

EPUB3.0 の仕様と日本語組版

2011/10/07

むらかみ しんゆう
村上 真雄



Twitter: [@MurakamiShinyu](https://twitter.com/MurakamiShinyu)
murakami@antenna.co.jp

1. はじめに～自己紹介

- 1990 年ごろ、[テキスト整形ツール「XTR」](#)を開発しフリーソフトウェアとして公開。それ以降、組版・印刷とマークアップ言語(SGML、XML、HTML)、スタイルシート言語(XSL、CSS)に関心を持つ。
- 1999 年、XML 自動組版ソフトの開発を[アンテナハウス](#)に企画提案、Antenna House Formatter として製品化。
- [AH Formatter](#) は、多言語の大量の XML データからの自動組版などで威力を発揮して、けっこう世界で使われています。
- XSL および CSS の仕様とその実装に携わり、その関係で電子書籍フォーマット EPUB の日本語組版拡張に関わる。

2. EPUB って何？

- リフロー型（ページが固定でない）電子書籍の世界標準仕様
- IDPF という標準化団体が策定
 - ◆ IDPF = International Digital Publishing Forum
 - ◆ 世界の出版社や ITC 企業を中心に構成される
- 文書内容は XHTML、レイアウトは CSS で指定（ウェブの標準技術を応用）、それに書籍の情報や目次データを合わせて ZIP 形式で圧縮したもの
 - ◆ W3C という標準化団体が、(X)HTML と CSS の仕様を策定
 - ◆ (X)HTML は 5 がこれからの標準。CSS は 2.1 が標準、それに拡張機能を追加するレベル 3（CSS3）の仕様がモジュール別に標準化が進む
- EPUB2.0 は、XHTML1.1 + CSS2 だった。縦書きが出来ないなど、日本語組版で問題があった。
- EPUB 仕様改訂 → EPUB 3.0 へ（HTML5 と CSS3 縦書き仕様など採用）

3. EPUB EGLS (EPUB の世界言語対応の動き)

- JEPA (日本電子出版協会) EPUB 研究会 (2009 年 11 月発足)
 - ◆ 技術主任: 村田 真 (@[muratamakoto](#): 標準化のエキスパート)
 - ◆ 2010 年 4 月、[EPUB 日本語組版要求仕様案](#)を公開
 - ◇ W3C JLTF [「日本語組版処理の要件」](#)を参考としている
 - ◇ IDPF の [EPUB 改訂の課題リスト](#)に組み込まれる
 - ◆ 2010 年 6 月 [「EPUB 仕様の日本語組版拡張を目指して」](#) (村田+村上)
- IDPF EPUB 改訂作業グループのサブグループとして、EGLS グループ
 - ◆ EGLS = Enhanced Global Language Support
 - ◆ コーディネーター: 村田 真
 - ◆ 日本、台湾、韓国、米国などのメンバー
 - ◆ 2010 年 8 月 札幌会議、10 月 台北会議、など
 - ◆ EPUB の縦書きは CSS3 縦書き仕様で実現することに

4. IDPF (EPUB) と W3C (CSS) の連携

- 2000 年ごろブラウザ IE5.5 で縦書きが実装されていた。当時マイクロソフトは W3C に CSS の縦書きなど国際化テキスト仕様の提案をしていた。
- その仕様提案に深く関わっていたのが、当時 MS Word の開発を担当していた石井宏治さん（そのあと MS を退社、独立）
- W3C の CSS3 テキストと縦書き仕様の担当者 Erika(Fantasai)、ボランティアを申し出てくれた石井さんと共同作業（2010 年 9～）。専門家たちが集結。
- EPUB と CSS の縦書き仕様推進への支援体制ができて、IDPF (EPUB) と W3C (CSS) の連携を後押し
 - ◆ 総務省 平成 22 年度新 ICT 利活用サービス創出支援事業「EPUB 日本語拡張仕様策定」（イースト、アンテナハウス、JEPA が共同提案）
 - ◆ 国際標準化戦略に関する検討チーム「次世代ブラウザのテキストレイアウトに関する検討会」（総務省）

5. EPUB3.0 はこうなった

- 最新の Web 標準技術を利用 : (X)HTML5 と CSS2.1 + CSS3 (一部)
- [CSS3 Writing Modes](#) : 縦書き、縦中横
 - ◆ まだドラフトのためプロパティ名に -epub- プレフィックス付きで採用 :
-epub-writing-mode など
- [CSS3 Text](#) : 禁則、行揃え、圏点など
 - ◆ 同じく -epub- プレフィックス付きで採用 : -epub-text-emphasis など
- ルビ (HTML5 ruby)
- 段組 (CSS3 Multi-column)
- 外字は? まず Unicode です。Unicode に無い文字はフォント埋め込みで
- SVG (ベクターグラフィック)、MathML (数式)
- マルチメディア (音声、動画)
- DAISY との連携 : だれにでもアクセシブルな書籍・教科書へ

