

# Meet dwani.ai

dwani.ai: Knowledge from Curiosity

Accessible, Private, Open Source, Free



dwani.ai

# Anita's Story

- Anita cannot use Voice Tech
- English Only, Not Private



# Why We Need dwani

- Ever tried asking Siri or Alexa something in Kannada ? They don't get it !
- Over 5 crore people speak kannada, but most voice tech ignores them
- dwani uses Open Source Tools to fix this - built for YOU !

# What Makes dwani Special ?

- Its **FREE** for anyone to use or tweak-like a community project !
- Runs on your premise, keeping your chats private.
- Uses Open-weight Models released by AI4Bharat, Google Deepmind, OpenAI

# dwani.ai Team

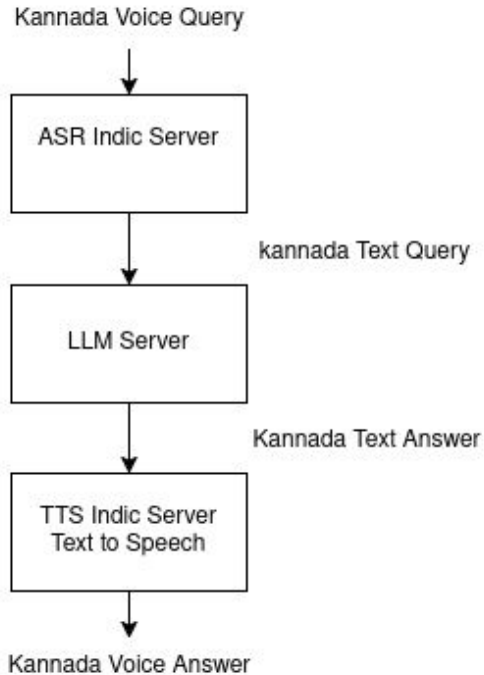
- Sachin Shetty - Software Engineer, 10+ year S/W , 2+ year GenAI
- Sahana Shetty - Robotics Engineer, “Human-Assisting Apparatus”, Indian Patent - 562119
- Nithish S - Embedded Engineer, Outreach and Community Management

# Traction and Recognition

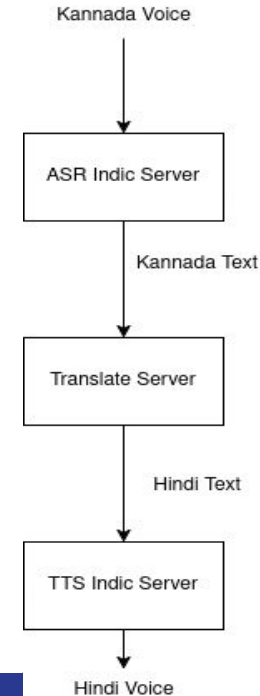
- **Zerodha** Open Source - Rs. 2.3 Lakhs Grant
- Part of **Nvidia Inception Program**
- Selected for
  - a. **Lambda Labs Startup Program** - USD 7,500 GPU Credits
  - b. **Google for Startups Program** - USD 25,000 Cloud credits
  - c. **Yotta Labs- Rudra Startup Program** - USD 50,000 GPU Credits

