

出版システムのワンソースマルチユース

— **HTML5+CSS** 組版について —

村上 真雄 (Shinyu MURAKAMI)

株式会社ビブリオスタイル (Vivliostyle Inc.)

for better styling on web, ebooks and
print books with open web standards

目次

1. ワンソースマルチユース出版の必要性
2. 文書構造と体裁の分離
3. スタイルシート仕様 CSS の進化
4. HTML5+CSS の出版への活用と課題
5. Vivliostyle HTML5+CSS 組版の取り組み
6. まとめ

ワンソースマルチユース 出版の必要性

ひとつのソースで、いろいろな媒体に展開 =すべての人にアクセシブルに

- 普通の印刷本
- 大活字本
- 白黒反転
- 音声
- 点字
- 電子書籍
- Web
- などなど

文書構造と体裁の分離

ワンソースマルチユース出版の基本

= 文書構造と体裁の分離

- 文書構造は共通に、体裁（スタイル）はそれぞれの出力媒体用に調整できるようにする。
- そのための技術として、
 - マークアップ言語（HTML など）
 - スタイルシート（CSS など）

マークアップ言語の標準：XML と HTML

- 1980年代に生まれた SGML（標準汎用マークアップ言語）から進化して：
 - XML（拡張可能マークアップ言語）1998年に W3C 勧告
 - さまざまな応用がされている
 - オフィス文書形式、数式、組版ソフト、などなど
 - HTML（ハイパーテキスト・マークアップ言語）
 - XHTML（XML 構文の HTML）
 - HTML5（最新版の HTML）2014年に W3C 勧告
 - 文書構造を表現しやすく（**section** 要素など）
 - 数式やベクターグラフィックが標準で埋め込める（**math, svg**）
 - 動画などマルチメディア埋め込みが標準に（**video, audio**）
 - あらゆる端末で動くアプリのプラットフォーム

簡易マークアップ

- XML や HTML のタグを書くのはめんどろ、そこで、より簡単に
- 代表的な簡易マークアップは `markdown` (マークダウン)
 - ほぼ普通のテキストで、見出しなどに簡単な印をつけるだけ
 - HTML への変換が簡単
 - いろいろな Web のシステムでのテキスト入力に使われている

markdown 記法の例	対応する HTML
# 見出し1	<code><h1>見出し1</h1></code>
## 見出し2	<code><h2>見出し2</h2></code>
### 見出し3	<code><h3>見出し3</h3></code>
* リスト	<code>リスト</code>
> 引用	<code><blockquote>引用</blockquote></code>

Web 標準化団体 W3C について

- マークアップ言語の標準仕様 XML や HTML、スタイルシート仕様 CSS、その他 Web の各種技術の標準化を行っている国際的な非営利団体 **World Wide Web Consortium**、略称 **W3C** (ダブリュースリーシー)。
- Web の発明者であるティム・バーナーズ＝リーが1994年に創設。
- ブラウザベンダーや大手 IT 企業はじめ世界中から参加 (現在約420社)。
- 仕様策定の議論は公開され、メーリングリストなどで誰でも意見できる。
- EPUB など電子出版の技術の標準化については、**国際電子出版フォーラム (IDPF)**が W3C 標準を利用する形で進められてきたが、両方の標準化団体は密接に連携していて、最近では、両組織の統合 (2017年1月を目指して、IDPF を W3C に統合する計画) が検討されていることが発表されている。



スタイルシート仕様 **CSS** の進化

CSS = Web 標準のスタイルシート仕様

Web の初期の時代（「ネットスケープ」がブラウザの主流になる1995年ごろ）、本来は文書構造を表す意味がある HTML のタグが濫用され、

- 見出しでないのに文字の大きさを変えるために見出しタグを使う
- テーブル（表）ではないのにレイアウトの目的でテーブルタグを使う

などが広まり、

- 音声読み上げで読むのが難しくアクセシビリティが欠如、
- 特定のブラウザでうまく表示されても別のブラウザ環境では崩れる

など問題だった。そこで、CSS（Cascading Style Sheets）が提唱される。

- CSS（カスケーディング・スタイルシート）の名前の由来は、複数のスタイルの指定を重ね合わせること（カスケーディング）から。
- 各ブラウザ（あるいはリーディング・システム）には、デフォルトのスタイルシートが組み込まれているために、コンテンツの制作者がスタイルシートを指定しなくても、デフォルトのスタイルシートでコンテンツが表示される。たとえば、デフォルトのスタイルシートに次のようにあることで、見出し **h1** が本文の2倍の文字サイズで表示される：

```
h1 { font-size: 2em; }
```

- コンテンツの制作者は、スタイルシートを指定して、デフォルトのスタイルを上書きすることができる。例えば、

```
h1 {  
  font-size: 50pt;  
  font-family: "ユニークで豊かな書体";  
}
```

- ブラウザのデフォルト・スタイルシートと制作者のスタイルシートのほかに、ユーザー（閲覧者）の側でもスタイルシートを指定するしくみがあり、これにより、ユーザーによって読みやすいフォントで読むことなどできる。例えば、

```
h1 {  
    font-family: "私の読みやすい書体" !important;  
}
```

