**Proposal for Assessing FCV Sensitivity in Climate-Related Projects Using LLMs**

**Project Title:**

Leveraging AI to Classify FCV Sensitivity in Climate Co-Benefit Projects

**Objective:**

To design and implement a protocol for classifying and assessing the extent of Fragility, Conflict, and Violence (FCV) sensitivity in climate co-benefit tagged activities in Project Appraisal Documents (PADs) from the World Bank Projects database. This project aims to leverage large language models (LLMs) to analyze PAD content and determine FCV-sensitive characteristics, incorporating a scoring mechanism to evaluate confidence and accuracy.

**Background:**

The intersection of climate resilience and FCV considerations is critical in ensuring sustainable development. Many climate-related projects involve indirect references to FCV-sensitive components, making it challenging to assess their true impact. This project seeks to address this gap by automating the analysis of PADs using LLMs, guided by a structured classification protocol.

**Scope of Work:**

**1. Data Collection and Preprocessing**

* **Scraping World Bank Projects Database**:
  + Develop a web scraper to extract PADs and associated metadata from the [World Bank Projects database](https://projects.worldbank.org/en/projects-operations/projects-home).
  + Store extracted text and metadata in a structured database.
* **Content Extraction**:
  + Use tools such as PyPDF2, Apache Tika, or pdfminer.six to extract and clean text from PADs.
  + Preprocess text to remove noise and irrelevant information.

**2. Classification Protocol Development**

* **Framework Design**:
  + Develop a classification protocol based on a Climate and FCV-Sensitivity Framework.
  + Include criteria to assess FCV-sensitive coverage using yes/no and multiple-choice questions.
* **Scoring Mechanism**:
  + Assign weights to questions based on importance.
  + Incorporate confidence scores from LLM responses to evaluate reliability.
* **Benchmarking**:
  + Annotate a sample of PADs manually to serve as a benchmark for evaluation.

**3. LLM Implementation**

* **Model Selection and Fine-Tuning**:
  + Use a pre-trained LLM (e.g., OpenAI GPT) fine-tuned with climate and FCV-related documents.
* **Prompt Engineering**:
  + Develop prompts for LLMs to process PAD content and respond to protocol questions.
  + Example: "Does this project include conflict-sensitive approaches? Provide Yes/No and confidence level."
* **Integration**:
  + Design a pipeline to extract text from PADs, classify FCV-sensitive content, and calculate confidence scores.

**4. Scoring and Classification**

* Develop a methodology to:
  + Summarize and classify PAD activities based on FCV-sensitive characteristics.
  + Use embedding techniques to identify indirect references to FCV issues.

**5. Evaluation**

* **Performance Metrics**:
  + Use precision, recall, F1-score, and accuracy to evaluate model output.
  + Compare LLM classifications with annotated ground truth.
* **Human Validation**:
  + Conduct periodic reviews by domain experts.

**6. Deliverables**

1. Web scraper for PAD extraction.
2. Structured database of PAD content.
3. Classification protocol with a scoring mechanism.
4. Integrated LLM pipeline for analyzing PADs.
5. Evaluation report with performance metrics and recommendations.

**Methodology:**

1. **Data Collection**:
   * Scrape PADs from the World Bank Projects database and preprocess the content.
2. **Protocol Development**:
   * Collaborate with domain experts to finalize the classification framework.
3. **LLM Implementation**:
   * Fine-tune an LLM to answer protocol questions and classify FCV-sensitive content.
4. **Scoring and Evaluation**:
   * Implement a scoring mechanism to assess classification confidence.
   * Validate model outputs against a benchmark dataset.
5. **Iteration and Feedback**:
   * Incorporate feedback from domain experts to refine the protocol and model.

**Challenges and Mitigation Strategies:**

* **Indirect References in PADs**:
  + Mitigate by using advanced LLMs with embeddings to understand nuanced text.
* **Model Confidence**:
  + Introduce a confidence threshold to flag uncertain classifications for manual review.
* **Data Availability**:
  + Address data limitations by focusing on publicly available PADs and collaborating with relevant stakeholders for additional resources.

**Expected Outcomes:**

1. A scalable protocol to classify and assess FCV sensitivity in climate-related projects.
2. An AI-powered tool to automate PAD analysis, reducing manual effort.
3. Insights into the integration of FCV-sensitive considerations in climate co-benefit activities.
4. A validated scoring mechanism to evaluate LLM outputs.