

デジタル組版の基本用語

最終更新日: 2009/04/25

アンテナハウス株式会社

自動組版と DTP

テキストや画像などの素材をコンテンツという。テキストに指定するフォント名、文字の大きさ、行中の文字配置方法、約物の詰め処理、行頭・行末処理、行の間隔などを総称してスタイル指定という。本文領域外の余白・柱・ページ番号、本文領域内の段落・表・図など、様々な組版対象の配置指定をレイアウト指定という。

組版とはコンテンツを、スタイル指定とレイアウト指定に従って、見栄え良く紙面の上に配置する処理である。活字の時代には職人の技であったが、情報化の進展に伴い、徐々に職人の技がコンピュータ・プログラムによって置き換えられている。

DTP と WYSIWYG

DTP はデスクトップ・パブリッシングの頭文字であり、従来は専門の制作会社によって行われていた組版処理をデスクトップ・コンピュータで行うこと (DTP 組版) またはそのためのソフトウェア (DTP ソフト) を意味する。Macintosh 上の PageMaker を発売したアルダス社の社長であるポールブレナール氏が提唱した言葉とされている。DTP 組版では、オペレータが、DTP ソフトを使ってパソコンの画面上にテキストや画像などのコンテンツを配置しながらページを組み上げて行く。オペレータがコンテンツの入力、スタイル指定、レイアウト指定を同時並行で対話的に編集するので木目細かいページを実現できる。DTP

と同時に流行った言葉に WYSIWYG がある。

WYSIWYG は What You See Is What You Get の頭文字であり、編集画面上の組版結果を画面に表示されるとおりに印刷結果として得られることを意味する。

自動組版

DTP 組版はオペレータが画面上で対話的に組版処理を行うのに対して、自動組版は組版処理をコンピュータによる自動処理で行う。自動組版ではコンテンツはデータベースからダイナミックに取出し、あるいは、予め別に準備されているものを利用する。スタイル指定やレイアウト指定はプログラムやパラメータとして予め準備しておく。DTP 組版オペレータの代わりにプログラムとパラメータが担うことになる。

但し、DTP 組版もレイアウト指定、スタイル指定、文字や行単位の組版処理には DTP ソフトに組み込まれたプログラムの支援を利用するので半自動組版といっても良いだろう。自動組版システムには DTP ソフトに組み込まれた組版プログラムを使って、最小限の自動処理機能をプラグインで追加して実現する方式と、組版処理すべてを自力で行う自動組版専用ソフトを使う方式がある。DTP 組版と自動組版ではコスト構造が根本的に異なることに注意したい。DTP 組版ではオペレータが 1 ページ単位に人手で組版処理を行うのでコストがページ数に比例する傾向がある。自動組版はスタイル指定、レイアウト指定を予めプログラムしておくため初期投資がかかる

が、ページ数が多くなってもそれに比例してコストが増えるわけではない。DTP組版はオペレータの知恵を必要とする柔軟な紙面構成の出版物、自動組版は所定の同じスタイル・レイアウトを繰り返して使用する出版物に向いている。

XML

狭い意味では1998年にW3Cの勧告となったExtensible Markup Language仕様(XML勧告)のことを意味する。XML勧告に定められた条件を満たすデータのことをXML文書という。概念的な意味でXML文書の応用をXMLということもある。XML文書本文はテキストで論理構造をマークアップ表記で表す。図形等バイナリファイルを外部において参照できる。次は日々の交通費のデータをマークアップ表記で示した例である。

<交通費>

```
<データ record="1">
  <日付>5月10日</日付>
  <勘定科目>交通費</勘定科目>
  <金額>1,260</金額>
  <備考>B社まで、タクシー</備考>
</データ>
```

```
<データ record="2">
  <日付>5月10日</日付>
  <勘定科目>交際費</勘定科目>
  <金額>15,000</金額>
  <備考>B社Aさんを接待</備考>
</データ>
```

</交通費> <データ record="2">、<日付>など、< >でくられた表記をタグという。タグには開始タグ(< >)と終了タグ(</ >)がある。開始タグから終了タグまでを要素、タグで囲まれた範囲を要素の内容、recordを属性、"2"を属性値という。

XML文書のマークアップ表記は①文書全体を囲む唯一のルート要素があること、②ある要素は必ず別の要素の内部になること、③開始タグと終了タグが常にペアになることなどの条件を満たす。

あるXML文書に使用するタグの種類(タグセット)を決めるのは、文書の雛形を決めることに相当する。これを文書型定義という。XML勧告は文書型を定義する方法、XML文書処理するパーサというソフトが処理すべきことなども定めている。

