# MINDKRAFT EXPO PROJECT [TEAM A]

## ➢ Problem Statement

Students with temporary physical injuries such as fractures or post-surgery conditions face difficulty in writing exams manually. Dependence on human scribes affects privacy and independence, while existing exam systems lack flexible support for such students. Hence, a secure and voice-based examination system is required to ensure equal academic opportunities.

## ➢ Proposed Solution

A **Speech-to-Text Examination Assistant** enables completely hands-free exam taking. Students use **face recognition for secure login**, followed by **IVR-style voice navigation** (numbered menu options) to dictate answers. The system automatically reads questions aloud using **eSpeak TTS**, processes speech via **Vosk/Whisper STT**, formats answers intelligently using **Llama LLM**, and operates in **fullscreen kiosk mode** to prevent cheating.

## ➢ Features of the System

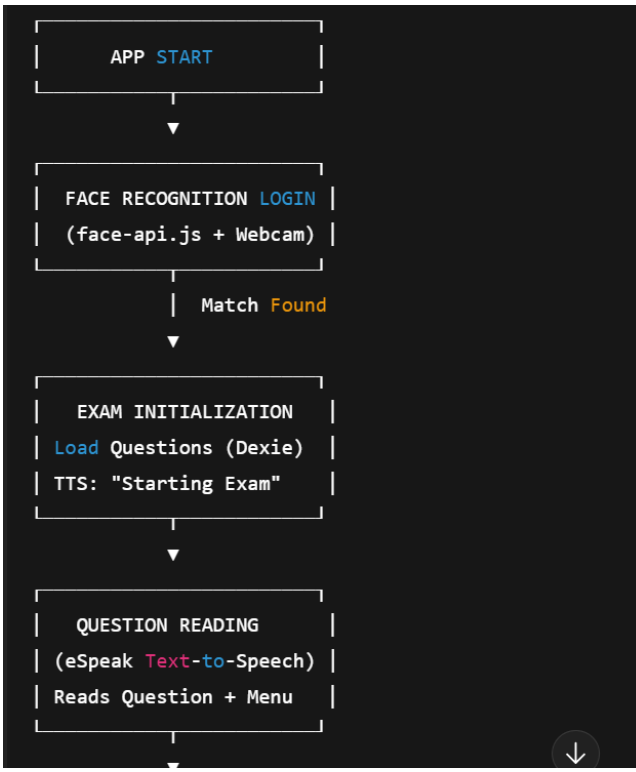
1. **Face Recognition Login** - Secure biometric authentication using webcam
2. **IVR-Style Voice Navigation** - Numbered menu system (1=Answer, 2=Read again, 3=Next, etc.)
3. **Automatic Question Reading** - eSpeak TTS reads entire question paper sequentially
4. **Real-time Speech-to-Text** - Vosk for commands, Whisper/Vosk for answers
5. **AI Text Formatting** - Llama3.2 LLM adds punctuation, capitalization, grammar correction
6. **Hands-free Operation** - Zero keyboard/mouse required from login to submission
7. **Fullscreen Kiosk Mode** - Locks other applications, prevents Alt+Tab
8. **Auto-save with Timestamps** - Prevents data loss, audit trail for authorities
9. **Offline Operation** - All components work without internet
10. **Multi-language Support** - Vosk models + multilingual LLM
11. **Activity Logging** - Complete audit trail with timestamps

## ➢ TECH STACK

| Category | Technology / Tool | Purpose / Usage in System |
|---|---|---|
| **Frontend UI** | **React 18 + Vite** | Builds fast, responsive user interface for exam screens |
| **Styling** | **Tailwind CSS** | Provides high-contrast, accessible design for disabled users |
| **State Management** | **Zustand** | Manages exam state (current question, answers, menus) |
| **Face Recognition** | **face-api.js (MediaPipe)** | Performs biometric login using webcam face matching |
| **Voice Commands** | **react-speech-recognition** | Captures spoken number commands (1,2,3,0) |
| **Offline Storage** | **Dexie.js (IndexedDB)** | Stores questions, answers, and logs locally |
| **Desktop Platform** | **Electron** | Runs app in fullscreen kiosk mode |
| **Speech Recognition (Commands)** | **Vosk-node** | Converts spoken numbers into commands (high accuracy) |
| **Speech Recognition (Answers)** | **Whisper.cpp / Vosk** | Converts spoken answers into text |

