

Como usar Pandoc

Enno Nagel

CoAITI, Hotel Premier Maceió, 28 Outubro 2017

pandoc: conversor entre linguagens de marcação

O que é uma *Linguagem de Marcação* ?

Uma *linguagem de marcação* é uma linguagem para formatar o texto e as imagens de uma página. Por exemplo, para formatar uma página visualizada pelo navegador, tal como o Firefox,

- a linguagem usada é HTML e a sua sintaxe, por exemplo:

"É `Importante!`"

- a linguagem Markdown compila a HTML e usa uma sintaxe mais intuitiva, por exemplo:

"É `**Importante!**`"

O resultado da formatação \rightsquigarrow "É **Importante!**"

O que é pandoc ?

pandoc é um programa (de linha de comando) que converte um arquivo de uma linguagem de marcação a outra: De (entre outras)

- markdown, HTML, LaTeX, Microsoft Word docx, reStructuredText, textile, OPML, Emacs Org-Mode, Txt2Tags, DocBook, EPUB ou Haddock markup

...

a (entre outras)

- HTML: XHTML, HTML5, e eslaides por Slidy, reveal.js, S5 ...
- Processador de Textos: Microsoft Word docx, LibreOffice e OpenOffice ODT, OpenDocument XML
- PDF via LaTeX
- Ebooks: EPUB version 2 or 3, FictionBook2
- Documentação: DocBook, GNU TexInfo, Groff man

- TeX: LaTeX, ConTeXt, eslaides LaTeX Beamer

O que faz pandoc para mim ?

Converte arquivos do formato

- markdown

...

ao formato

- HTML5
- Microsoft Word docx , LibreOffice ODT, OpenDocument XML
- LaTeX Beamer eslaides
- PDF via LaTeX

Pandoc versus Programas Especializados: O bem.

Acessibilidade:

Código markdown é texto de fácil leitura.

Em comparação com outras linguagens de marcação:

- A sintaxe markdown é mais simples do que a sintaxe (La)TeX (O inventor de TeX, Donald Knuth, espanta-se por quê demorou tanto para substituir LaTeX por outra linguagem de marcação leve que se compila a TeX).
- Em particular, a sintaxe Markdown é muito mais simples do que a sintaxe de LaTeX Beamer.
- Fórmulas matemáticas são mais facilmente escritas em Markdown que em Microsoft Word ou LibreOffice,
- É especialmente apta para artigos curtos em HTML, como postagens de blog.

Pandoc versus Programas Especializados: O mal.

- Funções específicas à linguagem de marcação, em pandoc,
 - não podem ser usadas, ou
 - só podem ser usadas invalidando o output em outras linguagens. (\implies a sintaxe de markdown é tão reduzida quanto a base comum entre as linguagens de marcação.)

...

- Pandoc está ainda sendo desenvolvido, por isso:
 - a documentação é incompleta,
 - o ecossistema de programas para pandoc, tais como editores, é ainda incompleto: Por exemplo, comparando a TeX:

- * pode-se pular da posição do cursor no arquivo TeX à posição correspondente no arquivo pdf compilado, e vice-versa. Não há para pandoc: markdown compila a TeX e depois a pdf.
- * O plugin de Vim para markdown é recente e básico em comparação àquele para LaTeX que é estável e potente.

sintaxe markdown: documentação = cola

Cola

Syntax Summary - Markdown 1.0.1 and PHP Markdown Extra 1.0.1 All table text highlights features added by PHP Markdown Extra

SPAN ELEMENTS

Links

Inline style
This is [an example](http://example.com/ "Optional Title") inline link.

Reference style
Link definition:
[id]: http://example.com/ "Optional Title Here"
Link:
This is [an example][id] reference-style link.

Automatic links
~http://example.com/~
<address@example.com>

Images

Inline style
![Alt text](/path/img.jpg "Optional title")

Reference style
Image definition:
[id]: url/to/image "Optional title"
Link:
![Alt text][id]

Emphasis

** tag**
asterisk _underscore_

** tag**
asterisk __underscore__
No underscores in the middle of words

Inline code
Enclose within backticks
Use the `htmlspecialchars()` function.

BLOCK ELEMENTS

Paragraph
One or more consecutive lines of text, separated by one or more blank lines.

**
 line break**
End a line in a paragraph with two or more spaces.

Code blocks
Indent 4 spaces or 1 tab:
This is a normal paragraph:

```
This is a code block.
```

Inline HTML
1. Indent opening tag by no more than three spaces.
2. Inside a list, indent to the same level as the list.

Horizontal rules
Three or more asterisks, dashes or underscores on their own line:
*** OR --- OR ___

Markdown inside HTML blocks
`<div markdown="1">`
This is "true" markdown text.
`</div>`

Headers

This is an H1 This is an H1
This is an H2 This is an H2
This is an H4 <small>-----</small>

Header id attribute
Header 2 ## [#header2]
OR:
Header 1 [#header1]
=====

Link back to header
[Link back to header 1][#header1]

Tables

First Header	Second Header
Content Cell	Content Cell
Content Cell	Content Cell

OR

First Header	Second Header
Content Cell	Content Cell
Content Cell	Content Cell

Text alignment within columns
NB: Uses the HTML ALIGN attribute:

Item	Value
Computer	\$1600

Left aligned

Item	Value
Computer	\$1600

Centre aligned

Item	Value
Computer	\$1600

Right aligned

Item	Value
Computer	\$1600

Blockquotes
> Lorem ipsum dolor sit amet, consectetur adipiscing elit.
> Donec sit amet nisl. Aliquam semper ipsum sit amet velit.

