

情報知識学会誌

Vol.12 No.3

第7回 SGML/XML研修フォーラム特別号

目次

XML 入門	原田 隆史	1
Web サービス	中村 祐一/浦本 直彦	15
セマンティックウェブの技術的位置づけと課題		
- 周辺技術から見たセマンティックウェブ -	飯島 正	21
eCampus - SGML/XML の教育システムの応用 -	海田 茂	33
XML の標準化	小町 祐史	37
トピック・マップ	内藤 求/Lars Marius Garshol	45
XML 署名	杉山 高弘	53
地質情報と XML		
- 国土交通省の電子納品における地質情報に係わる XML -	中田 文雄	61
XML 電子カタログと電子商取引の考察	藤岡 慎弥	69
お知らせ		
・新編集規程	安永 尚志	77
・論文募集	国沢 隆	89
資料		
情報知識関連新刊図書一覧	平田 周	90
会告		
平成 15~16 年度役員選出について		91

は し が き

情報知識学会主催の「SGML/XML 研修フォーラム」は、会員、参加者各位の支援をえて、今回第7回を開催する運びとなった。政府は長期化、深刻化する不況脱出のための有力な方策として、わが国社会のあらゆる側面における IT の普及、浸透を図っている。すなわち「e ジャパン」関連施策であり、その中では、ユービキタス・ネットワークという標語のもとに、情報通信網の格段の高速化、高機能化を促進している。そして、この通信網を介してコンテンツを大量、円滑に流通させるための方式・技術の主体をなすものとして XML がある。こうした状況下、各種の応用分野に即して、様々の XML 関連規格が提案、実装されつつあるのは周知のとおりである。

第7回「SGML/XML 研修フォーラム」は、基調テーマとして「XML と Web サービス入門」を掲げて、改めて XML の基本を押さえた上で、今後急速に普及、応用が進展すると期待される「Web サービス」に注目したプログラムを編成した。これは、ネットワーク上に分散する各種の関連情報を目的に応じて動的に統合し、利用者の情報要求に適確に対応することにより、社会的規模における情報・知識の交換・共有を実現するシステム統合技術であって、次世代インターネットの中核技術のひとつであるといえる。

本フォーラムでは、まず、XML、そして Web サービスについて、研究者・推進者により、基礎的事項および最新動向の整理・鳥瞰を与えてもらう。そしてこの上に立って、各種の技術分野、応用分野に関して、その現況と今後の動向の解説をいただくこととした。すなわち、XML、セマンティック Web、トピック・マップ等の標準化や実装技術に関する解説であり、また、教育システム、電子商取引、電子納品、XML 署名等への応用事例の紹介と将来への展望である。これら講演を踏まえて、参加者・講演者間での活発な意見交換が行われ、わが国における「高度情報通信ネットワーク社会」の形成が一層促進されることを、本フォーラム企画者一同として、大いに期待する次第である。

2002年10月

SGML/XML 研修フォーラム実行委員長

根岸正光

情報知識学会主催・第7回 SGML/XML 研修フォーラム 「XML と Web サービス入門」

[プログラム]

<< 第1日 >> 2002年10月22日(火)

9:30~ 受付開始

9:55~10:00 開会挨拶

実行委員長：根岸正光／国立情報学研究所教授

【午前の部】

座長：根岸正光／国立情報学研究所教授

10:00~11:30 XML 入門

原田隆史／慶應義塾大学文学部助教授

11:30~12:30 { 昼食 60分 }

【午前の部】

座長：高橋靖明／凸版印刷(株)

12:30~14:00 Web サービス

中村祐一／日本 IBM 東京基礎研究所
浦本直彦／日本 IBM 東京基礎研究所

14:00~15:00 セマンティックウェブの技術的位置付けと課題
—周辺技術から見たセマンティックウェブ—

飯島正／慶應義塾大学工学部助教授

15:00~15:20 { 休憩 20分 }

15:20~16:20 e Campus —SGML/XML の教育システムへの応用—

海田 茂／ネクストソリューション(株)

<< 第2日 >> 2002年10月23日(水)

9:30~受付開始

【午前の部】

座長：細野公男／慶應義塾大学教授

10:00~11:00 XML の標準化

小町祐史／松下電送システム
／ISO/IEC JTC1 SC34 委員長

11:00~12:00 トピック・マップ

内藤求／シナジー・インキュベート取締役
Lars Marius Garshol／Ontopia AS

12:00~13:00 { 昼食 60分 }

【午後の部】

座長：菊田昌弘

／シナジー・インキュベート代表取締役

13:00～14:00 XML 署名

杉山高弘／NECソリューションズ

／ニューメディア開発協会XML署名委員長

14:00～15:00 地質情報とXML —国土交通省の電子納品における地質情報に係わるXML—

中田文雄／川崎地質(株)情報企画管理室

15:00～15:20 { 休憩 20分 }

15:20～16:20 XML電子カタログと電子商取引の考察

藤岡慎弥／沖電機工業／シニアコンサルタント

16:20～16:25 閉会挨拶

実行委員：菊田昌弘

／シナジー・インキュベート代表取締役

〔SGML/XML 研修フォーラム実行委員会〕

委員長	根岸 正光	国立情報学研究所 教授
副委員長	細野 公男	慶應義塾大学 教授
副委員長	石塚 英弘	筑波大学 教授
委員	岩渕 幸雄	日本創造学会 評議員
委員	菊田 昌弘	(株)シナジー・インキュベート 代表取締役
委員	高橋 靖明	凸版印刷(株)Eビジネス推進本部
委員	白鳥 裕	大日本印刷(株)C & I 事業部

XML 入門

慶應義塾大学文学部 原田隆史(*ushi@slis.keio.ac.jp*)

I. XML とは

XML とは eXtensible Markup Language の略で「拡張可能なマークアップ(印を付けた)言語」という意味を持つ。マークアップとは、本体となる文字列の中に、特定の機能を持った文字(タグ)を使って意味や機能を書き込むことを意味する。その際に重要なことは、本文の文字と機能を示す記号が、同じ種類の文字によって書かれているということである。これによって、データ本体だけではなく、その意味や機能も同時にひとつの文書中で表現することが可能となる。

マークアップ言語としては、Web ページの記述に使われる HTML が有名であるが、XML も HTML と同様にタグでマーク付けしながら記述することになる。ただし、XML が HTML と最も異なるのは、自由にタグを設定できるという点である。HTML は、Web ブラウザで表示するためのタグのみしか設定できなかったのに対し、XML では意味のある単位でタグを自由に好きなだけ定義することができる。

このテキストで記述された文書であるということは、XML に「わかりやすい」「取り扱いやすい」というメリットをもたらすこととなった。たとえば、XML は従来のデータ表現と比較して以下の特徴を持つ。

- 1) 人間が直接内容を理解しやすい
- 2) タグ付けされているためにデータ構造の変更に柔軟でアプリケーションの処理がしやすい。
- 3) データが特定の環境に依存しない
- 4) テキストエディタをはじめとした、テキ

ストファイルを扱う多くの既存ツールが利用可能

特に、データが特定の環境に依存しないということは、資産を有機的に組み合わせて利用できることを意味している。

従来の一般的なアプリケーションが使用するデータは、そのアプリケーションのみに対応するものであった。また、データ交換用に用いられる CSV(Comma Separated Value)はデータの構造を持たず、出現順序のみに依存する体系を持っているため、効果的な運用するのは困難な場合があった。しかし、データが XML 文書になっていれば、記述の整合性のチェックやデータ構造を変換したいときなどに XML に対応した汎用のツールが利用できる。このことは、データが単一のアプリケーションだけに限定されたものではなく、プラットフォームの枠を越えて、多くのソフトウェアで、また多くの業務で共通して利用可能になったことを意味する。

また、アプリケーション側にとっても大きなメリットがある。たとえば、各アプリケーションで取り扱う機能が増えてグラフィックスや数式、音声などといったデータを取り扱う場合、利用者の要求に応じた形でこれを処理するためには、多くの機能を開発し提供する必要がある。しかし、データが XML 文書形式であれば、アプリケーション開発者は、独自にグラフィックスや数式や音声を表現するためのデータ形式を考え出す必要はなく汎用のものを使用すればよいだけとなる。XML でそれらのデータを記述する方法が、SVG や

MathML や VoiceXML といった仕様として存在している。これらをデータ形式として流用すれば、簡単に短期間に、しかも完成度の高いものを実現することができる。

また、タグを自由に記述することができるということは、データの表現力を豊かにすることにもなった。たとえば、「赤いきつねと緑のたぬき」という表現がある場合、「赤い」「緑の」という修飾部と「きつね」「たぬき」という被修飾部の関係が重要となるが、従来の HTML などではこれを表現することが困難であった。そのため、「赤いたぬき」「緑のきつね」といった関係ないものも検索されることとなっていた。これが XML を使用して、修飾部および被修飾部それぞれに適切なタグを付与することで、確実に求める情報にアクセスすることができるようになる。XML 用の検索言語も、すでに XML の関連規格として XML でコンテンツを記述した Web サイトを対象に検索を実行する検索言語 XQL(XML Query Language)の検討が進んでいる。

このような考え方は昔から存在していた。しかし、データ量が増加することや、文字以外のデータは変換が必要になるなど、コンピュータ処理を行う上での効率が低下することから、今までは実現のための努力はほとんどはらわれてこなかったのである。

コンピュータの処理能力に余裕が生まれてきた現代においては、かつてのようにデータを1ビットでも圧縮して記録し、処理することよりも「扱いやすい」「わかりやすい」ことが重要視されるようになってきた。また、コンピュータが取り扱う範囲が増え、一部の天才的職人に依存してはニーズを満たせないほどコンピュータの利用は増加の一途をたどっている。XML は、こういう時代の背景に合致する技術として注目を集めている。

2. XML 制定の歴史

インターネットの普及に、HTML と Web

ブラウザの果たしてきた役割は非常に大きいものがあつた。HTML のプレゼンテーション能力は画期的なものであり、従来のテキストベースでの単色で地味なビジネス文書の世界に、カラフルでデザイン化された文書を表現する能力と、リンクを使用した自由な移動可能性を持ち込むこととなった。これによってインターネット上でのデータの利用が花開いたのであり、もしこの両者が登場していなければ、インターネットはここまで普及しなかったであろう。

しかし、Web が発達してくると、単なる綺麗な情報提供方法という役割だけでは不満が発生するようになった。特に、HTML がスタティックな情報を表現するための手法であり、機能上の限界があることが問題となってきた。

商取引をはじめとして、実際のビジネス社会において交換されるデータの種類は多岐にわたる。どのデータが何を指し示しているのかを明確に表現することは、ビジネスの第一歩でもある。しかし、HTML は Web ブラウザのための文書記述用マークアップ言語であり、World Wide Web コンソーシアム(W3C)の HTML WG(ワーキンググループ)がタグや属性の名前と意味を業界標準として開発している。したがって、利用者が自由にタグを拡張して使用することはできない。

つまり、HTML では、書式指定以外のタグを使いたいという利用者のニーズには応えることはできないのである。このことは、商用アプリケーションで HTML を使用する場合に決定的な欠点となる。

そこで、この HTML の限界を解消するため、1996年に Web で SGML を使う手法の検討が W3C でスタートした。そして、6月には W3C 内に XML WG の前身である SGML WG が正式にスタートし、11月には XML 仕様の第一版が提出された。この時提出された言語は、HTML の固定化したタグ仕様と比較し

て、エンドユーザでもタグの拡張が可能であることから「eXtensible Markup Language」(略称 XML)と名付けられた。そして、1998年2月10日に、Extensible Markup Language (XML) 1.0 は、W3C の正式勧告として発行され、その仕様書が W3C の Web サイトに公開された¹⁾。

