

# EPUB 3 Solutions Landscape

Bill McCoy

Executive Director, IDPF

November 30, 2013



**International Digital Publishing Forum**

International Trade and Standards Organization  
for the Digital Publishing Industry

# Agenda

- Authoring tools
- Content production tools
- Distribution services
- Open source developments

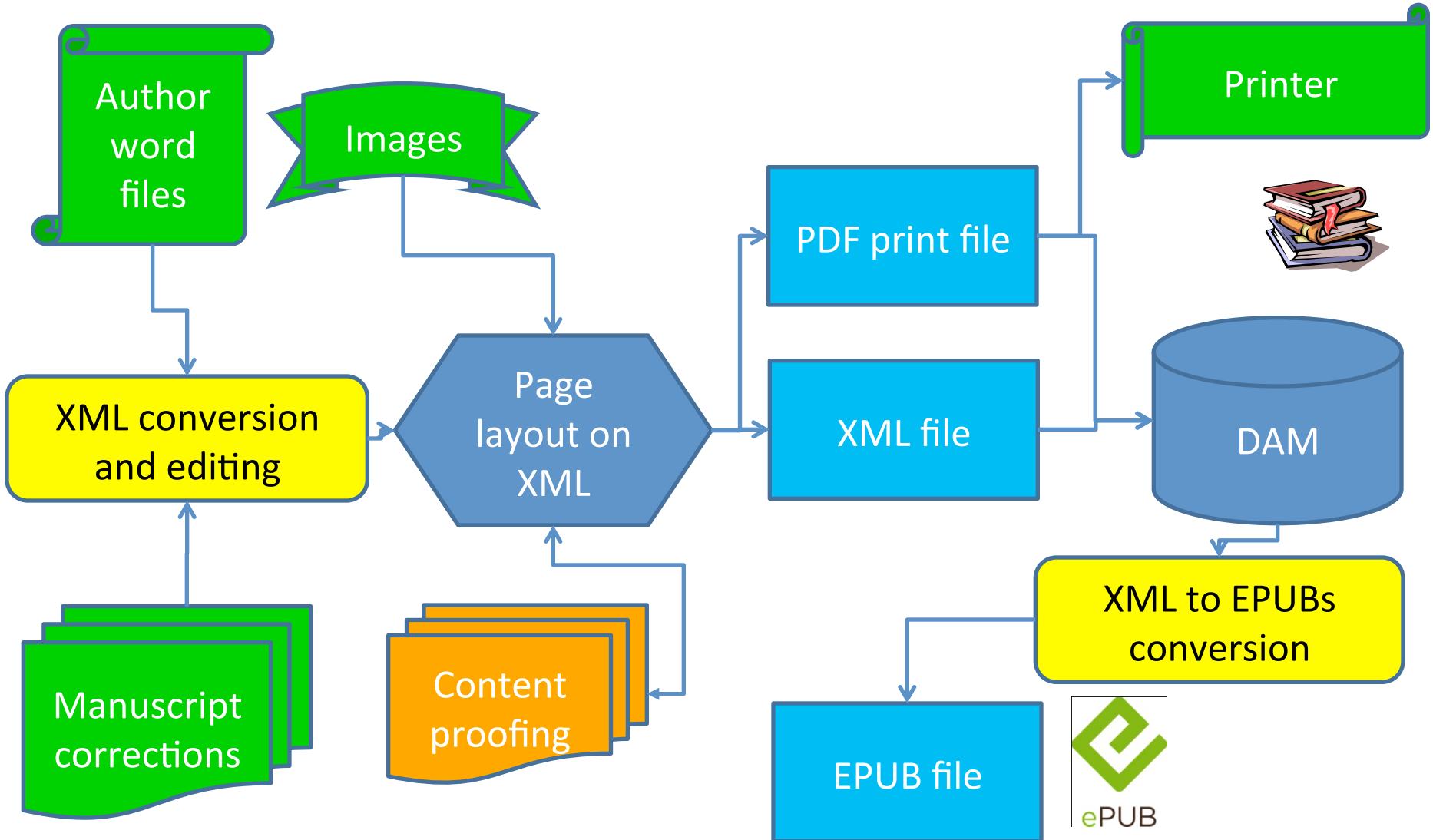
# Strategic Questions

- How much to do yourself vs. outsource to a content conversion partner?
- What inputs do you have available (InDesign files, PDFs, XML database)
- Three possible workflows
  - Print, then digital (“PDF first”)
  - Parallel development (ideally from shared asset DB e.g. “XML first”)
  - Digital, then print (“HTML first”)

