TextEvolve

Turn Ink into Intelligence

Team Name	:	Dynamic Dreamers
NMNT ID	:	NMNTSTD7116020
College Name	:	Kathir College of Engineering

Theme: Artificial Intelligence

Problem Statement:

How might we develop an AI or OCR solution to digitize and convert handwritten, old registered documents into a readable and accessible format in regional languages, improving public access and readability of historical records? The solution should allow the output to be downloaded in various formats such as PDF and Word, enabling wider distribution and accessibility.

Existing System:

- Existing manual transcription methods are slow, costly, and error-prone.

- Standard OCR models often struggle with accuracy for handwritten text, regional languages, and historical scripts.

- Physical document degradation and limited accessibility are significant challenges.

Proposed System:

- **Dual OCR Approach**:

  Custom ML model for script-specific recognition

  Google Vision API, Amazon Textract and Azure AI Vision for improved accuracy

- **Smart Language Tools**:

  Google Translate API for seamless language conversion

  Gemini API for grammar & spelling correction

- **Key Features**:

  Real-time results & progress tracking

  Fast, scalable, and user-friendly
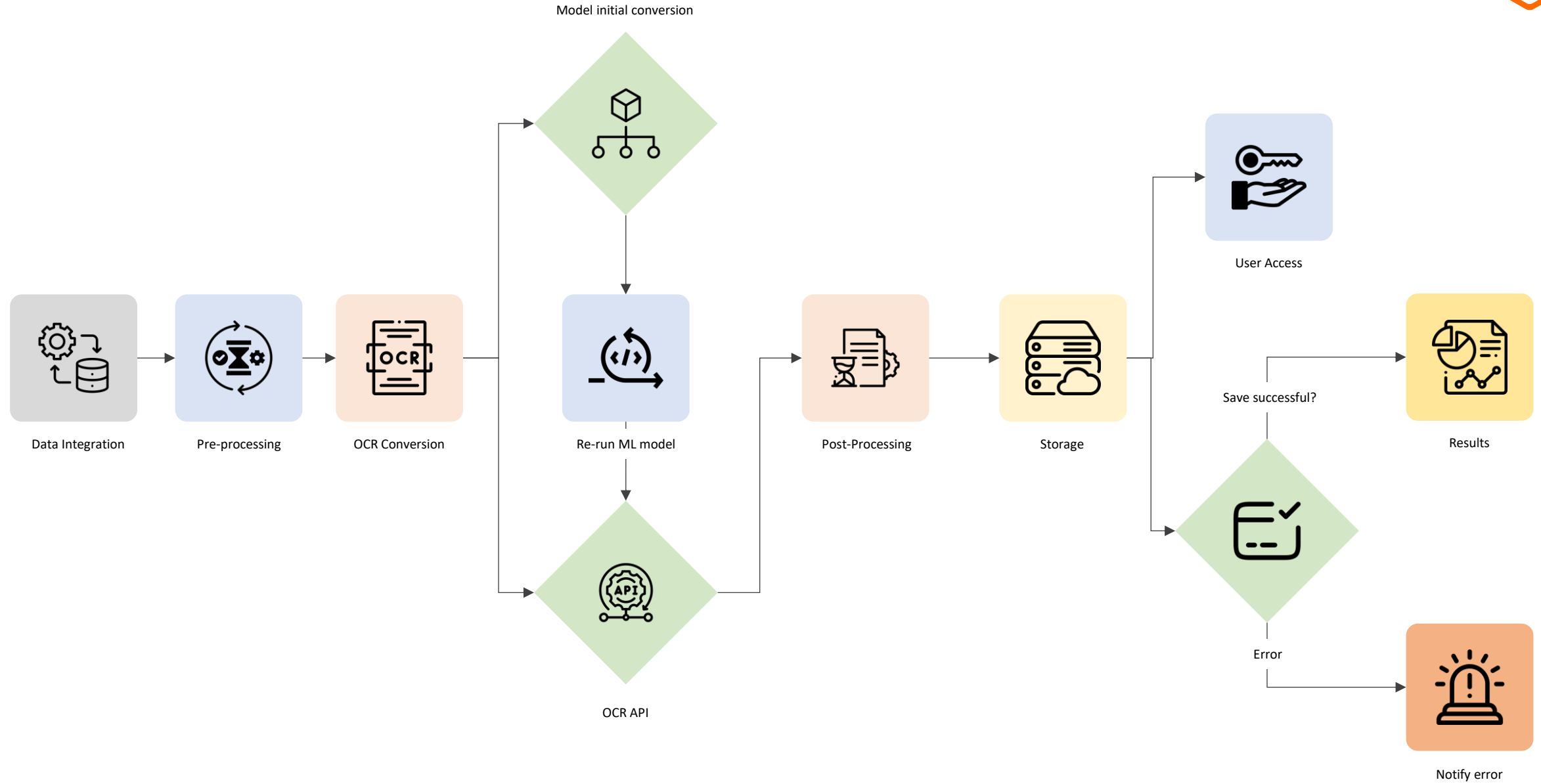
Programming Languages
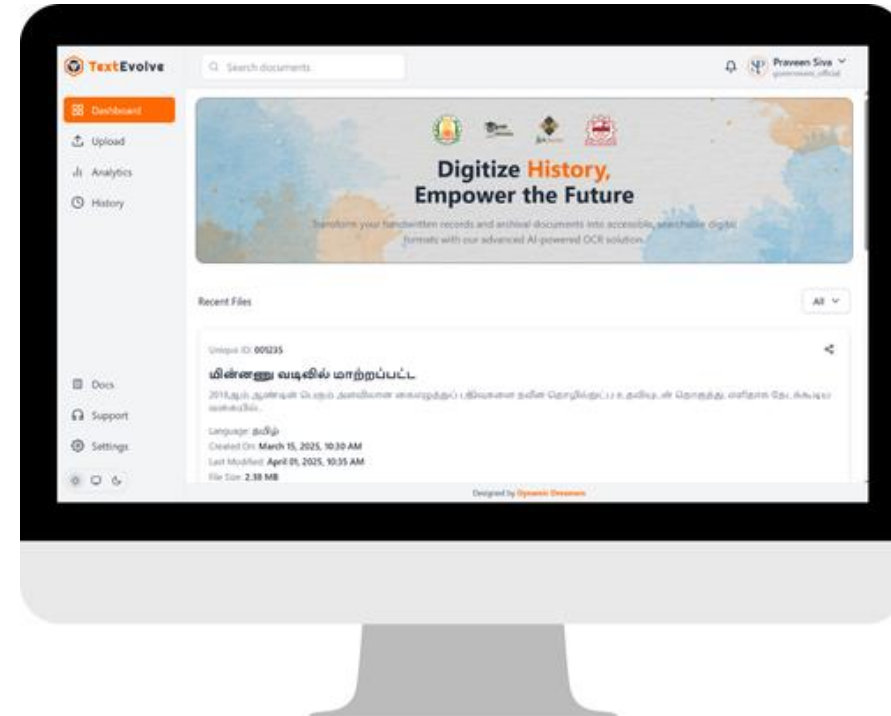
Database & Deployment

