

# OpenSource.com: Pandoc Cheat Sheet

BY JASON VAN GUMSTER

Pandoc is the “Swiss Army Knife” of document conversion. With it, you can convert files from one document format to a whole suite of possible other formats. This cheat sheet covers some of the most common cases.

## BASIC CONVERSIONS

reStructuredText to Markdown

```
pandoc -s -f rst FILE.rst -o FILE.markdown
```

Markdown to DOCX

```
pandoc FILE.markdown -o FILE.docx
```

Multiple Markdown files to EPUB

```
pandoc -s -o FILE.epub FILE01.markdown FILE02.markdown FILE03.markdown
```

ODT to RTF

```
pandoc -s FILE.odt -o FILE.rtf
```

EPUB to plain text

```
pandoc FILE.epub -t plain -o FILE.txt
```

DocBook to HTML5

```
pandoc -s -f docbook -t html5 FILE.xml -o FILE.html
```

MediaWiki to DocBook 4

```
pandoc -s -f mediawiki -t docbook4 FILE.wiki -o FILE.xml
```

LaTeX to PDF

```
pandoc FILE.tex --pdf-engine=xelatex -o FILE.pdf
```

Website to Markdown

```
pandoc -s -f html https://opensource.com -o FILE.markdown
```

## KEY COMMAND LINE OPTIONS

**--standalone (-s)**: In most formats, Pandoc generates a document fragment, rather than a self-contained single document. Use this flag to ensure appropriate headers and footers are included.

**--from=FORMAT (-f)**: Usually Pandoc can infer the file format from context or its file extension. Use this flag to remove any guesswork. See the section below for the formats Pandoc supports.

**--to=FORMAT (-t)**: Just like with the **-f** flag, this option allows you to explicitly specify the output format coming from Pandoc.

**--output=FILE (-o)**: If you want your output to go to a file instead of standard out, make sure to include this option

**--template=FILE**: You can specify a template file for your output document using this flag if you want to control the look and appearance of the converted file.

**--toc**: Enable this option to automatically generate a table of contents in your output document.

**--highlight-style=STYLE|FILE**: If your converted output incorporates code that should be syntax-highlighted, use this option to use a predefined style (e.g. **pygments**, **breezeDark**, **espresso**, **haddock**, **kate**, **monochrome**, **tango**, and **zenburn** – the default is **pygments**) or a style theme that you define in a particular file.

## OPTIONS FOR SPECIFIC OUTPUT FORMATS

**--self-contained**: Add this option if you’re generating an HTML document or HTML-based slides and you want to have no external file dependencies.

**--number-sections (-N)**: If you’re working on a document (like an academic paper) that requires numbered sections, make use of this flag

**--css=URL (-c)**: This option allows you to link to a specific CSS file for styling your output document. Pandoc tries to use sensible defaults, but if you want to give your EPUB or HTML output a custom look, this is the way to go.

**--epub-cover-image=FILE**: Use this flag to specify a cover image for your EPUB book. If your input format is Markdown, you can define this in a metadata block instead of using the command line option.

**--epub-metadata=FILE**: If you don’t have metadata specified in your input document, you can use this flag to let Pandoc know of a file where that metadata is located.

**--pdf-engine**: Use this option to stipulate which backend software you’d like to use to generate your output PDF. The default option is **pdflatex**, but other options include **context**, **lualatex**, **pdfroff**, **prince**, **weasyprint**, **wkhtmltopdf**, and **xelatex**, assuming you have those backends installed.

**--mathjax**: Pandoc defaults to using pretty simple styling for mathematical equations. Enable this option to make use of MathJax Javascript to render your equations and formulas.

## INPUT/OUTPUT FORMAT OPTIONS

### INPUT FORMATS SUPPORTED BY PANDOC:

- **commonmark** (CommonMark Markdown)
- **creole** (Creole 1.0)
- **docbook** (DocBook)
- **docx** (Microsoft Word .docx)
- **epub** (EPUB)
- **gfm** (GitHub-flavored Markdown)
- **haddock** (Haddock markup)
- **html** (HTML)
- **json** (JSON version of native AST)
- **latex** (LaTeX)
- **markdown** (Pandoc’s extended Markdown)
- **markdown\_mmd** (MultiMarkdown)
- **markdown\_phpextra** (PHP Markdown Extra)
- **markdown\_strict** (original unextended Markdown)
- **mediawiki** (MediaWiki markup)
- **native** (native Haskell)
- **odt** (LibreOffice/OpenOffice text document)
- **opml** (OPML)
- **org** (Emacs Org mode)
- **rst** (reStructuredText)
- **t2t** (txt2tags)
- **textile** (Textile)
- **tikiwiki** (TikiWiki markup)
- **twiki** (TWiki markup)

### ALL OF THE ABOVE FORMATS ARE AVAILABLE FOR OUTPUT, PLUS THE FOLLOWING:

- **asciidoc** (AsciiDoc)
- **beamer** (LaTeX beamer slide show)
- **context** (ConTeXt)
- **docbook** or **docbook4** (DocBook 4)
- **docbook5** (DocBook 5)
- **dokuwiki** (DokuWiki markup)
- **dzslides** (DZSlides HTML5 and Javascript slide show)
- **epub2** (EPUB v2 ebook)
- **epub** or **epub3** (EPUB v3 ebook)
- **fb2** (FictionBook2 ebook)
- **html4** (XHTML 1.0 Transitional)
- **html** or **html5** (HTML5/XHTML polyglot markup)
- **icml** (InDesign ICML)
- **jats** (JATS XML)
- **man** (groff man page)
- **opendocument** (OpenDocument)
- **plain** (plain text)
- **pptx** (PowerPoint slide show)
- **revealjs** (reveal.js HTML5 and Javascript slide show)
- **rtf** (rich text format)
- **s5** (S5 HTML and Javascript slide show)
- **slideous** (Slideous HTML and Javascript slide show)
- **slidy** (Slidy HTML and Javascript slide show)
- **tei** (TEI Simple)
- **texinfo** (GNU Texinfo)
- **zimwiki** (ZimWiki markup)

Detailed documentation is available on the Pandoc website: <http://pandoc.org/MANUAL.html>