# What can Dwani Do?

## Chat In Kannada



## Translate on the Fly



# Current Status - July 8, 2025

Answer -

Kannada,  
German,  
English

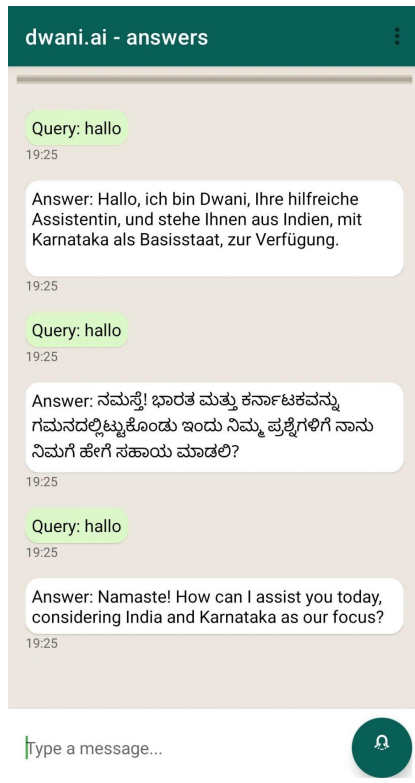
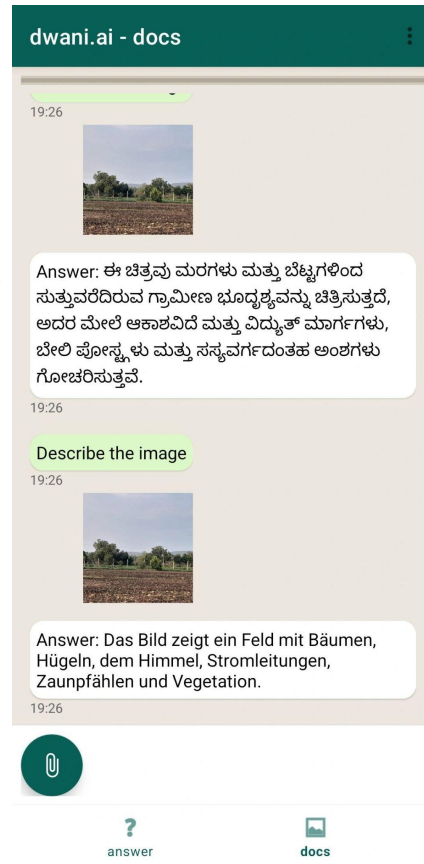


Image Query





# dwani.ai - python-sdk

- pip install --upgrade dwani
- Credentials to use dwani.ai API
  - DWANI\_API\_BASE\_URL = `https://beta-api.dwani.ai`
  - DWANI\_API\_KEY = `email_id + _dwani`
  - Ex : `email_id` : `user-email@example.com`
  - DWANI\_API\_KEY = `user-email@example.com_dwani`
- Languages Supported
  - Hindi, Kannada, Malayalam, Marathi, Odia, Punjabi, Tamil, Telugu
  - English, German

## [dwani.ai](https://dwani.ai) - Android App



# 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
- HuggingFace Account
- Fork and clone

[GitHub](#)

[Hugging Face](#)

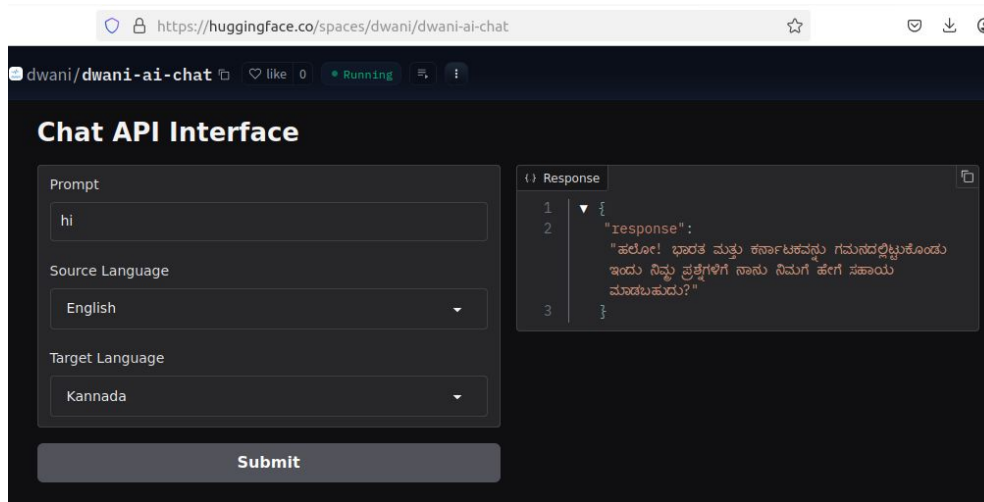
<https://github.com/dwani-ai/workshop>

# Workshop Plan

- Walk through of dwani.ai GitHub Projects for reference
- Build UX with Gradio/Python to explore Use-cases for Voice AI with Kannada + Indian Languages

