```
jarvis-friday-ai/
│
├── frontend/
│   └── interface.html              # All HTML/CSS/JS in one file
│
├── backend/
│   ├── app/
│   │   ├── __init__.py
│   │   ├── main.py                 # FastAPI app entry point
│   │   ├── config.py               # Configuration
│   │   │
│   │   ├── api/
│   │   │   ├── voice_api.py        # Endpoints: /voice/*
│   │   │   │   # Requires: Whisper (STT), ElevenLabs (TTS - paid) or pyttsx3 (free)
│   │   │   │
│   │   │   ├── vision_api.py       # Endpoints: /vision/*
│   │   │   │   # Requires: OpenCV, YOLOv8 (free) or AWS Rekognition (paid)
│   │   │   │
│   │   │   ├── command_api.py      # Endpoints: /command/*
│   │   │   │   # Requires: OpenAI API (paid) or Llama 2 (free)
│   │   │   │
│   │   │   ├── biometric_api.py    # Endpoints: /biometrics/*
│   │   │   │   # Requires: HeartPy (free) or Empatica API (paid)
│   │   │   │
│   │   │   └── env_api.py          # Endpoints: /environment/*
│   │   │       # Requires: OpenWeatherMap (free/paid), Google Maps API (paid)
│   │   │
│   │   ├── services/
│   │   │   ├── stt_service.py      # Speech-to-text service
│   │   │   │   # Options: Whisper (free), AssemblyAI (paid), Google Speech-to-Text (paid)
│   │   │   │
│   │   │   ├── tts_service.py      # Text-to-speech service
│   │   │   │   # Options: ElevenLabs (paid), pyttsx3 (free), Google TTS (free)
│   │   │   │
│   │   │   ├── vision_service.py   # Computer vision
│   │   │   │   # Options: YOLOv8 (free), MediaPipe (free), Azure Computer Vision (paid)
│   │   │   │
│   │   │   ├── nlp_service.py      # Natural language processing
│   │   │   │   # Options: GPT-4 (paid), Llama 2 (free), Rasa (free)
│   │   │   │
│   │   │   └── biometric_service.py # Health analysis
│   │   │       # Options: HeartPy (PPG analysis), OpenBCI (EEG - paid hardware)
│   │   │
│   │   └── utils/
```

```
│   │   ├── audio_utils.py      # Audio processing helpers
│   │   ├── video_utils.py      # Video frame processing
│   │   └── api_client.py       # External API wrappers
│   │
│   ├── models/
│   │   ├── yolo/               # Pretrained YOLO models
│   │   ├── whisper/            # Whisper model files
│   │   └── llama/              # Local LLM files (if using)
│   │
│   ├── tests/
│   ├── requirements.txt        # Python dependencies
│   └── Dockerfile
│
├── data/
│   ├── audio_samples/          # Voice recordings
│   ├── face_profiles/          # Known face embeddings
│   └── env_data/               # Cached environmental data
│
├── docs/                       # Documentation
│   └── API_REFERENCE.md        # API endpoints
│
└── 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):

OpenAI 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)