

平成 7 年度

**ユーザーの視点で《SGML》を
検証する
～CALISの重要要素技術SGMLの検討～**

平成 8 年 3 月

標準化動向調査部会

ドキュメント・データ標準の課題WG

【目 次】

エグゼクティブ・サマリー

WGメンバー

第1章 はじめに

- 1. 1 背景
 - (1)標準化の背景
 - (2)ドキュメント電子化の背景
- 1. 2 活動目的
- 1. 3 WGの役割と検討項目
- 1. 4 今後の展望として

第2章 WGの活動内容

- 2. 1 SGML導入効果
 - (1)SGML化による期待効果
 - (2)SGMLの導入効果のある事例
 - (3)SGML化のための投資と労力
 - (4)SGML導入効果の評価
- 2. 2 ワープロとSGMLの連携
- 2. 3 マルチメディア対応ツールとしてのSGML
- 2. 4 SGMLの業務利用
 - (1)J2008 のDTD
 - (2)データベースとSGML
 - (3)EDIとSGML
- 2. 5 HTMLとSGML
 - (1)HTMLの特色
 - (2)SGMLとHTMLの共存
- 2. 6 SGML関連の動き
- 2. 7 公開されているDTDのリスト

第3章 見学会・勉強会の報告

- 3. 1 日本IBMでの調査報告
 - (1)「ドキュメント・データ標準」化の現状認識とIBMへのヒヤリング目的
 - (2)SGML化(ドキュメント・データの標準化)へのIBMの取り組み
- 3. 2 SGMLツールの実習

- (1) 文書構造
- (2) SGML文書の作成法
- (3) MS-Wordを使用したSGML文書作成
- (4) SGMLエディタの実習
- (5) 感想

第4章 まとめ

[参考:略語一覧]

エグゼクティブ・サマリー

「ドキュメント・データの標準」という場合、それが企業の中で、すなわち、仕事の中で使われることを想定してみていただきたい。これまでの系列を中核にした日本のビジネスは、図1のAS-IS(現状)モデルで現せるが、そこでは、企業内部(系列を含む)での効率化を図るために独自の仕事のやり方と言葉(用語)が生まれ一つの商慣行に基づく文化圏を形成してきた。コンピュータシステムとネットワークの導入により、この独自の商慣行がアプリケーションプログラムとして具現化され、その企業内でのデータ標準が確立されたのである。

企業は、データの蓄積と流れをコントロールすることで配下の企業を囲い込み、系列を電子的に固定することになったが、時代とともに系列外との取引が拡大され、系列ごとに異なる伝票やコンピュータアプリケーション、データ形式やネットワークの隔たりを乗り越えることが求められるようになってきた。顕著に行われるようになってきた外部との情報のやり取りに応えるために、電子データ交換(EDI)のための標準が制定されることになったが、この段階では各々の企業の社内システムにはメスを入れることなく、シンタックスルールとして定められたデータ項目をインタフェースとして利用するものであったゆえに、仕事のやり方とデータ標準の統合に対する根本的な解決策とはなりえない、という意味で限界があった(図1「仕事のやり方とデータの関係」参照)。

図1 仕事のやり方とデータの関係

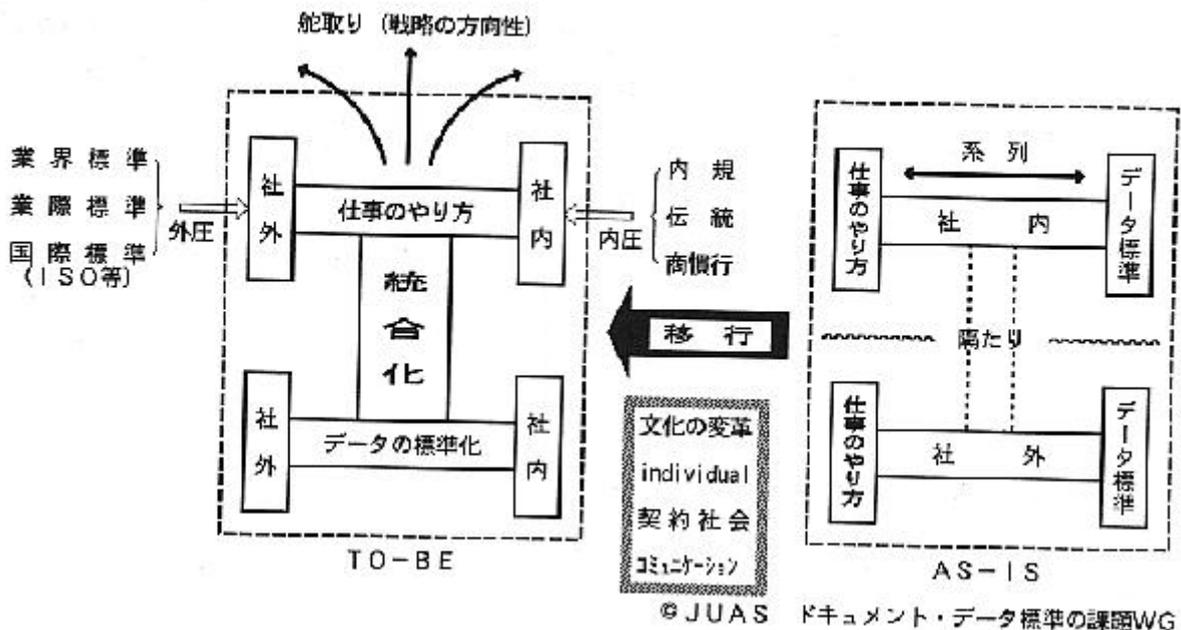


図1「仕事のやり方とデータの関係」

一方、ドキュメントという「紙」ベースが前提となっていたが、情報技術の発展とともに、「紙」という媒体に代わり、あらゆる情報を電子化しドキュメントデータとして扱い、それを高速大容量のネットワークで瞬時に遠隔地に電送したり、データベースで管理するという仕掛けが現実のものになりつつある。

この情報化の起爆剤として期待されているものがインターネットを通じたエレクトロニック・コマースとCALS(Continuous Acquisition And Lifecycle Support)プロジェクト(生産・調達・運用支援統合情報システム技術研究組合)である。これらにより実現される効果は、情報交換の効率化と情報の共有化である。情報を保管する媒体が紙から電子ファイルへ移行することは、情報の管理業務のみならず、業務そのもののやり方に影響を与えることになった。

従来の情報システムは、そのアウトプットとして大量の帳票(「紙」)を生み出してきたが、競争の激しい時代、型にはめられた成果物としての情報では、企業は身動きができない状況になってきている。

競争関係の激化とともに早くなってきたペースに追い付くために、企業では、いかに効率よく情報を再利用することができるか、が求められるようになってきた。このため、情報システムを有効に活用しつつ仕事のやり方を見直すBPR(Business Process Reengineering)(紙ベースから電子ベース)が脚光を浴び、社内外を通じて仕事のやり方の標準化とそれに適合したデータ標準を基にした新たなビジネス(TO-BE)モデルが求められるようになってきたのである(図1)。このビジネスの世界では、社内外の情報やリソースを電子的に自由に活用しながら舵とりをしていくことが求められる。

今後は、ペーパーレス時代にどう生き残るかという発想で企業の再構築を考え、まとめ、その中でドキュメント・データの標準化にどう取り組んでいくかが課題となっていくであろう。

すなわち、ドキュメントの観点からいうと、仕事のやり方に統合されたドキュメントの構造の標準化、ドキュメント管理の自動化、およびドキュメント情報の交換や共有化が、ドキュメントのライフサイクルにわたり、いかに効率的に行える仕組みを構築できるかである。この仕組みのメリットを十二分に享受するには、ドキュメントを扱うすべての仕事のやり方の変革が必要となる。また、それにつれて、商売のやり方も従来と大きく異なってくることになる。

その仕組みの要素技術としてSGML(Standard Generalized Markup Language)をとらえることができるのである。

その際、企業がドキュメントを電子化する動機付けをするトリガーが必要になる。そのドライバーとしては、以下のようなドキュメントを電子化して扱うことによる省力効果が考えられる。

- 社内外で情報交換が行われるもの
- ドキュメント量が大量なもので、検索や改訂が頻繁に行われるもの
- 将来の再利用を前提として長期間にわたり保存されているもの
- 共同著作が行われるもの
- 翻訳が行われるもの
- 紙、CD-ROM等の複数のメディアで利用されるもの

これらに合致したドキュメントとして、例えば技術ドキュメントやマニュアルを挙げることができる。

しかし、日本では、こうした電子化に積極的に推進していくメリットを理解しきっていないのが現状である。これは、一般的に企業がドキュメント作成・維持・管理に使っている費用を正確に捉えることが難しい状況にあり、それゆえ、電子化の費用ばかりが目立ちすぎている状況にあるといえよう(※参照)。

また、ペーパーレスといっても、依然として「紙」が存在するわけで、その部分も含んだ取り扱いが課題となってきている。しかも、この点をクリアしたとしても、現状認識としてビジネスがどう行われ、どういう情報がどういう意味を持つかの分析と標準化して理解する上での曖昧さをなくすというプロセスが新たな関門としてそびえ立っている。その際、どう情報を活用するかという視点からどう電子ドキュメントを設計するかのノウハウがユーザー側にないことも問題である。

