

基于 Markdown 的论文写作方法

作者 作者 2

2020 年 12 月 11 日

摘要

本文介绍了基于 Pandoc 使用 Markdown 写作论文的一种方案。Markdown 是一种简单易学的标记语言。相对于 L^AT_EX, 使用 Markdown 作为写作语言, 能够让作者专注于内容而非排版, 排版工作则由 Pandoc 基于预定义的 L^AT_EX 模板生成 PDF。本方案既能获得 L^AT_EX 排版效果, 又避免了 L^AT_EX 的复杂性, 尤其适合 L^AT_EX 新手, 对于熟悉 L^AT_EX 的作者也不失为不错的替代选择。另外, 由于 Markdown 内容和 L^AT_EX 模板分离, 作者只需维护一份源码, 通过编译生成不同样式的文档, 降低了维护成本。

关键词: L^AT_EX, PanBook, Markdown, 论文排版

1 使用方法

本章简要介绍 panbook¹ 的安装, 基本使用方法及 Pandoc's Markdown 语法。建议读者阅读 PanBook 使用手册^{2[1]} 获取更详细的内容。

1.1 快速入门

以 Windows 10 为例, 演示如何使用。首先需要安装依赖软件。

- 安装 msys2 (Linux 及 OS X 请忽略此步骤)
- 安装 texlive 2018 或以上版本
- 安装 Pandoc 2.7.3 或以上版本
- 下载 pandoc-crossref 对应版本安装到 path 目录下 (建议和 Pandoc 放同一目录)

¹<https://github.com/annProg/PanBook>

²<https://panbook.annhe.net/pub/PanBook-book-ctexbook-pc.pdf>

然后下载 PanBook。打开终端 (msys2)，假设工作目录为 /d/dev，克隆代码并设置环境变量：

```
$ cd /d/dev
$ git clone https://github.com/annProg/PanBook
# 将 PanBook, TeXLive 及 Pandoc 加入环境变量
$ tail -n 1 ~/.bashrc
export PATH=$PATH
↪ :/d/texlive/2018/bin/win32:/d/dev/PanBook:/c/Users/myname/AppData/Local/Pandoc
```

完成环境变量设置之后，在任意空目录下执行 `panbook thesis`，会自动初始化写作环境，生成示例源码。然后在 `src` 目录下开始写作。目录规范见代码 1。

代码 1: PanBook 目录规范

```
.
|-- extensions          # 自定义扩展
|-- fonts              # 自定义字体
|-- src                # Markdown 源码目录
|   |-- images         # 插图目录
|   |-- metadata.yaml  # 书籍元数据文件
|   |-- frontmatter.md # 前言文件
|   |-- backmatter.md  # 后记文件
|   |-- 100-chapter1.md # 正文，命名须保证能按正确章节顺序列出
|   |-- 200-chapter2.md
|-- styles             # 自定义风格
|-- templates         # 自定义模板
|-- build             # 电子书构建目录
```

1.2 指定风格 (style)

风格 (style)，也可以理解为模板 (为了和 pandoc 模板区分)，是预定义的文档样式。通过 `--style` 参数指定，缺省时，会使用默认风格 (thesis)，通过命令 `panbook thesis -l` 查看模块支持的风格列表：

```
$ panbook thesis -l
thesis
```

支持和计划支持的学校论文模板如表 1 所示。

表 1: 支持的论文模板

学校	style 名称	状态
demo	thesis	完成
湖南大学	hnu	WIP
清华大学	thu	WIP
上海交通大学	sjtu	WIP
南京大学	nju	WIP
山东大学	sdu	WIP
中国科学院大学	ucas	WIP
电子科技大学	uestc	WIP
北京航空航天大学	buaa	WIP

1.3 调试模式

加 `-d` 选项, 会输出详细的 `latexmk` 编译过程, 加 `--trace` 选项, 可以输出更多的调试信息。

1.4 注意事项

- Markdown 源码文件需要使用 UTF-8 编码
- Pandoc 扩展的 Markdown 语法要求在标题前留出一个空行, 因此按章节拆分的多个 Markdown 文件, 开头需要空一行, 否则 pandoc 不能正确识别标题
- 请勿将正文文件命名为 `frontmatter.md` 或者 `backmatter.md`, 这 2 个文件有特殊用途

1.5 论文元数据

在 `src/metadata.yaml` 中使用 Yaml 语言³ 定义书籍的数据及可用的模板变量, 示例见代码 2。

查看模板文件, 可以获取模板支持的所有变量 (形如 `var`)。也可以通过修改模板来添加自定义的变量。

³<http://www.ruanyifeng.com/blog/2016/07/yaml.html>

代码 2: Metadata

```

---
title: 基于 Markdown 的论文写作方法
etitle: A Thesis Writing Method Based On Markdown
author:      # 作者 (数组)
  - 作者
xuehao: 201907131122
college: 计算机学院
supervisor: \LaTeX 教授
date: \today # 日期
lof: false # 是否生成插图列表页
lot: false # 是否生成表格列表页
# 不引用, 但是显示在参考文献列表里。通配符 @* 表示全部在列表里显示
nocite: |
  @pandocManual
...