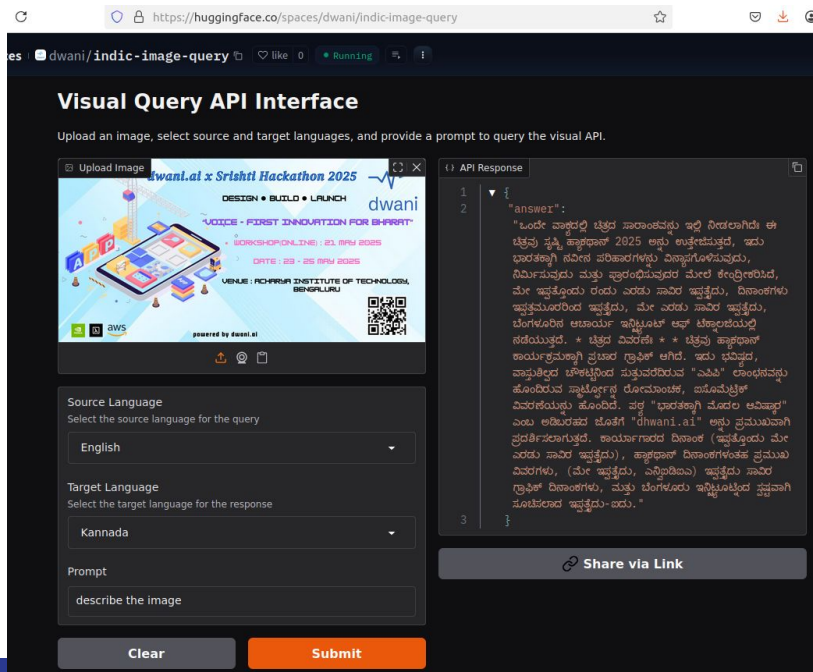
# Kannada LLM Chat

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




# Vision Query

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



Visual Query API Interface

Upload an image, select source and target languages, and provide a prompt to query the visual API.

Upload Image: 

Source Language: English

Target Language: Kannada

Prompt: describe the image

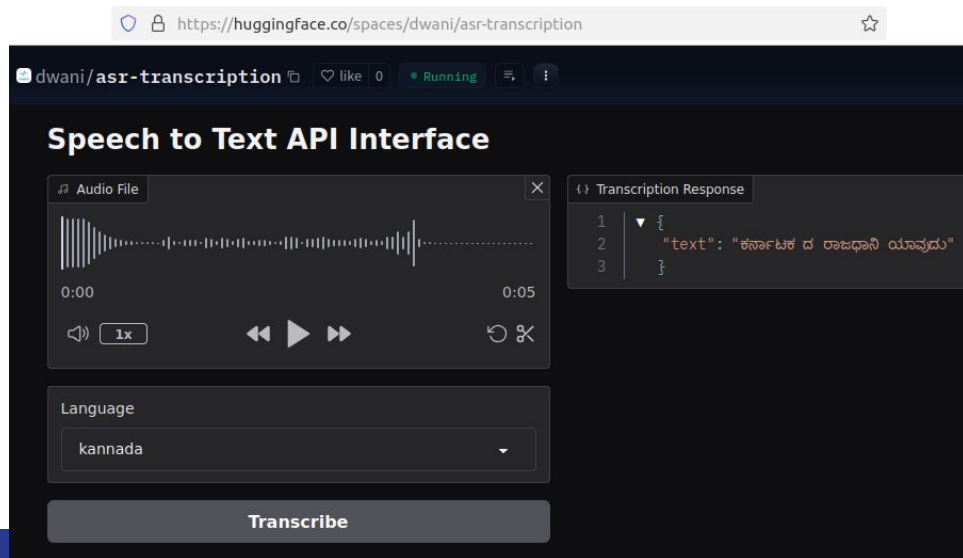
API Response:

```
1 {
2   "answer":
3 }
```

