The `tex4ebook` package

Michal Hoftich*

Version v0.4
2024-02-23

Contents

---

*michal.h21@gmail.com

1 Introduction

`TeX4ebook` is a tool for conversion from LaTeX to ebook formats, such as EPUB, MOBI and EPUB 3. It is based on `TeX4ht`[1], which provides instructions for the actual LaTeX to HTML conversion, and on `make4ht`[2].

The conversion is focused on the logical structure of the converted document and metadata. Basic visual appearance is preserved as well, but you should use custom configurations if you want to make the document more visually appealing. You can include custom CSS or fonts in a configuration file.

`TeX4ebook` supports the same features as `make4ht`, in particular build files and extensions. These may be used for post-processing of the generated HTML files, or to configure the image conversion. See the `make4ht` documentation to see the supported features.

1.1 License

Permission is granted to copy, distribute and/or modify this software under the terms of the LaTeX Project Public License, version 1.3.

2 Usage

Run on the command line:

```
tex4ebook [options] filename
```

You don't have to modify your source files unless you want to use commands defined by `tex4ebook` in the document, or when your document uses a package which causes a compilation error.

If you want to use `tex4ebook` commands, add this line to your document preamble:

```
\usepackage{tex4ebook}
```

But it is optional. You shouldn't need to modify your TeX files

2.1 Available commands

- `\coverimage[<graphicx options>]{coverimage.name}` – include cover image to the document. You can pass the same options as to `\includegraphics` command in the optional argument.

For example:

---

[1] `https://tug.org/tex4ht/`
[2] `https://ctan.org/pkg/make4ht?lang=en`

```
\thispagestyle{empty}
\begin{document}
\coverimage[scale=0.8]{coverimage.name} % include scaled cover image
...
\pagestyle{headings}
```

- \epubpage – add page break to the document.

3   Command line options

-a,--loglevel Set message log level. Possible values: debug, info, status, warning, error, fatal. Default: status.

-c,--config specify custom config file for TeX4ht

example config file: File `sample.cfg`

```
\Preamble{xhtml}
\CutAt{section}
\begin{document}
\EndPreamble
```

run

```
tex4ebook -c sample filename.tex
```

This config file will create `xhtml` file for every section. Note that this behaviour is default.

-e,--build-file (default nil) Specify make4ht build file[3]. Default build file filename is `filename.mk4`, use this option if you use different filename.

-f,--format (default epub) Output format. Possible values are `epub`, `epub3`, `mobi`, `azw` and `azw3`.

-j,--jobname Specify the output file name, without file extension.

-l,--lua Use LuaLaTeX as TeX engine.

-m,--mode (default default) This set `mode` variable, accessible in the build file. Default supported values are `default` and `draft`. In `draft` mode, document is compiled only once, instead of three times.

-s,--shell-escape Enable shell escape in the `htlatex` run. This is necessary for the execution of the external commands from your source files.

---

[3]`https://github.com/michal-h21/make4ht#build-file`

`-t,--tidy` clean-up output HTML files using the `HTML tidy` command[4].

`-x,--xetex` Use xelatex for document compilation

`-v,--version` Print the version number.

4   Configuration

`TeX4ebook` uses `TeX4ht`[5] for conversion from LaTeX to html. `TeX4ht` is highly configurable using config files. Basic config file structure is

```
\Preamble{xhtml, comma separated list of options}
...
\begin{document}
...
\EndPreamble
```

Basic info about command configurations can be found in a work-in-progres *TeX4ht tutorial*[6], *TeX4ht documentation*[7], and in series of blogposts on CV Radhakrishnan's blog: *Configure part 1*[8], *Configure part 2*[9], *Low level commands*[10]. Available options for `\Preamble` command are listed in the article *TeX4ht: options*[11]. *Comparison of tex4ebook and Pandoc output*[12]

A great source of tips for `TeX4ht` configuration is *tex4ht tag on TeX.sx*[13]. There is also a *tag for tex4ebook*[14].

Examples of interesting questions are *including images and fonts in ebooks*[15] or *setting image size in em units instead of pt*[16].