この経緯でも明らかなように、XML は HTML の拡張ではない。データの交換という用途に用いられることもあるが、現在同様に使われている CSV の拡張でもない。XML が HTML の限界という認識のもとにスタートしたのは事実である。しかし、XML は、SGML の複雑なオプション群から、不必要なものを削除し、HTML の経験をふまえて追加すべきと考えられたものを追加するという過程を経て SGML から派生したものである。

XML が言語として HTML と異なる最大のポイントは、XML が SGML と同じく「メタ言語」であるという点にある。メタ言語とは、言語を分析・記述するために用いる言語(記号体系)である。つまり、XML とは特定の文書を記述する言語を記述するための言語なのである。

それに対し、HTML は SGML というメタ言語の文法を利用して開発された、Web ブラウザに文書を表示するためのアプリケーション言語である。言い換えれば、HTML は SGML の文法によって定義された一つの文書型(Document Type)であるということができる。

日本でも 2001 年に XML コンソーシアムが設立されるなど、その普及が進んできている²⁾。また、W3C の仕様も各地で日本語化され公開されている³⁾⁴⁾⁵⁾。

3. XML を構成する仕様

XML 規格の基本仕様は、あくまで、文書の構造を規定したものにすぎない。XML 文書は、人間も読むことができるファイルではあ

るが、それだけで何かができるというものではない。そこで、実際に使用するにあたっては様々な仕様が必要となる。

XML 文書の書き方を定めた XML 1.0 を中心に、さまざまな XML に関する仕様が存在し、その全体で XML の利用範囲を広げている。たとえば、XML 文書をほかのデータ形式にコンバートするスクリプト言語の XSLT、XML 文書間のハイパーリンクを記述するための XLink、HTML を XML 文法で記述した XHTML などがこれにあたる。

この XML 関連仕様は、大きく 3 つに分類することができる。1 つめは XML 文書の記述法そのものにかかわる仕様。2 つめは XML 文書进行操作するためのスクリプトや API などにかかわる仕様。そして、3 つめは XML によって定義された応用言語などである。

XML の機能を拡張するために、多くの関連技術が現在も開発され続けており、その量は膨大なものとなる。その代表的なものを以下にあげる。

3.1 XML 文書を記述するための仕様群

XML 文書を記述するための仕様には、XML 1.0、XML 名前空間などがある。XML の最も基礎となる仕様である。

1) XML 1.0 (Second Edition)

XML のタグの形式や DTD(Document Type Definition: 文書型定義)の形式など、必要最小限の文書に関わる仕様を規定するもの。XML 文書の論理構造、物理構造、記述法などを定めている。現在のところ正式バージョンは 1.0 しか存在しないが、2002 年 4 月に XML1.1 がワーキングドラフトとして公開されている。

2) XML Schema, DTD, RELAX

XML 文書の構造、内容、意味を定義するためのスキーマ言語。どんなタグが使い、タグの中にどんなデータが入るべきなのかといった XML 文書の形式を定義できる。

XML1.0では、その元となったSGMLと同様にXML1.0の仕様の中にDTDと呼ばれる基本的なスキーマ言語を含んでいる。しかし指定できるデータ型に制限があることや、文法がXMLと無関係である、名前空間にうまく対応できないなどの欠点が指摘され、もっと複雑で高度な文書形式を表現するためにW3Cによって策定されたXML Schemaや日本で開発されたRELAXといったスキーマ言語を使用する方法が注目を集めている。

3) XML 名前空間(Name Space)

XMLでは自由にタグの名前を決めることができる。そのため複数のXML文書を結合する場合やデータを共有する場合などにおいては、同じ名前のタグが重複してしまう可能性がある。特に、NAMEやTITLEなどといった名前のタグは同一文書の中でも重複してしまう可能性が高い。しかし、だからといって誰も使わないような特殊な名前のタグばかりを使うと、本来の要素の意味を連想できない、非常に読みづらい文書になってしまう。そこで、この問題を解決するために導入されたのが、オブジェクト指向型プログラミング言語などでも用いられる名前空間(Name Space)という考え方であり、XMLのタグを書き分けるための方法として規定されている。

4) XLink, XPointer, XPath

XLinkは、XML文書中に他のXML文書などへのハイパーリンクを挿入する時の仕様である。もともとはXLL(eXtensible Linking Language)として策定され、これからXpointerが分離され、さらにXPathが分離された。

3.2 XML文書进行操作するための仕様群

XML文書はHTMLとは異なり、それ自身では単に構造化されたデータを記述するだけであってレイアウト情報は含まれていない。そのため、人間が見やすい形で表示するためには、XMLをどう表示させるかというスタイ

ルを別途指定する必要がある。

また、アプリケーションからXML文書进行操作(内容の抽出、追加など)するための方法も標準的なAPIが定められている。これらのAPIをサポートしたソフトウェアを一般に「XMLパーサ」と呼んでいる。XMLパーサを利用してXML文書进行操作しているプログラムは移植性が高く、またプログラミングのノウハウも蓄積しやすい。

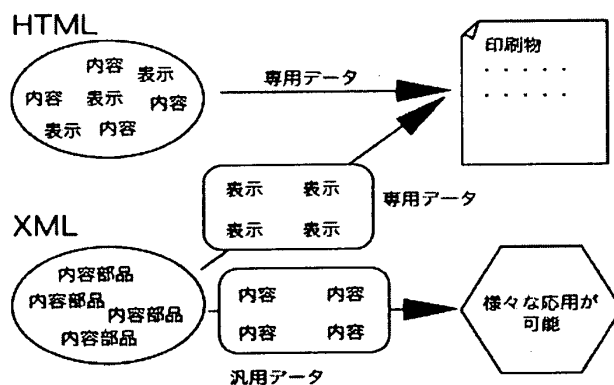


図1. データ内容と表示レイアウトの分離

1) スタイル言語(CSS, XSL)

XMLに対応したスタイル言語としては、CSS(Cascading Style Sheets)、XSL (eXtensible Style Language) などがある。CSSはHTMLやXMLに対応して全体のレイアウトや、それぞれのタグのテキストをどのように表示するかを指定するスタイルシートとして利用される。

一方、XML用のスタイルシートとして提唱されたものがXSLである。XSLはスタイルシートのための規格の総称で、XML文書に記述された中から必要とする部分を選択して取り出し、必要に応じて別の記述形式に変換するXSLT(XML Stylesheet Language Transformation)と、書式の情報を記述するXSLFO(XSL Formatting Object)で構成されている⁹⁾。

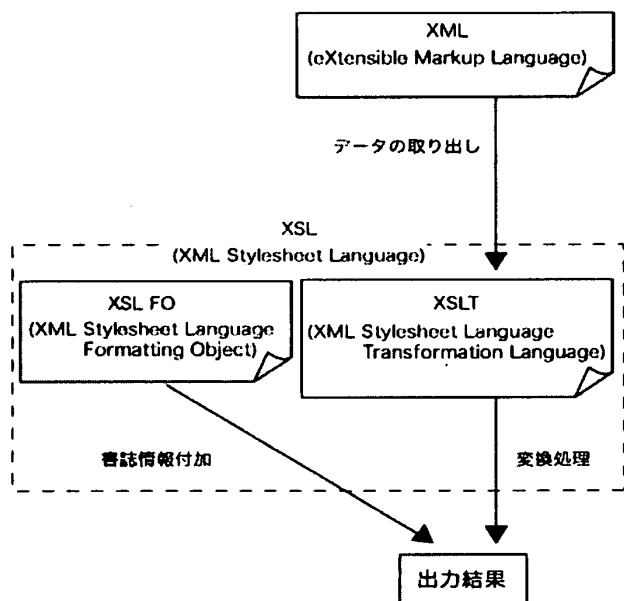


図2. XSL, XSL FO, XSLT の関係

このように XSLT と XSL FO が分離された背景には、XSL の仕様策定までに時間が必要とされたため、XSL の処理の中で元の XML 文書の構造を別の構造に変換する部分だけを抜き出して仕様を確定したという歴史的経緯がある。

その結果策定された XSLT は XML 文書データを別の構造の文書に変換する汎用的な仕組みと見なすこともできる存在となった。すなわち、XSLT プロセッサを使用すれば、各種の変換用 XSL ファイルを用意することで XML 文書を HTML 文書や PDF(Portable Document Format)に変換したり、テキストファイルを生成するといった変換が自在にできることになったのである。

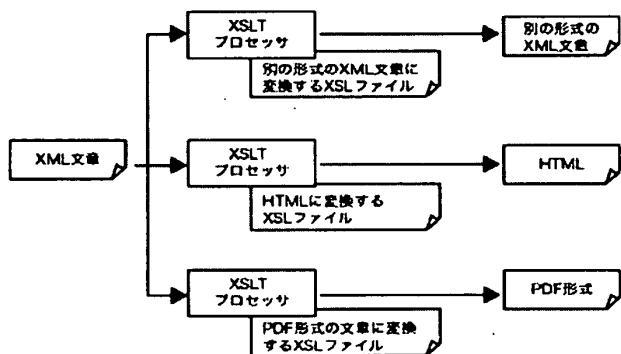


図3. XSLT プロセッサによるデータの変換

XSL FO の仕様が固まるまでに時間がかか

ったこともあり、現在は XML 文書のスタイルシートを作成する場合に書式設定の方法として XSL FO を仕様するのではなく HTML と CSS を使用する例が多くなっている。

2) XML 文書进行操作するための API

(Application Program Interface)

XML は、属性参照や名前空間、文字符号化方式などを処理するための多くの機能をもっている。そのため、これらの機能に対応できるプログラムを作成するためには大きな労力が必要となる。

そこで、XML 文書进行操作のために、これらの機能を実装したプログラムとして XML パーサ(XML Parser)が開発された。この XML パーサを使うことで XML の機能に関する箇所のプログラムを作成する負担を軽減することができる。

この XML パーサを使って XML 文書进行操作するための主な API として、DOM (Document Object Model) と SAX(The Simple API for XML)がある。

DOM は、XML 文書をツリー構造として解釈しアクセスするための API 仕様である。プログラムは、この DOM ツリーに対して DOM インタフェースを介して操作を行うこととなる。DOM は XML 文書をツリー構造に解析してメモリ上に展開するため、使用するメモリ量が多くなり大きな文書の操作には向いていないが、XML 文書内のデータの出現順序に関係なくアクセスすることができるなど高機能である。

一方、SAX は XML 文書を先頭から順に読み込んでいき、そこで発生したイベントを伝えるための API 仕様である。XML ドキュメントを読み込みながら、要素の開始や終了、キャラクターデータの出現というイベントをプログラムに通知するイベント駆動型の API で、XML 文書を順次処理していくため高速だが文書中をあちこち参照するような複雑な処理には向いていない。

3.3 XML を使って定義された言語群

XML 文法を使った多くのデータフォーマットが仕様として定められている。言い換えれば、特定の用途のデータを記述するのに必要な要素や属性の名前と構造を定めることによって様々な XML のサブセット仕様を作成しているという表現もできる。

たとえば、数式や画像などを組み込んだ文書を出版しようとした場合、これらに対応した XML 仕様が必要となる。このように様々な分野に対応した XML のサブセット仕様を用いて言語をすることで、XML という統一的な枠組みの中で全データを記述し、単一のプログラムで処理することが可能となる。以下に示すのは、その一部であり、見積書や発注書のようなビジネス情報から、書籍情報、気象情報などといった業界ごと、アプリケーションごとにさまざまな言語が考えられ、開発が進められている。

1) XHTML

XHTML は、eXtensible Hyper Text Markup Language の略である。XML と HTML を混ぜたようなような名前の通り HTML の代わりを XML のサブセットにさせようという仕様を意味している。すなわち、XML 文法に従って HTML を記述し直したものであり、Web ページを記述するという用途に特化した XML のサブセット仕様と言うことができる。