# XML first in novels from books to ebooks

**Light Semantics  
Content driven**

**Hachette Livre  
start in 2001**



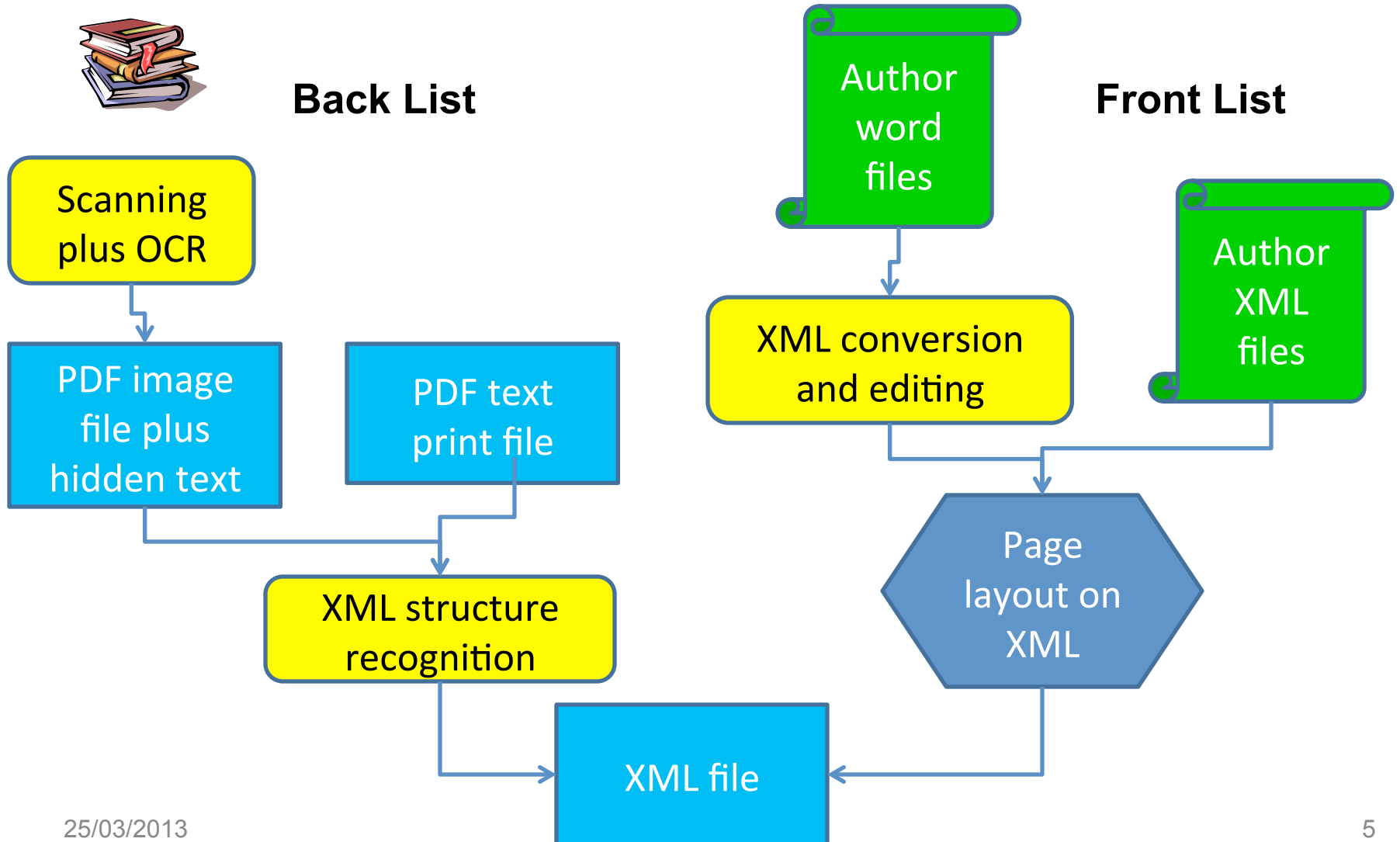
# XML process : Where do you do XML?