4.1   Provided configurations

`tex4ebook` provides some configurations for your usage:

```
\Configure{UniqueIdentifier}{identifier}
```

Every EPUB file should have unique identifier, like ISBN, DOI, URI etc. Default identifier is URI, with value `http://example.com/\jobname`.

---

[4]if it is available on your system
[5]http://www.tug.org/tex4ht/
[6]https://github.com/michal-h21/helpers4ht/wiki/tex4ht-tutorial
[7]http://www.tug.org/applications/tex4ht/mn11.html
[8]https://web.archive.org/web/20180908234227/http://www.cvr.cc/?p=323
[9]https://web.archive.org/web/20180908201057/http://www.cvr.cc/?p=362
[10]https://web.archive.org/web/20180909101325/http://cvr.cc/?p=482
[11]https://web.archive.org/web/20180813043722/http://cvr.cc/?p=504
[12]https://github.com/richelbilderbeek/travis_tex_to_epub_example_1
[13]http://tex.stackexchange.com/questions/tagged/tex4ht
[14]http://tex.stackexchange.com/questions/tagged/tex4ebook
[15]http://tex.stackexchange.com/a/213165/2891
[16]http://tex.stackexchange.com/a/195718/2891

```
\Configure{@author}{\let\footnote\@gobble}
```

Local definitions of commands used in the \author command. As contents of \author are used in XML files, it is necessary to strip away any information which don't belongs here, such as \footnote.

```
\Configure{OpfScheme}{URI}
```

Type of unique identifier, default type is URI. It is used only in the EPUB format, it is deprecated for EPUB 3.

```
\Configure{resettoclevels}{list of section types in descending order}
```

Configure section types which should be included in the NCX file. Default value is the whole document hierarchy, from \part to \paragraph.

```
\Configure{DocumentLanguage}{language code}
```

Each EPUB file must declare the document language. It is inferred from babel main language by default, but you can configure it when it doesn't work correctly. The language code should be in ISO 639-1 form.

```
\Configure{CoverImage}{before cover image}{after cover image}
```

By default, cover image is inserted in <div class="cover-image"> element, you may use this configuration option to insert different markup, or even to place the cover image to standalone page.

```
\Configure{CoverMimeType}{mime type of cover image}
```

Default value is image/png, change this value if you use other image type than png.

If you don't want to include the cover image in the document, use command

```
\CoverMetadata{filename}
```

in the config file.

```
\Configure{OpfMetadata}{item element}
```

Add item to <metadata> section in the OPF file.

```
\Configure{OpfManifest}{maifest element}
```

Add item to <manifest> section in the OPF file.

```
\Configure{xmlns}{prefix}{uri}
```

Add XML name space to the generated XHTML files. Useful in EPUB 3.

4.1.1 Example config file

```
\Preamble{xhtml}
\begin{document}
\Configure{DocumentLanguage}{de}
% Use following lines if your document has ISBN:
% \Configure{OpfScheme}{ISBN}
% \Configure{UniqueIdentifier}{3-0000-1111-X}
% Another possibility is URI that points for example to the ebook homepage:
\Configure{OpfScheme}{URI}
\Configure{UniqueIdentifier}
{https://de.wikipedia.org/wiki/Der_achte_Sch&ouml;pfungstag}
\Configure{CoverMimeType}{image/jpeg}
% If you don't use \coverimage in the document text,
% add cover image using this command:
\CoverMetadata{coverimage.jpg}
% You can also add more authors to your ebook metadata:
\Configure{OpfMetadata}
{\HCode{<dc:publisher>Deutscher Bücherbund</dc:publisher>}}
\Configure{OpfMetadata}
{\HCode{<dc:contributor>Image Artist</dc:contributor>}}
\Configure{OpfMetadata}
{\HCode{<dc:contributor>Trans Lator</dc:contributor>}}
\Configure{OpfMetadata}
{\HCode{<dc:date opf:event='original-publication'>1888</dc:date>}}
\EndPreamble
```

Remarks:

- Leading percent signs in the `.cfg` file introduce comments

- If the unique identifier is a URI which contains diacritical characters, the
  equivalent HTML code needs to be inserted. `UTF8` is not recognized at that place.