いずれにしても、「1億総情報化武装時代」とも呼べるような時代背景にあって、先に述べたように企業が競争激化の時代に生き残り、かつ情報の電子化という“仕組み”のメリットを享受するためには、克服すべき難題が山積していることを踏まえていただくとともに、本報告が、SGML導入を検討しているユーザーの参考になれば幸いである。

※ 一例としてエンジニアリング業のドキュメントハンドリング電子化の定量評価と定性評価として以下のような報告がある。

ドキュメント (社内調整分)	15,000枚
ドキュメント (社外調整分)	159,000枚
承認に要する時間	21日
コピー～メール～配布の時間	5日
変更回数 (社内ドキュメント)	2.9回
変更回数 (社外ドキュメント)	1.1回
マンアワー(種類検索)	2.7%
マンアワー(コピー)	2.9%
マンアワー(ファイリング)	4.1%
総計	9.7%
その他 事務所費 メール代 通信費・出張費	

(出展:【IBM CALS セミナー】CALS のプラント・エンジニアリングへの適用)

第1章 はじめに

標準化動向調査部会の活動は、今回で3年目を迎えた。初年度は“標準化の課題はどこに、どれだけ存在し、ユーザーとしてはどのようにしてほしいのか”等をまとめ、次年度は3WGに分割し初年度の報告書から、ユーザーにとって重要と考えられるテーマを選び、より具体的な内容で報告書を作成した。今年度は、ドキュメント・データ標準の課題WGが活動を継続して行い、CALISの最重要要素技術の1つであるSGMLに活動の焦点を絞り込み、検討を続けた。

活動に先立ち、なぜ今SGMLなのか、検討しなければならないのかについて、メンバー相互に意見交換を行い、検討課題をまとめた。

1.1 背景

ドキュメントデータの標準化に関連する背景とドキュメント電子化に係る背景は、以下のようなものあげられる。

(1)標準化の背景

- システムに依存せずにデータ交換をしたいというニーズがあるが、オープン、オープンといわれる割にデータの流通ができない。
- データ形式が社内業務やコンピュータシステムに依存しており、自由な交換の妨げとなってきた。
- 経営資産としての情報の有効利用と再利用で、いかに付加価値を生むかが課題となってきた。
- 情報技術と情報通信ネットワークの進歩により企業内、企業間のデータ交換が可能となってきた。
- 標準化の進行の中でドキュメントデータ標準化に目の目があたってきている。

(2)ドキュメント電子化の背景

- 外圧(海外からの電子化要求)
- 経営のやり方(情報を交換するニーズ)
- ツールおよび電子媒体の普及
- 印刷や改版の効率化
- 執筆の電子化
- 翻訳作業の効率化

1.2 活動目的

一昨年度の「標準化動向の現状とユーザーの問題意識－ユーザーサイドから見た標準化の問題点の検証－」にあるように、ワープロやパソコン上のワープロソフトウェアの利用により電子的に調製されたドキュメントが流通されるようになってきた。

これに伴い、異なった機種間でのデータの交換や共有が大きな課題となって浮かび上がってきた。異機種間でのデータの互換性が完全にとれていない現状では、例えばプリンタでドキュメントとして紙に出力し再入力する以外、ドキュメントのデータを利用することができない。ドキュメントそのものの互換性がとりにくいという点と標準化の動向、とりわけ技術分野でのドキュメント・データの標準化として、CAL(S (当時は Computer-aided Acquisition and Logisitric Support) に係る活動が活発になっていることが報告されている。

今日のビジネスにおいて、効率的な経営と並んで、経営資産としての「情報」の有効かつ高付加価値を生むような活用が望まれるようになってきている。さらに、電子出版やインターネットの普及によりビジネスの種としての電子化された情報に係る需要が急に伸びてきている。

電子化されたドキュメントを扱う標準としては、SGML (Standard Generalized Markup Language)、ODA (Open Document Architecture)、OPENDOC等があるが、ツール群が整いつつあるSGMLがCAL(Sブームの中で脚光を浴びてきている。

SGMLとは、ドキュメントの記述形式の国際標準(ISO8879)である。SGMLでは、ドキュメントの構造を記述するためにドキュメントデータのうちアプリケーションソフトウェアに固有な体裁情報を捨象し、ドキュメントデータのみをテキストデータとして扱いつつ、「表題」「目次」「段落」といったドキュメントの論理構造に係るデータを「タグ」として各々対応するドキュメント文書部分の前後に埋め込む構造となっている(注1参照)。

それゆえ、SGMLは、上述したドキュメント・データ標準に係る課題への一つの回答となりえるかもしれないドキュメント・データの有効的活用を促すドキュメント処理の新しい形とも言えよう。本年度の当WGの活動としてSGMLの調査を通してドキュメントデータ標準に関する課題の報告を行うことにする。

(注1)SGMLの詳細については、平成6年度活動報告書「標準化推進委員会・標準化動向調査部会報告書 ドキュメント・データ標準化に向けて」を参照。

1.3 WGの役割と検討項目

WGでは、限られた時間の中でドキュメントデータ標準とSGMLの役割を検討した。主な項目を挙げると、ドキュメントデータ標準の観点から、

- ドキュメントの構造について
- ドキュメント構造とSGMLの位置付け
- SGML勉強会を通じてのSGMLの理解と体得

- モデルSGML利用環境手順の紹介
- SGMLの利用普及状況:
導入業種、対象文書、文書量、管理方法、公開されているDTD、導入効果、今後の予想
- SGMLツール群の位置付けとチェックポイント
- SGML実践的活用法:
テンプレートの活用、Internet (HTML)との連携等を取り上げた。

また、活動スケジュールは以下のとおりである。

第 1 回研究部会	7 月 2 0 日	キックオフミーティング
第 2 回研究部会	8 月 3 1 日	今後の進め方
第 3 回研究部会	9 月 2 7 日	I B M大和 見学会
第 4 回研究部会	1 0 月 1 2 日	構造化文書
第 5 回研究部会	1 1 月 7 日	延期
第 6 回研究部会	1 1 月 1 3 日	文書の SGML 化 1
第 7 回研究部会	1 1 月 2 8 日	文書の SGML 化 2
第 8 回研究部会	1 2 月 7 日	SGML 化の実践応用編
第 9 回研究部会	1 2 月 1 2 日	本年度の活動のまとめについて
第 1 1 回研究部会	1 月 2 3 日	D y n a T e x t の実践事例
第 1 2 回研究部会	1 月 2 5 日	本年度の活動のまとめについて
第 1 3 回研究部会	2 月 2 2 日	本年度の活動のまとめについて
第 1 4 回研究部会	3 月 2 8 日	本年度の活動のまとめについて
第 1 5 回研究部会	4 月 2 6 日	最終打ち合わせ

1.4 今後の展望として

ユーザーの視点から受動的な立場で、ドキュメント・データ標準に対してSGMLの活用の可能性についての見解を出そうとして活動してきた。

SGML導入のポイントとしては、ドキュメント・データ標準から見ると、以下のようなことを押さえておく必要があると思われる。

1. SGML導入効果
2. ワープロとSGMLの連携
3. マルチメディア対応ツールとしてのSGML
4. DBとSGMLの連携
5. HTMLとSGML

6. SGML関連の動き
7. 公開されているDTDのリスト

第2章 WGの活動内容

2.1 SGML導入効果

ここでは、SGML導入により期待される効果、導入効果の上がる事例、SGML導入のために必要な投資／労力、そして導入効果の評価について述べる。

(1) SGML化による期待効果

SGML化による期待効果はさまざまなところに広がる。電子化による文章(情報)の作成の効率化はもちろんのことではあるが、大きな効果は、

- 文章作成の生産性向上が可能
- 再利用性向上による効率化が可能
- ワープロなどの機種依存性がなく、長期にわたってもデータの保存が可能
- 大量のデータを規則的に処理することが可能(品質の向上)
- 企業内はもちろん、企業間の共通データフォーマットとして利用が可能
- ドキュメントのオンライン配布(ネットワーク、CD-ROM)化が可能

などに現れる。

(2) SGMLの導入効果のある事例

一般的にSGML導入の効果があると思われる事例としては、保存と再利用をするドキュメント類がある。例えば、社内規定、法律条文、ISO9000のマニュアル、コンピュータのマニュアル類などで、それら配布と更新、分散環境でのドキュメント管理(協調作業環境、グループオーサリング 注1参照)、翻訳などに効果をもたらす。一方、定型文であっても、単なる打ち合わせ文書、会議議事録などは、それがそのまま再利用されることは稀であり、SGML化することが必要かどうかは難しいところである。もちろん、議事録などであっても、電子化され、あるレベルの管理がなされる場合には、SGML化が有効なのは言うまでもない。

この他、効果があると思われる事例としては、DoDにおけるドキュメント問題や国際調達の電子化(グローバル市場での電子取引、インターネット上でのEC)が考えられる。しかし、安易な動機で「その場しのぎの場当りの対症療法」としての導入では、本来の効果(仕事のやり方とデータの融合)は期待できない場合が多い。

