



ePUB

# ePub: What, Why, and How?

STC Summit 2011 - Sacramento, CA

---

17 May 2011 - Scott Prentice, Leximation, Inc.

# 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.

# What is an ePub?

---

- ❖ An electronic publication or eBook
- ❖ A file format, like a CHM (HTML Help)
- ❖ A collection of files that define the content, organization, and formatting
- ❖ Must be viewed with a reader application or device
- ❖ Just one of many eBook formats (AZW, MOBI, PDF, HTML, TXT, ...)
- ❖ Technically cased “EPUB,” but I prefer “ePub” so that's what you'll see here

# 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
- ❖ ePub 3.0 first public draft released in Feb 2011
- ❖ Combination of the following specifications:
  - ❖ Open Publication Structure (OPS)
  - ❖ Open Packaging Format (OPF)
  - ❖ Open Container Format (OCF)



# So, what's under the hood?

---

- ❖ Underlying technologies are XML, XHTML, and CSS
- ❖ Supports images, standard formatting, links (well, mostly)
- ❖ Content intended to be re-flowable to fit the constraints of the rendering device or application
- ❖ Most eBook readers work on a screen-based “paged” concept
- ❖ Content may include DRM or be open

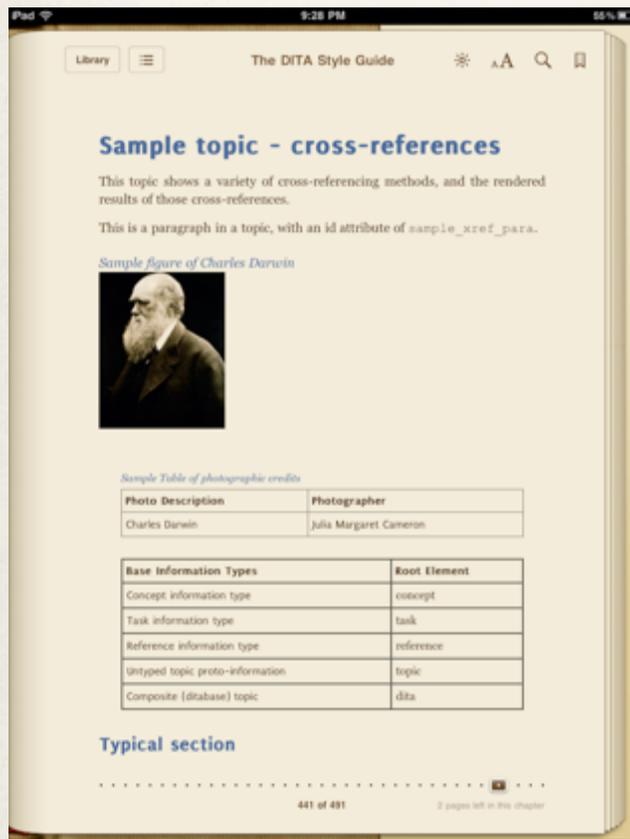
# What does an ePub look like?

---

- ❖ Well, it depends ..
- ❖ Fundamentally, an eBook reader application is just a web browser, so in general, it looks a lot like a simple web page
- ❖ Readers and applications will render the same ePub differently
- ❖ It's the "browser wars" all over again .. but worse
- ❖ The following samples show the same content on different devices

*Samples from the "DITA Style Guide" by Tony Self, published by Scriptorium Press, used with permission.*

# Tablet / dedicated readers



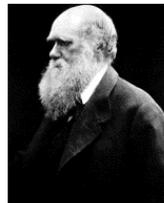
iPad 1 - iBooks

## Sample topic - cross-references

This topic shows a variety of cross-referencing methods, and the rendered results of those cross-references.

This is a paragraph in a topic, with an id attribute of `sample_xref_para`.

