ePub: What, Why, and How?

WHAT

- ePub = Electronic Publication, aka eBook
- Official casing of ePub is "EPUB" (but I like "ePub")
- ePub file delivers an eBook via numerous readers and applications
- Just one of a number of eBook formats (MOBI, DJVU, PDF, HTML, TXT ...)
- Specifies the format and structure of the deliverable .. like a CHM, HLP, PDF or HTML file
- Requires an application and device to actually render and view the content
- Designed to be reflowable to fit the constraints of the rendering device or application.
- Most ePub readers work with a screen-based "paged" concept (trying to emulate a book).
- Underlying file format is XHTML and CSS
- May include DRM (Digital Rights Management)
- ePub specification maintained by IDPF (International Digital Publishing Forum)
- ePub 2.0 Became an official standard in September 2007, superseding the older Open eBook standard from 1999
- ePub 2.0.1 was approved in May 2010 and is the current stable release
  http://idpf.org/epub/20/spec/OPS_2.0.1_draft.htm
- ePub 3.0 first public draft released in Feb. 2011
- Combination of the following specs:
  - Open Publication Structure (OPS) - a standard for representing the content of electronic publications
  - Open Packaging Format (OPF) - defines the structure and semantics as well as the mechanism by which the various components of an OPS publication are related
  - Open Container Format (OCF) - defines the mechanism by which all components of an electronic publication are packaged into a single deliverable (ZIP archive)

WHY

- eBooks have become very popular over the past few years
  http://idpf.org/about-us/industry-statistics
- Many options for end users .. dedicated devices, or desktop applications, even in your web browser.
  http://www.epubbooks.com/ebook-readers
- Dedicated eBook readers projected to exceed 11 million units in 2011
  http://digitalbook.wikispaces.com/eReader+Sales+Figures
- ePub format is supported by all eBook readers and applications except the Kindle
- Especially nice for books that have a limited life span
- Seems best for "linear" books, but can be used for references as well
- Fixed reading path may complicate reference material, provide a good TOC and possibly an index (issues)
- Possible caveats
  - tables may not work well on small screens
  - links don't always work
  - reader applications don't all support the same level of the specification .. few (none?) are totally compliant
  - spec does not support index, but one can be created as a page with links

- Hardware devices:
  - Apple iPad, iPhone, iPod Touch
  - Sony Reader Pocket Edition, Touch Edition, Daily Edition
  - Barnes and Noble NOOK, and NOOK Color
  - Kobo eReader
  - Bookeen Cybook Opus
  - iRiver eReader
  - Elonex eReader
  - Tablet computers .. Samsung Galaxy, Motorola Xoom, HP Slate, others...
- Software applications:
  - EPUBReader - Firefox
  - Adobe Digital Editions - Windows, Mac
  - Stanza - Windows, Mac, iOS
  - Aldiko - Android
  - FBReader - Windows, UNIX (Open source)
  - Calibre - Windows, Mac, Linux
- Online storage and browser-based readers:
  - Ibis Reader
  - Bookworm Online ePub eBook Reader
  - Book Glutton

HOW

- ePub is just a collection of XHTML, CSS, and XML files. You could create one "by hand," but don't, use a tool.
- Your current authoring tool may export to ePub, if not, there are many conversion utilities available.
- As with the readers, tools for creating an ePub will vary in their support of ePub features; try many before choosing one

Authoring/conversion tools:

- Adobe RoboHelp 8 or 9 - $1000, Windows. Import from FM or Word.
- Adobe InDesign - $700; Windows, Mac
- Apple iWork Pages - $80; Mac
- eCub - $0; Windows, Mac, Linux, FreeBSD, Solaris. Good for converting existing HTML or text content into an eBook
  http://www.juliansmart.com/ecub
- Jutoh - $39; Windows, Mac, Linux, FreeBSD, Solaris. eCub + WYSIWYG authoring interface, plus more features
  http://www.jutoh.com/
- Atlantis Word Processor - $35, Windows.
  http://www.atlantiswordprocessor.com
- Sigil - $0; Windows, Mac, Linux (open source). WYSIWYG eBook editor. *http://code.google.com/p/sigil/*

Desktop conversion tools:

- Calibre - $0; Windows, Mac, Linux (open source). Not an editor; a conversion and management tool. Input formats: CBZ, CBR, CBC, CHM, EPUB, FB2, HTML, LIT, LRF, MOBI, ODT, PDF, PRC, PDB, PML, RB, RTF, SNB, TCR, TXT.
  http://calibre-ebook.com/
- DITA Open Toolkit + DITA4Publishers plugin - $0; Generate EPUB from DITA. *http://sourceforge.net/projects/dita-ot/*
- eScape ODT2ePub converter - $0; Windows, Linux.
  http://www.infogridpacific.com/igp/AZARDI/eScape%20-ODT2ePub/
- ODT to ePub - $55; Windows, Mac, Linux (Java)
  http://www.pincette.biz/odftoepub/index.xhtml
- PDF to ePub - $40; Windows. *http://www.pdftoepub.com/*

- epub-tools - $0; Windows, Mac, Linux (open source). Utilities for converting to ePub from: Word, RTF, DocBook, TEI, FictionBook. *http://code.google.com/p/epub-tools/*
- epubcheck - $0; Windows, Mac, Linux (open source). Validation tool. *http://code.google.com/p/epubcheck/*

Online conversion tools:
- Feedbooks - *http://feedbooks.com/share/*
- BookGlutton API - *http://www.bookglutton.com/api/*
- EasyEPUB - *http://www.easypress.com/products/aqpep/*
- Epub2Go (PDF to ePub) - *http://www.epub2go.com*

What should be in an ePub?
- ePub is a "ZIP" of the content and packaging files
- Content files (OPS specification):
  - Must be valid XHTML 1.1 (and valid XML 1.0)
  - Not all XHTML 1.1 elements and attributes are supported
  - Formatted with a subset of CSS 2 markup
  - Unicode is supported but readers are not required to support all glyphs
  - Supports the following image types: GIF, JPG, PNG, SVG
- Packaging documents (OPF specification):
  - All files must be valid XML 1.0
  - Unique identifier for the OPS Publication as a whole.
  - Publication metadata (title, author, publisher, etc.).
  - List of files (documents, images, style sheets, etc.) that make up the publication.
  - Arrangement of documents providing a linear reading order.
  - Set of references to fundamental structural features of the publication, such as table of contents, foreword, bibliography, etc.

About Scott

- 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.
- Coined the term "AIR Help" in 2007 after learning about Adobe's new AIR technology.
- Interested in creating innovative ways to provide user assistance that is actually used.

Questions?

- Scott Prentice <info AT leximation DOT com>
- Leximation - *http://www.leximation.com*