図2-1-1 SGMLか対象の特性

○大量のデータがあるもの

- マニュアルなど言葉の整合性が必要なもの
- ハイパーリンクの必要なもの
- 再利用するもの
- 長期保存するもの
- 多種のメディアへ出力するもの

などなど

(3) SGML化のための投資と労力

前述のような、SGML化による効果を引き出すためには、人的にも金銭的にも、ある程度の投資が必要である。この投資／労力と、それに対しての期待効果を考え合わせて、SGML化を決定する必要がある。

まず、トップを含めた管理者層のきちんとした理解／意識改革が前提でないと、その導入は円滑に行われぬ。現状体制の見直し、新しい環境に対しての教育などは管理者層の理解なしには進められない。

また、SGML導入にはいくつかの山が存在し、その山を越えたときに徐々に効果が表れる。ある部分については、近未来には効果の恩恵にあずかれず、将来その効果を見出せる部分も存在する。したがって、あらかじめ投資の大きさを見積り、すぐに効果が出ないからと言って騒ぐようなことのないようにしないといけない。

SGMLと直接関係ある話ではないが、人間は常に環境に対しては保守的であり、現状の環境を変えることには抵抗を示す。この抵抗がSGMLの速やかな導入を阻害する要因になるので、SGML導入によるメリットを常に見せて、理解を得ながら進めていくとよい。

また、SGMLを利用する上で鍵になるのがDTDの選択／定義である。自分の業務を的確に分析し、最適なDTD(注2参照)を選択したり、定義したりできるノウハウを自社内にもっているか、もしくはコンサルテーションを受けられるかが、SGML導入の成功の鍵を握っていることにも留意すべきである。

(4) SGML導入効果の評価

さてSGMLを実際に導入した場合の効果は、どのように評価したらよいだろうか。以下のような項目を検討する必要があると考えられる。

- 現状のドキュメント関連にかかっているコスト
- SGML導入による効果(どのような効果を、どの程度期待しているか)
- SGML導入のための投資

現状のドキュメント関連にかかっているコストとしては、原稿の作成作業はもちろんのことながら、その他にも校正作業、レイアウト作業、印刷作業、翻訳作業、配送作業、改訂作業な

さまざまな部分(図2-1-2参照 IBM CALSセミナー資料)について考える必要がある。広い意味では、電話、FAXなどによるやりとりのコストについてもとらえておくのが有用であろう。これらを総合的にとらえて、次の「SGML導入による効果をどこに求めるか」を検討する必要がある。

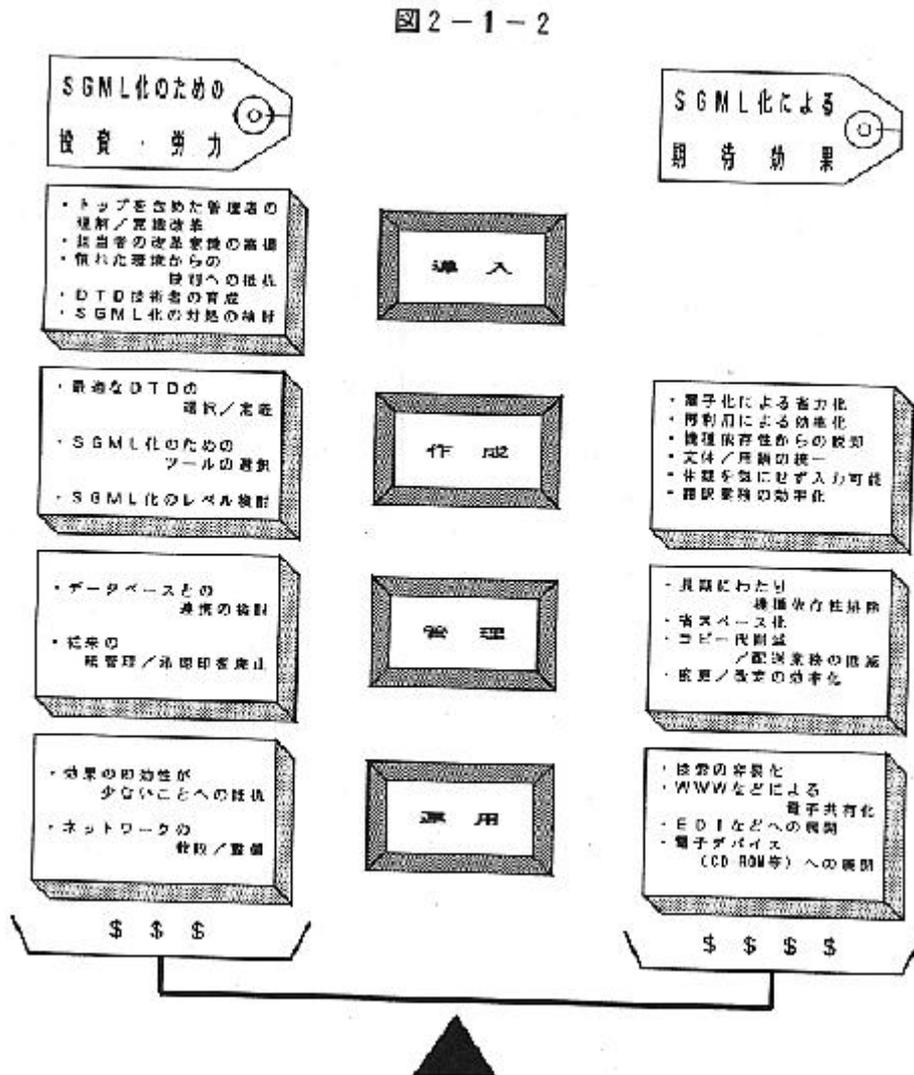


図2-1-2

SGMLをただ、安易に流行りだからといって導入しようとしても失敗するのは目に見えている。SGMLの特性をきちんと理解し、現状の問題点、将来的な形をきちんととらえた上で、SGMLを導入することによる効果をどの程度、どの部分に求めるのか目標を立てる必要がある。

いずれにせよ、SGMLの導入には、「自分の業務をどうにかしたい」というユーザーの問題意識が大前提として必要であり、この前提が存在した上でのSGML導入は大きな効果をもたらすことになる。

(注1)グループでの作業効率を高めるための仕組みとして、グループウェア・アプリケーションも増えているが、これらは依然として社内での効率化を目指したものであり、会社間までも含めた仕組みにはなっていない。

(注2)SGMLの規格であるISO8879(JIS X4151)に含まれる General DTD や、NCALSの中で決まりつつあるDTDなども参考にできる可能性がある。これ以外にも WWW、nonyous FTPなどで、さまざまなDTDが公開されている。

2.2 ワープロとSGMLの連携

ワープロはだれでも容易に扱うことができるため、企業や家庭(個人)で広く使われるようになった。主に ペンやタイプライターの代わりに利用され、いかに美しい文字やレイアウトができるか、各社各様の形式でボトムアップ的に発達してきた。結果的に内部形式の標準化や文書構造等の情報処理としての側面が考慮されない形になり、ワープロ間の情報交換や文書の共有に大きな問題を残すことになった。

SGMLは、各種文書を電子的な媒体で交換したり、共有したり、データベース化するのを容易にするという目的で作られた国際規格である。SGML化に際し「ドキュメント構造、構造化」が重要なポイントとなる。文書構造には物理構造と論理構造とがあり、罫線、書体、文字サイズ、割り付け等といった紙の上での見栄えをよくするための要素による物理構造よりも、題名、著者名、章、節、項等の要素による論理構造の方が文書にとってより本質的で普遍的な情報となる。その論理構造を具現化するための手段としてDTDがあり、その中でタグが大切な役割を果たす。すなわち、タグの名前とその中味の定義および構造化が大切であるということである。

SGMLドキュメントの作成は、そのドキュメントの性格に従いその構造を定義し、それを表すDTDを作成し、DTDに基づいたタグ付けをしながら実際のドキュメントを作成していかなければならない。しかし、すべての文書をSGML化すればよいわけではなく、伝票等の定型書類、技術書や操作手引書等のマニュアル類のように文書構造が明確なものが摘要の対象となり、メモ等、思い付きで書いた文書構造が曖昧なものや文書構造を持たないものは適さない。ドキュメント管理の側面からいって個別なDTDを作るのではなく、共有を意識した設計が必要となり、その意味ではシステムインフラストラクチャ的な色彩を帯びている。すなわち、オープンなハードウェア環境においてアプリケーション間でのドキュメントデータの相互運用性の確保という観点から真のオープンシステムということになる。また、DTDを各社で個別に定義し公開しては負担が大きく、企業間で交換をする場合にDTDの解釈に時間がかかる。そこで、EDIと同じように業界単位で標準DTDを定義し公開することが望まれる。

ワープロとSGMLの連携(保存と再利用)として、テンプレートを活用することが考えられる。社内ユーザーに対してのワープロ講習会で、社内定型文のテンプレートを使わせるようにしたり、社外とのやり取りでテンプレートを配布して使ってもらえれば、後日、テンプレートを使って書かれた文書をSGML化する場合に、変換スタイル情報とSGMLのタグ対応表を

使って比較的簡単に変換できる。また、Microsoft のRTF文書や FrameTechnology のMIF文書のように文書構造がテキスト形式で開示されている場合はこれを利用してSGMLに変換することもできる。SGML化しておけば紙だけでなくCD-ROM等の他の媒体への転記も可能となる。最近ではノーツのようなグループウェア製品に、作成したコンテンツをほぼ自動的にHTML形式に変換できる機能やWeb機能が付き始めた。