*Sample figure of Charles Darwin*

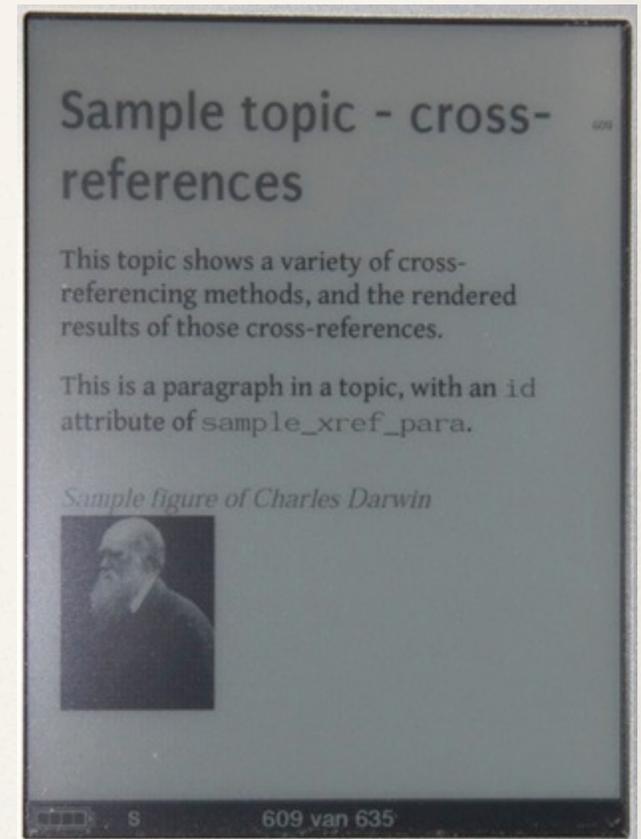


### Sample Table of photographic credits

Photo Description	Photographer
Charles Darwin	Julia Margaret Cameron

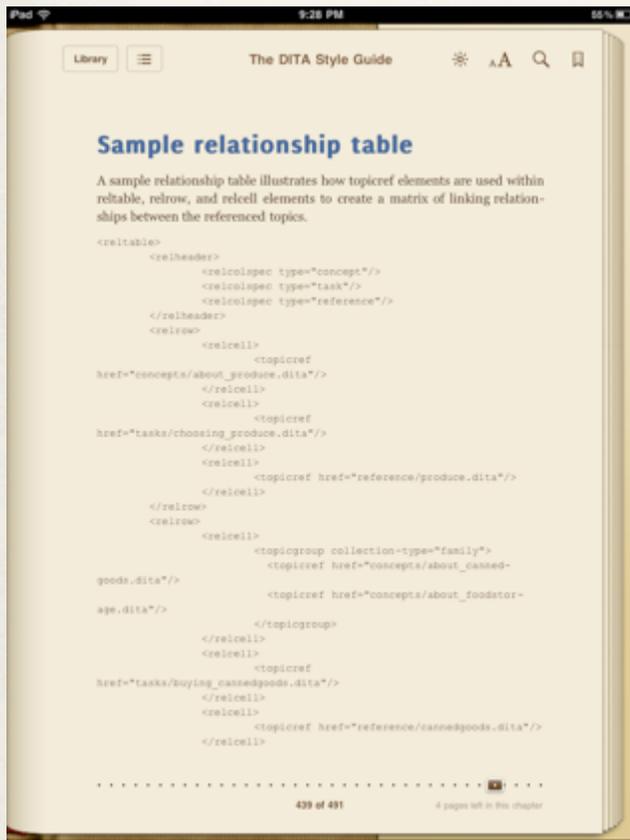
79%

Kindle 3

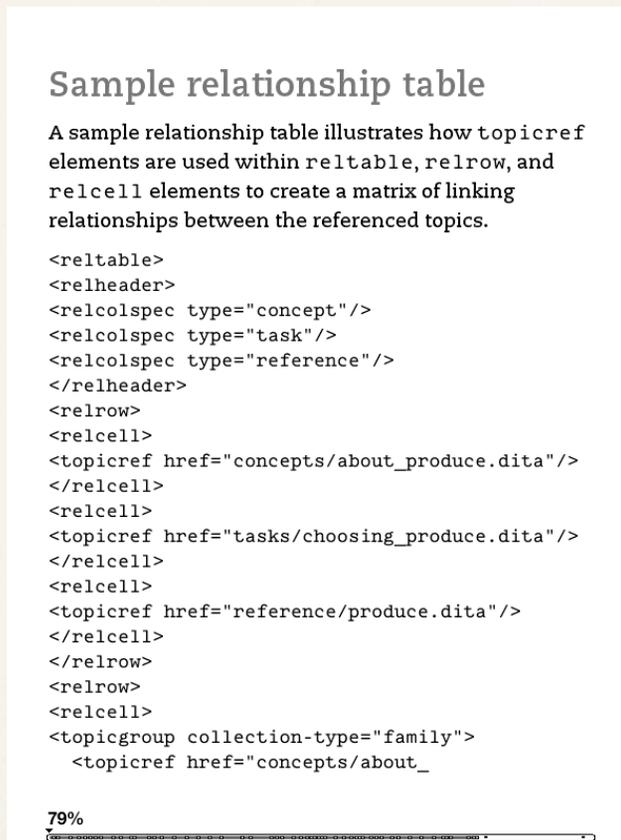


Sony PRS-600

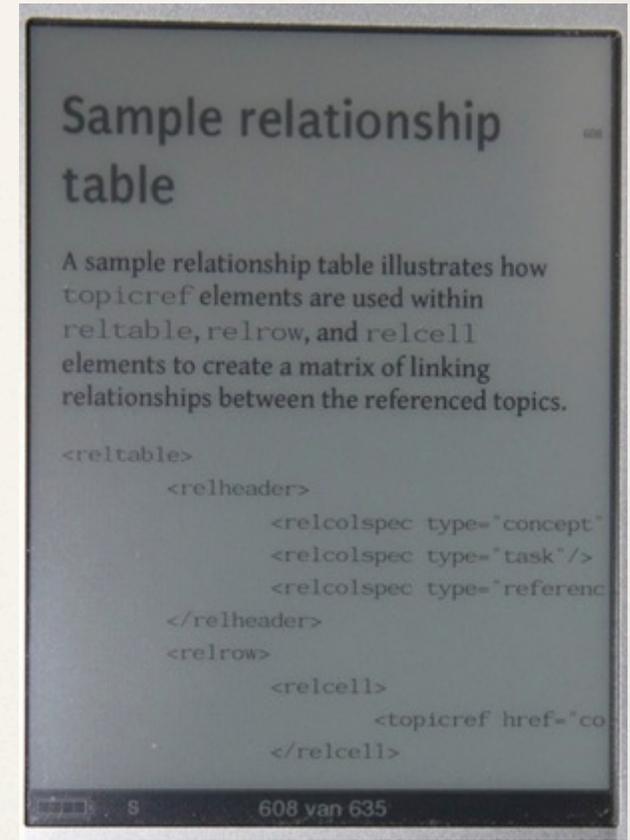
# Tablet / dedicated readers



iPad 1 - iBooks



Kindle 3



Sony PRS-600

# Mobile phones

AT&T 6:05 PM

## Sample topic - cross-references

This topic shows a variety of cross-referencing methods, and the rendered results of those cross-references.

This is a paragraph in a topic, with an id attribute of `sample_xref_para`.

Sample figure of Charles Darwin



iPhone (1G)  
Stanza

## Sample topic - cross-references

This topic shows a variety of cross-referencing methods, and the rendered results of those cross-references.

This is a paragraph in a topic, with an id attribute of `sample_xref_para`.

