

XMLによるマニュアル情報の発信

根岸 寛明 (negishi@td.ed.fujitsu.co.jp)

富士通株式会社

電子デバイス事業推進本部技術標準部ドキュメント技術部

【要旨】

マニュアルの SGML 化に続き XML 化を行い, XML が SGML 化された情報の提供, 表示のツールとして「使い物になる」レベルに達している。しかし, 資産としてのドキュメントの形式としては今後の規格やサポートツールの状況を見守る必要があると判断している。

今後, XML の時代に対応し, 情報構造を意識したコンテンツ作成技術の開発が急務である。

Trial of Manual Delivery by XML Format

Hiroaki Negishi (negishi@td.ed.fujitsu.co.jp)

Technical Communication Dept. Electronic Device

FUJITSU LIMITED

【Abstract】

Through the operation of SGML-ize and the trials of XML-ize of our semiconductor manuals, the XML, in the present condition, reaches the usable level as the format for delivery and display of the SGML-ized documents. Though, as the format for the assets of enterprises, the XML, in the present condition, has not reached the available level. We must watch the growth and improvement of standards relate to XML and the XML support applications.

We, the contents providers, should develop and apply the techniques for contents construction conforming to the age of XML contents, that must be well structured.

1 はじめに

富士通の半導体部門の一般顧客向け製品情報は, 小さな物は HTML 化, 小冊子類は PDF 化し, 既に当社の情報公開ウェブより公開している。「マニュアル」と呼ぶ, 製品の詳細仕様や使用方法等に関する厚手のドキュメントは SGML 化を進めており, 一部は CD-ROM ドキュメントとして出版している。このマニュアルも当社の情報公開ウェブから公開することを目指し, 出版形態として, 今後, XML を採用することにし, 現在, XML 化を進めている。

本稿ではマニュアルの SGML 化とそれに続く XML 化の状況, 工夫を要した点, 電子出版形式として見た XML の評価について報告する。

2 マニュアルの SGML 化状況

マニュアルは一般に, 構造が類似していること, 1冊あたりのページ数が多くかつ機能追加等による改版が頻繁であること, 別製品/新製品でも類似機能が多く, 原稿データの部品化や流用が進んでいること, 前書, 奥付等, 定型部分があること等々, SGML 化に適したドキュメントである。富士通の半導体部門も当初から SGML 化を前提にした電子化を進めてきた。

2.1 DTD

マニュアルの DTD は半導体のマニュアルに固有のものである。紙マニュアルの時代より「構造化」と呼ぶ独自の作り方[1]を採用しており、それに合わせた DTD(「構造化マニュアル DTD」と呼んでいる)を作成し、使用している。96年2月の試行版、同9月にCD-ROM 出版のための第1版、98年3月にDTPフォーマッタ対応のための第2版を経、98年10月にXML 対応のための第3版を作成し、現在に至っている。

構造化マニュアル DTD の特性は以下のとおりである。

- 「本」としての構造(章,節,項,…)を採用している。
- データ(原稿)作成の便宜を図るため、インスタンスを細分して作成可能にしている。即ち、章,節,項, …をそれぞれ独立した SGML 文書として取り扱えるように考慮している。これにより、部品化/共通化が容易になると同時に、これらのレベルでのハイパーテキスト構造のドキュメントも作成可能になっている。
- 「本」内のリンクは SGML の ID/IDREF 参照、「本」をまたぐリンクは HyTime リンクを、図データの参照等に用いる外部エンティティ参照は SGML で標準的に採用されている属性値のエンティティ名参照を採用している。運用上エンティティ宣言を1ファイルに集中させており、SGML 処理系間で発生するサポートデータ形式の非互換を、データ変換とエンティティ宣言ファイルの一括変換を組み合わせることにより解消している(図1)。
- 表は CALS 表を採用している。