- `UTF8` characters may be used in the `OpfMetadata` sections.

4.2 TeX4ht options

TeX4ht supports lot of options, that change produced HTML code without need to use configurations. Their list is available in the TeX4ht documentation. You can pass options to `tex4ebook` in the argument that follows filename:

```
tex4ebook filename.tex "option1,option2"
```

Alternatively, they can be put in the `\Preamble` command in the config file:

```
\Preamble{xhtml,option1,option2}
```

4.2.1 Options provided by `tex4ebook`

`no-cut` By default `tex4ebook` splits document to separate HTML pages
on `\chapter` command when it is available. Othervise, it splits on
`\section`. This can be changed using the `\CutAt` command or numeric
options, but you need to use the `no-cut` option to prevent fatal error.

4.3 Commands available in the config file

`\OpfRegisterFile[filename]` register file in the `OPF` file. Current output
file is added by default.

`\OpfAddProperty{property type}` add `EPUB3` property for the current file.
See *EPUB3 spec*[17]

`\OpfGuide[filename]{title}{type}` Add file to the `<guide>` section in the
`OPF` file. See *Where do you start an ePUB and what is the `<guide>` section
of the `.OPF` file?*[18] for some details. Note that `<guide>` is deprecated in
`EPUB 3`.

4.4 Build files

`tex4ebook` uses `make4ht`[19] as a build system. It provides support for build
files written in Lua. These build files can be used to call additional commands,
like BibTeX or Makeindex, post-process generated HTML files, change the
way how images are created, or to modify parameters of the conversion.

   Sample build file can look like this:

```
if mode=="draft" then
  Make:htlatex {}
else
  Make:htlatex {}
  Make:htlatex {}
  Make:htlatex {}
end

Make:image("png$",
"dvipng -bg Transparent -T tight -o ${output} -D 170  -pp ${page} ${source}")
```

---

[17]http://www.idpf.org/epub/301/spec/epub-publications.html#
sec-item-property-values
[18]http://epubsecrets.com/where-do-you-start-an-epub-and-what-is-the-guide-section-of-the-opf-file.
php
[19]https://github.com/michal-h21/make4ht

The `mode` variable holds value of the `--mode` argument to `tex4ebook`. The `draft` mode is used for faster compilation, it calls LaTeX only once.

The `Make:image` function can configure handling of images created by extraction from the DVI file. It can be complex math, TikZ or PSTricks pictures, and so on. The `${<name>}` placeholders are filled by `tex4ebook` with parameters like current page number of the image DVI file, or output image name.

You can compile your document with a build file using the `-e` option:

```
tex4ebook -m draft -e build.lua filename.tex
```

See `make4ht` documentation for more details on configuration files.

4.5 `.tex4ebook` **configuration file**

`tex4ebook` supports a default build file, which is loaded automatically without need to use the `-e` option.

4.5.1 Location

The configuration file can be saved either in `$HOME/.config/tex4ebook/config.lua` or in `.tex4ebook` in the current directory or it's parents (up to `$HOME`).

See the `make4ht` documentation for an example and more information.

5 Troubleshooting

5.1 Kindle formats

`tex4ebook` uses `kindlegen` command for the conversion to Kindle formats (`mobi`, `azw` and `azw3`). Unfortunatelly, Amazon discontinued this command, so we use also `ebook-convert` provided by Calibre if `kindlegen` fails.

5.2 Fixed layout EPUB

The basic support for the Fixed layout EPUB 3 can be enabled using the following configurations:

```
\Configure{OpfMetadata}
{\HCode{<meta property="rendition:layout">pre-paginated</meta>}}
\Configure{OpfMetadata}
{\HCode{<meta property="rendition:orientation">landscape</meta>}}
\Configure{OpfMetadata}
{\HCode{<meta property="rendition:spread">none</meta>}}
\Configure{@HEAD}
{\HCode{<meta name="viewport" content="width=1920, height=1080"/>\Hnewline}}
```

Modify the dimensions in the `<meta name="viewport">` element according to your needs.

