# Summer Training TR-103 Prompt Engineering

## Day 7 Report

The seventh day of the training focused on introducing Speech-to-Text technology and its integration into prompt engineering. Participants explored a Python-based implementation using 'sounddevice', 'scipy', 'ipywidgets', and the OpenAI API with the 'whisper-1' model, featuring three interactive buttons: "Record Fixed Duration," "Record with Stop Button," and "Stop Recording." The session emphasized real-time audio capture, transcription, and accuracy considerations.

## Introduction to Speech-to-Text

Speech-to-Text is a technology that converts spoken language into written text, enhancing prompt engineering by enabling voice-based inputs for AI models. This session implemented a system using 'sounddevice' for audio recording, 'scipy' for file handling, 'ipywidgets' for an interactive interface, and the OpenAI 'whisper-1' model for transcription.

- The ultimate goal is to develop effective voice prompts that result in accurate text outputs.

- **Significance of Speech-to-Text in Prompt Engineering:**
  - Enables real-time voice-to-text conversion.
  - Enhances interactivity with AI systems.
  - Supports dynamic prompt generation.

- **Real-World Use Cases:**

1. **Voice Input:** Record and transcribe commands.
2. **Interactive Tools:** Use in real-time applications.
3. **Testing:** Evaluate transcription accuracy.

## Role of Libraries and Widgets in Code

- **sounddevice:** Captures audio data for recording.
- **scipy:** Handles saving recorded audio as files.
- **ipywidgets:** Creates buttons for interactive control.
- **whisper-1 (OpenAI):** Transcribes recorded audio into text using the OpenAI API.

## Functionality of the 3 Buttons

- **Record Fixed Duration:** Starts a set-duration recording.
- **Record with Stop Button:** Allows manual start and stop of recording.
- **Stop Recording:** Halts the recording process.

## Accuracy

- Accuracy, including punctuation, depends on audio quality and model training.
- Paid versions of the OpenAI API offer higher accuracy for punctuation and more.

## Conclusion

The seventh day of training provided a comprehensive introduction to Speech-to-Text within prompt engineering. Participants developed an understanding of how audio inputs are captured and transcribed using ‘sounddevice’, ‘scipy’, ‘ipywidgets’, and the OpenAI ‘whisper-1’ model, with hands-on practice using three interactive buttons. Key principles such as clarity, specificity, and effectiveness were emphasized, along with accuracy considerations, including the advantage of paid versions. This knowledge equips participants to create precise voice-based prompts, laying the groundwork for leveraging Speech-to-Text as a valuable skill.