このようにマークアップ表記の規則を細かく規定したものをマークアップ言語といい、起源は30年以上前に一般化マークアップ言語GMLに始まる。1986年に標準一般化マークアップ言語SGMLが作られたが複雑すぎてあまり普及しなかった。この反省を踏まえて、1998年のXML勧告はSGMLと比べて簡素化され、またWebで利用しやすいように決められた。XMLはシステム間・企業間のデータ交換需要が高まった時代背景もあり短時間で普及した。

DTDとスキーマ

XMLの文書型を定義する方法は幾つかある。伝統的な方法はDTD(Document Type Definition 文書型定義宣言)を使うものでSGMLの時代から引き続いている。ただし、DTDには、①XMLになって導入された名前空間が規定できない、②データの型の表現が不十分、などの限界がある。

このため新しくXML Schemaという定義方法が標準化された。XML SchemaはXSD(XMLスキーマ定義)を使って文書型を定義するが、例えば数値の型や数値の範囲などの規定が可能となっておりDTDよりも強力である。但し、オフィス用の文書処理に応用す

る文書型定義のためには DTD の機能で充分と考えられている。

DITA

DITA (Darwin Information Typing Architecture の頭文字) は、XML の文書型の一種類を規定する国際標準である。DITA は文書を部品化して記述し、再利用する目的で設計された。

DITA では文書をトピックという基本単位で記述して、多数のトピックをマップという構成情報を利用して組み立ててマニュアルなどを作成する。マークアップ言語はもともと文書の再利用のために発案されたが、特に DITA では一歩進んで文書処理のシステム化を志向していると言える。

トピックはタイトルとある種の形式の内容をもつ情報の単位であり、ひとつの主題・質問の回答として意味をもつ範囲で最も短く、かつ一単位として編集するのに適当な長さをもつようにする。トピックの文書型として topic、task、reference などの既定の文書型が幾つか用意されている。この既定の文書型は徐々に増えている。また、既定の文書型をユーザが拡張することもできる。DITA で作成したコンテンツは、HTML に変換して Web ページとして利用したり、PDF に変換して出版物として利用する。

画像と SVG

ドキュメントのコンテンツとしての画像はラスター画像とベクトル画像に分けることができる。ラスター画像の代表例は携帯電話の画像交換によく使われる JPEG である。ベクトル画像は線や折れ線、多角形を組み合わせる図形を表現したもので Adobe Illustrator などアプリケーション専用形式や Windows 用

の WMF/EMF など多数ある。SVG は Web などで使うためのベクトル画像の標準形式として提唱されており、XML 文法を使って定義されているのが特徴である。SVG には、携帯電話や組み込み系のシステムで使うための SVG Tiny という小規模セットも提唱されている。Web 上では当初アドビシステムズの SVG Viewer が無償配布されている。また、FireFox などのブラウザは SVG をある程度表示できるが、マイクロソフトの Internet Explorer が SVG をサポートしていないなど、Web での普及が遅れている。

なお、SVG の中には image タグを使って JPEG、PNG を指定することでラスター画像を含むこともできる。

XSLT と XSL-FO

XSL-FO (Extensible Stylesheet Language - Formatting Objects) は XML を印刷するために開発された標準仕様である。SGML の時代には同じ目的の DSSSL という国際標準があったが、XML の時代になって XSL-FO が DSSSL にとって代わる存在となっている。

XSL-FO の仕様はページマスター、ページ上の領域、ブロック・オブジェクト、インライン・オブジェクト、表などの様々な組版対象オブジェクト (Formatting Object) を規定している。また、ページマスターの出現順序、ブロック前後の強制改ページの有無、ブロック間の空き量、ブロック内での改ページ禁止など、自動組版の際のオブジェクトの振る舞いを事前にプロパティとその値で指定することができる。

XSL-FO を使って XML を印刷するときは、第一段階で XML ファイルをレイアウト指定済みの XSL-FO ファイルに自動変換する。第二段階で XSL-FO ファイルのレイアウト指

定を解釈しながらページ上にオブジェクト配置処理する。第二段階が組版処理であり、組版処理された結果は PDF に出力することが多い。

なお、第一段階で XML から XSL-FO に変換するとき使用するのが XSLT というスタイルシートである。XSLT (XSL Transformations) は XML から XSL-FO 変換のために導入されたが、前半の変換だけでも有用なので XSLT のみで独立した標準仕様となっている。

PDF

組版した結果を Web 経由で大勢が共有する場合、または、組版した結果を印刷会社に受け渡すためには、組版結果を保持することができる電子ファイル、すなわちページを表現できる電子ファイルが必要である。

この機能を実現する電子ファイル形式の中で最もポピュラーなものが PDF である。PDF は 1990 年代初頭にアドビシステムズが発明したものであるが、PDF の仕様書は 2008 年に ISO の 32000-1 という国際標準仕様として刊行された。