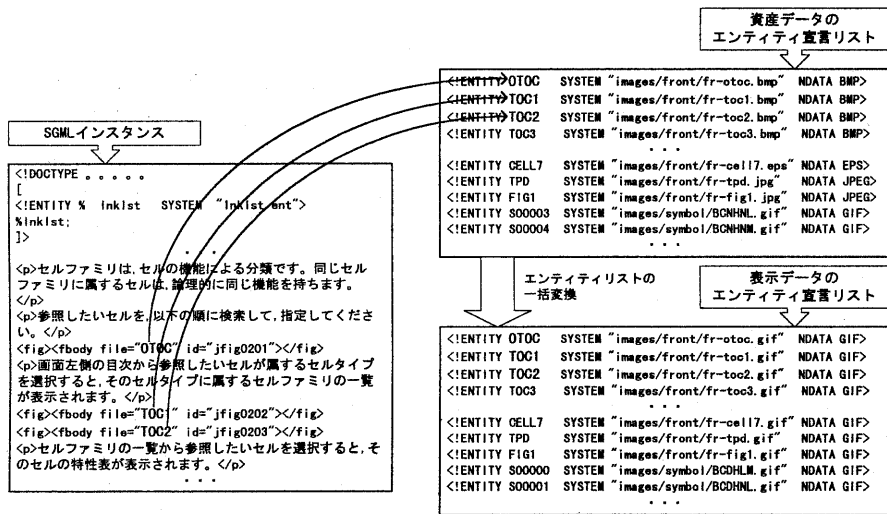


図1 属性値のエンティティ名参照とエンティティ宣言ファイル

採用している SGML サポート DTP, SGML 出版ツール, SGML フォーマッタ等, SGML 処理系の間に DTD の制約があった。これらの制約については、DTD の改造や複数構造のサポート、データ変換スクリプトによる処理、および各処理系の機能(スタイル等)の活用で対応している。

- DTP 側では他ページ参照に対し、参照元には参照先タイトルキャプション等の情報を持たず、印刷時に引き出して挿入するという一般的な仕様になっており、サポートできる DTD も参照元の要素を“EMPTY”としなければならなかった。一方、SGM 出版ツール側ではアクセス数を少なくする

ため、リンク先(参照先)を示すアンカーの存在が前提になっていた。双方に対応するため、参照元の要素の直後にアンカーを置けるようにし、DTD を切り換えてアンカーを有効にしたり無効にしたりできるようにした。

- DTP では表内表(2重の表)が使えないため、DTD より表内表を削除した。
- その他、図や表のキャプションの位置や要素名の制約等々があり、DTD でそれらに対応した。

SGML 処理系間の制約のいくつかは排他的なものであったため、構造化マニュアル DTD は第 2 版以降、「作成/保管用 DTD」と「SGML 処理用 DTD」の 2 本建てになっている。マーク付セクションを利用して DTD を切り換えるようにし、SGML で許されているタグの省略機能を利用することで、同一インスタンスで複数の構造に対応できるようにしている。

また、情報を扱う人間側にも、情報構造として採るべき構造と表示等により情報を使用する際に理解しやすい構造に差がある(図 2)。

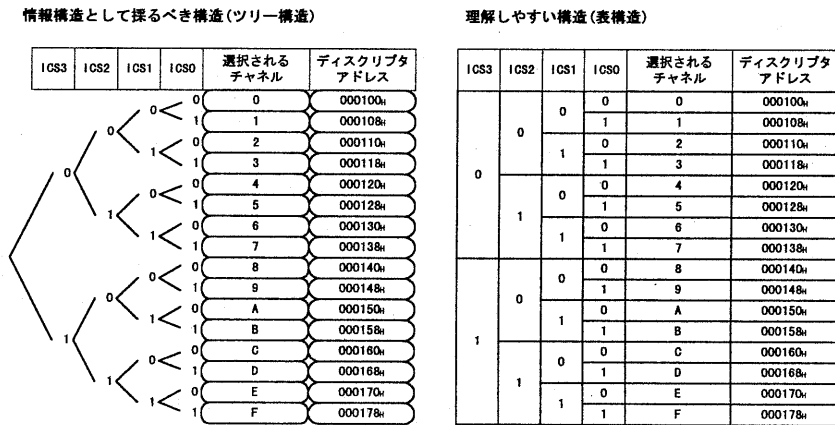


図2 本来の情報構造と理解しやすい情報構造の例

情報の本来の構造と情報使用時の構造の差については、原稿データ作成時に使用を前提にした構造で(例えば図 2 の例では表で)原稿を作成するのが通例になっていたため、構造化マニュアル DTD での対応は行っていない。将来的にこの種の情報が本来の構造で記述されるようになったとしても、DTD ではなくスタイルで対応すべきものであろう。

