

XHTML

新XHTMLの全貌

神崎正英
www.kanzaki.com

異なる環境で情報を共有するための表現手段であるHTMLは、ブラウザの競争によって複雑なものとなった。独自の言語を採用したiモードなどの携帯電話サービスが普及したことで、コンテンツ制作者の混乱はさらに深まっている。この状況を解決するためにW3Cが開発中のXHTMLのモジュール化とXHTML Basicに注目したい。

XHTML 1.0は新世代の始まり

文書とデータを統合した XHTML

初期のHTMLは、定義されている要素（注1）型も20種類程度と少なく、誰でも簡単に情報共有ができるマークアップ言語としてスタートした。その後ブラウザの競争によってレイアウト効果のための独自要素が多数追加され、1つのコンテンツを提供するのにブラウザごとに別々のHTMLが必要になるという本末転倒の事態が生じたのは周知のとおり。

この状況を整理し、国際化やアクセシビリティのようなHTMLに求められる要素を追加して、HTML 4.0が1997年にW3C勧告となった。

HTML 4.0ではデザイン機能はスタイルシートに移され、HTMLは文書の意味構造を示すという役割が再確認されている。ハードやソフトの環境にかかわらず情報を共有するという、HTML本来の姿を取り戻す道筋が付けられたのだ（図1）。

一方ビジネス界では、データをインターネット経由でコンピュータに処理させるニーズが高まっていく。この目的に添って設計された言語の枠組みがXMLだ。XMLでは、HTMLには用意されていない「商品名」「価格」といった要素を定義したり、多くの処理をコンピュータで自動化したりできる。

このXMLの利点をHTMLに取り入れ、2000年1月にXHTML 1.0が勧告された。XHTMLを使えば、人間がブラウザで読むだけでなく、コンピュータにも処理しやすいコンテンツを記述できる。冷蔵庫の内蔵プログラムが買い物情報を収集、表示するという世界の実現には、XHTMLが大きな役割を果たすはずだ。

注1 要素：HTMLでは、文書を見出し（h1～h6）段落（p）などの「要素」の集まりとして捉えて役割を明示する。使える要素のタイプ（型）はマルチメディアなどを扱うために拡張されてきた。「タグ」と呼ばれることが多いが、正確にはh1、pなどの要素型名を<>で囲んだものがタグ。

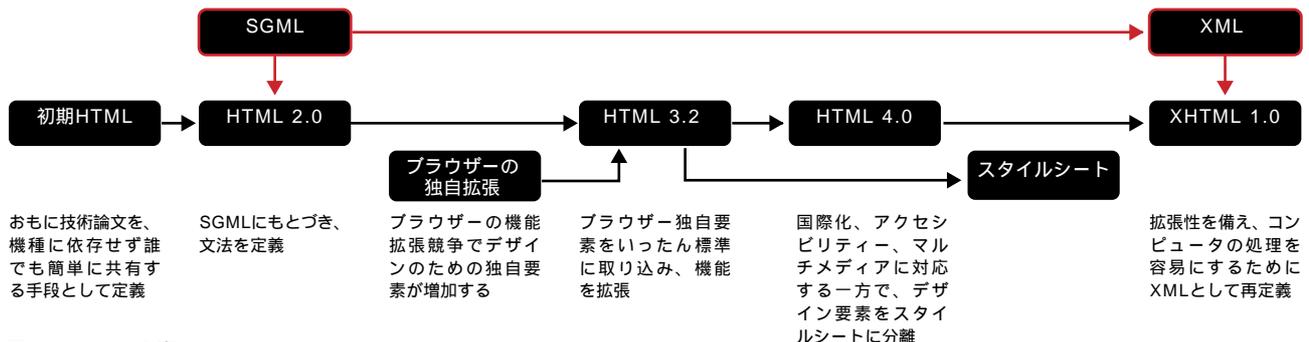


図1 HTMLの変遷

NonPC用言語の混乱

PC以外に対応できない HTML 4.0

かつては、WWWの情報はPC上のブラウザで閲覧されるものと考えてよかった。文字表示のみのブラウザや音声ブラウザも存在するが、いずれもPC上で動作し、共通の入力端末としてキーボードを想定することもできた。しかし、今やPDAや携帯電話はもちろん、テレビやゲーム端末、情報家電など、PCとは形も機能もまったく異なる環境が、ウェブの世界でも大きな比重を占め始めている。共通の基盤として想定できる前提が、がらりと変化しつつあるのだ。

HTML 4.0は、「環境に依存しない」というHTML本来の姿を取り戻そうとするものだった。したがって、HTML 4.0を使えばPCから携帯電話までどんな環境でも利用できるコンテンツを提供できるようになるはずだ。

