

文档编号: COMPANY-DEPARTMENT-00000000

Markdown 书写 PDF 输出优雅的方案

0.1

Markdown 书写 PDF 输出优雅的方案

DOMINIC

COMPANYNAME

2022-08

修订记录

序号	修订人	修订时间	修订记录
V0.1	Dominic	2022-08-19T00:00:00.000Z	创建文档

目录

1 折腾背景	1
2 基础环境配置	1
2.1 Pandoc 的安装	1
2.2 LaTeX 的安装	1
3 模板配置	2
3.1 配置 Pandoc 模板	2
3.2 配置 LaTeX 模板	2
3.3 字体设置	2
4 如何生成 PDF	3
4.1 PDF 文件指定 metadata 信息	3
4.2 Markdown 其他编写要求	4
4.3 生成 PDF 文件	4
5 摆脱命令行，优雅的 VSCode 书写转换方案	4
5.1 VSCode 与插件安装	4
5.2 使用 VSCode 书写 Markdown	4
5.3 配置 MPE 使用 Pandoc 导出	5
6 常见问题解决	8
6.1 LaTeX 相关错误	8
6.1.1 Markdown 转换 LaTeX	8
6.1.2 LaTeX 编辑器打开，以 TexStudio 为例	8

1 折腾背景

Markdown 的简便性是 LaTeX 无法替代的，LaTeX 对排版的精准控制能力又是 Markdown 无法比拟的。一直在寻找一种能够将 Markdown 优雅地转换成 PDF 的方案，虽然早就听说也使用过 Pandoc 这把瑞士军刀，但是它太过强大，以致于一直都没用明白。只会简单的转换命令，但是实际效果并不好，最近学会了使用 LaTeX 模板的功能，这才让我眼前一亮，这才是我想要的结果。

2 基础环境配置

Markdown 生成 PDF 主要需要使用 Pandoc 和 Latex 两个工具，具体安装方式如下：

2.1 Pandoc 的安装

Pandoc 是由 John MacFarlane 开发的标记语言转换工具，可实现不同标记语言间的格式转换。

- Windows 下的安装：
 - 下载[安装包](#)直接安装即可
 - 如果安装了 Chocolatey: `choco install pandoc`
 - 如果安装了 winget: `winget install pandoc`
- Linux/FreeBSD 下的安装：
 - Pandoc 已经包含在大部分 Linux 发行版的官方仓库中，直接使用诸如[apt](#)/[dnf](#)/[yum](#)/[pacman](#)之类的安装工具直接安装即可
- MacOS 下的安装：
 - `brew install pandoc`

详细的安装说明参见：[官方安装文档](#)

2.2 LaTeX 的安装

LaTeX 工具, 建议安装 texlive。

- Windows 下的安装：
 - [参考该文章](#)下载完整 texlive，注意安装后需要再安装 cjk, cjk-fonts 等相关 package

- Linux/FreeBSD 下的安装：
 - 使用 `apt/dnf/yum/pacman/pkg` 等安装工具安装 `texlive`、`texlive-latex` 等相关软件包
- MacOS 下的安装：
 - 使用 HomeBrew 安装 `texlive` 即可

3 模板配置

3.1 配置 Pandoc 模板

为保证生成的 pdf 格式（自动插入封面、目录页、页眉页脚等信息），在本地环境中安装模板，具体步骤是：

- 下载[MPPL: Markdown to PDF with Pandoc via Latex](#)仓库
- 将`template/mppl.latex`拷贝到`*/pandoc/templates`目录下
 - Window 下：`C:/Users/USERNAME/AppData/Roaming/pandoc/templates`，如果 `Roaming` 没有 `pandoc` 目录，请手动创建！
 - Linux/FreeBSD/MacOS：`~/.pandoc/templates/`