XHTML は HTML と同様に Web ブラウザで表示できると同時に XML 文書でもあるため、XSLT によって変換したり、XML パーサで整合性をチェックするといった XML 文書としての柔軟性も持ち合わせる。また、後述の SVG や MathML と組み合わせることで、HTML における IMG タグのように外部に構成要素が配置される形ではなく、処理を行う統一的な枠組みの中に位置づけることができる。逆にこのような枠組みを想定しなければ HTML をわざわざ XML で再定義した意味は

ないとも言えるだろう。

2) SVG (Scalable Vector Graphics)

XML の構文でベクトルグラフィックスを記述する言語である。線や円といったベクトル情報によってグラフィックスを記述する。静的な図だけではなく JavaScript で動きを加えることも可能である。ただし、現時点ではレイアウトが固定されてしまうため、あまり実用的ではないという指摘もある⁷⁾。

3) MathML

XML 文書によって数式を記述するための言語。従来から Web ページ上で学術論文を記述しようとした場合に、数字の位置関係やルート、べき乗といった数式によく見られる複雑な演算子などが表現しづらく、どうしても表現しようとした場合には画像ファイルとして貼り付けるしかないことが問題となっていた。しかも、画像データはサイズが大きくなるだけでなく修正も面倒であった。そこで、XML のサブセットとして作成されたのが MathML である。単体で使用されるほか、XHTML 文書中に埋め込んで使用されることもある。

4) XBRL (eXtensible Business Reporting Language)

XML 文書によって企業の財務・経営・投資など様々な用途に使用する情報を記述するための言語。たとえば、財務情報は各企業や業種ごとに、また年度ごとに文書の構造や項目などが異なることが多い。このため、企業ごとに別のシステムを開発することが多かったのであるが、これではデータの共通化や二次利用は困難となる。

XBRL は、財務情報の作成・流通・分析・変換などに対応した XML による標準規約であり、財務情報の提供者（公的機関、民間信用調査機関等）と利用者（与信管理システムなど）の共有情報基盤になると予想されている。

5) XML Signature

Web サービスを行う場合などには、受信者が送信者の身元とデータの信頼性を確認できることが非常に重要なこととなる。XML Signature は、そのために必要な XML 文書を電子署名するための標準的な手法である。

さらに、送信途中での信頼性の確保も重要であり、メッセージにスクランブルをかけて、受信者に届くまでの間に傍受されないようにするための暗号化技術も必要であり、XML Encryption が用意されている。また、XML アプリケーションが署名と暗号化プロセスに必要なキーを取得するための XML Key Management も開発中である。

4. XML 文書の構造

4.1 検証済み XML(valid XML)と整形 XML(Well-Formed XML)

XML 文書は、以下の3つの基本的な構成要素からなっている。

- 1) XML 宣言 : XML 文書であることを宣言
- 2) DTD(Document Type Definition) : XML インスタンスで使う要素やその構造を定義
- 3) XML インスタンス(XML データ) : XML 文書の本体であり、タグ付けされたデータの集まり

このうち、DTD については必ずしも必要なものではない。この3つの要素を全て持つものを「検証済み XML」と呼び、DTD が無いものを「整形 XML」と呼ぶ。

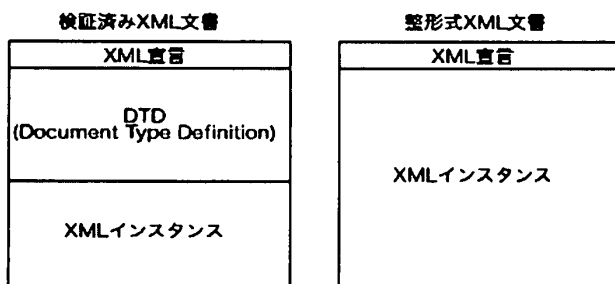


図 4. 検証済み XML と整形 XML

たとえば、XML 文書を HTML に変換して

Web 上で公開するような場合には、XML 文書の構造は XML 文書処理するプログラムだけが知っていればよい。このような場合には XML データの要素などが文法どおり表現されていれば、DTD がなくても処理に問題はない。

なお、HTML では、少々記述が間違ってもブラウザは無理やり表示していたが XML では以下の規則に基づいていない文書は有効ではない。

- 1) 開始タグと終了タグの対応が正しい
- 2) 木構造の根(ルート)に対応するタグがただ 1つだけ存在する
- 3) 親子・兄弟の関係が正しく入れ子の状態になっている

4.2 XML 宣言

XML 宣言は、XML 文書の先頭に記述され、その文書が文字通り XML であるということを宣言するものである。また同時に、XML のバージョンやエンコード方式なども属性で伝えることができる。

なお、XML 宣言自体は省略可能であるが、UNICODE(ISO/IEC 10646)の UTF-8 および UTF-16 以外のエンコード方式を用いる場合にはエンコード方式の指定が必要になるので省略はできない。

4.3 XML インスタンス(XML データ)

XML インスタンスは、XML 文書の本体のことであり、タグのついたデータの集まりである。XML インスタンスで示すデータは木構造をもち、その中に各要素を記述することになる。

4.4 DTD(文書型定義)

DTD は XML インスタンスで記述する各要素について、個々の要素の定義、要素間の関係(構造)の定義、外部にある DTD やデータの参照、画像データなど非 XML データの参照

などを行うものである。すなわち、DTD は文書の要素や属性などの相互的な関係を表すものであり、言うなれば、XML で使われる要素などのルールを制定するためのものである。

グループで統一された文書を書くときには、それぞれの記述者が適当に要素を付けることは大きな混乱のもととなる。そこで、基本的なルールを DTD で定め、DTD に従って XML 文書を書くことにより統一された一貫性のある文書の作成をはかるのが一般的である。逆に、その XML 文書を自分しか使わないのであれば DTD(や他のスキーマ言語)は不必要である。すなわち、DTD はルールをコミュニケーションするための言語だと言える。

5. XML 記述の実際

XML は簡単な技術であり、XML 文書を書く方法を修得するのは容易である。

最も簡単な形の XML 文書では XML 宣言も DTD も記述する必要はない。ルートをもつだけ持った XML インスタンスが存在すれば十分である。

したがって、きちんと記述された HTML もまた XML として成立することとなる。

DTD 付の XML は以下のように表現される。

```

1: <?xml version='1.0'?>
2: <!DOCTYPE 図書館目録[
3: <!ELEMENT 図書館目録 (図書 DATA)+>
4: <!ELEMENT 図書 DATA (書名,定価,著者)>
5: <!ELEMENT 書名 (#PCDATA)>
6: <!ELEMENT 定価 (#PCDATA)>
7: <!ELEMENT 著者 (#PCDATA)>
8: ]>
9: <図書館目録>
10: <図書 DATA>
11: <書名>XML 入門</書名>
12: <定価>1200</定価>
13: <著者>情報太郎</著者>

```

```

14: </図書 DATA>
15: <図書 DATA>
16: <書名>Web サービス入門</書名>
17: <定価>3500</定価>
18: <著者>知識次郎</著者>
19: </図書 DATA>
20: </図書館目録>

```

ここで、1 行目が XML 宣言、2~8 行目が DTD、そして 9~20 行目が XML インスタンスである。

一般に DTD は XML 文書の外側に置いて参照することが多いが、ここでは内部に埋め込む形で記述している。2 行目の<!DOCTYPE ... [が DTD の宣言の開始を、そして 8 行目の]> が DTD の終わりを示している。

3~7 行目は、要素の並び順を制約するために記述されている。これによって定義されていない要素名を書いたり、許されない順番で要素を記述したときに、それをエラーとして判定することが可能となる。

9 行目以降が文書の本体である。この XML インスタンスでは<図書館目録>要素がこの文書のルート要素となっている。このルート要素から枝分かれする形で各要素が記述される。

要素の中には、別の要素(子要素)やテキストを入れることができる。例では、「図書館目録」の中に「図書 DATA」要素が入り、さらに「図書 DATA」要素の中に「書名」要素と「定価」要素および「著者」要素が含まれている。

要素や属性の名前には、半角英数字、漢字、全角カタカナ、ひらがななども使用できる。ただし、名前の 1 文字目には半角数字を使用できない、1 文字目、2 文字目以降に関わらず、半角カタカナと全角英数字は使用できないなどといったいくつかの制限がある。

また、名前空間付きの整形 XML 文書の

場合は以下のようになる。

```

1: <?xml version='1.0'?>
2: <book xmlns:BK="http://www.x.com/y">
3:   <BK:name>図書館概論</BK:name>
4:   <BK:price>2699</BK:price>
5:   <personal xmlns:PL=
      "http://www.x.com/z">
6:     <PL:name>原田隆史</PL:name>
7:     <PL:height>172</PL:height>
8:   </personal>
9: </book>

```

ここで、「name」要素は図書の書名として用いられるほか、著者の氏名としても用いられている。したがって、名前空間を用いないで記述した場合には、「name」要素が重複してしまい、正しい解釈を行うことができない。

2行目および5行目が名前空間の宣言であり、上記の例では2行目で図書についての名前空間、5行目で著者についての名前空間を宣言している。このことによって、図書についてはBK:、著者についてはPL:が設定され、混乱を解消することができる。

6. XML技術の用途

XMLの仕様が策定された当時は、XMLの利用法として電子出版のような、ドキュメントの表現力向上を主な用途として想定していた。しかし実際には、XMLは出版関係だけでなく、企業内でのデータの統合的な利用、企業間での取引情報、アプリケーション間の通信など、さまざまな分野で使われるようになった。

6.1 Webパブリッシング

HTMLの表現力の向上にともなって、Web上で情報提供をすることが誰にでも可能となった。現在のインターネット利用の中でWorld Wide Webによる情報発信が占める比率は非常に大きい。

しかし、変化する情報に対応したページを維持するためには、CGIのような仕組みを設けて動的に更新するなどの工夫が必要である。また、携帯電話の出現によりiモードに対応したWebページの作成も求められるなど、Webページを作成・維持していくためには従来よりも多くの努力が必要とされるようになってきている。

これを解決する手法としてXMLを利用することができる。すなわち、コンテンツをXML文書にしておき、必要に応じてXSLTを使用することで、ひとつの文書からさまざまな形式への変換が用意に可能になる。たとえば、Webブラウザから参照されたときにはXML文書をHTML形式に変換する。iモードなどの携帯電話から参照された場合にはCHTMLなどに変換する。さらにはPDFに変換することで書籍や雑誌などの紙メディアとして出版することも可能となる。また、データの変更結果を直接的にWebページに反映させることもできる。

6.2 データの共用・多面的な利用

XMLは、そのデータ構造が規定され、多くの組織で共通に利用できる点に大きな特徴がある。たとえば、現在、気象情報は公的機関、公共団体、民間企業など多くの機関で収集されている。しかし、収集された情報はこれまで各機関内で利用されているだけということが多く、共用や交換されている例はごくわずかであった。

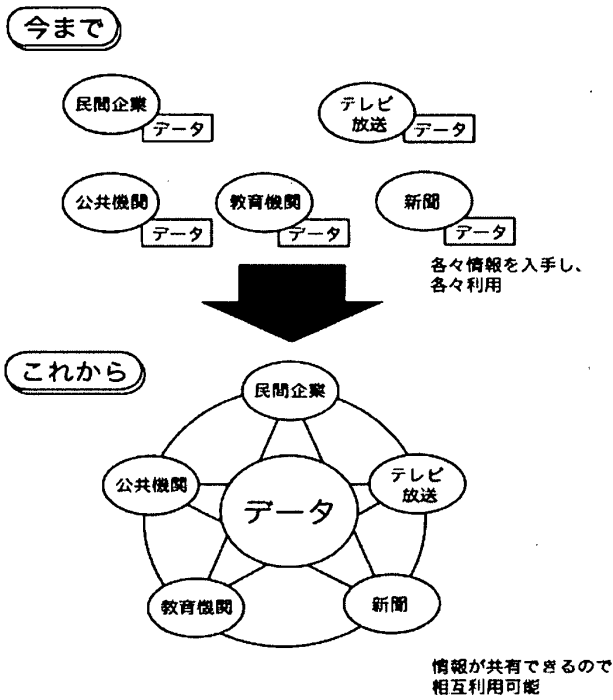


図 5. データの共有と相互利用

データの共有が進まない理由としては、各機関の事情もあるが、データの表現法が各機関で異なっていたことがあげられる。

しかし、気象情報をあらゆる統一フォーマットを作成し、それにのっとって XML で気象情報を記述すれば、必要なあらゆる機関で利用可能となる。このようなデータの共有を行うことが有効な分野は気象情報だけに限らず、主として公共の利益に関わる分野で無数に存在する。実際に、いくつかのグループや機関でその活動も開始されている。

