

マーク付け言語関連規定のこれまでの議論と今後の展望

Overview and Perspectives for Markup Languages and Related Standards

小町祐史

Yushi KOMACHI

大阪工業大学

Osaka Institute of Technology

〒573-0196 大阪府枚方市北山 1-79-1

E-mail: komachi@y-adagio.com

ISOがSGMLを制定することにより、マーク付け言語が国際的な舞台に登場した。HTMLの構文としてSGMLが採用されてからSGMLはインターネット上での圧倒的な利用者を獲得して、それはXMLの開発に繋がった。XMLの公表と共にXML関連規定が充実し、マーク付け言語の第二世代が始まる。ここにDTDに代わるスキーマ言語の開発競争が激化し、この分野での日本のリーダーシップが表面化すると共に、DSDLの国際標準化が進んだ。これらの議論を紹介し、今後の展望に言及する。

国際規格, マーク付け言語, SGML, XML, HTML, DSDL, ODF, OOXML
International Standard, Markup Language, SGML, XML, HTML, DSDL, ODF, OOXML

1. まえがき

情報知識学フォーラムがSGML/XMLフォーラムとして開始されたとことに基づき、それに係わるテーマとして"SGMLから始まるマークアップ言語の歴史と今後"についての講演を依頼された。

筆者は"XMLの標準化"と題して2002年9月に第7回SGML/XML研修フォーラム[1][2]にて講演を行い、その時期までのXMLの標準化の経緯を紹介した。そこで本稿では、この"XMLの標準化"に対して、XMLの前身であるSGMLに関する記述を加え、さらに、SGMLを標準化したISO/IEC JTC1のSC18/WG8がその後SC34となって、その作業グループのWG1の主要検討課題として議論が続いている文書スキーマ定義言語(DSDL)について加筆する。

SC18/WG8, SC34で開発された幾つか文書関連規格、およびW3C(World Wide Web Consortium)のHTML/XML関連の主要規定は、ちょうど国内で標準情報TRの制度が始まった時期に開発され公表された。そこで、この制度(後に、標準報告書TR, 標準仕様書TS)を利用して、それらの翻訳が公表されて利用者のレビューを受けた。レビュー結果は直ちに国際にフィードバックされてCorrigendaやErrataとして発行され、次の改訂に反映されるという、国際と国内との同期的

協調関係が成立し、国際への寄与を継続した。そこで本稿では関連する国内でのSGML/XML標準化活動にも言及する。

XMLは著しい普及を示し、XMLを利用してデータ構造を記述する数多くの規定等が開発されている。それらに関する紹介は他稿に譲ることとし、ここでのSGML/XML応用は、JTC1/SC34の中で検討されたものにとどめる。

2. SGML

文書の文字列中へのマーク付けは、タイプセット(植字機)用の命令コードに始まる[3]。この機器依存の命令コードは、やがて文書を構成する論理的要素に対するマーク付けに発展し、ここに至って文書の論理構造を扱うシステムとフォーマタとが分離された。

マーク付けはさらに、ある文書クラスの論理的要素を共通に識別するようなタグ集合へと一般化され、共通マーク付け(generic markup)と呼ばれた。要素に関する属性記述をもタグに含めて、多様なアプリケーションに対応できるようにしたマーク付けも行われ、一般化マーク付け(generalized markup)となった。

ISO/IEC JTC1/SC18/WG8はこのようなタグ集合の定義方法を国際的に取り決め、言語としての体系付け

を行って、標準一般化マーク付け言語(Standard Generalized Markup Language, SGML)を規定する ISO 8879[4]とその Amendment 1[5]を開発し、それらはそれぞれ 1986 年と 1988 年に制定された。この国際規格 SGML によって、さまざまな文書やアプリケーションに対してマーク付けが定義可能となり、利便性の向上のために各種の補助機能も用意された。SGML を規定する JIS X 4151 は、ISO 8879 とその Amendment 1 とを含めた規定内容となっている。

その後、SGML 開発者達は Amendment 2 を目指してさまざまな拡張機能を議論したが、普及のためには規定内容を安定化する必要があるとの基本方針に阻まれて Amendment 2 が公表されることはなかった。

SGML の関連する規定としては次の SGML 支援ファシリティが開発され、ISO から出版された。

(a) 公開テキスト所有者識別子

ISO/IEC 9070: Registration Procedures for Public Text Owner Identifiers

(b) SGML 文書交換様式

ISO 9069: SGML Document Interchange Format

(c) SGML 構文主導形編集システムの指針

ISO/IEC TR 10037: Guidelines for SGML Syntax-Directed Editing Systems

(d) 数学及び科学用公開実体集合

ISO/IEC TR 9573-13: Public Entity Sets for Mathematics and Science

(e) 非ラテン系アルファベット用公開実体集合

ISO/IEC TR 9573-15: Public Entity Sets for Non-Latin Based Alphabets

SGML は米国国防総省の調達仕様に続いて、米国出版協会(AAP), EC 出版局(OPOCE)などに採用されたが、最も大きな普及は Web 文書の記述に用いる HTML への採用であった。HTML では構文として SGML が用いられ、比較的簡素な構造をもつ文書の記述のための文書型定義(DTD)が規定される。

3. HTML

WWW(World Wide Web)は、スイスの CERN(欧州素粒子物理研究所)において、所内の研究者間の研究成果の共有を支援することを目的として、1990 年に分散形広域ハイパテキストシステムの構築のためのプロジェクトによって考案された。このハイパテキストでは、文書にアンカ(端点)を定義して、アンカ間の関係としてハイパリンクを規定している。ハイパリンクを