3.2 配置 LaTeX 模板

模板定制主要修改模板最前面的**模板基础配置**相关内容，主要可修改的包括：

- 公司和组织，目前默认是“MPPL”
- 正文缩进，目前默认是 2em（2 个中文字符，4 个英文字符）
- 主要中文字体和英文字体：目前都是微软雅黑
- 页眉、页脚展示内容，目前是：
 - 左页眉：title
 - 右页眉：“企业机密-禁止外传”
 - 左页脚：company
 - 右页脚：页码

3.3 字体设置

目前页面默认的字体是微软雅黑，对于非 Windows 系统，可能不存在该字体，则有以下两种解决方案：

1. 手工安装微软雅黑字体（需要 msyh,msyhbd 两个文件）
2. 修改为其他字体，如苹方、文泉驿等

若需要多个团队共同使用，建议采用方案一。

4 如何生成 PDF

4.1 PDF 文件指定 metadata 信息

在每个 markdown 最前面增加以下主要 metadata 信息，metadata 内容开始行内容为三个“-”，结束行为三个“.”，示例如下：

```
1 title: "MPPL"
2 version: "0.1"
3 subtitle: "Markdown to PDF with Pandoc via LaTeX"
4 date: "2022-08"
5 author: "Dominic"
6 company: COMPANYNAME
7 file-code: COMPANY-DEPARTMENT-00000000
8 logo: true
9 logo-url: ./img/logo.png
10 history:
11   - version: V0.1
12     author: Dominic
13     date: 2022-08-19
14     desc: 创建文档
```

其他可选配置项目如下：

- header-left: 左页眉
- header-right: 右页眉
- footer-left: 左页脚
- footer-right: 右页脚
- CJKmainfont: 主要中文字体
- mainfont: 主要字体
- lot: 是否创建表格目录
- lof: 是否创建图片目录

可选配置项中，建议除了 subtitle 外，全部在模板中定制，不在 markdown 文件中定制

4.2 Markdown 其他编写要求

Pandoc 默认使用的 `pandoc_markdown` 格式，为避免 markdown 转 pdf 格式异常，在编写 markdown 的时候有几个原则要求：

- 每个标题前后都必须有空行
- 每个表格前后都必须有空行
- 每个代码块前后都必须有空行
- 每个列表前后都必须有空行

总之，不同文本类型之间都要有空行。

4.3 生成 PDF 文件

```
1 pandoc --listings --pdf-engine=xelatex --template=mppl.latex README.md  
  -o README.pdf
```

