

統一中間フォーマットの要件

齋鹿 尚史*

キーワード フォーマット；HTML；XML；スタイル記述；国際標準化；IEC62448

統一フォーマットの技術的意義

統一中間（交換）フォーマット（以下、統一フォーマット）を策定することになった経緯については、別稿に詳しく述べられているのでここでは繰り返さないが、その技術的な意義について述べる。

国内のテキスト系コンテンツフォーマットとして、現状ではボイジャーのドットブック(.book)、シャープのXMDFが双璧となっている。ドットブックはHTMLを拡張した仕様となっており、一方XMDF（記述フォーマット）はXML形式で定義されている。それぞれのタグ記述例を図1に示す。いずれも、「横書き、本文のフォントはOsaka、MSゴシックの順に優先して使用」「“はじめに”という文字列を中央揃えで表示」という記述を行う例である。

いずれのフォーマットもタグ形式で定義されている点では同じであるが、両者のタグや属性は、機能が一致しているものでも、図1でわかるように、名称は必ずしも一致していない。したがって、2つのフォーマットでコンテンツを作成する

には、何らかの方法で、それぞれの仕様に基づいて個別にタグ付けを行う必要がある^{*1}。統一フォーマットの最大の意義は、このような、フォーマットが分かれているために生じている電子出版制作上の、非効率性を除去することである。

ドットブック、XMDFは、それぞれの開発の経緯や各社の意図によって、機能についても一致しない部分がある。このような機能面での相違の扱いについては後に述べる。

なお、統一フォーマットの仕様は公開が前提となっており、誰もが使用することができるものであることを改めて強調しておきたい。また、統一フォーマットから、閲覧フォーマットや流通フォーマットに変換することは、各社のビジネス領域であり、今回の標準化（統一フォーマット）で規定する範囲には含まれない（図2）。

中間フォーマットを変換する出力先は、特定の流通フォーマットに限られるものではない。コンテンツの権利者の許諾が得られ、変換ツールが整備されているのであれば、中間フォーマットから、いかなるフォーマットに変換して配信することも

【XMDF】	【ドットブック】
<pre><text_default_attribute baseline="right" /> <flow_default_font fontname="Osaka, MSゴシック"/> : <p align="center">はじめに</p></pre>	<pre><T-Time stroke="horizontal" honmonface="Osaka, MSゴシック"> : <center>はじめに</center></pre>

図1 タグの例。左がXMDF、右がドットブック

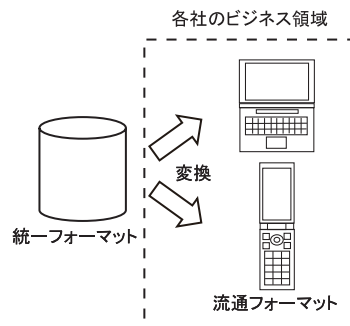


図2 流通フォーマットへの変換

* SAIGA, Hisashi

シャープ株式会社

システムソリューション事業推進本部電子出版事業推進センターコンテンツシステム開発室 係長

〒632-8567 奈良県天理市櫛本町 2613-1

saiga.hisashi@sharp.co.jp

可能である。このような点については、中間フォーマットの標準化では取り扱わず、各社がビジネスとして取り組むことになる。図2で、「各社のビジネス領域」と書かれているのはこのような意味である。

統一フォーマットの技術的内容

統一フォーマットの仕様については現在未定の部分が多く、方針も含めて、現時点ではあくまで案の段階であることをお断りしておく。

XML フォーマット

統一フォーマットは、別稿で説明されているような、XML フォーマットとして策定する予定である。XML を扱うツールの充実などから考えて、統一フォーマットをXML で定義するのは技術的に自然なことと考えられる。また、今回の統一フォーマットは国際標準化を視野に入れており、この面からもXML 形式で定義することには異論は少ない。

既存フォーマットとの関係

統一フォーマットは、その策定の目的から、ドットブックおよびXMDF の機能を包含したものとする必要がある^{*2} (図3)。なお、図中の「ミニマムセット」については後に述べる。

