

Creating EPUBs: What's the Best Tool for Me?



STC Summit 2012 - Chicago, IL

21 May 2012 - Scott Prentice, Leximation, Inc.

Introduction

— [Scott Prentice, President of Leximation, Inc.

— [Specializing in FrameMaker plugin development as well as structured FrameMaker conversions, consulting, and development. FrameMaker user/developer since 1991.

— [Developed DITA-FMx, a FrameMaker plugin for efficient DITA authoring and publishing.

— [Developer of custom Help systems and creative/functional web applications.

EPUB basics

- [EPUB is an ebook file format that can be read on almost every device and platform
- [Requires a “reader” application or dedicated reader device
- [A collection of XHTML, XML, CSS, and media files wrapped up in a “zip” archive
- [Best way to learn about the EPUB format is to open your EPUB and review the contents

Structure of an EPUB file

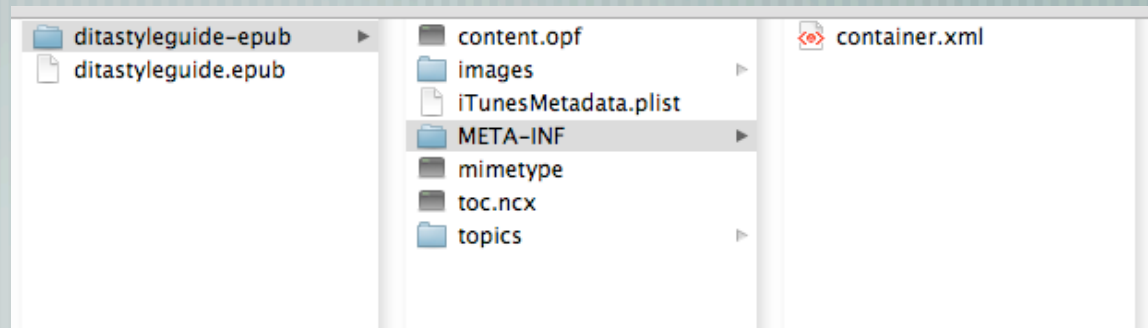
— [mimetype file (at root) contains “application/epub+zip”

— [META-INF/container.xml points to the OPF file

— [OPF file contains metadata, manifest, and spine;
points to an NCX file

— [NCX file is the “TOC”

— [XHTML, CSS, and
media files are referenced by the OPF file



EPUB2 or EPUB3?

— [IDPF (International Digital Publishing Forum) released EPUB3 in October 2011; replaces 2.0.1

— [EPUB3 adds support for HTML5, SVG, MathML, and more

— [Tool/reader support for EPUB3 is lacking

— [For now use EPUB2 unless you know that your target device/application supports EPUB3

Creating an EPUB

- [Create content in an authoring tool then export to EPUB
- [Convert to EPUB from existing content (PDF, HTML, DITA, etc.)
- [People in techcomm will likely convert from existing content
- [Each tool has its methods for authoring/conversion
- [If possible, use named styles since these will typically map to CSS selectors in the XHTML
- [No .. iBooks Author is not an EPUB authoring tool

Popular Authoring Tools

Authoring Tool	Platform	Cost
Adobe InDesign	Windows/Mac	\$700
Adobe RoboHelp	Windows	\$1,000
Adobe Technical Communication Suite	Windows	\$1,900
Apple iWork Pages	Mac	\$70
Atlantis Word Processor	Windows	\$35
eCub	Windows/Mac/Linux/FreeBSD/Solaris	\$0
Jutoh	Windows/Mac/Linux/FreeBSD/Solaris	\$40
MadCap Flare	Windows	\$1000
oXygen XML Editor	Windows/Mac/Linux	\$488
Scrivener	Windows/Mac	\$40/\$45
Sigil	Windows/Mac/Linux	\$0

Authoring Tool + Plugin

Authoring Tool + Plugin	Platform	Cost
Adobe FrameMaker + ElmSoft EPubFm	Windows	\$50
Open Office Word Processor + Writer2ePub	Windows/Mac/Linux	\$0
Open Office Word Processor + Pincette ODF to ePub	Windows/Mac/Linux	\$53

Conversion Tools

Conversion Tool	Platform	Cost
Aspose.Words Express Imports: DOC, DOCX, RTF, HTML, ODT	Windows	\$0
Calibre Imports: CHM, EPUB, HTML, LIT, MOBI, ODT, PDF, RTF, TXT, ...	Windows/Mac/Linux	\$0
DITA Open Toolkit + DITA for Publishers plugin Imports: DITA (XML)	Windows/Mac/Linux/ others (Java)	\$0
DocBook EPUB XSL stylesheets Imports: DocBook (XML)	Windows/Mac/Linux/ others (Java)	\$0
Pincette ODF to ePub Batch Converter Imports: ODT (DOC, DOCX, RTF, STW, WPD, PSW, TXT, HTML, ...)	Windows/Mac/Linux (Java)	\$400
WebWorks ePublisher Imports: FM, MIF, DOC, DOCX, RTF, DITA	Windows	\$800

Selecting a Tool

— [No “perfect” tool for everyone

— [Authoring specifically for EPUB or multiple outputs?