※**!important** の指定は、制作者の指定よりもこちらを優先させるため。

CSS 以外の「スタイルシート」について

- 文書の要素とスタイルを関連付ける「スタイルシート」機能は、CSS 誕生より前からワープロや DTP ソフトに存在して、次のようなことができる：
 - 「見出し1」「見出し2」「見出し3」など定義済みのスタイル名にそれぞれの体裁を設定して文書全体の見出しの体裁など一括で変えられる。
 - HTML で **h1**, **h2**, **h3** などのタグを使い、CSS でそれらの体裁を指定することに相当する。
 - 「段落スタイル」や「文字スタイル」を名前を付けて定義して、文書中のその特別なスタイルにしたいところに適用する。
 - HTML のブロック要素やインライン要素に **class** 属性でクラス名を与え、CSS でその体裁を指定することに相当する。

※ワープロや DTP ソフトでこの「スタイルシート」機能を活用することは、HTML・CSS への変換・移行をしやすくするために有用である。

- W3C で標準化されたスタイルシート仕様には、CSS のほかに、XML 用のスタイルシート言語として XSL (Extensible Stylesheet Language) がある。

XSL 仕様は、XML 文書の構造を変換する XSLT (XSL Transformations) と、組版を行うための XSL-FO (XSL Formatting Objects) とに分かれる。

このしくみを使う XML 組版の処理は次のようになる：

- XML 文書 + XSLT(変換) ⇒ (XSL-FO 組版) ⇒ PDF・印刷

また、XSLT は、XML 文書を別形式（別の XML 形式や、HTML）に変換するためのツールとして有用である：

- XML 文書 + XSLT(変換) ⇒ HTML (+CSS で各出力媒体へ)

※W3C は、XSL-FO 仕様の標準化作業はすでに終了していて、W3C でのスタイルシートについての標準化は CSS に集中している。

進化する CSS 仕様。「CSS 3」とは

現在最新の HTML 仕様が HTML5 と呼ばれているように、最新の CSS 仕様は「CSS 3」と呼ばれることが多い。しかし、HTML の「5」はバージョン番号であるのに対して、CSS の番号は、基本的な機能か、より高度な機能かを表す「レベル(Level)」である。

CSS レベル2 (1998年に W3C 勧告、2011年にその改訂版の CSS 2.1) まではひとつの仕様書にまとまっているが、より高度な機能や新しい機能はモジュールごとに仕様書が作られて、それぞれ独立にレベル番号が付与される。「CSS3」という用語は、それら新しい CSS 仕様を大雑把に指して使われるバズワードであり、実際の CSS モジュールに付与されているレベルは「レベル4」のもの (例: [Selectors Level 4](#)) もあれば「レベル1」のもの (例: [CSS Ruby Level 1](#)) もあり、その仕様の完成度はまちまちである。現時点で完成している CSS 仕様は「[CSS Snapshot 2015](#)」という文書にまとめられている。

日本語組版関連 CSS 標準化

- 「W3C 技術ノート 日本語組版処理の要件 (JLReq)」
 - W3C 標準仕様に日本語組版に必要な要件を反映させるために、W3C と日本語組版の専門家からなる合同チーム「W3C 日本語組版タスクフォース」によってまとめられた文書（略称「JLReq」）。英語版と日本語版が同時に作られている。
 - JLReq 第1版が2009年、第2版が2012年に公開された。
 - この文書は W3C 標準仕様（とくに CSS 仕様）に日本語組版処理に関係する機能を反映させるために参照されているほか、ブラウザ開発者をはじめ世界の開発者たちが日本語組版処理を実装するのに役立っている。
 - 電子出版の標準として、縦書きはじめ日本語組版に必要な機能が入った EPUB 3.0 の仕様が作られたときの要件としても JLReq が使われた。

- W3C では JLReq に続き、世界中の言語の組版要件をまとめつつあり、すでに 欧文、韓国語、中国語、インド系言語、チベット語、アラビア語、エチオピア語の各言語の組版要件のドラフト文書がある。
- 日本語や東アジア言語の組版にとくに関係する CSS 仕様
 - CSS Writing Modes : 縦書き
 - CSS Text : 禁則処理など
 - CSS Text Decoration : 圏点、傍線など
 - CSS Ruby : ルビの体裁
- 「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」
 - 日本からの標準化支援活動として、電子書籍端末メーカー、印刷事業社、通信事業者などの民間企業等により構成され、総務省が支援。
 - CSS 縦書き関連仕様と実装の課題整理、実装とテスト作成支援、CSS 縦書きを利用した Web コンテンツの普及促進などの活動。

HTML5+CSS の出版への 活用と課題

Web の標準 = 電子出版の標準へ

- 電子出版の標準形式である EPUB は、コンテンツとその体裁の表現に Web の標準である HTML と CSS を利用する。Amazon の Kindle 版の電子書籍も基本的に同様である。

