Shiksha Setu: Multilingual AI-Powered Offline Education Platform

Team ID :

1. Project Overview

- **Platform Goal:** To create a multilingual, AI-powered, offline-first platform connecting teachers and students in rural India, ensuring accessible and high-quality education regardless of connectivity.
- **Target Users:** Students (learner-centric features, offline study, gamification) and Teachers (content upload, live classes, AI lesson/PPT generation, training/evaluation).
- **Core Value Proposition (USPs):**
    - **22-Language Support:** Full content and UI support for all 22 official Indian languages.
    - **True Offline Functionality:** Lessons, quizzes, PPTs, summaries, and TTS audio are fully available and functional offline.
    - **AI-Generated Lessons:** Ability for students to request and receive entire, comprehensive lessons from the AI, not just summaries of teacher content.
    - **AI Interview Evaluation:** A specialized module for students to practice job/academic interviews, receiving detailed scorecards and improvement plans from the AI.
    - **Smart Sync & P2P Sharing:** Includes Smart Delta Sync for efficiency and Offline P2P sharing (Bluetooth/Hotspot) for lesson distribution.

2. Project Timeline (Estimated)

Phase	Duration	Focus Area	Key Deliverables
Phase 1: Foundation & MVP	Week 1	Core Platform Setup & Basic Functionality	* Frontend (PWA/React Native) setup with 4-zone UI structure. * FastAPI Backend & Firestore integration. * Core Teacher Module (Content Upload) and Student Module (Basic Lesson Viewing). * Authentication and initial ASR (Whisper Tiny) integration for transcripts.
Phase 2: AI/ML & Multilingual Engine	Week 2-3	Full AI Pipeline Integration & Offline Hardening	* Full AI Pipeline deployment (IndicTrans2, GPT-4/Gemini for generation). * Full 22-Language translation for all content artifacts (quizzes, summaries). * Implementation of the **Offline-First System** using SQLite. * **Smart Delta Sync** and resumable upload/download functionality.
Phase 3: Advanced Features & Polish	Week 4	Live Classes, Evaluation, Gamification & Security	* Live Class Module (with auto-record and instant lesson generation). * AI Interview Evaluation and Teacher Training Module. * Gamification implementation (XP, Leaderboards, Learning Games). * Security hardening (TLS, AES encryption) and final performance optimization (lightweight goal).

3. AI/ML Pipeline (Detailed)

This pipeline runs automatically in the background (Worker Engine) whenever a teacher uploads a lesson, creating all necessary multilingual artifacts.

Step	Model / Technology	Process / Function	Output Artifacts
1. Transcription (ASR)	Whisper Tiny/Small (Client Preview) / Whisper Medium (Server)	Converts uploaded video/audio content into text.	Timestamped Transcript.
2. Summaries & Keypoints	GPT-4 / Gemini (High Quality)	Processes the transcript to extract core concepts and structure the lesson.	5-Line Summary, Detailed Keypoints, Lesson Plan.
3. Translation	IndicTrans2	Translates all generated text artifacts (transcript, summary, quiz text) into 22 official Indian languages.	22-language versions of all text assets.
4. Quiz & Model Papers	GPT-4 / Gemini Nano	Generates structured questions based on the lesson content.	JSON format for MCQs, True/False, Short Answers, and Full Model Question Papers.
5. PPT Generation	Custom AI Service (Gemini/GPT-4 orchestrated)	Structures the lesson into a professional slide deck format.	PPTX file containing Slide Titles, Bullet Points, and Speaker Notes.
6. Interview Evaluation	ASR + NLP Scoring + Audio/Camera Analysis	Assesses student/teacher mock interviews based on technical and soft skills.	Scorecard (Fluency, Clarity, Confidence, Answer Quality) + Improvement Plan.

4. Challenges & Solutions

Area	Challenge	Proposed Solution
Offline Reliability	Ensuring reliable data access, student progress tracking, and quiz functionality in a truly offline state.	**Solution:** Use **SQLite** for robust on-device storage of all lessons (transcripts, summaries, quizzes, PPT JSON).
Data Synchronization	Efficiently syncing large volumes of multilingual data (22 versions) in low-bandwidth rural environments.	**Solution:** Implement **Smart Delta Sync** to only update files that have changed. Utilize Resumable Uploads/Downloads for large files.
Multilingual Quality	Maintaining high quality and cultural accuracy for AI-generated content (quizzes, interview feedback) across 22 Indic languages.	**Solution:** Utilize **IndicTrans2** for specialized Indic language translation and rely on high-quality LLMs (GPT-4/Gemini) for the initial generation step.
Performance	Achieving the project goal of an **extremely lightweight & RAM-efficient** application, especially on older mobile devices.	**Solution:** Optimize the React Native bundle, use compressed JSON for PPTs/Quizzes, and utilize smaller models like **Whisper Tiny** for on-device operations when possible.

5. Implementation Plan / Execution Strategy

- **Frontend Strategy (Offline-First Design):**
    - Build the application using **PWA** (for browser compatibility) and **React Native** (for high-performance mobile app distribution).
    - Implement the mandatory **4-Zone Interface** for structured learning: Top Bar (Search/Language), Left Panel (Syllabus), Center Panel (Viewer), Right Panel (Tools).
- **Backend and Data Strategy:**
    - Utilize **FastAPI (Python)** for its performance and compatibility with Python-based AI/ML libraries.
    - Deploy a dedicated **Worker Engine** (asynchronous processing) to handle the long-running AI tasks (ASR, translation, generation) without blocking the main API endpoints.
    - Use **Firestore** for real-time cloud data (user profiles, analytics, metadata) and **SQLite** for local, offline content storage.
- **Security & Compliance:**
    - Enforce **TLS** for all data communication.

- Ensure all data stored locally on the device via SQLite is secured with **AES Encryption**.