Elisa Kirja
ePub support
Version 1.0

17 December 2012

CONTENTS

elisa

17 December 2012

1   What is epub?

Epub is a free, open e-book industry-standard file format defined by the International Digital Publishing Forum. It is designed for reflowable content instead of fixed content, which means that devices with different screen sizes can display the content optimally.

Epub has become widely adopted around the globe and book sellers such as Barnes and Noble, Waterstone and Google, Amazon being the exception. Project Gutenberg, which hosts over 30 000 free ebooks, is using epub as one delivery format. Also most eReader devices today support the epub standard, including mobile phones using iPhone OS, Android, Windows Mobile and Symbian.

Epub consists of three specifications: **Open Container Format (OCF), Open Publication Structure (OPS) and Open Packaging Format (OPF)**. It's basically a subset of xhtml and css, with some new custom properties added. It's packaged in a zip-format and the extension of the file is epub. Although epub is a free and open standard, it can contain an optional DRM protection. Adobe provides an implementation of such protection and it's becoming a de-facto standard at the industry.

2   Ebooks, metadata and their requirements for Elisa eReading service

Elisa is supporting valid epubs in the webshop. That means supporting the standard of **EPUB 2.0.1**. Full specifications of these can be found from: http://idpf.org/epub/201 and later on we will give some recommendations how to build and validate your own epub.

In general, to provide a good working epub which works in Elisa's service, one needs to remember that these are books, not websites:

- ePub 3.0 is *not* supported
- No JavaScript, no multimedia
- External HTTP links are supported but they won't work on hardware readers without web access. The current version of Elisa Kirja application supports internal or external links.
- Simple tables only – no nested tables, no fixed-width tables, no complex tables
- Keep your HTML page templates simple – complex and dynamic layouts will often fail
- No background images
- External images are *not* supported in epub
- Hotspots in images are not supported by most epub readers (both hardware and software)

Validate your epub and test it in Adobe Digital Editions and a real device!

elisa

17 December 2012

Elisa is providing two different digital rights protection solutions for epub, watermarking and Adobe's DRM (Adobe Content Server 4)**.** In Adobe's DRM this means that the electronic books are encrypted and can only be opened using the Adobe ID used by the buyer. This ID is generated for every user of the eReading service.

If the protection is provided by Adobe's DRM called ADEPT, the epub and pdf support for the devices are also based on Adobe's support, more specifically **Adobe Digital Editions.** More information about Digital Editions and it's supported features and devices can be found from http://www.adobe.com/products/digitaleditions

Adobe's InDesign also provides **protection for embedded** fonts when exporting to epubs, so they will be encrypted and cannot be extracted from the package. Adobe's Digital Editions software is then needed to open documents with the correct fonts. Elisa's supported ereader devices and Elisa Kirja application will also support this protection.

In addition to the actual content, the epub also needs metadata to tell about its content. Required fields in the metadata are: **title, author(s), identifier, language, publishing date and publisher.** These are used in the service to display and sort the books in the ereader device. A separate metadata is required to display the book in the webshop and to handle the statistics and pricing of the products.

3   What makes a good epub?

- Table of Contents
This is metadata of the epub and it provides user an access to a certain part of the book. eReaders support the display of TOC in different ways. Using the toc.NCX is recommended, even though epubs also support xhtml-based TOC.

- Metadata
To be able to list the books in webshop, in the device's book listing and different eReaders, the metadata of the book needs to be supplied. This means isbn-identifier (of the specific ebook), title, author, publisher, rights, publishing date, description, category and language.

- Layout fits to multiple screen sizes
Ebooks are usually read using small screen devices, so a good layout means designing for multiple resolutions. For example the widths of tables and images should be defined in percentage-values instead of pixels. Also keep in mind that reader can change the font size at his will. Defining image size in em (relating image size to font size) usually results in poor rendering quality in most Adobe Mobile Reader based devices and applications, including Elisa Kirja. This definition is only usable with small logo images, which should be the same size as text.

- Cover page
The eBook simply looks better if it has a cover page. Cover page is also used in eReaders such as Adobe Digital Editions to display the thumbnail of the book.

