

PDF to Markdown Test Document

This is a test PDF file for pdf2md.py

Created for testing the o2md PDF converter

Chapter 1: Introduction

This document demonstrates the PDF to Markdown conversion capabilities.

The converter extracts text from each page and renders pages as images.

It supports both PNG and SVG output formats.

Features:

- Text extraction from PDF pages
- Page rendering as PNG or SVG images
- OCR fallback using manga-ocr
- Integration with o2md.py dispatcher

Chapter 2: Usage

To convert a PDF file:

```
python pdf2md.py input.pdf -o output_dir
```

Or use the unified o2md.py:

```
python o2md.py input.pdf -o output_dir
```

Options:

```
--format png      Output images in PNG format
--format svg      Output images in SVG format (default)
-v, --verbose     Enable verbose output
```