5.3 Math issues

Note that while `mobi` is supported by Amazon Kindle, most widespread ebook reader, it doesn't support `MathML`. This means that math must be represented as images. The same issue is true for the EPUB format as well. This is problematic especially for the inline math, as you may experience wrong vertical alignment of the math content and surrounding text. If your ebook contains math, a better solution is to produce the `epub3` format, as it supports `MathML`. The issue with EPUB 3 is that majority of `e-ink` ebook readers don't support it. Reader applications exists mainly for Android and Apple devices. For books which contains mainly prose, all formats should be suitable, but EPUB 3 supports most features from web standards, such as `CSS`.

5.4 Compilation errors

When compilation of the document breaks with error during `LaTeX` run, it may be caused by some problem in `TeX4ht` configuration. Comment out line `\usepackage{tex4ebook}` in your source file and run command:

```
htlatex filename
```

if same error as in `tex4ebook` run arises, the problem is in some `TeX4ht` configuration. Try to identify the source of problem and if you cannot find the solution, make minimal example showing the error and ask for help either on *TeX4ht mailing list*[20] or on *TeX.sx*[21].

5.5 Validation

In case of successful compilation, use command line tool `epubcheck`[22] to check whether your document doesn't contain any errors.

Type

```
epubcheck filename.epub
```

5.5.1 Common validation issues:

- WARNING: filename.epub: item (OEBPS/foo.boo) exists in the zip file, but is not declared in the OPF file

Delete the `filename-(epub|epub3|mobi|azw|azw3)` folder and `filename.epub`. Then run `tex4ebook` again.

---

[20]http://tug.org/mailman/listinfo/tex4ht
[21]http://tex.stackexchange.com/
[22]you need to install it separately, see https://github.com/IDPF/epubcheck

- WARNING(ACC-009): hsmmt10t.epub/OEBPS/hsmmt10tch17.xhtml(235,15): MathML should either have an alt text attribute or annotation-xml child element.

  This is accessibility message. Unless you use some macro with annotations for each math instance, you will get lot of these messages. Try to use `epubcheck -e` to print only serious errors.

6 Changes

- 2024/02/23

    – version `0.4` released.

- 2023/10/30

    – added the `\epubpage` command

- 2023/10/17

    – added the `--build-dir` command line option.

- 2023/10/15

    – don't set PNG as image format explicitly.

- 2023/06/02

    – prevent fatal errors in the `zip` command executable detection.

- 2023/05/29

    – save author and date globally.

- 2023/05/23

    – removed spurious numbers from TOC in Epub 3.

- 2023/03/17

    – released version `0.3j`.
    – fixed bug in generating of TOC in the NCX file.

- 2023/03/02

    – released verision `0.3i`.
    – check if file with TOC exists before processing.

- 2023/01/12

- – added all mimetypes supported in Epub.

- 2023/01/09

  - – fixed handling of metadata for filenames with accented characters.

- 2022/12/15

  - – use lower case file extensions for mimetype matching.

- 2022/12/01

  - – set destinations for `\label` used inside of footnotes.
  - – fixed handling of multiple tables of contents in Epub 3.

- 2022/11/20

  - – added support for the `fn-in` option in the Epub 3 output.

- 2022/03/29

  - – fixed support for TeX filenames that contain dot.
  - – removed spurious comma that was introduced earlier.

- 2022/03/28

  - – documented the `no-cut` option.

- 2022/02/23

  - – extended the documentation about build files.
  - – deprecated the `--resolution` CLI argument.

- 2022/02/22

  - – print `\subsubsection` in the Epub TOC.

- 2022/02/18

  - – released version `0.3h`.

- 2022/01/13

  - – fixed issue where child TOC elements were inserted into `<a>` element.

- 2021/12/07

  - – print space after section number in Epub 3 TOC.

- keep original elements in Epub 3 TOC.

- 2021/12/04

  - fixed support for appendix chapters in Epub 3.

- 2021/11/08

  - released version `0.3g`
  - bug fix: removed spurious `0` character from the NCX file.

- 2021/11/05

  - released version `0.3f`
  - fixed spurious numbers in NCX TOC caused by wrong use of `\cs_if_exist_use:cTF` command.

- 2021/10/08

  - fix for `\author` support in `amsart` class.