5 摆脱命令行，优雅的 VSCode 书写转换方案

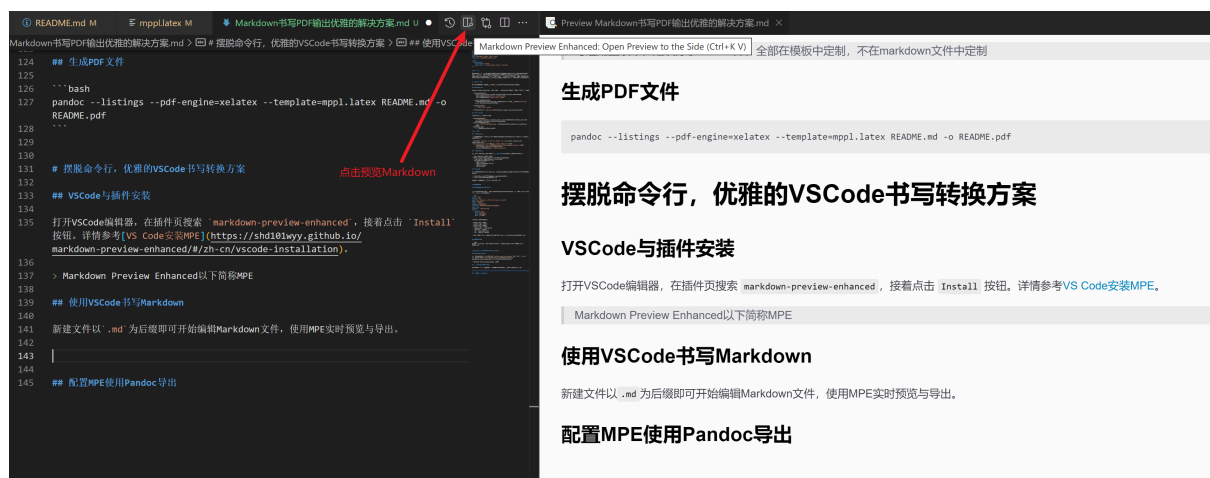
5.1 VSCode 与插件安装

打开 VSCode 编辑器，在插件页搜索 `markdown-preview-enhanced`，接着点击 `Install` 按钮。详情参考[VS Code 安装 MPE](#)。

Markdown Preview Enhanced 以下简称 MPE

5.2 使用 VSCode 书写 Markdown

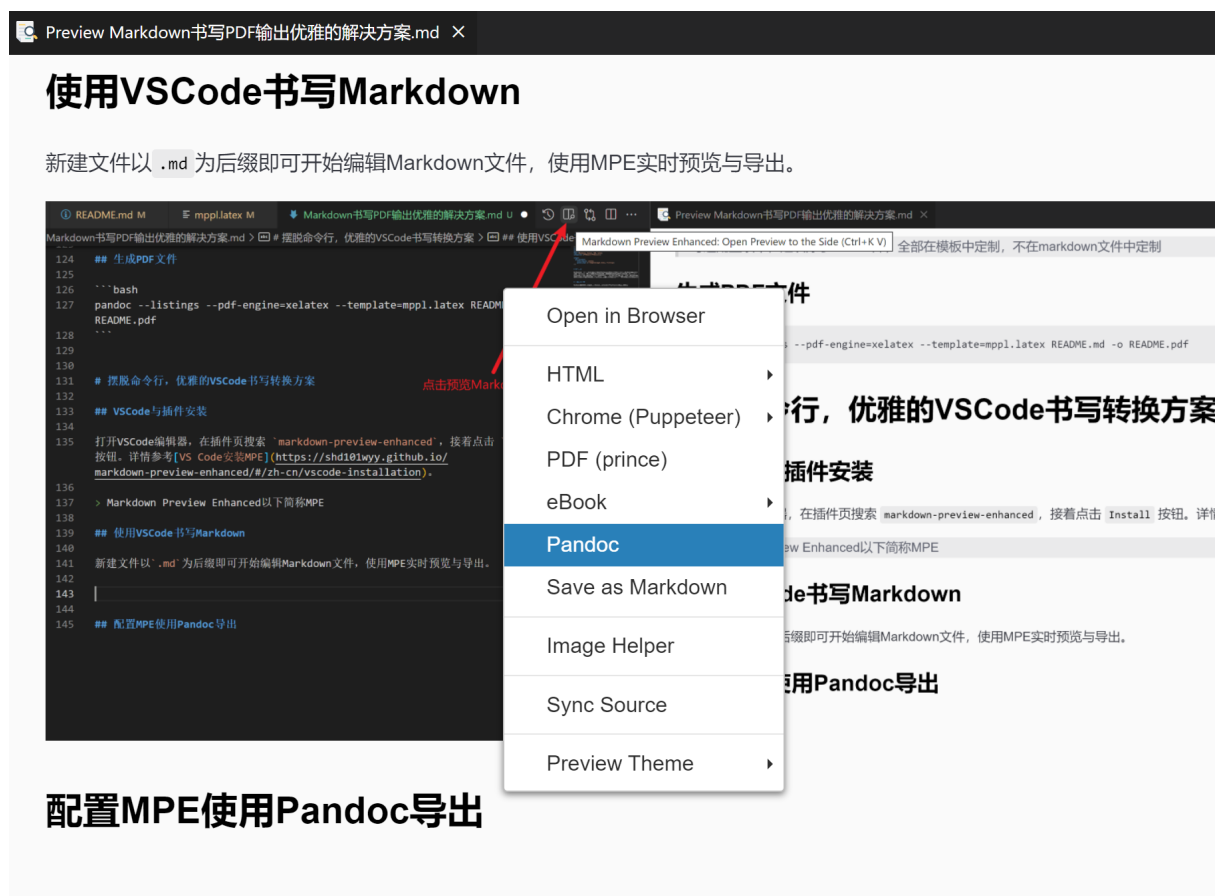
新建文件以 `.md` 为后缀即可开始编辑 Markdown 文件，使用 MPE 实时预览与导出。