Sample figure of Charles Darwin



Nexus / Android  
Aldiko

AT&T 6:07 PM

## Sample relationship table

A sample relationship table illustrates how `topicref` elements are used within `reltable`, `relrow`, and `relcell` elements to create a matrix of linking relationships between the referenced topics.

```
<reltable>
  <relheader>
    <relcolspec type="concept"/>
    <relcolspec type="task"/>
    <relcolspec type="reference"/>
  </relheader>
  <relrow>
    <relcell>
      <topicref href="concepts/about_prod" />
    </relcell>
    <relcell>
      <topicref href="tasks/choosing_prod" />
    </relcell>
    <relcell>

```

iPhone (1G)  
Stanza

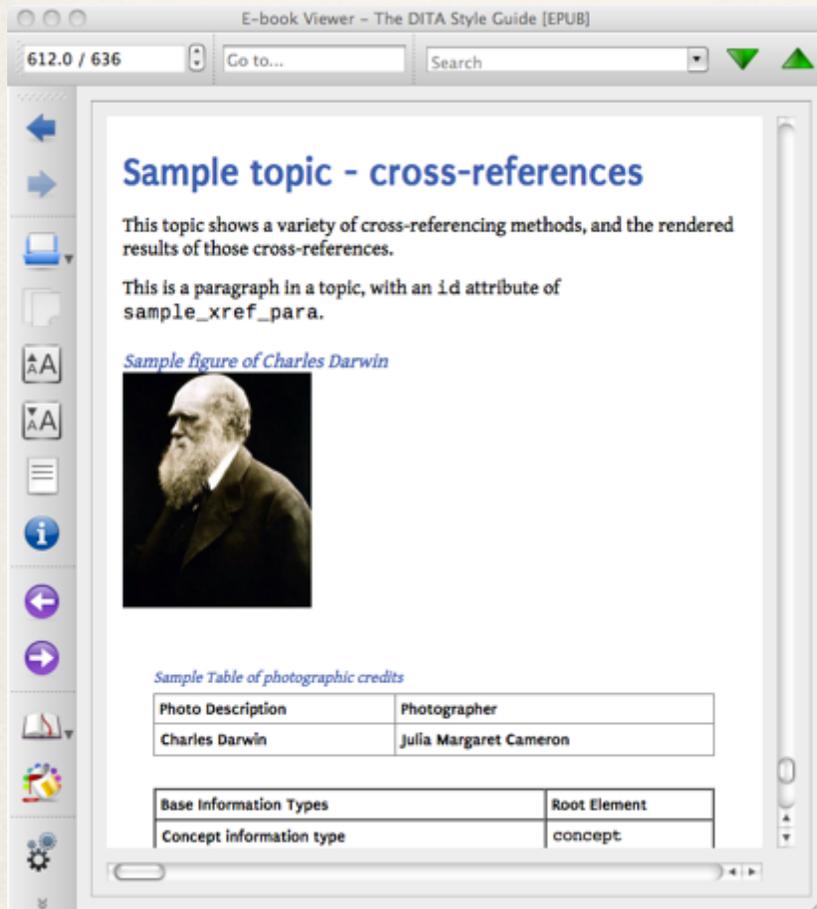
## Sample relationship table

A sample relationship table illustrates how `topicref` elements are used within `reltable`, `relrow`, and `relcell` elements to create a matrix of linking relationships between the referenced topics.