しかし現実には、HTML 4.0はリッチな表現と高度な機能を盛り込んだため巨大な規格と化した。パワーのない機器では、スクリプトやマルチメディアを含むHTML 4.0に完全対応することは困難だ。また、HTMLは機器に特有な機能をうまく扱うための拡張性を備えていない。携帯電話なら電話番号を一貫した方法で扱いたいだろうし、テレビではチャンネルという概念が大切だが、これらはHTMLでは想定していないものだ。

仕様の乱立でコストがかさむ

HTML 4.0が使いつらいこともあってか、携帯電話の世界では、PCと異なる言語どころか事業者ごとに異なる言語が定義されてしまった。iモードではコンパクトHTMLをアレンジした独自HTMLが採用され、EZwebでは携帯電話の特性を活かす「カード」などの概念を取り込んだHDML (Handheld Device Markup Language) という別の言語が使用されている。さらにその他のモバイル端末向けサービスでもそれぞれ別仕様が用いられている(表1)。

携帯電話向けだけでもこの始末なのに、たとえばBSデジタル放送向けのコンテンツを作成するには、BML (Broadcast Markup Language) というまったく異なる言語を使わ

なければならぬ。標準がない世界では、このように仕様が乱立し、すべてのユーザーに利用してもらうためには、同じコンテンツでも各機種向けに別々のマークアップをほどこさなければならなくなる。情報提供者はコストがかかるし、利用者はコンテンツ選択の幅が狭まってしまい、どちらにとってもありがたい状況だ。

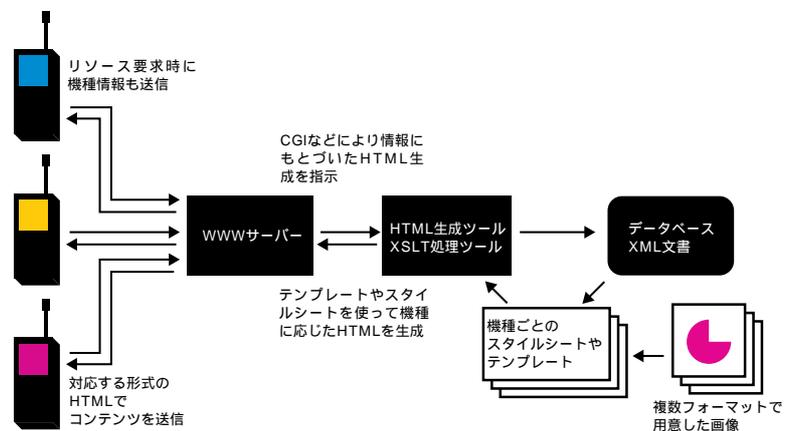
求められる標準化

コンテンツを何種類も作り分けるのは負担が重いので、すべての端末にサービスを提供する場合には、何らかのツールを使って1つのソースを各言語用に変換することになる。データベ

表1 携帯電話用の言語

事業者	サービス	言語	内容
NTTドコモ	iモード	iモード用HTML	小型端末向けに提案されたコンパクトHTMLにもとづき、独自の拡張をほどこしたHTML。画像はGIFのみ。
au	EZweb	HDML	Handheld Device Markup Language。「カード」「デッキ」など携帯の特性を活かす仕様だがHTMLとはかなり異なる。画像はBMPを使う。
Jフォン	J-スカイウェブ	MML、J-スカイウェブ用HTML	MML (Mobile Markup Language) はモバイル向けの独自言語。一般的には独自拡張のHTMLで記述する。画像はPNGのみ。
アステル	ドットi	ドットi用HTML	iモードとよく似たHTMLのサブセット。画像はGIFとPNGに対応している。
DDIポケット	オープンネットワークコンテンツ	独自形式のテキスト	ウェブコンテンツをPメールDXで受信するための規格。HTML文書はタグが取り除かれ、改行処理だけが行われてテキストとして表示される。

携帯電話系ウェブサービスの言語だけでも、事業者によって異なったものを採用している。通常のHTMLでコンテンツを書けば、サーバー側でそれなりに変換してくれる場合もあるが、画像形式の違いなど、単純にはいかない点が多い。



端末から送られてくる機種情報を使って、ページ生成ツールなどで機種ごとに異なる形式のマークアップを施して送信することは可能だが、それなりに技術と設備が必要になる。

図2 マルチデバイス対応の現状

ース系のサービスなら端末ごとにテンプレートを用意してページを生成したり、XMLで記述した情報ならXSLT(注2)で変換出力したりするなどの方法を探るわけだ(図2)。しかし、それぞれの端末の特性を活かした変換にはかなり技術を要するうえ、画像形式まで異なっているので、誰もが対応できるわけではない。端末の特性を活かしつつ、共通の基盤でコンテンツを共有できる標準が強く求められているのだ。