辿ることによって、必要な文書の必要な箇所へのアクセスを物理的なページ概念を使わずに容易にした。

WWW のプロジェクトができた当初は、CERN において特定マシン上のラインモードブラウザが用意されただけであったが、その後 CERN 以外にも WWW の利用が可能になり、X ウィンドウシステム用のブラウザが開発された。イリノイ大学は MOSAIC を発表して文書中の画像表示を可能とし、やがてマルチメディアのサポートが可能になって、Web の普及に拍車をかけた。

Web の普及と共にハイパテキストの構造記述の標準化への問題意識が高まり、IETF(Internet Engineering Task Force)において、HTML(HyperText Markup Language, ハイパテキストマーク付け言語)と HTTP(Hypertext Transfer Protocol)の作業グループが設立されて本格的な標準化作業が開始され、その後 HTML の標準化作業は、W3C に移された。

XML の普及の後、構文として XML を用いて、充実した XML のサービス機能を HTML 環境へ導入することを容易にした XHTML が開発された。その結果、たとえば HTML への機能追加が必要になっても、その拡張 HTML を規定する DTD を開発することなく、拡張部分に対応するモジュールの追加によって対応できる。

ISO(国際標準化機構)は、HTML を国家規格(National Standard)として採用したいとする参加国(National Body)の要求、ISO 内部での文書交換様式として HTML を使いたいとする要求等に対処するため、W3C の HTML 勧告の追認作業を進めてきた。その結果、HTML4.0 のサブセットが ISO/IEC 15445 として承認され、それは国内では JIS X 4156 として制定されている。

HTML のこれらのバージョンを、表 1 に整理する。

表 1 HTML のバージョンアップ

HTML バージョン	勧告・規格番号	公表/改正
HTML 2.0	IETF RFC1866	1995-11
HTML 3.2	W3C REC-html32	1997-01-14
HTML 4.0	W3C REC-html40	1997-12-18/1998-04-24
	ISO/IEC 15445	2000-05
	JIS X 4156	2000-12/2005-03
HTML 4.01	W3C REC-html401	1999-12-24
XHTML 1.0	W3C REC-xhtml1	2000-01-26/2002-08-01
XHTML 1.1	W3C WD-xhtml11	2000-01-05

4. XML

HTMLの大量普及の結果、そのスコープを越えた利用が行われ、HTMLでは記述できない、または記述しにくい文書がクローブアップされることとなり、HTMLと同様の手軽さでSGMLと同様の文書記述を行いたいというユーザ要求が強まってきた。この要求に応えることを目的としてW3Cが開発した記述言語がXML(Extensible Markup Language, 拡張可能なマーク付け言語)である。

HTMLがSGMLで記述された文書型であるのに対して、XMLはSGML(ISO 8879/Cor.2まで含める)のサブセットに位置付けられる。つまりXMLは、HTMLでは扱えない文書をもサポートでき、これまでのHTML処理系と同様に、DTDが与えられなくても処理可能としている。W3Cは1996年末に最初のXMLドラフトを発表すると共にその更新を続け、1998年2月にその勧告を制定した。その後は、幾つものXML関連規定の開発を行っている。

W3CにおいてXMLは、次のドラフトを経てXML 1.0 (Second Edition)の公表に至った。

- <http://www.w3.org/TR/WD-xml-961114>
- <http://www.w3.org/TR/WD-xml-lang-970331>
(この版のハードコピー版は、XMLソースに対してDSSSLによってスタイル指定され、Jadeによってレンダリングされ印刷されていた。)
- <http://www.w3.org/TR/WD-xml-lang-970630>
- <http://www.w3.org/TR/WD-xml-970807>
- <http://www.w3.org/TR/WD-xml-971117>
- <http://www.w3.org/TR/PR-xml-971208>
- <http://www.w3.org/TR/1998/REC-xml-19980210>
- <http://www.w3.org/TR/2000/REC-xml-20001006>
(Second Edition)

96年11月版は、公表直後にBostonで開催されたGCAのSGML Conferenceで配布され、参加していた国内の関係者達(村田、小町、他)によって直ちにかつ独立に翻訳されて、それぞれのWeb等で国内に紹介された。

XMLに関する国内専門家の活動を効率化するため、日本規格協会の情報技術標準化研究センター(INSTAC)で既に活動を行っていた"マルチメディア/ハイパーメディア調査研究委員会"は、その中にWG4/XML-SWGを組織して、国内の専門家達に参加

をお願いし、XMLの標準情報(TR)化作業を開始した。

標準情報(TR)は、1996年8月に制定された標準情報(TR)制度実施要綱[6]に基づく制度であり、JIS化に至る前段階における技術標準等の状況を積極的に公表することによって、オープンな議論を推進し、関係者間の幅広い意見を集めるものである。

ISOでもIECでもないW3Cの勧告を国内規定として承認するためには、まさに時機を得た制度であった。XMLのTR X 0008:1998を皮切りとして[注1]、その後、6.に示すとおりW3CのXML関連規定が次々とTR化された。