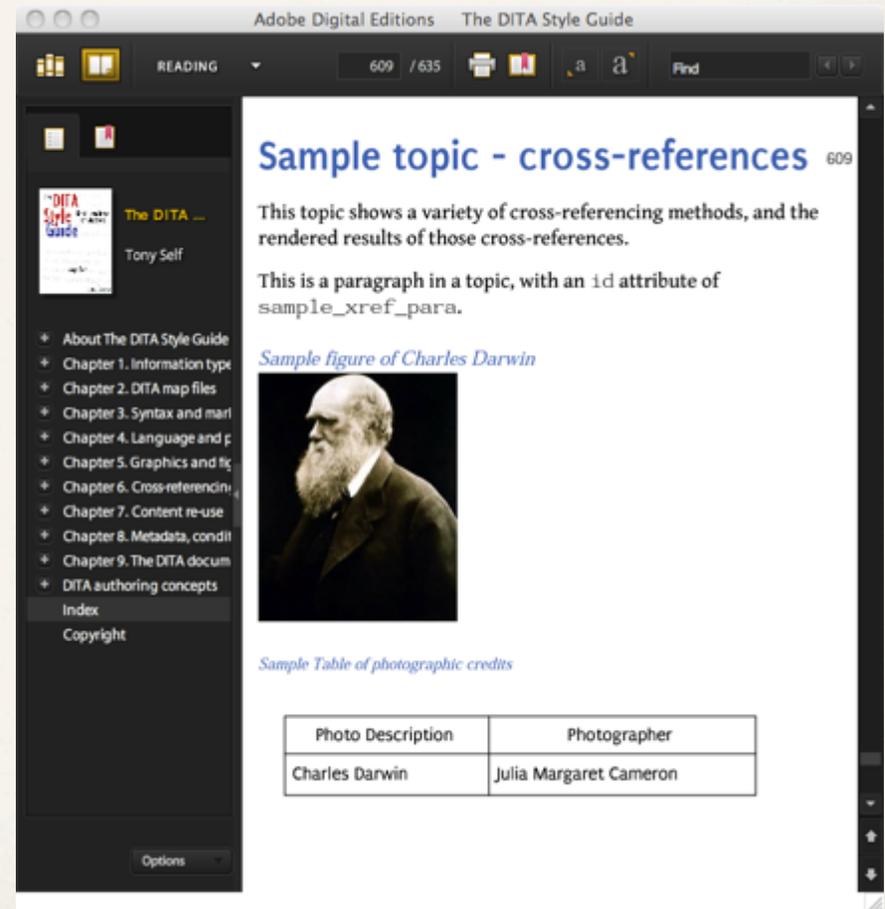
```
<reltable>
  <relheader>
    <relcolspec
type="concept" />
    <relcolspec
type="task"/>
    <relcolspec
type="reference"/>
  </relheader>
  <relrow>
    <relcell>
```