5.3 配置 MPE 使用 Pandoc 导出

右击 MPE 的预览区域，可以看到 MPE 提供多种导出 PDF 的方案，如使用 Chrome 的 Puppeteer 导出，Prince 导出，Pandoc 导出等等。

在未使用 Pandoc 前，我也一直使用 MPE 提供的 Chrome 方式导出，但是导出的 PDF 排版总是不尽如意。现在介绍如何使用 Pandoc 方式导出。



创建 PDF 文档，你需要在 markdown 文件中的 front-matter 里声明 pdf_document 的输出类型：

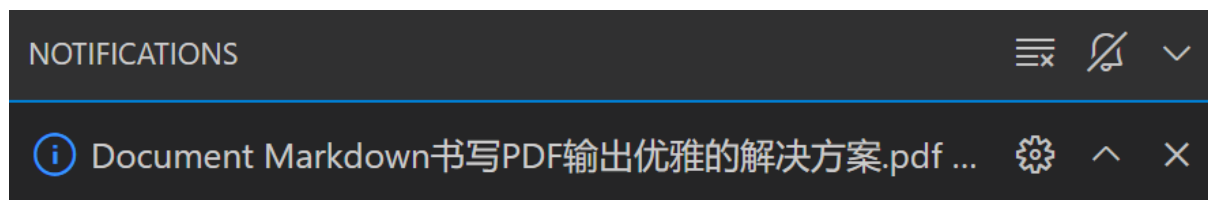
```
1 ---
2 output:
3   pdf_document:
4     latex_engine: xelatex
5     pandoc_args: [--template=mppl.latex,--listings]
6 ---
```

- front-matter: 文章的最开头，也就是上文元数据放的地方。和元数据放在一起即可，如图所示：

```
1  ---
2  output:
3    pdf_document:
4      latex_engine: xelatex
5      pandoc_args: [--template=mppl.latex,--listings]
6
7  title: "Markdown书写PDF输出优雅的方案"
8  version: "0.1"
9  subtitle: "Markdown书写PDF输出优雅的方案"
10 date: "2022-08"
11 author: "Dominic"
12 company: COMPANYNAME
13 file-code: COMPANY-DEPARTMENT-00000000
14 history:
15   - version: V0.1
16     author: Dominic
17     date: 2022-08-19
18     desc: 创建文档
19 ---
20
21
22 # 折腾背景
23
24 Markdown的简便性是LaTeX无法替代的，LaTeX对排版的精准控制能力又是Markdown无法比拟的。一直在寻找一种能够将Markdown优雅地转换成PDF的方案，虽然早就听说也使用过Pandoc这把瑞士军刀，但是它太过强大，以致于一直都没用明白。只会简单的转换命令，但是实际效果并不好，最近学会了使用LaTeX模板的功能，这才让我眼前一亮，这才是我想要的结果。
25
26
27 # 基础环境配置
28
29 Markdown生成PDF主要需要使用Pandoc和Latex两个工具，具体安装方式如下：
```