Web、HTML と XHTML

DTP、自動組版、XSL-FO、PDF といった技術とその応用は印刷をベースとして形作られたものである。これに対して、Web の起源は印刷ではなくまったく異なった領域である。すなわち、Web は分散して存在する文書など情報をインターネット経由で利用するという目的で発明された。その後、パソコンの画面の上に視覚的に情報を表しながら Web を閲覧する Web ブラウザが現れたことで爆発的に普及した。

HTML (Hyper Text Markup Language) は、

Web 上に情報を表現し、散在する情報を利用するために SGML 文法を利用して定義された文書型である。HTML を XML 文法によって定義しなおしたものが XHTML である。

CSS

CSS は Cascading Style Sheets の略であり、HTML でコンテンツを表現し、CSS を使ってレイアウトを表す役割分担と共に提案されている。スタイルを適用する HTML のタグ名とそのタグをどのように表示するかを、次の例のように指定する。

```
h1 {font-size: 2em; font-weight: 900;}
```

(h1 要素のフォントの大きさ、太さを指定)

```
h1 a { background-color:rgb(240,240,212);}
```

h1 要素内の a 要素に背景色を指定)

上の例では h1 要素の内容を可視化するときは文字を本文の 2 倍の太字で表すことになり、さらに、h1 要素内の a 要素の内容を可視化するときは色を変える。CSS ではスタイルを適用する対象を選択するセレクタ (h1、h2 など) と、それに適用するスタイルを表すルール ({ } 内の記述) が基本である。CSS は、①ブラウザや CSS 組版ソフト (UA) に内蔵、②CSS ファイルとして HTML にリンク、③HTML の <style>~</style> 要素の内容として指定、④タグの style 属性に指定など分散して配置できる。ひとつの要素に多数のセレクタが該当するとき、④ > ③ > ② > ①の優先度でスタイルを適用する。これをカスケード方式という。CSS が提唱された背景には、Web ブラウザが普及して、Web を画面で対話的に操作するようになる段階で、HTML に視覚的なレイアウト指定をする機能が追加され、そ

の利用が増えたことがある。ブラウザの表示媒体であるコンピュータの画面は、大きさ、縦横比、解像度などは千差万別であり、Web ページの情報を記述する HTML にレイアウト指定をも埋め込んでしまうと、多様な画面表示に対応し難いなどの問題が生じるので、コンテンツとスタイルを分離することが推奨されている。

なお、CSS は HTML のみでなく、XHTML や XML 文書のレイアウト指定もできる。

ツール

DTP では、データは使用する DTP ソフト毎に独自の形式で作成されて外部のファイルとして保存される。ドキュメントを記述する形式がアプリケーション・ソフトウェアから切りはなされていないため、InDesign や QuarkXpress などの特定のツールで作成した文書ファイルを他のツールで利用するには 1 対 1 対応のコンバータを使う必要がある。

これに対して、XML、SVG、XSL-FO、DITA、HTML、CSS は標準仕様である。誰でも、また様々なツールを使って標準仕様に準拠するデータを作成することができる。そして、標準仕様で作成されたデータを異なるツールを利用している企業間で受け渡すのも比較的簡単である。

XML を初め、これらの標準仕様が普及する背景には、こうしたオープンな環境を誰でも簡単に手に入れることができることがある。

オーサリング・ツール

ドキュメントを編集するツールのことはエ

ディタという。最近では、コンピュータによるマルチメディアの編集機能の充実に伴い、グラフィックス、映像、音楽、Web サイトなどの様々な情報の制作をコンピュータで行なうことができるようになってきた。このようなコンピュータ上で様々な情報を編集して、作品を制作するツールをオーサリング・ツールという。

自動組版ツール

XSL-FO や HTML+CSS はレイアウト指定を行なった状態である。レイアウト指定された情報をページの中に配置していく処理は、自動組版の一種である。XSL-FO を PDF に変換するツールは 20 種類以上ある。但し、日本語組版のできる製品は少ない。

HTML + CSS はこれまで Web ブラウザでのサポートが先行してきた。CSS は現在レベル 3 の標準化が進んでおり、レイアウト指定機能が強化されている。これに伴い、HTML +CSS3 で印刷組版のレベルまで可能になる時代が近づいている。

Web と印刷の技術は起源がまったく異なっており、現在は、Web 制作と印刷物の制作過程は全く別のプロセスになっている。しかし、将来は HTML と CSS3 によって Web 表示物と印刷物の制作過程が一元化され、コンテンツも統合される可能性が大きい。

初出

・「印刷雑誌」2009 年 5 月号 印刷学会出版部 発行