- 2021/09/30

  - released version `0.3e`
  - better detection if `kindlegen` was found.

- 2021/09/23

  - use `ebook-convert` for convertsion to Kindle formats if `kindlegen` fails.

- 2021/08/22

  - fixed cross-referencing issue related to unnumbered equations.

- 2021/07/26

  - released version `0.3d`

- 2021/05/29

  - renamed `DeclareLanguage` to `\DeclareLanguageEbook`. Fixes issue 78.

- 2021/05/15

  - use `assert` in checking of existence of the `zip` command.

- replace colons in `OPF` id attributes and add trailing `x` if the `id` starts with number. It should fix some validation issues.

- 2021/05/02

  - added more examples of configuration to the documentation.

- 2020/11/09

  - set exit status

- 2020/11/09

  - don't redefine `PicDisplay` configuration.

- 2020/11/06

  - bug fix: remove custom elements from the NCX file in the Epub 3 format.
  - released version `0.3c`

- 2020/09/07

  - released version `0.3b`

- 2020/08/26

  - fixed hiearchical structure in NCX TOC for chapters in backmatter and appendix
  - load `common_domfilters` extension by default.

- 2020/07/09

  - addded `AZW` and `AZW3` format support.

- 2020/06/21

  - save `\title` element
  - save contents of `\author` in macro directly

- 2020/06/15

  - remove child elements from elements that don't allow them in the OPF and NCX file.

- 2020/03/14

  - explicitly list supported section types in the NCX table

- 2019/11/01

    - released version `0.3a`
    - added `tex4ebook-` prefix to the output formats.
    - removed unused files.

- 2019/11/01

    - released version `0.3`

- 2019/10/20

    - fixed the `TOC` cleanup in the `ePub 3` mode.
    - added support for the `page-spread-left` and `page-spreat-right` properties.

- 2019/10/20

    - addapted to use the `make4ht` logging mechanism.

- 2019/10/06

    - fixed bug with void elements parsing in the OPF file.
    - undo `\XeTeXcharclass` for the `:` character when the OPF file is generated.

- 2019/09/16

    - make the default build sequence before loading of the extensions. Some extensions need to modify the build sequence.

- 2019/08/28

    - added support for reading input from `STDIN`.

- 2019/08/27

    - added support for the `--jobname` command line option.

- 2019/08/25

    - use the `mkparams.get_args` function to retrieve the command line options

- 2019/07/24

    - added support for the `\author` command with an optional argument

- 2019/05/09

  - added support for the `\title` command with an optional argument

- 2019/04/04

  - register appendix chapters and section in the OPF file list
  - fixed handling of appendices in the NCX table of contents

- 2019/03/21

  - released version `0.2c`

- 2019/03/07

  - use `Luatexbase` package in the documentation because of `Microtype` error
  - remove `<guide>` element even in ePub 2
  - added `encoding` attribute to XML declaration in the NCX and OPF files
  - clean the temporary directory (`filename-format`) before file packing

- 2019/01/21

  - added `\Configure{@author}`

- 2019/01/10

  - released version `0.2b`

- 28/11/2018

  - added support for appendix sections to the NCX file

- 27/11/2018

  - use the `uni-html4` option by default. It will convert some math characters as Unicode chars

- 13/11/2018

  - added –xetex option to the README

- 30/10/2018

  - use the original section numbering in TOC in the Epub 3 output. The numbering of the `<ol>` list is disabled by CSS.

- 18/10/2018

  - fixed the executable installation

- 03/09/2018

  - updated the `--help` message

- 30/08/2018

  - removed spurious `\NoFonts` command in the footnote configuration, it caused formating issues in the document following a footnote

- 22/06/2018

  - added support for the output directory selection

- 09/05/2018

  - added support for Polyglossia language codes
  - released version `0.2a`

- 03/05/2018

  - fixed output format handling

- 16/04/2018

  - don't run Git if the Makefile is executed outside of Git repo

- 09/04/2018

  - released version `0.2`

- 06/04/2018

  - documented the configuration file

- 02/03/2018

  - added support for `.tex4ebook` configuration file

- 28/02/2018

  - added media overlays handling

- 19/10/2017

  - added support for XeTeX

- 06/10/2017

    - added support for Make4ht extensions

- 27/04/2017 Version 0.1e

    - pack the accumulated changes for distribution
    - set version number from git tag

- 17/01/2017

    - process duplicate images only once
    - check the OPF table for duplicated id attributes

- 11/01/2017

    - pass `settings` table to the build file.