テンプレートと同様の汎用のDTDが組み込まれる時期がきて、ユーザーが「意識しなくてもワープロ感覚でSGML化できるツール」が出てくることが望まれる。

2.3 マルチメディア対応ツールとしてのSGML

一般にドキュメントにはテキストデータだけでなく図、表などが含まれる。SGMLではSGMLの記述とは独立にファイルとして格納される図、さらには写真等の静止画、動画、音声等の情報全てを外部データとして扱うことができ、マルチメディアに対応している。

ここでは外部データファイルとして格納されている図のSGMLでの扱いについて一例を示す。そのDTDの定義例を以下に示す。

```
<!ELEMENT      :   -0      EMPTY
<!ATTLIST  figure  id      ID      #IMPLIED
           figure  file    CDATA   #IMPLIED
           type    CDATA   #IMPLIED
           sisex   CDATA   #IMPLIED
           sizey   CDATA   #IMPLIED>
<!ELEMENT  figure -0      EMPTY>
<!ATTLIST  figure  refid   IDERF   #IMPLIED>
:
```

ここでエレメント名 figure は属性として識別子値である id、外部データとして図が格納されている ファイル名を示す file、そのファイルがどのような形式で格納されているかを示す type、さらに図の大きさを示す sisex、ならびに sizey から 構成されると定義されている。また、同じ図をドキュメント中で別の場所でも参照できるようにエレメント名 figref が定義されている。

一方、このDTDに従ってSGMLのテキスト中では、例えば以下のように記述することで該当の図を挿入できる。

```
・・・<figure id=fig1 file="b:¥figure¥figure1.pict" type=PICT>・・・
```

ここで挿入される図は、“b:¥figure¥figure1.pict”として格納される PICT 形式のファイルである。

テキスト中の箇所でのこの図を参照する場合には、該当箇所での下記のとおり記述すればよい。

……<figref refid=fig1>……

マルチメディア情報の大きな特徴として時間の制御がある。SGMLは上記例に示すとおり外部データとして扱うことで、2次元空間上に文字、写真、図版等を組み込むことができる。マルチメディアではさらにこれに音声、動画等時間軸上にある要素を付加したものと考えられる。言い古されてはいるが、マルチメディア化により情報そのものが2次元空間(紙)から3次元の空間へ広がって作成、転送、蓄積できるようになるということである。

これは従来のドキュメントに時間の概念を入れる、すなわち時間の経過とともに変遷する情報の移り変わりそのものを扱うことができるということを意味しており、このようにして作成された情報は、その変遷そのものが貴重な情報であるということから、情報の陳腐化が起きないという大きな特徴を有することになる。こうなると極論すれば、そのドキュメント自体を未来永劫容易に利用しておく必要があり、その標準化がますます重要となる。

従来のSGMLでもこれらマルチメディアのデータを扱うことができるが、さらに文書間と文書内にハイパーリンクを指定し、時間的および空間的なマルチメディア情報のスケジューリングを行うための標準的な機構を提供することでハイパー文書の交換を可能にする規格、すなわちSGMLのマルチメディア向け拡張版として [HyTime](#) が開発されている。

HyTimeについては、本WGの昨年度の報告書で報告しており、これを参照していただきたい。

SGMLをマルチメディア情報の表現、処理手段として利用するだけでなく、さらに知的所有権や肖像権、著作権の管理にも利用し、従来にないきめの細かいサービスを提供するマルチメディア情報検索システムの開発がNTTで進められている。

従来、マルチメディア素材をアクセスする場合、マルチメディア素材に対してすべての知的財産権を買い取ったものか、あるいはすべて許諾されたものが、専用のビデオサーバーやCD-ROM等に格納され、これをアクセスしていた。しかしながら、このようなやり方では、利用できるマルチメディア素材そのものに限りがあり、この限界から普及の大きな妨げとなっている。

あるマルチメディア素材について、限られた一部ならば自由にアクセスが可能でも、その他の部分はアクセスに対価を支払う必要がある。しかもその対価は、さらに部分によって異なるなど、一般的にマルチメディア素材のアクセス条件は極めて複雑であるとするべきであり、これらのアクセス条件に対応してきめの細かい制御ができれば非常に有効である。

ここでいうアクセス条件とは、そのマルチメディア素材に対して必要に応じて分割された個々の部分に対するID、パスワード等の使用者資格情報、肖像権、著作権等の知的財産権に関する各種規定、ロイヤリティ、イニシャルペイメント等の使用料金、製作元問い合わせ先情報、使用上の注意事項や、警告文等の使用者へのメッセージ等の制御情報を基にしてアクセスできるかどうかを判断する条件のことである。

NTTで開発を進めているシステムは、図2-3-1に示すとおり、これら制御情報を該当のマルチメディア素材と対して管理する手段としてSGMLを利用している。SGMLはテキストを含むマルチメディアデータを自由に管理できるので、一度マルチメディア素材と制御情報を対してSGMLで記述すれば自然にパッケージ化され、格納される媒体に依存せ

ず流通、保存が容易となる。しかも制御情報自体がテキストで書かれているので、実状の変化に合わせた制御情報の修正、変更を容易に行うことができるという利点もある。

図2-3-1

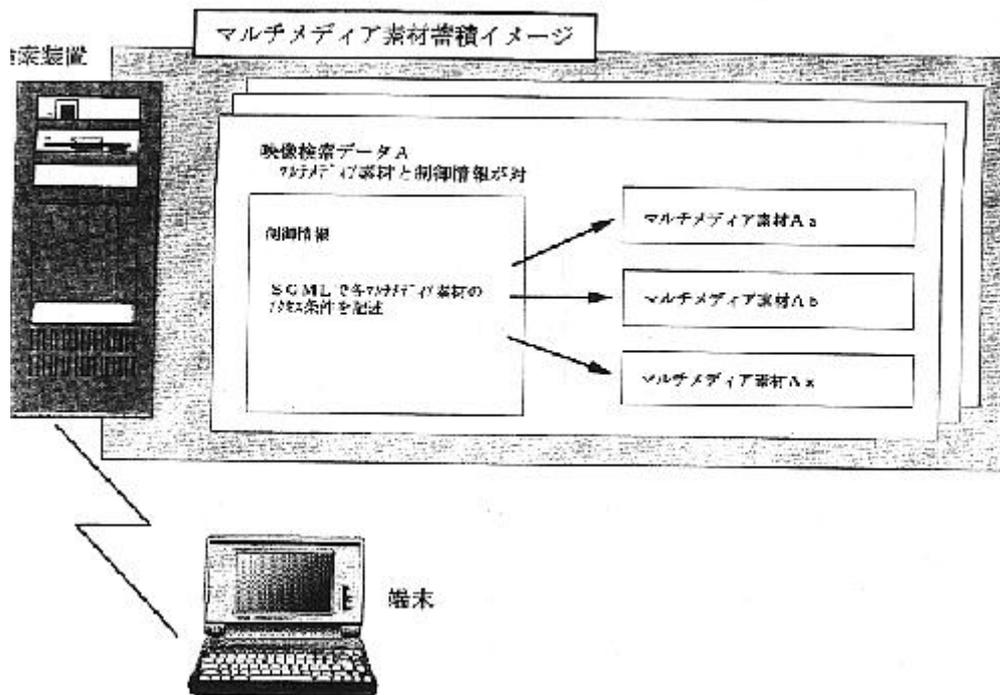


図2-3-1

SGMLで書かれたこの制御情報を検索装置がインタープリットしていき、例えば使用者資格の問い合わせを端末に表示、利用者の入力された情報がアクセス条件を満たしていれば、該当のマルチメディア素材を端末装置に表示するといった処理を通して、まずはさわりの部分だけを表示して、全体を見たい場合にはそこで対価を要求するといったサービス、あるいは、あらかじめ入力された使用者資格の条件に従って、部分的に表示をカットするサービス等アイデア次第で魅力的なマルチメディア素材検索サービスを自由に実現できるようになる。

2.4 SGMLの業務利用

SGMLは、文書の内容や書式に触れることなく文書の構造を表現する規格として発展してきた。しかし規格の中には、タグを付けるにあたってどのような目的にそれを用いるかという制約はなく、タグで何を表現するかは利用者の裁量に任されている。近年のSGMLの動向を見ると、SGMLの考案された当初の趣旨からは逸脱するかもしれないが、むしろ積極的

に文書の内容を表現するタグをつけ、文書中のデータ項目を識別させることによって、業務アプリケーションの中で活用していこうという動きが目立ってきている。

(1) J2008 のDTD

SGMLのタグの中に、文書の持っているデータ項目を反映させていこうという傾向の一つの例が、J2008と呼ばれている、自動車の保守マニュアルを記述するためのDTDである。このDTDは、次に示すとおり文書を管理するための書誌情報や、文書の中で使われるデータ項目をエレメントとしてタグ付けし、実現値はタグに含まれる属性(アトリビュート)として管理している。データ項目をエレメントとして識別するか、アトリビュートとして表現するかについては、微妙な判断になるが、最近のSGMLの動向を示す、一つの例であると言えるであろう。

(2) データベースとSGML

先に述べたように、SGMLのタグは書誌情報やデータ項目を識別するという用途に利用することができる。このことはデータの受渡しに際して、SGMLが有効に活用できることを示しているが、データが利用者の手元に届き、蓄積・活用されようという場面では、SGML文書をそのまま保管していたのでは効率が悪い。

