found and web scraped. This API is popular for building music-related applications, helping users discover songs and learn about artists.

- **Get Lyrics:** Retrieve the lyrics for a specific song, allowing users to sing along accurately.
- **Get Song Based on Lyrics:** Search for songs by providing a snippet of lyrics, helping users find the correct song they want to sing.
- **Get Artist Information and metadata:** Fetch detailed information about artists and the song, including their biography, song metadata.

User Stories for each corresponding feature:

1. As **Stan**, I want to see the lyrics for each artist. Due to my limited budget, I cannot afford to subscribe to music streaming services that offer lyrics support. It would be great for me to see the lyrics for free while having fun at the same time.

2. As **Sarah**, I want to make the best out of my limited time. The ability to look up songs by just lyrics, would be greatly helpful for me. As I only ever get to hear bits of songs here and there, and do not get the chance to listen to music on my own. Hence I do not know many song names, only pieces of lyrics.
3. As **Linda**, I want to get more info on artists. I am greatly interested in the music world and artists, but rarely get to go to events to discover new and upcoming artists due to my disability. Having the ability to understand more about artists and songs will be greatly valuable for me.

OpenAI API: This is the most commonly used LLM (large language model) api, developed by OpenAI and used for ChatGPT. It can be used to generate text from a prompt, and can return almost any kind of response.

Features:

1. **Translate Lyrics:** Use the API to translate song lyrics into different languages, allowing users to understand the meaning of non-native songs in real time, allowing them to improve on their foreign language skills.
2. **Romanization/Phonetics:** Use an additional prompt to convert the characters into their phonetic (roman) equivalents so it can be easily read by someone used to Roman script. This allows someone the opportunity to sing along in a new language even if they can't read the characters.
3. **Basic Lyric Analysis:** Analyze song lyrics to provide insights or themes, enhancing users' understanding and engagement with the music.

User Stories for each corresponding feature:

1. As **Alex**, a graduate student in linguistics, I want to improve my **Korean language skills** in a fun and engaging way. The ability to **translate song lyrics into English in real time** would be incredibly helpful. This would allow me to **understand the meaning of the songs I'm singing** without relying solely on textbooks or language exchange partners.
2. As **Mike**, I want to foster stronger bonds within my diverse team. I believe that adding **romanized** versions of non-English song lyrics will provide a fun and inclusive way for everyone to participate in karaoke sessions and have every team member sing the songs from each of their cultural backgrounds, even if everyone else doesn't speak the original language. This will help build a more cohesive team and create a positive work environment.
3. As **Emma**, I want to deepen my understanding of the songs I love. Analyzing song lyrics would help me **uncover hidden meanings and appreciate the artistry behind the music**. This feature would also be a valuable tool for **teaching my students about song structure, storytelling, and emotional expression**.

Spotify API [Backup]: Spotify's own web API for developers. It gives access to data from the user's activity. It can be used to search for content and also control the playback.

Features:

1. Get access to the **user's playlist**: This feature will be used so users can pick songs from their playlist to sing along to.
2. **Audio preview** for songs (replacement for getting lyrics): This is so that users can get a feel of how a song sounds before choosing it.
3. Get **Song recommendations** based on current songs: To personalize the user's experience. After singing along to a few songs, similar songs will be recommended to the user.

User Stories for each corresponding feature:

4. As **Emma**, I want to use songs from my own playlist. I have a diverse set of songs that I enjoy, and I do not remember every single song's name. So the ability to sync my playlist and choose my favourite songs will be valuable to me.
5. As **Alex**, I want to preview a song before choosing it. I am always learning new languages, and I find singing along to songs to be helpful in my learning journey. However, some songs are challenging for me as they are either too fast or advanced. Therefore, being able to preview songs before choosing would be helpful. I also want to confirm if it is the correct song I want to sing, so that I don't accidentally choose a song that I don't know.
6. As **Sarah**, I want to be recommended songs. I do not have time to research and find songs that I enjoy. Therefore, being able to choose similar songs to the ones that I do enjoy, would be useful.

Web Audio API [Backup]: This is an alternative if we need additional features not mentioned in the previous API's. This API is commonly used for the handling of Audio inside of JavaScript and is provided by Mozilla.

Features:

1. **Autotune**: This feature can be used to correct the pitch of the song in real time. This can be used to enhance the user's singing or as a feature just for fun.
2. **Play songs on a webpage**: Connect the sound to the webpage with a nice UI including a play/pause button.
3. **Visualize the music playing**: Display a progress bar that adjusts based on the audio that is playing.

User Stories for each corresponding feature:

1. As **Mike**, features that are able to bring my team opportunities to interact without feeling uncomfortable such as voice changers, would be extremely beneficial to the experience. This will drive more engagement to our team.
2. As **Linda**, considering my visual disability that prohibits me to see clearly, ease of use of the UI is essential. I often struggle with new technology as well so ease of use and navigation controls are imperative to my experience.
3. As **Sarah**, the ability to see how much longer is left in each song is essential as I am often running on a super tight schedule.