機能の包含の仕方については、

- 1) 共通または類似した機能は、同一のタグや属性にまとめる。
- 2) 双方のタグをできるだけそのまま使えるようにする。

という、異なる方針が考えられる。後者の方針は、

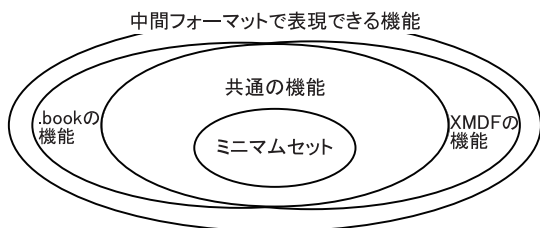


図3 フォーマット間の関係

各フォーマットとの互換性は高いという利点がある反面、統一フォーマット自体の仕様は煩雑になりがちであり、前者の特質はその裏返しとなる。このあたりも今後の検討によって決定することになる。

スタイル記述の分離

統一フォーマットは、個々の端末に縛られるようなものであってはならず、そこから作成されたコンテンツが長期の利用に耐えるものでなくてはならない。すなわち、今後の端末の発展に、交換フォーマットとして利用し続けることができるものである必要がある。

これを保証するためには、

◇端末の仕様 (解像度など) に依存した、「見え方」に関する部分 (スタイル)

◇端末によって変わらない部分 (内容)

を分離するのが良いと考えられる。これにより今後、現在予想されていないような画面を持った端末に遭遇したとしても、最悪でもスタイル記述のみ変更するだけでコンテンツを利用し続けることができる。すなわち、現在の端末仕様に縛られず、長期にコンテンツを利用できることになる (図4)。

スタイル記述と内容を分離する考え方は、たとえばWeb でも導入されてきている。Web では、よく知られているように、HTML でコンテンツが記述されるが、スタイル記述は、スタイルシートと呼ばれる、別のファイルに記述することが可能になっている。

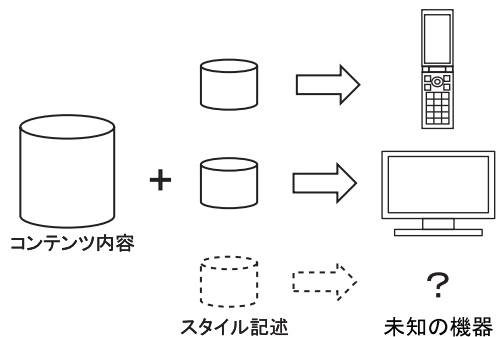


図4 スタイル記述の分離

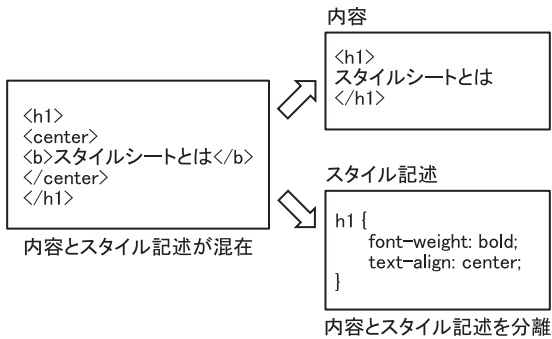


図5 スタイルシートのご概念

HTMLの各タグに対して、どのように表示すべきかを指定するのが、Webにおけるスタイルシートの基本的な考え方である。図5に内容とスタイル記述が混在しているHTMLの例と、スタイル記述をスタイルシートに分離した例を示す。

図5左側の例では、ボールド体であることを示すタグ、センタリング（中央寄せ）を行うことを示す<center>タグは、内容（「スタイルシートとは」という文字列）がレベル1の見出しであることを示す<h1>タグと同様に同一のファイルの中に記述されている。すなわち、スタイルの記述は内容と分離されていない。

これに対して、図5右側の例では、ボールド体であること、センタリングを行うスタイル記述はfont-weight, text-alignとして記述され、<h1>タグで挟まれた内容とは別のファイルに取られている。

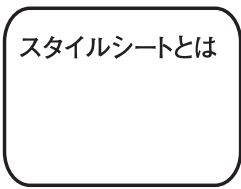


図6 図5のスタイルシートの表示例

ここで例に用いたスタイル記述の仕様はCSS（Cascading Style Sheet）と呼ばれており、HTMLのスタイル記述としては主流となっている。図6は図5で示した記述に対応する表示の例であり、「スタイルシートとは」という文字列がボールド体で、行の中央に表示されている。

統一フォーマットはHTMLではないが、ここで述べたスタイル記述の考え方を適用することで、各社のフォーマットの仕様も踏まえつつ、スタイル記述と内容との分離を図る。

ミニマムセット

統一フォーマットの制定にあたっては、日本語を表すのに最低限必要なタグの集まりであるミニマムセットを同時に定義することが考えられている。

統一フォーマットは、すでに日本語コンテンツで実績を持つ2つのフォーマットを機能的に包含しているため、このような日本語を表すのに最低限必要な機能は当然含まれることになる。したがってミニマムセットの機能は、図3に示した通り、統一フォーマットの機能のサブセットである。そのタグ名や属性名などの仕様については、

“電子書籍”に踊らされることなかれ!


