

JustPerform

"A better karaoke experience"

Team D1

Kiera Cawley, Aleks Watkins, Hugo Martinez



Use Case

The Problem:

Karaoke at-home experience is fragmented. Users must secure equipment, search different platforms to look for songs that may have suboptimal audio, and do not receive feedback.

Our Solution:

JustPerform creates one seamless karaoke experience that removes audio from users' personal music library and analyzes audio and motion to provide feedback for improvement.

Use Case Requirements

Objectives:

Microphone	10ms latency
Speaker	10ms latency
Song accuracy	80% accurate vocal removal 98% accurate lyrics
User feedback system	90% user agreement in testing
Program latency	<10s latency to start system <10s latency for feedback

Design Requirements

Speaker:

- Less than 15 lbs
- Capacity for ≥ 80 dB
- Filter latency < 10 ms

Microphone:

- Bluetooth functionality for mobility
- Connection latency < 10 ms

UI:

- Lyric sourcing < 3 s
- Lyric timing < 200 ms of delay
- Less than 200 ms of delay on audio scoring
- Less than 200 ms of delay on motion scoring

Solution approach

Feature	Motivation	Considerations
Connect to user's personal music library Auto populate and display lyrics Remove vocal track in real time	Unite the tedious parts of karaoke experience into one location	Appropriate sizing for those with visual disabilities Non-resource intensive solutions
Process audio for singing timing and volume	Focus on typical karaoke experience and gamification aspect; pivoted from improvement feedback	Accessible to those with different accents
Process motion for activity levels and beat matching	Promote performance and gamification experience beyond just vocals	Accessible to those with mobility issues, as long as they can hold the microphone

Audio Processing Plan

Vocal Removal

- By subtracting left and right channels, we can remove vocals without affecting the backing track
- We will then filter additionally to help remove any remaining artifacts with a simple bandstop filter.

Final Output Generation

- We can sum the microphone signal to the track with vocals removed.
- This will then be outputted to our speaker.

Performance Scoring (Audio)

- By subtracting the final generated signal (user vocals with original vocal-removed track) our resulting signal carries information about how accurate the performance was.
- This will be performed via a sliding window, that can look at sections of audio and on our software side.

Tradeoffs

Design Choice	Pros	Cons
Left and Right Channel Subtraction	Allows for real time processing, effective, simple	Versatility reduced to only songs in stereo
Direct original / final signal comparison	Provides a metric that can be analyzed in real time	Complex (buffering input signal and processing remnants), simplified representation
Accelerometer System	Lower latency Acceleration (x, y, z) and angular velocity (x, y, z) values given to you directly	Adds more hardware Potentially complex wiring

Implementation Plan

We're Buying

- Speaker

We're Downloading

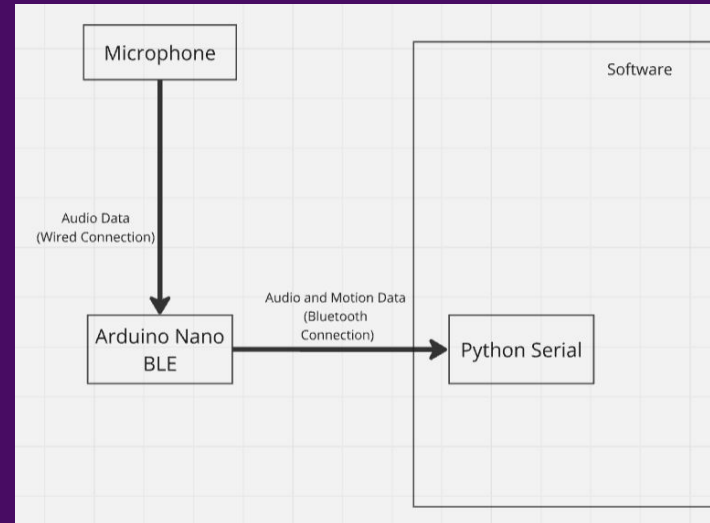
- Song lyrics (scraping lyricsify.com)

