

Geelan Assaedi, Raeda Azkoul, Fatima Jabr, Ghofran Al-Siyami, Dalya Al-Amri.

Supervised by: Dr. Areej Alfraih.

Project ID: UQU-CS-2021S-07

Computer science department, College of computers and information systems,
Umm-Al-Qura University, Makkah, KSA, 2022.

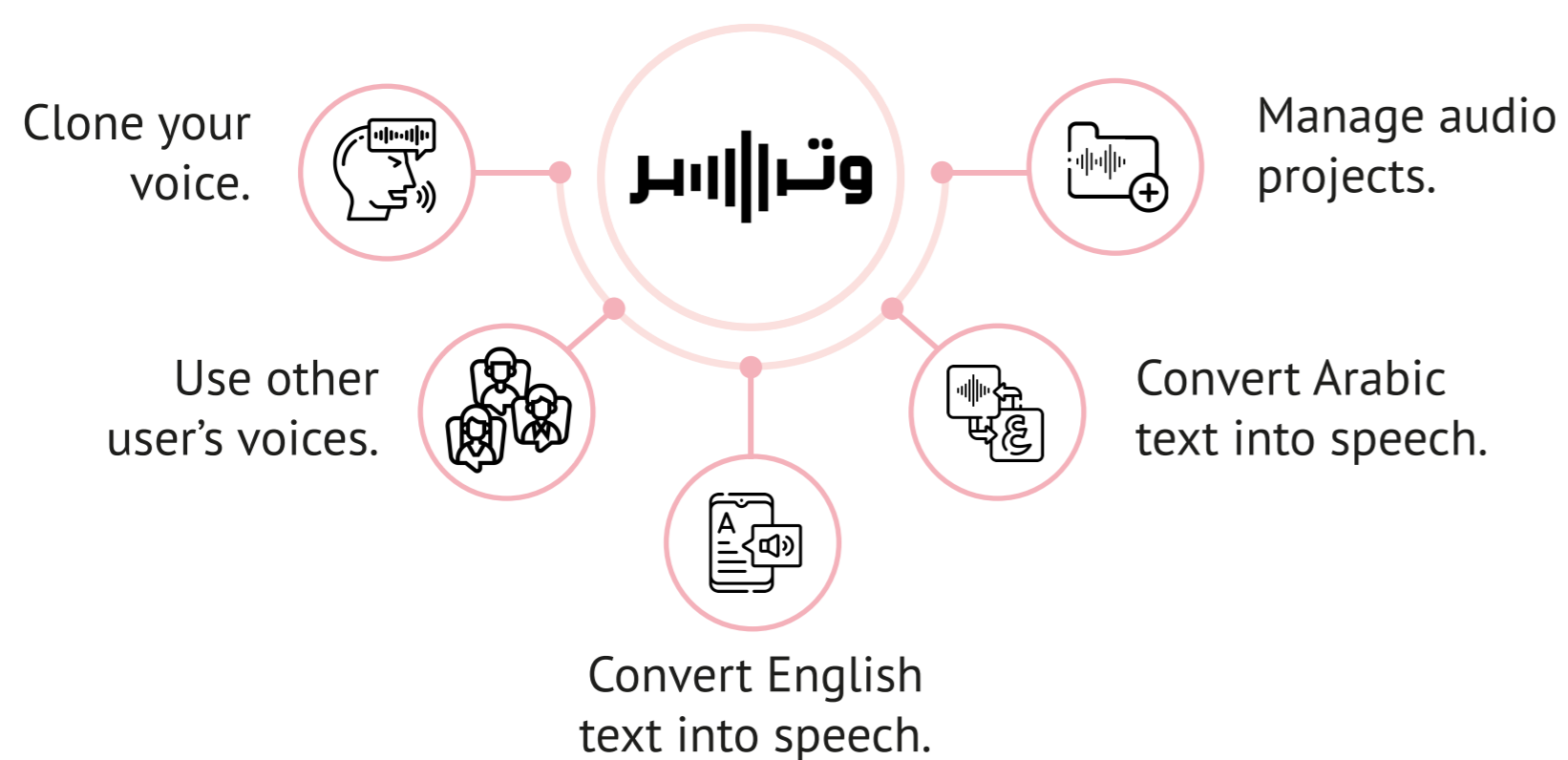
Abstract

"Watar" is the Arabic word for a string element that vibrates to produce sounds, in relation to how Watar platform performs, it can be considered a string that executes to produce speech. Watar platform provides a real-time voice cloning service to the public and in return presents mute people with a voice and facilitates the creation of audio content. To ensure such a platform is achievable surveys were conducted, a website was developed, and AI models were integrated. Thus, Watar has been evaluated by possible users of the website where it was rated 3.75 out of 5. And so Watar enables its users to manage their audio projects, manage their profiles, and it has a user-based voice library where users can publish their voices.