elisa

17 December 2012

Supported formats are gif, jpg, png and svg (but no animations!). Good size for cover page is usually about 700 x 1000 pixels and jpg is the most safe image file format. Depending on aspect ratio, image resolution should be following, width: 600-800 pixels and height: 800-1280 pixels. Reference for creating cover pages: http://blog.threepress.org/2009/11/20/best-practices-in-epub-cover-images/. SVG-wrapped cover page is a one good way of creating cover that scales properly to your eBook device:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<title>Book Title</title>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
<link href="style.css" rel="stylesheet" type="text/css"/>
<style type="text/css">
          body {margin:0; padding: 0;  border-width: 0;}
          @page {margin: 0; padding: 0; border-width: 0;}
          div {margin:0; padding: 0;  border-width: 0;}
</style>
</head>
<body>
<div>
<svg version="1.1" xmlns="http://www.w3.org/2000/svg"
          xmlns:xlink="http://www.w3.org/1999/xlink"
          width="100%" height="100%" viewBox="0 0 700 1000"
          preserveAspectRatio="xMidYMid meet">
          <image width="700" height="1000" xlink:href="images/cover.jpg" />
</svg>
</div>
</body>
</html>
```

- Good image format and packing
Jpg-format is good for all images with more than 256 colors. Use very high, high or medium quality packing depending on image file size constraints (below 500 kilobytes per image). Png-8 packing is usually good for images with few colors, below 256 (black/white, halftone images, logos, etc.). Maximum image resolution is 800 x 1280px or 1280 x 800px depending on aspect ratio.

- Good indexing and use of hyperlinks
