Name: Pratham Bhanushali

Nationality: Indian

Location: Mumbai, India

University: University of mumbai

Email: prathambhanushali71@gmail.com

Degree/Field of Study: B.Tech in Information Technology (3rd year)

Github: pratham2888

Project Title: "Extension of Apertium Web API for Cross-Language Information Retrieval"

Abstract: Apertium is an open-source rule-based machine translation platform for translating between various languages. It provides a Web API that can be used to translate text from one language to another. However, there is still room for improvement in the existing Web API. This project aims to extend the Apertium Web API by adding new functionalities to make it more useful for cross-language information retrieval.

Goals:

Addition of Language Identification Feature: The Web API will be extended to support language identification of input text. This feature will enable the API to automatically detect the input language and translate it to the desired language. This will make the API more user-friendly and improve its accuracy.

Integration with External Resources: The Apertium Web API will be integrated with external resources like wordnets, ontologies, and other language resources. This will enable the API to provide more accurate and context-aware translations.

Addition of Query Expansion Feature: The Web API will be extended to support query expansion. This feature will enable the API to expand the user's query with relevant synonyms and related terms in the target language, thus improving the relevance of the search results.

Integration with Existing Translation Memory: The Apertium Web API will be integrated with existing translation memory tools like OmegaT and other Translation Memory (TM) systems. This will enable the API to provide better translations by leveraging the translations already present in the TM.

User-Friendly Web Interface: A user-friendly web interface will be developed to demonstrate the newly added functionalities of the Apertium Web API. The interface will enable users to input text and view translations along with additional information like language identification and query expansion.

Deliverables:

- Language identification feature added to the Apertium Web API.
- Integration of external resources with the Apertium Web API.
- Query expansion feature added to the Apertium Web API.
- Integration of the Apertium Web API with existing translation memory tools.
- User-friendly web interface developed to demonstrate the new functionalities.

Timeline:

Week 1-2: Set up the development environment and become familiar with the Apertium

Web API.

Week 3-4: Add language identification feature to the Apertium Web API.

Week 5-6: Integrate external resources with the Apertium Web API.

Week 7-8: Add query expansion feature to the Apertium Web API.

Week 9-10: Integrate Apertium Web API with existing translation memory tools.

Week 11-12: Develop a user-friendly web interface to demonstrate the new functionalities.

Week 13-14: Test and debug the system.Document everything so far.Fix any remaining bugs.

Week 15-16: Documentation and final project submission.Buffer period Ensure that all work is completed by now for safety.