Nexus / Android  
Aldiko

# Desktop reader applications

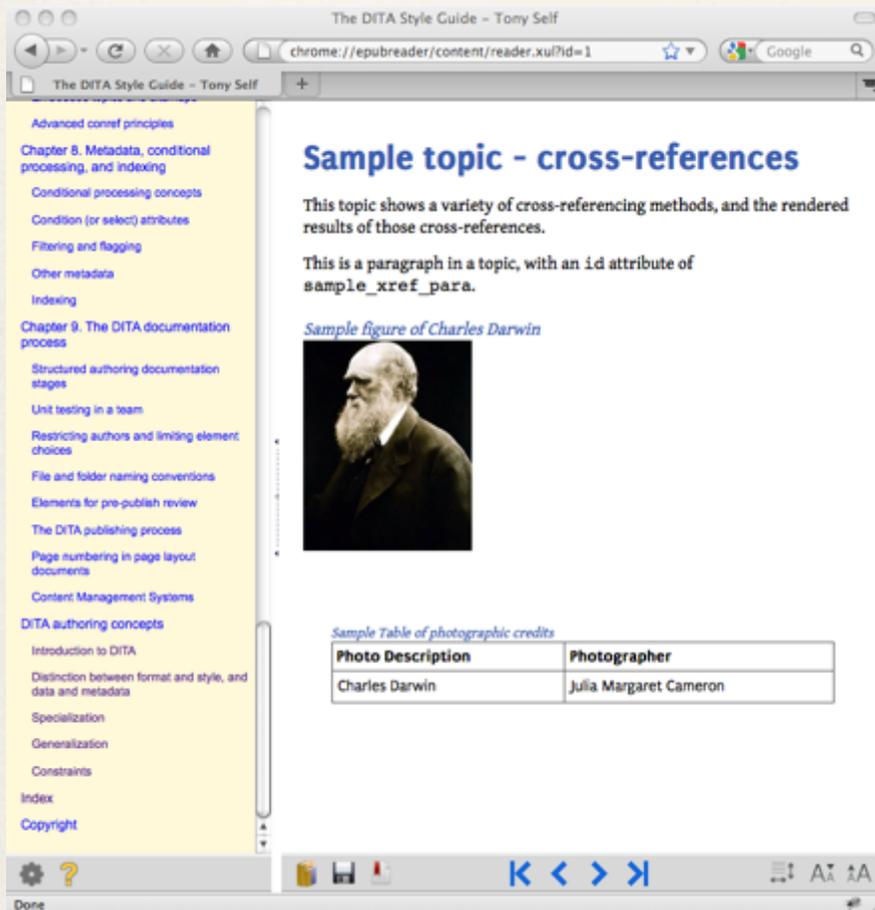


Calibre - Mac

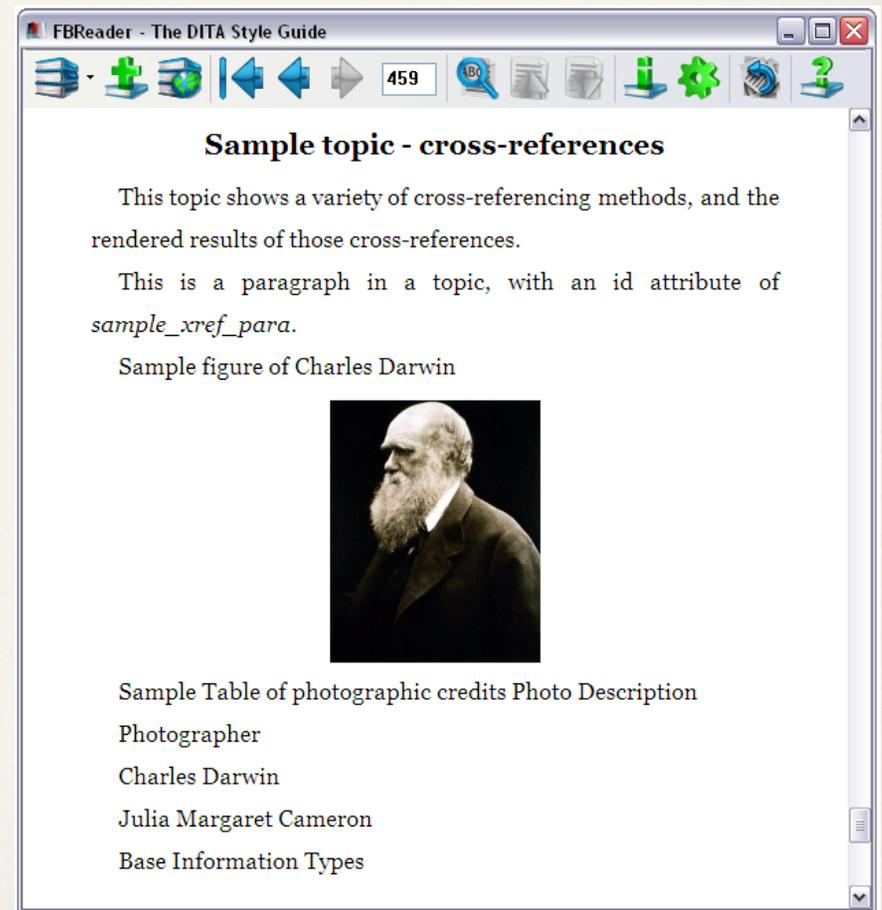


Adobe Digital Editions - Mac

# Desktop reader applications



Firefox - EPUBReader plugin



FBReader - Windows

# Why should I care?

---

- ❖ ePub is the most widely accepted eBook format by dedicated readers and applications
- ❖ Make your documentation available on all devices and platforms
- ❖ Association of American Publishers, April 14, 2011 ..

## **Popularity of Books in Digital Platforms Continues to Grow, According to AAP Publishers February 2011 Sales Report**

*Thursday, 14 April 2011 | Andi Sporkin*

### **E-Books Rank as #1 Format among All Trade Categories for the Month**

April 14, 2011; New York, NY-- Powerful continuing growth of books on digital platforms--both e-Books and Downloaded Audiobooks--are highlights of the February 2011 sales report of the Association of American Publishers, which is being released today.

The report, produced by the trade association of the U.S. book publishing industry, tracks monthly and year-to-date publishers' net sales revenue in all categories of commercial, education, professional and scholarly books and journals.

