Hackathon Project Phases Template for the **AutoSage App** project.

Hackathon Project Phases Template

Project Title: AUDIO2ART (Transforming voice prompts into visual creations using transformers)

AutoSage App Using Gemini Flash

Team Name:

(ECE CHAMPS)

Team Members:

- A.GANGASAGAR
- B.ANUSHA
- G.MANASA
- G.SAHITHI
- M.BHARSHARANI

Phase-1: Brainstorming & Ideation

Objective:

- ❖ Audio2Art aims to generate unique, aesthetically appealing artwork that captures the essence of the audio input.
- ❖ Audio2Art is to transform auditory experiences into visual representations using advanced AI techniques.

Key Points:

1. **Problem Statement:**

Audio2Art addresses this challenge by leveraging AI to analyze audio characteristics—such as tone, rhythm, and emotion—and generate visually compelling artwork that encapsulates the essence of the sound.

2. **Proposed Solution:**

The solution utilizes deep learning models, including generative adversarial networks (GANs) and transformer-based architectures, to interpret audio characteristics and translate them into dynamic visual compositions.

3. **Target Users:**

a) **Digital Artists & Graphic Designe**rs – Creators looking to explore AI-generated visuals based on sound inputs for unique artistic expressions.

b) **Filmmakers & Content Creators** – Professionals who need dynamic audio-reactive visuals for movies, social media, advertising, or interactive media..

4. **Expected Outcome:**

- **Seamless Audio-to-Visual Conversion** – Users can effortlessly generate visually compelling artwork from any audio input, creating a unique fusion of sound and imagery.
- **Real-Time & Customizable Visuals** – The system will offer both real-time and pre-rendered visualizations with customizable artistic styles, catering to different creative needs.
- **New Opportunities in Art & Technology** – Bridges the gap between AI, music, and visual art, fostering innovation in creative industries.

Phase-2: Requirement Analysis

Objective:

To convert audio into dynamic visual art using AI, enhancing creativity, immersion, and accessibility.

Key Points:

1. **Technical Requirements:**

   - **Programming Language**: **Python**
   - **MODEL_NAME** = "microsoft/git-large-textcaps"
   - **CLIP_MODEL_NAME = "openai/clip-vit-large-patch14"**
   - **AUDIO_FILE = "audio_prompt.wav" #**
   - **Your audio file DEVICE = "cuda"**
   - **IMAGE _SIZE=224**

2. **Functional Requirements:**

   Audio input and processing –

   Goggle ai studio

   ○ Provide **real-time vehicle maintenance tips** based on seasons.

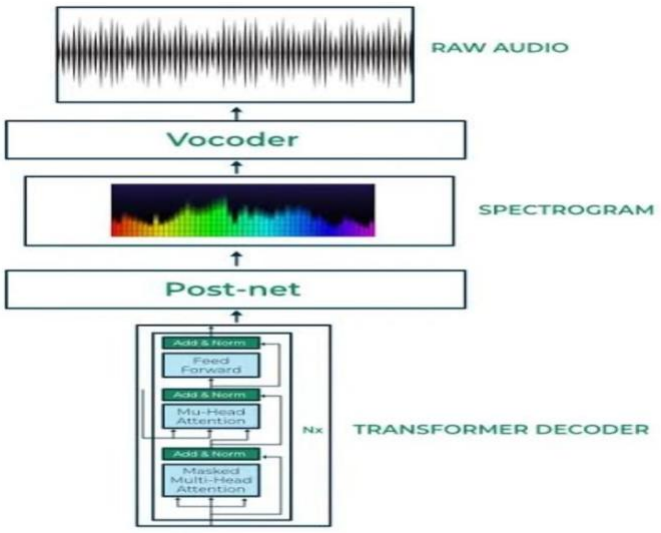
   ○ Allow users to **search eco-friendly vehicles** based on emissions and incentives.

3. **Constraints & Challenges:**

   Audio visual maping complexity

   Style adaptation challenges

Phase-3: Project Design :



Key Points:

1. **modular Architecture**
2. **User interface**
3. **Audio Processing**
4. **AI Based image generation**
5. **Output and storage module**

Phase-4: Project Planning (Agile Methodologies)

objective: Develop a system that transforms voice/audio into visual art using transformers.

Tech Stack: Python, TensorFlow/PyTorch, Google Colab, GenAI models (e.g., Stable Diffusion, DALL·E).

Project Definition

- **Project Title**: "Audio to Art: Exploring Sound Through Visuals"

- **Objective**: To transform audio data into visually captivating art forms, creating a unique representation of sound.
- **Scope**:
    - Convert various types of audio (e.g., music tracks, spoken words, nature sounds) into corresponding visual art.
    - The final product could be in forms like digital images, animations, physical installations, or interactive visualizations.

2. Concept Development

- **Artistic Vision**:
    - Decide what kind of visual style or medium you want to use (e.g., abstract art, generative visuals, minimalist, psychedelic, etc.).
    - Determine how you want the audio to influence the visual elements (color changes, movement, form, texture, etc.).
- **Medium**:
    - Digital visuals (e.g., using tools like Processing, TouchDesigner, or Adobe After Effects)
    - Physical pieces (e.g., sculptures, light installations)
    - Interactive installations (e.g., sound-reactive screens or physical art)

3. Audio Selection & Processing

- **Types of Audio**:
    - Music tracks, sound effects, or environmental sounds.
    - Spoken word or recorded voice.
- **Audio Processing Tools**:
    - Audio analysis software (e.g., Pure Data, Max/MSP, or Python libraries like librosa) for extracting features like frequency, amplitude, and tempo.
    - Use of MIDI data for music-driven visuals.