— [Level of formatting required

— [Source file formats and OS requirements

— [Level of expertise

— [Frequency of use

— [Budget

Adobe InDesign

— [Probably the most popular EPUB authoring tool

— [Imports DOC, PDF, XML

— [Numerous EPUB-specific features

— [CS6 supports both the 2.0.1 and 3.0 EPUB formats

— [Great for carefully crafted EPUBs

— [Mac and Windows

Adobe RoboHelp

— [Nice option for single sourcing

— [Imports DOC, DOCX, RTF, FM, PDF, DITA, XML, HTML

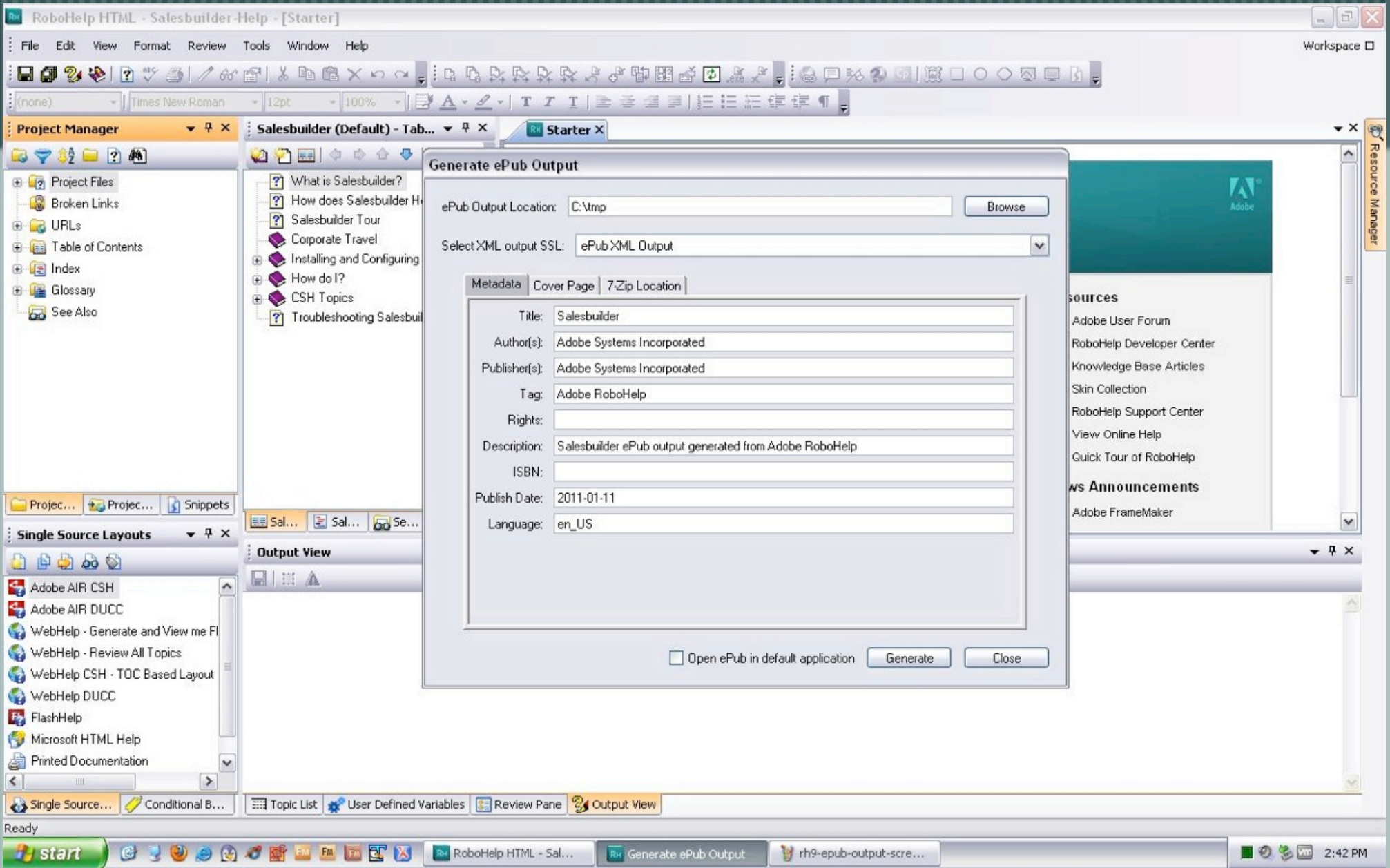
— [Just EPUB 2.0.1 (for now)

— [Allows setting of cover image

— [Limited specification of metadata

— [Can be used for scripted processing

— [Windows only



Adobe TCS

- [Allows authoring in FrameMaker
- [Imports DOC, DOCX, RTF, DITA, XML, ..
- [Publishes EPUB through RoboHelp
- [Just EPUB 2.0.1 (for now)
- [Set cover image in RoboHelp
- [Limited specification of metadata
- [Windows only