注1: TR X 0008:1998の前には、DVD論理フォーマット、フォント情報処理用語、Java言語規定、規格文書用DTDなどが、TRとして公表されている。

4.1 XMLの標準情報(TR)

(1) 勧告案

WG4/XML-SWGは、XMLの各版に対して詳細レビューを行うと共に、97年3月版(WD)と97年12月の勧告案(PR)とに対して翻訳作業を行い、これらの活動の結果明らかになった問題点をその都度W3Cに対してフィードバックしてきた。

W3Cが1997年12月版を公表した後、XMLのTRに対する国内の業界要求が高まり、最終的な勧告(REC)を待ってからTR化するより、この段階でひとまずTR化の方が適切と判断して、97年12月版の翻訳をTR原案として1998年2月に当時の通産省工技院に提出した[7]。これは98年3月の通産省の審議委員会で承認され、TR X 0008:1998 拡張可能なマーク付け言語(XML)として、98年5月に公表された[注2]。

注2: この当時のTRの開発、承認、公表・発行の早さに注目されたい。

(2) 勧告

W3Cでは、97年12月版に対するコメントへの回答が98年2月10日に公開され、それを反映して変更を加えたW3Cの勧告"Extensible Markup Language (XML) 1.0"が98年2月に発行された。

INSTACは、98年4月に"高速Webにおける標準化に関する調査研究委員会"を設立し、その作業グループ(WG3)のXML特別作業グループ(XML-SWG)が、TR X 0008:1998のメンテナンスを担当して、W3C勧告のXML1.0の翻訳作業を開始した。その翻訳原案は、99年2月までのW3C正誤表の内容を反映して、TR X 0008:1998の改正案として同年2月末に工業技術院に

提出された[8]. これは, TR X 0008:1999 拡張可能なマーク付け言語(XML)1.0 として, 99年5月に公表された.

翻訳における訳語選定に際しては, SGML を規定している JIS X 4151 との整合を配慮した. しかしその後 JIS として出版された SGML 関連規格において, 適切な理由に基づいて JIS X 4151 の訳語を変更している用語については, なるべく新しい SGML 関連 JIS の訳語を採用している.

W3C の規定は, 必ずしも JIS 又は TR の様式には整合していないため, 多少の変更が必要になる. しかし TR の読者が原規定を参照する際の便を考慮すると, clause の構成はなるべく原規定のそれを保存することが望まれる. そこで, ごくわずかな修正(clause 番号の変更なし)だけを施して, TR 原案とした. この翻訳規定の構成は, JIS X 4151 の構成における反省に基づき, その後の文書関連 TR/JIS の構成に引き継がれている.

4.2 XML の JIS

INSTAC に 2001 年 4 月に設立された"次世代コンテンツの標準化に関する調査研究委員会"の作業グループ(WG2)は, TR X 0008 の公表等によって XML が国内に十分な利用者を獲得し, この規定内容に関するコンセンサスが得られたと判断して, その JIS 化作業に着手した.

2001 年 6 月には, W3C から XML1.0 の 2nd Edition が公表されたため, それを JIS 化対象の原規定とし, できるだけ新しい W3C 正誤表の内容を盛り込むことを目標とした. この JIS 原案(HTML 版)は, 2002 年 2 月に経済産業省に提出された[9]. 省内の都合によって, 同年 3 月の審議にはかけられず, 2002 年 6 月の審議によって承認された.

経済産業省は Web による JIS の電子公開をめざして, 2002 年 4 月から JIS 原案の MS Word フォーマットによる提出を要求した(この論旨の展開は, XML 関係者の注目を集めた.). その結果, この JIS 原案(HTML 版)は, 2 月に提出済みであったにもかかわらず, 5 月になって JIS テンプレートに従った Word フォーマットへの書き換えを求められることになり, JIS X 4159 拡張可能なマーク付け言語(XML)は 2002 年 10 月に制定・発行された.

その後, W3C が XML1.0 の 3rd Edition を公表したことに伴い, JIS X 4159 の改正作業が行われ, 改正原案は 2004 年 12 月に日本工業標準調査会標準部会の情

報技術専門委員会での審議を受けて承認された. これは 2005 年 3 月に発行されている. (この頃から, JIS の承認, 公表・発行の手続きに時間がかかるようになってきている.)

W3C は, W3C 文書の内容を, どんな目的のためにもどんな媒体でも, 報酬又は使用料なしに[注 3], 使用, 複写及び配布することを許可しているが, それは, 使用する文書又はその一部のすべての複写物が次を含む場合に制限されている.

(a) W3C 文書(原規定)へのリンク又は URI.

(b) 原著者の既存の著作権表示. それがない場合は, 次の形式の表示. "Copyright, World Wide Web Consortium, (Massachusetts Institute of Technology, Institut National de Recherche en Informatique et en Automatique, Keio University). All Rights Reserved. <http://www.w3.org/Consortium/Legal/>" (ハイパテキストが望ましいが, テキスト表現も可.)

(c) W3C 文書(原規定)の状態

TR X 0008 は, この要求に従った内容で発行されたが, XML1.0 の JIS 化に関して, W3C はさらに, 勧告のまえがきの記載内容を JIS のまえがきを含めることを要求し, 原案委員会の WG2 メンバと W3C 担当者との打合わせがもたれ, 具体的な様式が決められた. この様式は, その後の W3C 勧告に対応する JIS に継承されている.

注 3: 日本規格協会が発行する TR, JIS 等は有料で提供される. これは, W3C が要求する"どんな目的のためにもどんな媒体でも, 報酬又は使用料なしに"の要求を満たせないため, W3C 勧告に対応する TR, JIS 等の料金とは, それらに含まれる解説の料金であるとの解釈が採用されている.