Good indexing is valuable to reader since epubs don't really have page numbers. Use of relevant hyperlinks especially in professional books is strongly encouraged. Reference for improving indexes: http://blog.threepress.org/2009/11/08/improving-indexes-in-ebooks/ Building comprehensive hyperlinks is recommended.

elisa

17 December 2012

- It looks good and readable in the eReader!

Fastest way to verify the results is to use eg. Adobe Digital Editions, but to ensure that it looks good for the end user, verify the results in an eReader device. If possible, also test on a smartphone device such as the iPhone, iPad or Android based device. Elisa Kirja application should be one of your primary testing environment. Its Android version allows you to test digital books easily.

- Xhtml-files are small enough.

Uncompressed size of single xhtml file should be less than 300 kilobytes. To be on the safe size less than 100 kilobytes is good practise. One file for each chapter is good rule of thumb for a regular novel. Smaller files also load faster which gives reader a better user experience. If the book has a lot of images, you should split xhtml files even further to minimize the number of image references in single xhtml file. Reading software loads all referenced images into memory when loading xhtml file and with many images it takes longer. Combined size of single xhtml with all referenced images has to be smaller than 10 megabytes.

4   Converting to epub

Epub is html and css, so a good looking epub requires understanding of those underlying techniques. There are several conversion tools available, which can convert from formats such as html, MS LIT, MS word, pdf and InDesign, but none of these provide perfect results automatically. Since epub is reflowable, books designed with a fixed layout, especially with lots of figures embedded in the text, will usually provide unsatisfactory results. Font styles can be lost in the conversion, pictures end up being separated from text, line breaks are in wrong places and so on.