注2 XSLT: スタイルシート言語 XSL (Extensible Stylesheet Language) のうち、変換 (Transformation) 機能をつかさどる仕様。XML文書を一定のルールに従って別のXML文書やHTML文書などに変換する。XMLをブラウザで表示させるには、XSLTを用いてHTMLに変換するのが一般的。

「モジュール化」と「Basic」が解決策となる

柔軟な言語を作る XHTMLのモジュール化

XHTMLの狙いは、HTMLコンテンツをコンピュータで処理しやすいXMLで記述することに加えて、多様な端末向けのHTMLを設計しやすくするという点にもあった。互換性のない言語をあちこちで作るのではなく、広く普及したHTMLをベースに、相互運用性を確保しつつ必要な機能を追加したり、負荷の重い要素を省いたりした身軽な言語を設計できるようにしようというのだ。

このために導入の準備が進んでいるのが、XHTMLの「モジュール化」^{KJump01}である。HTML 4.01の要素型群を中心に、ウェブのコンテンツを構成する機能をいくつかの小さなグループ（モジュール）に分けて定義しておき、

その中から必要なモジュールを選んで独自のXHTMLを構築するというものだ。標準として提供されるモジュールを組み合わせることで、ゼロから言語を設計するよりも効率的に、しかも互換性のあるマークアップ言語を定義できる。

見出しやブロックテキスト、ハイパーリンクなどの基本要素は「必須モジュール」として採用を義務づけ、最低限の互換性を保証した。そのうえで、科学技術論文用なら数式を記述するMathMLのモジュールを追加することも、特別な端末の機能を生かすモジュールを独自に設計して追加することもできる（図3）。HTMLのモジュール化仕様はW3Cの勧告を目前に控えているが、すでにこの手法を利用して、XHTML 1.0に電子出版のための要素を加えて拡張したOpen eBook仕様^{KJump02}など、応

用の試みが始まっている。

携帯端末の標準となる XHTML Basic

XHTML 1.1はモジュールを使って定義される予定だ。XHTML 1.1はPC向けの標準HTMLであり、HTML 4.01やXHTML 1.0とほぼ同等だが、過渡期的な措置だったTransitional仕様（注3）はなくなる。

XHTML 1.1と並んで重要なのが、同じくモジュールを組み合わせて設計されるXHTML Basic^{KJump03}である。必須モジュールに基本的なテーブル、フォーム、画像などを追加したシンプルな言語で、携帯端末などでも十分その機能を活かせるように考案された仕様だ。携帯端末メーカーの集まりであるWAPフォーラムの定義するWML（Wireless Markup Language）も、次期バージョンはXHTML Basicに準拠するものと見られる。フォーラム会員の携帯電話各社がこの新WMLを採用すれば、モバイル向けコンテンツの混乱が一気に解決に向かう可能性もある。

XHTML Basicにはスクリプトやマルチメディア要素が含まれず、PC向けコンテンツを作成するには制約が多い。しかしこれを使えば、同じソースをPC向けとモバイル向けに共有できるという大きなメリットがある。XML文書でもあるから、XSLTなどの技術をすぐに適用でき、工夫次第で画面サイズやキー機能の違いなどにもうまく対応可能だろう。

XHTML Basicもモジュール化仕様と並んでW3C勧告の一手前まで来ている。携帯端末でのWWW利用がますます進む2001年、XHTMLから目が離せなくなりそうだ。

^{KJump01} www.w3.org/TR/xhtml-modularization/
^{KJump02} www.openebook.org
^{KJump03} www.w3.org/TR/xhtml-basic/

XHTML

HTML 4

```
html blockquote head link address title meta table  
script body h1 form ul p a input span style strong  
em object img div param ...
```

- 必須コアモジュール
- XHTML Basic
- XHTML 1.1

XHTMLモジュール



XHTMLモジュールは、HTML 4.0の構成要素を中心に、要素を機能ごとにグループ化して定義し、必要なものを選んで言語を設計できるようにした。SVGやMathMLのようなほかのXMLからモジュールを取り入れたり、独自のモジュールを設計して加えることも可能。コアモジュールが必須となっているので、異なる言語を作っても最低限の相互運用性が確保される。

図3 XHTMLモジュールの構成

注3 Transitional仕様：HTML 4.0では、文書の意味構造を示す働きに徹した「Strict」型のほかに、互換性のためにfontやcenterなどのレイアウト要素も使用可能とした「Transitional」型も認められた。これはXHTML 1.0にも引き継がれたが、1.1では廃止される。

標準化と独自機能の バランスが大切

テッド・ウゴフスキー

オープンウェブシステムズ社 テクノロジーグループ、チーフアーキテクト



photo : Nakamura Tohru

XHTML 1.0やXHTML Basic仕様書のエディターの1人であるオープンウェブシステムズ社(旧フォンドットコム社)のテッド・ウゴフスキー氏に、モバイルでのウェブ環境とXHTMLについて話を聞いた。