**Light Semantics**  
**Content driven**



**Back List**

**Front List**



# Fixed Layout production

## Layout driven

### InDesign to Fixed Layout

- InDesign Scripting and IDML export
- Conversion to HTML5 with absolute positioning



# Fixed Layout production

## Layout driven

### PDF to Fixed Layout

- SVG export with appropriate tools
- Use SVG directly in spine

08:44
94%

J'applique

### Connaître le vocabulaire de base - Choisir une représentation graphique adaptée

**EXERCICE RÉSOLU**

Une entreprise assurant le service « restauration » à bord des trains de voyageurs souhaite renforcer son offre. Elle effectue une étude statistique des distances, en km, parcourues en train par un panel de voyageurs. Le tableau ci-dessous présente les résultats de cette étude.

Distance en km	[0 ; 200[	[200 ; 400[	[400 ; 600[	[600 ; 800[
Fréquence (en %)	22	40	26	12

a) Quelle est la population étudiée ?  
 b) Quel est le caractère étudié ? Préciser sa nature.  
 c) Quelle est la représentation graphique adaptée à cette étude ?  
 d) Quelle est la tranche kilométrique de fréquence la plus élevée ?

a) La population étudiée est : les voyageurs en train.  
 b) Le caractère étudié est : les distances parcourues. Il est quantitatif continu.  
 c) La représentation graphique la plus adaptée à cette étude est l'histogramme.  
 d) La tranche kilométrique de fréquence la plus élevée est [200 ; 400[. (40 % de voyageurs).

**10** Pour prévoir les quantités nécessaires à son réapprovisionnement, une sandwicherie a réalisé une enquête sur le type de sauce accompagnant les sandwiches Kebab le samedi :

Sauce	Tomate	Mayonnaise	Blanche	Moutarde
Effectifs	55	67	35	18

a) Quelle est la population étudiée ?  
 b) Quel est le caractère étudié ? Préciser sa nature.  
 c) Quelle est la représentation graphique la plus adaptée à cette série ?

**11** Une étude statistique est représentée par le graphique ci-dessous.

a) Quelle est la population étudiée ?  
 b) Quel est le caractère étudié ? Préciser sa nature.  
 c) Peut-on représenter cette série par un diagramme en secteurs ? Justifier la réponse.  
 d) Combien d'élèves dépendent pour le cinéma entre 20 et 30 € par mois en moyenne ?

**12**

Frères et sœurs	0	1	2	3
Effectifs	124	245	268	256

Le caractère étudié est quantitatif discret. Justifier cette affirmation.  
 Peut-on représenter les résultats de cette étude par un histogramme ? Justifier.  
 c) Peut-on représenter cette étude par une autre représentation graphique ? Justifier.

**13** On donne la représentation graphique suivante : Répartition des différents types d'emplacements du camping « Les pieds dans l'eau »

a) Quelle est la population étudiée ?  
 b) Quel est le caractère étudié ?  
 c) Comment appelle-t-on ce type de représentation graphique ?  
 d) Que signifie le 47 % ?

**3 Calculer une fréquence**

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**8 Lire des informations à partir d'un graphique**

Un club de plongée compte 80 licenciés. Son administrateur a répertorié le nombre de plongées effectuées par plongeur pendant un an. Ses résultats sont indiqués à l'aide de l'histogramme ci-dessous :

Nombre de plongées effectuées par plongeur pendant un an

a) Combien de plongeurs ont effectué au moins 30 plongées ?  
 b) Combien de plongeurs ont effectué plus de 40 plongées ?

**9 Choisir une représentation graphique adaptée**

Voici les notes obtenues par un élève en sciences physiques pendant le premier semestre :

13,5 ; 12 ; 14 ; 8 ; 14 ; 7 ; 10.

Quelle(s) est(sont) la(les) représentation(s) graphique(s) la(les) mieux adaptée(s) pour représenter les résultats de cet élève ?

