Easier LaTeX with Pandoc and Markdown

Convenient Academic Writing

Nguyen, Duc Hieu

May 30, 2018

Markdown for Academic Writing

Criteria for writing tools

Structured document:

- Title page
- Abstract
- Table of contents
- Content
- Footnotes
- Citations

Additional Management:

- Images, tables and Formulas
- Managing Bibliography
- Collaboration

Markdown

- Created by John Gruber and Aaron Schwarz
- Lightweight markup language in plain text
- Convertible into HTML/XHTML
- One of the standard on the net:
  - README files
  - Forum & Blog Posts
  - Chat services
  - E-mails

Advantages of Markdown for Academics

Like LaTeX :

- Separation of content and form
  $\rightarrow$Focus on writing
- Less compatibility issues
- Platform independence
- Logging and merging with VCSs

Addition:

- Easier to read and write
- Easier Collaboration

LaTeX vs Markdown

Basic Markdown: Some things you can format

- Header
- Lists
- Paragraph & line breaks
- Emphasis
- Horizontal rules
- Code blocks
- Block quotes
- Links
- Images
- Tables

Header, Paragraph & Emphasis

LaTeX

```latex
\section{Header 1}

\subsection{Header 2}

\subsubsection{Header 3}

This is a paragraph \\
with a new line.

\textit{italic}
\textbf{bold}
```

Markdown

```markdown
[atx-style] [setext-style]
# Header 1,     Header 1
                ========
## Header 2,    Header 2
                --------
### Header 3

This is a paragraph \
with a new line.

*italic*, _italic_
**bold**, __bold__
```

Lists

LaTeX

```latex
\begin{itemize}
  \item some
  \item items
  \begin{itemize}
    \item sub
    \item items
  \end{itemize}
\end{itemize}
```

Markdown

```markdown
- some
- items
  - sub
  - items
```

```markdown
* some
* items
  * sub
  * items
```

Links and Images

LaTeX :

```latex
% Links
\usepackage{hyperref}
\href{http://example.net}{descrition}
% Images
\includegraphics{/path/to/image}
```

Markdown:

```markdown
<!-- Links -->
[link](example.net)
<!-- Images -->
![Alt](/path/to/image.jpg)
```

Tables: Example

Table 1: A simple Table

Left (default)	Centered	Right
First	row	12
Second	row	123
Third	row	2

Table in LaTeX generated by Pandoc

```
\usepackage{longtable}
\begin{longtable}[]{@{}lcr@{}}
  \caption{Sample Table}\tabularnewline
  \toprule
  Left (default) & Centered & Right\tabularnewline
  \midrule
  \endfirsthead
  \toprule
  Left (default) & Centered & Right\tabularnewline
  \midrule
  \endhead
  First & row & 1\tabularnewline
  Second & row & 123\tabularnewline
  Third & row & 22\tabularnewline
  \bottomrule
\end{longtable}
```

Similar Table in LaTeX

```
\begin{center}
  \begin{tabular}{l c r}
    \toprule
    Left (default) & Centered & Right \\
    \midrule
    First & row & 1 \\
    Second & row & 123 \\
    Third & row & 22 \\
    \bottomrule
  \end{tabular}
\end{center}
```

Sample Table in Markdown (1)

Simple Tables

```
Left (default)    Centered    Right
---------------- ---------- -------
First                row          12
Second               row         123
Third                row           2
```

Pipe Tables

```
| Left (default) | Centered | Right |
|:---------------|:--------:|------:|
| First          | row      | 12    |
| Second         | row      | 123   |
| Third          | row      | 2     |
```

Sample Table in Markdown (2)

Grid Tables

```
+----------------+----------+-------+
| Left (default) | Centered | Right |
+:===============+:========:+======:+
| First          | row      | 12    |
+----------------+----------+-------+
| Second         | row      | 123   |
+----------------+----------+-------+
| Third          | row      | 2     |
+----------------+----------+-------+
```

Multiline Tables (1)

```
-------------------------------------------------------------
 Centered   Default           Right Left
  Header    Aligned         Aligned Aligned
---------- ------- ------------ -------------------------
   First    row                12.0 Example of a row that
                                    spans multiple lines.

  Second    row                 5.0 Here's another one. Note
                                    the blank line between
                                    rows.
-------------------------------------------------------------

Table: Here's the caption.
It, too, may span multiple lines.
We might need further information about tables.
```

Multiline Tables (2)

Table 2: Here's the caption. It, too, may span multiple lines. We might need further information about tables.[1]

Centered Header	Default Aligned	Right Aligned	Left Aligned
First	row	12.0	Example of a row that spans multiple lines.
Second	row	5.0	Here's another one. Note the blank line between rows.