2.2 SGML インスタンスの作成/改版と出版

SGML インスタンスの作成/改版は、一部の手作業による直接タグ付けとデータベースからの SGML データ生成を除き、大多数は SGML サポート DTP からの変換、またはワープロと市販 SGML 変換ソフトの変換の組合わせで行っている。なお、SGML サポート DTP、市販の SGML 変換ソフトとも、属性値等、必要な値のすべてを設定することができないため、テキスト変換スクリプトを作成し、完全な形に再変換している。

SGML 出版(CD-ROM で出版)、およびイントラネット内のウェブ出版には OLIAS (OnLine Information Access System)、および OLIAS Web を、紙出版には、一部で SGML フォーマッタ、大多数は DTP/ワープロの出力を使用している。なお、OLIAS を用いた CD-ROM 出版は 1997 年 4 月に初版を出版し、現在も継続中である。

3 XML化

XML化にあたり、公開されていた非互換項目リストやインターネットに公開された情報、市販書籍等による事前評価を行い、SGMLからXMLへの変換が容易であること、XMLへの変換で失われる情報がないこと、さらにXMLからSGMLへの逆変換も容易であることを確認し、さらに現状(99年6月現在)の処理系が"使い物になる"ことも確認している。

3.1 事前評価と変換

XML化に先立ち、SGMLとXMLの26項目の非互換リスト[2]に基づき、机上でXML移行の可否の評価を行っている。また、該当規格のドラフト、各種公開情報等により、その後の変更をフォローしている。

事前評価の結果はおおよそ以下のとおりであり、これに基づき、XML化を開始した。

- SGMLで必要な、および使用している機能は、すべてXMLに継承されている。
- 既存のSGMLインスタンスの変更は、事実上、不要である。
- SGMLに対して加えられた制限の大半は、ブラウザ等の処理系の処理効率や、特にタグ省略時に発生する文書構造の曖昧さの除去という観点から、合理的である。
- ただし、XML化により、SGMLでは可能であった手作業によるタグ付けが、事実上、不可能になり、XMLインスタンスの作成には何らかのツールが必要になる。この点で、人間が扱うためのSGML、機械が扱うためのXMLという棲み分けが発生している。

DTD改版等を通じ、上記非互換項目についての対応は以下のとおりであった。

- 大半の項目は、現在の処理環境等では無用になっているものや通常では使用しない機能であり、対応不要であった。
- 以下の項目はDTDの変換で対応した。
 - ◆ 排他的な条件はマーク付セクション等により、複数DTDの切替えで対応した。
 - 名前に"_"と"."が使用可となり、特にXMLでは"."が必須となった。逆にこれらはSGMLではエラーになる。
 - DTD内で最小化パラメータ("o","-")の指定が不可になった。
 - ◆ 構造上の制限を緩和し、上記と同様、複数DTDの切替えで対応した。なお、構造が緩和されたDTDはXMLでの使用に限っている。
 - 内容モデルが混合モデル(mixed content)である場合、#PCDATAを必ず先頭に置く。(なおこの項目は、SGMLの場合でも、構造が曖昧になることが指摘されていた[3].)
 - 内容モデルの添加("+)、排除("-)が不可となった。一部の要素で、無限ネストの排除が困難になった。
 - ◆ 以下の項目は手作業で1件ずつ修正した。
 - 要素型宣言、属性宣言で要素名指定に名前グループの使用が不可となった。
 - コメントが"