Share via Link

# Automatic Speech Recognition (ASR)

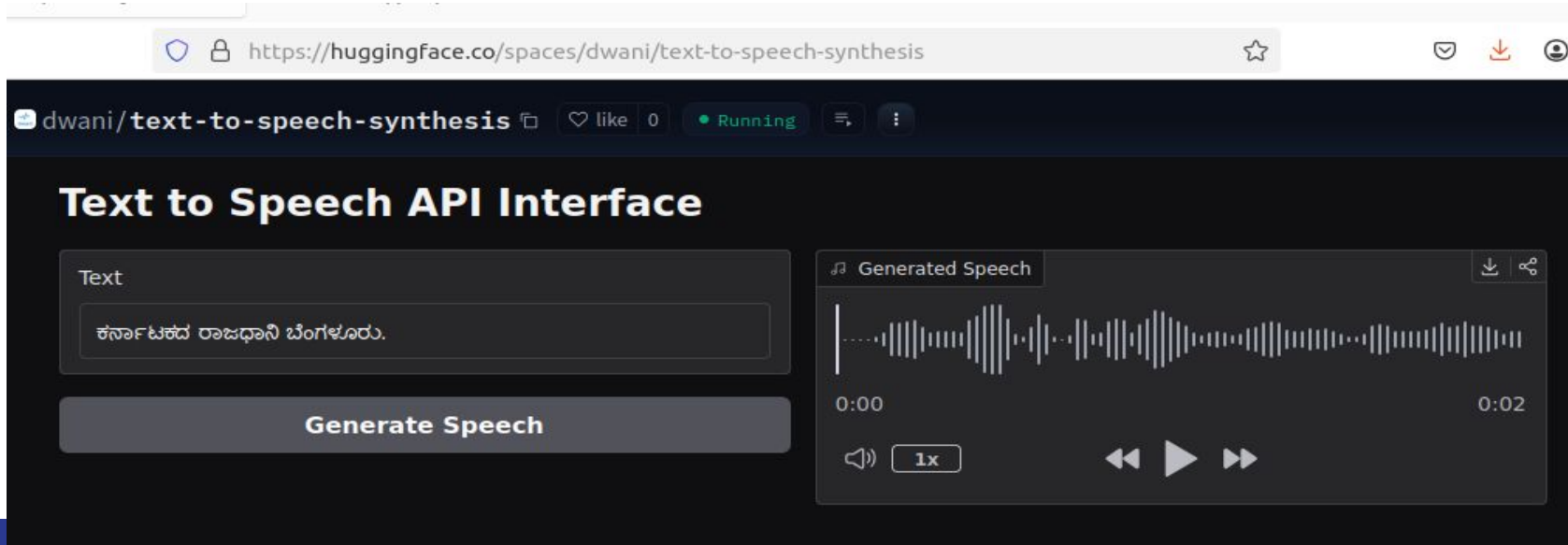
- <https://huggingface.co/spaces/dwani/asr-demo>
- <https://github.com/dwani-ai/workshop/blob/main/src/transcribe-dwani.py>





# Text to Speech

- <https://huggingface.co/spaces/dwani/tts-demo>
- <https://github.com/dwani-ai/workshop/blob/main/src/text-to-speech-dwani.py>



The screenshot shows a web browser window displaying the Hugging Face Text to Speech API Interface. The URL bar shows <https://huggingface.co/spaces/dwani/text-to-speech-synthesis>. The interface has a dark theme. At the top, it says "dwani/text-to-speech-synthesis" with a "like" button showing 0 likes and a "Running" status indicator. The main heading is "Text to Speech API Interface". Below this, there is a text input field labeled "Text" containing the Kannada text "ಕರ್ನಾಟಕದ ರಾಜಧಾನಿ ಬೆಂಗಳೂರು." (Karnataka's capital is Bengaluru). Below the input field is a large button labeled "Generate Speech". To the right of the input field is a section titled "Generated Speech" which displays a waveform of the generated audio. Below the waveform is a playback control bar showing a duration of 0:00 to 0:02, a volume icon, and a speed selector set to "1x".

# PDF Chat

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

https://huggingface.co/spaces/dwani/indic-pdf-chat-custom

dwani/indic-pdf-chat-custom like 0 Running

## PDF Custom Prompt Processor

Upload a PDF, specify a page number, enter a prompt, and select source and target languages.

Upload PDF AlphaEvolve.pdf 3.2 MB

Page Number 1

Custom Prompt List the key points

Source Language English

Target Language Kannada

Process

elaborate on any of these points, or perhaps focus on a specific aspect of the document?"

"Processed Page": 1,

"Translated Response":

