

PROJETO: Sistema de Dublagem Automática de Vídeos com IA

Objetivo:

Criar um pipeline que recebe um vídeo e gera uma versão dublada usando IA (em outro idioma ou na mesma língua), com sincronização simples.

Arquitetura Geral:

Vídeo → extração de áudio → transcrição (Whisper API) → tradução (LLM)

→ geração de nova voz (TTS) → ajuste de duração → remonte do vídeo (FFmpeg)

Estrutura das Pastas:

project/

■■■■ input/

■ ■■■■ video.mp4

■■■■ output/

■ ■■■■ audio_new.wav

■ ■■■■ video_final.mp4

■ ■■■■ transcript.srt

■■■■ scripts/

■ ■■■■ 01_extract_audio.py

■ ■■■■ 02_transcribe.py

■ ■■■■ 03_translate.py

■ ■■■■ 04_generate_voice.py

■ ■■■■ 05_align_audio.py

■ ■■■■ 06_merge_video.py

■ ■■■■ pipeline.py

■■■ requirements.txt

■■■ README.md

Funcionalidades do Projeto:

1. Extração do Áudio:

- Usa FFmpeg
- Salva em WAV 16kHz

Input: video.mp4

Output: audio_original.wav

2. Transcrição com Whisper API:

- Usa OpenAI Whisper via API
- Gera texto e arquivo .srt com timestamps

Output: transcript.srt, transcript.txt

3. Tradução via LLM:

- Mantém estrutura por frases
- Mantém timestamps do SRT

Output: transcript_translated.srt

4. Geração de Voz (TTS):

- Usa OpenAI TTS ou Coqui TTS API
- Mantém ritmo aproximado da frase original

Output: audio_new.wav

5. Ajuste simples de duração:

- Usa pydub
- Encaixa o áudio nos timestamps

Output: audio_aligned.wav

6. Remontagem do vídeo final:

- Usa FFmpeg
- Substitui áudio antigo pelo novo

Output: video_final.mp4

7. Automatização (Pipeline):

```
python pipeline.py --video input/video.mp4 --lang en
```

Tecnologias usadas:

Python 3, FFmpeg, Whisper API, GPT, TTS, Pydub, MoviePy

Diferenciais de Portfólio:

- Escolha de voz
- Idiomas múltiplos
- Modo voz parecida
- Interface Streamlit/Gradio

Escopo do MVP:

- Carrega vídeo
- Transcreve
- Traduz

- Gera voz

- Monta vídeo final

Versão Plus (opcional):

- Escolha de voz

- Lip-sync leve via Wav2Lip (Colab)