Apple Pages

— [One of the first tools for exporting to EPUB

— [Imports RTF

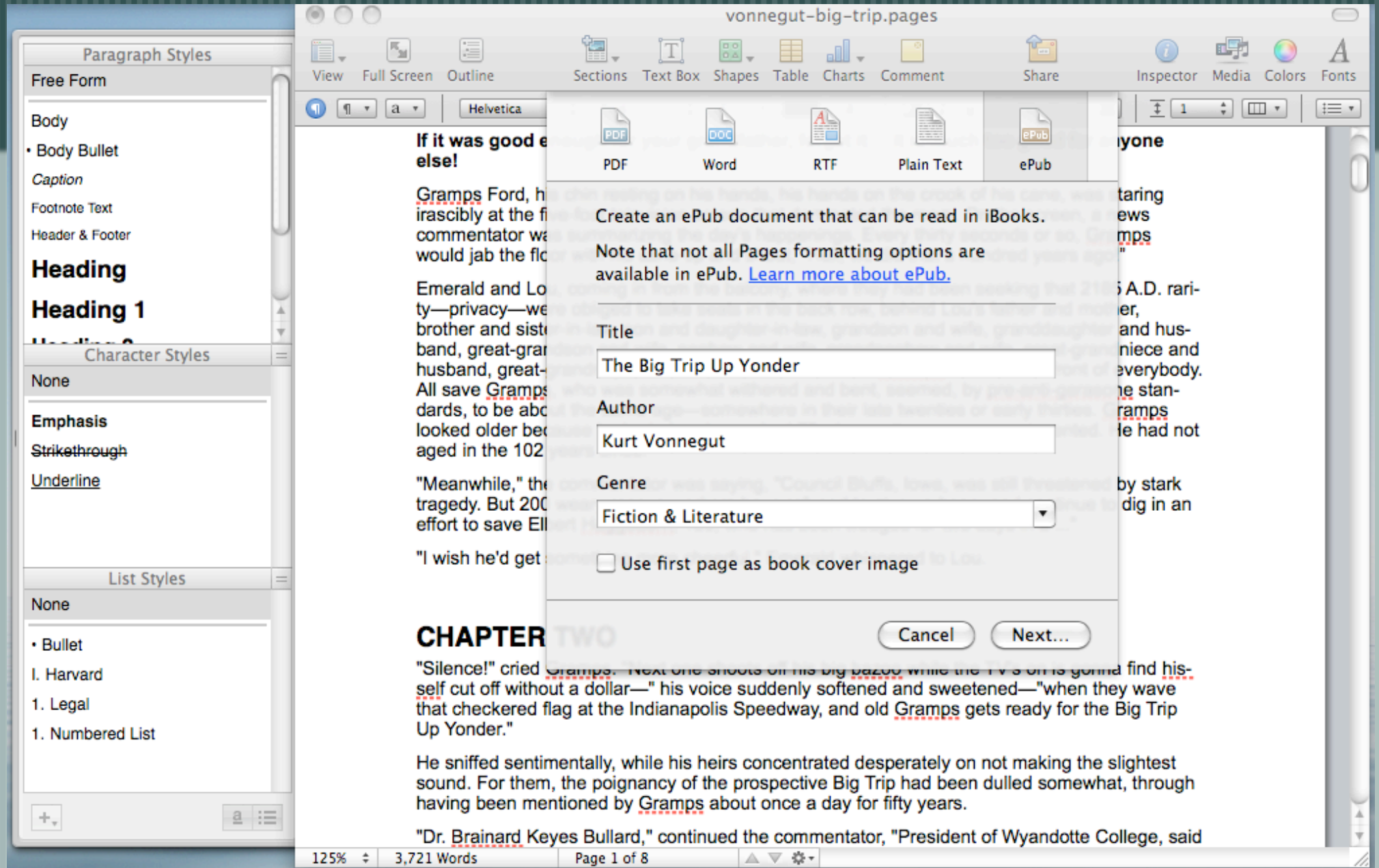
— [Currently just EPUB 2.0.1

— [Uses predefined styles for chapters and TOC entries

— [Option to use first page as cover

— [Limited specification of metadata

— [Mac only



Atlantis Word Processor

- [Low-cost “Word-like” word processor
- [Imports DOC, DOCX, ODT, RTF, COD, TXT
- [Supports EPUB 2.0.1
- [Uses first image as cover page
- [More metadata than some, but limited
- [Can perform batch conversions to EPUB
- [Windows only

eCub

— [Specifically designed for easy EPUB/MOBI creation

— [Imports TXT, HTML, or EPUB

— [Supports EPUB 2.0.1

— [Not WYSIWYG, just text or XHTML

— [Cover designer or from template

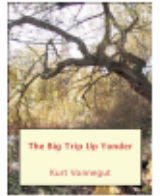
— [Can be used for batch conversions

— [Supports “all” platforms .. and it’s Free!