# Other benefits

---

- ❖ Instant gratification
- ❖ Lower cost (as compared to a physical book)
- ❖ Most readers offer full text search, bookmarks, and annotations
- ❖ Makes it possible for your documentation to ultra portable
- ❖ Particularly nice for books that have a limited life span
- ❖ Seems best for “linear” books, but can be used for references as well

# Caveats?

---

- ❖ Tables may not work well on small screens
- ❖ Formatting will look different on all readers
- ❖ 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 pages with links (as long as links are supported by the reader)
- ❖ Fixed reading path may complicate use of reference material

# eBook 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...

# eBook software applications

---

## **“Desktop” applications**

- ❖ EPUBReader - Firefox
- ❖ Adobe Digital Editions - Windows, Mac
- ❖ Stanza - iOS
- ❖ Aldiko - Android
- ❖ FBReader - Windows, Mac, Linux, Android, Motorola, etc.
- ❖ Calibre - Windows, Mac, Linux

## **Online storage and web-based applications**

- ❖ Ibis Reader
- ❖ Bookworm Online ePub eBook Reader
- ❖ Book Glutton - social eBook networking

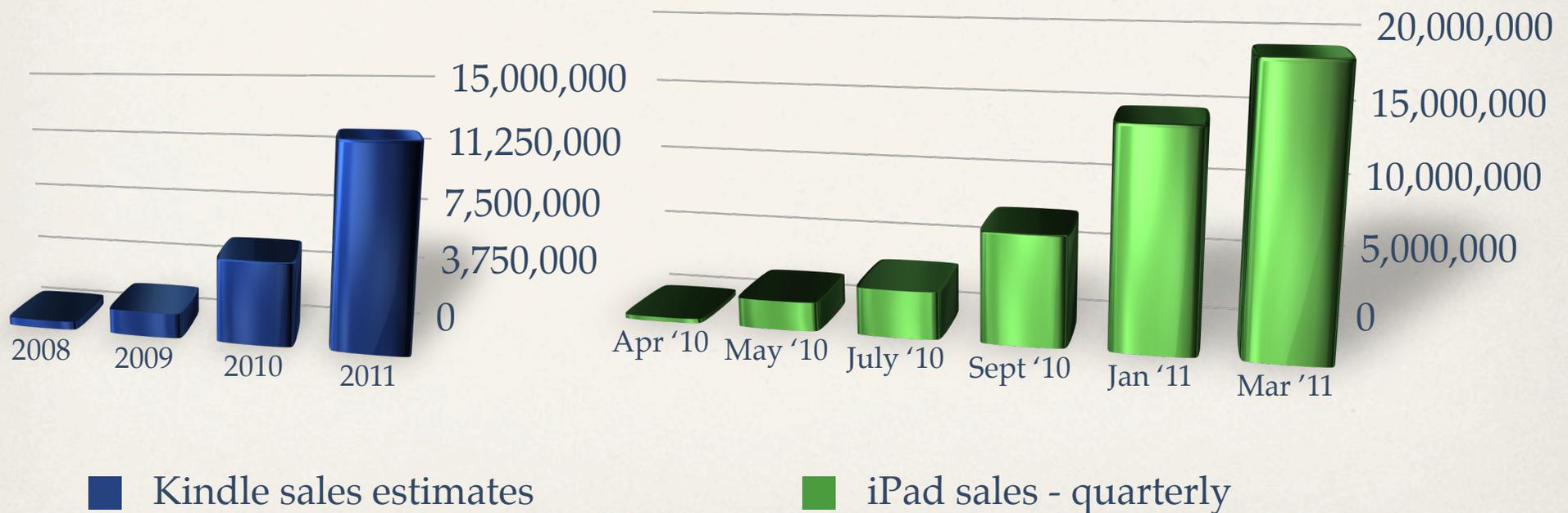
# Statistics and propaganda...

---

- ❖ Amazon claims the Kindle is the #1 best selling product on amazon. (Doesn't release sales figures.)
- ❖ “Nook passes Kindle sales in March” - CrunchGear.com, April 2010
- ❖ Nearly 18 million media tablets and 12 million e-book readers shipped in 2010, says IDC - digitimes.com, Mar 2011
- ❖ Global sales of 20-25 million eBook readers projected for 2011, says E Ink chairman

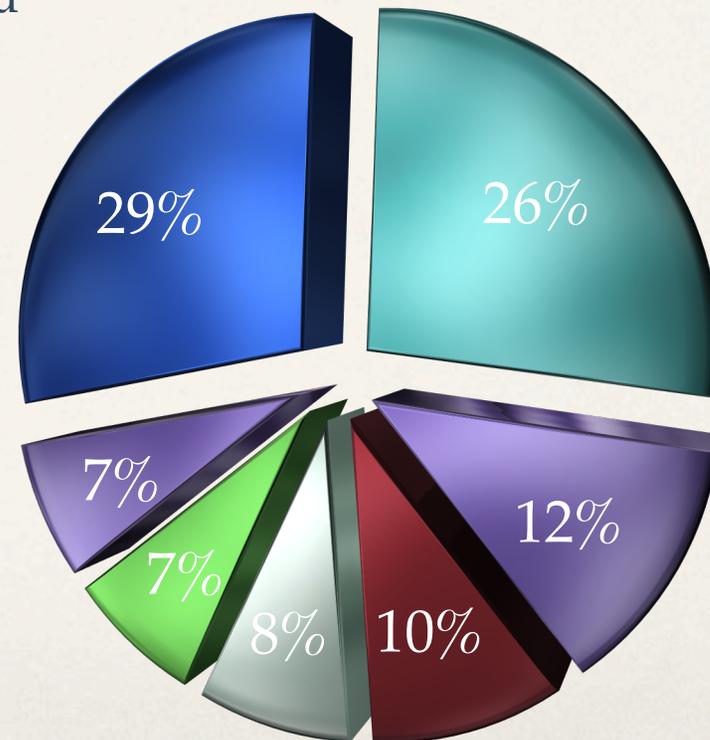
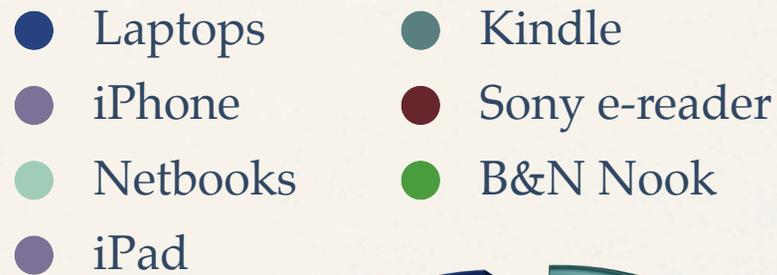
# Tablet and dedicated reader sales are huge