- `latex_engine`: 默认情况下 PDF 文档由 `pdflatex` 生成。你可以用 `latex_engine` 选项来定义你想用的引擎。支持的引擎有 `pdflatex`, `xelatex`, 以及 `lualatex`。这里需要使用 `xelatex`。
- `pandoc_args`: 配置 Pandoc 接受的一些参数，这里我们使用 `--template=mppl.latex` 和 `--listings` 来指定模板和使用 `listings`。这里配置的参数就是执行 Pandoc 时使用的参数，以后就不需要命令行输入了。这里使用上文的 `mppl.latex` 模板。

配置完之后，右击预览界面，选择 Pandoc 导出，稍等片刻，即可生成 PDF 文件。

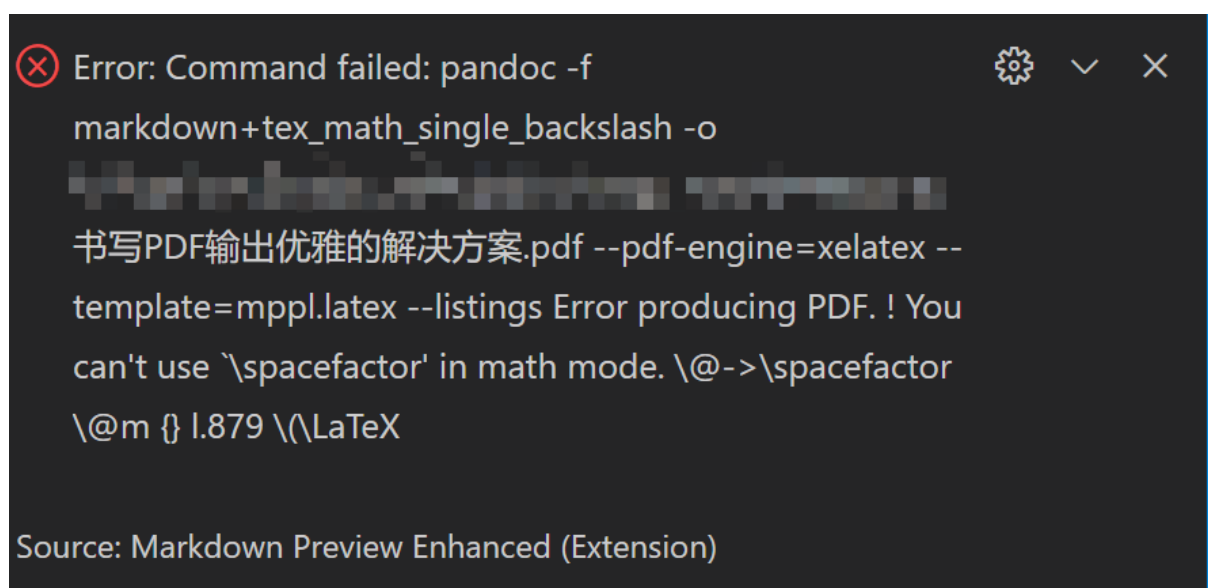


6 常见问题解决

6.1 LaTeX 相关错误

VSCode 导出出错时报错信息较短，并且常常不知道具体报错原因及位置，因为是 LaTeX 转换成 PDF 的过程中出现的错误。报错位置是 LaTeX 中间源码的位置，而不是 VSCode 中的位置。这时候我常用的方法是先将 Markdown 转为 LaTeX，然后再转为 PDF，在 LaTeX 编辑器里就可以看到错误位置了。

比如下面这个错误，我们能看到一些报错信息 `cant use \spacefactor in math mode`，但是并不知道具体哪里的错误。从信息里可以看出和 `\LaTeX` 有关，大概能推测出是使用了这个命令，因为文章里使用了这个命令的地方只有一处。但是如果有其他的错误，就很难确定了。