The Big Trip Up Yonder

Kurt Vonnegut



Files Log

The Project Gutenberg eBook of The Big Trip Up Yonder

CHAPTER ONE
CHAPTER TWO
CHAPTER THREE
CHAPTER FOUR
CHAPTER FIVE
CHAPTER SIX
CHAPTER SEVEN
CHAPTER EIGHT
CHAPTER NINE

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
```

```
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en"><head>
```

```
  <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1">
```

```
  <meta http-equiv="Content-Style-Type" content="text/css">
```

```
  <title>
```

```
    CHAPTER ONE
```

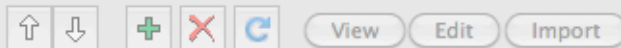
```
  </title>
```

```
  <style type="text/css">
```

```
p {margin-top: .75em; text-align: justify; margin-bottom: .75em;}
```

```
h1 {text-align: center; padding-left: 3em; text-indent: -3em;}
```

```
h2 {text-align: left;}
```



Title: CHAPTER ONE

File: chapter1.htm

Guide type: <none>

Level: 2 Show in: TOC NavMap

Save

Reload

Read

Convert

Jutoh

— [Specifically designed for easy EPUB/MOBI creation

— [Imports ODT, HTML, EPUB, TXT

— [Currently just EPUB 2.0.1

— [Nice WYSIWYG editor

— [Cover page designer with templates

— [Extensive output options

— [Supports “all” platforms!

Organizer

Projects Favourites Explorer

The Big Trip Up Yonder

- Content
 - CHAPTER ONE
 - CHAPTER TWO
 - CHAPTER THREE
 - CHAPTER FOUR
 - CHAPTER FIVE
 - CHAPTER SIX
 - CHAPTER SEVEN
 - CHAPTER EIGHT

CHAPTER THREE

Heading 1 16 Times


CHAPTER THREE

He was instantly offered eleven pens, one from each couple.

"Not *that* leaky thing," he said, brushing Lou's pen aside. "Ah, *there's* a nice one. Good boy, Willy." He accepted Willy's pen. That was the tip they had all been waiting for. Willy, then—Lou's father—was the new favorite.

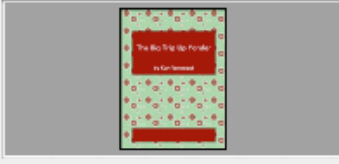