[1] https://pandoc.org/MANUAL.html#tables

Pandoc

Pandoc

- Created in 2006 by John MacFarlane
- Commandline tool for converting document formats
- Intention:
  - Markdown originally designed for HTML generation
  - Pandoc is designed for different output formats
  - Initial document should be written in Markdown

Pandoc: Universal Converter

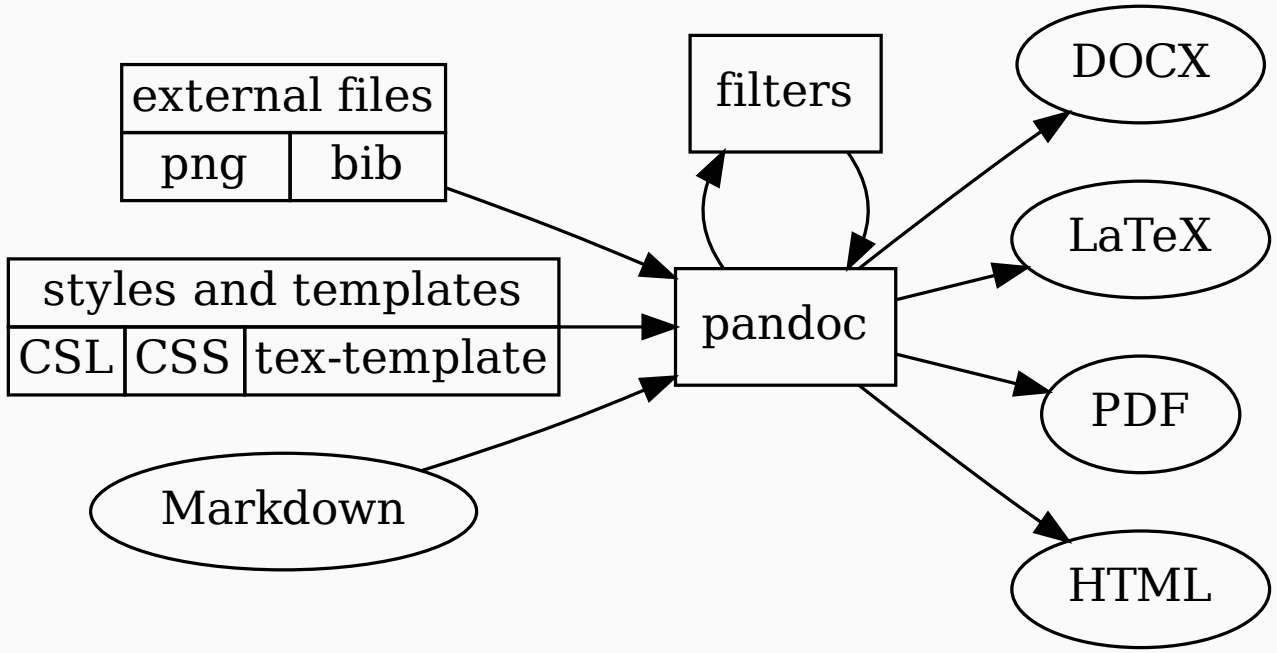
Input: 28 formats (6 markdown flavours)

commonmark (CommonMark Markdown), creole (Creole 1.0), docbook (DocBook), docx (Word docx), epub (EPUB), fb2 (FictionBook2 e-book), gfm (GitHub-Flavored Markdown), haddock (Haddock markup), html (HTML), jats (JATS XML), json (JSON version of native AST), latex (LaTeX), markdown (Pandoc's Markdown), markdown_mmd (MultiMarkdown), markdown_phpextra (PHP Markdown Extra), markdown_strict (original unextended Markdown), mediawiki (MediaWiki markup), muse (Muse), native (native Haskell), odt (ODT), opml (OPML), org (Emacs Org mode), rst (reStructuredText), t2t (txt2tags), textile (Textile), tikiwiki (TikiWiki markup), twiki (TWiki markup), vimwiki (Vimwiki)