- ◆ "OMITTAG"が"NO"に固定となり、開始タグ,終了タグの省略が不可になった。
- ◆ タグ名等で大文字と小文字が区別されるようになった。
- ◆ "SUBDOC"が"NO"となり、副文書が使用不可になった。
- ◆ 属性値指定の際の属性名,引用符等の省略ができなくなった。

これにより、第3版以降,構造化マニュアル DTD は「作成/保管用 DTD」と「SGML 処理用 DTD」に「XML 処理用 DTD」を加えた3本建てになっている。これもマーク付セクションを利用して DTD を切り換えるようにしている。

3.2 変換プログラムとスタイル

上記に基づき,SGML から XML に変換するテキスト変換スクリプトを開発している(図3)。

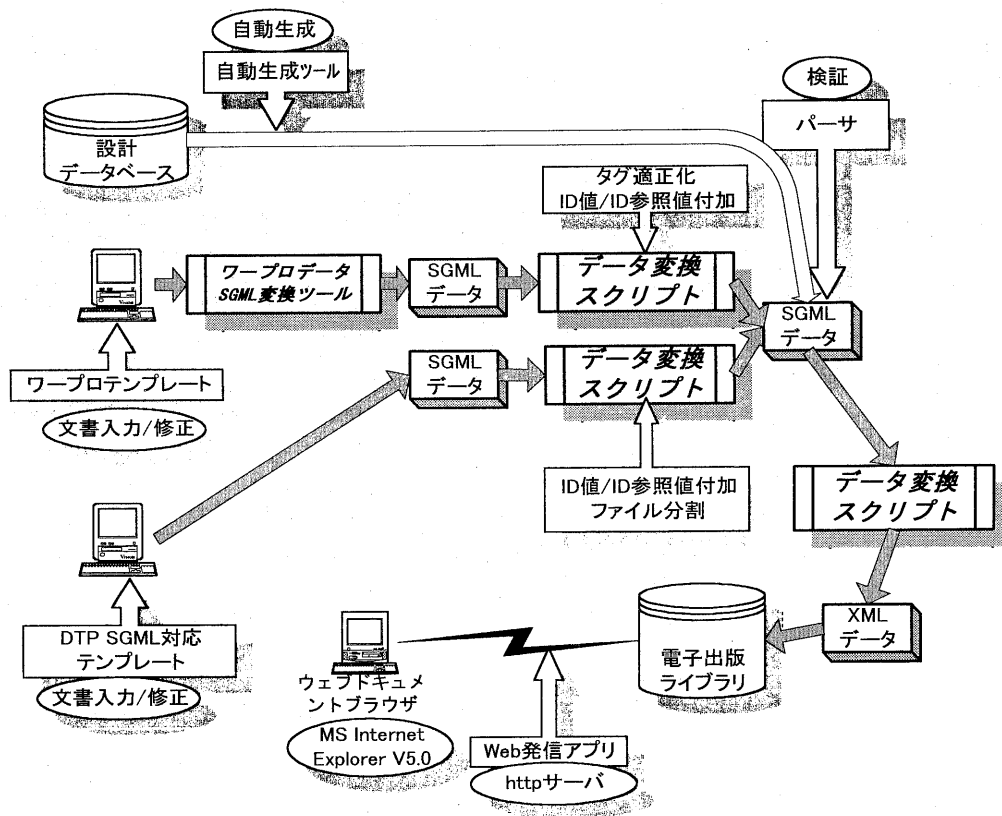


図3 SGML/XML 化環境

変換のエンジンには著名なフリーウェアである SP[4]を使用し、パース,フルタグ化(タグの省略の補完),XML 化を行い、さらにスクリプトにより以下の変換を行っている。情報提供者の責務として、提供する XML データは"Valid"でなければならないと考えており、このスクリプトで変換した XML データは、上記「XML 処理用 DTD」に基づいた"Valid"な XML データとなるようにしてある。

