A screenshot of a project layout

AI-generated content may be incorrect.

A screenshot of a white background

AI-generated content may be incorrect.

**Chatbot 2.0 — AI in Healthcare (Vision + Voice)**

**📌 Project Overview**

Chatbot 2.0 is an AI-powered healthcare assistant that combines **multimodal vision and voice interaction**. The chatbot simulates a doctor–patient conversation using:

* Image inputs (e.g., scans or charts)
* Patient speech (via microphone)
* Text-to-speech responses (doctor’s voice)

**📁 Project Layout**

**Phase 1 – Setup the Brain of the Doctor (Multimodal LLM)**

* Configure **GROQ API Key** for fast inference.
* Convert input image to required format.
* Load and run **Multimodal LLM** (e.g., LLaMA 3 Vision).

**Phase 2 – Setup Voice of the Patient**

* Implement audio recording using ffmpeg & portaudio.
* Use **OpenAI Whisper** for speech-to-text transcription (STT).

**Phase 3 – Setup Voice of the Doctor**

* Use **Text-to-Speech (TTS)** models like:
  + gTTS (Google Text-to-Speech)
  + ElevenLabs for realistic synthesis
* Convert model-generated text into voice output.

**Phase 4 – Setup UI for the VoiceBot**

* Build an interactive interface using **Gradio**.
* Capture image uploads and audio inputs.
* Display responses in both text and voice formats.

**🧰 Tools and Technologies**

| **Component** | **Tool / Framework** |
| --- | --- |
| AI Inference | **Groq API** |
| Vision Model | **LLaMA 3 Vision (Meta)** |
| Speech-to-Text | **OpenAI Whisper** |
| Text-to-Speech | **gTTS, ElevenLabs** |
| UI | **Gradio** |
| Language | **Python** |
| IDE | **Visual Studio Code (VS Code)** |

**⚙️ Technical Architecture**

rust

CopyEdit

[Patient]

|

|-- Phase 2: Audio Recorder --> Whisper STT

|

|-- Transcribed Text ----------> Vision + LLM (Groq)

|

[User Image] --> Upload -----> Vision Model (Meta LLaMA 3)

|

LLM Response (Text)

|

|-- Phase 3: Text --> TTS --> Audio Output

|

[Doctor Voice] --> Gradio UI (Phase 4)

A diagram of a technical architecture

AI-generated content may be incorrect.

**✅ Setup Instructions**

**1. Clone the Repository**

bash

CopyEdit

git clone https://github.com/your-repo/healthcare-voicebot.git

cd healthcare-voicebot

**2. Create Virtual Environment**

bash

CopyEdit

python -m venv venv

source venv/bin/activate # On Windows: venv\Scripts\activate

**3. Install Dependencies**

bash

CopyEdit

pip install -r requirements.txt

**4. Run Locally**

bash

CopyEdit

python app.py # or gradio\_app.py depending on UI implementation

**🚀 Future Enhancements**

* Add chatbot memory for context retention
* Multilingual support
* HIPAA-compliant data handling
* Integration with EHR/EMR systems