Output: 45 formats

asciidoc (AsciiDoc), beamer (LaTeX beamer slide show), commonmark (CommonMark Markdown), context (ConTeXt), docbook or docbook4 (DocBook 4), docbook5 (DocBook 5), docx (Word docx), dokuwiki (DokuWiki markup), epub or epub3 (EPUB v3 book), epub2 (EPUB v2), fb2 (FictionBook2 e-book), gfm (GitHub-Flavored Markdown), haddock (Haddock markup), html or html5 (HTML, i.e. HTML5/XHTML polyglot markup), html4 (XHTML 1.0 Transitional), icml (InDesign ICML), jats (JATS XML), json (JSON version of native AST), latex (LaTeX), man (groff man), markdown (Pandoc's Markdown), markdown_mmd (MultiMarkdown), markdown_phpextra (PHP Markdown Extra), markdown_strict (original unextended Markdown), mediawiki (MediaWiki markup), ms (groff ms), muse (Muse), native (native Haskell), odt (OpenOffice text document), opml (OPML), opendocument (OpenDocument), org (Emacs Org mode), plain (plain text), pptx (PowerPoint slide show), rst (reStructuredText), rtf (Rich Text Format), texinfo (GNU Texinfo), textile (Textile), slideous (Slideous HTML and JavaScript slide show), slidy (Slidy HTML and JavaScript slide show), dzslides (DZSlides HTML5 + JavaScript slide show), revealjs (reveal.js HTML5 + JavaScript slide show), s5 (S5 HTML and JavaScript slide show), tei (TEI Simple), zimwiki (ZimWiki markup)

How Pandoc works



Requirements

- Text editor: Notepad++, Geany
- Commandline Terminal
  - Windows: Powershell
  - MacOS: Terminal, iTerm
  - Linux: Terminal(gnome-terminal), Konsole, xterm
- Pandoc: https://pandoc.org/installing.html
- LaTeX :
  - Windows: MiKTeX (http://miktex.org/)
  - MacOS: MacTeX (http://www.tug.org/mactex/)
  - Linux: TeX Live (http://www.tug.org/texlive)

Command

To generate PDF documents:

```
# for articles
pandoc input.md -o output.pdf
# for beamer presentations
pandoc -t beamer input.md -o output.pdf
```

Additional flags:

- `-s`, `--standalone`
- `--filters` [FILE]
- `--highlight-style=`[FILE]
- `--template=`[FILE]
- `-t` [TARGET FORMAT]

For more: `man pandoc`

Pandoc: YAML-Header

In Pandoc metadata for a document are written in YAML (usually at the top of Markdown documents):

```
---
title: Title of your work
author: Name of Author
date: 11.11.2011
tags: [markdown, writing]
abstract: |
  Abstract text here.
---
```

Raw TeX (1)

Inline TeX commands will be preserved
and passed unchanged to the LaTeX writers:

```
You can use \LaTeX\ to create
\textbf{bold} or \textit{italic} text.
```

Renders:

You can use LaTeX to create **bold** or *italic* text.

Raw TEX (2)

Detailed TEX Tables are easily added if necessary:

```
\begin{tabular}{|r|l|}
  \hline
  7C0 & hexadecimal \\
  3700 & octal \\ \cline{2-2}
  11111000000 & binary \\
  \hline \hline
  1984 & decimal \\
  \hline
\end{tabular}
```

7C0	hexadecimal
3700	octal
11111000000	binary
1984	decimal

TeX Math

TeX Math is written between two $-signs

```
<!-- Inline math -->
Here we see some inline math: $a^2 + b^2 = c^2$
<!-- displayed equation -->
And some displayed equation:
$$ \sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$
```

both render:

Here we see some inline math: $a^2 + b^2 = c^2$

And some displayed equation:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