- オンライン表示に適した形式への変換
 - ◆ 表示単位へのファイル分割
 - ◆ 上記各表示単位へのナビゲート生成
 - リンク,外部エンティティ参照の変換
 - ◆ ID/IDREF 参照,HyTime リンクの XLink への変換
 - ◆ 外部エンティティ参照の URL 参照への変換
 - XML ブラウザ(Microsoft Internet Explorer V5.0)での表示のための属性値付加
 - ◆ CALS 表を HTML 表に変換して表示するための属性値の付加
- スタイルでは,各要素を HTML 要素に対応付ける他,以下の変換を行っている。

- XLink の HTML リンク(~)への変換
- 外部エンティティ参照の埋込み画像要素()への変換

なお,スタイル定義では CSS の使用は 1 箇所のみで,他はすべて HTML 要素への変換でスタイル付けを行っている。これは,今後予想される HTML 要素への対応付けから現在策定中の Formatting Object への対応付けへの変更を容易にするためである。

XML 化し,現在,イントラネット中で公開している XML マニュアルの画面例を以降に示す(図 4,5)。

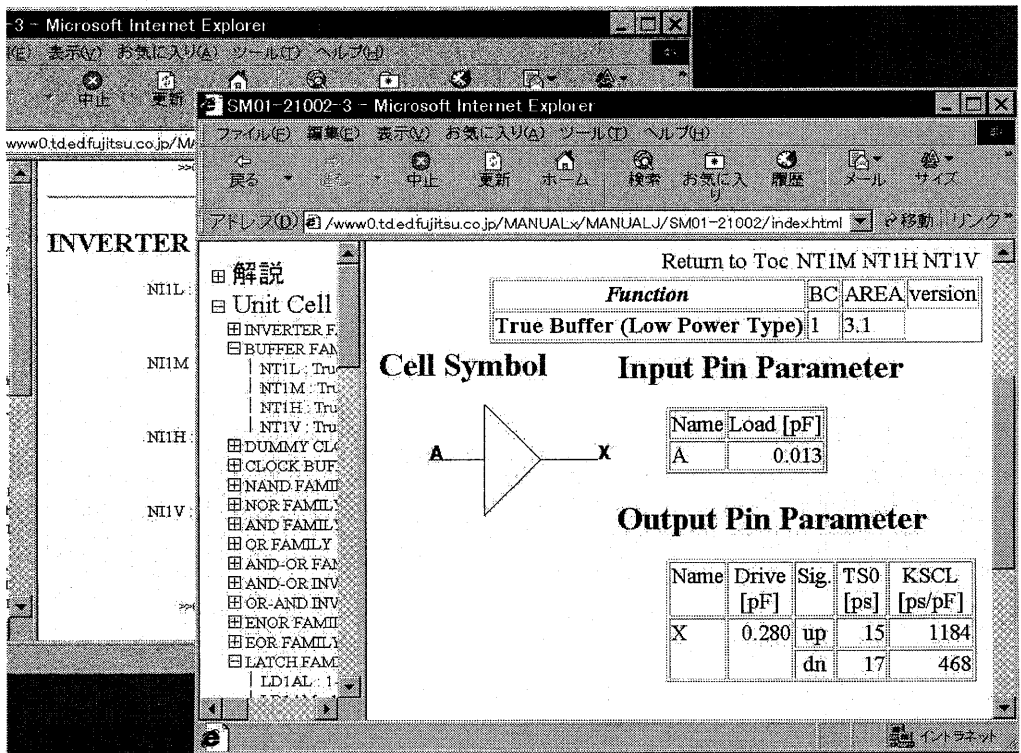


図4 XML マニュアル表示例-1



図5 XML マニュアル表示例-2

XML 自体は勧告になっているが、関連規格の Xlink(検討時点では Working Draft 3 [5]),XSL(同,Working Draft 16 [6]),XSLT(同,Working Draft 21 [7])等がまだ勧告になっていないこと、これらをサポートするブラウザ等の処理系がない(現状は Microsoft Internet Explorer V5.0 のみであり、かつサポート範囲は XSLT の一部)ことより、この XML への変換プログラムとスタイルは、今後、規格制定の進捗、および処理系の規格サポート状況に対応して改造を加えていく。