6.3 BtoB

BtoB とは Business to Business の略で企業間のネットワーク(主としてインターネット)を介した電子商取引を意味する。XML の登場によって最も活性化すると予想されているのがこの BtoB であると言われている⁸⁾。

電子商取引においては、企業間で取り交わされる製品情報や発注情報といった複雑な情報を、柔軟に、かつ共通形式で取り扱う必要がある。企業間でやりとりされるデータの交換フォーマットとして XML が利用可能にな

り、情報の標準化が一気に促進されたのである。

共通の DTD に基づいた XML によって製品情報を交換することで、Web 上の情報が、各企業の枠を越えた巨大なデータベースとなり統合的な検索や利用が可能となる。また、XML の自由なタグは HTML よりもはるかに高度な検索を可能とし、統合された高度な応用システムの作成への道をひらくこととなる。

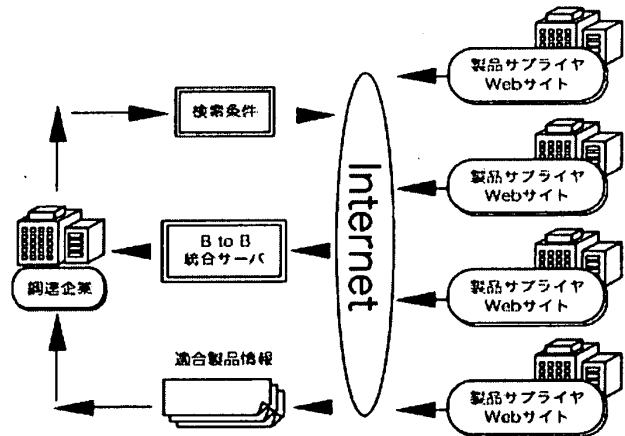


図 6. BtoB の概念図

このような XML を基礎とする BtoB は、すでに実験段階から実用段階へとつつろうとしている。BtoB サーバと呼ばれるカテゴリの製品も数多く発売されるほか、業界ごとに DTD を定めようとするグループも多い。たとえば、ebXML 策定グループは EDI(Electronic Data Interchange: 企業間の電子データ交換)を XML で代替するだけでなく、より進んだサービスにするための検討と提案を行っている。

また、CALS(Commerce At Light Speed)は XML を用いた BtoB と、その基本的な考えはほぼ同様であり、この基礎的な仕組みとして XML の適用も検討されるなど⁹⁾、このような活動は急速にその数とバリエーションを増している。

6.4 EAI

(Enterprise Application Integration :

エンタープライズ・アプリケーション統合)

企業内では、経理システム、人事システム、受発注システム、在庫管理システムなど種々のアプリケーションが稼働している。また、そのプラットフォームも UNIX や Windows で動いているものから独自 OS 上で稼働するものまで多岐にわたることがある。

変化の激しい社会の中で、企業が効果的な経営管理を行うためには、これらのデータを統合して、顧客管理、売り上げ管理、人事管理などに関わる意志決定が行えるようにする必要がある。また、企業の戦略を考える上でも種々のデータを用いたシミュレーションは欠かせない。EAI は、このようなアプリケーション間のデータを統合して利用するものである。

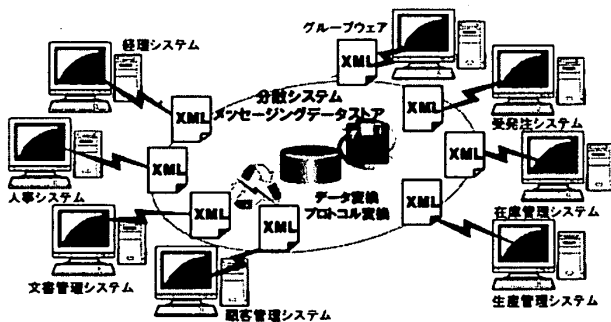


図 7. EAI サービスの概念図

(NEC ソフト. EAI サービス¹⁰⁾)

<http://www.necsoft.com/solution/xml/xml09.html> より)

このようなアプリケーションシステム間でのデータ連係を行うためには、システムで異なるデータフォーマットを統一して相互にデータの交換が行えるようにする必要がある。

XML は、こうしたアプリケーション間で異なるデータフォーマットの間言語としての役割を果たし得る存在である。たとえば、アプリケーションのデータを変換し蓄積する EAI サーバ上で、各アプリケーションが使用するデータはいったん XML 文書に変換して蓄積され、また別のアプリケーション用に変換されて提供するという形が考えられる。

また、将来的には各アプリケーションおよび各種のデータを管理する管理システムが直

接 XML を取り扱うということも考えられるだろう。すなわち、XML は EAI におけるデータ変換のコア技術であるということができる。

6.5 Web サービス

現在、XML 関連仕様の中で最も注目されているものは Web サービスであることは間違いのないだろう⁸⁾。

Web サービスは、WWW 上で行われる情報サービスのことを指すのではなく、SOAP(Simple Object Access Protocol : XML を利用した通信プロトコル。メッセージの XML 形式を定義した仕様¹¹⁾)、WSDL(Web Services Description Language : Web サービスのインタフェースを記述するための言語)、UDDI(Universal Description, Discovery and Integration : インターネット上の Web サービスを検索するための仕様)を使った分散処理システムを意味する。

従来の Web 上のサービスは HTML で記述され、人間に対してサービスを行うのが一般的であった、それに対して、Web サービスは XML を利用することにより、ほかの Web サービスに対して人間を介することなくやりとりをすることが可能となることが大きな特徴である¹²⁾。

このことは、単純に利用者が複数のサイトを個別にアクセスする必要がなくなったということだけではなく、様々な Web サービスを組み合わせることで、新たな付加価値サービスを提供することが可能になったことを意味する。すなわち、同種のサービスを横断的に検索するという枠にとどまらず、複数の異種サービスを統合的に利用することによって新たなサービスが生み出される可能性がある。たとえば、地図提供サービス、時刻表サービス、気象情報サービス、レストランガイドなどを組み合わせて、旅行プランサービスを提供するような組み合わせがこれにあたる。こ

のような新たな高付加価値サービスの可能性は大きく広がっており、ビジネスチャンスも広がると考えられる。

このように、XML と Web サービスはインターネット利用の中心となる可能性を秘めている。そのため、マイクロソフト、サン・マイクロシステムズをはじめとして、主要なソフトウェアベンダはこぞって Web サービス対応を打ち出している。

7. まとめ

XML の仕様は日々進化しているとはいえ、XML 関連仕様の基礎となっている XML 1.0、XML 名前空間といった部分はほとんど変化していない（もちろん、ここが不安定なようではしっかりした仕様とはいえないわけだが）。

一方で、XML 関連技術としての Web サービス、BtoB のための応用言語など、その応用分野は広がってきている。企業における Web サービスの利用も導入スピードは鈍化したものの、着実に増加し続けている¹³⁾。

eXcelon 社の Joe Bellini 氏によれば、Web サービスの導入を待つまでもなく、XML の導入だけでも、以下の3つのメリットがあり、十分に投資の回収が可能であるという¹³⁾。

- 1) シリアルではなく、パラレルでの処理が可能となり、ビジネスプロセスのサイクルタイムが短縮されること
- 2) 統合作業を行う対象が個々のアプリケーションではなく XML のバックボーンに集中するため、システム統合にともなうメンテナンスやサポートのコストが削減できること
- 3) XML によりミドルウェアを Linux に置き換えるなどしてハードウェアコストが削減できること

XML を学ぶことは、単に一つの言語を習得するというのではなく、データの取り扱いについての新たな考え方を学ぶことであるということも言える。従来のデータベース管理

システムを中心に据えたデータ管理からアプリケーション・サービス主体のサービスへの転換は、営利・非営利を問わずビジネス現場、研究教育現場をはじめとする社会全てに大きなインパクトを与える力を持っている。XML はそのための基礎技術として発展していくだろう。

参考文献・引用文献

- 1) W3C. Extensible Markup Language (XML). [2002/09/10] <http://www.w3.org/XML>
- 2) XML コンソーシアム. XML Consortium. [2002/09/18] <http://www.xmlconsortium.org/>
- 3) Infoteria Corporation. XML ノート:基本仕様・応用技術・関連サイト.[2002/09/17] <http://www.infoteria.com/jp/xmlnote/>
- 4) どちら猫本舗.XML 関連仕様書. [2002/09/10] <http://www.doranekeo.org/xml/>
- 5) SUGAI, Manabu. やむをえず XML. [2002/09/01] <http://members11.tsukaeru.net/msugai/xml/index.html>
- 6) 井上孝司. Web コンテンツ作成のための XSLT 入門. 329p. 2002.
- 7) atmarkIT.SVG(Scalable Vector Graphics) [2002/09/23] <http://www.atmarkit.co.jp/aig/01xml/svg.html>
- 8) 新野淳一.新入社員のための XML 入門講座. [2002/09/01] http://www.atmarkit.co.jp/fxml/tanpatsu/17freshman/xml_basic.html
- 9) 山本隆彦. PAGE2000 建設 CALS/EC と XML. [2002/09/23] <http://www.jagat.or.jp/page/2000/rejume/B1/yamamoto/>
- 10) NEC ソフト. EAI サービス. [2002/09/10]. <http://www.necsoft.com/solution/xml/xml09.html>
- 11) Roger Wolter. とにかくも SOAP. [2002/09/13]

- <http://www.microsoft.com/japan/msdn/columns/xml/xml10152001.asp>
- 12) Fujitsu. Web サービス. [2002/09/10].
<http://xml.fujitsu.com/jp/tech/web/index.html>.
- 13) 末岡洋子. XML による統合のメリットを示すエクセロン. [2002/09/15]
<http://www.atmarkit.co.jp/news/200207/31/exc.html>
- 25) 勝見浩之. XML って何? -XML を理解すれば次世代 Web が見えてくる-. [2002/09/13]
<http://www.hnes.co.jp/documents/xml/>
- 26) A.Okazaki. Studying XML beginners. [2002/09/13].
<http://www.asahi-net.or.jp/~ps8a-okzk/xml/>
- 27) 川俣晶. XML を学ぼう. [2002/09/19]
<http://www.atmarkit.co.jp/fxml/rensai/xmlwomanabou01/learning-xml.html>
- また全般にわたって以下の文献を参考にした。
- 14) XML/SGML サロン. 標準 XML 完全解説. 技術評論社. 354p. 1998.
- 15) 村田真. XML 入門 - HTML の限界を打ち破るインターネットの新技术. 日本経済新聞社. 216p. 1998.
- 16) 坂田健二. Windows で学ぶ XML. 技術評論社. 367p. 2002.
- 17) 尾内恭輔. XML がわかる本. 毎日コミュニケーションズ. 255p. 2002.
- 18) 岡部 恵造. 図解 そこが知りたい! XML がビジネスを変える!. 翔泳社. 332p. 2000.
- 19) 屋内恭輔. XML がわかる本. 毎日コミュニケーションズ. 255p. 2002.
- 20) 檜山正幸. XML 用語集. [2002/09/04]
<http://www.ascii.co.jp/pb/ascii/special/old/XML/>
- 21) 新野淳一. XML 技術者のための情報提供/情報交換フォーラム. [2002/09/05]
<http://www.atmarkit.co.jp/fxml/>.
- 22) 石塚英弘. WWW ちょっと知っている人の XML 入門. [2002/09/05]
<http://www.ulis.ac.jp/library/SEMINAR/>
- 23) 中沢実. XML(eXtensible Markup Language)の習得. [2002/09/11]
<http://www.infor.kanazawa-it.ac.jp/~nakazawa/kougi/xml-jp/>
- 24) 赤坂玲音. XML 入門. [2002/09/13]
<http://black.sakura.ne.jp/~third/xml.html>