5. XML 日本語プロフィール

5.1 標準情報(TR)

XML は, 符号化文字集合として JIS X 0221 及び Unicode 2.0 を採用しており, これは日本語文字をすべて含む. 文字符号化スキームとしては UTF-8 及び UTF-16 を推奨し, これらの実装を義務付けている. 既存の文字符号化スキームも, Unicode 2.0 の文字だけを扱う限りオプションとしてすべて許容している.

しかし, XML の勧告では, 日本語文字の交換に広

く使われてきた既存の文字符号化スキームはほとんど説明されてなく、オプションの一つとして許容されているに過ぎない。SMTP 及び HTTP などのプロトコル並びに情報交換用ファイルで、どの文字符号化スキームを用いるかについても、特に定められてはいない。

既存の文字符号化スキームと JIS X 0221 及び Unicode 2.0 との対応も不明確である。相互に異なるいくつかの変換表が用いられており、複数の XML プロセッサが異なる結果を出力する場合がある。

これらの問題点を明確にするため、TR X 0008:1998 はその解説の中に、"3. 日本語プロファイル"を設けて、全角英数字及び半角片仮名、情報交換用ファイル中の XML 文書、HTTP による XML 文書の配送、メールによる XML 文書の配送などに関する記述を含めた。

TR X 0008:1998 を改正して TR X 0008:1999 の原案を作成する際、原案委員会である"高速 Web における標準化に関する調査研究委員会"の作業グループ (WG3/XML-SWG)はこの問題の重要性を再確認して、この日本語プロファイルを独立した標準情報(TR)とすることにした。TR 原案は99年2月に提出され[8]、TR X 0015 XML 日本語プロファイルとして1999年5月に公表された。

5.2 W3C Note

XML 文書の中で日本語を使う利用者は、必ずしも国内の利用者に限定されるわけではない。そこで原案委員会の WG3/XML-SWG は、TR X 0015 を英訳し、W3C に対して Note として提案を行った[10]。規格協会も当時の工業技術院も W3C のメンバではなかったため、W3C への提案は、XML SWG の主要メンバおよび議論に参加した W3C メンバが属している次の組織によって行われた。

Submitting organizations:

Xerox
Panasonic
Toshiba
GLOCOM
Academia Sinica
Alis Technologies
Sun Microsystems

Submission 要求は99年12月に送付され、AC Rep からの Confirm を受けて、この英訳は、XML Japanese Profile, W3C Note 22-12-1999 として公表された。

5.3 TR の改正

W3C に Note として提案するための議論の中で、TR X 0015:1999 の内容に対して部分的修正が施された。W3C Note と TR との一致を図るため、この修正を反映した TR X 0015 の改正原案が、INSTAC の"次世代コンテンツの標準化に関する調査研究委員会"によって作成され、02年2月に経済産業省に提出された[9]。これは、2002年6月に公表されている。

6. XML 関連規定の標準情報(TR)化

TR X 0008:1998 の公表の後、関連する多くの W3C 勧告が INSTAC における幾つかの委員会で翻訳され、TR 原案として通産省/経済産業省に提出されて、承認を受けた後、次に示す TR として公表された[11]。この活動は2002年度末までに accessibility と security を除く W3C の主要な勧告を TR として公表することを目指して行われた。その後、これらの TR の幾つかは、部分的修正を加えて JIS として制定・発行されている。

これらの原案作成に際しては、word-by-word の詳細レビューが行われ、そこで明らかになった原規定(勧告、勧告案など)の問題点は、W3C にフィードバックされている。

6.1 文書構造

(1) XML 関連

TR X 0023:1999, XML 名前空間 [Namespaces in XML]

TR X 0076:2003, XML リンク付け言語 [XLink 1.0, XML Linking Language (XLink) Version 1.0]

TR X 0054:2002, XML スキーマ 第0部 基本 [XML Schema Part 0: Primer]

TR X 0063:2002, XML スキーマ 第1部 構造 [XML Schema Part 1: Structures]

TR X 0064:2002, XML スキーマ 第2部 データ型 [XML Schema Part 2: Datatypes]

(2) HTML 関連

TR X 0033:2000, ハイパテキストマーク付け言語 (HTML) 4.0 [HyperText Markup Language (HTML) 4.0 Specification]

TR X 0033:2002, ハイパテキストマーク付け言語 (HTML) 4.0 [HyperText Markup Language (HTML) 4.0

(4.01) Specification]

TR X 0037:2001, 拡張可能なハイパertextマーク付け言語 XHTML 1.0 [XHTML 1.0: The Extensible HyperText Markup Language]

TR X 0051:2001, XHTML 基本 [XHTML Basic]

TR X 0056:2002, XHTML のモジュール化 [Modularization of XHTML]

TR X 0080:2003, XHTML 1.1 - モジュールに基づく XHTML [XHTML 1.1 - Module-based XHTML]

(3) DOM/RDF 関連

TR X 0022:1999, 資源記述の枠組み(RDF) モデル及び構文規定 [Resource Description Framework (RDF) Model and Syntax Specification]

TR X 0019:1999, 文書オブジェクトモデル(DOM)水準 1 規定 [Document Object Model (DOM) Level 1 Specification]

TR X 0065:2002, 文書オブジェクトモデル(DOM)水準 2 コア規定 [Document Object Model (DOM) Level 2 Core Specification]

