

dwani's Goals

To integrate and enhance the following models and services for Kannada:

- **Automatic Speech Recognition (ASR):** To convert spoken Kannada into text.
 - a. Listen : Understands spoken Kannada-like a friend who gets you
- **Text-to-Speech (TTS):** To convert Kannada text into natural-sounding speech.
 - a. Speak - Talks back in kannada, naturally
- **Translation Services:** To enable translation between Kannada and other Indian languages.

Workshop - Prerequisites

- Python 3.10 (mandatory)
- Ubuntu 22.04 (Preferred)
- GitHub Account - <https://github.com>
- HuggingFace Account - <https://huggingface.co>
- Fork and clone <https://github.com/dwani-ai/dwani-workshop>

Workshop Plan

- Walk through of Dwani GitHub Projects for reference
- Build UX with Python to explore Use-cases for Voice AI with Kannada

Kannada LLM Chat

- <https://huggingface.co/spaces/dwani/dwani-ai-chat>
- <https://github.com/dwani-ai/dwani-workshop/blob/main/src/chat-dwani.py>

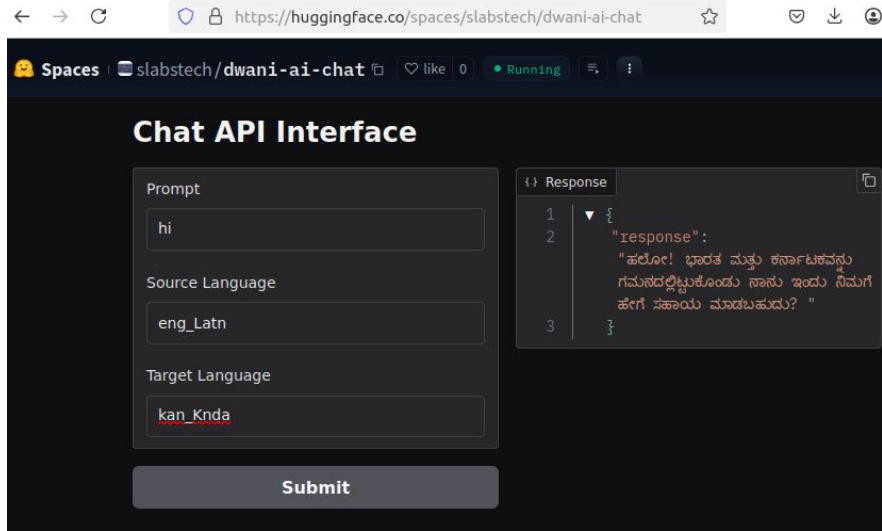
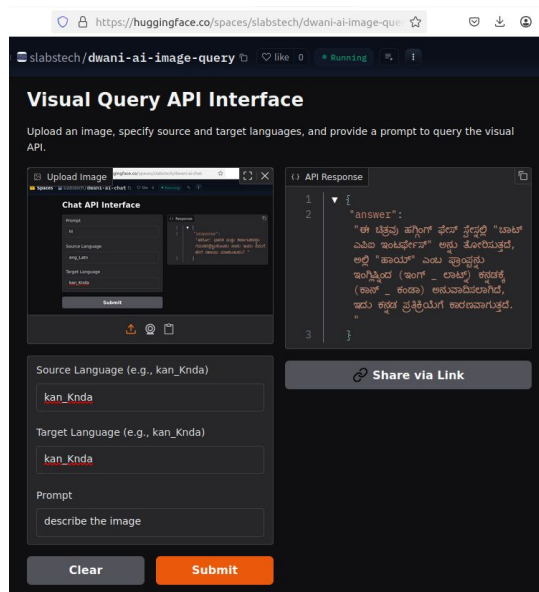