Web サービス

中村 祐一, 浦本 直彦

日本 IBM(株)東京基礎研究所

{nakamury, uramoto}@jp.ibm.com

1. はじめに

Web サービスは企業間取引や、大企業内でのアプリケーション統合を実現するための技術基盤である。XML やインターネット技術に基づいていることから、多くのソフトウェアベンダがその考えに賛同し、製品でのサポートを進めている。

本稿では、Web サービスの概念について述べ、そこで必要となる基本技術として SOAP^[1]、WSDL^[2]、UDDI^[3]を紹介し、その利点を概説する。そして、実用化への課題として、セキュリティとパフォーマンスの問題を取り上げる。

2. Web サービスとは何か？

Web サービスは、広義には、「自立、自己記述的、モジュラーなアプリケーションであって、Webの標準技術を用いて記述し、公開/検索し、呼び出すことができるもの (webservices.org)」ということができる。つまり、Web サービスそのものは非常に一般的な概念であり、CORBA、RMI、Agentといった従来技術にも当てはまる。しかしながら、通常は、Web サービスは XML、SOAP、WSDL、UDDI を技術基盤とするものを指すので、本稿でもこれらの技術を中心として説明する。

3. Web サービスの構成要素

Webサービスの特徴は、サービスプロバイダ、リクエスタ、ブローカの3者で、サービスの公開(publish)、検索(find)、それに呼び出し(bind)を行う点にある。ここでは、このようなこのモデルを支える3つの仕様である SOAP、WSDL、UDDI について簡単に述べる^[4]。

3.1 SOAP

XMLに基づくトランスポート独立なプロトコル及びメッセージ・エンベロープのための仕様が、2000年5月にMicrosoft、Lotus、IBMなど5社が中心となって W3C に提案した、Simple Object Access

Protocol (SOAP) 1.1 である。SOAP は XML の構文を用いたきわめて単純なメッセージ・エンベロープ構造と、同じく XML の構文を使った汎用のデータ直列化を定義した仕様である。SOAP1.1 エンベロープは Header と Body からなり、Body の中身がこのメッセージのペイロードとなる。Header は XML の特徴を活かして拡張可能になっており、経路情報、トランザクション情報、セキュリティ情報(例えば、電子署名)などを格納することができる。現在は、SOAP 1.2 の仕様が公開され、標準化に向けて協議中である。

3.2 WSDL

個々の Web サービスに関する情報(入出力のスキーマや接続方法など)を記述するのが Web Service Description Language (WSDL)である。分散オブジェクト指向プログラミングでは IDL を用いてサービス(関数)を定義するが、Web サービスでは XML(SOAP)メッセージを用いて入出力を行い、サービスの型は XML スキーマで規定される。

3.3 UDDI

Universal Description, Discovery, and Integration (UDDI)は、の名の通り、サービスの記述/発見/統合を自由に行うための基盤(<http://www.uddi.org>)である。その第一段階として公開ビジネスディレクトリを構築し2000年11月16日からサービスを提供している。このビジネスディレクトリはWebサービスをビジネス名、ビジネスの業種(カテゴリ)で検索できるだけでなく、サービスの型(tModel)でも検索できるところに特徴がある。現在は、UDDI Version 3 仕様が公開されている。

4. Web サービスの利点

最初に述べたように、多くのソフトウェアベンダが Web サービスのサポートを表明し、製品の「Web サービス化」を進めている。それでは、なぜソフト

ウェアベンダがこれほど Web サービスに着目しているのか考えてみよう。

- 相互運用性. 企業間取引はもちろん, 企業内においても様々なアプリケーションが稼動しており, それらを統合することは重要な技術課題である. そこでの統合は異なるプラットフォーム上に構築されたアプリケーションを含んでおり, 従来の同一のプラットフォームを前提とした統合は機能しない. Web サービスでは, XML/SOAPにより, 最低限のプロトコルを決めることにより, 「とりあえずつながる」ことを保障している. これにより, 例えば CORBA と. Net をつなぐような話も容易になり, アプリケーションの統合が格段に進化すると期待できる.

- 既存システムの継承. いくら Web サービスが優れた技術でも, 企業が既存のシステムを全て作りかえる必要があるなら, 全く受け入れられないであろう. Web サービスの真髄は, アプリケーション固有の API を XML/SOAP によるデータ交換という形でラップすることにより, アプリケーションの Web サービス化が簡単にできる点にある. 一旦 Web サービス化されてしまうと, アプリケーションの実装がどうなっているかに関しては, 呼び出し側はまったく関知する必要がなくなる. さらに, この考え方を発展させると, 新たに定義された XML の API を UDDI を通してビジネス・パートナーに公開することもできる.

以上のように見てくると Web サービスの利点は「手軽さ」にあるようにも見える. もちろん, それも重要な要素であるが, より重要な概念は「非集中化 (decentralization)」である. 現状の Web では, 世界中のコンピュータが様々なプラットフォームで稼動し, 緩やかにつながっている (loosely-coupled). このような状況は, 分散コンピューティングの観点から, 実は画期的である. というのも, 従来は同一のプラットフォームを用いて分散システムを構築することが当然と思われてきたからである. Web で実現された「非集中化」の世界で分散コンピューティングを実現することが, Web サービスの究極的な目的となるはずで

ある.

5. Web サービスの課題

現在, 既存のアプリケーションを Web サービス化して外部に公開したり, 企業内のアプリケーション統合を Web サービスによって実現しようとする動きが徐々にではあるが, 起こりつつある. ところが現状ではそれらは主に Web サービスの「手軽さ」という面に着目したものであり, ビジネス上の取引までは視野には入っていないものが多い. ここでは, セキュリティとパフォーマンスをとりあげ, その問題点を考える.

Web サービスでは, 非集中の環境下でアプリケーション同士がやり取りをするので, これまでとは違うセキュリティの考え方が必要になる. Web サービスセキュリティは, IBM, Microsoft, VeriSign が共同で提案した Web サービスに関するセキュリティの枠組みならびに仕様である. その中では, ロードマップとして, 7つのエリアを同定し, それらの仕様の策定を進めている. 例えば, メッセージセキュリティとして, メッセージの暗号化や署名の方法などが規定されている. これらが出揃うと, ビジネス上の取引もかなり安全に行えるようになると考えられる.

パフォーマンスは Web サービスにおける最大の問題点である. XML をやり取りすることは相互運用という点では大きな利点であるが, メッセージのサイズや処理時間に関して大きなオーバーヘッドとなる. Apache SOAP などの処理系を使って SOAP/HTTP を実行すると, 単純な RMI よりも何十倍も遅いという実験結果もある. パフォーマンスを向上させると同時に, 既存の高速な技術と適宜組み合わせることでアプリケーションを統合することが重要と考えられる.

以上のような課題はあるが, Web サービスは着実に浸透していくと思われる.

参考文献

- [1] 高瀬 SOAP, 情報処理学会誌, Vol.42, No.9, 2001.
- [2] 高瀬 UDDI と WSDL, 情報処理学会誌, Vol.42, No.9, 2001.
- [3] 丸山, 小坂, 浦本, Web Services による動的な電子商取引の実現, 情報処理学会誌, Vol.42, No.7, 2001

Webサービス

2002年10月22日

中村 祐一
日本IBM東京基礎研究所
nakamury@jp.ibm.com

目次

- ▽ Webサービスとは？
- ▽ Webサービスを構成する技術
 - ▽ SOAP, UDDI, WSDL, ...
- ▽ Webサービスの利点
 - ▽ CORBA, RMI, DCOM, ...とはどう違うのか？
- ▽ 課題
 - ▽ セキュリティ, パフォーマンス, ...

Webサービスとは？

狭義: 標準的な技術でアクセス可能なアプリケーション
HTTP, XML, etc. -> SOAP

サービスリクエスタ ↔ サービスプロバイダ

Webサービス

Webサービス

広義: WSDL (Web Services Description Language)で記述されたアプリケーション

Webサービスに関わるXML標準

登録 (UDDI) (WSDL) → プロバイダ (Publish)

検索 (UDDI) ← プロバイダ (Find)

プロバイダ ← 接続 (SOAP) ← リクエスタ (Bind)

- ▽ XML (1998.2.10)
 - ▽ 標準データフォーマット
- ▽ SOAP (2000.5.8)
 - ▽ XMLを運ぶ封筒
- ▽ UDDI (2000.9.6)
 - ▽ サービス情報の登録と検索
- ▽ WSDL (2001.3.15)
 - ▽ サービス接続情報

SOAP 1.1

Simple Object Access Protocol

- ▽ XML電文を運ぶための封筒
 - .. マイクロソフト, ロータス, IBM, 他による共同提案 (5/8/2000)
 - .. <http://www.w3.org/TR/SOAP/>
 - .. トランスポート独立(HTTP, SMTP, MQSeries, ...)
 - .. 軽量, 拡張性 (cf. ebXML TRP)
- ▽ 2000年9月, W3C WG発足 (XML Protocol)
- ▽ 3つの主要部分
 - .. メッセージエンベロープ
 - .. フォルト処理
 - .. データの直列化とリモート手続き呼び出し
- ▽ 現在 V1.2の標準化が進行

SOAPメッセージの例

エンベロープ(封筒)

```


<SOAP-ENV:Envelope
  xmlns:SOAP-ENV="http://schemas.xmlsoap.org/soap/envelope/"
  >
  <SOAP-ENV:Header>
    <t:Transaction
      xmlns:t="some-URI"
      SOAP-ENV:mustUnderstand="1">5</t:Transaction>
    </SOAP-ENV:Header>
    <SOAP-ENV:Body>
      <m:GetLastTradePrice xmlns:m="Some-URI">
        <symbol>DIS</symbol>
      </m:GetLastTradePrice>
    </SOAP-ENV:Body>
  </SOAP-ENV:Envelope>
  
```

ヘッダ (付加情報)

ボディ (本文)

UDDI

Universal Description, Discovery, and Integration



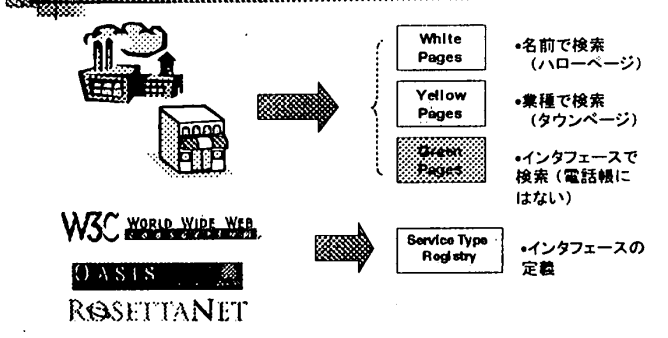
- Webサービスの相互運用のためのサービス仕様規格
 - http://uddi.org/
 - Ariba, IBM, MSが中心となり策定、33社で発表
 - 2002年7月現在、220社以上が参加
 - 2000年9月 UDDI V1.0
 - 2001年6月 UDDI V2.0
 - 2002年7月 UDDI V3.0
- グローバル・ビジネス・ディレクトリ
 - IBM, MS, SAPの3社で独立運営
 - 内容はそれぞれコピーされる

Yahoo!は人のために
UDDIは機械のために

UDDIでできること

- ビジネスに関する記述
 - この会社の名前, 住所, 代表者...
- Webサービスの記述
 - このWebサービスは何ができるのだろう?
 - インターフェース (DTD, プロトコル, etc)は何かな?
- Webサービスの発見
 - XXという情報を提供するWebサービスは?
 - 我々のアプリケーションと直接話せるWebサービスは?

