

Hackathon Project Phases Template for the **Audio transcription app** using **OpenAI Whisper** project.

Hackathon Project Phases Template

Project Title:

Audio transcription app using OpenAI Whisper

Team Name:

Mindsparks

Team Members:

- M.Ankhitha
 - L.Rishanka
 - M.Pooja
 - T.Richa
-

Phase-1: Brainstorming & Ideation

Objective:

The objective of an **Audio Transcription App using OpenAI Whisper** is to convert spoken language from audio files or live recordings into accurate text using Whisper's powerful speech-to-text capabilities.

Key Points:

1. Problem Statement:

- Many industries require accurate and efficient audio transcription, but existing solutions are costly, inaccurate in noisy environments, and lack multilingual support.
- Manually transcribing audio is slow, prone to errors, and often lacks support for multiple languages.

2. Proposed Solution:

- Develop an Audio Transcription App powered by OpenAI's Whisper model to provide highly accurate, multilingual, and cost-effective transcription services.
- The app will allow users to upload audio files or record speech in real-time.

3. Target Users:

- Content creators (podcasters, YouTubers, journalists)
- Businesses (meeting transcriptions, customer support)
- Students & Researchers (lecture notes, interviews)

4. Expected Outcome:

- The audio transcription app using OpenAI Whisper will support live and batch processing, speaker identification, and noise-resistant transcriptions.

Phase-2: Requirement Analysis

Objective:

Define the technical and functional requirements for the Audio transcription app using OpenAI Whisper.

Key Points:

1. Technical Requirements:

- Programming Language: **Python**
- Backend: **FastAPI or Django**
- Frontend: **React.js or Flutter**
- Database: **Not required initially (API-based queries)**

2. Functional Requirements:

- Audio Input: Allow users to upload or record audio files.
- Transcription Service: Automatically transcribe audio using OpenAI Whisper.
- Multilingual Support: Detect and transcribe multiple languages.

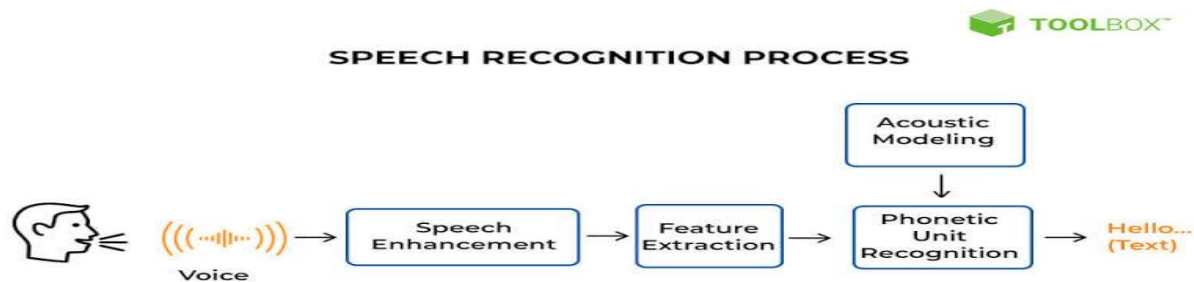
3. Constraints & Challenges:

- Using OpenAI's API may have usage limits and costs, impacting scalability.
- Data Privacy & Compliance: Handling audio files raises concerns about GDPR, HIPAA, or other data protection laws..

Phase-3: Project Design

Objective:

Develop the architecture and user flow of the application.



Key Points:

1. System Architecture:

- Direct client-API integration
- Input handling
- Preprocessing
- Post-performance

1. User Flow:

- Step 1: App launch
- Step 2: File selection
- Step 3: Transcription process
- Step 4: post-Transcription Actions
- Step 5: Error Handling







2. UI/UX Considerations:

- **Minimalist, user-friendly interface**
- **Clear visual hierarchy**
- **Dark & light mode** for better user experience.

Phase-4: Project Planning (Agile Methodologies)




Objective:

Break down development tasks for efficient completion.



Sprint	Task	Priority	Duration	Deadline	Assigned To	Dependencies	Expected Outcome
Sprint 1	Environment Setup & API Integration	 High	6 hours (Day 1)	End of Day 1	L.Rishanka	openAi whisper API key,python,Flask	API connection established & working
Sprint 1	Frontend UI Development	 Medium	2 hours (Day 1)	End of Day 1	M.Pooja	API response format finalized	Basic UI with Upload and record options
Sprint 2	Audio processing and Transcription	 High	3 hours (Day 2)	Mid-Day 2	T.Richa & M.Ankhitha	API response,Audio processing modules ready.	Transcription functionality Implemented.
Sprint 2	Error Handling & Debugging	 High	1.5 hours (Day 2)	Mid-Day 2	L.Rishanka & T.Richa	API logs	Improved API stability
Sprint 3	Testing & UI Enhancements	 Medium	1.5 hours (Day 2)	Mid-Day 2	M.Ankhitha&M.pooja	API response, UI layout completed	Responsive UI, better user experience
Sprint 3	Final Presentation & Deployment	 Low	1 hour (Day 2)	End of Day 2	Entire Team	Working prototype	Demo-ready project

Sprint Planning with Priorities



Sprint 1 – Setup & Integration (Day 1)

- ( **High Priority**) Set up the **environment** & install dependencies.
- ( **High Priority**) Integrate **OpenAI Whisper API**
- ( **Medium Priority**) Build a **basic UI** with input fields.

Sprint 2 – Core Features & Debugging (Day 2)

- ( **High Priority**) Implement **audio transcription functionalities**.
- ( **High Priority**) Debug API issues & handle **errors** in transcription processing.

Sprint 3 – Testing, Enhancements & Submission (Day 2)

- ( **Medium Priority**) Test API responses, refine UI, & fix UI bugs.
- ( **Low Priority**) Final **demo preparation & deployment**.

Phase-5: Project Development

Objective:

Implement core features of the AutoSage App.

Key Points:

1. Technology Stack Used:

- **Frontend:** React.js/Flutter
- **Backend:** FastAPI
- **Programming Language:** Python

2. Development Process:

- Implement **API key authentication** and **Whisper API integration**.
- Develop **real-time and batch transcription logic**.
- Optimize **speaker identification and timestamping**.



3. Challenges & Fixes:

- **Challenge:** Delayed transcription processing for large files.
- **Fix:** Enable catching and partial transcription to minimize API calls
- **Challenge:** Background noise affecting accuracy.
- **Fix:** Integrate noise reduction preprocessing before sending to whisper.

Phase-6: Functional & Performance Testing

Objective:

Ensure that the **Audio transcription app using OpenAi Whisper** works as Expected.

Test Case ID	Category	Test Scenario	Expected Outcome	Status	Tester
TC-001	Functional Testing	Transcribe clear English audio input	Accurate transcription should be displayed.	 Passed	pooja
TC-002	Functional Testing	Transcribe noisy background speech	Transcription should still be accurate.	 Passed	ankhitha

TC-003	Performance Testing	Transcription multi speaker conversation.	Different speakers should be distinguished correctly.	⚠ Needs Optimization	Tester 3
TC-004	Bug Fixes & Improvements	Transcribe non-English language..	Correct transcription in the respective language.	✅ Fixed	Developer
TC-005	Final Validation	Transcribe 60-minute long audio file.	Should complete within expected time without errors.	passed	Richa
TC-006	Deployment Testing	Transcribe mumbling.	should provide best effort transcription with markers	🚀 Deployed	DevOps

Final Submission

1. Project Report Based on the templates
2. Demo Video (3-5 Minutes)
3. GitHub/Code Repository Link
4. Presentation