Hackathon Project Phases Template for the **Audio2Art** project.

Hackathon Project Phases Template

Project Title:

Audio2Art - Transforming Voice Prompts into Visual Creations

Team Name:

Debug Thugs

Team Members:

- E.Usha
- K.Lakshmi
- Mamatha Kumari

Phase-1: Brainstorming & Ideation

Objective:

Develop an AI-driven system that converts voice prompts into images using Generative AI.

Key Points:

1. **Problem Statement:**
    - Many users struggle to visualize ideas quickly.
    - Artists and designers need a fast way to generate concept art.

2. **Proposed Solution:**

    ◯ A system that captures voice input, transcribes it into text, refines the text using AI, and generates images.

3. **Target Users:**

    ◯Education & E-Learning.
    ◯Gaming & Entertainment
    ◯Children's Creativity & Learning.

4. **Expected Outcome:**

    ◯ A functional AI-powered platform for converting voice descriptions into visuals

---

Phase-2: Requirement Analysis

Objective:

Define the technical and functional requirements for the Audio2Art.

Key Points:

1. **Technical Requirements:**
    ◯ Programming Language: **Python**
    ◯ Speech-to-Text: **OpenAI Whisper**
    ◯ Text Processing**:** Text Refinement
    ◯ Image Generation**: DALL-E or Stable Diffusion**
    ◯ Frontend**: Streamlit**
2. **Functional Requirements:**

    ◯ Convert speech to text accurately
    ◯ Process text and generate refined image prompts
    ◯ Display AI-generated images

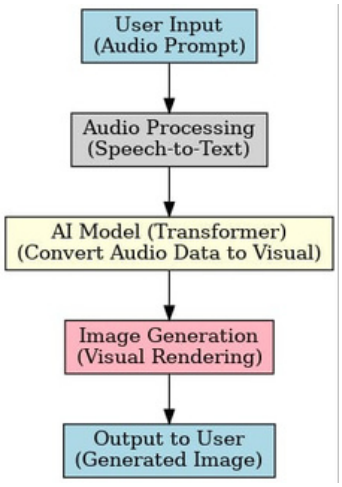
3. **Constraints & Challenges:**

    ◯ Speech Recognition Accuracy
    ◯ Processing Speed
    ◯ Noise Interference
    ◯ Subjectivity in Art Interpretation
    ◯ Consistency Issues.

Phase-3: Project Design

Objective:

. Develop the architecture and user flow of the application.



Key Points:

1. **System Architecture:**
   - ◯ User speaks into the microphone.
   - ◯ Whisper transcribes speech to text.
   - ◯ DALL-E or Stable Diffusion generates an image

2. **User Flow:**
   - ◯ Step 1: User drops a audio file into the web app
   - ◯ Step 2:The audio is converted into text and the text is converted into image
   - ◯ Step 3: It generates a AI image according to the according to the audio

.

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	Member 1	Google API Key, Python, Streamlit setup	API connection established & working
Sprint 1	Frontend UI Development	🟡 Medium	2 hours (Day 1)	End of Day 1	Member 2	API response format finalized	Basic UI with input fields
Sprint 2	Speech to text conversion	🔴 High	3 hours (Day 2)	Mid-Day 2	Member 1& 2	Whisper model , audio input	convert voice into accurate test
Sprint 2	text processing, image generate	🔴 High	1.5 hours (Day 2)	Mid-Day 2	Member 1&4	Whisper output text , GPT-4 text ,	generate image using DALL.E
Sprint 3	Error handling,Testing & UI Enhancements	🟡 Medium	1.5 hours (Day 2)	Mid-Day 2	Member 2& 3	API response, UI layout completed	improved API stability,better Ux resonsive UI
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 for Audio2Art

Sprint 1 – Setup & Integration (Day 1)

(🔴 High Priority) Set up the development environment and install dependencies.

(🔴 High Priority) Integrate OpenAI Whisper, GPT-4, and DALL·E/Stable Diffusion APIs.

(🟡 Medium Priority) Build a basic UI with input fields for voice input.

Sprint 2 – Core Features & Debugging (Day 2)

(🔴 High Priority) Implement voice-to-text conversion using Whisper.

(🔴 High Priority) Process transcribed text with GPT-4 to generate refined image prompts.

(🔴 High Priority) Generate images using DALL·E or Stable Diffusion.

(🔴 High Priority) Debug API issues and handle errors in queries.

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

(🟡 Medium Priority) Test API responses, refine UI, and fix UI bugs.

(🟢 Low Priority) Final demo preparation and project deployment.

Phase-5: Project Development

Objective:

Implement core features of the Audio2Art.

Key Points:

1. **Technology Stack Used:**
   - ◯ **Frontend:** Streamlit
   - ◯ **Backend:** OpenAI Whisper,GPT-4 and DALL.E
   - ◯ **Programming Language:** Python

2. **Development Process:**
   - ◯ Implement **API key authentication** and integrate OpenAI APIs
   - ◯ Develop voice to text processing using Whisper. refine textual prompts using GPT-4 for better generation.Generate image using DALL.E
   - ◯ Optimize response times for smooth real-time image generation.
3. **Challenges & Fixes:**

   - ◯ **Challenge:** Delayed API response times.
     Fix: Implement **caching** to store frequently queried results.
   - ◯ **Challenge:** Accuracy issuesin text to image conversion
     Fix: Enchance prompt structing using GPT-4 for better detail

---

Phase-6: Functional & Performance Testing

- ◯ Test audio transcription accuracy.
- ◯ Optimize text processing for relevant image prompts.
- ◯ Validate image quality and accuracy.

Final Submission

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