Image Query

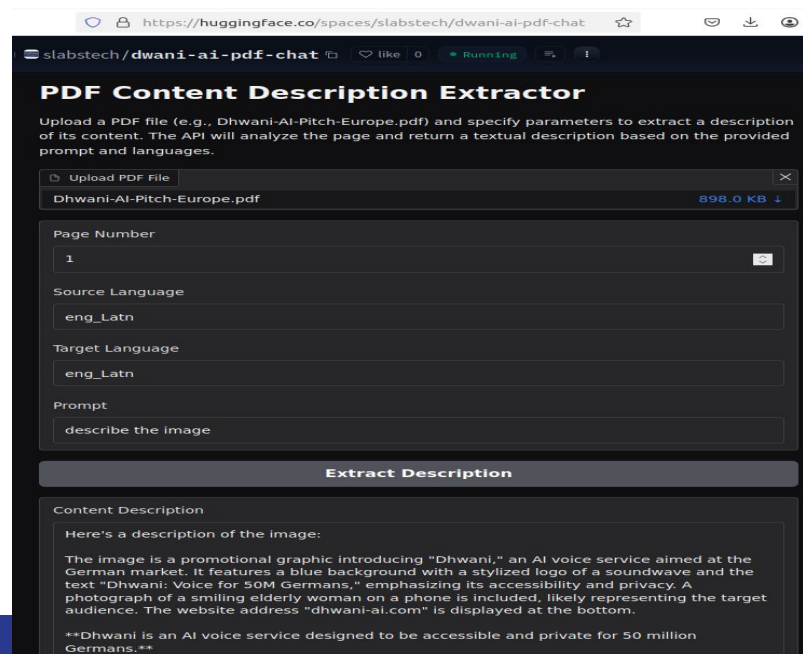
- <https://huggingface.co/spaces/dwani/dwani-ai-image-query>
- <https://github.com/dwani-ai/dwani-workshop/blob/main/src/image-query.py>



PDF Chat



- <https://huggingface.co/spaces/slabstech/dwani-ai-pdf-chat>
- <https://github.com/slabstech/dwani-workshop/blob/main/src/pdf-chat-dwani.py>

A screenshot of a web application titled 'PDF Content Description Extractor'. The interface is dark-themed. At the top, it says 'Upload a PDF file (e.g., Dhwani-AI-Pitch-Europe.pdf) and specify parameters to extract a description of its content. The API will analyze the page and return a textual description based on the provided prompt and languages.' Below this, there's a file upload section with a button 'Upload PDF File' and a file named 'Dhwani-AI-Pitch-Europe.pdf' (898.0 KB). There are input fields for 'Page Number' (set to 1), 'Source Language' (eng_Latn), and 'Target Language' (eng_Latn). A 'Prompt' field contains the text 'describe the image'. A large 'Extract Description' button is below these fields. At the bottom, a 'Content Description' box shows the result: 'Here's a description of the image: The image is a promotional graphic introducing "Dhwani," an AI voice service aimed at the German market. It features a blue background with a stylized logo of a soundwave and the text "Dhwani: Voice for 50M Germans," emphasizing its accessibility and privacy. A photograph of a smiling elderly woman on a phone is included, likely representing the target audience. The website address "dhwani-ai.com" is displayed at the bottom. **Dhwani is an AI voice service designed to be accessible and private for 50 million Germans.**'

Kannada - PDF Extraction, Translation and Creation

- <https://huggingface.co/spaces/slabstech/kannada-pdf-prompt>
- <https://github.com/slabstech/dwani-workshop/blob/main/src/vllm/kannada-pdf.py>

slabstech/kannada-pdf-prompt

like

0

Running

1

Kannada - PDF Query and translation

Upload a PDF, specify a page number, prompt, and source language. The PDF is sent to an API, and both the input and output PDFs are previewed.

Upload PDF

ITA-software-travel-complexity.pdf1.1 MB

Page Number

1

Prompt

list the points

Source Language

eng_Latn

Clear

Submit

Input PDF Information

Input PDF Info:

- Number of pages: 51

Input PDF Preview (First Page)

Output PDF Information

Output PDF Info:

- Number of pages: 1

Output PDF Preview (First Page)

Automatic Speech Recognition (ASR)

- <https://huggingface.co/spaces/slabstech/asr-transcription>
- <https://github.com/slabstech/dwani-workshop/blob/main/src/transcribe-dwani.py>