**2** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**3** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**8** Lire des informations à partir d'un graphique

Un club de plongée compte 80 licenciés. Son administrateur a répertorié le nombre de plongées effectuées par plongeur pendant un an. Ses résultats sont indiqués à l'aide de l'histogramme ci-dessous :

Nombre de plongées effectuées par plongeur pendant un an

a) Quel est le pourcentage d'élèves portant des baskets Puma ?  
 b) Est-il vrai que 37,5 % des élèves portent des baskets Adidas ?

**9** Choisir une représentation graphique adaptée

Voici les notes obtenues par un élève en sciences physiques pendant le premier semestre :

13,5 ; 12 ; 14 ; 8 ; 14 ; 7 ; 10.

Quelle(s) est(sont) la(les) représentation(s) graphique(s) la(les) mieux adaptée(s) pour représenter les résultats de cet élève ?

**2** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**3** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**8** Lire des informations à partir d'un graphique

Un club de plongée compte 80 licenciés. Son administrateur a répertorié le nombre de plongées effectuées par plongeur pendant un an. Ses résultats sont indiqués à l'aide de l'histogramme ci-dessous :

Nombre de plongées effectuées par plongeur pendant un an

a) Quel est le pourcentage d'élèves portant des baskets Puma ?  
 b) Est-il vrai que 37,5 % des élèves portent des baskets Adidas ?

**9** Choisir une représentation graphique adaptée

Voici les notes obtenues par un élève en sciences physiques pendant le premier semestre :

13,5 ; 12 ; 14 ; 8 ; 14 ; 7 ; 10.

Quelle(s) est(sont) la(les) représentation(s) graphique(s) la(les) mieux adaptée(s) pour représenter les résultats de cet élève ?

**2** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**3** Calculer une fréquence

Le Parthénon s'inscrit dans un rectangle doré, c'est-à-dire tel que le rapport de la longueur à la hauteur est égal au nombre d'or :  $\phi$ .

$\phi = 1,6180339887$

**8** Lire des informations à partir d'un graphique

Un club de plongée compte 80 licenciés. Son administrateur a répertorié le nombre de plongées effectuées par plongeur pendant un an. Ses résultats sont indiqués à l'aide de l'histogramme ci-dessous :

Nombre de plongées effectuées par plongeur pendant un an

a) Quel est le pourcentage d'élèves portant des baskets Puma ?  
 b) Est-il vrai que 37,5 % des élèves portent des baskets Adidas ?

**9** Choisir une représentation graphique adaptée

Voici les notes obtenues par un élève en sciences physiques pendant le premier semestre :

13,5 ; 12 ; 14 ; 8 ; 14 ; 7 ; 10.

Quelle(s) est(sont) la(les) représentation(s) graphique(s) la(les) mieux adaptée(s) pour représenter les résultats de cet élève ?

# Fixed Layout production

## Layout driven

### PDF to Fixed Layout Images

- High quality image export form PDF
- Embed images in HTML5 files





# Unstructured PDF

- ❖ PDF *does* encode precise, fixed-layout drawing instructions for glyphs, shapes and images, enabling a consistent appearance
- ❖ PDF *does not* (usually) encode reading order, words, paragraphs, columns, lists, tables, stories, chapters, text roles, shape/image groupings, shape/image roles, etc.

# Structure Reconstruction

Analysis	Structure
Border shapes	Tables, Callouts, Articles
Clusters of glyph positions	Words
Clusters of word positions	Alignment
Vertical white strips	Line breaks, Columns, Lists
Text line spacing & alignment	Paragraphs
Shape grouping	Diagram

# Authoring EPUB 3

- There is not and may never be a “one size fits all” EPUB authoring tool
- Authoring EPUB is creating a special kind of website
  - Infinite range of potential expression
  - Wide range of suitable approaches
  - Specialized tools may be best for piece parts (widgets, animations)
- EPUB reading system implementations have differing levels of feature support

# Classifying EPUB 3 Authoring Tools

- EPUB native - EPUB is output format
- WYSIWYG - code-level – batch (converter)
- Reflow - fixed-layout
- WP-ish - interactivity-centric
- CMS/cloud-based - standalone
- Workflow-oriented = single-user
- (Available to all-comers vs. requires a business relationship)