- 19/12/2016

    - new configuration: `\Configure{resettoclevels}{list of sectioning levels to be included in the NCX}`

- 31/10/2016

    - added tidyconf.conf to the Makefile

- 22/10/2016

    - added support for new Make4ht command `Make:add_file`

- 22/08/2016

    - fixed incorrect `<dc:creator>` generated by `\author`
    - clean ids in the OPF file if they contain invalid characters at the beginning

- 31/07/2016

    - use monospace font for command line options in order to prevent double hyphens to become dashes by ligaturing

- 22/07/2016

    - removed debugging message when `--lua` option is used

- 07/04/2016

    - Cut `\part` commands to standalone pages

- 05/04/2016

  - bug fix: pages which contained math were arranged before other pages

- 31/03/2016 Version 0.1d

  - bug fix: Unicode wasn't used by default
  - bug fix: Formats weren't preserved

- 06/12/2015 Version 0.1c

  - added `--help` and `--version` command line options

- 25/11/2015

  - added missing language codes
  - added `\Configure{DocumentLanguage}` for the cases when document language inferencing doesn't work

- 17/11/2015

  - use mkparams for cli arguments handling.
  - fixed inconsistencies between recent changes in make4ht and tex4ebook

- 24/09/2015 Version 0.1b

  - info about new workaround for `fontspec` package

- 25/08/2015

  - fatal error happened with epub3 when the document didn't contain the TOC

- 23/08/2015

  - simple cleaning of the ncx file if tidy command isn't available

- 05/07/2015 Version 0.1a

  - `tex4ebook` script was missing in the distribution zip file
  - all links moved to footnotes in the documentation

- 29/06/2015 Version 0.1

  - fixes in documentation

- 22/06/2015

  - changes moved from README.md to CHANGELOG.md

- 18/06/2015

  - replaced `--mathml` option with `--mode`. For `mathml` support, use `mathml` option for `tex4ht.sty`.
  - lot of stuff was fixed in `epub3` support.
  - new command `\OpfAddProperty`

- 14/01/2015

  - thanks Volker Gottwald for guide on [installing and using
    tex4ebook](https://d800fotos.wordpress.com/2015/01/19/create-
  e-books-from-latex-tex-files-ebook-aus-latex-tex-dateien-erstellen/)
    on Windows

- 23/11/2014

  - added new command `\OpfGuide`, for adding items to `<guide>`
    section in the `opf` file. This is useful for `epub` and `mobi` formats.

    Usage:

        \OpfGuide[filename]{title}{reference type}

    `filename` is optional, current file name is used when empty. See
    [epub
    secrets](http://epubsecrets.com/where-do-you-start-an-epub-
  and-what-is-the-guide-section-of-the-opf-file.php)
    article

- 20/10/2014

  - fixed issues with starred sections

  - files created with starred sectioning commands (`\chapter*`,
    `\section*`) should be included in correct reading order now

- 16/09/2014

  - new features added
  - new configuration file for `--tidy` option, mathml and html5 elements
    are supported. This means that many validation errors in `mathml`

output can be fixed with `--tidy` option
  - added configuration for all languages supported by `babel`
  - `woff` and `ttf` fonts are supported
  - added inline footlines in `epub3` format
  - added `no-cut` command line option for breaking sections and chapters
    into standalone pages
  - Please support [iniciative for improving mathml
    support](http://www.ulule.com/mathematics-ebooks/) in Gecko and Webkit
    engines. This will hopefully improve also Epub3 readers.

- 10/08/2013

  - [`make4ht`](https://github.com/michal-h21/make4ht) is now standalone
    application which `tex4ebook` depends on. You must [install

it](https://github.com/michal-h21/make4ht#instalation) in order to use
current tex4ebook version.