電子出版の構図

実体のない書物の行方

植村八潮 著

12年前から繰り返される“電子書籍元年”への軌跡とは。
出版、印刷、図書館、そしてマスメディア……
メディアの未来は過去を知らずして語れない。

◎四六判 280ページ / 定価 2,100円（本体 2,000円+税）



話題沸騰中!

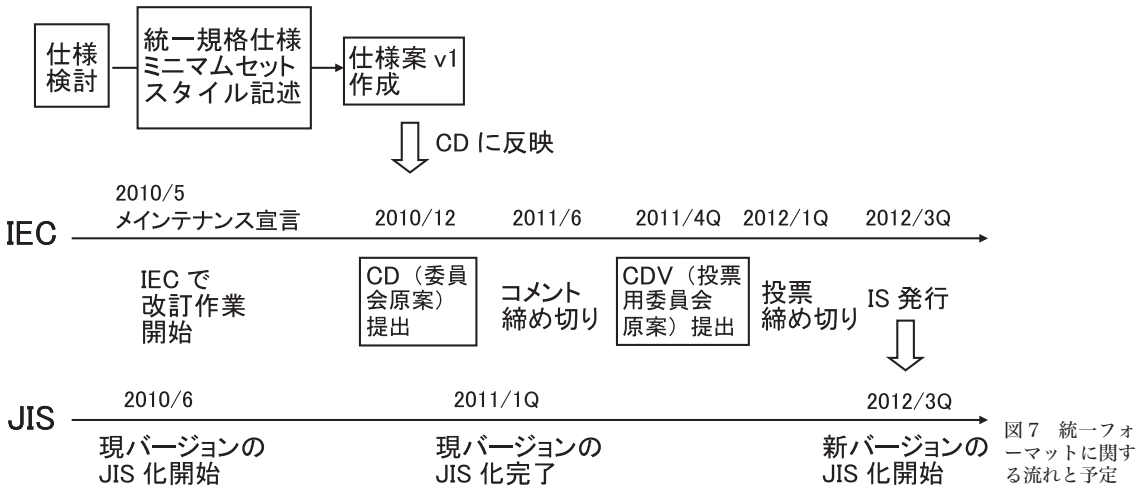
●直送の場合、送料が別途かかります

〒104-0032 東京都中央区八丁堀4-2-1

株式会社 印刷学会出版部

TEL.03-3555-7911 / FAX.03-3555-7913
info@japanprinter.co.jp / www.japanprinter.co.jp

2010年 > 2011年 > 2012年



既存フォーマットや、IEC62448 との関係も踏まえて、どのような形が望ましいかを検討することになっている。

仕様策定の動き

2010年10月に仕様案第一版を策定することを目標としている。また国際標準化も同時に進めることを想定しており、IEC62448の改訂に合わせ、仕様案をCD(委員会原案)に盛り込むことになる。国際標準化提案にあたっては、必要に応じて、多国語対応のための仕様調整や拡張が行われる。

現状では不確定要素も多いが、2012年中にこ

の統一フォーマット仕様が盛り込まれた国際標準を発行することを目標に、国際標準化活動も並行して進める予定である。(図7)

統一フォーマットの策定にあたって、ご協力、ご指導頂いている各位に深く感謝いたします。■

※1：一方から他方に機械的に変換するソフトウェアも存在するが、完全なものではない。

※2：特殊な端末や用途に特化した仕様についてはかならずしもこの限りではない。またそれ以外の機能については今後検討の必要がある。

本の未来、情報の未来—そして印刷の未来

我、電子書籍の抵抗勢力たらんと欲す

中西 秀彦 著

電子端末の台頭、紙の消えゆく社会……。

紙文化の未来はどこへ向かうのか。

デジタル情報化社会の奔流に立ち向かう老舗印刷屋最前線!!

◎四六判 206ページ/定価 1,680円(本体 1,600円+税)

発売即重版!



●直送の場合、送料が別途かかります

〒104-0032
東京都中央区八丁堀4-2-1

株式会社 印刷学会出版部

TEL.03-3555-7911 / FAX.03-3555-7913
info@japanprinter.co.jp / www.japanprinter.co.jp