| Category | Technology / Tool | Purpose / Usage in System |
|---|---|---|
| **AI Processing** | **Ollama + Llama3.2:3B** | Formats raw speech text into proper exam answers |
| **Text-to-Speech (TTS)** | **eSpeak CLI** | Reads questions and menus aloud |
| **Database** | **SQLite3** | Stores student, exam, and answer records |
| **Audio Recording** | **MediaRecorder API** | Captures microphone input for STT processing |
| **Inter-Process Communication** | **Electron IPC** | Enables communication between UI and AI backend |
| **Security Mode** | **Electron Kiosk Mode** | Locks system to prevent cheating |
| **Activity Logging** | **Custom Logging System** | Records timestamps and user actions |
| **Document Generation** | **PDF Generator** | Creates audit logs and submission reports |

## ➢ Workflow Diagram




**1.App Start**

Student opens the exam app.

**2.Face Login**

Webcam scans face → verifies student → allows entry.

**3.Exam Initialization**

Questions load → system announces start.

**4.Question Reading**

System reads:

- Question
- IVR menu options

**5. IVR (Interactive Voice Recognition) Menu**

Student hears:

“Press 1 to answer, 2 to repeat…”

**6. Voice Command**

Student speaks number → Vosk detects.

**7. Two Possible Paths**

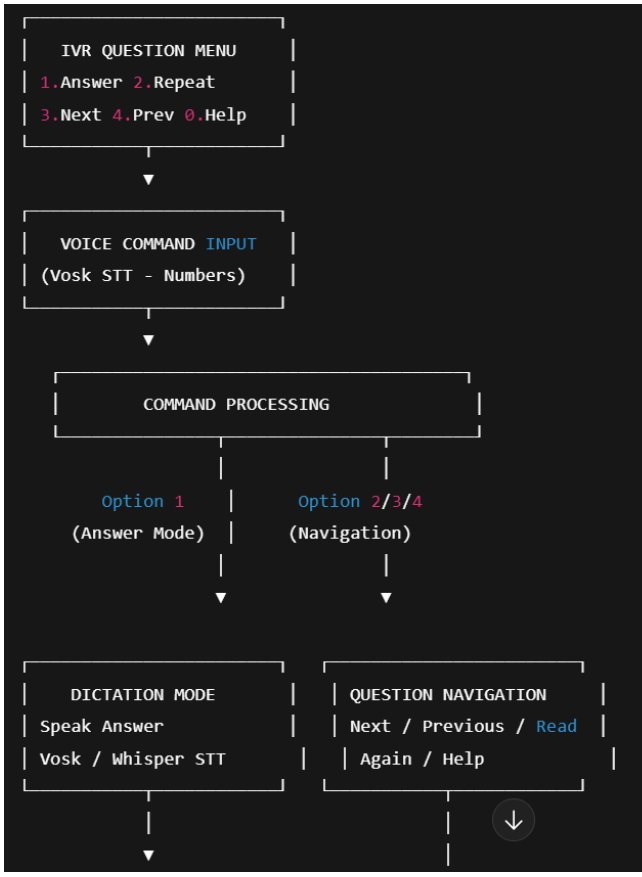
▶ If Navigation

- Next
- Previous
- Repeat

→ Go back to question reading.

▶ If Answer

→ Dictation mode starts.




**8. Dictation Mode**

Student speaks answer → STT converts.

**9. AI Formatting**

Llama improves grammar & structure.

**10. Auto Save**

Answer saved with time + ID.