SGML文書はあくまで唯のテキストファイルであって、なんらかの検索をかける場合には、タグを活用し効率をあげるとはいっても、所詮は全文検索(フルテキストサーチ)の域を出ない。そこで業務系のシステムを考える場合にはデータベースとの連携について考えておかなければならない。

SGMLを議論するときに、暗黙のうちに報告書や論文といった構造化された文・段落構造を持った文書を引き合いに出すことが多いが、一つのケースとして製品カタログをSGMLで表現する場合を考えてみよう。例えば次のようになる。

これはカタログを表したSGMLテキストであり、適切な印刷プログラムやビューアソフトを用意すれば、しかるべき形式のカタログとして再構成することができる。と同時に、このSGMLテキストをデータベースの更新プログラムに対する入力データとして用いることもできる。

例えば購入者は、同じDTDでカタログが表現されてさえいれば、上記XX株式会社の製品のみならず、株式会社YYやZZ商店のカタログを集めて、容易に購買品目のデータベースを構築することができる。

データベースを更新して以降は、例えば次のような形でSGML文を保存しておけば、新しいデータが登録されても、常に最新のカタログとして参照することができる(作成日・有効起源・前書などの各文章要素との整合性の保持は、応用プログラム側で処理すべき問題である)。

(3) EDIとSGML

SGMLを用いた実業務上のデータの交換に際して、データベースとの連携と同時に考慮しておくもう一つの問題は、EDIとの連携である。EDIは、発信側に対する送達・開封通知や発信者・受信者の認証問題など、商業利用を目的とした通信のプラットフォームとして、今後の役割拡大が予想される。例えば提案書がSGMLで記述され、入札が電子化されるといような姿を考えると、EDIのプロトコルに従ってSGML文が送達されるようになる日も遠くはない。

EDIの実現方法は、オンライン・オフラインの区別を含めて、いくつかのパターンが考えられるが、基本となるのは、文書の持っている取引プロセス上のステータス(提案書なのか要求仕様書なのか発注書なのか)や発信者・受信者の指定などの商業利用上の必須項目を、データの一部としてではなく、ヘッダー情報としてEDIのプロトコル自身が管理しようとする考え方である。

提案書を例にとって考えると、その中には「この文書が提案書である」ことを示す情報や、提案者・宛て先・日付・対応する提案要請(RFP)など、先に述べた「商業利用上の必須項目」となるべき情報が記載されているはずである。その提案書がSGML化されており、その必須項目がエレメントまたはアトリビュートとしてタグ付けされていれば、SGML文書から機械的にEDIのヘッダー情報を生成することができる。

このことからわかるとおり、DTDの設計には文書の構造のみではなく、文書の使われ方までを含んだ業務モデル全体を視野においた、幅広い分析が必要とされるようになってきている。

[参考文献]

『はじめてのSGML』

(株)日本ユニテックSGMLサロン編著 1995.11.25. 技術評論社

2.5 HTMLとSGML

SGMLの応用として現在最も注目を浴びているのが、WWWのキーテクノロジーの一つHTMLであろう。むしろ、CD-ROMや電子出版という一部の業界に適用されてきた規格であるSGMLが広く市民権を得てきた背景の一つに、日本ではCA LSと並んでHTMLが挙げられるとすら言えるであろう。

ただしSGMLとHTMLの2つは、生まれは同じであってもそれぞれ別の進化の途をたどっている。SGMLが前項で述べたようにデータ項目の識別という、業務モデルに深くかかわった使い方をされ始めているのと反対に、HTMLはSGMLがかつて切り捨ててきたもう一つの文書の側面である書式定義を表現するように進化し始めている。

(1)HTMLの特色

SGMLと比較した場合のHTMLの長所は、比較的ビューイングの機能に重点を置いた標

準のDTDが存在し、そのDTD用に特化したビューア(Netscape など)やワードプロセッサからのコンバータが存在し、そしてWWWというわかりやすい応用分野が存在していることであろう。最近ではデータベースやグループウェアといった、別のジャンルのソフトウェアとの間のデータ連携も整備される傾向にあり、利用環境はますます充実してきている。

対するSGMLは、まず文書構造や文書利用業務を分析し、標準化を行い、DTDを作り、データの作成・蓄積・検索・表示・印刷などのアプリケーションを自ら開発するか、あるいは(幸運なケースでも)市販のパッケージソフトをカスタマイズする必要がある。

この大きな差は、目的を「表示」に限定してDTDを標準化してしまうというHTMLの思いきりのよさに起因しているが、同時にそれがHTMLの限界にもなっている。

HTMLはDTDがあらかじめ定められており、しかもその内容がビューイングの機能に重点を置いたものになっているため、SGMLが最近目指し始めた データ項目に対するタグ付けといった、文書の利用形態を反映した設計はとり込めない。わずかに可能性があるのは、HTMLは未知のタグを無視するという規約がある点であるが、このルールに依存したタグ付けをするのは(とくにHTMLの規約自体が進化しつつあることを考えると)、堅実な方針とは言えないであろう。

(2) SGMLとHTMLの共存

SGMLがCALISやECのツールとして脚光を浴び始め、社会インフラの中に組み入れられていくようになると、データベースやEDIプロトコルとの整合的な運用がますます求められるようになってくる。残念ながらHTMLは、そのような要請に応える規格ではない。しかし一方でHTMLは、WWWを構成する デファクトスタンダードの規格として、SGMLの応用分野として類例を見ない普及を遂げている。とにかく文書がHTMLになりさえすれば、WWWにのせて 全世界に情報発信できるわけであるから、CALISやECの立場からしても、これが利用できることには無視できないメリットがある。

例えば、提案要請(RFP)をWWWを通じて世界に公開し、未知の企業との取引に窓を開く。これは、長期的な取引関係を前提として成立してきた従来のビジネス慣行を崩すという両刃の剣であるが、副作用を抑えるうまいやり方が見つければ、受注者・発注者双方にとってビジネスチャンスの拡大につながり、世界的な規模での生産活動の最適化をもたらすきっかけになる可能性も秘めている。

このような可能性を視野に入れると、SGMLとHTMLは、どちらが正当的な技術で、どちらが実用的な技術であるかというような捉え方をするよりも、違った応用分野に向けて特化しつつある2つの技術として使い分けていく必要がある。

実際的には個別の業務モデルに依存してDTDが設計されるSGMLの方が、DTDがあらかじめ決まっているHTMLよりもタグの持っている情報量が多いため、業務上利用しているSGMLを必要に応じてHTMLに変換するというやり方が現実的なものになるであろう。

試みに先ほど示したカタログの情報をHTML化した場合、図2-5-1のようになる。

図2-5-1 WWWビューアから見たカタログ

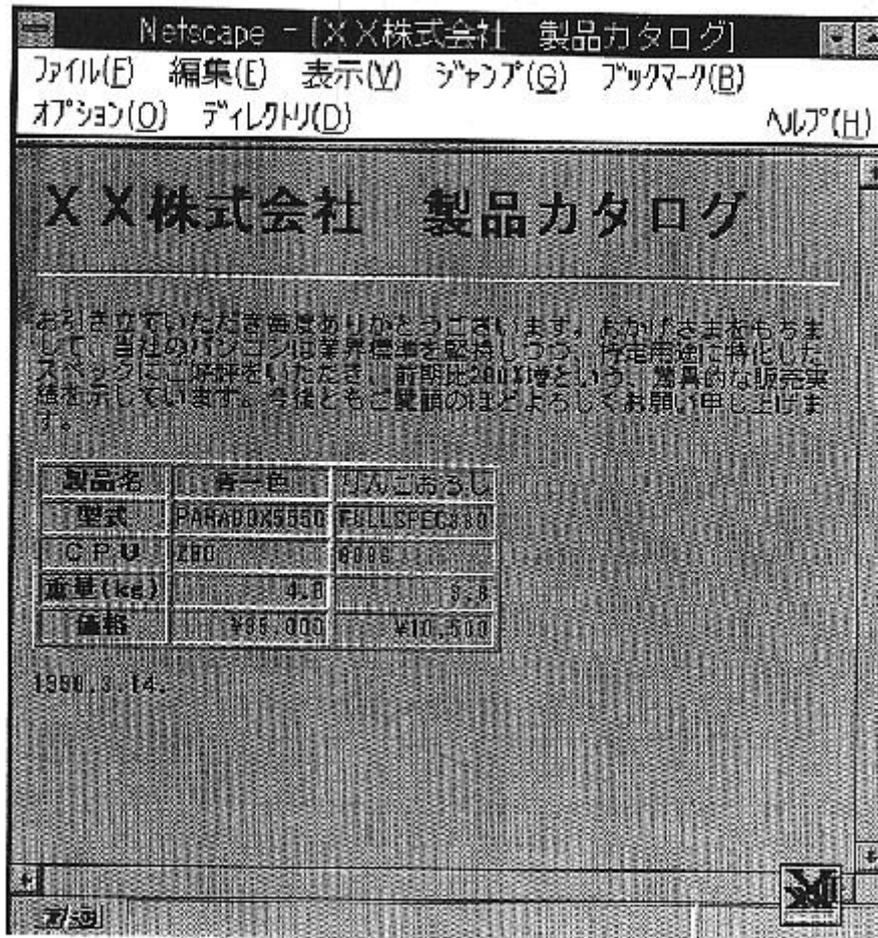


図2-5-1

2.6 SGML関連の動き

SGMLの今後の展開を占う上で、CALsの動向は見逃せなくなっている。周知のとおり、SGMLはCALsの起源であるDoDのCALsの中で、MIL-M-28001(電子印刷出力とSGMLテキスト交換のためのマークアップ要求)として規定されており、そのDTDとしては、DoD-STD-2167Aなどがある。