6. EPUB3.0 で使えるようになった CSS3 の日本語組版機能の例

- [CSS3 Writing Modes](#) : 縦書き、縦中横

縦書きの指定 :

```
html { -epub-writing-mode: vertical-rl }
```

縦書きで欧文や数字を直立に (デフォルトは横倒し) :

```
html { -epub-text-orientation: upright }
```

縦中横の指定の例 :

```
span.縦中横 { -epub-text-combine: horizontal }
```

- [CSS3 Text](#) : 圏点、禁則、行揃えなど

圏点の指定の例 :

```
em.ゴマ点 { -epub-text-emphasis: sesame }
```

```
em.蛇の目 { -epub-text-emphasis: double-circle }
```

圏点
ゴマ点
蛇の目
など

EPUB3.0
で縦中横

たてちゅうようこ

強い禁則の指定：

```
html { -epub-line-break: strict }
```

弱い禁則の指定：

```
html { -epub-line-break: loose }
```

□ ルビ (HTML5 ruby)

グループルビの例：

```
<ruby>秋桜<rt>コスモス</rt></ruby>
```

熟語ルビの例：

```
<ruby>秋<rt>あき</rt>桜<rt>ざくら</rt></ruby>
```

モノルビの例：

```
<ruby>秋<rt>あき</rt></ruby>の<ruby>桜<rt>さくら</rt></ruby>
```

秋あき
のの
桜ざくら
はは
秋あき
桜ざくら
かか
秋コス
桜モス
か。
。

7. EPUB3.0 仕様には含まれなかった日本語組版に関すること

EPUB3.0 では日本語組版に最低限必要な機能が入ったということで、まだ課題は残っている。

EPUB 仕様に取り入れられるためには、CSS 仕様の標準化が進むことが必要。

□ 文字組みアキ量設定に相当：CSS3 Text の [text-spacing](#) プロパティ

- ◆ 和欧文間スペース：ABC伊呂波123 ⇔ ABC 伊呂波 123

```
text-spacing: ideograph-alpha ideograph-numeric
```

- ◆ 連続する約物の処理：

「『約物（括弧など）』。」 ⇔ 「『約物（括弧など）』。」

- ◆ 行頭と行末の約物の配置（全角／半角など）

```
text-spacing: trim-start /* 行頭約物を半角取りに */  
               trim-end   /* 行末約物を半角取りに */
```

□ 句読点のぶら下げ等 : CSS3 Text の [hanging-punctuation](#) プロパティ

◆ 句読点のぶら下げ :

```
hanging-punctuation: allow-end /* ぶら下がり有り */  
hanging-punctuation: force-end /* ぶら下がり強制 */
```

◆ 段落 1 字下げ・起こし食い込み :

```
text-indent: 1em; hanging-punctuation: first
```

□ 自動縦中横 :

CSS3 Writing Modes の [text-combine-horizontal](#) プロパティ

```
text-combine-horizontal: digits 2 /* 2桁までの数字を自動で縦中横に */
```

□ グリッドにもとづく行の配置の機能

◆ 行取りの指定ができること

◆ 行長が字送りの整数倍になるようにすること

いまの仕様では行長は画面サイズとマージンにより決まるので、字詰めが半端になり、行揃えで文字間があく。(一昔前の DTP でもあった問題)

□ 論理方向プロパティ

CSS 仕様では「margin-left」は横書きでも縦書きでも絶対方向で左側を空ける指定になるため、横書きなら左のインデントだが縦書きではブロック（見出しや段落）と次のブロックとの間のアキの指定ということになる。縦書きではインデントの指定に「margin-top」としなければならない。

このため、縦書き用のスタイルシートと横書き用のスタイルシートの両方を用意するのに手間がかかる、また、縦書き用にレイアウトしたものを縦書き非対応の環境で見たときにレイアウトが大きく崩れるという問題がある。

これを解決するため、縦書きでも横書きでも共通に使える論理方向のキーワード start（行頭側）、end（行末側）、before（前側）、after（後側）を使う論理方向プロパティ仕様が提案されている。「margin-start」で、横書きでも縦書きでも行頭側を空けるインデントとするなど。

