# LiveKit Voice Pipeline with Audio Length Validation

**Objective**

This project aims to create a voice assistant by leveraging LiveKit's Voice Pipeline Agent. A crucial part of the process is ensuring that the estimated duration of the audio generated doesn't exceed 60 seconds before being processed by the Text-to-Speech (TTS) engine. If the duration surpasses this limit, the backend server trims the audio to its central portion before sending it back.

**Key Features**

* LiveKit Voice Pipeline Agent integration
* Validation callback before the TTS step
* Flask backend for audio duration verification
* Trimming long audio outputs
* Text summarization via Hugging Face Transformers (bonus)
* Exposing Flask backend with Ngrok for remote access
* User interface (UI) for voice interaction and demonstration

**Implementation Overview**

## LiveKit Voice Pipeline Agent Setup

* The LiveKit Voice Pipeline Agent demo was cloned from the LiveKit GitHub repository and set up.
* All dependencies were installed, and the agent was confirmed to be working.
* A before\_tts\_cb callback was integrated to intercept text prior to TTS processing.

## Before-TTS Validation

* Extracts the input text from the voice stream.
* Estimates the expected audio length from the text.
* Sends the estimated length and text to the Flask server for validation.
* If the server returns modified text (e.g., trimmed or summarized), the original input is replaced before TTS conversion.

## Flask Api

* Developed a Flask-based REST API.
* API Endpoint: /validate\_audio to receive text and estimated audio length.
* If the audio length exceeds 60 seconds, the server trims the middle section of the text.
* Incorporates summarization model (t5-small to further optimize long responses.
* Returns the updated text to the LiveKit agent.

## Interaction Between LiveKit Agent and Flask Api

* Communication between the LiveKit agent and Flask server is handled using Python's requests library.
* ngrok is employed to expose the Flask server for remote accessibility during testing and development.

1. Voice Interaction Interface

* The LiveKit Voice Pipeline Agent is configured to manage voice interactions smoothly.
* Ensures the user experience is concise and clear.