特に日本ではCALsに関する大きな動きとして、通商産業省が主導して95年5月に生産・調整・運用支援統合情報システム技術研究組合(NCALs)、並びにCALs推進協議会(CIF)が発足して、日本の産業界におけるCALs標準の作成、並びにその普及活動を開始している。その成果も一部ではあるがすでに現れてきている。

95年の10月30日から11月2日までの間、CALs推進協議会が主催して東京・晴海の国際見本市会場で開催されたCALs Pacific '95においてVE2006のデモが実施され、非常に注目を浴びた。VE2006とは、西暦2006年にデジタル革命があらゆる局面に行き渡り、バーチャル・エンタープライズが日常的に実施されていることを想定し、その予想されるバー

チャル・エンタープライズを現在の情報技術やインフラストラクチャーを使って実現させたものである。

このVE2006のデモでは自動車のリアランプ部分のマイナーチェンジを想定してそのプロセスをCALSで実現したものであるが、もちろんこのデモの中でSGMLは非常に重要な役割を担っていた。

企画からメンテナンスまでの全工程で発生するドキュメントの作成・閲覧に関する各種文書について、時間的な制約から共通のDTDを作成・利用するまでには至らなかったが、SGMLフォーマットの利用によりデータ交換はかなり円滑に行うことができたと報告されている。96年以降のCALS PacificではCALS技術研究組合の成果が順次デモされる予定と聞いており、目が離せない状況である。

一方、工業技術院配下のINSTACにおいてソフトウェアの生産をCALSの枠組みで実現しようとする試み、並びにその場合の標準仕様の作成の場としてソフトウェアCALS標準化委員会が95年11月に発足している。この委員会の配下のプロダクトデータ交換形式検討分科会においてSGMLが検討されてきている。

当然のことながら、この分科会の検討スコープにはソフトウェアの開発において特有であるCASEデータ等の交換も含まれているが、要求仕様書等の企業間で取り交わされるドキュメントのSGML化は重要な検討課題となっている。95年度は、参加各社の要求仕様書を目次項目レベルで予備的に検討した結果、記述項目に大きな差が見られないということがわかり、96年度以降、要求仕様のDTD案の検討を開始することになっている。このとき国際連携のソフトウェア開発を視野に入れ、世界に通用する仕様の開発を行うという立場から先に挙げたDoD-STD-2167A、NASA-SFW-DID-08、IEEE/ANCI830-1984等も参考とする予定になっているとのことである。

またCASEデータ、図式や画像等のデータについてもそのデータ形式自体は、それぞれに適したものとするが、SGMLの外部参照の機能を利用して全体を統一的に扱えるようにしていくことを考えている。

このようにCALSの動きに支えられて、SGMLの産業界への利用が非常に注目されるようになってきている。しかし、本来標準仕様としての役割を果たすためには、業種・業界横断的にDTD等を維持、管理していく組織が必要になると考えられるのに対し、残念ながらSGMLに関するそのような組織は現在のところ存在しない。CALSの基本技術であるSTEP、EDIともにSTEP推進センター、あるいはCIIなど業界を束ねた組織がすでにあり、日本における、あるいは国際連携上重要な役割を担っているのに対し、SGMLについては今のところそのような動きが見受けられないのは非常に残念である。

ここでは、図2-6-1に示すようなSGML(DTD)の普及を図る組織、仕組みが必要であるということが、本節の最も重要な提言であるという我々WGの強い共通認識があったことを最後に示す。

図2-6-1

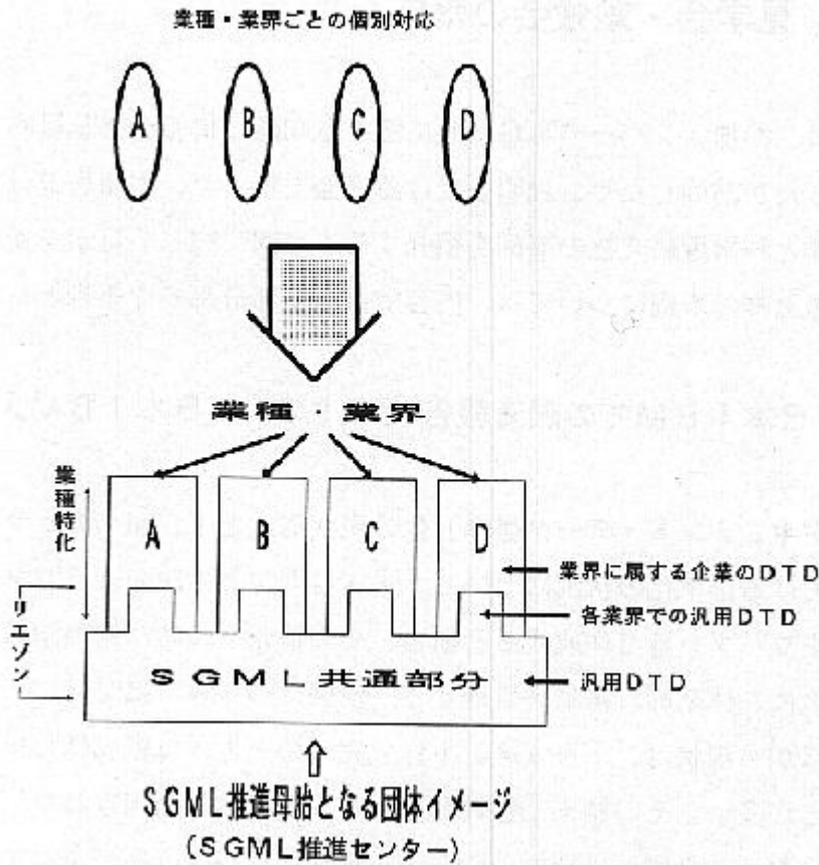


図2-6-1

2.7 公開されているDTDのリスト

WGでは、公開されている主なDTDを一覧表にまとめたが、この他にも多くのDTDがインターネット上等で公開されていることを付記しておく。

第3章 見学会・勉強会の報告

WGでは、参加メンバーのSGMLに関する知識と情報を得る目的で、限られた企業ではあったが訪問したり、説明を受ける機会を持った。本報告では、日本アイ・ビー・エム(株)と株式会社学習研究社の事例を紹介する。今回、同じく見学を受け入れていただいた日本電気(株)の事例については、内容で重なる部分が多く報告としては記載しない。

3.1 日本IBMでの調査報告(95.9.27 : 日本IBM大和事業所)

(1)「ドキュメント・データ標準」化の現状認識とIBMへのヒヤリング目的

日本における標準化の状況について、本WGの認識は次のようである。

「文書をワープロ等で作成する習慣と、その情報を瞬時に遠隔地に電送したり、データベースに、体系的に蓄積、管理する“仕掛けづくり”はできつつある。

しかしながら現状は、ドキュメント作成完了のときには紙媒体に出力して一件落着ということが多い。その結果、整理しきれない紙媒体と活用されない磁気媒体が氾濫することになる。紙媒体は数枚の資料から適切な情報のみをピックアップして、それに基づいて判断することにはふさわしいが、大量の情報をナビゲート、検索して、判断しなければならないときには不便である。

ドキュメント・データの標準化、ドキュメントの電子化、さらにはこのドキュメントから再利用するデータベース情報(コンテンツ)を効果的に構築できるようにすることが、ペーパーレス化へのキーポイントである」

ドキュメント量が大量で、改訂、検索が頻繁であり、記述内容が長期にわたり有効であるものとして、コンピュータのソフトウェアのマニュアルの開発保守があり、この分野はドキュメント・データの電子化への橋頭堡として有力である。

アメリカ国防総省のCALCでは、この考えをIETM(対話型電子技術マニュアル)として推進しているようである。

そのときの基本モデルとして、多くのページからなる技術文書を電子文書化してシステム全体の効率アップを図ることを取り上げたことは、IBMのコンピュータ・マニュアルの実例が大きな影響を与えているとの認識から、その電子マニュアル開発の最先端をヒヤリングして要旨を報告する。

(2)SGML化(ドキュメント・データの標準化)へのIBMの取り組み

IBMには生きている(改訂版発行の可能性のある)ドキュメントとして、膨大なコンピュータ・マニュアルがあり、それが不断に改訂、出版されているので、その製品のライフサイクルに応

じて管理されねばならない。

さらに世界規模で事業を展開しているので、同期をとった各国語への翻訳が伴うことを考えねばならない。

この枠組みにそった仕組みが、インフォメーション開発システムとして稼働、運用中である。このシステムをマニュアル開発元から見ると、最も重要な電子文書化ニーズは、既作成のマニュアル情報の再利用にあるとのことである。

