Publishing eBooks with AsciiDoc

REVISION HISTORY			
NUMBER	DATE	DESCRIPTION	NAME

Contents

It’s easy to write and publish books in EPUB and PDF formats using AsciiDoc.

Here are three examples: The first is a minimal example introducing the AsciiDoc format, the second The Brothers Karamazov is a rather long multi-part book and the third The Adventures of Sherlock Holmes includes a front cover image and customized page styling.

The examples presented below were created on a PC running Xubuntu Linux 10.04 but should be applicable to any UNIX-based operating system.

Note
A number of *asciidoc* and *a2x* features used in this article require newer versions of AsciiDoc — version 8.6.5 or better is recommended. The version of DocBook XSL Stylesheets used was 1.76.1.

Tip
If you are experiencing *xsltproc* errors when you run *a2x* take a look at Francis Shanahan’s Fixing the ePub problem with Docbook-XSL/A2X/Asciidoc blog post.

Chapter 1

Minimal Book

This didactic example contains a title and two chapters. The AsciiDoc source is a plain text file named `minimal-book.txt`:

```
= The Book Title

== The first chapter
Nec vitae mus fringilla eu vel pede sed pellentesque. Nascetur fugiat
nobis. Eu felis id mauris sollicitudin ut. Sem volutpat feugiat.
Ornare convallis urna vitae.

Nec mauris sed aliquam nam mauris dolor lorem imperdiet.

== The second chapter
Ut suspendisse nulla. Auctor felis facilisis. Rutrum vivamus nec
lectus porttitor dui dapibus eu ridiculus tempor sodales et. Sit a
cras. Id tellus cubilia erat.

Quisque nullam et. Blandit dui tempor. Posuere in elit diam egestas
sem vivamus vel ac.
```

To generate the EPUB formatted book file `minimal-book.epub` run AsciiDoc's *a2x* toolchain wrapper command from the command-line:

```
a2x -fepub -dbook minimal-book.txt
```

The `-f` option specifies the output format, the `-d` option specifies the document type (book, article or manpage). The optional `--epubcheck` option tells *a2x* to validate the EPUB file using epubcheck. To generate a PDF version (`minimal-book.pdf`) with *dblatex* run:

```
a2x -fpdf -dbook minimal-book.txt
```

The distributed example PDFs were built using FOP — if you prefer FOP over dblatex use:

```
a2x -fpdf -dbook --fop minimal-book.txt
```

You can also generate an HTML formatted book, either using DocBook XSL Stylesheets:

```
a2x -fxhtml -dbook minimal-book.txt
```

or directly using the *asciidoc* command:

```
asciidoc -dbook minimal-book.txt
```

Note

The a2x toolchain wrapper uses the following programs (most will be available prepackaged for your Linux distribution):

- xsltproc
- DocBook XSL Stylesheets
- dblatex (PDF file generator)
- FOP (alternative PDF file generator)
- epubcheck (optional EPUB file validator)

Chapter 2

The Brothers Karamazov

The Brothers Karamazov is an example of a multi-part book. To generate the AsciiDoc source I downloaded the Project Gutenberg plain text source and edited it manually (this took around 15 minutes). To generate the `brothers-karamazov.epub` EPUB file run this command:

```
a2x brothers-karamazov.txt
```

Table 2.1: Brothers Karamazov source files and eBooks

EPUB	PDF	AsciiDoc source	Project Gutenburg text
brothers-karamazov.epub	brothers-karamazov.pdf	brothers-karamazov.zip	http://www.gutenberg.org/-etext/28054

Here's the start of the AsciiDoc source file showing the AsciiDoc specific meta-data:

```
//
// Text from Project Gutenburg http://www.gutenberg.org/etext/28054
//
// Formatted for AsciiDoc by Stuart Rackham.
//
// a2x default options.
//    a2x: -f epub -d book --epubcheck
// Suppress revision history in dblatex outputs.
//    a2x: --dblatex-opts "-P latex.output.revhistory=0"
// Suppress book part, chapter and section numbering in DocBook XSL outputs.
//    a2x: --xsltproc-opts
//    a2x: "--stringparam part.autolabel 0
//    a2x: --stringparam chapter.autolabel 0
//    a2x: --stringparam section.autolabel.max.depth 0"
//

= The Brothers Karamazov
:author: Fyodor Dostoyevsky
:encoding: iso-8859-1
:plaintext:

...........................................................................
Translated from the Russian of Fyodor Dostoyevsky by Constance Garnett
The Lowell Press New York
```

```
 :
 :

***START OF THE PROJECT GUTENBERG EBOOK THE BROTHERS KARAMAZOV***
.........................................................................


= PART I

== The History Of A Family

=== Fyodor Pavlovitch Karamazov

Alexey Fyodorovitch Karamazov was the third son of Fyodor Pavlovitch
...
```

- The book, book part and chapter titles have been edited to conform to AsciiDoc title conventions.
- Book part titles must be level zero titles:

  ```
  = Book title (level 0)
  = Part one title (level 0)
  == Chapter 1 title (level 1)
     :
  ```

- Sub-chapter levels can be accommodated with ===, ==== and ===== title prefixes.
- In this example the *Project Gutenberg* front matter was put in the untitled preface inside an AsciiDoc literal block.
- The comment lines starting with `// a2x:` contain *a2x* command options for this document (it’s easier to put them in the document once than type them every time on the command-line).
- A number of *xsltproc* options are passed to *a2x* to suppress book part, chapter and section numbering.
- Setting the AsciiDoc *plaintext* attribute suppresses most of AsciiDoc’s text formatting and substitution conventions, this allows large amounts of text to be imported with little or no manual editing.
  - Once you have defined the AsciiDoc *plaintext* attribute the only requisite manual editing will be to format the titles and rectify the odd paragraph being interpreted as some other AsciiDoc block element.
  - In the combined 49 thousand lines of imported *Sherlock Holmes* and *Brothers Karamazov* text I only encountered two misinterpreted list items, neither affected the rendered output or necessitated editing.
  - You can selectively enable and disable the *plaintext* attribute throughout your document using AsciiDoc attribute entries.