Lists

Unordered lists
Asterisks, plus signs or dashes:
* Red + Red - Red
* Green + Green - Green
* Blue + Blue - Blue

Ordered lists
1. Bird
2. McHale
3. Parish

Definition lists
Term Definition
* Definition

Multiple definitions
Apple
*Pomaceous fruit of plants of the genus Malus in the family Rosaceae.
* An american computer company.

Multiple terms
Term 1 Definition
Term 2 Definition

LITERAL CHARACTERS
Use backslash escape to insert the following characters literally:
\\ backslash
* asterisk
_ underscore
{} curly braces
[] square brackets
() parentheses
hash mark
+ plus sign
- minus sign (hyphen)
. dot
! exclamation mark
: colon
| pipe

Markdown: daryn@uball.net/projects/markdown/hyrtex
PHP Markdown Extra: www.michelf.com/projects/php-markdown/extra/
THIS .DOC: www.php.net/manual/2006/01/markdown.html#head-1head

Figura 0.1: Documentação (= Cola) da sintaxe de Markdown

Itens alistados enfatizados:

- *peso*
- *altura*

Itens enumerados negritos:

1. **mãe**
0. **pai**

Tabela

```
|          | mãe   | pai   |
|-----|-----|-----|
| peso    | 100 kg | 200 kg |
| altura  | 1,20 m | 2,10 m |
```

~>

Itens alistados enfatizados:

- *peso*
- *altura*

Itens enumerados negritos:

1. **mãe**
2. **pai**

Tabela

	mãe	pai
peso	100 kg	200 kg
altura	1,20 m	2,10 m

compilar pandoc: Makefile e Mainfile

Parâmetros Pandoc

General options:

--from = FORMAT

Specify input FORMAT such as markdown, rst, ..

--to = FORMAT

Specify output FORMAT such as html, LaTeX, ..

--output = FILE

Write output to FILE instead of stdout.

General writer options:

--standalone

Produce output with header and footer.

--table-of-contents

Include a generated table of contents in output.

--self-contained

Produce standalone HTML file without external dep.

Makefile

O Makefile serve para:

- Compilar o input: por exemplo, o comando
 - make pdf gera o documento pdf,
 - make docx gera o documento docx,
 - make html gera o documento HTML.

...

- Olhar o output: por exemplo, o comando
 - make run_pdf mostra o documento pdf, e.g. em Zathura,
 - make run_docx mostra o documento docx, e.g. em LibreOffice,
 - make run_html mostra o documento HTML, e.g. em Firefox.

```
NAME = palestra
```

```
FILES = $(NAME).pandoc intro.md conteudo.md resumo.md
```

```
PANDOC_OPTIONS = --stand-alone --filter pandoc-citeproc
```

```
PANDOC_LATEX_OPTIONS = --include-in-header header.tex
```

```
all: latex pdf
```

```
latex: $(FILES)
```

```
    pandoc $(PANDOC_OPTIONS) $(PANDOC_LATEX_OPTIONS) \  
        --from markdown --to latex \  
        $(NAME).pandoc $(FILES) --output $(NAME).tex
```

```
pdf: latex
```

```
    latexrun $(NAME).tex
```

```
run: run_pdf
```

```
run_pdf: pdf
```

```
    $(PDFVIEWER) $(NAME).pdf >/dev/null /2> &1 &
```

Main file

O main file, o arquivo principal, estabelece

- o título,
- o autor, e
- a data.

...

Além disso,

- o idioma, pela opção lang que controla por exemplo o nome do resumo e das referências, e

- varias opções de TeX, tal como:
 - o tipo de documento, e
 - tamanho da fonte.

Exemplo

```
% Pandoc é bem massa!  
% [Enno Nagel](mailto:epn@fsfe.org)  
% CoALTI, Hotel Premier Maceió, 28 Outubro 2017  
  
---  
lang:                pt  
  
# latex:  
documentclass:      scrartcl  
classoption:        final,DIV = calc,bibliography = toctoc  
fontsize:            12pt  
bibliography:        pandoc.bib  
...
```

Configurando o editor Vim

Compilação automática

Para o editor Vim compilar após guardar o arquivo, acrescente ao arquivo `~/.vim/after/ftplugin/markdown.vim` a linha:

```
autocmd BufWrite <buffer> silent make!
```

Por exemplo, se o output é

- pdf (via TeX), então o amostrador zathura automaticamente recarrega o arquivo modificado,
- html, então a extensão autoreload do navegador Firefox automaticamente recarrega o arquivo modificado.

Site e Repositório

Estão disponíveis

- os eslaides,
- um resumo e
- a cola

em <https://konfekt.bitbucket.io/talks/pandoc>,

e

- o Makefile e

- o código fonte (simples!) para compilar estes eslaides
no repositório <https://www.bitbucket.org/konfekt/talks/pandoc>