**11.Loop**

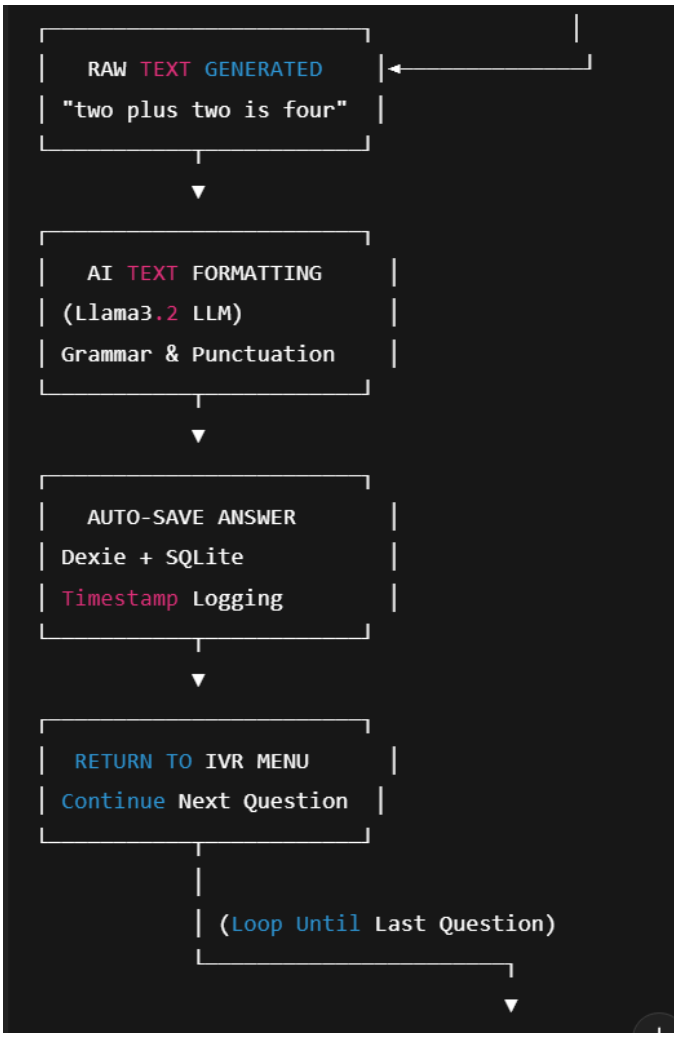
System goes back to menu for next question.

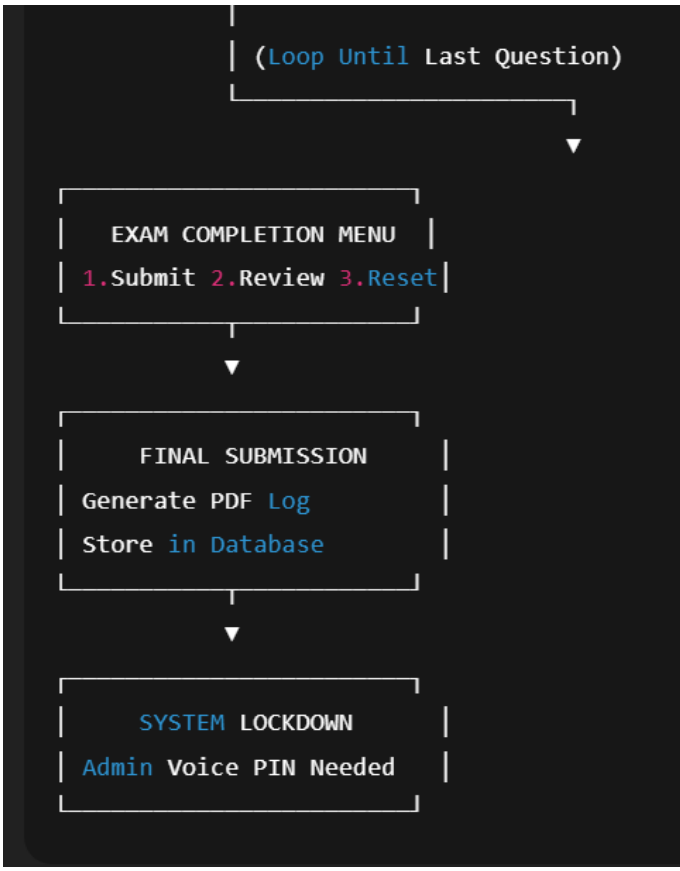
**12. Submission**

At end → student submits.