4 まとめと今後

XML は「使い物になる」レベルに達していると判断しており、現在、XML マニュアルを情報公開ウェブから社外公開するため、社内の関連部署と協議中である(本稿発表時点で既公開となっているかも知れない)。懸案であった、SGML 文書をウェブ公開するための「一般的」な仕組みがやっとできたと考えている。

関連規格の制定状況と処理系の規格準拠状況に依存するが、当面、XML は情報提供と表示のための形式に留まると考えている。マニュアル等のドキュメント類は資産であり、管理され、頻繁な改版、流用に耐えるものでなければならず、現在の規格、処理系では不十分である。資産そのものの XML への移行についてはペンディングとしている。

情報流通の人間との接点を担うコンテンツをの立場では、今後、XML の発展、浸透のためには、

XLink,XSL,XSLT,スクリプト等,XML 関連規格の成立とともに,Microsoft Internet Explorer を始めとする XML ブラウザが制定された規格に準拠していくことが必須である。現状の HTML ブラウザのような, ブラウザ間の非互換を考慮しながらコンテンツを作成するような状態は好ましくない。

XML に限らず, オンライン情報検索に対するコンテンツ作成や評価のガイドラインの検討が進んでいる [8,9]。アイデアレベルであるが, 既にガイドラインの断片が提起されている。「本」と同様のナビゲートの必要性等が提起されており, XML および XSL,Xlink 等の関連機能, ブラウザ等の機能でどのように実現していくかを今後検討していく必要がある。

また, XML の時代に対応し, 情報構造を意識したコンテンツ作りの技術の開発を進める必要がある。15 年程前に紙のドキュメントの作成技術をまとめた [10] が, これらをオンラインドキュメントの作成技術としてブラッシュアップしていく必要がある。

【参考文献】

- [1] "構造化マニュアルの作成教育" 杉田裕子,小松美智代,岡崎善博,木村和峰,塩見さち子
90/11/17 (社)電子情報通信学会 研究会 ET 90-101
- [2] "Extensible Markup Language (XML) W3C Working Draft 14-Nov-96WD-xml-961114"の
"A. XML and SGML" (<http://www.w3.org/TR/WD-xml-961114.html#secA>)
- [3] "SGML An Author's Guide to the Standard Generalized Markup Language" Martin Bryan
Addison-Wesley Publishing Company, Inc. 1988
該当箇所は邦訳版"SGML 入門" 山崎俊一監訳/福島誠訳 アスキー出版 1991 p.270
- [4] SP (<http://www.jclark.com/sp/index.htm>)
- [5] XML Linking Language (XLink) Working Draft 3-March-1998
(<http://www.w3.org/TR/1998/WD-xlink-19980303>)
最新版は XML Linking Language (XLink) Working Draft 26 July 1999
(<http://www.w3.org/1999/07/WD-xlink-19990726>)
- [6] Extensible Stylesheet Language (XSL) Version 1.0 Working Draft 16-December-1998
(<http://www.w3.org/TR/1998/WD-xsl-19981216>)
- [7] XSL Transformations (XSLT) Specification Version 1.0 Working Draft 21 Apr 1999
(<http://www.w3.org/TR/1999/WD-xslt-19990421>)
最新版は XSL Transformations (XSLT) Version 1.0 Working Draft 9 July 1999
(<http://www.w3.org/1999/07/WD-xslt-19990709>)
- [8] "平成 10 年度 オンラインドキュメンテーション標準化調査研究委員会報告書"
平成 11 年 3 月 (財)日本規格協会 情報技術標準化研究センター
- [9] "平成 10 年度 電子マニュアル評価ガイドラインの適正標準化に関する調査研究"
平成 11 年 3 月 (財)ニューメディア開発協会
- [10] "ドキュメンテーションエンジニアリングによるマニュアル作成" 根岸寛明,吉田哲三
86/1 情報処理学会第 27 回プログラミングシンポジウム
"富士通におけるソフトウェア品質保証の実際" 久保宏志監修 日科技連出版社 1989
第 6 部:ユーザ・マニュアルの品質保証の実際(pp.515~602)
および,関連社内資料,社外講習会テキスト等