- ❖ Cumulative sales figures for two top selling devices .. Kindle and iPad



# Most popular eBook devices

\* Nov 2010 Wired article quoting a Forrester research survey on the most popular way to read an eBook...



# Many dedicated readers...



Asus DR900



Azbooka 516



B&N  
Small Nook



Bookeen  
Cybook 3 Gen

# Too many...



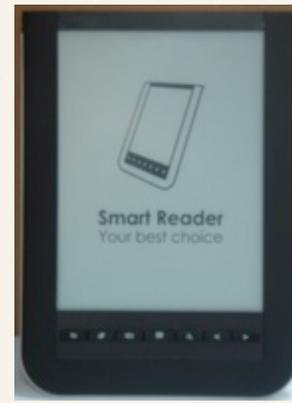
Bookeen  
Cybook Opus



Bookeen  
Cybook Orizon



Egriver  
Ideo



Egriver  
Touch



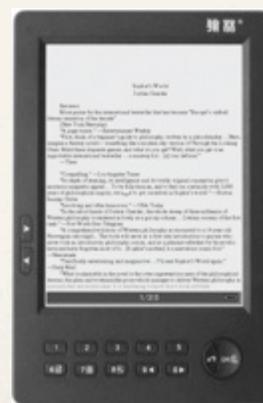
enTourage  
eDGe



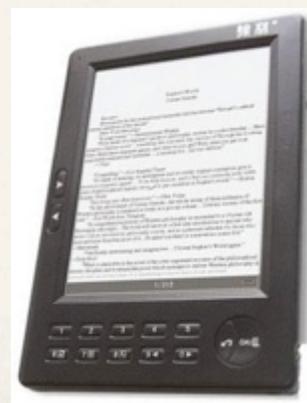
Fujitsu  
FLEPia



HanLin  
eBook A9



HanLin  
v3 Plus



HanLin  
v5



Hanvon  
HandyBOOK

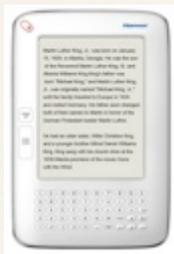


Hanvon  
WISereader E920

# Really, too many...



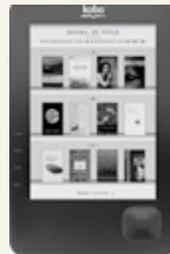
Hanvon  
WISereader  
518



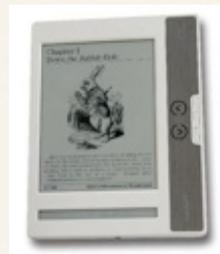
Hanvon  
WISereader  
526



iRiver  
Story



Kobo  
Wireless  
eReader



NUUTbook  
NUUT2



Onyx  
Boox M90



Onyx  
Boox 60



PAGEone



Pixelar  
MReader6  
Touch



Pixelar  
MReader6



Pixelar  
MReader9



Pocketbook  
301 Plus



Pocketbook  
302



Pocketbook  
360



Pocketbook  
903



Sagem  
Binder



Sony  
PRS-950



Sony  
PRS-350



Sony  
PRS-650



Spring Design  
Alex



Viewsonic  
VEB620



WISereader  
B630



WISereader  
N610



WISereader  
N618

# How do I make an ePub?

---

- ❖ An ePub is just a collection of XHTML, CSS, and XML files. You could create an ePub “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.
- ❖ Easiest method is to save to HTML, then use one of the free conversion tools to convert from HTML to ePub.
- ❖ Keep it Simple!