TR X 0060:2003, 文書オブジェクトモデル(DOM)水準 2 イベント規定 [Document Object Model (DOM) Level 2 Events Specification]

TR X 0078:2003, 文書オブジェクトモデル(DOM)水準 2 ビュー規定 [Document Object Model (DOM) Level 2 Views Specification]

TR X 0082:2003, 文書オブジェクトモデル(DOM)水準 2 スタイル規定 [Document Object Model (DOM) Level 2 Style Specification]

TR X 0083:2003, 文書オブジェクトモデル(DOM)水準 2 たどり及び範囲の規定 [Document Object Model (DOM) Level 2 Traversal and Range Specification]

(4) XML 応用

TR X 0014:1999, 同期化マルチメディア統合言語 (SMIL) 1.0 [Synchronized Multimedia Integration Language, SMIL 1.0]

TR X 0093:2003, 同期化マルチメディア統合言語 (SMIL 2.0) [Synchronized Multimedia Integration Language (SMIL 2.0)]

TR X 0077:2003, X フォーム 1.0, [XForms 1.0]

TR X 0094:2003, SMIL アニメーション, [SMIL Animation]

TR X 0095:2003, 変倍ベクタ図形 (SVG) 1.0, [Scalable Vector Graphics (SVG) 1.0 Specification]

TR X 0047:2001, XML による画像参照交換方式 [Picture Reference Exchange by XML] [注 4]

注 4: XML による画像参照交換方式については, まず TR X 0047 が国内で公表され, その英訳が, Embedding Glyph Identifiers in XML Documents と題する W3C NOTE になった.

TR X 0096:2003, ウェブ文書記述関連用語 [Glossary of Terms Used in Web Documents Description] [注 5]

注 5: この規定は, 13 件の W3C 関連規定から用語を抽出してまとめた規定である.

6.2 文書スタイル指定

TR X 0011:1998, 段階スタイルシート 水準 1(CSS1) [Cascading Style Sheets, level 1 (CSS1)]

TR X 0032:2000, 段階スタイルシート 水準 2(CSS2) [Cascading Style Sheets, level 2 CSS2 Specification]

TR X 0048:2001, XSL 変換 (XSLT) 1.0 [XSL Transformations (XSLT) Version 1.0]

TR X 0088:2003, 拡張可能なスタイルシート言語 (XSL) 1.0, [Extensible Stylesheet Language (XSL) Version 1.0]

TR X 0089:2003, XML パス言語 (XPath) 1.0, [XML Path Language (XPath) Version 1.0]

TR X 0059:2002, XSLT ライブラリ [XSLT Library] [注 6]

注 6: この TR は, INSTAC の委員会で独自に開発された XSLT 関連技術の規定である.

7. XML の ISO/IEC への影響

7.1 SGML Cor.2

SGML を開発した ISO/IEC JTC1/SC18/WG8(その後, JTC1/SC34)は, XML が発表されると, XML を厳密に SGML のサブセットに位置付けるため, ISO 8879:1986 の Corrigendum 2 の編集作業に着手した. SGML のエディタからの email による呼びかけに応じて各国の専門家が議論に参加した.

この議論は, 97 年 5 月の JTC1/WG4 Barcelona 会議 (SC18/WG8 から SC34 への移行期間での審議は, JTC1/WG4 として行われた.)での審議を経て, SC18 N5763 として投票にかけられ, 99 年 11 月に ISO 8879/Cor.2 として発行された.

ISO 8879/Cor.2:1999 は翻訳されて, JIS X 4151:2000

SGML 追補 2 として制定されている[注 7].

注 7: ISO 8879/Cor.1 は拡張命名規則に関する修正を規定するものであり, 1996 年 12 月に発行された. その翻訳は JIS X 4151:1998 SGML 追補 1 として制定されている.

7.2 XML を参照する規格等

SGML/XML は, データの構造記述に用いられると共に, その構文を使って規格等の厳密記述, 交換フォーマット記述等に使われている. その場合, 引用規格として XML を参照する必要があるが, XML そのものが国際規格(つまり ISO, IEC の規格, ITU の勧告)ではないため, 引用規格で参照することが困難である場合が多い. この問題の形式的解決策として, ISO 8879:1986 と ISO 8879/Cor.2:1999(又は JIS X 4151:2000) とを引用して, XML を参照したことと等価とすることが行われている. その具体例として, 次の規格等がある.

Amd.3 to ISO/IEC 9541-1, Multilingual extensions to font resource architecture

IEC 62318, Multimedia home server systems - Home server conceptual model and Systems

8. DSDL

データ構造の枠組みを規定するために, SGML においては, 独自の構文を用いる DTD(文書型定義)が使われてきた. 構文に XML を用い, データ型の記述を容易にすると共に, モジュール化への配慮をも施して, データ構造の枠組みを規定する言語がスキーマ言語である.

W3C における XML Schema の議論の膠着状態を打開するため, 日本は国内で開発された類似の規定である RELAX をまず標準情報(TR)として公表[12]し, それを ISO/IEC JTC1 に Fast-track 手続きを用いて提案した. この提案に驚いた W3C は, RELAX の投票期限直前に突然, XML Schema の勧告を公表した.

文書記述言語における重要技術の国際的合意を日本からの提案に奪われることを嫌った欧米は, RELAX の公表を阻止しようとした[9]が, 規定内容として洗練されしかも正規の手続きに従って進められたこの日本提案を阻止できず, その提案内容は ISO/IEC TR 22250 として発行[13]された.