Objectives

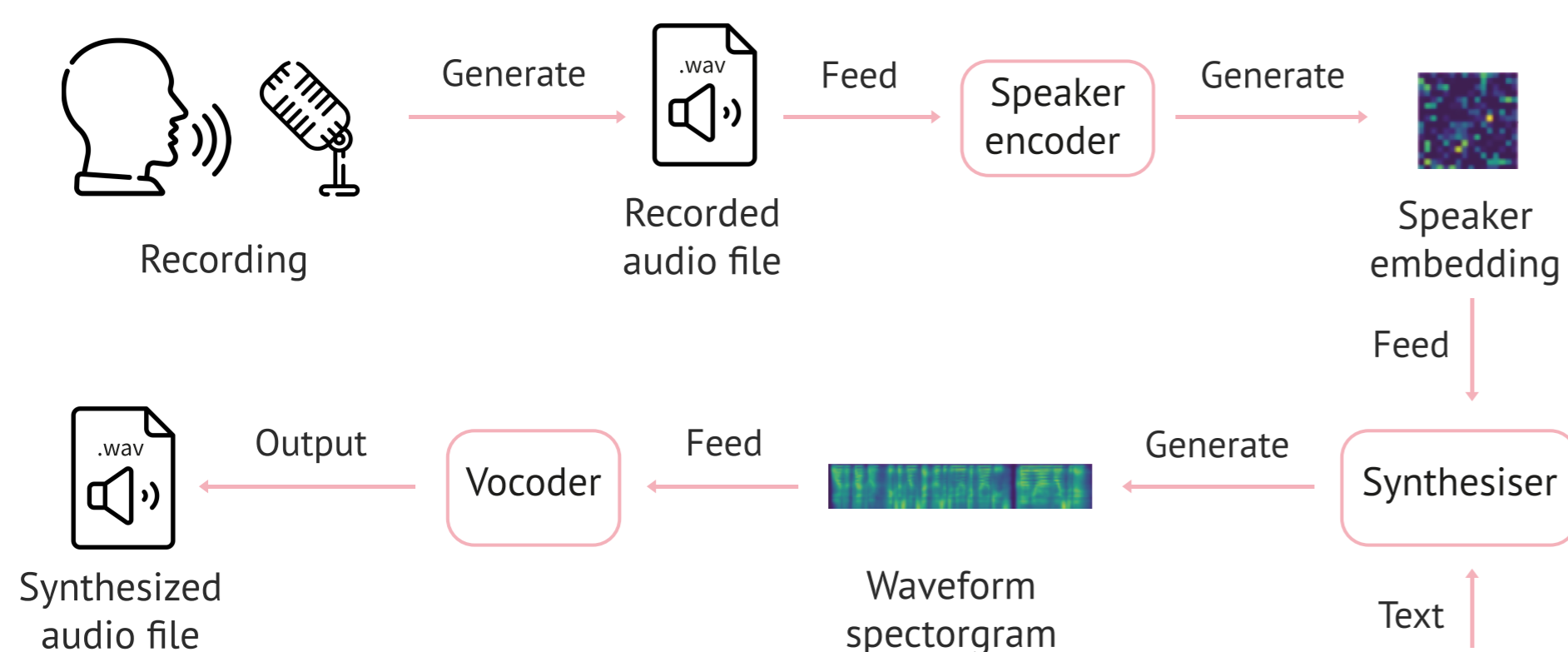
- To build a web platform which has several services to facilitate the creation of audio content.
- To build a real-time voice-cloning web platform using artificial intelligence technologies.

What can you do with WATAR?



Methodology

The voice cloning process starts with the user recording their voice, the waveform will be saved as a ".wav" audio file and fed to the model. The output will be a generated audio file of an artificial version of the user's voice speaking the input text.