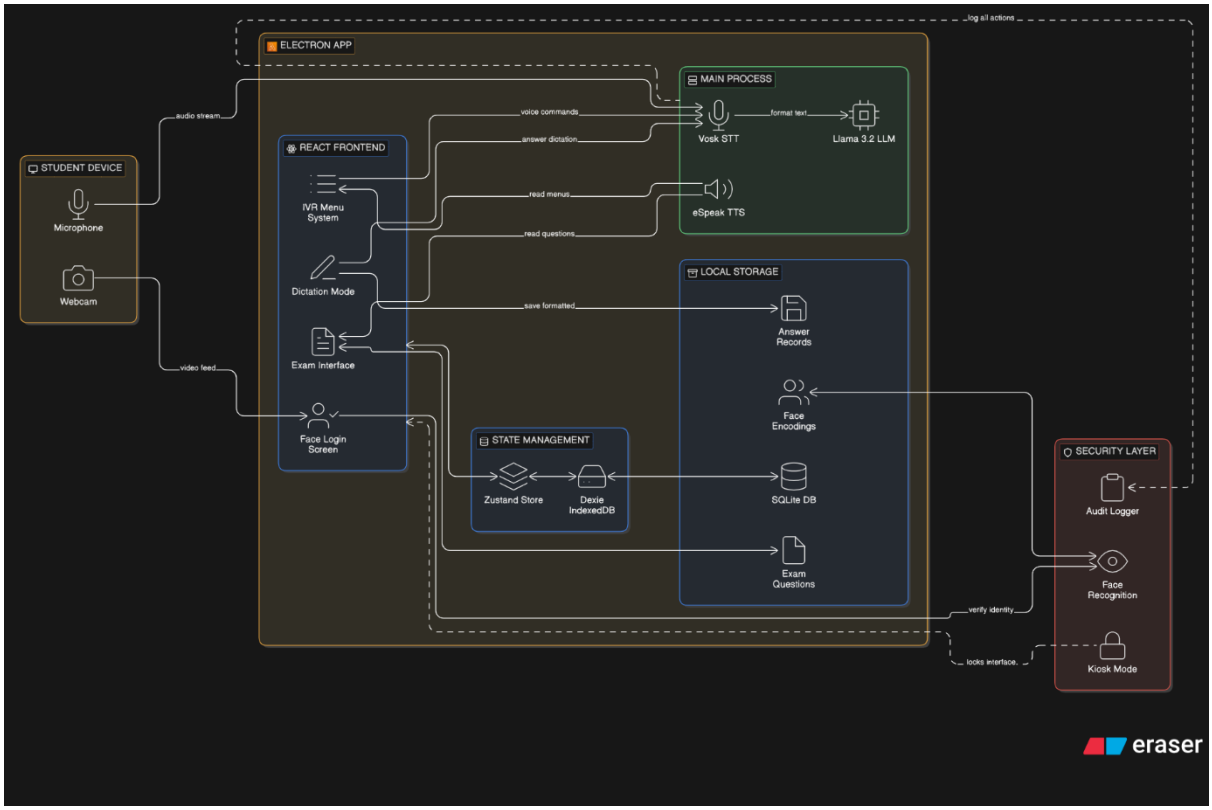
**13. Lock System**

App locks → only admin can exit.



## ➢ System Architecture




## ➢ Sample IVR Menu Structure

text

QUESTION MENU:

1. Answer question

2. Read question again

3. Next question

4. Previous question

0. Help

DICTATION MENU:

1. New line

2. Delete last word

3. Done (format & save)

0. Cancel

SUBMIT MENU:

1. Submit exam

2. Review all answers

3. Restart current exam

0. Main menu