組み方向によって margin 指定など変えなくてはならない例

横書き

標題

段落。これは段落です。これは段落です。これは段落です。これは段落です。

引用。これは引用です。これは引用です。これは引用です。

段落。これも段落です。これも段落です。これも段落です。

縦書き

標題

段落。これは段落です。これは段落です。これは段落です。これは段落です。

引用。これは引用です。これは引用です。これは引用です。

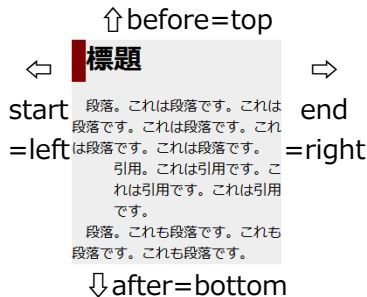
段落。これも段落です。これも段落です。これも段落です。

```
h1 {
  border-left: 12pt solid;
  margin-bottom: 12pt;
}
blockquote {
  margin-left: 3em;
}
```

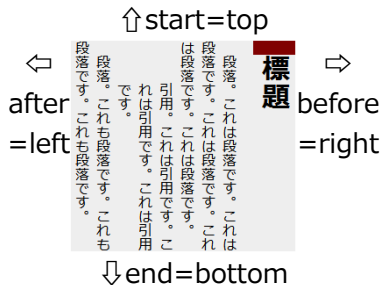
```
h1 {
  border-top: 12pt solid;
  margin-left: 12pt;
}
blockquote {
  margin-top: 3em;
}
```

論理方向プロパティが使えるら

横書き



縦書き



```
h1 {
  border-start: 12pt solid;
  margin-after: 12pt;
}
blockquote {
  margin-start: 3em;
}
```

同じ margin/border/padding プロパティを両方の組み方向で使うことができる。

実装によってプレフィックス付きで今でも使えます（非標準ですが）。
-webkit-margin-start など。

8. 電子書籍の組版をよいものにするために

- 美しく読みやすい組版の技術は、電子書籍の組版にも引き継がれるべきもの
- 電子書籍だからこの程度の品質、ということはなくっていく
 - ◆ 表示デバイスの高精細化
 - ◆ 処理速度の向上 → 高度な組版処理が、リフロー型の電子書籍でも可能に
- 必要なことは
 - ◆ EPUB や CSS の組版に関する標準仕様が進み、それが実装されること
 - ◆ 制作者側が、電子書籍の特性を理解して、ふさわしい作り方をすること
- 組版の世界が、活字→電算写植→DTP と移り変わっていったとき、移行期には、新しい道具も未熟、それを使う人も未熟で、組版の質がよくなかったりした。電子書籍でも同じことがいえる。
- 道具（EPUB や CSS の標準仕様と、その実装）が良くなっていかなくてはならないし、制作者側も道具をよりよく使うために学んでいかなければならない。

- 道具を良くしていくには、利用者からのフィードバックが欠かせない。
- そのための集まりとして6月に東京と京都にて「[次世代電子出版と Web 表現技術フォーラム](#)」が開催された。(主催：次世代 Web ブラウザのテキストレイアウトに関する検討会、後援：W3C/慶應・総務省)。このような取り組みが継続発展する必要がある。
- W3C 標準化はオープンであり、だれでも意見を出すことができる。HTML5 と CSS3 仕様に関する日本語のメーリングリストもあります：
[HTML5 Japanese Interest Group](#)

9. リフロー型の電子書籍の特性を生かした本作りを

□ 作り手だけで全てをコントロールできない

- ◆ 1 行の文字数、1 ページの行数は、固定でない
- ◆ どう表示されるかは、作り手、ビューア（デバイス、ソフトウェア）、読み手の 3 者とも関係する。
- ◆ よい EPUB ビューアは、作り手の意図を尊重しながら、美しく読みやすく表示するための工夫があって、読み手の好みやニーズによって表示の仕方が選べるものになるだろう。

□ アクセシビリティが重要

- ◆ 大きな文字にしても読みやすいように。（文字を大きくすれば 1 行の文字数が少なくなることには注意）
- ◆ 音声読み上げで読めるように。

- ◇ 漢字にルビをふるのに、ルビで小書きの仮名を使わないという従来の組版ルールでは、正しく読みが分からない問題がある。

例：「不^ふ詳^{しょう}」と「不^ふ使用^{しょう}」

データでは小書きの仮名を入力すべき。

```
<ruby>不詳<rt>ふしょう</rt></ruby>
```

表示の設定で、小書きの仮名を大きな仮名に変換して表示することができる。

そのための指定が EPUB3.0 で可能

```
rt { text-transform: -epub-fullsize-kana }
```

◆ 縦書き／横書きの両方に対応すること

- ◇ 読字障害で、横書きは読めるが縦書きは読むのが難しい（あるいはその逆）がある。縦書き横書きを切り替えられることは意味がある。
- ◇ EPUB3.0 では縦書き用と横書き用の両方のスタイルシートを設定することができる。