Storyboard of the Application Interface/Features

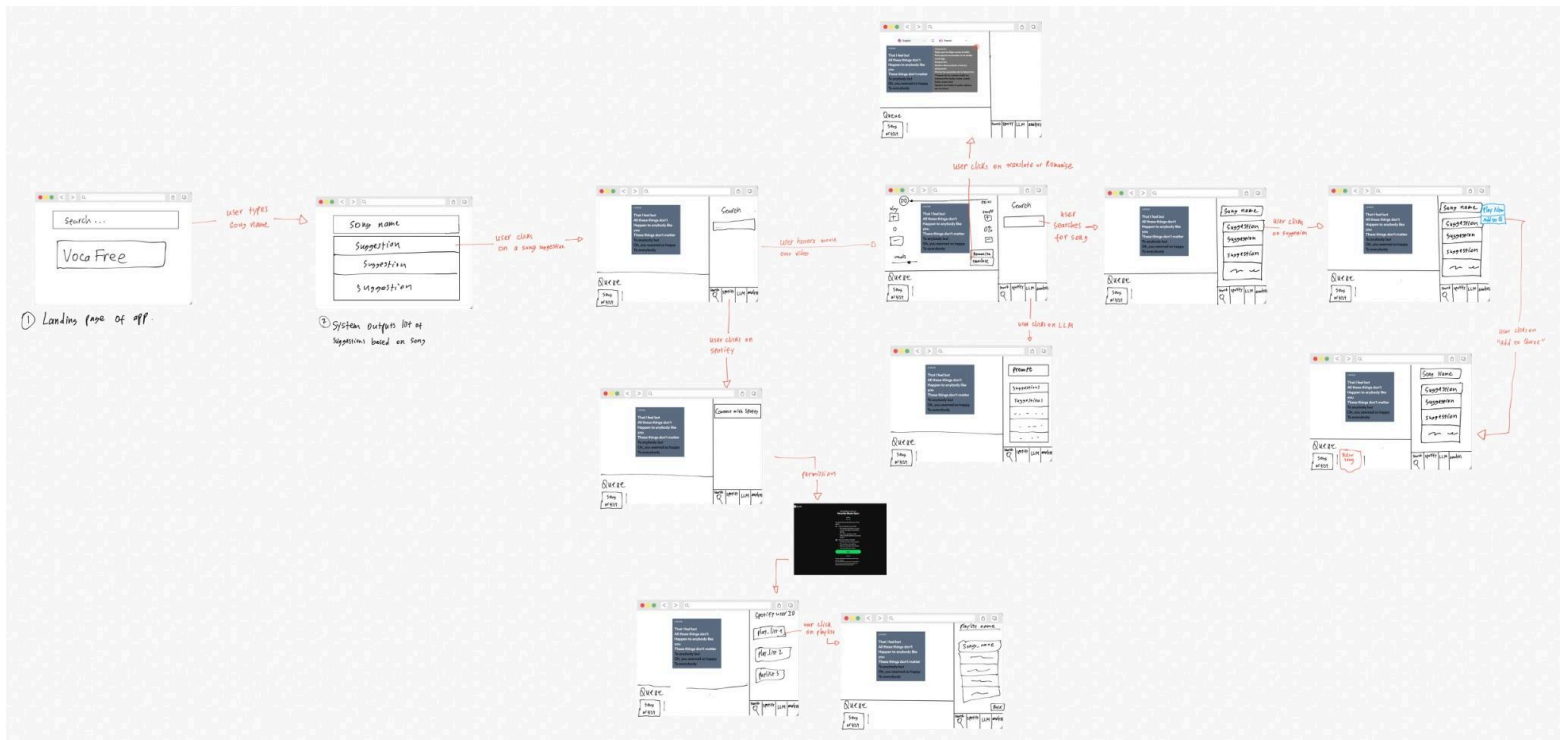


Image 1. Low-fidelity of the application interface (more details can be seen by opening this image in our repository).

Front-End Technology Stack

Javascript: React.js

HTML/CSS: Tailwind CSS (or Material-UI)

Testing: Jest, React Testing Library

CI/CD Pipelines: GitHub Actions

Deployment: Vercel

- **Ease of Use:** React's component-based architecture and declarative syntax make it relatively easy to learn and use, even for those with limited frontend development experience.
- **Abundant Resources:** There's a vast ecosystem of tutorials, documentation, and community support available for React, making it easy to find answers and solutions to common challenges.
- **Team Experience:** Some members of our team have prior experience with React, which can accelerate development and knowledge sharing, especially since we have a significant amount to accomplish. Using a framework that we are already familiar with will allow us to work more efficiently and focus on building key features right away.
- **Resume Building:** React is a widely used and respected framework in the industry. Learning and using React can enhance your skills and make your resume more competitive.

By choosing React, we're investing in a technology that offers a strong foundation for building scalable and maintainable web applications.

****Extra features not expected to be marked, but for the TAs/Prof/project markers to know****

On the backend, we'll leverage **Flask**, a lightweight Python web framework, to handle server-side logic and potential web scraping. Flask's flexibility and simplicity make it well-suited for tasks like **audio separation** and other Python-based operations. This will be used in particular with the **vocal extractor in the audio separator**. We will test with **pytest** and deploy with **Heroku**. However, we may just containerize the front and back end with docker instead for ease of deployment.