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

Hackathon Project Phases Template

Project Title: Thoughts2Arts

Thoughts2Arts using Diffusers & StableDiffusionPipeline

Team Name: *Team Eternal*

Team Members:

- Mohammed Yahya (TL)
- Siddu Yadagiri
- Mohammed Aslam

Phase-1: Brainstorming

Objective:

Develop an AI-powered model(Web) that listens to your audio and generates the art based on Designers, Developers, Artists,etc.. Voice prompt

Key Points:

1. **Problem Statement:**
    - **Designers, Developers, Artists,** etc are struggling to convert their **thoughts** to **arts**, some of them **forget** their thoughts or ideas with in mins/hrs/days…
    - Also if they remember, it **makes difficult** for them to bring it into **real world** art piece.

2. **Proposed Solution:**

    ○ An **AI-Powered Web** app using **Diffusers & StableDiffusionPipeline** is introduced that helps Designers, Developers, Artists, etc to bring their **thoughts/ideas** into real world with in two **couple of minutes**

    ○ This Web app is **simple** and **easy** to use to the users with good **User Interface**

3. **Target Users:**

    ○ Designers, Developers, Artists
    ○ Editors
    ○ General Public

4. **Expected Outcome:**

    ○ A **Web App** that records your **voice prompts** as **audio** and **shows** the prompts in t**ext format**, if **accepted** by user it **generates image** according to prompt

Phase-2: Requirement Analysis

Objective:

Defining the requirements for this AI-powered Model

Key Points:

1. **Technical Requirements:**

    ○ **Python** – For backend processing (handling AI models, audio processing, image generation)
    ○ **HTML, CSS, JavaScript** (via Gradio) – Used in the frontend for the web interface.
    ○ **Gradio** – A Python library that simplifies building interactive web UIs for machine learning models.

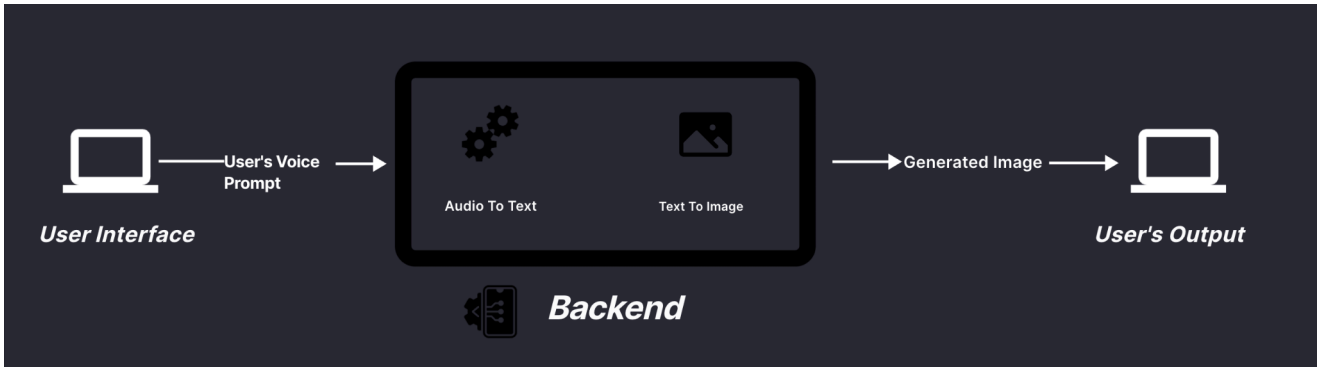
2. **Constraints & Challenges:**

    ○ **Technical** – High performance systems that includes **Nvidia cuda**.
    ○ **Implementation** – AI-friendly hosting.

Phase-3: Project Design

Objective:

Develop the architecture and user flow of the application.



Key Points:

1. **System Architecture:**

    ○ User gives audio File.
    ○ Audio is converted to text and later into art/image .
    ○ Google recognizes the audio file.
    ○ Frontend Displayes users prompt along with image/art generated.

2. **User Flow:**

    ○ Step 1: User **uploads/records** a voice prompt (eg. Hundreds of students participating GenAI Hackathon)
    ○ Step 2: **Backend** converts **audio prompt** to **text prompt** and **generates image** according to prompt given.
    Step 3: **Generated Images** is displayed on **user's monitor**.

3. **UI/UX Considerations:**

- **Minimalist, user-friendly interface** for seamless navigation.
- **Downloading** of generated images made easy
- **Dark & light mode** for better user experience.

Phase-4: Project Planning

Objective:

Break down development tasks for efficient completion.

Task	Priority	Duration	Deadline	Assigned To	Dependencies	Expected Outcome
Environment Setup & API Integration	High	2-3 hours (Day 1)	End of Day 1	Member 1 & TL	Google Translator,Whisper, API's, Python, GitHub,Google Speech Recognition, StableDiffusionPipeline	API connection established & working
Frontend UI Development	Moderate-High	4-5 hours (Day 1)	End of Day 2	Member 2	HTML, CSS, JS	Best UI with input fields
Audio - Text -Image Generation (Programming)	High	5-8 hours (Day 1)	End of Day 2	TL & Member 1	Google Translator API, Whisper, Stable Diffusion API's	Recognizes audio files and converts into images
Error Handling & Debugging	High	1-2 hours (Day 1)	End of Day 2	TL & Member 1	API logs, UI inputs	Improved API stability & proper input values
Testing & UI Enhancements	Moderate	1-2 hours (Day 2)	End of Day 2	TL & Member 1	API response, UI layout completed	Responsive UI, better user experience
Final Presentation & Deployment	High	1 hour (Day 2)	End of Day 2	Entire Team	Working prototype	Demo-ready project

Phase-5: Project Development

Objective:

Implement core features of the **Thoughts2Arts** WebApp.

Key Points:

1. **Technology Stack Used:**

    ○ **Frontend:** HTML, CSS, JS
    ○ **Backend:** Python and Frameworks
    ○ **Programming Language:** Python

2. **Development Process:**

    ○ **Environment setup** and **Implementation** of **API's.**
    ○ **AI-Powered** Model to convert **Voice Prompt** to **Text based Prompt** and **Generate Image**.
    ○ Implementation and **Testing** of the **Model**.

Phase-6: Functional & Performance Testing

Objective:

Ensure that the AutoSage App works as expected.

Test Case ID	Category	Test Scenario	Expected Outcome	Status	Tester
TC-001	Functional Testing	Audio File (Group of Students participating GenAI Hackathon)	Group of Students participating GenAI Hackathon Image Generated	✅ Passed	Team Leader
TC-002	Performance Testing	Record or Upload Voice Prompt	Image Generated According to Voice Prompt	✅ Passed	Team Leader
TC-003	Final Validation	Audio File Uploaded or Voice Prompt Recorded	UI accepts input audio file or asks to record the voice prompt, Image Generated according to prompt	✅ Passed	Team Member 3
TC-003	Deployment Testing	Host the app using Streamlit Sharing	App should be accessible online.	Deployed	Team Member 2

Final Submission

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