Pandoc

First Notes

These steps tested on an Debian Linux (e.g. Ubuntu) environment. Steps may be similar for windows but the installation may differ.

Installing pandoc

The source files or installation program can be found at https://pandoc.org/

for Linux, there is a `.deb` file available to download, which can be opened directly with `QApt` or your local `.deb` package manager.

To test the installation worked, you can try running pandoc. This is largely a CLI (command line interface) program, so you'll need to use your terminal or command line.

For your terminal or commandline, test the install with:

```
pandoc --version
```

You should get something similar to the return below.

```
>>>pandoc 2.2.3.2
Compiled with pandoc-types 1.17.5.1, texmath 0.11.0.1, skylighting 0.7.2
Default user data directory: /home/osboxes/.pandoc
Copyright (C) 2006-2018 John MacFarlane
Web:  http://pandoc.org
This is free software; see the source for copying conditions.
There is no warranty, not even for merchantability or fitness
for a particular purpose.
```

Converting to `.docx`

first, you'll need to `cd` to your latex directory (or `dir` for windows). once there, you can instruct pandoc to convert your document.

The format of the command is:

```
pandoc [source_file] -f [source_format] -t [target_format] -s -o [target_file]
```

`source_file`: the name of your existing file

`-f [source_format]`: **from** - what format we're converting from.

`-t[target_format]`: **to** - what format we're converting to. the list of supported formats can be found in the manual here

`-s`: **standalone**- this flag states that you want the output document to be a single complete file, rather than a fragment - usually recommended, but not always required.

`-o [target_file]`: **output**: states that you want to output into a file, and what it's called.

In our case we want to convert from LaTeX to .docx, so we would run:

```
pandoc document.tex -f latex -t docx -s -o document.docx
```

As ever, format conversion is an inexact science, especially with figures and images, so mileage may vary. As LaTeX structuring is explicit, it is common to find more issues when converting to LaTeX than from.

Gotchas

There are a few unintuitive details about how pandoc works:

1. Because windows has different text encoding (how characters are represented in memory) to how Linux and MacOS do, you'll need to switch the encoding explicitly to `UTF-8` in your terminal before working with pandoc in windows. if you are using cmd then run:

   ```
   chcp 65001
   ```

   If you are using powershell, you shouldn't need to run this step.

2. some formats require additional metadata like header and body blocks - such as pdfs and html. recent versions of pandoc should automatically default to correcting for this but you may need to include `-s` in your pandoc command to ensure the document can be used alone. The only reason that you wouldn't want these is if you are embedding the documents within another website or document, which is usually if you're creating a document-serving application. You probably don't want to do this.