それでもなおスキーマ言語においてリーダーシップ

を取ろうとする欧米は, 別のスキーマ言語規格を作る課題提案(NP)を行って ISO/IEC TR 22250 を事実上使用できないものにしようと画策した. その NP 投票が承認されて DSDL(Document Schema Definition Languages, 文書スキーマ定義言語)プロジェクトができると, さらに同様の繰返しを回避するため, 欧米は, 類似規格の課題提案 NP を事前にチェックする手続きを JTC1 Directives に導入して日本を牽制した.

最初に UK から DSDL プロジェクトに提案された素案をレビューした結果, それが技術的に貧弱なものであったため, 日本はそれに全面的に反対して素案の取り下げを求めると共に, DSDL のマルチパート化を提案し, その主要パートを, RELAX をさらに改良した RELAX NG 関連規格にすることを目論んだ.

その結果, 表 2 のパート構成の中で特に重要なパートの Part 2, 4 が日本提案の国際規格となった. Part 2 の RELAX NG は, 今では ISO, IEC ではもちろんのこと, W3C の関連規格の構造記述にも広く利用されている. Part 2, 3, 4 は既に翻訳され JIS(それぞれ JIS X 4177-2, -3, -4)として制定されている.

表 2 DSDL の主要パート

パート	表題	発行(ステータス)
1	Overview	Martin Bryan (CD)
2	Regular-grammar-based validation -- RELAX NG	2003-12
2 Amd.1	RELAX NG Amendment 1: Compact syntax	2006-01
3	Rule-based validation -- Schematron	2006-06
4	Namespace-based validation dispatching language	2006-06
7	Character repertoire description language	(FCD)
8	Document semantics renaming language (DSRL)	(FDIS)
9	Namespace and datatype declaration in Document Type Definitions (DTDs)	(FDIS)

9. SGML/XML 応用

9.1 HyTime

SGML を用いてマルチメディア情報の構造記述を

行おうとする活動が Steven Newcomb, Charles Goldfarb らによって開始され、まず SMDL(標準音楽記述言語)のプロジェクトが SC18/WG8 に設立された。その議論の過程で、マルチメディアに共通する技術要素が抽出され、それを記述するために、体系形式(Architectural Form)という概念が導入されて、HyTime(ハイパメディア及び時間依存情報の構造化言語)が独立した規格 ISO/IEC 10744:1992 として発行された。

HyTime はマルチメディア/ハイパメディアを扱うために必要な多くの斬新な技術を含み、原案審議段階でその分野の多くの技術者から注目を集めた。新規分野を扱う国際規格にありがちなように、多くの要求とコメントとが集まった結果、エディタはそれらを規格の中に反映することを余儀なくされ、HyTime の規定内容は膨大なものになって、その実装が広く普及することはなかった。しかし HyTime で投入されたハイパリンクの扱いは、HTML のハイパリンクに引き継がれて World Wide Web の大規模な普及につながり、さらに XLink へと発展した。時間情報の扱いは SMIL 等の関連規格に引き継がれ、ロケーションモデルの扱いは XPath に発展した。

9.2 Topic Maps

トピックマップは HyTime のアプリケーションとして考案され、HyTime を対象とする初めての国際コンファレンス "First International Conference on the Application of HyTime (IHC'94)"において各国の注目を集めた。このコンファレンスは、今も継続している XML Conference と同様に、米国の GCA (Graphic Communications Association, 現在の IDEAlliance)が主催する国際コンファレンスであり、カナダのバンクーバーで 1994 年の 7 月に開催された。

この IHC '94 で発表された講演課題には、その後の関連議論が直接的または間接的に SC34 等の国際標準化団体に引き継がれてそこでの主要な標準化課題となり、国際規格等として発行されたものが多く含まれている(表 3 参照)。このコンファレンスは講演件数が 16 件という小規模なものではあったが、その後の国際標準化に寄与する重要な議論が行われた価値あるコンファレンスであった。

表 3 国際規格等に反映された IHC '94 での講演課題

(番号) / 講演課題[講演者] / その課題の議論が国際標準化団体に引き継がれ、反映された国際規格等

- (a) / Interactive electronic technical Manuals [B. K. Caporlette] / ISO/IEC 13240:2001 Interchange Standard for Multimedia Interactive Documents (ISMID)
- (b) / Conventions for the Application of HyTime [M. Biezunski] / ISO/IEC 13250:2000 SGML Applications – Topic Maps
- (c) / The electronic library project at EDF's DER [M. Biezunski] / ISO/IEC 13250:2000 SGML Applications – Topic Maps
- (d) / A representation method for multilingual documents [Y. Komachi] / ISO/IEC TR 19758:2003 DSSSL library for complex compositions
- (e) / Making HTML a HyTime Application [W. E. Kimber] / ISO/IEC 10744:1997 Hypermedia/Time-based Structuring Language (HyTime) 2nd edition
- (f) / Using HyTime for external references [H. A. Tucker] / W3C Rec. XML Linking Language (XLink)

M. Biezunski は、表 3 に示される彼の 2 件の講演(b), (c)の中でトピックマップの原型となる概念を公表した。なお当時は、トピックマップは、"Topic Maps"ではなく"Topic Navigation Maps"と呼ばれ、HyTime を構文として用いていた。