したがって、HTML と CSS を用意できれば、同じ電子出版コンテンツを、ワンソースマルチユースで、Web、EPUB、Kindle へと展開可能。

- EPUB で利用できる HTML と CSS は、Web で利用できるものよりも制約はある。XHTML 形式 (XML 構文の HTML) を使う必要があること、利用できる CSS3 の機能が限られていることなど。しかし、Web の標準を作る W3C と電子出版の標準を作る IDPF とは密接な連携から、統合へと向かっており、いずれ Web と電子出版の標準の違いはなくなる方向である。

👉 ワールドワイド・ウェブ・コンソーシアム W3C と国際ナショナル・デジタル・パブリッシング・フォーラム IDPF が両組織の統合を検討中

印刷用 CSS Paged Media 仕様

- デジタルファーストのワンソースマルチユース出版で、同時に紙の本の制作も行う場合、問題になるのは印刷用の組版をどうするかである。

原理的には、同じ HTML+CSS で印刷用の組版までできれば、効率的なはずである。

- CSS でのページ媒体の組版に関して、以下のドラフト仕様などがある：
 - **CSS Paged Media Module Level 3**
 - ページサイズと余白、ページヘッダーとフッター、ページ番号など
 - **CSS Generated Content Module Level 3**
 - ページ番号参照などクロスレファレンス、欄外見出し生成など
 - **CSS Page Floats**
 - ページ上のフロート図版配置

ワンソースマルチユース出版に **HTML+CSS** 組版を利用している有名出版社の例

- 米国の技術書出版社 O'Reilly の書籍は HTML₅+CSS で制作されて、印刷・PDF、EPUB、Kindle、Web (online) へと多媒体展開。

※参照: “The Case for Authoring and Producing Books in (X)HTML₅”, Sanders Kleinfeld, O'Reilly Meida, Presented at Balisage: The Markup Conference 2013.

(HTML₅ をソースと出力フォーマットの両方に使っていること、CSS3 Paged Media 対応の組版エンジンを印刷用 PDF 生成に利用していること、Web ベースで WYSIWYG のオーサリングについて、HTMLBook という HTML₅ で本を書くしくみについて紹介)

- 米アシェット社は一般本を HTML+CSS で制作、電子版・紙版を同時出版。

※参照: “Production as if Digital Mattered: Making books with HTML and CSS”, Dave Cramer, Hachette Book Group

HTML+CSS の出版への活用における課題

- 従来ワークフロー（紙の組版用のDTPソフト中心）からの移行は、一般にはハードルが高い
- 印刷用の CSS 組版仕様は未完成のドラフト仕様であり、それを部分的に実装した処理系はあるが標準的ではないし、仕様が変わる可能性がある。
- 一般の Web ブラウザには印刷用の CSS 組版仕様は実装されていない。また、印刷用の CSS 組版を実装した処理系には、Web ブラウザで利用できる最新の CSS3 の機能が必ずしも実装されていない。したがって、印刷用と Web（画面）用とで共通の CSS で同じレイアウトを実現するということができない。
- 印刷だけでなくブラウザの画面上でもページネーションを行う必要性は高まっているが、そのための CSS 仕様の標準化はまだ初期の段階にある。

Vivliostyle HTML5+CSS

組版の取り組み

Vivliostyle (ビブリオスタイル) とは

- vivlio (βιβλίο /vivlío/ = ギリシャ語 〈本〉) + style
 - 本の作り方や読み方の新しいスタイルを、世界の誰でも自由に使えるオープンな標準技術で。
 - 紙でも電子でも、
 - 表現したいように表現できる自由と、
 - 読みたい形で読める自由を (=アクセシビリティ)

ということを目指して Vivliostyle と命名。

- 2014年8月に株式会社ビブリオスタイル設立、研究開発を開始。

Web ブラウザベース HTML5+CSS 組版

- 標準 Web ブラウザ上で本格的な組版を可能にする JavaScript ライブラリ Vivliostyle.js をオープンソースで開発。
 - 次世代の CSS 組版仕様の標準を、世界に先駆けて実装。
 - 標準の Web ブラウザベースなので、その最新の Web 標準技術 (HTML5/CSS3/JS/etc.) も利用可能。
- W3C メンバーとして、Web 標準技術を推進。とくに次世代 CSS 組版仕様の標準化とその実装に積極的に関わる。
- 製品として Vivliostyle Viewer (画面表示 JS ライブラリ) と Vivliostyle Formatter (PDF 生成エンジン) をリリースしている。製品内容、利用事例、組版サンプルなどは Web サイト <http://vivliostyle.com> を参照。



Vivliostyle サンプル紹介ページから

固定ページサイズ



可変ページサイズ



Adaptive Layout



まとめ

HTML5+CSS がワンソースマルチユース出版の鍵

- 文書構造は HTML5 で、体裁は CSS で指定するのが電子出版の標準。
- スタイルシート仕様 CSS は進化して組版能力は高まっているが、印刷用のページ組版仕様はまだ未完成。
- CSS ページ組版仕様のドラフト実装を使って本格的なワンソースマルチユース出版を行っている事例が一部にはある。
- ブラウザの画面上でも印刷でも HTML5+CSS で本格的な組版を実現するために、W3C CSS 組版仕様の標準化とその実装の取り組みが進行している。