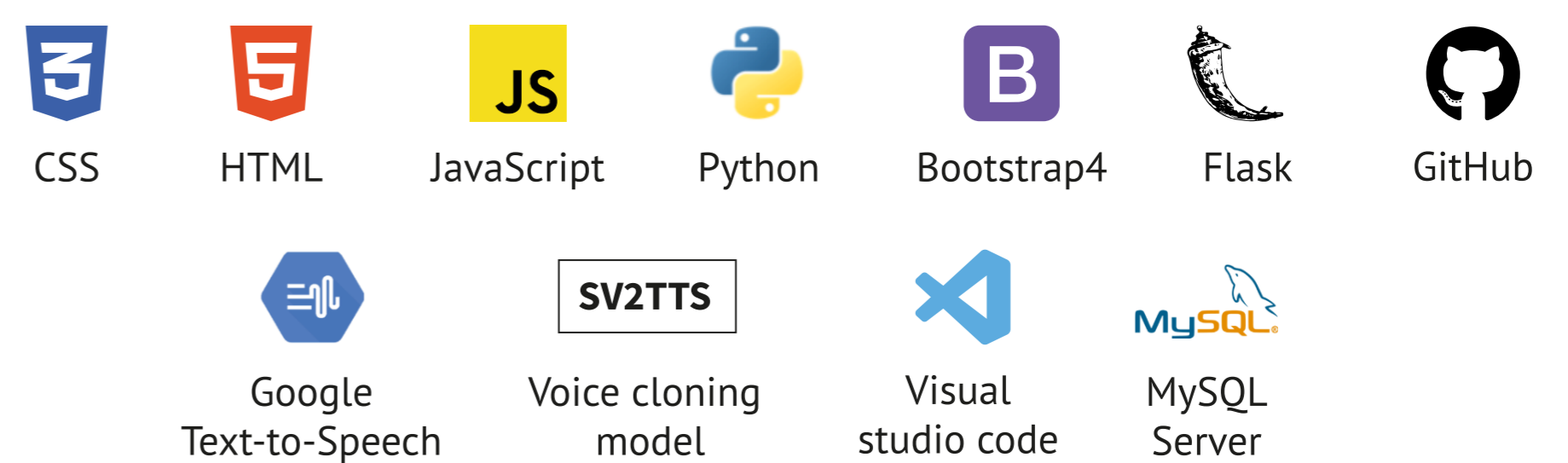


Related work



These high-rated systems are similar to Watar, and although they clone the user's voice, they do not execute in real-time. In addition to the fact that watar has an audio gallery based on users' contributions and even has fully supported Arabic interfaces.

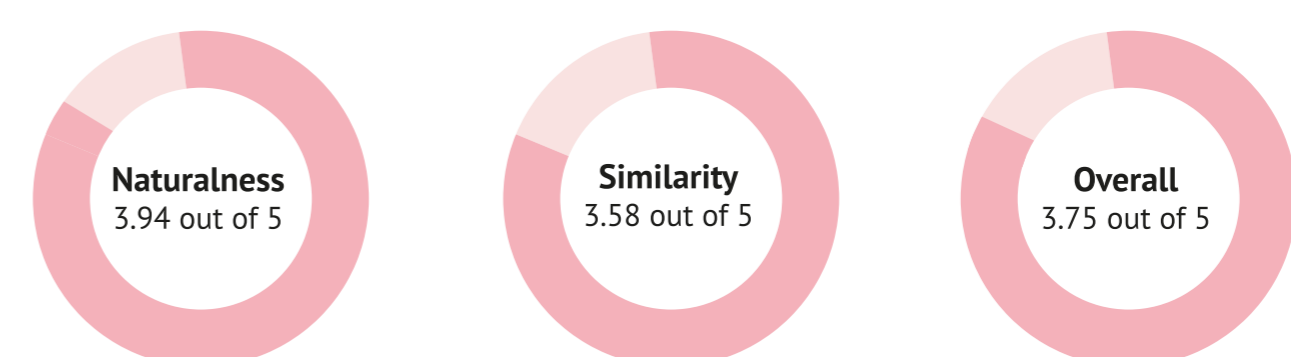
Development tools



Results

The quality results of the voice cloning model were collected by conducting a survey, and it shows that the average overall quality rating for Watar's voice cloning was 3.75 out of 5. The test contained two parts: naturalness test and similarity test.

The naturalness test average result was 3.94 out of 5
The similarity test average result was 3.58 out of 5.



Conclusion & Future work

This project implements WATAR application using Machine Learning and Web technologies to ease content creation by utilizing the most recent technologies, such as voice-cloning and Text-To-Speech for both the Arabic and English languages. This is only the start, the platform could be further improved by:

- Fully supporting Arabic voice cloning model.
- Adding speech recognition for security purposes.
- Adding new feature for selling and buying voices.
- Providing more control over the voice characteristics, such as tone and pauses.
- Adding an option to import Docx, PDF, or image to extract text and convert it into speech.

Contact

Team's leader: Raeda.azk@gmail.com
Supervisor: Asfraih@uqu.edu.sa

References



Useful link