Chapter 3

The Adventures of Sherlock Holmes

This book includes a front cover image and a customized page design. To generate the `adventures-of-sherlock-holmes.epub` EPUB file run this command:

```
a2x adventures-of-sherlock-holmes.txt
```

Table 3.1: Sherlock Holmes source files and eBooks

EPUB	PDF	AsciiDoc source	Project Gutenberg text
adventures-of-sherlock-holmes.epub	adventures-of-sherlock-holmes.pdf	adventures-of-sherlock-holmes.zip	http://www.gutenberg.org/-etext/1661

Here's a screenshot of the first page of the first chapter (rendered by the Firefox EPUBReader addon):

images/epub.png

The AsciiDoc source Zip file contains the following files:

`adventures-of-sherlock-holmes.txt`
The AsciiDoc source (derived from the Project Gutenberg plain text source).

`adventures-of-sherlock-holmes.jpg`
The front cover image.

`adventures-of-sherlock-holmes-docinfo.xml`
A docinfo file containing DocBook markup for the front cover image and the Project Gutenberg frontmatter. DocBook XSL Stylesheets identifies the book cover image by the `role="cover"` attribute in the DocBook `mediaobject` element.

`adventures-of-sherlock-holmes.css`
CSS rules for styling the page layout. The design is based on the ePub Zen Garden *Gbs* style. Because this file is not named `docbook-xsl.css` the name must be specified explicitly using the *a2x* `--stylesheet` option.

`underline.png`
A title underline image that is used by the CSS stylesheet. This resource has to be explicitly passed to *a2x* because *a2x* only searches HTML files for resources.

Here's the start of the AsciiDoc source file showing the AsciiDoc specific meta-data:

```
//
// Text from Project Gutenberg http://www.gutenberg.org/etext/1661
//
// Formatted for AsciiDoc by Stuart Rackham.
//
// a2x default options.
//    a2x: -f epub -d book -a docinfo --epubcheck
//    a2x: --stylesheet adventures-of-sherlock-holmes.css
// Suppress revision history in dblatex outputs.
//    a2x: --dblatex-opts "-P latex.output.revhistory=0"
// Suppress book part, chapter and section numbering in DocBook XSL outputs.
//    a2x: --xsltproc-opts
//    a2x: "--stringparam part.autolabel 0
//    a2x: --stringparam chapter.autolabel 0
//    a2x: --stringparam section.autolabel.max.depth 0"
// Additional resources.
//    a2x: --resource underline.png
//

= The Adventures of Sherlock Holmes
:author: Sir Arthur Conan Doyle
:encoding: iso-8859-1
:plaintext:

== A Scandal in Bohemia

To Sherlock Holmes she is always THE woman. I have seldom heard
...
```

The manual editing of imported plain text involves formatting the title and chapter names as AsciiDoc titles (in this example we’ve used the *single line* title format). Sherlock Holmes only has two title levels:

```
= The Adventures of Sherlock Holmes
== A Scandal in Bohemia
== The Red-Headed League
== A Case of Identity
== The Boscombe Valley Mystery
== The Five Orange Pips
== The Man with the Twisted Lip
== The Adventure of the Blue Carbuncle
== The Adventure of the Speckled Band
== The Adventure of the Engineer’s Thumb
== The Adventure of the Noble Bachelor
== The Adventure of the Beryl Coronet
== The Adventure of the Copper Beeches
== Colophon
```

Chapter 4

Styling your books

You can customize the appearance of a document with CSS. CSS file. Either create your own or copy and edit the existing default `docbook-xsl.css` file (located in the AsciiDoc `stylesheets` configuration directory) then place it in the same directory as your AsciiDoc source file. Use the *a2x* `--stylesheet` option if you want to use a different stylesheet file. Take a look at the `adventures-of-sherlock-holmes.css` CSS file.

There are some great examples of EPUB book styles at ePub Zen Garden.

Chapter 5

Including embedded fonts

See this FAQ.

Chapter 6

Reading EPUB documents

My current favorite software epub reader is the Firefox EPUBReader addon. EPUBReader honors the book’s CSS styling rules and renders the page as the author intended (many EPUB readers only process a sub-set of CSS styling rules).

Browsers are a natural fit for EPUB (EPUB is just a bunch of zipped XHTML files) — I’d love to see browsers support EPUB natively.

As of writing this article most eBook readers (with the notable exception of Amazon’s Kindle) support the EPUB format.

Chapter 7

Non-fiction Books and Manuals

AsciiDoc supports a rich set of markup conventions and can generate reference works and technical manuals complete with tables, illustrations, indexes, bibliographies and appendices. All the examples on the AsciiDoc web site can be formatted as EPUB eBooks.