Footnotes (1)

```
The next sentence has a note.
To look up about footnotes:[^1]
Another sentence that has a long note.[^longnote]

[^1]: <https://pandoc.org/MANUAL.html#footnotes>

[^longnote]:
    Note with multiple blocks.
    { some.code }
    The whole paragraph can be indented,
    or just the first line.
    In this way, multi-paragraph footnotes work like
    multi-paragraph list items.
```

Footnotes (2)

The next sentence has a note. To look up about footnotes:[2]
Another sentence that has a long note.[3]

[2] https://pandoc.org/MANUAL.html#footnotes

[3] Note with multiple blocks. { some.code } The whole paragraph can be indented, or just the first line. In this way, multi-paragraph footnotes work like multi-paragraph list items.

Citations (1)

Bibliographies are managed with Bib-files.

```
# mybib.bib
@article{macfarlane2013pandoc,
    title={Pandoc: a universal document converter},
    author={MacFarlane, John},
    url={http://pandoc.org},
    year={2013}
}
```

To use it in your document:

```
John MacFarlane's Pandoc [@macfarlane2013pandoc]
```

Citation (2)

Citations are generated through an external filter:

`pandoc-citeproc`

```
pandoc --filter pandoc-citeproc input.md -o output.pdf
```

Bibliographies are either added as metadata into YAML

```
---
bibliography: mybib.bib
---
```

Or are added as argument:

```
pandoc --bibliography mybib.bib ...
```

Intermediate features

Templates

Generate default templates for further customization:

```
pandoc -D [FORMAT] > [filename]
pandoc -D latex > template.tex
pandoc -D beamer > template.beamer
```

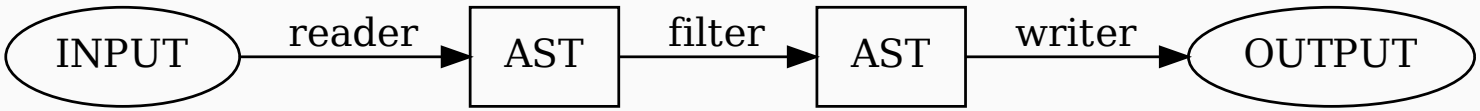
To use it: `--template my-template.tex`

Reuse for multiple projects move to:

- Unix, Linux, macOS: `~/.pandoc/templates/`
- Windows XP:
  `C:\Documents And Settings\USERNAME\Application Data\pandoc`
- Windows Vista or later:
  `C:\Users\USERNAME\AppData\Roaming\pandoc`

Filters

Filters: Programs for manipulating Pandoc's representation of the document: AST – the "Abstract syntax tree". Users can create their own for their specific needs.



Use them with `--filter`:

```
pandoc --filter filter.py input.md -o output.pdf
```

Limitations of Markdown

- No further customization of Tables
  - cannot add lines between rows and columns
  - cannot span over rows and columns
- Free nesting of LaTeX and Markdown not possible
- Pandoc generates cross references with hyperlinks instead of `\label` and `\ref`
- Math is only inline or display expressions (latter as `\displaymath`)
  - not possible to specify other environments: `equation`, `gather` etc.

Summary and Links

Summary

- Pandoc Markdown features as alternative to LaTeX for academics
- Pandoc can read and write countless kinds of formats
- Personalized templates and filters extend already powerful tool

Links

- Markdown:
  - https://programminghistorian.org/en/lessons/getting-started-with-markdown
- Latex:
  - https://en.wikibooks.org/wiki/LaTeX
  - https://www.cs.princeton.edu/courses/archive/spr10/cos433/Latex/latex-guide.pdf
- Pandoc:
  - https://pandoc.org
  - https://pandoc.org/MANUAL.html
- Panflute:
  - http://scorreia.com/software/panflute/
- My Presentation:
  - https://github.com/ndhieu1994/digi-philo-pandoc-markdown/

References

Dominici, Massimiliano. 2014. "An Overview of Pandoc." *TUGboat* 35 (1): 44–50. https://www.tug.org/TUGboat/tb35-1/tb109dominici.pdf.

Gruber, John. 2004. "Daring Fireball: Markdown." https://daringfireball.net/projects/markdown/.

MacFarlane, John. 2013. "Pandoc: A Universal Document Converter." https://pandoc.org.