

Text-to-Speech App - Eleven Labs

Overview

The **Eleven Labs - TTS App** is a lightweight **HTML-based web tool** designed to convert text into realistic speech using the **ElevenLabs Text-to-Speech API**.

It provides an intuitive user interface for selecting voices, adjusting audio generation parameters, and downloading the generated speech output.

This app is primarily used for **testing, prototyping, and generating synthetic voice clips** directly from a browser.

Architecture

High-Level Flow

1. User Interface (Frontend):

- Built using **HTML, CSS, and basic JavaScript**.
- Allows the user to input text and configure TTS parameters (voice, style, stability, similarity boost, etc.).

2. Backend / API Integration:

- The app connects to **ElevenLabs TTS API**.
- It sends the text input and parameters as a JSON payload.
- The API returns an **audio file (MP3/WAV)** that can be played or downloaded.

3. Database Connection (optional):

- A minimal database is used to store **API keys, logs, or past generation history**.
- Example technologies: **MySQL** or **SQLite**.

System Diagram

```
graph TD
    A["[User Interface (HTML Page)]"] --> B["[JavaScript Fetch Request]"]
    B --> C["[Backend Script or Direct API Call]"]
    C --> D["[ElevenLabs TTS API]"]
    D --> E["[Generated Audio File (.mp3)]"]
    E --> F["[Playback / Download in Browser]"]
```

Tech Stack

Component	Technology Used
Frontend	HTML, CSS, JavaScript
API	ElevenLabs TTS REST API
Backend (optional)	Node.js / PHP (for API key management)
Database	MySQL or SQLite
Output Format	MP3 / WAV

User Interface Elements

Element	Function
---------	----------

Select Voice	Dropdown to choose available ElevenLabs voices (e.g., Clyde).
Similarity Boost	Slider (0–100%) controlling how closely the voice matches the reference tone.
Stability	Slider (0–100%) controlling consistency across speech samples.
Style	Slider (0–100%) adjusting expressiveness.
Use Speaker Boost	Boolean toggle for enhancing voice clarity.
Text Area	Input field for any text to be converted into speech.
Generate Button	Sends the text and settings to the API to create speech.
Download Button	Allows downloading the generated speech file.

Setup Instructions

1. Prerequisites

- **ElevenLabs API key** (sign up at <https://api.elevenlabs.io>)
- Basic web server (e.g., Node.js, PHP, or local HTML host)
- Optional: MySQL or SQLite database

2. Configuration

Add your ElevenLabs API key in your JavaScript or backend `.env` / config file:

```
const API_KEY = "YOUR_ELEVENLABS_API_KEY";
```

-

Configure API endpoint:

```
https://api.elevenlabs.io/v1/text-to-speech/{voice\_id}
```

-

3. Run the App

- Open the `index.html` file in a browser or host it on a local web server.
 - Input text, select desired parameters, and click **Generate**.
 - Wait for the generated audio and optionally click **Download**.
-

API Request Example

```
fetch("https://api.elevenlabs.io/v1/text-to-speech/Clyde", {
  method: "POST",
  headers: {
    "Content-Type": "application/json",
    "xi-api-key": API_KEY,
  },
  body: JSON.stringify({
    text: "Hello, this is ElevenLabs speaking.",
    voice_settings: {
      stability: 0.5,
      similarity_boost: 0.5,
      style: 0.5,
      use_speaker_boost: true
    }
  })
})
.then(res => res.blob())
.then(audioBlob => {
  const audioUrl = URL.createObjectURL(audioBlob);
  const audio = new Audio(audioUrl);
  audio.play();
});
```

Output

- The generated output is an **audio file (.mp3)**.

- The file can be **played directly in the browser** or **downloaded** using the “Download” button.