For these problems we have studied several different conversion routes and all of them need manual work and validation of the results. But for normal novels with text content, the conversion can be done in couple hours by an experienced user. In a more complex book, it takes significantly longer, depending on the required quality.

17 December 2012

4.1   Workflow

The best source format for epub conversion is InDesign document. The next best option is to start with a Word document and the most difficult alternative is to have pdf as the source format. In general, it is best to convert formats from the most complex (and rawest) formats to the simpler ones.

In the long term, the optimal option would be to move from InDesign to a xml-based, automated workflow of publishing, where epub could be automatically generated from structured content such as DocBook. This is the approach O'Reilly, for example, uses. See: The Truth About eBooks: Devices, Formats, Pirates. This makes it easier to support in the future also other formats than epub.

Recommended InDesign based workflow
For epub-conversion the following workflow is suggested:
1) In InDesign, make the document work according to the tips below in *InDesign tips*
2) Export the epub from InDesign
3) Cleanup the epub
4) Validate the epub
5) Test the epub on different platforms
6) Publish

The skills required for epub conversion include:
* experience in web styling (css, html, xhtml)
* xml understanding (schemas, dtd:s, validation)
* graphics design / layout (Adobe InDesign, typography)

4.2   InDesign tips

Use only version CS4 or newer of InDesign for generating epubs! The quality of epub exporting in previous versions of InDesign is not acceptable. CS5.5 InDesign exports images much better than previous versions so you should upgrade to that if you make books with lot of images.
