Document Translation Test

This is a test document created to verify the translation functionality of the PyMuPDF-YOLO integrated document processing pipeline. The document contains various types of content that should be extracted and translated correctly.

Section 1: Technical Overview

The enhanced document processing system combines PyMuPDF for high-fidelity text extraction with YOLO for layout analysis. This integration provides intelligent content mapping and processing routing based on document characteristics. Pure text documents use fast PyMuPDF-only processing, while mixed content documents use coordinate-based extraction.

Section 2: Translation Features

The translation service supports multiple languages and includes advanced features such as context-aware translation, glossary management, and markdown-aware processing. The system preserves document structure and formatting while providing accurate translations.

Key Features:

• High-fidelity text extraction using PyMuPDF

• Layout analysis with YOLO neural networks

• Intelligent processing strategy routing

• Context-aware translation services

• Markdown structure preservation

• Multi-format output generation

Conclusion

This document demonstrates the capabilities of the integrated translation system. When processed, it should generate translated output in the target language while preserving the original structure and formatting. The system should create a Word document, processing report, and extract any non-text elements to a separate images folder.