

# Process Definition Document (PDD)

## LaTeX Document Translator Automation

---

### Document Information

<b>Document Title</b>	Process Definition Document - LaTeX Translator
<b>Project Name</b>	TranslatorEvaluator
<b>Version</b>	1.0.0
<b>Date</b>	January 14, 2026
<b>Author</b>	Leustean Robert George & Kurti Mark
<b>Status</b>	Final

---

### Table of Contents

1. Process Overview
  2. Business Context
  3. Scope
  4. Process Description
  5. Inputs and Outputs
  6. Business Rules
  7. Exception Handling
  8. Process Metrics
  9. Stakeholders
  10. Glossary
  11. Appendices
- 

## 1. Process Overview

### 1.1 Process Name

**LaTeX Document Translator** (TranslatorEvaluator)

### 1.2 Process Purpose

This automation process translates LaTeX (.tex) documents from a source language to a target language while preserving the complete LaTeX structure, formatting, mathematical formulas, and special environments. The process leverages the DeepL translation API for high-quality neural machine translation.

## 1.3 Process Type

- **Automation Type:** Attended Automation
- **Trigger Type:** User-Initiated (Manual)
- **Execution Mode:** Sequential Processing

## 1.4 Business Value

Benefit	Description
<b>Time Savings</b>	Reduces manual translation time from hours to minutes for academic and technical documents
<b>Quality Assurance</b>	Preserves LaTeX syntax integrity, preventing formatting errors post-translation
<b>Consistency</b>	Ensures uniform translation quality across documents using professional-grade APIs
<b>Cost Reduction</b>	Eliminates need for specialized LaTeX-aware translators
<b>Caching</b>	Avoids redundant API calls through intelligent translation caching

## 2. Business Context

### 2.1 Current State (As-Is Process)

Translating LaTeX documents manually involves: 1. Opening the LaTeX source file 2. Manually identifying translatable text segments (excluding commands, math, code) 3. Copying text to external translation tools 4. Carefully replacing translated text while preserving LaTeX formatting 5. Validating the document structure 6. Compiling to verify no syntax errors were introduced

**Pain Points:** - High risk of breaking LaTeX syntax during manual translation - Time-consuming process for large academic documents - Inconsistent handling of special characters and environments - No preservation of translation history for repetitive content

### 2.2 Future State (To-Be Process)

The automated process: 1. Reads the complete LaTeX file 2. Validates document structure 3. Intelligently extracts only translatable text segments 4. Translates via DeepL API with caching support 5. Reconstructs the document with translated content 6. Saves the translated file with appropriate naming

### 2.3 Process Frequency

- **Execution Frequency:** On-demand, user-initiated
- **Expected Volume:** Variable (1-50 documents per day)

- **Average Processing Time:** 1-10 minutes depending on document size
- 

### 3. Scope

#### 3.1 In Scope

Item	Description
LaTeX File Reading	Read and parse .tex files with UTF-8 encoding
Structure Validation	Validate document has proper LaTeX structure
Text Extraction	Extract translatable text while preserving structure
Translation	Translate text using DeepL API
Caching	Cache translations to optimize API usage
File Generation	Generate translated LaTeX file
Error Handling	Handle and report validation/translation errors
Logging	Comprehensive logging of all operations

#### 3.2 Out of Scope

Item	Reason
PDF Compilation	Requires LaTeX compiler installation
Image Translation	Images within documents are not translated
Multiple API Providers	Currently only DeepL is supported
Batch Processing	Single file processing only
BibTeX Translation	Bibliography files not processed
Custom Package Handling	Specialized package commands may need manual review

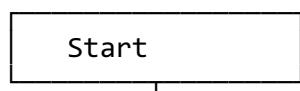
#### 3.3 Supported Languages

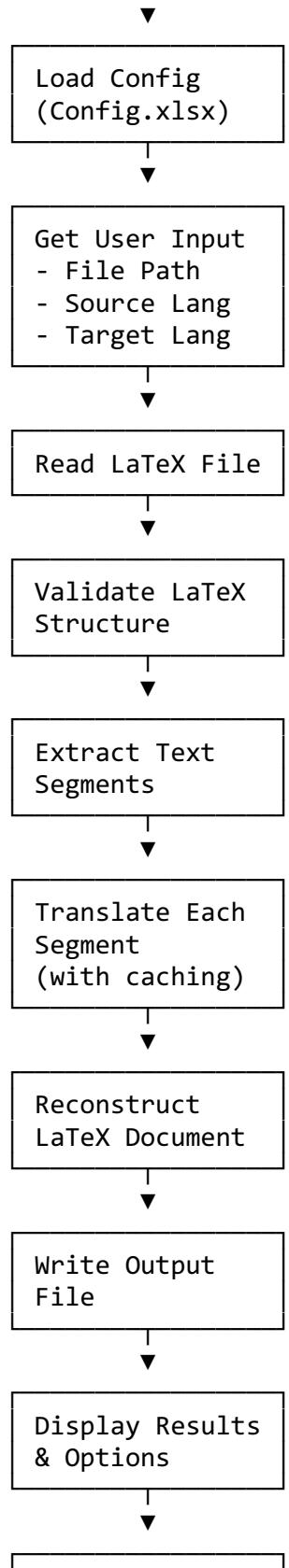
The automation supports all languages available through the DeepL API, including: - English (EN) - Romanian (RO) - German (DE) - French (FR) - Spanish (ES) - Italian (IT) - Portuguese (PT) - Dutch (NL) - Polish (PL) - Russian (RU) - Japanese (JA) - Chinese (ZH) - And others supported by DeepL

---

### 4. Process Description

#### 4.1 High-Level Process Flow





End

## 4.2 Detailed Process Steps

### *Step 1: Initialize Settings*

Attribute	Value
<b>Workflow</b>	Framework.xaml
<b>Description</b>	Loads configuration from Config.xlsx
<b>Input</b>	Config.xlsx file path
<b>Output</b>	Dictionary of configuration settings

Configuration settings include: - DeepL API Key - Default source/target languages - Log folder path - Cache settings

### *Step 2: Get User Input*

Attribute	Value
<b>Workflow</b>	PresentationLayer.xaml
<b>Description</b>	Prompts user for file path and languages
<b>User Actions</b>	Select .tex file, confirm source/target languages

User input validation: - File must exist and have .tex extension - Source and target languages must be different - Invalid inputs trigger re-prompt

### *Step 3: Read LaTeX File*

Attribute	Value
<b>Workflow</b>	DataLayer.xaml
<b>Description</b>	Reads the complete .tex file content
<b>Encoding</b>	UTF-8 (Code 65001)

### *Step 4: Validate LaTeX Structure*

Attribute	Value
<b>Workflow</b>	BusinessLayer.xaml
<b>Description</b>	Validates required LaTeX components

Validation checks: - ✓ \documentclass command present - ✓