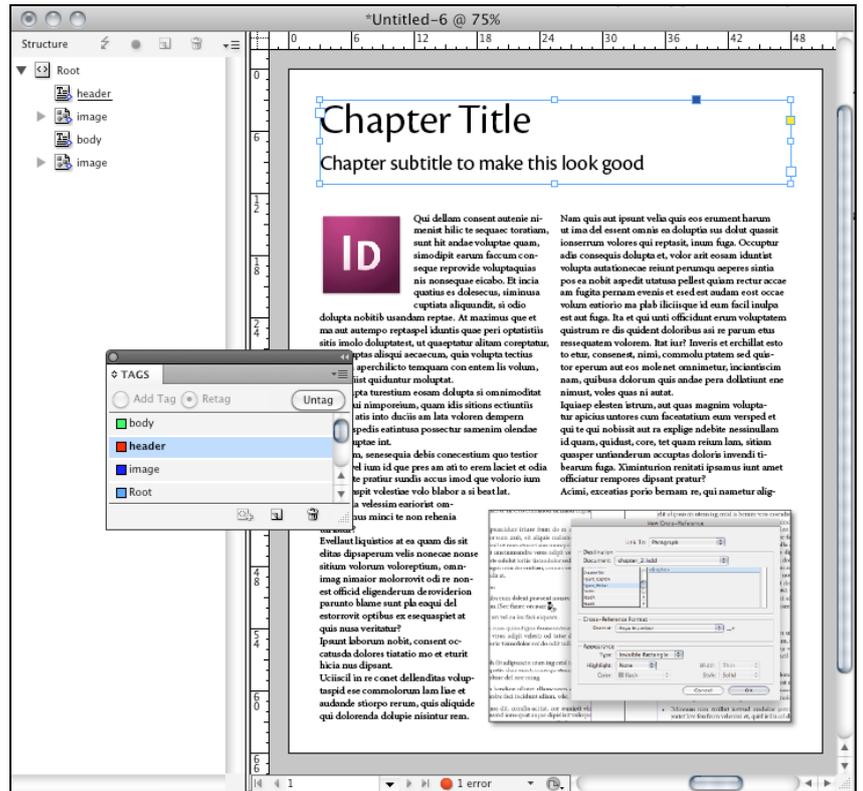
* Organize your content flow before ePub export.
  Options for controlling the order of your content:
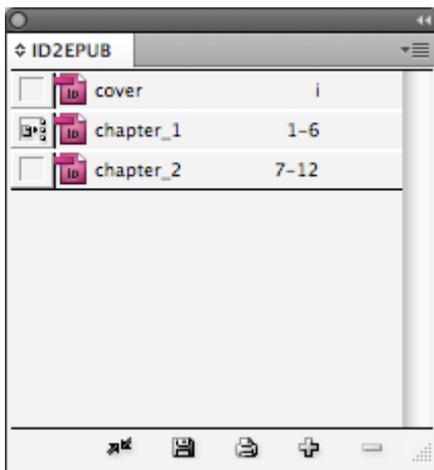
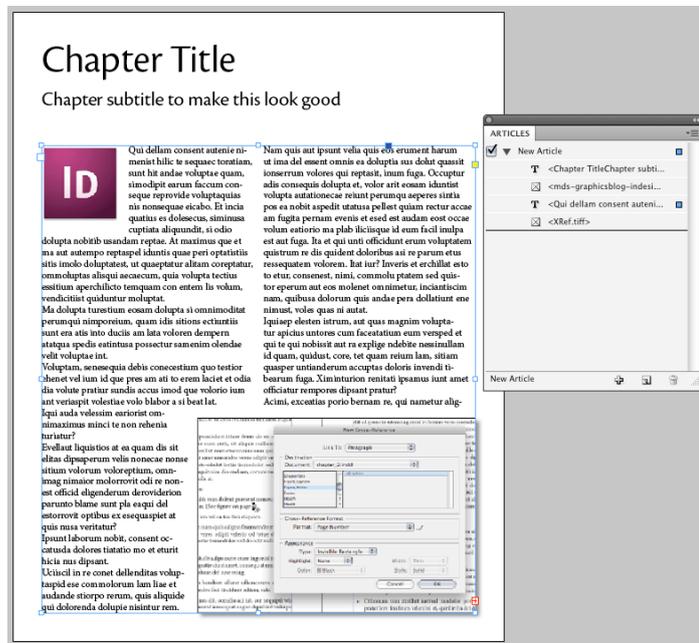17 December 2012
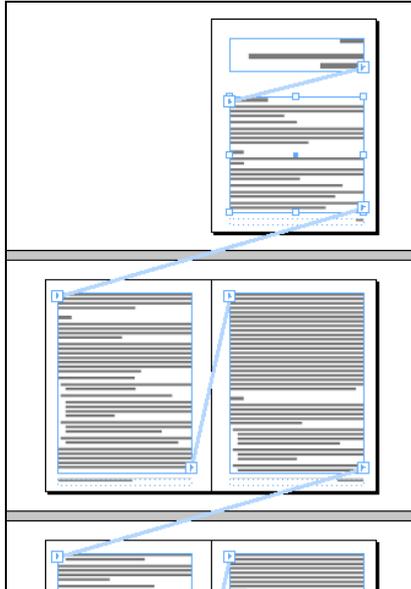
1. One text flow,
   all text frames linked together

2. XML structure

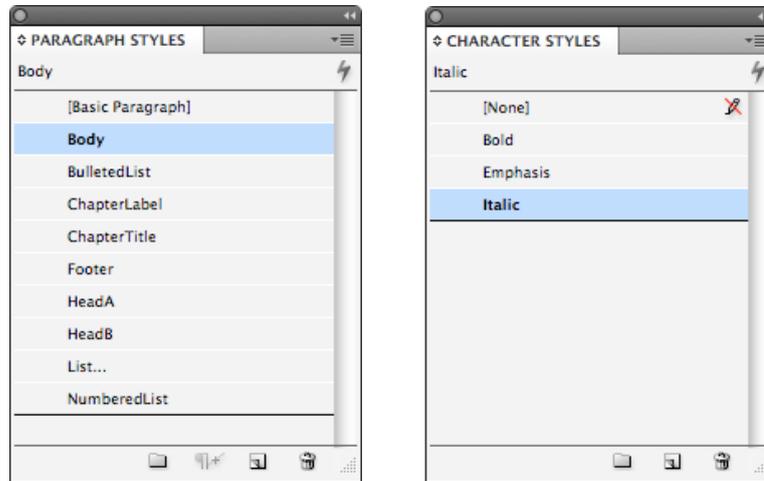3. Book file for more than
   one Indesign file

4. Articles panel

17 December 2012

- Always create and use paragraph and character styles to styled text

- Check that all images are anchored or inline to text frames. If you are using content flow option 2 or 4, then you can also place images straight to the xml structure or articles panel. Doing otherwise, images may end up towards the beginning or end of the document. It is good practice to check all images visually from exported epub and make corrections manually if needed. This important because the position of an image in printed book is often determined by layout and content. For example an image is at the bottom of the page because it looks nice and the text related to that image is before the image. In epub it might be better if image is before the text which is related to the image. Now we have greater probability that reader can see the image and text at the same time or at least picture gets explained. This is really depending on content and manual work is needed. Indesign CS5.5 also gives you better options for object and image export. Custom image rasterization, alignment and spacing, and Alt text.
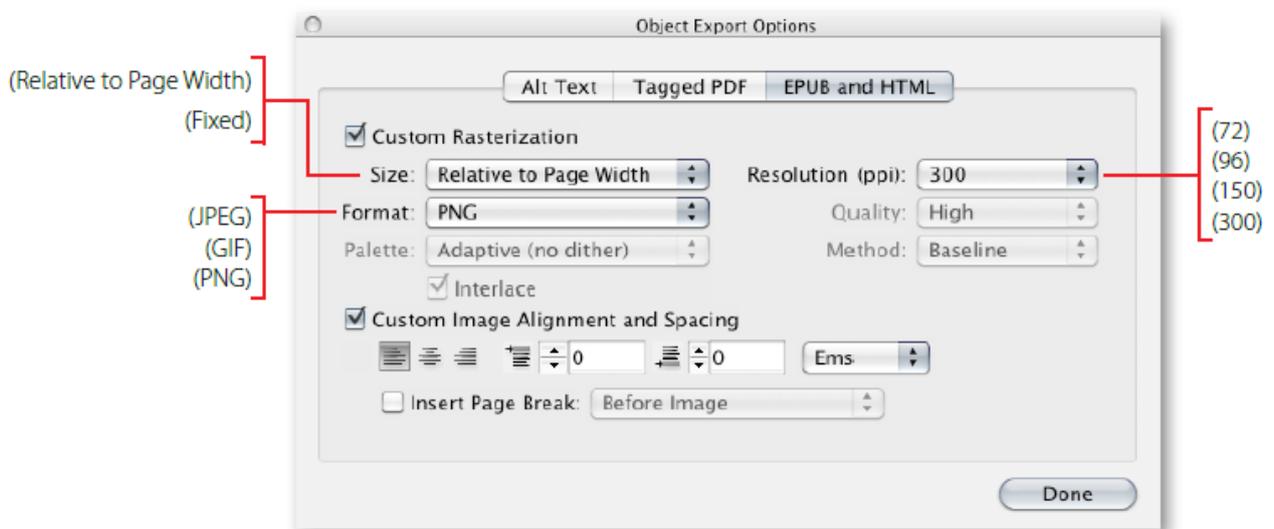
17 December 2012

- Create "Table of Contents" for the book using "Layout / Table of Contents …"

This table of contents created from paragraph styles will be used automatically in the epub TOC.
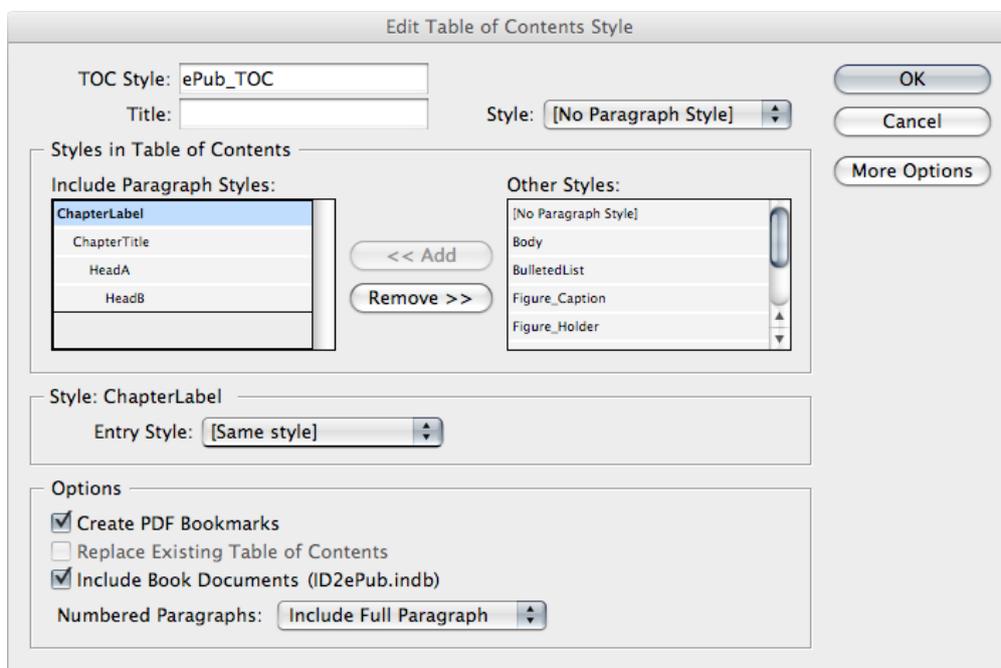