Frameworks & Libraries

Development Tools

System Architecture & Workflow

Model initial conversion

Data Integration

Pre-processing

OCR Conversion

Re-run ML model

OCR API

Post-Processing

Storage

User Access

Save successful?

Results

Error

Notify error

www.textevolve.in

Degradation:

Physical documents deteriorate over time.

Inaccessibility:

Limited access to rare or regional texts.

Handwriting Variability:

Handwritten texts present difficulties in standard OCR processing.

Language Barriers:

Many OCR solutions are not optimized for regional languages.

Cultural Preservation: Safeguard historical documents for future generations.

Enhanced Accessibility: Allow researchers, educators, and the public to easily access digitized content.

Improved Research: Facilitate the study of regional languages and historical texts.

Scalability: A robust system that can grow to handle vast archives.

TextEVolve is an AI-driven solution that digitizes handwritten historical

documents with high accuracy using a **custom ML model** and **Google OCR API**.

By combining innovation with technology, it preserves cultural heritage, enhances

accessibility, and supports research — ensuring valuable historical knowledge is

safeguarded for future generations.

? Ask Questions

Mrs. Kavitha M

Faculty Guide

Yuva Nandhini M

AI/ML Developer

Praveenkumar S

AI/MERN Developer

Sibi Siddharth S

AI/MERN Developer

Uma Maheswari P

AI/ML Developer

Preserving the Past,
Empowering the Future