Willy, who looked almost as young as Lou, though he was 142, did a poor job of concealing his pleasure. He glanced shyly at the daybed, which would become his, and from which Lou and Emerald would have to move back into the hall, back to the worst spot of all by the bathroom door.

Gramps missed none of the high drama he had authored and he gave his own familiar role everything he had. Frowning and running his finger along each line, as though he were seeing the will for the first time, he read aloud in a deep portentous monotone, like a bass note on a cathedral organ.



Control Panel

Build Files



Configuration:
Epub

Compile Check Launch

Log

----- Processing configuration Epub -----

Building Epub/Mobi file at Fri 18 May 16:42:26 2012...


Epub compilation successful at Fri 18 May 16:42:31 2012.

Created /Users/leximation/Shared/Data/ePubs/vonnegut-big-trip/chapters/The Big Trip Up Yonder.epub.

Finished building /Users/leximation/Shared/Data/ePubs/vonnegut-big-trip/chapters/The Big Trip Up Yonder.epub at Fri 18 May 16:42:31 2012.

Errors

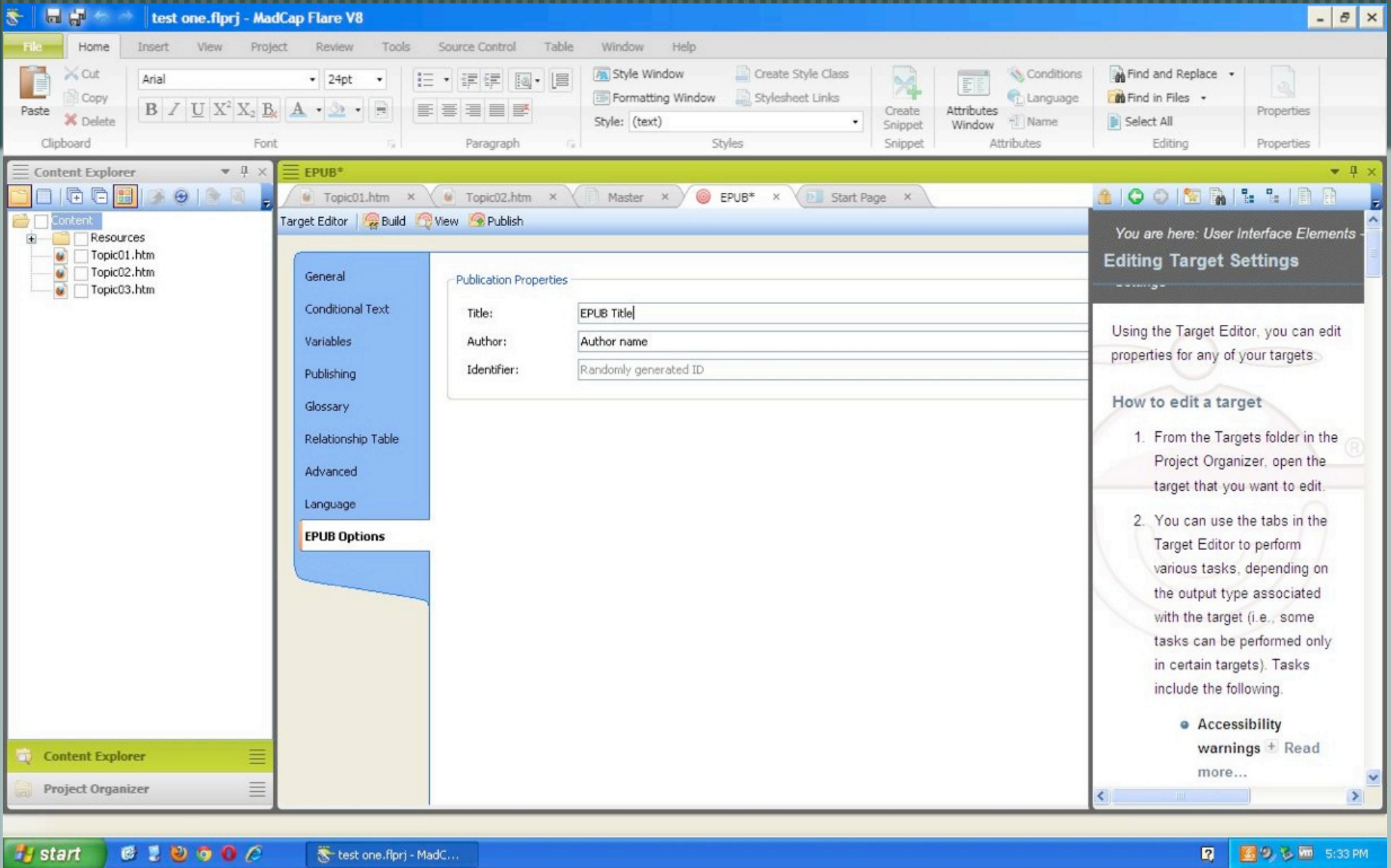
No errors for The Big Trip Up Yonder.epub at Fri 18 May 16:42:31 2012



Launch

MadCap Flare

- [Another option for single sourcing
- [Imports DOC, DOCX, FM, DITA, HTML
- [Currently just EPUB 2.0.1
- [No cover page support, no chapter support
- [Very limited specification of metadata
- [Can be used for scripted processing
- [Windows only



oXygen XML Editor

— [Open and edit EPUB files (great for cleanup task)

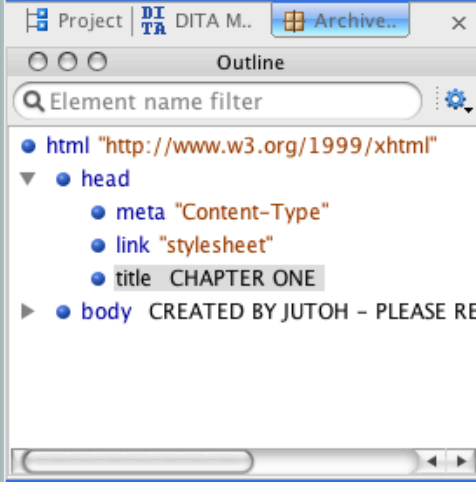
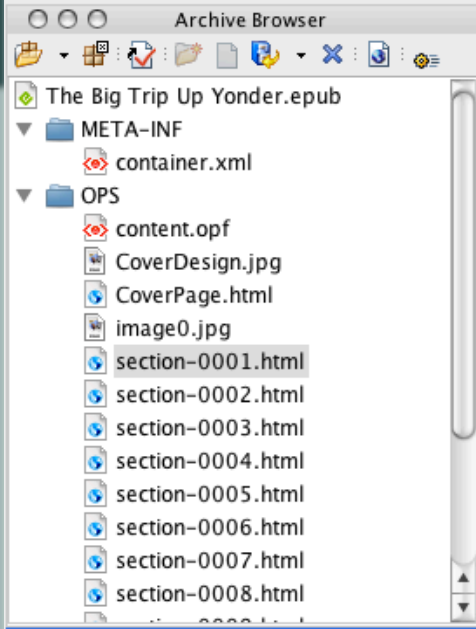
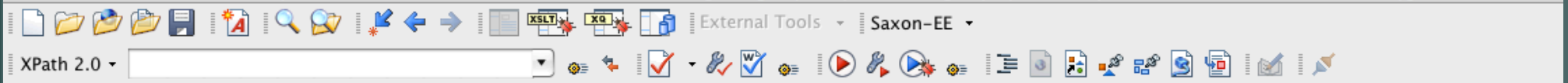
— [Supports EPUB 2.0.1 and 3.0

— [Create EPUB from template (EPUB 2.0.1)

— [WYSIWYG and code views

— [Powerful XML editor features

— [Mac/Windows/Linux