- Hyperlinks and cross-references

If your hyperlinks and cross-references are properly set up, then they will translate seamlessly to clickable hyperlinks when exported to ePub.

- The first page of the document should be the image of the book cover. Please note that even CS5.5 version of InDesign can't export cover page correctly in some cases. You should check the cover page carefully and replace the cover page image with correct one if needed. Usually cover page image can be taken from pdf file of the cover.

17 December 2012

- Use "File / File Info" to define metadata fields for the epub:
– Document Title
– Author
– Date (In CS5.5, the date info is added automatically)

File Information for epub-test-281009.indd

| Description | IPTC | Camera Data | Video Data | Audio Data | Mobile SWF | Categories | Origin | DIC ▶ ▼ |

Document Title: Epub test document.

Author: eReader team

Author Title: editor

Description: This is a document for testing ePub conversion and device compatibility.

Rating: ★ ★ ★ ★ ★

Description Writer:

Keywords:

ⓘ Semicolons or commas can be used to separate multiple values

Copyright Status: Copyrighted ▼

Copyright Notice: Elisa

Copyright Info URL: [Go To URL...]

Created: 28.10.2009 – 11:37:07   Application: Adobe InDesign 6.0

Modified: 28.10.2009 – 13:56:47   Format: application/x-indesign

Powered By **xmp**

[Import... ▼] [OK] [Cancel]

elisa

17 December 2012

- Export with "File / Export…"
– Select "Include Document Metadata". This grabs metadata from File Info.
Use ISBN number on Unique Identifier.
– Select EPUB cover rendering option
– Select used content ordering and formatting option
– Images: default settings
– Contents: Format for EPUB Content: "XHTML"
– Select "Include InDesign TOC Style"
– Base for CSS Styles: Defined Styles
– Select "Include Embeddable Fonts"
