Accessible PDF Viewer with AI-Powered Content Analysis

Overview

This project focuses on developing an **accessible PDF viewer** that can be embedded into web pages, designed specifically for blind users or those with visual impairments. The viewer uses AI-powered offline tools to analyze and convert complex PDF content—such as tables, formulas, and figures—into structured data. This structured data is then processed to provide clear explanations, making the content more comprehensible and navigable for users relying on assistive technologies like screen readers.

Core Features

1. PDF Viewer Integration

- **Embedded Viewer**:
  - A web-embedded PDF viewer that integrates seamlessly into any webpage.
  - Basic controls for navigation, zooming, and page management.
- **Responsive Design**:
  - Optimized for different devices and screen sizes to ensure usability on desktops, tablets, and mobile devices.

2. Offline AI-Based Data Extraction

- **AI-Powered Parsing**:
  - Analyze PDFs offline using AI packages like MinerU or llma-parse.
  - Extract and structure key content types:
    - **Text**: Process main content and headings.
    - **Tables**: Convert tables into machine-readable formats, such as JSON or plain text.
    - **Figures and Images**: Generate descriptive labels and detailed explanations for visual content.
    - **Mathematical Formulas**: Translate formulas into accessible textual descriptions or MathML.

- **Efficient Processing**:
  - Optimize the AI models for fast and accurate offline performance.

3. Interactive Content Explanation

- **Clear Explanations**:
  - Convert structured data into accessible, easy-to-understand descriptions.
  - Examples:
    - Explain a table’s content row by row in plain language.
    - Describe the purpose of a chart or figure and summarize its key findings.
    - Translate mathematical formulas into step-by-step textual explanations.
- **Multi-Language Support**:
  - Provide explanations in multiple languages to cater to diverse audiences.

4. Accessibility Features

- **Screen Reader Compatibility**:
  - Fully compatible with screen readers like JAWS, NVDA, and VoiceOver.
  - Dynamically generated tags and descriptions are accessible for text-to-speech output.
- **Text-to-Speech Support**:
  - Integrate text-to-speech functionality for real-time reading of extracted and explained content.
- **Structured Data Presentation**:
  - Present extracted content in a structured, logical format for easy navigation:
    - Tables rendered as lists or grids.
    - Figures with alt-text and detailed captions.
    - Formulas explained step-by-step.

Key Objectives

1. **Enhance Accessibility**:
   - Make complex PDF content more accessible to blind users by converting it into structured, understandable formats.
2. **Empower Independent Understanding**:

    - Provide clear, AI-driven explanations of tables, figures, and formulas to foster comprehension.
3. **Ensure Privacy and Performance**:
    - Perform all data extraction and processing offline to protect user privacy and enable faster results.

Benefits

- **Improved Accessibility**: Transform inaccessible PDFs into structured content that assistive technologies can process.
- **Contextual Understanding**: Enable blind users to comprehend complex data and visuals in a meaningful way.
- **Seamless Integration**: Embed the viewer into any webpage for universal access.

Deliverables

1. **PDF Viewer**:
    - An embeddable, responsive PDF viewer with basic navigation controls.
2. **Offline AI Processing**:
    - Integrated AI tools for analyzing and structuring PDF content.
3. **Accessibility Integration**:
    - Features compatible with screen readers and text-to-speech technologies.