UDDI V1 = グローバル・ビジネス・ディレクトリ



- White Pages
 - 名前検索 (ハローページ)
- Yellow Pages
 - 業種で検索 (タウンページ)
- Green Pages
 - インタフェースで検索 (電話帳にはない)
- Service Type Registry
 - インタフェースの定義

W3C WORLD WIDE WEB
OASIS
ROSETTANET

ビジネスの記述例

```
<businessEntity
  businessKey="9BB0ADC6-CEB3-11D3-849F-0050DA1803C0">
  <name>Webシューズ.com</name>
  <description>世界中の靴をあなたへ</description>
  <contacts>
    <contact useType="sales contact">
      <personName>靴売 増太郎</personName>
      <phone>03-1234-5678</phone>
    </contact>
  </contacts>
  <categoryBag>
    ...ビジネスカテゴリ(業種)情報...
  </categoryBag>
</businessEntity>
```

サービスの記述例

- あるビジネスは複数のサービスを持てる

```
<businessService
  serviceKey="762095AC-3F44-4742-92DE-12E8E39D08F9"
  businessKey="9BB0ADC6-CEB3-11D3-849F-0050DA1803C0">
  <name>婦人靴小売販売</name>
  <description xmlLang="JA">婦人靴の調達、小売販売</description>
  <bindingTemplates>
    ... サービス接続のための接続情報 ...
  </bindingTemplates>
</businessService>
```

接続情報の記述例

- あるサービスは複数の接続情報を持てる

```
<bindingTemplate
  bindingKey="00790B5C-E54E-40A8-BA43-7F9448A09BCC"
  businessKey="9BB0ADC6-CEB3-11D3-849F-0050DA1803C0">
  <description>HTTP/SOAP Access Point</description>
  <accessPoint>
    http://www.webshoes.com/order/
  </accessPoint>
  <ModelInstanceDetails>
    <ModelInstanceInfo>
      <ModelKey="uuid:E34D22FF-GDC7-11D3-A1DC-00A0C9CD766C">
        <description>靴業界ML (ShoeML)</description>
      </ModelInstanceInfo>
    </ModelInstanceDetails>
  </bindingTemplate>
```

インタフェース仕様のID

WSDL

Web Services Description Language

- Webサービスの記述言語
 - http://www.w3.org/TR/wsdl
 - IBMとMicrosoftの共同提案
 - サービスの型を記述するが、意味は記述しない

The diagram shows a central box labeled 'Web サービス' with a 'port' label. Arrows point to it from boxes labeled '入力のメッセージ型' (Input message type) and '出力のメッセージ型' (Output message type). An arrow points from the 'Web サービス' box to a box labeled 'オペレーション' (Operation).

抽象的な型の定義

```

<message name="GetLastTradePriceInput">
  <part name="body" element="xsd:TradePrice"/>
</message>

<message name="GetLastTradePriceOutput">
  <part name="body" element="xsd:TradePriceResult"/>
</message>

<portType name="StockQuotePortType">
  <operation name="GetLastTradePrice">
    <input message="tns:GetLastTradePriceInput"/>
    <output message="tns:GetLastTradePriceOutput"/>
  </operation>
  ...
</portType>
    
```

入出力メッセージの型

抽象インターフェースの定義

型とプロトコルのバインディング

- この例はSOAPのメッセージ型バインディング
- 他にもRPC型がある

```

<binding name="StockQuoteSoapBinding" type="tns:StockQuotePortType">
  <soap binding style="document" transport="http://schemas.xmlsoap.org/soap/http"/>
  <operation name="GetLastTradePrice">
    <soap operation
      soapAction="http://example.com/GetLastTradePrice"/>
    <input>
      <soap body use="literal"/>
    </input>
    <output>
      <soap body use="literal"/>
    </output>
  </operation>
</binding>
    
```

SOAP over HTTP を使う

入出力電文は SOAP BODYに入る

WSDLツールの利用

- 既存のソフトウェア部品からWSDLを生成
 - 既存アプリ → WSDLを生成 → WSDL → UDDIで公開 → 検索
- WSDLからスタブとスケルトンを生成
 - WSDLを検索 → WSDL → スケルトンを生成 → スタブ → スケルトン

Webサービスの利点

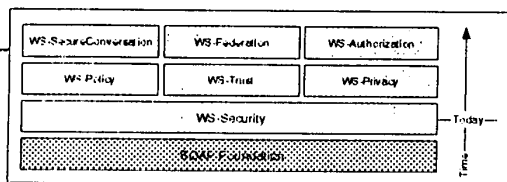
CORBA, RMI, DCE/RPC, DCOMなどとどこが違うのか?

- 分散プログラミング(Distributed Prog.)から、非集中化プログラミング(Decentralized Prog.)へ
 - 全体を管理する実体がない
 - システムは独立なコンポーネント群の緩やかな結合
 - WWWは史上最大の分散プロジェクト・かつ非集中化
- アプリケーション統合が密結合(Tightly Coupled)から疎結合(Loosely Coupled)へ
 - 統合のための要件(OS、言語etc)が小さい
 - いつも密結合が良いとは限らない

Webサービスの課題

- シビアなビジネスをやるためには不十分?
- セキュリティ
 - Webサービスにおけるセキュリティの標準化
 - IBM, MS, VeriSignの提案 (SUNも支持)
 - その後、OASISに提出
- パフォーマンス
 - XMLの処理が大きなオーバーヘッド
 - RMIより数十倍遅い!

WSセキュリティ・ロードマップ



概要

- ▼ 分かりやすいセキュリティモデル
 - ・既存のセキュリティ・インフラを抽象化したフレームワーク
- ▼ 現在, WS-Security仕様のみ出ている
 - ・近いうちに, WS-Policyなどが公開される予定

まとめ

- ▼ 現状
 - ▼ 基本的な標準はそろった(SOAP, WSDL, UDDI)
 - ▼ 次の段階の標準化が進行中(セキュリティ, トランザクション, ワークフロー)
- ▼ 課題
 - ▼ パフォーマンスには大きな課題が残っている
 - ▼ Webサービスを利用して本当に大規模開発はできるのか?

セマンティックウェブの技術的位置づけと課題

～ 周辺技術から見たセマンティックウェブ ～

慶應義塾大学 理工学部 管理工学科 飯島 正

1. はじめに

WWW(World Wide Web)技術の次世代版と目されているセマンティックウェブ(Semantic Web)[1][2]はいかなる技術であろうか。セマンティックウェブに関しては、米国の DAML(Darpa Agent Markup Language)プロジェクトにおいて、2000年から5年間(2004年まで)で7000万ドルもの資金提供がなされ、W3Cによって標準化がすすめられている。

The Semantic Webは、The Internetのように、ある種、固有名詞であるが、同時に semantic という形容詞を Web に付けた一般名詞でもある。その目指すところは、ドキュメントの字面だけに左右されずに計算機(ソフトウェア・エージェント)が意味内容を理解して利用できる Web, 名の通りセマンティクス(semantics; 意味論)を扱うことのできる Web 技術にある。その技術は、Tim Berners-Lee の示したアーキテクチャのマップ(図1)の上に積み上げられつつある。しかし、そうした個々の要素技術を拾い上げて理解することは重要ではあっても、その方法では、限られた時間の中で要素技術の発展を追いかけることに終始してしまい、なかなか全体像はつかみきれないという問題がある。

本稿では、セマンティックウェブのインサイダでも熱狂的なウォッチャーでもない立場から、次世代Web技術の一つと目されているセマンティックウェブの全体像に関連技術の中で位置づけて、その課題を論じることを目的としている。具体的には、Web 技術を中心に、それを取り巻く、ドキュメント、情報検索、データベース、自然言語、人工知能(AI)といった周辺分野を取り上げていく。

セマンティックウェブは人工物であるので、そこには設計思想がある。その前提となる主要原理として6つの原理があげられている[2]。こうした前提なる原理自体の正当性とそれが含む難しさに関しても、周辺分野との関係で位置づけていく中で確認していきたい。

【原理1】 すべては URI で識別できる

Everything can be identified by URIs
 ※URI(Universal Resource Identifier)
 (RFC2396:1998年8月,
 RFC3305:2002年8月)

【原理2】 資源とリンクは型を持ちうる

Resources and links can have types

【原理3】 情報が不完全でも困らない

Partial information is tolerated

【原理4】 情報が絶対的に正しいとは限らない

There is no need for absolute truth

【原理5】 発展性を支えられなくてはならない

Evolution is supported

【原理6】 できるだけ単純なことからはじめる

Minimalist design

これらの原理は WWW 上の資源を取り扱い、それを使ったアプリケーションを構築していく上での基盤を作るには重要なことであるのは直感的に感じ取ることができる。原理1と原理2は、情報構造を規定しており、原理5と原理6は将来に渡って「使える」ものにするための基本方針を述べている。しかし、原理3や原理4といったWWW上の情報の性質は、その「情報を知識に持ち上げていく」上で、多くの解決すべき困難な課題を含んでいることも同様に直感的に気づくことであろう。我々は、WWWの上に大量の情報があることを知っているだけで

なく、その中には誤った情報、古くて現状とあわない(賞味期限切れ)情報、情報の欠落(リンク先のページが見つからないことなんて日常茶飯事)、価値観の相違、矛盾のある情報に満ち溢れていて、権威者もしくは大量の価値観を同じくする者(非権威者かもしれない)による内容の保証はなされていないし、そのための枠組みがない。

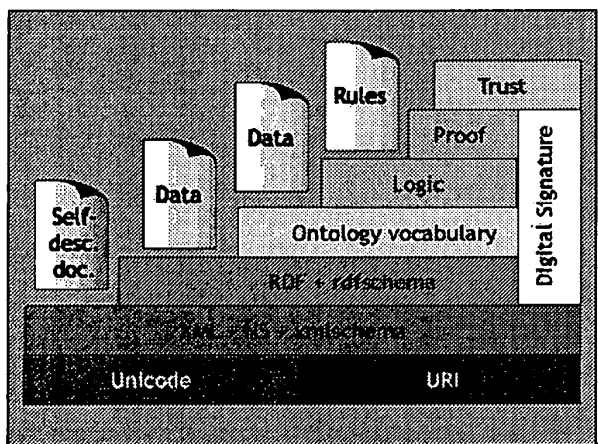


図 1.セマンティックウェブのアーキテクチャ[1]

そこで図 1 の 7 階層 (レイヤ) アーキテクチャを再び見たとき、既存の WWW の枠組みの上に素直にのっている下 2 層に加えて、概念的知識と推論、信頼性保証、価値といった層が加えられていると見ることができる。

2. セマンティックウェブのイメージ

サイエンティフィック・アメリカン[3]に掲載されたセマンティック・ウェブの一般向け解説をもとに多少の改変を加えて、利用者から見たイメージを提供する例題を提示する。イメージを提供するのと一緒のこの章では、セマンティックウェブの原理 1 について若干の考察を加えてみたい。

セマンティックウェブが提供してくれる未来像は例えば以下のようなものである。休日に車を運転していて急に歯が痛くなったとする。そのとき、携帯情報端末や車載情報機器の中のソフトウェア・エージェント(代理人; 個人秘書のようなもの)が、現在走行中の地点からもっとも近くて休日診療してくれる歯医者さんを探し出し、自動的に予約をとってくれる。なおかつ、本人のアレルギー等の必要な病歴情報なども(必要に応じてかかりつ