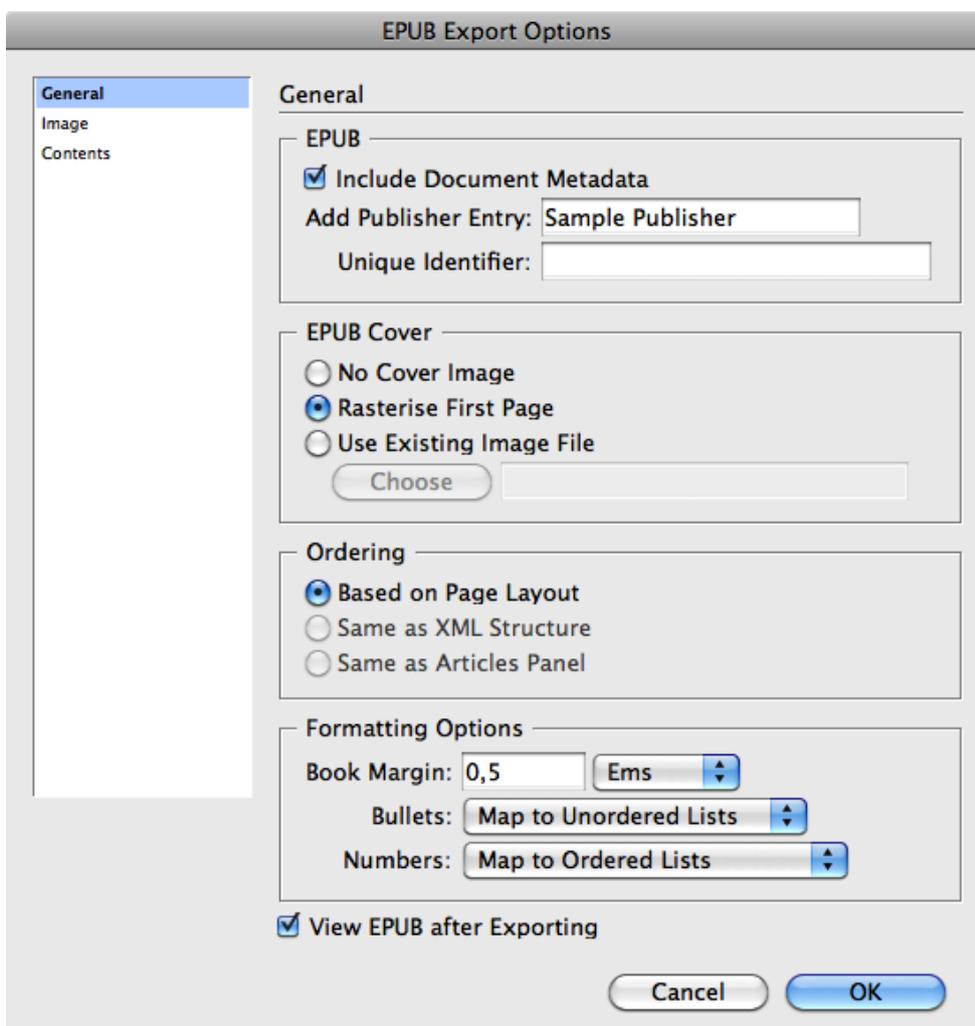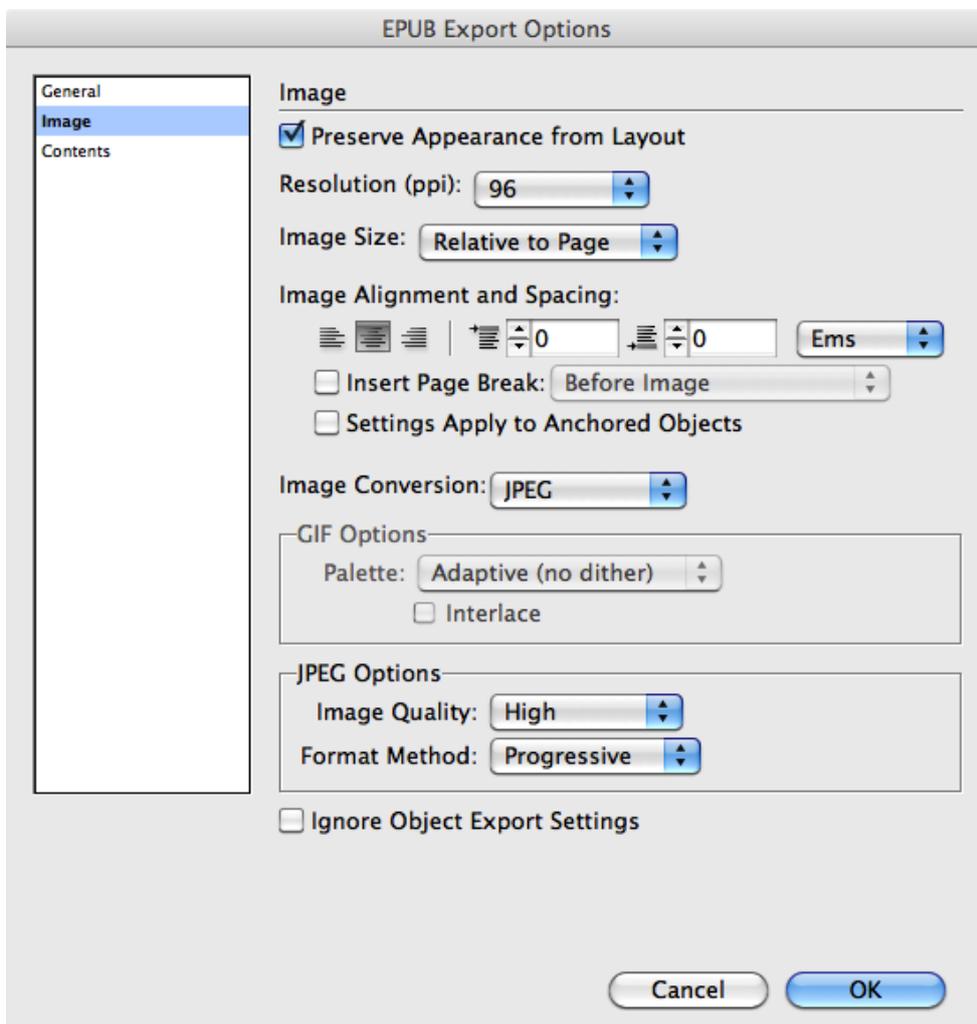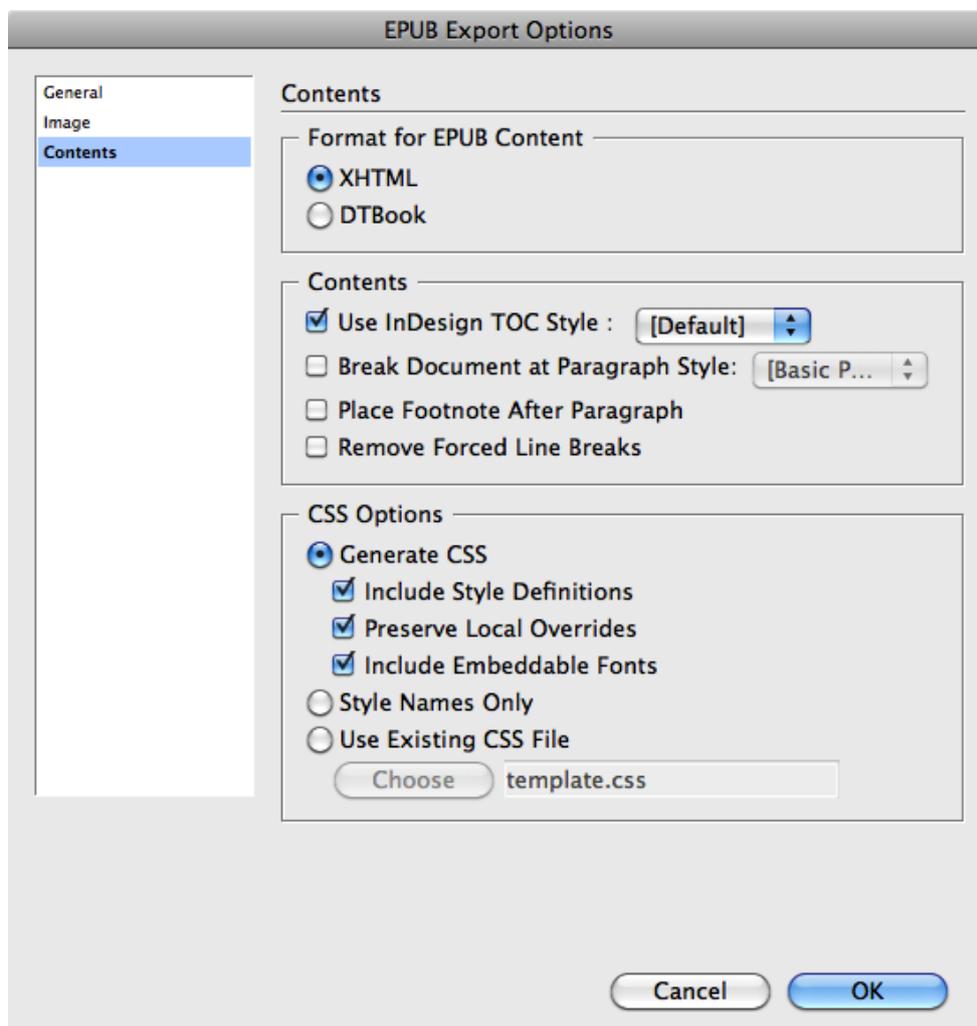
17 December 2012

Image export in Adobe InDesign CS5.5 is a lot better than in previous versions. Resolution of 96 or 150 PPI is usually good for various images. If your book has a lot of images, you should process them manually and optimize the size and jpeg compression for each picture.

EPUB Export Options

General
Image
Contents

Contents

Format for EPUB Content
- ◉ XHTML
- ○ DTBook

Contents
- ☑ Use InDesign TOC Style : [Default] ▽
- ☐ Break Document at Paragraph Style: [Basic P... ▽
- ☐ Place Footnote After Paragraph
- ☐ Remove Forced Line Breaks

CSS Options
- ◉ Generate CSS
  - ☑ Include Style Definitions
  - ☑ Preserve Local Overrides
  - ☑ Include Embeddable Fonts
- ○ Style Names Only
- ○ Use Existing CSS File
  - (Choose) template.css

(Cancel)   (OK)

With InDesign CS5.5 you can easily split your book into smaller xhtml files using Break Document at Paragraph Style -option. It might cause quite a lot of unnecessary breaks if your headings are separated to different lines using manual line breaks, so be sure that they are removed before ePub export.

- Embedding fonts

InDesign supports OpenType fonts in font embedding. Also, most TrueType fonts should work, according to Adobe. Often InDesign gives an error that some fonts could not be embedded. Replace these fonts with other font types (for example, from postscript to opentype) to have the fonts show up correctly. Liza Daly has published further tips on embedding fonts in epubs, see How to embed fonts in epub files. Embedded fonts work in Adobe Mobile Reader environment, including Elisa Kirja application. Some reading software and hardware doesn't recognize embedded fonts. For this reason, specify fonts also as generic family (sans-serif, serif, monospace):

