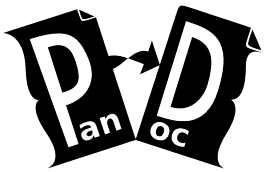
PDF Generation with Pandoc and GitHub Actions

How to generate good looking PDFs from Markdown files using Pandoc and how to automate the process with GitHub Actions

Nicolas Rocq



2024-03-27

Contents

1 Introduction

This repository is an exploration of how to generate a PDF from a Markdown file using Pandoc.

1.1 Requirements

- Pandoc[1]
- LaTeX[2]
- Latex packages : `sudo tlmgr install adjustbox footnotebackref pagecolor csquotes mdframed zref needspace sourcesanspro sourcecodepro titling selnolig lualatex-math` (for Mac users)

1.2 Contributors

- Nicolas Rocq[3]

[1] https://pandoc.org/
[2] https://www.latex-project.org/get/
[3] https://github.com/Nishogi

2 Usage

There are three levels of complexity:

2.1 1. A simple PDF file with no special formatting.

To generate the PDF, run the following command:

```
pandoc README.md -o README.pdf
```

2.2 2. Usage of a template to generate a PDF with a title page and better looking.

We are using the template eisvogel.tex from the pandoc-latex-template[4] repository.

To generate the PDF, run the following command:

```
pandoc -s -o document.pdf -f markdown_strict+backtick_code_blocks+
    pipe_tables+auto_identifiers+yaml_metadata_block+implicit_figures+
    table_captions+footnotes+smart+escaped_line_breaks+header_attributes --
    data-dir=pandoc --template eisvogel.tex --toc --listings --columns=50
    --number-sections --dpi=300 --pdf-engine xelatex -M date="$(date +%Y-%m
    -%d)" md/HEADER.YAML md/*.md
```

2.3 3. Automate the generation of the PDF using GitHub Actions.

See the **GitHub Actions workflow file** bellow for more details.

```
name: Build PDF Document

on:
  push:
    paths:
      - '.github/workflows/*'
      - 'docs/md/*'

jobs:
  build:
    runs-on: ubuntu-latest

    container:
      image: knsit/pandoc:latest
      options: --user root

    steps:
      - name: Checkout code
        uses: actions/checkout@v4

      - name: Build PDF
        run: |
          pandoc -s -o document.pdf -f markdown_strict+pipe_tables+
              backtick_code_blocks+auto_identifiers+yaml_metadata_block+
              smart+implicit_figures+table_captions+footnotes+smart+
              escaped_line_breaks+header_attributes --data-dir=pandoc --
```

[4] https://github.com/Wandmalfarbe/pandoc-latex-template

```
            template eisvogel.tex --toc --listings --columns=50 --number-
            sections --dpi=300 --pdf-engine xelatex -M date="$(date +%Y-%
            m-%d)" md/HEADER.YAML md/*.md

      - name: Upload artifact
        uses: actions/upload-artifact@v4
        with:
          name: PDF-Document
          path: document.pdf
```

And the artifacts generated by the workflow are available in the Actions tab[5] of this repository.

[5]https://github.com/Nishogi/pdf-from-markdown/actions/new

3 Sources

3.1 Markdown

- Markdown Syntax - https://www.markdownguide.org/basic-syntax/

3.2 Pandoc

CI with pandoc - https://gitlab.com/pandoc/pandoc-ci-example

3.3 Templating

Templating with pandoc - https://github.com/alexeygumirov/pandoc-for-pdf-how-to

3.4 Github Actions

Github Actions - https://docs.github.com/en/actions