G. A Human-in-the-Loop Environment for Arabic Machine Translation[2]

Introduction and Objective

Translation is a cornerstone of cross-cultural communication, knowledge transfer, and access to global resources. While fully automated machine translation has advanced significantly, it often struggles with the nuances, cultural contexts, and domain-specific terminology of languages such as Arabic. Human translators remain essential in ensuring accuracy, readability, and trustworthiness.

This project aims to design and implement an **LLM-assisted translation environment** that supports human translators in producing high-quality English–Arabic and French–Arabic translations. By combining **large language models (LLMs)** with a carefully curated translation interface, the project seeks to enhance productivity, improve translation quality, and streamline workflows. The solution will integrate cutting-edge translation technologies with a human-centered interface, optimised for usability, safety, and cultural relevance.

Project Description

The system will be built on three main pillars: **(1) data collection and curation**, **(2) evaluation and fine-tuning of LLMs**, and **(3) a translator-oriented environment**.

1. **Data Collection and Curation:** The project will construct parallel corpora of English–Arabic and French–Arabic texts. These will include curated bilingual documents, translations contributed by professional translators, and a domain-specific database of in-context standard translations. (The last two will be provided to you if possible; otherwise, you will b=develop them.) This corpus will serve as both training and evaluation material.

2. **LLM Selection and Evaluation:** A set of free/open-source LLMs will be evaluated for Arabic translation performance. From these, **three models** will be selected and fine-tuned to improve quality, focusing on accuracy, style, domain adaptation, and robustness.

3. **Human Translation Environment:** A user-friendly application will be developed to support professional translators. It will allow document ingestion (Word/PDF), segmentation into logical units (the largest possible chunks as chapters, sections, or paragraphs), simultaneous generation of translations from the three LLMs, and interactive editing features (copy-paste, replace, recombination). The environment will facilitate comparison and refinement, enabling the translator to choose a reference translation, edit it efficiently (using the available translations), and export the final version.

Core Components

1. **Data Collection and Curation**
   - Collect English–Arabic and French–Arabic corpora from public sources and professional translators.
   - Build a **domain-specific translation memory** to capture standard terms and in-context expressions. The translator should also be able to enrich the translation memory upon approval of some validated translation entries.
   - Apply preprocessing, annotation, and quality control to ensure clean, usable data.

2. **Model Development and Training**
   - Evaluate multiple free/open-source LLMs for translation performance.
   - Select **three top-performing models** and fine-tune them on curated corpora.
   - Apply **prompting strategies** and optimisation techniques to improve translation quality, coherence, and context retention.

[2] This project is done in collaboration with the Algerian Academy for the Arabic Language.

3. **Translation Environment**
    - **Input Handling:** Upload and segment Word/PDF documents into chapters, sections, or paragraphs.
    - **Parallel Translation Views:** Generate translations from the three fine-tuned LLMs, displayed side-by-side in the interface.
    - **Interactive Editing:** Allow translators to select a reference translation, copy-paste from alternatives, and apply edits directly.
    - **Export Functionality:** Generate the final translation as a polished Word/PDF output.
    - **Optional Enhancements:** Translation memory suggestions, terminology highlighting, and adaptive prompts for specific domains.

4. **Optimisation for Deployment**
    - Adapt the system for efficient use on **limited infrastructures** (lightweight deployment, model compression, reduced latency).
    - Support both **web-based deployment** (for collaborative work) and **desktop applications** (for offline/local translation).

5. **Safety and Trustworthiness**
    - Ensure that translations remain **culturally sensitive, safe, and domain-appropriate**.
    - Implement safeguards to filter harmful or biased outputs.

System Evaluation

- Evaluate translation quality using both **automatic metrics** (e.g., BLEU, etc.) and **human evaluation** by professional translators.
- Assess usability through translator feedback on interface efficiency, ease of editing, and workflow support.
- Compare baseline machine translation vs. the fine-tuned LLMs to measure quality improvements.

Expected Outcomes

- Curated English–Arabic and French–Arabic parallel corpora and translation memories.
- Fine-tuned, domain-adapted LLMs for high-quality Arabic translation.
- A **user-friendly environment** that seamlessly integrates machine translation with human editing.
- A formal **technical report/research paper** (10–15 pages) presenting the methodology, results, evaluation, and discussion.

Expected Impact

The project will contribute a **practical and innovative translation tool** for professional use, empowering human translators with advanced AI assistance. By integrating LLMs into the translation process while preserving human oversight, the system ensures accurate, trustworthy, and culturally relevant outputs. The environment will improve productivity, reduce repetitive translation tasks, and help institutions, publishers, and businesses overcome language barriers. Furthermore, this approach can serve as a **model for hybrid human–AI translation systems** in other language pairs and domains.