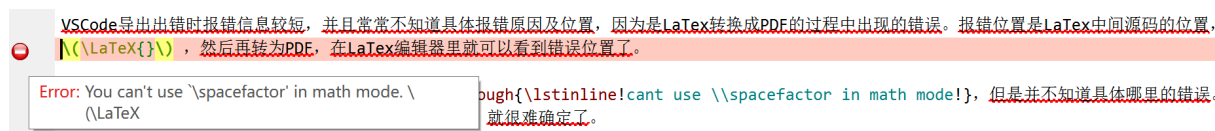
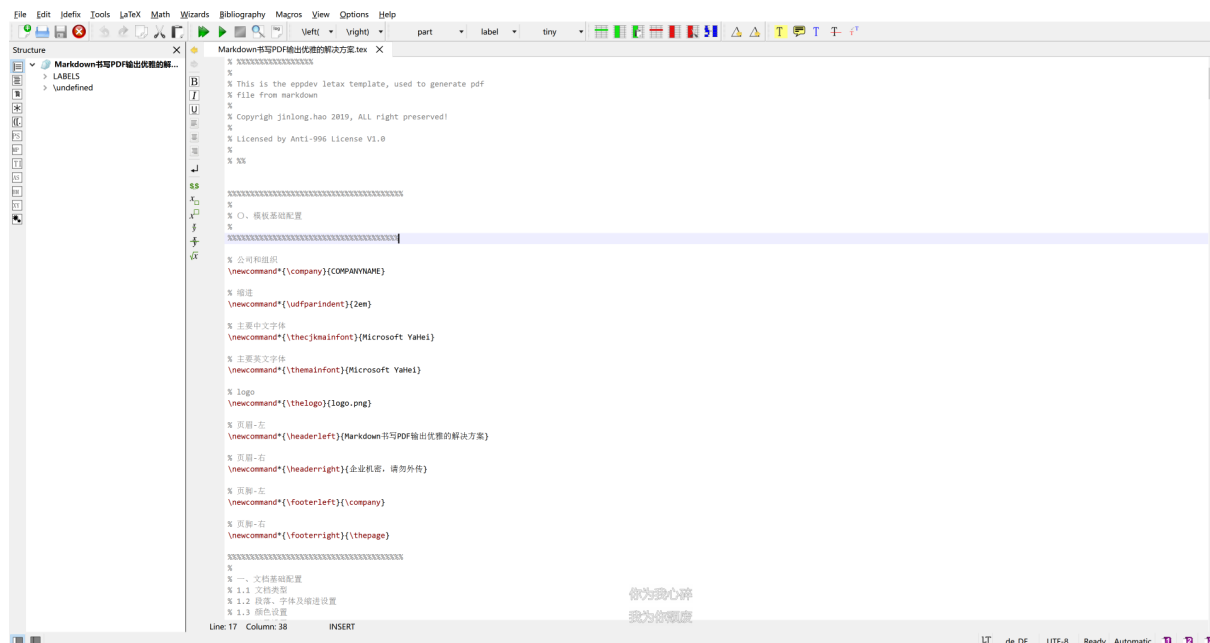
6.1.1 Markdown 转换 LaTeX

这里还是以模板仓库的 `README.md` 为例，当然这个文件是可以正常转换 PDF 的，不会报错。这里只是拿 `README.md` 做一个如何使用命令的演示。

```
1 pandoc --listings --template=mppl.latex -s README.md -o README.tex
```

6.1.2 LaTeX 编辑器打开，以 TexStudio 为例

打开 `README.tex` 文件，编译：



我们可以快速的定位到问题出现的位置，只要搜索相关问题即可。

\LaTeX 这个宏不能用在数学模式下。但是因为我在 Markdown 里必须使用美元符号 $\$$ 才能表示 LaTeX 环境，才能正确输出 LaTeX 符号，而 Markdown 转换成 LaTeX 源码时，这个宏就会被包裹在数学环境里，就会报错。如果我想在 PDF 里显示这个符号，那就在 Markdown 里不使用美元符号 $\$$ ，而是直接输入 $\backslash\text{LaTeX}$ 即可，再导出 PDF 时就不会报错。