CALSでの文書標準形式であるSGMLは、Goldfarb 博士が開発したIBMの Generalized Markup Language (GML)をベースとしたものであり、GMLと一連のインフォメーション開発システムについては、IBMでは 1970 年代からマニュアル用として長い使用実績がある。

その結果、現在、マニュアル・テキストとして約 900 万ページのデータベースが蓄積され、BookMaster/GML形式(検索などに便利のように圧縮された独自のメインフレームのデータベース)として、メンテナンスされているシステムになっている。900 万ページというのは、ざっと 1000 ページの 書籍 9000 冊分ということになる。

この情報は基本的に、コンピュータのマニュアルという集約されたジャンルに構築された章節などによる文書構造に着目した論理構造物であり、技術用語 index も整備され、本文自身の記述内容も平易で、記述文体もわかりやすく、1つの体系を構成しているものである。

CALSでSGMLが標準として採用され、その成りゆきが注目を惹いているのは、1つには、ドキュメント・データの標準化が、データの蓄積を地球規模で 行い、その集められたシステムティックなデータをコンピュータによる検索、再利用に結びつけ、ネットワークの大容量化、低コストの下で、文書というものが 紙媒体から完全に解放され、画期的な進歩をもたらしてくれるのではないかという期待からであり、「ワープロとかDTP<デスクトップ出版>での電子文書 化の目的が、レイアウト情報を付加して人間への見やすさ中心であるのとは、大きな格差がある」と強調している。

BookMaster/GMLを主体としたマニュアル作成 (Information Development) はホスト中心であり、マニュアル作成時には、WYSWYGよりも、専門家向けの簡便さを求め、VMという OSの下で動く XEDITエディタを利用しているとのことであった。また一部はワークステーションのAIXシステム上のベンダーのDTP製品 (Interleaf) を利 用して、インフォメーション・ソースを作成しているようだ。

SGML化のポイントである文書の論理的な構造を決める文書型定義<DTD>は、IBM IDDOCとして、IBM社の全社外用技術ドキュメントの構成を決める単一の文書構造フォーマット体系になっている。このDTDは、近々、一般にもオープン化される予定と聞いた。

BookMaster はマニュアル作成用のタグ言語であり、もともとは英語のマニュアル用として製品化されている。ただし、日本語マニュアル作成用も BookMaster4.0 として製品化されているとのこと。

IBM大和事業所では、マニュアルの現地語化(翻訳)をしている。これはSGM Lが狙うドキュメント・データの再利用例になるといえる。

この場合、インフォメーション・ソースは、常に対象となるマニュアルの英語版からもってきて、その都度現地語化している。

すでに一度翻訳出版した現地語のマニュアルの旧版を、新版のベースにするわけではない。この辺り、議論の分かれるところか？

アジアでは日本語、韓国語、中国語(本土、台湾)で、タイ語、ベトナム語は検討中である。このすべての現地語化も日本IBMで行っており、BookMaster のシステムで対応している。

BookMaster によって作成された情報から、変換ツール、後処理レベルで自動的にハードコピー(版下ドラフトはAFPプリンタで高速印刷、印刷版下は Postscript 出力する)、ソフトコピー(オンラインマニュアル、CD-ROM)、ヘルプを作成できる。

SGML標準化の動きと合わせてか、現 BookMaster/GML形式の世界からISO標準のSGML文書に変更して、オープン化とか社内外の使いやすいツールを積極的に利用することへ変換中であるという。これによって、より一層効率化へのアプローチをしたいといっている。インフォメーション開発の戦略(現行システムのどの部分を改革するか)は、マニュアルの現地語化(英文テキストから日本語への翻訳)のコスト低減を果たすことであった。そのためには、英語でのインフォメーション・ソース(原文のこと)を、より翻訳しやすいものにすることがポイントである。

IBMのマニュアル作成でのテキストそのものは、専門ライターによる記述で作られる。その他、マルチプラットフォーム、マルチランゲージ化(ホストとWS/パソコン、自社ツールと、市販DTP、SGMLツールの併用を図ることでもある)が、目下の急務であるとのこと。現状はマニュアル開発システムがホスト主体で運用され社内専用生産治具になっているので、SGML化の機会にこれも実現したいとの意向であった。

インフォメーション・ソースの再利用可能性の向上をどうやって図るか。システム設計側からはドキュメント記述の相対化、抽象化が望まれる。例えば章番号を3とか、6章参照などと絶対的な記述にしない。

マニュアル・ライターからは、その逆にドキュメント記述の絶対表示とか具体的な記述が要求される。

最後に本論からはずれるが、IBMの各マニュアルがオンラインマニュアルという形でインターネット経由でアクセスできるようになるというデモも見せていただいた。

3.2 SGMLツールの実習

SGMLは、文書交換、文書情報の共有、文書の長期保存のための有力な手段であるが、まだ認知度が高いとはいえず、広く普及しているというわけでもない。そこで、われわれはSGMLを適用したドキュメント作成作業とはどんなものなのか、ワープロを使用する場合と何が違うのかなどについて感触を得るため、㈱学習研究社の協力のもとに、4回にわたり、以下のような勉強会・実習を行った。

1. 文書構造
2. SGML文書の作成法
3. MS-Wordを使用したSGML文書作成

4. SGMLエディタの実習

(1) 文書構造

SGMLを理解する上で基本となる文書構造について、以下の項目を学習した。

- SGMLの特徴、メリット、適用対象文書の特性、構造タグと体裁タグ、リストの種類、タイトルと見出し・キャプション・図表・クロスレファレンス意味と使い方、章番号・リストの項識別記号等の自動生成、ワープロ・DTP文書の構成素
- Interleafの印刷イメージとAscii形式のファイル、Wordのテンプレートとテンプレートの限界
- DTDの概要、テンプレートとSGMLの比較、図表のSGMLによる記述例、アトリビュート(属性)の記述例

(2) SGML文書の作成法

SGML文書の作成法について、以下の項目を学習した。

- リスト、表、図、リンクの張り方、外字の利用法
- ファイルに作成した図を表示するためのリンク生成法とスタイルシートの作成法

(3) MS-Wordを使用したSGML文書作成(図3-2-1参照)

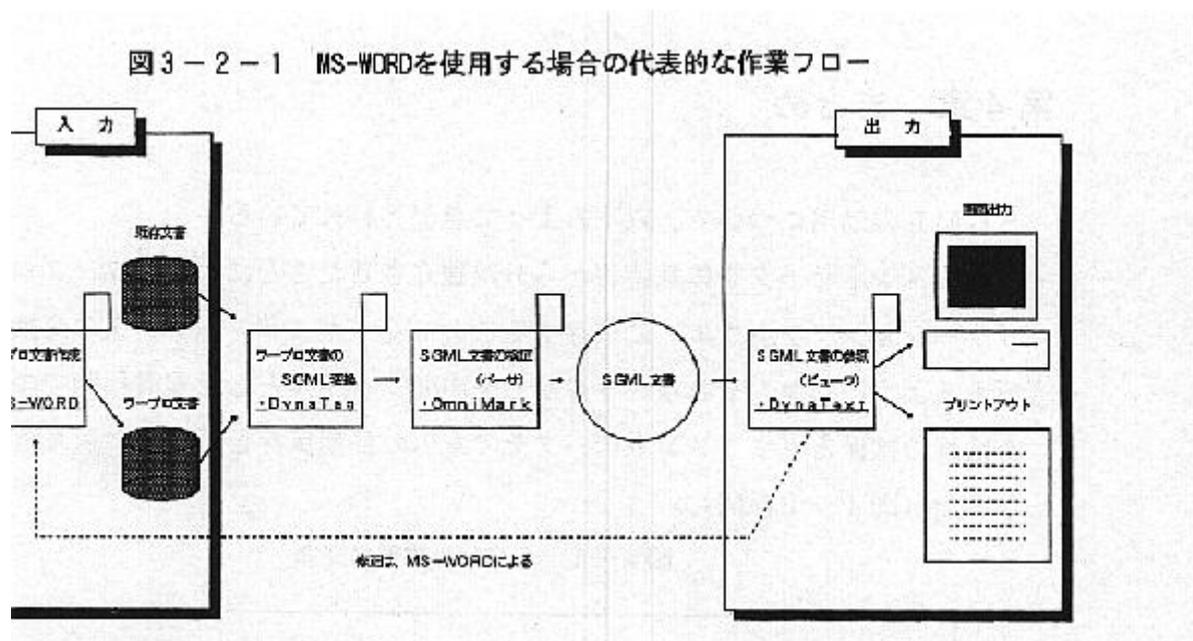


図3-2-1

SGML文書を作成する方法は大きく分けて2つある。1つは、ワープロの造化機能を利用して文書を作成後、SGML変換する方法。他の1つは、SGMLエディタを用いてSGML文

書を直接作成する方法である。

ここでは、MS-Word 6.0 のテンプレートによる構造化機能を使って既存文書(RTFファイル)をSGML変換するデモを見学し、質疑応答を行った。

- 通常の文書作成ではテンプレートを使用しないことが多い。SGML化するには、構造化情報が必要であるから、テンプレートを使用する習慣をつけること／強制することが必要である。
- 作成した文書ファイルをツール(DynaTag)により自動的にタグ付けを行い、タグ付けできなかった部分は編集機能によりタグ付けする。この手順でSGML文書が生成される。このとき構造チェック機能(パーサ)も必要である。またWP文書が持っているスタイル情報はスタイルシートに落とす。
- DynaTag は、Wordperfect、Interleaf、Framemaker 等にも対応している。
- マニュアルを数十万頁電子化した例もあるが、必ずしもSGML化していない。使用目的に対する効果と作成コストを考慮して、SGML化すべきかどうか判断する必要がある。