(1) 第 1 版

JTC1/SC18/WG8 では、Topic Navigation Map のプロジェクトを設立するための新作業課題提案(NP: New Work Item Proposal)が、委員会原案(CD: Committee Draft)と共に 1996 年 5 月の会議に提出され、審議の結果、それらを NP/CD 同時投票にかけることが決まった。最終 DIS(FDIS: Final Draft International Standard)テキストが作成される段階で、Topic Navigation Map は Topic Maps へと改められた。

投票の結果、FDIS テキストは承認され、1999 年 11 月の会議で、FDIS テキストに対するコメントへの対処を行って、それを反映した最終テキストが作成された。それは、2000 年 1 月に ISO/IEC 13250 として発行された。

(2) 技術訂正 1 と第 2 版

2000 年 12 月の会議では、第 1 版に関する誤り報告(Defect report)が提出されると共に、トピックマップの間合せ言語(TMQL: Topic Maps Query Language)と概念モデル(TMCM: Topic Maps Conceptual Model)の必要性が提案されて、それらの新プロジェクト提案(NP)が行われた。

国際規格第 1 版に対する Cor.1 の原案は、2001 年 9

月の投票で反対なしで承認された。日本、ノルウェー、英国、米国からのコメントに対する議論が行われ、2001年11月にはその結果を反映した Cor.1 が SC34 メンバに配布された。しかしその後、さらに Cor.1 の訂正内容に従って規格本体の修正が行われ、読み易さを考慮して、ISO はこれを ISO/IEC 13250 の第 2 版として 2003 年 5 月に発行した。

(3) 国際規格のマルチパート構成

2002 年 12 月にボルチモアで開催された SC34 会議では、トピックマップの再構成が検討され、次の機能を含めるように、国際規格(ISO/IEC 13250)のマルチパート化を図る NP が提案されて、2003 年 3 月を期限とする投票が行われた。

- データモデルの形式的指定
- 二つの標準的な交換構文の間の関係の記述
- 正準構文の利用による、トピックマップエンジンの適合性試験

投票の結果、この NP は承認され、その後もマルチパート構成の検討は続けられ、2005 年 11 月においては、次のパートが開発の対象になっている。

- パート 1(ISO/IEC 13250-1) 概要および基本概念
- パート 2(ISO/IEC 13250-2) データモデル
- パート 3(ISO/IEC 13250-3) XML 構文
- パート 4(ISO/IEC 13250-4) 正準化
- パート 5(ISO/IEC 13250-5) 参照モデル

さらに ISO/IEC 13250 に関連する規格として、次の課題も検討されている。

- Topic Maps の簡潔構文(CTM)
- Topic Maps の図形記法(GTM)
- 公開 subject 用メタデータ

(4) 国内での規格制定

国際及び国内の技術動向を踏まえ、通商産業省(当時)の工業技術院は、(財)日本規格協会 情報技術標準化研究センター(INSTAC)に対して 2000 年度の活動として、ISO/IEC 13250:2000 の JIS 化作業を委託した。INSTAC では"文書処理及びフォントの標準化調査研究委員会"(DDFD)がこの作業を担当し、国際規格を翻訳して 2001 年 8 月に JIS 原案を提出している。その後の審査を経て、この原案は、2002 年 8 月に JIS X 4157 SGML 応用トピックマップとして制定された。

JIS 本体においては、原規格の Introduction を 0. 導入とし、以降の clause 番号に関して ISO/IEC 13250 と JIS X 4157 との一致を保っている。原規格には、微妙

な英語表記が使われていて、翻訳によってはその意味を十分に伝えきれない可能性があった。そこで JIS として異例ではあるが、原規格の 5.を原文のまま附属書 1 として附属書 B の後に配置している。

ISO がトピックマップ国際規格の第 2 版を発行したことを受けて、国内ではその JIS への反映の議論が開始された。当初 SC34 のプロジェクトは Cor.(技術訂正)として活動していたこと、主要な修正が附属書 C の追加だけであることを考慮して、国内では修正内容だけを記述した追補による出版を行うことにした。そこで Cor.1 が承認された段階で、DDFD は、それを JIS 追補とするための翻訳作業に着手した。Cor.1 に技術的に一致する JIS X 4157 追補 1 原案は、2003 年 1 月に経済産業省に提出され、その後の審査を経て、2003 年 11 月に JIS X 4157 追補 1 として制定された。

9.3 ODF/OOXML

(1) ODF(オフィス応用のための開放形文書フォーマット)

マイクロソフトの Word によるワードプロセッサの独占状態を打開するため、IBM, Sun などは XML に基づく文書処理系の開発を進め、その交換様式を ODF(Open Document Format for Office Applications)として OASIS(Organization for the Advancement of Structured Information Standards)に提案した。OASIS での承認(OASIS 規格 ODF 1.0)の後、OASIS は ISO/IEC JTC1 が規定する PAS(Publicly available specification)の Fast-track 手続きを使って ISO/IEC JTC1 での ODF の追認を求めた。

2006 年 5 月を期限とする ISO/IEC DIS 26300(ODF)についての投票の結果、反対投票なしで DIS が承認された。これはマイクロソフト固有の閉じた仕様からオープンな規定への利用者要求の顕在化と見ることができる。その後 JTC1/SC34 は、JTC1 セクレタリアートの助言に従って投票結果対処会議の予定を取消し、OASIS の ODF TC がコメント対処を行って、改訂テキストとコメント対処とを SC34 セクレタリアートが各国に送付して、30 日デフォルト投票が開始された。デフォルト投票での承認を受けて、ISO/IEC 26300 が 2006 年 12 月に出版された。