```

1 <?xml version="1.0" encoding="utf-8" ?>
2 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN" "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
3 <html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
4   <head>
5     <meta http-equiv="Content-Type" content="application/xhtml+xml; charset=utf-8" />
6     <link rel="stylesheet" type="text/css" href="styles.css" />
7     <title>CHAPTER ONE</title>
8   </head>
9   <body>
10    <p class="P_Normal">CREATED BY JUTOH - PLEASE REGISTER TO REMOVE THIS LINE</p>
11    <h1 class="P_Heading_1">CHAPTER ONE</h1>
12    <p class="P_Body_Text_First_Indent">Gramps Ford, his chin resting on his hands, his hands on
13      the crook of his cane, was staring irascibly at the five-foot television screen that
14      dominated the room. On the screen, a news commentator was summarizing the day's
15      happenings. Every thirty seconds or so, Gramps would jab the floor with his cane-tip and
16      shout, "Hell, we did that a hundred years ago!"</p>
17    <p class="P_Body_Text_First_Indent">Emerald and Lou, coming in from the balcony, where they
18      had been seeking that 2185 A.D. rarity&#8212;privacy&#8212;were obliged to take seats in
19      the back row, behind Lou's father and mother, brother and sister-in-law, son and
20      daughter-in-law, grandson and wife, granddaughter and husband, great-grandson and wife,
21      nephew and wife, grandnephew and wife, great-grandniece and husband, great-grandnephew
22      and wife&#8212;and, of course, Gramps, who was in front of everybody. All save Gramps,
23      who was somewhat withered and bent, seemed, by pre-anti-gerasone standards, to be about
24      the same age&#8212;somewhere in their late twenties or early thirties. Gramps looked
25      older because he had already reached 70 when anti-gerasone was invented. He had not aged
26      in the 102 years since.</p>
27    <p class="P_Body_Text_First_Indent">"Meanwhile," the commentator was saying, "Council
28      Bluffs, Iowa, was still threatened by stark tragedy. But 200 weary rescue workers have
29      refused to give up hope, and continue to dig in an effort to save Elbert Haggedorn, 183,
30      who has been wedged for two days in a ..."</p>
31    <p class="P_Body_Text_First_Indent">"I wish he'd get something more cheerful," Emerald
32      whispered to Lou.</p>
33
34   </body>
35 </html>

```

Scrivener

- [Interesting “writing studio” interface:

- Corkboard, Outliner, Scrivenings

- [Imports RTFD, DOC, DOCX, ODT, TXT, FDX, FCF, OPML, URL ...

- [Supports EPUB 2.0.1

- [Extensive control over output settings (formatting, transformations, layout, etc.)

- [Mac and Windows

Sigil

— [Open and edit EPUB files

— [Imports TXT or HTML

— [Supports EPUB 2.0.1

— [WYSIWYG and code view authoring

— [TOC builder, HTML tidy, EPUB validator

— [Full metadata support

— [Mac/Windows/Linux .. and Free!

<Select heading>

Table of Contents

The Big Trip Up Yonder
 Table of Contents
 ▼ CHAPTER ONE
 CHAPTER TWO
 CHAPTER THREE
 CHAPTER FOUR
 CHAPTER FIVE
 CHAPTER SIX
 CHAPTER SEVEN
 CHAPTER EIGHT
 CHAPTER NINE

CoverPage.html

chapter1.htm

chapter3.htm

Book Browser

- Text
 - CoverPage.html
 - title.htm
 - TableOfContent...
 - chapter1.htm
 - chapter2.htm
 - chapter3.htm
 - chapter4.htm
 - chapter5.htm
 - chapter6.htm
 - chapter7.htm
 - chapter8.htm
 - chapter9.htm
- Styles
- Images
 - 001.png
 - CoverDesign.jpg
- Fonts
- Misc
 - toc.ncx
 - content.opf

CHAPTER THREE

HE was instantly offered eleven pens, one from each couple.

"Not *that* leaky thing," he said, brushing Lou's pen aside. "Ah, *there's* a nice one. Good boy, Willy." He accepted Willy's pen. That was the tip they had all been waiting for. Willy, then—Lou's father—was the new favorite.

Willy, who looked almost as young as Lou, though he was 142, did a poor job of concealing his pleasure. He glanced shyly at the daybed, which would become his, and from which Lou and Emerald would have to move back into the hall, back to the worst spot of all by the bathroom door.

Gramps missed none of the high drama he had authored and he gave his own familiar role everything he had. Frowning and running his finger along each