けの病院の情報システムから入手して)送っておき、ついでに、到着するまでの間のための医師の指示を聞くために電話をかけてくれる。もちろん、運転中なので、運転者が電話番号を押すようなことはさせない。電話で話をしている間は、車内のオーディオの音量を下げて、通話の邪魔にならないようにする。自動車のソフトウェア・エージェントは道路情報システムにアクセスし渋滞していなくて短時間で到着できる走行経路を割り出す。それを取った歯医者さんのエージェントはその走行経路上の薬局に鎮痛剤の処方箋を送って注文しておき、医院に到着する前にそれを受け取って飲んで置くように誘導する。もちろん、薬代の支払いは利用者のクレジットカードから引き落とされるように、エージェントが自動的にカード番号と有効期限を通知することもあるだろう。

いかがであろうか。ここには、非常に賢く気の利くソフトウェア・エージェントがいる。そのソフトウェア・エージェントは、ネットワークを介して病院情報や予約サービスにアクセスし、またオーディオ等の車載機器の状態を把握して制御して、利用者を助けてくれる。

ここでやっていることの中には、近くにあつて休日診療してくれる歯科医院を探すという情報検索がある。そうした情報が WWW 上にあり、その情報は人間が見るのではなく、ソフトウェア・エージェントがそれを読み理解し利用する。その際に、単に休日診療と書いて WWW ページをそのままキーワード検索するのではなく、「祝祭日休診、但し日曜は診療します」といった表現から、その当日に診療があるのかないのかを判断するといった賢い情報検索が必要とされる。セマンティックウェブのもっている第一のイメージはそれを可能にするものということになる。

しかもソフトウェア・エージェントが把握する情報は幅広い。ネットワークでつながれた車載機器の状態や、必要に応じてセンサを通して取得した情報なども推論や判断の対象として扱われなければならない。そして、そうした情報を元におこなった判断をフィードバックし、利用者に適切な形で提示し、意思決定をサポートし(一部を肩代わりし)、外部の情報サービスへのアクセスや、外部機器の制御といったアクションをとらねばならない。必要に応じて、複数のサービスを組み合わせることで、利用者の潜在的・顕在的な要求・欲求を満たさ

ねばならない。

このための大前提として、まず第一に、情報サービスや、機器、医師といった役割を持った個人をもソフトウェアが扱えるようになっていなければならない。そこで、原理1がある。【原理1】「すべては URI で識別できる」。これは WWW 空間上の資源に関しては問題ないにしても、実世界とのかかわりができたときに、難しくなる。機器を識別する URI に対しては IPv6 の導入がそれを押し進めることにもなるであろうが、個人を識別するための URI を与えることに関しては、ちょうど本稿執筆時期において住民基本台帳ネットワークの導入に付随する ID 番号の導入が、プライバシー保護のための法整備の遅れもあって、利便さを理解しつつも、国民の間で根強い嫌悪感(というより不安感)を引き起こしていることからみて難しい現状にあるかもしれない。本稿の趣旨ではないが ID の導入が個人情報の一元管理につながるという見方は短絡的であり、むしろ ID さえなければ分散した各種個人情報の結合(join)ができないかのような幻想を与え危険ですらある(1980年から1985年頃に公表された五木寛之氏の小説「ガウディの夏」では、IDもコンピュータパワーもなしにマンパワーでそうした情報のリンク付けをする話がポイントの一つとなっている)。しかし、セマンティックウェブ(DAML)の用途の中には、防諜、反テロリズム(counter-terrorism)も挙げられており[4]、テロリストの発見や監視も含まれているのかもしれない。

3. WWW 小史の中でのセマンティックウェブ

20世紀が残すところあと10年という時期に端を発しているWWW(World Wide Web)技術は、5年もたつたかないかというたないうちに大きな普及を遂げ、いままでもなくインターネット技術の中で、アプリケーションとして大きな成功を収めたものである。しかし、それは過去の成功にとどまらず、今なお、それを基盤としてさらなる進化を続け、発展を遂げようとしている。その大きな方向性が、静的な情報提供から動的コンテンツ提供へ進む流れの延長上に、サービスのプラットフォームとしてのWeb、そして、その上でのサービスの相互運用という流れである。人がインターネット上のサービスをより有効に活用していくためには、それを手伝ってくれるソフトウェア(エージェント)がいてくれると助かる。まずは、WWW

技術の歴史[5]の中から、そういう潮流を掴み出してみたい。

インターネットとは、いまさら歴史を紐解くまでもないと思うが、ネットワークの相互接続(inter-network)に端を発した用語であり、米国国防総省(DARPA)のプロジェクトをベースに構築された、定冠詞を付けた固有名詞The Internetとして(地球上?で唯一の)通信ネットワーク相互接続網を指している。もちろん当初は、WWW技術はなく、電子メールやネットニュース、FTPによるファイル交換といった利用が中心であった。しかし、インターネットという語が巷間に広まったのは、WWWという技術の普及によるものであるのは否めない。WWWにはWWWの流行した理由がある。

3.1. ドキュメントデータベースとしてのWeb

WWW技術は、1989年にスイスのジュネーブにあるCern(Council Européen pour la Recherche Nucléaire; 欧州合同原子核研究機関; the European Organization for Nuclear Research)においてイギリス人Tim Berners-Leeによって提案され、1990年にその基礎が構築された。WWWはデータサーバにファイル単位で保存された文書情報を開覧ソフト(ブラウザ)を備えたクライアントマシンから取り出すというサーバ/クライアント・モデルで実現された一種の分散型の文書データベースである。ここで、標準となる、文書のマークアップ言語HTML(HyperText Markup Language)と転送プロトコルHTTP(HyperText Transfer Protocol)が規定された。但し、その名に含まれているハイパーテキストとは、通常のテキスト文書の持つ線形構造以外にリンクという非線形構造を導入したものであり、そのリンクは他のサーバに置かれた文書の一部に対しても張ることができる。WWWの名前はこれに由来する。Webとは蜘蛛の巣、網のことであり、WWWとは世界中に張り巡らされた網ということになる。

もっともハイパーテキスト(ならびにテキスト以外の画像等の(マルチ)メディアに概念を拡張したハイパーメディア)の概念はWWWの登場以前からある。その源流は、第二次大戦中の1945年の7月にルーズベルト大統領の科学顧問であり6000人もの科学者を率いる立場であったVannevar BushがThe Atlantic Monthlyに掲載した"As We May Think"というタイトルの論文[6]の中で構

想している情報共有のためのシステム”Memex”とされている。Doug Engelbartの知識の共有を支援し人間の知的能力を増進(augment)させるためのシステムは、この構想に刺激されNLS(oN Line System)のデモに結実した。ここで情報を顕在化して計算機をメディアとして操作するということの重要性が認識され、更には、ハイパーテキストの概念につながる文献の参照関係を取り扱える技術も生まれた。Ted Nelson(Theodor Holm Nelson)は、1960年にハイパーテキストの着想を得てXanaduプロジェクトを開始した(HypterTextという名前を選んだのは1963年、出版したのは1965年、Xanaduという名前を選んだのは1967年とのことである[7]。

HTMLの標準化は、その後、HTML2.0(RFC1866; 1995年11月)、表構造のためのタグの追加(RFC1942; 1996年5月)、ブラウザ独自のタグの追加などを経て、HTML 4.01(W3C勧告; 1999年12月)が発行され現在に至っているが、ここで、ハイパーリンクという概念とマークアップという概念を分離して考えてみたい。

テキストのマークアップということに関しては、HTMLが構築される上での下地として、数学論文をはじめとする各種出版のための組版システムTeX(とそのファミリー)と電子化文書の論理構造のためのマークアップ言語SGML(Standard General Markup Language)があった。HTMLという名前の一部にもなっているMarkup Languageのマークアップとは、元々、印刷のために原稿に付ける活字書体・ページ割付などのための組版指示のことである。いろいろなマークアップ=メタデータの分類を考えた場合、見た目の制御のために原稿に注釈付けをするためのマークアップと位置づけることができる。TeXは、Donald E. Knuthが著書The Art of Computer Programmingの組版のために1980年前後に開発したシステム(ソフトウェア)である(そのTeXシステムのためのマークアップ言語という感覚でTeXという名前を使うこともある)。つまり、TeXはあくまでも印刷したときの品質(つまり見た目)を志向したものであるといえる。この見た目というのは重要な要素であり、WWWが普及したメリットの一つには、この見た目という効果が作成者にとって簡単に評価できフィードバックして改善していけるものであったということも大きい。これは、作成者に動機付けを与えてくれる要素の一つである。

一方、1986年ISOにより標準化されているSGMLは、特定用途のドキュメントには特定の型(タイプ)があり、見た目に左右されない論理構造(章立てなど)を物理的な体裁から独立に規定できるということをしこうしたものであった。HTMLの基本的な表現は、このSGMLから引き継いだものであるが、ドキュメントの型を規定し、スタイルと論理構造を分離するという点に関しては、(標準的な)HTMLでは一旦失われ、再びその発想が復活することは、XML(eXtensible Markup Language)の登場まで持ち越されることとなった。SGMLは、ドキュメントの型を文書型定義(DTD; document type definition)として定義することができる。DTDは、データベースのスキーマに相当する。スキーマは、データの構造と意味とを結び付けるメタデータである。これにより文書構造(内容面)の標準化と、構造に対応付けられ他見た目(スタイル)の標準化を、並行して図ることができるという期待を担っていた。しかし、その期待とは裏腹に、そのドキュメントの型の定義をすることの難しさから、広く一般に広まったというには程遠かった。それでも電子調達システム等で(データ項目の多様化と頻繁な変化に対応しなければならぬプログラムの負担を軽減するための)データ交換のためのメディアとしての用途への広がりもでてきて、文書の保存(データストア)、データ交換、見た目の制御というXMLの現在の主要な方向性が生まれている。XMLの標準化に関してはXML1.0(W3C勧告; 1998年2月)、同第二版(W3C勧告; 2000年10月)、1.1(W3C作業ドラフト; 2002年4月)といった経緯をたどっている。

ドキュメントの見た目を制御できるというマークアップと同時に、WWWの普及の大きな動機付けになったハイパーリンクに立ち戻ってみる。ハイパーリンクは、利用者がそれをたどっていくことで、関連する情報に到達するような、いわゆる「ナビゲーションな検索」機能を提供するものである。これは利用者にとってのメリットであると同時に、ドキュメント作成者にとっても大きく動機付けをしてくれた。というのは、自分で作成したドキュメント断片でなくても、WWW空間上にある他人の作成したドキュメントへのハイパーリンクを張ることでこの情報を参照することができるということである。これは、原著者の随時の改訂を許容しながら引用するといった、一種の情報の再利用を可能にしている。但し、著作権を保証するものと

は言い難い。

また、リンクに対して参照関係以上の「関係性」を明確に示すための標準的な方法や双方向性、多項関係などを明確に示す標準的な手段はHTMLには提供されていない。また直接関係付けられたリンクをナビゲートすることがあるが、いわゆる推論のように多段の関係を組み合わせた一つの複合的な関係とみなす標準的な手段も与えられていない。しかし、WWWの開発に当たってのTim Berners-Leeの提案書にあるハイパーテキスト(図2)[8]には、リンクに型があった。セマンティックウェブの【原理2】「資源とリンクは型を持ちうる」は既に、初期の構想の中に含まれていたものであった。この図をみて、データベースの概念設計のための実体関連図(ER図)、オブジェクト指向ソフトウェア設計のためのクラス図(UML図の一つ)、人工知能のための知識表現の一つである意味ネットワークを思い出す人も多いだろう。それらとの関係についてはまた後程論じる。

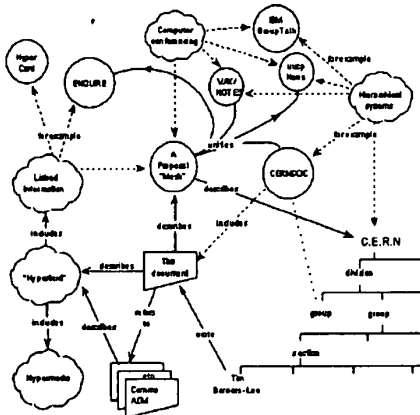


図 2.WWW 提案中のハイパーテキスト[8]