# Desktop authoring tools

---

- ❖ Adobe RoboHelp - \$1000, Windows
- ❖ Adobe InDesign - \$700, Windows/Mac
- ❖ Apple iWork Pages - \$70, Mac
- ❖ eCub - \$0, Windows, Mac, Linux, FreeBSD, Solaris
- ❖ Jutoh - \$40, Windows, Mac, Linux, FreeBSD, Solaris
- ❖ Atlantis Word Processor - \$35, Windows
- ❖ Sigil - \$0, Windows, Mac, Linux

# Desktop conversion tools

---

- ❖ Calibre - \$0, Windows, Mac, Linux
- ❖ DITA Open Toolkit + DITA for Publishers plugin - \$0
- ❖ eScape ODT2ePub converter - \$0, Windows, Linux
- ❖ Pincette ODT to ePub - \$55, Windows, Mac, Linux (Java)
- ❖ DNAML PDFtoePub - \$40, Windows.
- ❖ epub-tools ([code.google.com](http://code.google.com)) - \$0, Windows, Mac, Linux
- ❖ epubcheck ([code.google.com](http://code.google.com)) - \$0, Windows, Mac, Linux

# Online conversion tools

---

- ❖ Feedbooks ([www.feedbooks.com](http://www.feedbooks.com)) - author and develop eBook content directly in website
- ❖ EasyEPUB ([www.easypress.com](http://www.easypress.com)) - create ePubs from InDesign, Quark or MS Word files
- ❖ 2EPUB ([www.2epub.com](http://www.2epub.com)) - convert PDF, DOC, ODT, HTML, and eBook formats to EPUB, MOBI, and others
- ❖ Epub2Go ([www.epub2go.com](http://www.epub2go.com)) - free PDF to ePub converter

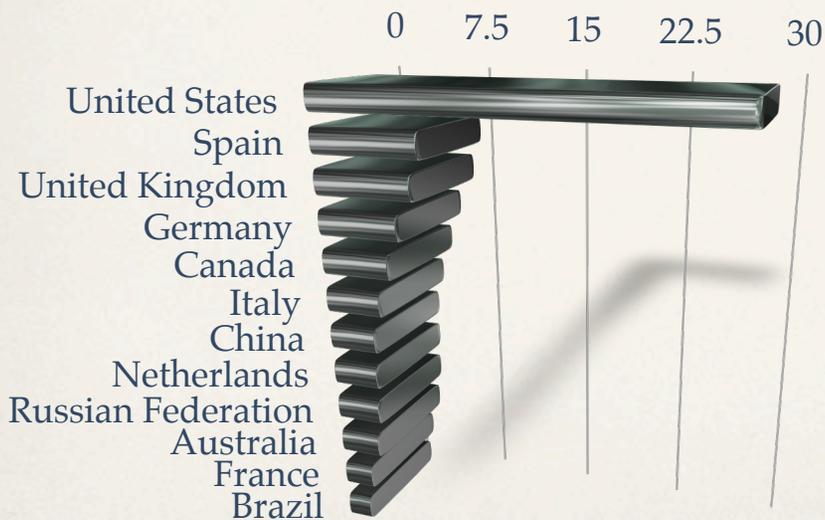
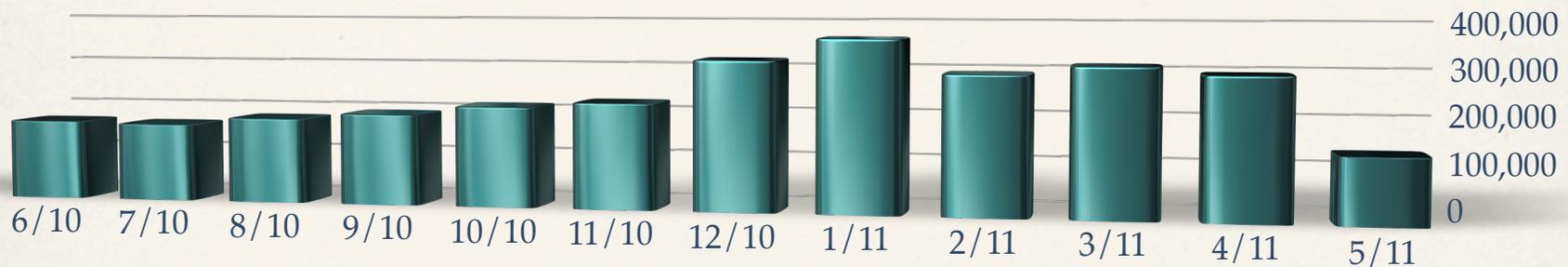
# Calibre .. a very nice tool

---

- ❖ Free, open source
- ❖ eBook library management tool (tag, catalog, comment, etc.)
- ❖ eBook conversion tool (to and from many formats)
- ❖ Server mode, view and download publications wirelessly
- ❖ Automatic news to eBook conversion
- ❖ Integrated eBook viewer

# Calibre usage statistics

■ New users per month



❖ 3,044,597 installations since Aug 2009

■ Percentage of users by country

# What's inside of 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 render all glyphs (placeholder is required)
  - Supports the following image types: GIF, JPG, PNG, SVG

# The ePub “package”

---

- ❖ 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 that make up the publication
  - Arrangement of documents providing a linear reading order
  - Definition of TOC hierarchy

# Questions?

---

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