(4) SGMLエディタの実習

2つ目のSGML文書作成方法であるSGMLエディタを使った文書策背を実習した(使用ツール:InContext)。作業項目は以下のとおりである。

- タグ付け、リストの付加、リンクの付加、画像ファイルの付加
- DynaText データの作成(Indexer)、スタイルシートの作成(InStEd)、作成データの閲覧(Browser)

(5) 感想

SGML文書作成は、やはり慣れ親しんだワープロ操作よりは面倒である。

- タグの種類と使い方を覚えなければならず、タグ投入操作が追加される
- 変換手順、操作法を理解しなければならない
- 見た目が印刷形態と異なるので、慣れないと違和感がある
- 複数のツールを組み合わせる必要がある、適用方法、利用環境も、選択したツールによって異なる
- ツールの完成度はまだ高くはない、価格も高い

が、実感である。

しかしツールの価格や完成度、操作性等の問題は、技術の進歩により遠からず解決されるであろう。残る重要な課題は、SGMLの利用目的と得ようとする効果の具体化、適用業務

の選定、業務設計などであり、結局のところ、BPR(業務の抜本的な革新)の一環として、着実にSGMLを推進することが必要と思われる。

第4章 まとめ

SGMLの効用について、以下のような点が言われている。

- ドキュメントデータを体裁情報から分離独立させたことにより、ワープロ等のアプリケーションソフトウェアに依存することなくドキュメントデータの交換を円滑に行えるようにした。それゆえSGMLを中間ファイルとした異機種間でのデータの交換性の確保とドキュメントデータそのものの長期保存と再利用に耐えうるようになった(図4-1参照)。
- SGMLでは参照ファイルとしてテキスト以外のデータ、例えばラスターデータ、音声、静止画等を参照でき、まさにマルチメディアを扱うのに適した標準でもある。
- SGMLタグと文書構造の定義を用いてドキュメントをツリー構造で管理することにより、ドキュメントを構成するコンテンツをオブジェクト的に扱うことができ、その部分の再利用や共有化が図れる。
- SGMLはドキュメントデータの国際標準として航空業界、製薬業界、半導体業界等で採用されている。また、その派生的なものであるHTMLもインターネットやイントラネットでのデファクト標準として利用されている。

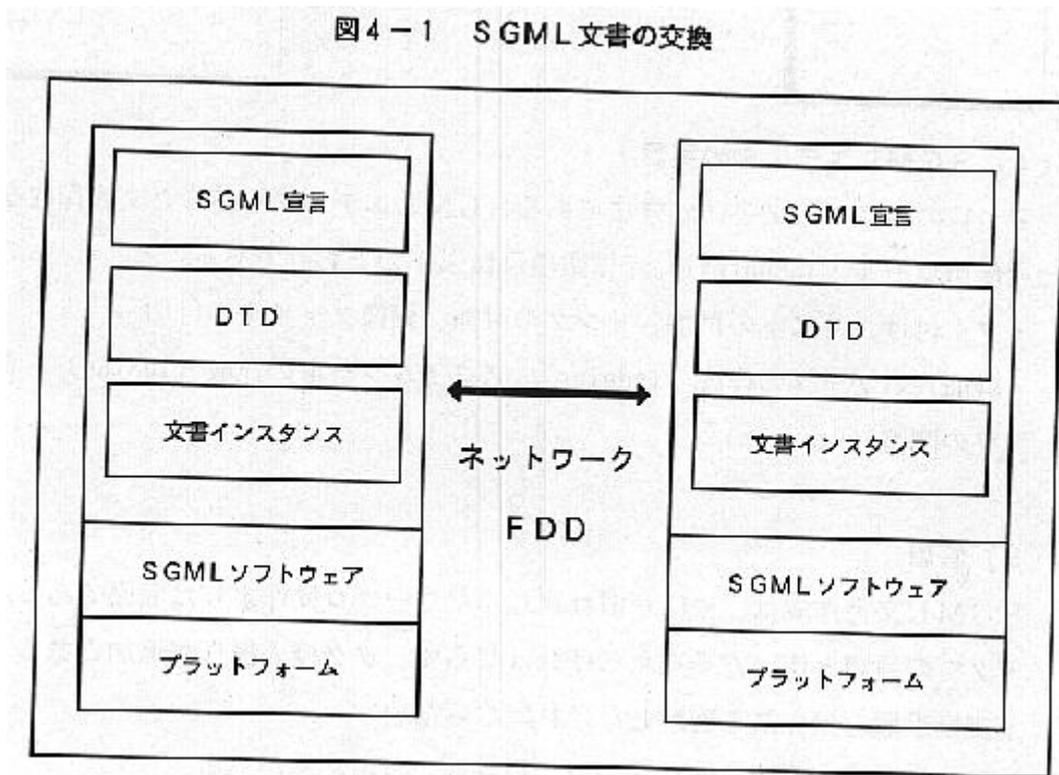


図4-1

これらは、いずれにしてもかなり総花的なものであり、WGで検討を行う中で「こういうところに使える」「こう使ってみては」とかいうSGMLの利用技術についての造詣を深めた。

SGMLの利用環境は各種そろってきている。これらのツールを用いることにより、例えばエディタからの直接入力したり、ワープロのテンプレート等から変換したりすることができる。しかしSGML化した後のプロセスや処理の標準が確定していないため、ツールに依存した状況になりかねない危惧がある。また、一般的に市場に出まわっている製品として完熟していないのが現状である。

製品モデルデータの交換に係る国際標準規約であるSTEPの中で各種製品モデルの定義が規定されている。この定義に従った製品モデルのデータを製品データ管理システム(PDM)として管理していくことが提唱されている。その際、製品モデルに関連する各種ドキュメントもその構造定義を、その製品モデルの定義とを関連づけることにより、データとドキュメントを対のものとして扱うことが可能となる。すなわち、製品データとそのデータの集合体としての情報が表現されるドキュメントのリンクとにより、データ・ドキュメントの交換と共有が円滑にはかれることになる。この点に関しては、先のJ2008DTDで言及したことでもある。

他方、インターネットの普及とともにフリーソフト的なブラウザーやツール群が普及し、またHTMLそのもののバージョンアップが行われると、HTML文化がデファクト標準化することも考えられる。むしろSGMLよりはHTMLが実用的、あるいは見掛けの利用上支障がないとしての普及が促進されるかもしれない。あるいは、SGMLの利用に関しては導入が始まったばかりで本来のSGMLの使われ方をされていないのかもしれない。

参考:略語一覧

本文中にも内容について記述を載せた略語で、ここでは説明を加えていないものもあるが、キーワードとして今後いろいろな場面で使用されると思われる主な略語について、改めて巻末にまとめて記載した。

参考にしていただければ幸いである。

- AS-IS モデル:現状モデル
- BPR : Business Process Reengineering
ビジネスプロセスリエンジニアリング
- CALS : Continuous Acquisition and Lifecycle Support
生産・調達・運用支援統合情報システム
この CALS の略語変遷は以下のとおりである。
 - Computer Aided Logistics Support
 - Computer-aided Acquisition and Logistic Support
 - Continuous Acquisition and Life-cycle Support
 - Commerce At Light Speed
- CIF : CALS Industry Forum, Japan CALS 推進協議会
- DoD : Department of Defense 米国国防総省
- DTD : Document Type Definition
文書型定義と呼ばれる SGML 形式の文書の構造定義部分
- EC : Electronic Commerce
電子通商
- EDI : Electronic Data Interchange
電子データ交換
- GML : Generalized Markup Language
1969 年 IBM が開発した構造化文書に係わる規定、後の SGML の母体
- HTML: Hyper Text Markup Language
インターネットの WWW のページを記述する言語
- HyTime : Hypermedia/Time-based Structuring Language
ハイパメディアおよび時間依存情報の構造化
- INSTAC :ソフトウェアCALSに関する調査研究委員会
- IETM : Interactive Electronic Technical Manual 対話型技術電子マニュアル
- J2008 : 米国 Society of Automotive Engineers が規定した DTD
- MIF : フレームテクノロジー社が策定した文書形式
- NCALS : Nippon CALS research partnership
生産・調達・運用支援統合情報システム技術研究組合

- ODA : Open Document Architecture
開放型文書体系に係る国際標準(ISO8613)
- OPENDOC : アップルコンピュータ社が提唱するドキュメントの統合作成環境
- PDM : Product Data Management 製品データ管理
- RFP : Request for Proposal 提案要求
- RTF : Rich Text Format マイクロソフト社が策定した文書交換ファイル形式
- SGML: Standard Generalized Markup Language
文書記述言語に係る国際標準規約(ISO 8879)
- STEP : Standard for Exchange of Product model data
製品モデルデータの交換に係わる国際標準規約(ISO 10303)
- TO-BE モデル:あるべき姿のモデル
- WWW : World Wide Web
欧州粒子物理学研究所が開発したインターネット上において情報を提供する一方法