```

48 </head>
49
50 <body>
51   <h1>CHAPTER THREE</h1>
52
53   <p class="cap"><span class="dcap">He</span> was instantly offered eleven
54   pens, one from each couple.</p>
55   <p>"Not <i>that</i> leaky thing," he said, brushing Lou's pen aside. "Ah,
56   <i>there's</i> a nice one. Good boy, Willy." He accepted Willy's pen. That
57   was the tip they had all been waiting for. Willy, then&mdash;Lou's
58   father&mdash;was the new favorite.</p>
59   <p>Willy, who looked almost as young as Lou, though he was 142, did a poor
60   job of concealing his pleasure. He glanced shyly at the daybed, which would
61   become his, and from which Lou and Emerald would have to move back into the
62   hall, back to the worst spot of all by the bathroom door.</p>
63   <p>Gramps missed none of the high drama he had authored and he gave his own
64   familiar role everything he had. Frowning and running his finger along each

```

Generate TOC from headings

Elmsoft EPubFm

— [Plugin for FrameMaker 6, 7.x, 8, 9, and 10

— [Exports FM files to EPUB 2.0.1 and MOBI

— [Various options for cover page

— [Nice conversion of FM index

— [Limited metadata support

— [Windows only

The screenshot displays the EpubFm software interface. A dialog box titled "EpubFm Project-C:\Demo\STC-2012\FMx-Auto-Sample\book\test.eprj" is open, showing the "Table of Contents" panel. The dialog box includes buttons for "New Project", "Open Project...", "Close Project", "Save Project", and "Save Project As...". The "Doc Title" is "FMx-Auto User's Guide". The "Source Documents" list includes "toc.fm", "usingautofmforautomation_xml.fm", "usingfmxauto_xml.fm", and "indexlist.fm". The "Table of Contents" panel shows a list of page tags: "=title.0", "==title.1", "===title.2", and "====title.3". The "Toc Level" is set to 4. The background shows the main interface with a menu bar (File, Edit, Add, Element, Format, View, StructureTools, DITA-FMx, Pubs-Tools, FLE, EpubFm, Window, Help), a toolbar, a file explorer showing "fm\title.fm", "toc.fm", "usingfmxauto_xml.fm", "installingfmxauto_xml.fm", "usingautofmforautomation_xml.fm", and "indexlist.fm", and a table of contents for the document. The table of contents includes entries like "authorization", "AutoFM installation", "creating AutoFM scripts", "manual testing", and "PDF, generating via print" and "via saveas". The status bar at the bottom shows "Flow: A ¶: title-index", "1", "11 (1 of 1)", "1", and "100%".

File Edit Add Element Format View StructureTools DITA-FMx Pubs-Tools FLE EpubFm Window Help

fm\title.fm
toc.fm
usingfmxauto_xml.fm
installingfmxauto_xml.fm
usingautofmforautomation_xml.fm
indexlist.fm

EpubFm Project-C:\Demo\STC-2012\FMx-Auto-Sample\book\test.eprj

New Project Open Project... Close Project Save Project Save Project As...

Doc Title FMx-Auto User's Guide

Panel: Selection Cover Meta Data **Toc**

Source Documents

toc.fm
usingautofmforautomation_xml.fm
usingfmxauto_xml.fm
indexlist.fm

Table of Contents

Add Tags from Sel Document... Clear All...

Pgfl Tags

=title.0
==title.1
===title.2
====title.3

<= Toc Level =>

Generate EPub... Close

authorization 36
AutoFM installation 56
creating AutoFM scripts 96
manual testing
PDF, generating
via print
via saveas

Pages: 11

VARIABLES CONDITIONAL TEXT

Flow: A ¶: title-index 1 11 (1 of 1) 1 100%

Writer2ePub

- [Extension for LibreOffice, OpenOffice, and NeoOffice
- [Supports EPUB 2.0.1 and MOBI (via KindleGen)
- [Limited metadata support
- [Cover image from first page or from file
- [Mac/Windows/Linux .. and Free!

vonnegut-big-trip.odt - OpenOffice.org Writer

Heading 1 Adobe Caslon Pro Bold 16 B I U

Illustrated by KOSSIN ¶

CHAPTER ONE ¶

If it was good enough for your grandfather, else! ¶

Gramps Ford, his chin resting on his hands, his irascibly at the five-foot television screen that commentator was summarizing the day's happenings would jab the floor with his cane tip and shout

Emerald and Lou, coming in from the balcony, rarity—privacy—were obliged to take seats in brother and sister-in-law, son and daughter-in-law, husband, great-grandson and wife, nephew and great-grandnephew and wife—a everybody. All save Gramps, who was somewhat gerasone standards, to be about the same age thirties. Gramps looked older because he had invented. He had not aged in the 102 years since

"Meanwhile," the commentator was saying, "C tragedy. But 200 weary rescue workers have made effort to save Elbert Haggdorn, 183, who has

"I wish he'd get something more cheerful," Emerald

Writer2ePub Preferences

File Split

- Split Files Before Headings 1
- Split Files Before Headings 2
- Split Files Before Headings 3

Split files every: 200 Kb

Fonts

Base font height: 100 %

Font: Georgia

Embed Font in the ePub file