Q. XHTMLの規格は、いつごろからどんな背景で検討されてきたのですか。

A. HTMLは規格がどんどん拡大され、HTML 4.0に至るとモバイル機器やテレビで完全に実装するには規模が大きすぎるものとなってしまいました。そこで、1997年にHTML 4.0を勧告した直後から、さまざまなデバイスに適應できる言語を提供するために、W3Cのワーキンググループでカスタマイズが可能なHTMLの開発が始まったのです。

Q. XHTMLファミリーの中でXHTML Basicはどのような位置付けて作られたのでしょうか。

A. XHTML Basicは、テレビやモバイル端末のような環境でのXHTMLファミリーのベースになることを念頭に置いています。それぞれの環境に適した付加機能を加えてマークアップ言語を設計するときも、XHTML Basicを基本にすることで、相互運用性を保つことができます。モバイル用言語の最大公約数という位置付けです。

Q. 共通の基盤という意味では、XHTMLには必須モジュールが定められていますが、これ

とは別にXHTML Basicを定義するのはなぜでしょうか。

A. XHTML Basicはテレビやモバイル端末といった具体的な環境を想定し、そこでの相互運用性を実現するためのベースです。XHTMLはもっと幅広い条件において利用され、Basicとは異なる構成要素が求められることもあります。幅広い範囲も含めた、ウェブでの情報記述のもっとも基本となる部分が、必須モジュールだと考えてください。

Q. 独自のXHTML言語が設計できると、また規格が乱立してしまい、情報の普遍性という目標から後退してしまいませんか。

A. 機種に依存しない情報共有という考え方は、いまでもW3Cの重要な目標です。ただ、それぞれのデバイスには独自の機能が備わっており、それを活かせるようにすることも利用者の視点からは大切です。これらのバランスを取りながら、情報共有のコストを小さくしていかなければなりません。

Q. せめて携帯電話の世界だけでも、言語の統一は進まないものでしょうか

A. WAPフォーラムでは、次世代のWMLをXHTML Basicをもとにして設計する方向で検討が進んでいます。私は個々の事業者がこれをどういう形で採用するかについて申し上げる立場ではありませんが、NTTドコモもWAPフォーラムのメンバーですし、これが次世代携帯電

話向けウェブの共通の基盤となることを期待しています。

Q. W3Cで検討されているプロファイル関連規格が採用されれば、モバイル向けコンテンツの制作ももっとシンプルになるでしょうか。

A. CC/PP(注4)などのプロファイル規格は、たとえばXSLTとの組み合わせなどにより、有効な解決策の一部になります。しかしXSLTでは機器のどのボタンが押されたかなどの挙動をうまく処理することができません。また同様の処理は今でもサーバー側スクリプトなどを使ってある程度は実現可能ですが、相応の技術力が必要です。

Q. 1つのソースで多様な端末に対応できるようにするには、XHTML Basicでコンテンツを作成しておけば安心と言えますか。

A. XHTML Basicは多くのベンダーから支持されています。これを1つの基準としてコンテンツを作成するのは賢明な選択です。将来のモバイル環境を視野に入れて、XHTML Basicを検討されることをおすすめします。

注4 CC/PP: 利用者とコンテンツ提供者がそれぞれプロファイルを用意し、最適な組み合わせを選択するための仕組みがW3Cで検討されている。CC/PPは端末の性能や利用者の好み(利用言語など)をサーバーに伝えるための規格で、Composite Capabilities/Preference Profilesの略。



[インターネットマガジン バックナンバーアーカイブ] ご利用上の注意

このPDFファイルは、株式会社インプレスR&D(株式会社インプレスから分割)が1994年～2006年まで発行した月刊誌『インターネットマガジン』の誌面をPDF化し、「インターネットマガジン バックナンバーアーカイブ」として以下のウェブサイト「All-in-One INTERNET magazine 2.0」で公開しているものです。

<http://i.impressRD.jp/bn>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、URL、団体・企業名、商品名、価格、プレゼント募集、アンケートなど)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真の撮影者、イラストの作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は収録されていない場合があります。
- このファイルやその内容を改変したり、商用を目的として再利用することはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用する際は、出典として媒体名および月号、該当ページ番号、発行元(株式会社インプレス R&D)、コピーライトなどの情報をご明記ください。
- オリジナルの雑誌の発行時点では、株式会社インプレス R&D(当時は株式会社インプレス)と著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

このファイルに関するお問い合わせ先

株式会社インプレスR&D

All-in-One INTERNET magazine 編集部

im-info@impress.co.jp