その後、日本でのこの規格の JIS 化作業に際して幾つもの技術的問題点が明らかになり、日本からのその指摘に基づいて ISO/IEC 26300 のプロジェクトエディタである P. Durusau が技術訂正案を用意して、SC34/WG1 からリエジンステートメントを OASIS に

送付した。しかし OASIS での対応がないため、SC34 の中に ODF メンテナンスのための作業グループを作ろうとする動きがある。

(2) OOXML(オフィス開放形 XML ファイルフォーマット)

ODF の ISO での追認に対抗するため、マイクロソフトは Word 処理系に基づく XML ベースの交換フォーマット OOXML(Office Open XML)を ECMA に提出して ECMA 規格とした後、ECMA から Fast-track 手続きを用いて ISO/IEC JTC1 に提出した。

この規格原案は ISO/IEC DIS 29500 として配布され、2007 年 9 月を期限とする DIS 投票が行われた。日本は次の技術的理由で条件付き反対を表明した。

- (a) ほとんどの XML 検証器で、規格の一部である W3C XML Schema が動作しない。
- (b) IETF の承認なしに pack URI スキームを利用している。
- (c) パッケージ中のパート名として ASCII 以外の文字が使えない。
- (d) 利用できるスキーマ言語が W3C XML Schema に限られている。

DIS 投票の結果、P メンバの賛成投票が 53.2%であったため原案承認に至らず、膨大な投票コメントに対する対処を検討するため(各国のコメントに基づいて ISO/IEC DIS 29500 を改善するため)の投票結果対処会議(BRM)が 2008 年 2 月にジュネーブで開催された。

この BRM において、日本から提出した主要コメントの中で、OOXML と ODF との協調要請についてはスコープ外として審議されなかったが、他の主要コメントについては満足すべき対処が行われた。他国からの主要コメントについても積極的な対処が検討され、アクセシビリティの改善、過去との互換性のための機構を切り出すフレームワーク、マルチパート化などに大きな改善がなされた。

BRM の対処結果を考慮して、2008 年 3 月を期限とする再投票が行われ、P メンバの賛成投票 75%を得て、ISO/IEC DIS 29500 は承認された。この再投票に際して、日本は賛成を投じている。しかし反対投票を投じた国がこの結果に不満を感じ、特に 4 ヶ国は承認手続きに対して不服申立て(appeal)を行っている。このような議論になると、もはや加速化手続きを使った意味はなくなる。ISO、IEC の通常手続きを用いて国際規格を開発するよりも短期間に国際規格を開発できるというデファクトスタンダード標準化組織の活動に対

する否定的な意見も出始めている。

10. むすび

SGML から始まったマーク付け言語は、HTML、XML の普及によって今や多くの人たちがそれを利用することを意識することなく利用している状態となった。それらの規格開発に関与した専門家達は ISO/IEC JTC1 が提供するデジュールスタンダードと W3C が提供するデファクトスタンダードの開発手続きを巧妙に使い分けて、利用者要求に応えると共に、規定の実装と普及をさまざまな戦略を展開した。

XML に基づく ODF と OOXML はその投票段階から多くの人たちの注目を集め、2007 年度末において、SC34 には 39 ヶ国の P メンバと 16 ヶ国の O メンバが参加するに至っている。これらの規格はいずれも JTC1 の加速化手続きを用いて提案されたが、その規定内容が膨大であったため、ODF はメンテナンスの手續きの問題を表面化させ、OOXML は加速化手續きにおける投票対処の問題を表面化させた。これらは JTC1 の問題として今後さらに検討されることになる。数千ページに及ぶ OOXML の規定内容は、規格出版・配布の形態にも見直しが求められよう。

文献

- [1] 小町祐史: XML の標準化, 第 7 回 SGML/XML 研修フォーラム, 2002-10.
- [2] 小町祐史: XML の標準化, 情報知識学会誌, Vol.12, No.3, pp.33-44, 2002.
- [3] Tucker, H.A. and Bogh, T.: SGML and ODA, Dansk Standiseringsrad, 1989.
- [4] ISO 8879, Standard Generalised Markup Language (SGML), 1986/10.
- [5] ISO 8879/Amd.1, Standard Generalised Markup Language (SGML) AMENDMENT 1, 1988/07.
- [6] 通商産業省: 標準情報(TR)制度実施要綱, 1996-08-01.
- [7] 1997 年度 マルチメディア/ハイパーメディア調査研究委員会報告書, 日本規格協会 INSTAC, 1998-03.
- [8] 1998 年度 高速 Web 環境における標準化に関する調査研究委員会報告書, 日本規格協会 INSTAC, 1999-03.
- [9] 2001 年度 次世代コンテンツの標準化に関する

調査研究委員会報告書, 日本規格協会 INSTAC, 2002-03.

[10] 1999年度 高速 Web 環境における標準化に関する調査研究委員会報告書, 日本規格協会 INSTAC, 2000-03.

[11] 小町祐史: インタネットメディア, 画像電子学会誌, Vol.33, No.6, 2004-11.

[12] TR X 0029:2000, XML 正規言語記述 RELAX コア, 2000-05.

[13] ISO/IEC TR 22250-1:2002, Information technology -- Document description and processing languages -- Regular Language Description for XML (RELAX) -- Part 1: RELAX Core, 2002-03.