10. 見出しやリードの行の折り返しの工夫

- 固定の行長を前提として強制改行で行の折り返しをしていると、行長が短いとき（小さな画面あるいは大きな文字にしたとき）、読みにくくなる。

見出しやリードの
行の折り返しの工夫 → 見出しやリード
の
行の折り返しの
工夫

- そうならないようにするには、強制改行による行の折り返しは避ける。
また、次の方法もありえる：

- ◆ 空白文字などでしか自動改行しない設定にする（-epub-word-break: keep-all）。そして、改行してもよいところに幅ゼロの空白文字（U+200B ZERO WIDTH SPACE）を入れる。（自動でできるとよい）

- いずれ、自動的に単語あるいは文節区切りだけで改行する設定が可能になるかもしれない（CSS3 Text に word-break: keep-word の提案があった）。

日本語組版は両端揃えが基本であるが、文節区切りなどで自動改行してラグ組みにする方法もありえる。そのように組んだ例：

わがはい 吾輩は猫である。	しかもあとで聞くと	なかったから別段
名前はまだ無い。	それは書生という	恐しいとも
どこで生れたかとなんと けんとう 見当がつかぬ。何でも	人間中で一番 ^{どうあく} 獰悪な 種族であったそうだ。	思わなかった。ただ彼の てのひら 掌に載せられて
薄暗いじめじめした所で ニャーニャー泣いていた	この書生というのは時々 我々 ^{つかま} を ^に 捕えて煮て	スーと持ち上げられた時 何だかフワフワした
事だけは記憶している。	食うという話である。	感じがあった
吾輩はここで始めて	しかしその当時は	ばかりである。
人間というものを見た。	何という考も	

11. 紙の本と電子書籍を同時に作るために

- 従来の DTP で紙の本を作り、そのデータを再利用して電子書籍にすることが多いでしょう。InDesign など DTP ソフトの EPUB 出力機能も今後充実していくと思われます。
- 新しい本の作り方もはじまっています。

ひとつのソース（形式は XML や HTML などタグ付きテキスト）から、EPUB の生成、自動組版で印刷のための PDF 生成をする方法です。

そのひとつの例として、[アンテナハウスの取り組み](#)を紹介します。

◆ [クラウド型汎用書籍編集・制作サービス CAS-UB](#)

"Cloud Authoring Service for Universal Book"

- ◇ Web 上で EPUB 電子書籍と印刷のための PDF の両方同時の制作ができるサービスです。コンテンツはクラウドサーバー上で管理され、編集者や著者のチームでの本づくりができます。

- ◇ CAS-UB の PDF 出力は、AH Formatter による自動組版です。
- ◆ AH Formatter は、XML や(X)HTML のコンテンツを、スタイルシートによるレイアウトの指定に従って自動組版するものです。
次の2つのスタイルシート標準仕様をサポートしています。

XSL (Extensible Stylesheet Language)

XML 文書を組版するための標準規格として世界で使われています。

CSS (Cascading Style Sheets)

ウェブ標準のスタイルシート仕様で、これまで説明したとおり EPUB 仕様もこれにもとづいています。主に(X)HTML 用ですが、XML のレイアウトにも使用可能です。ブラウザの画面表示だけでなく、ページ媒体の組版のための高度な機能も CSS3 ドラフト仕様があり、AH Formatter はそれを実装しています。

- CSS という標準のレイアウトの指定方法で、どの出力にも対応できるようになるということ
 - ◆ ウェブ用のスタイル、EPUB 用のスタイル、印刷用のスタイルは、ある程度は共通にできても、それぞれの媒体の特性に合ったものに変える必要はあるでしょう
 - ◆ CSS 仕様の標準化が進み、EPUB 仕様もそれとともに進化して、ウェブでも電子書籍でも、より美しく読みやすい組版が標準になることを期待。(AH Formatter は CSS3 ドラフト仕様の先行実装や独自拡張により CSS 組版を使えるものになっているのですが)

12. おわりに～参考情報

- epub café <http://www.epubcafe.jp/>
 - ◆ 総務省「電子出版環境整備事業」、「EPUB 日本語拡張仕様策定」のすべての成果物を公開。
「[EPUB 日本語拡張 コンテンツ制作者向けチュートリアル](#)」など
- IDPF <http://idpf.org/> ・ [EPUB3.0 仕様](#)が公開されてます。
- W3C CSS current work & how to participate
<http://www.w3.org/Style/CSS/current-work>
 - ◆ EPUB3 日本語組版に関係する [CSS3 Text](#) や [Writing Modes](#) など、最新のドラフト仕様があります。

私も、ツイッター [@MurakamiShinyu](#) や「[CSS 組版ブログ](#)」で、情報発信しています。どうぞよろしくをお願いします。

ありがとうございました。