"ಸರಿ, ಡಾಕ್ಯುಮೆಂಟ್ ಪ್ರಮುಖ ಅಂಶಗಳ ಪಟ್ಟಿ ಇಲ್ಲಿವೆ \* \* \* ಅದು ಏನು \* \* ಗೂಗಲ್ ಡಿವೈಡ್ ಅಭಿವೃದ್ಧಿಪಡಿಸಿದ ಕೋಡಿಂಗ್ ಏಜೆಂಟ್ ಆಲ್ಫಾವೆಲ್ಡ್ ಆಗಿದೆ. \* \* \* ಉದ್ದೇಶ: \* \* ಇದು ಸಂಕೀರ್ಣ ವೈಜ್ಞಾನಿಕ ಮತ್ತು ಕ್ರಮಾವಳಿ ಸಮಸ್ಯೆಗಳನ್ನು ನಿಭಾಯಿಸಲು ಮತ್ತು ಕಂಪ್ಯೂಟೇಶನಲ್ ಮೂಲಸೌಕರ್ಯವನ್ನು ಉತ್ತಮಗೊಳಿಸಲು ಎಲ್‌ಎಲ್‌ಎಂಗಳನ್ನು ಹೆಚ್ಚುತ್ತದೆ. \* \* ಇದು ಹೇಗೆ ಕಾರ್ಯನಿರ್ವಹಿಸುತ್ತದೆ: \* \* ಇದು ಸ್ವಾಯತ್ತ ವಿಶ್ಲೇಷಣೆಯ ವಿಧಾನವನ್ನು ಬಳಸುತ್ತದೆ, ಮೌಲ್ಯಮಾಪಕರಿಂದ ಪ್ರತಿಕ್ರಿಯೆಯ ಆಧಾರದ ಮೇಲೆ ಕೋಡ್ ಅನ್ನು ಪುನರಾವೇಶವಾಗಿ ಸುಧಾರಿಸುತ್ತದೆ. \* \* \* ಪ್ರಮುಖ ಸಾಧನೆಗಳು \* \* \* ಹೆಚ್ಚು ಪರಿಣಾಮಕಾರಿ ಡೇಟಾ ಸೆಂಟರ್ ಟೆನ್ಸೊರಿಂಗ್ ಅಲ್ಲಾರಿದಮ್ ಅನ್ನು ಅಭಿವೃದ್ಧಿಪಡಿಸಿದೆ, 56 ವರ್ಷಗಳ ನಂತರ ಸ್ವಾಸ್ಥ್ಯ ಅಲ್ಲಾರಿದಮ್ ಅನ್ನು ಮೀರಿಸಿದೆ. \* 4x4 ಸಂಕೀರ್ಣ-ಮೌಲ್ಯದ ಮ್ಯಾಟ್ರಿಕ್ಸ್‌ನ್ನು (48 ಸ್ತರೀಯ ಗುಣಕಾರಗಳು, 56 ವರ್ಷಗಳ ನಂತರ ಸ್ವಾಸ್ಥ್ಯ ಮೀರಿರುತ್ತದೆ) ಗುಣಿಸಲು ವೇಗವಾಗಿ ವಿಧಾನವನ್ನು ಕಂಡುಹಿಡಿದಿದೆ. \* ಸರಳೀಕೃತ ವೇಗವರ್ಧಕ ವಿನ್ಯಾಸವು ಹೊಸ ಹಾರ್ಡ್‌ವೇರ್ ಸ್ಕ್ರೂಟ್ ವಿಧಾನಗಳನ್ನು ಬಳಸುತ್ತದೆ."

# Research Goals / Collaboration

- Measure and improve the Time to First Token Generation (TTFTG) for model architectures in ASR, Translation, and TTS systems.

Tool	Source
Document Processing	<a href="https://github.com/dwani-ai/docs-indic-server">https://github.com/dwani-ai/docs-indic-server</a>
Automatic Speech Recognition : ASR	<a href="https://github.com/dwani-ai/asr-indic-server">https://github.com/dwani-ai/asr-indic-server</a>
Text to Speech : TTS	<a href="https://github.com/dwani-ai/tts-indic-server">https://github.com/dwani-ai/tts-indic-server</a>
Translation	<a href="https://github.com/dwani-ai/indic-translate-server">https://github.com/dwani-ai/indic-translate-server</a>
API Server	<a href="https://github.com/dwani-ai/dwani-api-server">https://github.com/dwani-ai/dwani-api-server</a>
Dwani Android App	<a href="https://github.com/slabstech/dwani-android">https://github.com/slabstech/dwani-android</a>

# Problem statements - Health

- Multilingual Symptom Checker
- Voice-Enabled Appointment Booking System
- Health Education Voice Platform
- Assistive Tool for Visually Impaired Patients
- Voice-Driven Telemedicine Connector
- Mental Health Support Voice Bot
- Voice-Based Rehabilitation Guide
- Multilingual Emergency Response Assistant
- Voice-Enabled Health Record Access
- Assistive Tool for Physically Challenged Patients

# Problem statements - Public Service and Government

- Multilingual Grievance Filing System
- Voice-Enabled Public Scheme Navigator
- Transparent RTI Query Tool
- Citizen Feedback Collector
- Voice-Driven e-Governance Chatbot
- Accessible Voting Information System
- Smart Public Announcement System
- Multilingual Legal Aid Assistant
- Voice-Based Public Transport Query System
- Citizen Safety Reporting Tool

## Q/A

- What are your feedback for Workshop ?
- What would you like to build?
- Slides Download - <https://tinyurl.com/dwani-ai-workshop>