Image quality (20%–100%) 75 % CheckBox

- Keep working files
- Hide messages
- Close document after conversion
- Mobi Export (Kindlegen required)

Metadata

Meta extraction from filename: A - T - P - Y
(A=Author T=Title P=Publisher Y=Year)

- Show Metadata window before creation
- Add W2E credits at the end of the book.

Reset Help Cancel OK

Version 1.1.16

Page 1 / 9 Default English (USA) INSRT STD Outline Numbering : Level 1 100%

Pincette ODF to ePub

— [Extension for OpenOffice 3.2 or later

— [Supports EPUB 2.0.1

— [Metadata extracted from source file

— [Splits on outline level 1 styles and “page break before”

— [Cover image from first page

— [Mac/Windows/Linux

Aspose.Words Express

— [Converts from DOC, DOCX, RTF, HTML, or ODT

— [Supports EPUB 2.0.1

— [Allows control over various formatting options

— [Windows only .. Free!

Calibre

- [Converts CHM, EPUB, HTML, LIT, MOBI, ODT, PDF, RTF, TXT, ...
- [Supports EPUB 2.0.1, as well as MOBI and KF8
- [Extensive control over various formatting options
- [Numerous cleanup and processing features
- [Can use XPath to define chapters and breaks
- [Mac/Windows/Linux .. Free!

Input format: RTF

Output format: EPUB



Metadata



Look & Feel

Heuristic
Processing

Page Setup

Structure
DetectionTable of
ContentsSearch
&
Replace

EPUB Output



Debug

Detect chapters at (XPath expression):



Chapter mark: pagebreak

 Remove first image Remove fake margins Insert metadata as page at start of book

The header and footer removal options have been replaced by the Search & Replace options. Click the Search & Replace category in the bar to the left to use these options. Leave the replace field blank and enter your header/footer removal regexps into the search field.

Insert page breaks before (XPath expression):



Fine tune the detection of chapter headings and other document structure.

Restore Defaults

Cancel

OK

DITA-OT + DITA 4 Publishers

— [Converts from DITA XML

— [Supports EPUB 2.0.1 and MOBI

— [Uses XSLT and CSS from XHTML DITA-OT target

— [Ideal option for creating EPUB from DITA

— [Mac/Windows/Linux .. Free!

DocBook to EPUB

— [Converts from DocBook XML

— [Supports EPUB 2.0.1 and 3.0 (beta)

— [Uses XSL (also Python and Ruby options available)

— [Best option for creating EPUB from DocBook

— [Mac/Windows/Linux .. Free!

Pincette ODF to ePub (batch)

— [Converts from ODF

— [Supports EPUB 2.0.1

— [Metadata extracted from source file

— [Splits on outline level 1 styles and “page break before”

— [Cover image from first page

— [Mac/Windows/Linux

WebWorks ePublisher

— [Converts FM, DOC, DOCX, RTF, DITA, and other XML models

— [Supports EPUB 2.0.1

— [Highly configurable formatting options

— [Mix and match input formats

— [Batch conversion through AutoMap

— [Windows only

How to choose?

- [Remember .. there's no "perfect tool"
- [Hand-crafted or conversion?
- [Try all that remotely meet your needs
- [Have a number of tools on hand
- [Watch for new tools

Conclusion

- [Create samples and test (and retest)
- [Test on as many devices and applications as possible
- [EPUBs may need to be “fixed”
- [Validate the output with epubcheck
- [Make use of scripting and batch processing

Tool resources

Adobe InDesign – <http://www.adobe.com/go/indesign/>

Adobe RoboHelp – <http://www.adobe.com/go/robohelp/>

Adobe Technical Communication Suite –
<http://www.adobe.com/go/technicalcommunicationsuite/>

Apple iWork Pages – <http://www.apple.com/iwork/pages/>

Atlantis Word Processor –
<http://www.atlantiswordprocessor.com/en/>

Aspose.Words Express – <http://www.aspose.com/categories/free-utilities-and-components.aspx>

Calibre – <http://calibre-ebook.com/>

DITA Open Toolkit – <http://dita-ot.sourceforge.net/>

DITA for Publishers plugin –
<http://dita4publishers.sourceforge.net/>

DocBook to EPUB XSL stylesheets – <http://sourceforge.net/projects/docbook/files/epub3/>

ElmSoft EPubFm – <http://elmsoftonline.com/>

eCub – <http://www.juliansmart.com/ecub/>

epubcheck – <http://code.google.com/p/epubcheck/>

Jutoh – <http://www.jutoh.com/>

MadCap Flare – <http://www.madcapsoftware.com/flare/>

oXygen XML Editor – <http://www.oxygenxml.com/>

Pincette ODT to ePub – <http://www.pincette.biz/odftoepub/>

Scrivener – <http://www.literatureandlatte.com/scrivener.php>

Sigil – <http://code.google.com/p/sigil/>

WebWorks ePublisher – <http://www.webworks.com/Products/>

Writer2ePub – <http://extensions.openoffice.org/en/project/Writer2ePub/>

Resources / questions

— [IDPF EPUB 2 and 3 specifications - www.idpf.org

— [Follow #eprdctn on Twitter

— [Resources - www.epubtest.com/resources.php

— [Questions?

— Scott Prentice <[scott AT leximation.com](mailto:scott@leximation.com)>

— Leximation, Inc. - www.leximation.com