# Adobe InDesign CC

- EPUB is output format
- WYSIWYG only
- Reflow only (for EPUB - app is oriented for complex page designs)
- WP-ish, standalone, single-user

# Blue Griffon EPUB Edition

- EPUB native (arbitrary input, round-trip)
- WYSIWYG and code-level
- Reflow only
- WP-ish, standalone, single-user

# Inkling Habitat

- EPUB is output format
- WYSIWYG
- Reflow
- WP-ish
- CMS/cloud-based, Workflow-oriented

# AerBook

- EPUB is output format
- WYSIWYG
- Fixed-layout
- Interactivity-centric
- CMS/cloud-based, multi-user but not workflow-oriented



# PressBook

- EPUB is output format
- WYSIWYG
- Reflow
- WP-ish (WordPress-based)
- CMS/cloud-based, multi-user but not workflow-oriented

# AerBook vs. PressBook

- Commonalities
  - EPUB is output format
  - WYSIWYG
  - CMS/cloud-based, multi-user but not workflow-oriented
- Differences
  - Reflow (PressBook) vs. Fixed (AerBook)
  - WP-ish (PressBook) vs. Interactivity-centric (AerBook)

# Oxygen XML Editor

- EPUB is output format
- Code-level only
- Reflow, WP-ish (fixed-layout / interactivity could be coded better in other apps)
- Standalone, single-user

# Magic EPUB

- EPUB is output format
- Batch converter (PDF input)
- Fixed-layout only (various options)
- CMS/cloud-based, some workflow features

# Flipick

- EPUB is output format
- WYSIWYG, Fixed-layout, interactivity-oriented
- Standalone, single-user (InDesign plug-in)

# PubCoder

- Demo

# Authoring Tool Observations

- Lots of EPUB 3 tools, but still significant capability gaps
- Situation evolving rapidly; best-of-breeds likely to be more clear in next 6 months (+ more EPUB 3 RS support)
- Track **#eprdctn** (note: that crew is hand-coding centric)
- Seek to streamline/focus on **your** unique content requirements
- A much better idea to “hire” experts via their tools rather than go with hand-coding by web developers
  - Unless you have truly unique content/experience requirements and the ROI pencils out
- Solutions directory will be available on new EPUB 3 microsite by EOY

- HTML 5
- CSS 2.1
- CSS 3 Selectors
- CSS 3 Media Queries
- CSS 3 Text
- CSS 3 Backgrounds and Borders
- CSS 3 Colors
- CSS 3 2D Transformations
- CSS 3 3D Transformations
- CSS 3 Transitions
- CSS 3 Animations
- CSS 3 Multi-Columns
- CSS Namespaces
- SVG 1.1
- WAI-ARIA 1.0
- MathML 2.0
- ECMAScript 5
- 2D Context
- WebGL
- Web Storage
- Indexed Database
- Web Workers
- Web Sockets Protocol/API
- Geolocation
- Server-Sent Events
- Element Traversal
- DOM Level 3 Events
- Media Fragments
- XMLHttpRequest
- Selectors API
- CSSOM View Module
- Cross-Origin Resource Sharing
- File API
- RDFa
- Microdata
- WOFF
- HTTP 1.1 part 1 to part 7
- TLS 1.2 (updated)
- IRI (updated)
- ...



# EPUB 3 Content Production Tools

- Content validation
  - EPUBCheck (<http://validator.idpf.org>)
- Content operability check (preflight)
  - DIY using iBooks, Kobo, Readius, ...
  - “CanIUse” for EPUB <http://epubtest.org>
- Convert to Mobi/KF8 format
  - KindleGen

# Distribution and Conversion Services

- Most larger US/EU publishers presently outsourcing most EPUB conversion/enhancement
  - Conversion specialists often assemble customized tool chains per content type
  - Ex: *MS Word* -> OpenOffice Writer -> *HTML* -> custom assembly tool -> *EPUB*
- Many providers offer to host content and offer downstream distribution services
  - Direct online retailing (w/ or w/out DRM)
  - Distribution to 3<sup>rd</sup>-party retail channels
  - P.O.D.
  - Integrated content editing/proofing

Thank you.

bmccoy@idpf.org  
@billmccoy



International Digital Publishing Forum