```

2 Pandoc Markdown 扩展语法

Pandoc 的目标与原始 Markdown 的最初目标有着方向性的不同。在 Markdown 原本的设计中, HTML 是其主要输出对象; 然而 Pandoc 则是针对多种输出格式而设计。因此, 虽然 Pandoc 同样也允许直接嵌入 HTML 标签, 但并不鼓励这样的作法, 取而代之的是 Pandoc 提供了许多非 HTML 的方式, 来让使用者输入像是定义列表、表格、数学公式以及脚注等诸如此类的重要文件元素。

Pandoc Markdown 语法介绍可以在 Pandoc 主页 找到。中文翻译请参考 PanBook 使用手册^{4[1]}。

2.1 交叉引用

Pandoc Markdown 对象 ID 语法格式形如 `{#label}`, 使用交叉引用, 还需遵循表 2 中的 ID 前缀要求。

基本的引用格式为 `[@label]`, 注意, 比如有一个图片的 ID 是 `{#fig:myfigure}`, 则引用为 `[@fig:myfigure]`。更复杂的引用格式请参考 PanBook 使用手册^[1]。

⁴<https://panbook.annhe.net/pub/PanBook-book-ctexbook-pc.pdf>

表 2: 交叉引用前缀规范

类型	前缀	示例
图片	fig:	{#fig:label}
表格	tbl:	{#tbl:label}
公式	eq:	{#eq:label}
代码	lst:	{#lst:label}
章节	sec:	{#sec:label}
定理	thm:	{#thm:label}
定义	def:	{#def:label}
引理	lem:	{#lem:label}
推论	cor:	{#cor:label}
命题	pro:	{#pro:label}

2.2 代码

普通代码块和原生 Markdown 语法一致, 如果需要包含 label 及 caption, 可用 `{#label .class caption="My Caption"}` 格式, `.class` 可以有多个, 一般第一个是代码语言类型。示例见代码 3。

代码 3: 示例代码

```
package main

import "fmt"
func main() {
    fmt.Println("hello world")
}
```

2.3 图片

直接使用 PanBook plot 扩展来展示图片, 详情请参考 PanBook 使用手册^[1]。效果见图 1。

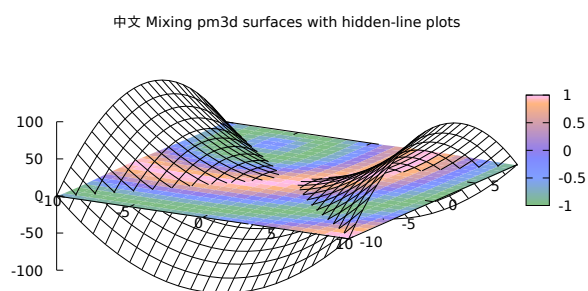


图 1: gnuplot 示例

2.3.1 子图

效果见图 2。

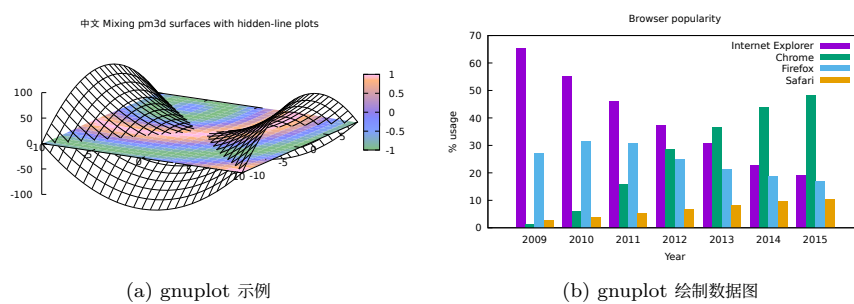


图 2: 子图示例

2.4 表格

效果请看表 2。

2.5 公式

以下是一个例子, 显示效果见公式 1:

```


$$a_1x + b_1y + c_1z = d_1$$


```

```
a_2x+b_2y+c_2z=d_2\\
a_3x+b_3y+c_3z=d_3\\
\end{cases} {#eq:math_demo}
```

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \quad (1)$$

2.6 参考文献

参考文献使用 biblatex 格式管理, 引文和引用格式化使用 Citation Style Language⁵, Zotero 样式库⁶ 可以下载到 csl 文件, 通过 PanBook 参数 `--csl` 指定 csl 文件。

文献引用放在方括号中, 以分号隔开。每一条引用都需要有一个 key, 由 @ 加上文献目录数据库中的文献 ID 组成, 并且可以选择性地包含前缀、定位以及后缀。引用键必须以字母、数字或 _ 开头, 并且可以包含字母数字、_ 和内部标点符号 (: . # \$ % & - + ? < > ~ /)。以下是一些范例:

```
Blah blah [@panbook].
Blah blah [see @doe99, pp. 33-35; also @smith04, ch. 1].
Blah blah [@doe99, pp. 33-35, 38-39 and *passim*].
Blah blah [@smith04; @doe99].
```

更多信息请参考 PanBook 使用手册^[1]。

参考文献

[1] HE A. PanBook 使用手册 [M]. <https://panbook.annhe.net/pub/PanBook-book-ctexbook-pc.pdf>; PanBook, 2019.

[2] MACFARLANE J. Pandoc User Manual[J]. <https://pandoc.org/MANUAL.html>, 2019.

[3] KING H. Test Dummy Book[M]. China, 1321.

⁵<https://www.zotero.org/styles>

⁶<https://citationstyles.org/>