Audio Breakdown:

- Isolate different audio components like rhythm, melody, and harmony for varied interpretations.

4. Tools & Technology

- **Art Creation Tools**:
    - For visual design: Processing, P5.js, TouchDesigner, Blender (for 3D art), or Adobe After Effects.
    - For generative art: OpenFrameworks, Max/MSP, or Sonic Pi.

    - For physical installations: Arduino, Raspberry Pi (for interactive or kinetic pieces).
- **Sound-to-Visual Software/Algorithms**:
    - Sound-to-visual conversion algorithms (e.g., FFT analysis for frequency and pitch mapping to colors or shapes).
    - Real-time sound-reactive visuals software.

5. Team Roles and Responsibilities

- **Project Manager**: Oversee the project timeline, resources, and coordination.
- **Audio Specialist**: Curate and process the audio files for analysis.
- **Visual Artist**: Responsible for creating the visual representation of the audio.
- **Tech Developer**: Implement the technology and software solutions for sound-to-visual conversion.
- **Producer/Designer**: To ensure the concept is followed and the final product matches the desired outcome.

6. Design and Development Phases

1. **Research and Inspiration**:
    - Research existing works in audio-to-art transformation, such as works by generative artists like Ryoji Ikeda or programs like Audiosurf or Sound Sculpture.
    - Gather inspiration from art styles that could represent audio visually (abstract art, fractals, etc.).
2. **Prototyping**:
    - Start by experimenting with basic visualizations (e.g., basic sound waveforms, frequency spectra, or particle systems).
    - Try different audio features like pitch, rhythm, and loudness to see how they influence visual movement.
3. **Refining Concept**:
    - After prototyping, decide on a final direction for the project.
    - Create a cohesive visual narrative that translates across all types of audio.
4. **Asset Creation**:
    - Develop final visual assets, such as textures, backgrounds, animation loops, or physical sculptures.
    - If it's an interactive piece, ensure that sound triggers specific visual elements.

7. Timeline & Milestones

- **Phase 1: Research & Ideation (2 Weeks)**:

    - Concept development and research of sound-to-visual tools.
- **Phase 2: Audio Processing & Prototype (4 Weeks)**:
    - Select audio samples.
    - Build initial sound-to-visual prototypes.
- **Phase 3: Art Development & Refinement (6 Weeks)**:
    - Develop the main art pieces based on audio analysis.
    - Adjust visuals based on feedback and experimentation.
- **Phase 4: Testing & Feedback (2 Weeks)**:
    - Test the final work with a small group for feedback.
    - Adjust based on user interaction or feedback.
- **Phase 5: Finalization & Presentation (2 Weeks)**:
    - Prepare for the exhibition or final presentation.

8. Budget Planning

- **Software/Tools**: Subscription fees for software like Adobe Suite, Processing libraries, or Touch Designer licenses.
- **Hardware**: For physical installations, costs may include sensors, projectors, screens, and computers.
- **Team Compensation**: Payment for team members depending on their roles (artists, developers, etc.).
- **Venue/Exhibition Costs**: If presenting in an exhibition or gallery, budget for the space, setup, and promotion.

9. Evaluation & Feedback

- After presenting the artwork, gather feedback to evaluate the impact of the visual art and the effectiveness of the audio-to-visual transformation.
- Use feedback to refine the work for future iterations or projects.

10. Promotion & Sharing

- Promote the final project on social media, art forums, and through press releases or blogs.
- Create a behind-the-scenes video or an article explaining the creative process to engage the audience further.

11. Possible Expansion

- After completing the initial project, consider how to expand the concept, such as:
    - Incorporating more interactive elements, where viewers can modify the visuals with their own sound inputs.

- Expanding to include more diverse audio types (e.g., live recordings or environmental sounds).
- Turning the project into a series of exhibitions or even an online interactive experience.

Sprint Planning with Priorities

Sprint 1 – Setup & Integration (Day 1)

(🔴 High Priority) Set up the **environment** & install dependencies.
(🔴 High Priority) Integrate **Google Ai studio** code in google colab
(🟡 Medium Priority) Build a Audio to image using hugging face

Sprint 2 – Core Features & Debugging (Day 2)

(🔴 High Priority) Implement **search & comparison functionalities**. **(🔴 High Priority)** Debug API issues & handle **errors in queries**. **Sprint 3 – Testing, Enhancements & Submission (Day 2)**

(🟡 Medium Priority) Test API responses and generate image using audio to image.

Phase-5: Project Development

Objective:

Implement core features of the AutoSage App.

Key Points:

1. **Technology Stack Used:**
    - **Frontend:** Streamlit
    - **Backend:** Google Gemini Flash API
    - **Programming Language:** Python
2. **Development Process:**
    - Implement **API key authentication** and **Gemini API integration**.
    - Develop **vehicle comparison and maintenance tips logic**. ○ Optimize **search queries for performance and relevance**.
3. **Challenges & Fixes:**
    - **Challenge:** Delayed API response times.
    - **Challenge:** Limited API calls per minute.
      Fix: Optimize queries to fetch **only necessary data**.

Phase-6: Functional & Performance Testing

Input handling

Transformation process

Performance testing: speed and latency

scalability

Final Submission

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