We're Designing + Developing

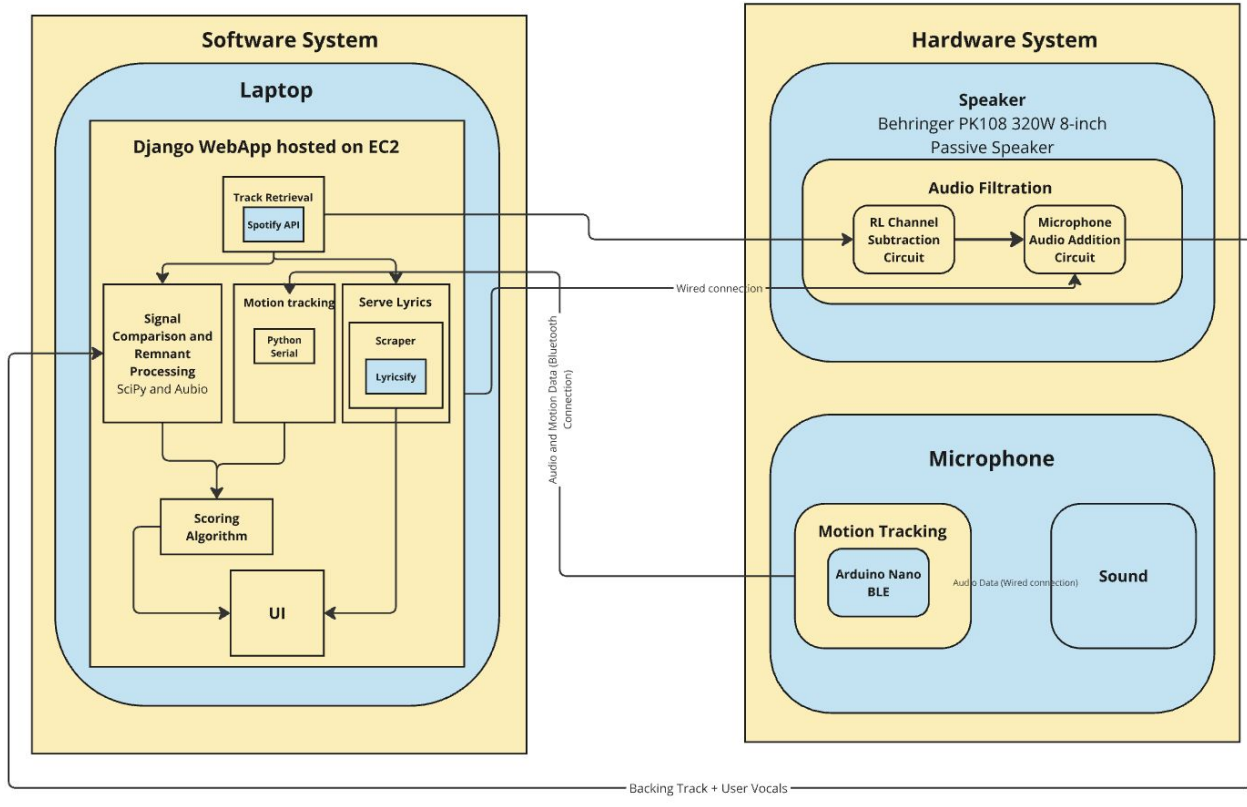
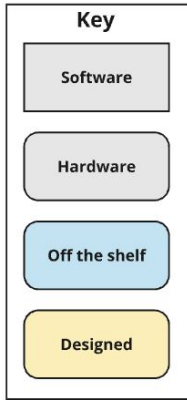
- Filter to separate lyrics from vocals
- Algorithm to score users based on vocals and movements
- Web application + UI

We're Assembling

- Enhanced microphone (captures audio and motion data)



Solution Approach



Testing Verification and Metrics

Singer vocals

90% Accurate

We should be able to directly compare the user's vocal signal + the backing track to the original audio and give them a score with reasonable accuracy.

Contingencies: perform more simple comparison (ensuring the user is singing at the right time or volume)

Singer movement

90% Accurate

The microphone will be moved according to a series of simple movements (direction and velocity changes) to check accurate decoding of sensor readings.

Contingencies: change accelerometer type, multiple accelerometers, basic CV

Music Library

80% Accurate

We should be able to source the track, lyrics, and remove audio for 80% of English speaking songs on Spotify with more than one primary singer with minimum 3 million listens.

Contingencies: pivot lyric source, manual upload of track, manual upload of timed lyrics for proof of concept

JustPerform

WebApp

- Set up application
- Generate basic HTML pages
- Connect with Spotify API and music l...
- Playable music with Spotify API
- Scraper for lyricsify.com
- Connect scraper with Spotify output
- Lyric UI Build
- Sync Kiera's mechanism + audio
- Finetune general UI (Figma)
- Deployment
- Slack

Motion Tracking

- Set up accelerometer within microph...
- Connect accelerometer to computer
- Noise filtering
- Decode basic movements
- Accuracy detection for specific mov...
- Sync prompt for user with music
- Finetuning and testing
- Slack

Audio Processing

- Isolate singer signal from track
- Stereo vocal attenuation finetuning
- Hardware filtering construction 1
- Boosting backing/cleaning artifacts
- Combining audio outputs
- Software side audio comparison
- Testing and Finetuning audio analysis
- Slack

