

Translate LaTeX documents with DeepL

Degree programme : BSc in Computer Science

Specialisation : Computer Perception and Virtual Reality

Thesis advisor : Prof. Claude Fuhrer

Expert : Dr. phil. nat. Igor Metz (Glue Software Engineering AG)

The translation of LaTeX documents is a complex process that involves the preservation of both content and structure. LaTeX poses a translation challenge due to the integration of text with control sequences for formatting, environments, equations and references. The objective of this project is to create solutions that guarantee seamless translation by utilizing the DeepL API while maintaining the technical structure of the document.

Goal

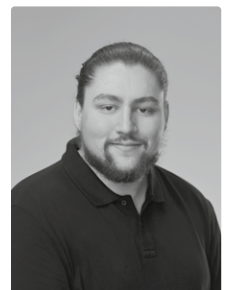
The aim of this project is to overcome the difficulties associated with translating LaTeX documents by creating a command-line interface and a graphical user interface. The purpose of these tools is to simplify the translation process while maintaining structural fidelity and content.

Design

The development of a Python library specifically designed to manage the complexities of LaTeX translation is the primary focus of the design. This library is the core of the project, executing wrapping, translation through the DeepL API and reconstruction of LaTeX documents. A CLI tool incorporates the library for direct use from the command line, and Flask is used to serve an API. The GUI is a modest Vue.js web application that seamlessly integrates with the API to provide a more interactive and accessible user experience. This modular approach ensures that the tools are adaptable and extensible.

Outcome

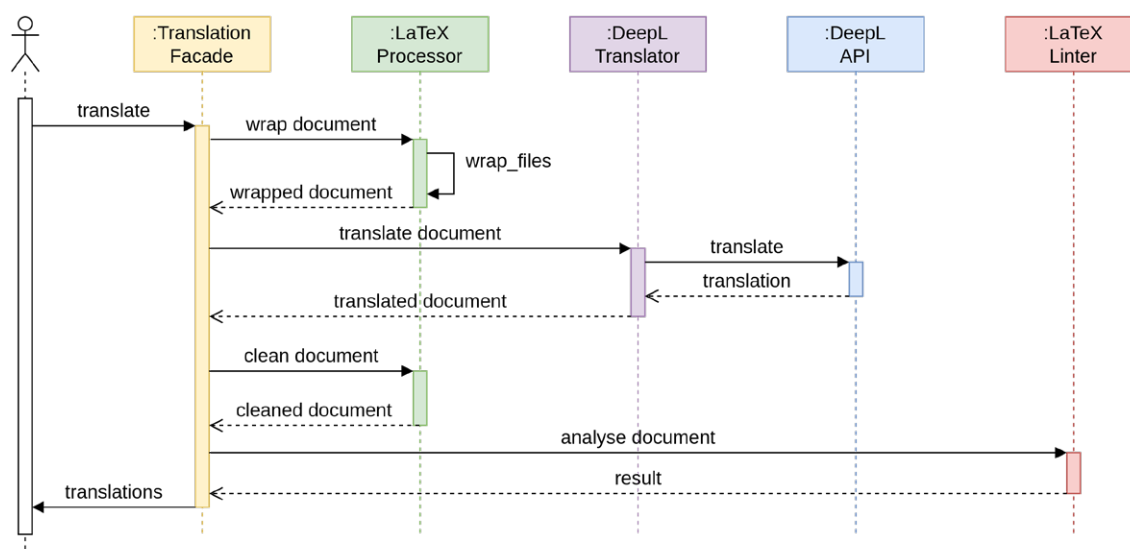
The project effectively developed the CLI and GUI, allowing users to translate LaTeX documents. They have the ability to exclude environments from translation, guaranteeing the preservation of specific content such as code or mathematics. In addition, the tools include a linter that identifies anomalies in the translated documents. The tools also facilitate partial translation by allowing users to select individual TeX files from documents. Although the library is able to translate most of the elements in a document, it has not yet achieved complete coverage. It mainly manages recurring and frequently used control sequences.



Karim Mathias Kummer

Future work

Future work will refine and extend the tools developed in this project. Improvements that could be made include adding comments directly to LaTeX files to help guide and change the translation process, combining the translation process with DeepL Write to get better results, and making it possible to handle more commands.



Sequence diagram of a document translation