```
jarvis-friday-ai/
   – frontend/
      interface.html
                            # All-in-one HTML/CSS/JS
      assets/
                           # Static files (optional)
    backend/
      - app/
        — main.py
                         # FastAPI app
                           # Configurations
         config.py
         – api/
             voice.py
                             # /voice/*
            # STT: Whisper (free), AssemblyAl (paid), Deepgram (paid)
            # TTS: pyttsx3 (free), ElevenLabs (paid), PlayHT (paid)
            — vision.py
                             # /vision/*
            # Object Detection: YOLOv8 (free), Azure CV (paid)
            # Face Recognition: FaceNet (free), AWS Rekognition (paid)
            # Gestures: MediaPipe (free)
                                # /commands/*
             commands.py
            # NLP: GPT-3.5/4 (paid), Llama 2 (free), Claude (paid)
            # Dialog: Rasa (free), Dialogflow (paid)
                                # /environment/*
             environment.py
           # Weather: OpenWeather (free), Tomorrow.io (paid)
           # Location: Mapbox (free tier), Google Maps (paid)
           services/
            — speech.py
                              # STT/TTS service
            # Recommended: Whisper + ElevenLabs combo
                            # Computer vision
            vision.py
            # Recommended: YOLOv8 + MediaPipe
                           # NLP processing
             – nlp.py
            # Recommended: GPT-4 + local Llama 2 fallback
            integrations.py # External APIs
           # API clients for weather, maps, etc.
          - utils/
                       # Audio processing
            audio.py
```

```
# Video processing
            video.py
                            # Generic utilities
             helpers.py
       - models/
                             # Local Whisper models
         — whisper/
                           # YOLO weights
          – yolo/
          – Ilama/
                            # Local LLM files
                          # Frontend build
       - static/
       - requirements.txt
                               # Python deps

    Dockerfile

                            # Container config
                                # Documentation
    - README.md
Key API Integrations:
1. Voice Processing:
Free:
Whisper (STT)
pyttsx3/espeak (TTS)
Paid (
0.006
0.006-0.02 per 1k chars):
ElevenLabs (best quality TTS)
PlayHT (alternative TTS)
Deepgram (low-latency STT)
2. Computer Vision:
Free:
YOLOv8 (object detection)
MediaPipe (face/pose/hands)
```

Paid (\$1-5 per 1k images):

AWS Rekognition
Google Vision AI
Clarifai
3. Natural Language:
Free:
Llama 2 (local)
Rasa (dialog)
Paid (0.002
0.002-0.12 per 1k tokens):
OpenAl GPT-4
Anthropic Claude
Google Gemini
4. Biometrics:
Free:
HeartPy (webcam PPG)
OpenFace (facial metrics)
Paid:
Empatica E4 API (\$1k+ hardware)
Biostrap API (wearable data)
5. Environment:
Free:

OpenWeatherMap (1k req/day)
AirVisual (air quality)
Paid:
Tomorrow.io (\$20+/mo)
Google Maps Geocoding (\$5/1k req)