```
font-family: Arial, sans-serif;
font-family: Times, serif;
```

elisa

17 December 2012

Adobe has some further tips in [Producing epub Documents from InDesign](#) for using InDesign for epub creation.
Another document worth reading is the [Epub Best Practices](#) by Adobe.

Metadata in content.opf and toc.ncx -files should be checked and edited manually if required. Adobe InDesign doesn't set metadata correctly.

Setting metadata correctly is helpful afterwards when particular book has to be found from huge amount of books. Especially setting ISBN in metadata is strongly recommended.

Here is an example of metadata from content.opf:
```
<dc:title>Sample book</dc:title>
<dc:creator opf:role="aut">Elisa Kirja</dc:creator>
<dc:date>2012-05-31</dc:date>
<dc:identifier opf:scheme="ISBN">978-010-1010-10-1</dc:identifier>
<dc:identifier id="BookId">7FA35952-57D7-4AA5-DF71-
10F9392C882D</dc:identifier>
<dc:language>fi</dc:language>
```

Here is an example of metadata from toc.ncx:
```
<meta name="dtb:uid" content="7FA35952-57D7-4AA5-9065-
10A4192B440D"/>
```

4.3   Editing epub files

Epubs are basically zip packages with a special structure, so for a technically inclined person, editing epub directly is possible. If this sounds too difficult, use some epub editing software instead.

Epub-archive exploded:
demobook
    META-INF
        container.xml – *defines the location of epub content-file*
    OEBPS
        images - *subfolders are not necessary, all files can be listed under OEBPS*
            img0001.jpg – *images which are referenced from the html*
            img0002.jpg
        styles
            style001.css – *style definitions*
        text
            content001.xhtml – *actual content file including the text*
            content002.xhtml
        content.opf – metadata *of the book, also creates references to items*
        toc.ncx – *table of contents in xml-format*
    mimetype – *text file containing the mime type of epub*

17 December 2012

To edit an epub directly, one may change the file suffix of the file form *epub* to *zip* and then uncompress it. Read some of the tutorials in the Resources list below to get a grasp of how to edit the epub. By compressing the file with zip and then changing the suffix back to epub, the file becomes again a valid epub document. Note! Mimetype file must be first in the zip file, uncompressed, with no spaces or any other characters after final character "p" (application/epub+zip). The *"Archive…"* command of Os X does not create zip files that are compatible with epubs.

4.4 Epub tuning

The resulting epub from InDesign conversion is still typically far from perfect. Further editing of the epub can be done either by using a basic text editor or a special epub editor software.

- The epub conversion process loses typically spacing of the elements, such as headers. One may use the css-style in *<style type="text/css">* for specifying the styles of the elements.

- Some elements may have wrong styles or may not be correctly identified. These can be fixed by editing the xhtml-code of the epub document manually.

- By default, InDesign gives css styles names such as "paragraph-style7". These can be changed to something more descriptive with Search/Replace for more convenient editing.

- Make the book index use hyperlinks (since the page numbers are mostly meaningless in epub, see Improving indexes in ebooks

- Always verify your work in ADE and in Elisa Kirja application.

Note! Using Calibre is not recommended, since it alters dramatically document styling!

4.5 Validating epubs

To make sure that the epub is correct in structure, the epubcheck tool should be used. There is also a web based interface for the validator. Remember that epubcheck doesn't check all flaws, like CSS-errors or file size restriction.

Make sure that you use the latest version of epubcheck tool.

elisa

17 December 2012

4.6   Testing and proofreading

Follow these steps to test that the book is working correctly.

1) Make sure that the book content is correct and shows up ok, see: Is epub "ugly"?.
Checklist:

- Line breaks are correct
- Words aren't run together
- Chapter titles and subtitles show up correctly
- Special characters are displayed properly (Scandinavian characters, dashes, fractions, drop caps)
- See that local styles work correctly (emphasized words, indentations etc)
- Check that images show up in correct places and with their captions
- Check that ordered and unordered lists show up correctly

2) Test the layout and design. Checklist:

- Are the fonts correct?
- Does text align ok (centered, left, right)?
- Do chapters start on their own pages?
- Headings aren't separated of their paragraph beginnings
- Is there enough space between paragraphs?
- Is there enough space in the beginning of chapters?
- Is there extra content in the beginning and end of the book?
- Is the cover image showing up ok?

3) Go through the epub elements. Checklist:

- Book metadata **(title, author, language, date, publisher, uuid and ISBN**) are correct
- See that the table of contents is correctly nested and shows ok

4) Test the ebook. Test the books both on an ebook reader device and in the Adobe Digital Editions desktop application to get an accurate feel of how the book will show on the various platforms. Consider testing also on other platforms and in other devices.

5) Compare the result to the printed book. Checklist:

- Is there still something missing?

6) Validate the ebook. This is important to make sure that no technical mistakes were made when editing the epub and so that it will work correctly on different devices.

17 December 2012

5   Examples of good epubs

O'Reilly ebooks are examples of good epubs. They don't use DRM, so it's easy to inspect their contents and their html and css files. Things to note:

- they have divided each chapter into sub chapters and corresponding files. One html-file is about 1-4 pages and usually 2-50KB in size.
- their html and css code is clean and human readable
- the metadata is filled so the book can be listed correctly
- images reserve the whole screen width, there's no text next to them. This looks good, and works easily with html divs and css styling
- the ebooks pass the epub validator test

6   Notes about specific epub features

The following features may be present in epub standard but may not have support in all epub reading systems, or are only supported by Adobe Digital Editions and not part of the epub standard:

- File size restrictions: epub max. size 100MB, xhtml/html chapter max. size 300KB, image (jpg/png/gif) max. size 500KB and xhtml/html file with its linked images max. size 10MB

- Image resolution restrictions: Maximum image resolution is 800 x 1280px or 1280 x 800px depending on aspect ratio.

- For full page images use CSS "width:100%" if image is wider than taller and "height:100%" if image is higher than wider. Use "max-width:100%" on images, when you want to keep them in original size, but scale them full page if they don't fit on the screen. Don't use fix-size image size definitions.

- SVG: Adobe recommends that epub covers be produced in SVG, but beware that SVG is not supported on old web-based readers. At least use jpeg or png as fallback to SVG.

- Adobe ADEPT DRM ebooks will only be supported in the Adobe Digital Editions desktop software and on the eReader and similar eink devices (like the Sony Reader). Elisa Kirja application also supports Adobe DRM ebooks on the iOS and Android devices.

- Embedded fonts: not supported on some eReaders which allow users to change their own fonts. InDesign's font embedding will not produce valid epubs. Proceed with caution.

- Text alignment: most ereaders allow users to adjust text justification. We recommend not justifying text as it doesn't produce good results on small devices without hyphenation.

17 December 2012

- CSS: small-caps not supported on ADE. If you need small-caps in your document, you must type text in UPPERCASE and adjust font-size to 80%.

- The ADE page-template (XPGT) file is only supported in ADE / Adobe Reader Mobile devices, and is not part of the epub spec. However it does not cause epubs to be invalid.

- Use "text-decoration: none; color: #000000;" style definition for "a" tags in css, if you want to remove blue underlining in hyperlinks. Doesn't work in all reading applications and devices.

- The Adobe "page-map" extension is not valid and is not part of the epub spec.

- ADE does not handle <br> tags that are children of <body>. While this is not valid XHTML, it is still extremely common usage.

- ADE doesn't support the non breaking hyphen unicode character. These should be removed from epub-files.

- ADE doesn't render <object> or <form> tag well (often crashing); those should be avoided in ebooks anyway.

- ADE doesn't handle img tags that have & in their src attribute

- Good size for body text is usually between 0.9em – 1.0em.

7 References

Full specifications of the epub standard can be found in IDPF-site: http://idpf.org/epub/201

Threepress, which provides consulting and creates software for publishers, has an online epub validator: http://threepress.org/document/epub-validate/

Listing of several different conversion tools and epub creators: http://www.lexcycle.com/faq/how_to_create_epub

A good technical introduction of epub format can be found in Epub Format Construction Guide.

Build a digital book with EPUB by Liza Daly is worth reading.

Creating eBooks: An epub Tutorial gives an step by step introduction on working with epub files on Mac Os X.

Keith Fahlgren has collected a fine list of resources related to making epubs.
Liza Daly gives some useful tips in Five ebook conversion tips