Proposal for PowerPoint Accessibility Enhancement Project

**Project PI- Ibne Farabi Shihab**

# Goal:

The primary objective of this project is to automate the accessibility enhancement of PowerPoint presentations using AI-driven tools, including Vision-Language Models (VLMs) and Large Language Models (LLMs). This project will focus on:

* Generating Alt Text Automatically for images.
* Ensuring Readable and Accessible Fonts & Contrast.
* Simplifying Complex Text to improve accessibility.
* Evaluating and Scoring Accessibility based on canvas (Ally).

# Scope:

This project will leverage AI-based accessibility tools to analyze and improve PowerPoint slides, ensuring compliance with accessibility standards such as WCAG 2.1. We will use Vision Transformers (ViTs), CLIP models, and LLMs to develop a framework that:

1. Processes PowerPoint slides and extracts text, images, and formatting information.
2. Generates alternative text (Alt-Text) for images using AI models.
3. Checks and adjusts font size & contrast based on WCAG 2.1 standards.
4. After all the steps are done, a update on PowerPoints and save that.
5. Generates a comparative evaluation report of accessibility improvements based on canvas Ally.

The final goal is to produce an **AI-powered website** that automates this process and compares **accessibility scores** before and after enhancements. One alternative of this is using ChatGPT which also can be explored.

The following tasks will be undertaken for the project.

**Task 0:** **Collect at least 50 PowerPoint files for training and testing, ensuring a balanced dataset with diverse accessibility issues. The dataset will be split into 2/3 for training and 1/3 for testing. Extract text, images, and metadata such as font sizes and contrast, annotating key accessibility challenges such as missing alt text, poor contrast, and small font sizes. Preprocess and structure the extracted data for AI model input.**

**Task 1**: Implement automated text and image extraction using off-the-shelf AI models. Store extracted content in a structured format and validate the accuracy of extracted text, images, and metadata. Develop initial scripts to evaluate font sizes, contrast levels, and missing alt text in the PowerPoint dataset, ensuring all extracted metadata is compatible with AI models for further analysis.

**Task 2:** Develop a pipeline integrating Vision-Language Models (Villa, VideoLLM-Online, GPT-4) to generate alternative text for images, detect contrast and font size issues, and classify accessibility compliance based on WCAG 2.1 standards. Implement an accessibility scoring mechanism aligned with Canvas Ally and optimize model performance by balancing accuracy and computational efficiency.

**Task 3**: Develop a user-friendly web interface using Streamlit, allowing users to upload PowerPoint files and receive accessibility evaluations. Integrate the backend pipeline to process uploaded files and display before-and-after accessibility scores. Implement visual indicators highlighting the detected accessibility issues and improvements while ensuring processed PowerPoints are saved with enhanced accessibility features.

**Task 4:** Conduct model evaluation and failure mode analysis, refining AI models based on feedback from Touchpoints 1 and 2. Perform a detailed comparison of model accuracy, processing speed, and computational cost, ensuring trade-offs are well understood. Validate accessibility score improvements using test datasets and iterating on necessary refinements.

**Task 5**: Finalize the model and pipeline, delivering a comprehensive report detailing improvements, challenges, and recommendations. Provide a final presentation summarizing the findings, comparing different AI models used, and demonstrating the effectiveness of the automated accessibility enhancement system. Ensure deployment readiness and explore potential future enhancements such as ChatGPT-based processing for broader applications.

# Timeline:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Feb 06** | **Feb 13** | **Feb 20** | **Feb 27** | **Mar 06** | **Mar 13** | **Mar 27** | **Apr 03** | **Apr 10** | **Apr 17** | **Apr 24** | **May 01** |
| *Task 0* Data Collection & Preprocessing |  |  |  |  |  |  |  |  |  |  |  |  |
| *Task 1-* Model Implementation |  |  |  |  |  |  |  |  |  |  |  |  |
| *Task 2-*  Web dev and Fine-Tuning |  |  |  |  |  |  |  |  |  |  |  |  |
| *Task 3-* Accessibility Scoring & Evaluation |  |  |  |  |  |  |  |  |  |  |  |  |
| Task 4: Evaluation and validation |  |  |  |  |  |  |  |  |  |  |  |  |
| Task 5: Final Report & Presentation |  |  |  |  |  |  |  |  |  |  |  |  |
| **Touchpoints FAST-TrAC and Pelmorex** | **0** |  |  |  | **1** |  |  |  | **2** |  |  | **3** |

**Touchpoint 0-**Initial model running to figure out if everything works roughly according to the WCAG 2.1

**Touchpoint 1-** A presentation will be delivered halfway through Task 1 to review preliminary findings and facilitate a discussion on any arising questions.

**Touchpoint 2-**The outcomes of Task 4 will be presented to the class, and feedback on the methodologies employed will be solicited. This feedback will be instrumental in refining the models further.

**Touchpoint 3** Final presentation of results, comparison tables, and AI model effectiveness summary. Upon completing this phase, discussions may commence on possible next steps.

**Total Cost of the Project: The only cost will be API cost in case we plan to use the Chat GPT API**