3.2. 情報検索としての WWW

WWWの普及のための第三の動機付けは、ロボット型検索エンジンの普及である。ネットワーク上に張り巡らされた情報の網(Web)の上を蜘蛛のごとく這い回って、予め与えたキーワードではなく、ドキュメントからのキーワード抽出と自動索引作成(ディレクトリ作成)を行う spider もしくは crawler(這い回るもの)によって、高速な全文検索のような感覚で使うことができた。しかし、そうした WWW 検索では、基本的には意味を扱うものはそう多くはない。つまり、その検索に使えるキーワードの柔軟性はせいぜい大文字と小文字、全角と半角を区別しないといった字面上の程度であって、それ以上の表記の揺れや同じ意味の別の表現(同義語への言い換え、外国

語への翻訳)、シソーラス(thesaurus)が与えてくれる上位概念/下位概念といった意味的な包含関係(包摂関係)には対応してくれるものはそう多くはない。しかし従来の検索の延長として、そうした「意味」を扱って欲しいという要求は、ごく自然なものである。セマンティックウェブでは、図1の階層アーキテクチャの中で、[Ontology vocabulary]層で扱われることとなる。オントロジ語彙とは、主に名詞的(体言的)な概念階層構造のことを指すと言ってもよいだろう。同意語や概念の上下関係は、これによって提供される。セマンティックウェブの活動の中では、DAML+OIL というオントロジ記述言語が推進されてきたが、それをベースに OWL という規格に変わりつつある。特定の専門分野(ドメイン)の専門用語のオントロジ記述は、その分野の専門家がそうしたオントロジ記述言語を使って記述することになるが、WWW 上の情報を活用しようとなるとそうしたものばかりではなく、自然言語処理/理解のための電子化辞書、たとえば EDR 辞書、WordNet / EuroWordNet などを使うことで、より広い一般語もカバーしうる。しかし、そのように複数のソースからオントロジを構築した場合、当然のことながら複数のオントロジが乱立することとなり、それらの中の相互変換、上位オントロジの利用などが必要とされるようになってくる。

また、情報検索においては曖昧な表現に対する解釈も重要である。歴史上の人物には、生年月日が不明だったり、諸説あるというケースもある。「生年月日 が1900年頃」と「1900年～1903年の生まれ」、「1992年生まれ」、ひいては「生誕 100年」といった表現を単一化/同一視できるようにするための知識も必要となる。「数名」とか「若干名」、「2-3名」といった表現も必要となる。これらは、図1のセマンティックウェブの階層アーキテクチャの中では[Logic]層で推論規則として与えられることとなる。

もっとも、どこまで厳密に推論する必要があるかはコストとのトレードオフであり、アプリケーション(エージェント)に任せることができる部分も多い。例えば、医院の検索において複雑な推論を行って「今日が本当に臨時休業でないか」まで決めることは必要がない。そもそも、【原理 3】と【原理 4】に基づいて、推論に必要な情報が欠落していたり、古くて間違った情報が含まれていた

するので、そこにコストを度外視して注力しても(たとえば複雑な推論をしてレスポンスが遅くなると)かえって利用者満足度を下げる結果になりかねない。その日が本当に臨時休業でないのかは、ある程度絞り込めさえすれば、予約(契約)するエージェントが医院のエージェントと交渉する際に確認すればよいことである。

ここで、情報検索の対象について考えてみたい。これまで、WWW 上のドキュメントの検索としては、ドキュメントそのもの(全文検索)ならびに、そのドキュメントから自動的に抽出されたキーワードを考えてきた。このキーワードのように、ドキュメントの本文そのものではなく、そこから抽出した情報を対象とすることで、検索(選択、収集)の処理を速めることができる。たとえば、図書館の図書や論文検索システムでは、本のタイトル、著者名、出版社名、出版年、キーワード、アブストラクトなどを対象に検索が行われる。こうした情報はメタデータ(meta data)もしくはメタ情報(meta information)と呼ばれる。メタデータとは「データに対するデータ」という意味である。セマンティックウェブの文脈では、元のデータを資源と置き換えて、「資源に関するメタデータ」というような言い方をする。このメタデータという言葉は、独り歩きしていて頻繁に使われている割に分かりにくい。上記の文脈では、検索という処理を前提にして、その処理で使うための情報を抽出したものである。タイトルや著者名といった情報は、一般に元の資源(論文の本文)に含まれているものである(もちろん一部が省かれていることはあるが)。このタイプのメタデータを標準化する活動に、Dublin Core Metadata Initiative の DublinCore メタデータ要素集合という基本となる 15 の要素タイプ(バージョン 1.1)がある。

しかし、メタデータとは、そればかりではない。その対象となるデータを「処理する主体」(それをエージェントと呼ぶこととする)の「処理」のために使われるデータ、それも特定の個々の処理のレベルには限定しないが、データの利用を想定した利用法のデータという方が、すくなくとも現状の使われ方をよくあらわしている。例えば、コンテンツの内容を評価し利用の仕方を規定するメタデータ PICS(Platform for Internet Content Selection)や個人情報利用に関するメタデータ P3P(Platform for Privacy Preference Project)、機器の機能と連携性や利

用者の嗜好のためのメタデータ CC/PP (Composite Capability/Performance Profile) などがある。こうしたメタデータは、元データ(例えば、ドキュメントの本文)には含まれず、別途与えられるものである。

メタデータが与えられる単位も重要である。必ずしも一つの資源に一連のメタデータということではなく、より小さな単位を想定した方が的確なこともある(その場合は、より粒度の小さい副資源と考える方が順当ともいえる)。例えば、ある語に対して、ボールド体という指定を行うことは、印刷という視覚化処理の際の振る舞いを規定するメタデータといえる。このように、メタデータの与えられる単位は資源単位と考えるほうがより一般的である。

もう一つ、重要なメタデータがある(前者を注釈的メタデータとすると、構造的メタデータと位置づけられる)。文書の型(DTD)やデータベースのスキーマは、その文書がどのような構造をもっているか、その構造中のその要素(例えばデータベースの第三カラム)がどのような意味(フィールド名)を持っているかと規定するものである。そのような「構造と意味を関連付けるメタデータ」もセマンティックウェブの中には数多く現れている。

3.3. データベースとしての WWW

データベースとしての WWW もセマンティックウェブへつながる道筋の一つである。HTML というよりも、XML はデータを交換したり格納するためのスキーマとしても広く使われており、ファイルシステムをベースにした WWW システムではなく、最近では XML をリレーショナルモデルに変換(マッピング)して、RDB(リレーショナルデータベース)へ格納したり、XML ドキュメントをそのまま格納し検索することを目指したデータベースも作られ、実用化されている。

図 2 において、Tim Berners-Lee の当初のハイパーテキストの概念図が実体関連図に似ているのを見たが、データベースの概念設計を格納のためのデータ物理設計を意識せずに扱うことができれば、効率はともかく柔軟性は高まる。スキーマ(構造と意味をつなぐメタデータ)をあまり厳密に規定せず柔軟性を持たせたデータモデルに半構造データモデル(Semi-structured Data Model)がある。このモデルにも幾つかのタイプがあるが

木構造をベースにしたものは XML 表現に非常によく対応する。また、図 2 はオブジェクト指向ソフトウェア分析設計に使う UML(Unified Modeling Language)のクラス図ともよく似ている。そうしたクラスをデータスキーマとし、アクセスメソッドと組み合わせるモデルにオブジェクト指向データモデル(Object Oriented Data Model)がある。これらは、図 1 の階層アーキテクチャの中で [RDF+rdfschema]層に相当する。さらには、推論機能でデータを補完する演繹データベース(deductive database)と組み合わせた演繹オブジェクトデータモデル(例えば、F-Logic)のようなものもある。こうしたモデルを選択することで、[Logic]層とも連携がよくなる。WWW 空間上のリソースの間に関連付けを行うと同時に、リソースの内部のプロパティを利用者ならびにソフトウェアから扱えるようにすることで、エージェントによる自動処理への道が開けることになる。

ここで、セマンティックウェブの発想とデータベースの発想を少しくらべてみたい。セマンティックウェブでは、【原理 3】において Web のリンク切れといった情報の不完全性の問題(不完全情報、部分情報とはリンク切れのことだけを指すのではないが)を許容することを宣言している。WWW 空間という不特定多数がかかわるオープンな空間においては、無矛盾性を強制することは極めて困難である。通常のデータベースを考えれば、そうしたことは考え難い。実用上は NULL 値の扱いには微妙なところがあるが、各種の一貫性制約は、データベースという閉じた世界(closed world)における無矛盾性を志向している。

3.4. サービス・プラットフォームとしての WWW

Cern という一研究所内(閉じた組織内)の文書管理システム/知識管理(ナレッジ・マネジメント)システムであった WWW が世界中に広まるにつれて、インターネット=Web というような認識さえ、一般利用者の間にはあった。一時期は、一般利用者の間で「インターネットする」といえば、WWW のブラウジング(閲覧のこと、通称ネットサーフィン)を指していた。その時点でのハードウェア媒体(メディア)は主に自宅や職場に設置された PC(パーソナルコンピュータ)であった。

しかし、近年、ADSL 等によるブロードバンドの普及、

無線 LAN をつけた駅などの公共施設におけるインターネット接続サービスの開始、携帯電話による電子メールが広く普及することで、さらに裾野を広げると同時に利用する場所を選ばなくなりつつあり、インターネット技術と社会とのかかわりが変化しつつある。情報端末は、社会の中に浸透し、特別に計算機(コンピュータ)として意識されるものではなくなっている(どこにでもコンピュータが遍く存在するという概念を意味するユビキタス・コンピューティングにかわり、パーベイシブ・コンピューティングという言葉も一般化してきた)。そうしたメディア(ハードウェア媒体)面から引き起こされた大きな変化と同時に、コンテンツの面からも一つの変革をしようとしてつつある。

その一つが本来の Web サーバの目的であったドキュメント(静的なコンテンツ)を提供するためのサーバから、もっと多様なサービス(動的なコンテンツ)を提供する Web サービスへの変化である。もともと、それは、Web サービスが話題になる前から、CGI(Common Gateway Interface)を介したデータベースアクセス、ASP(Active Server Pages)や JSP(Java Server Pages)、さらには EJB(Enterprise Java Beans)といった形で、オンラインショッピングや、バックエンドに置かれた DB 検索(列車時刻表)という現実のビジネスを通して育ってきたものである。そのように、サービスへのアクセスに必ず Web を使うことで、クレジット番号などの個人情報の管理や、利用履歴の管理が容易となり個人の嗜好性の蓄積利用など利便性の向上に役立ってきた。また単に個人の志向性をつかむだけではなく、同じ価値観を持った大量の利用者の志向を活用することで有用な助言システム(ソーシャルフィルタリングによるリコメンデーションシステム)が可能となった。

しかし、それが、携帯電話、携帯情報端末というハードウェアを得ることによって、さらに外出中も持ち歩いて認証・決済の機能を備えた外部サービスへの利用者インタフェースと発展してきている。

Web サービスは、UDDI(検索サービス)、WSDL(インタフェース定義言語)、SOAP(XML をベースとした通信プロトコル)という3要素をもって語られることが多い。現状の技術の下での狭義の定義といえるかもしれない。しかし、IIOP→SOAP、IDL→WSDL、Trader→UDDI という

図式を考えると、オブジェクト指向の概念でクライアント/サーバモデルを分かりやすく構築することを目指した分散オブジェクト技術 CORBA を Web プラットフォーム上に移行し一般に広げたものという見方もできる。その意味で、Web サービスというのは、過去の遺産を引き継いだ分散コンピューティングのための極めて堅実な発想といえる。一方で、Web サービス(W.S.)と対比させて W.S. セマンティックウェブ(S.W.)を考えると、セマンティックウェブ自身にはチャレンジングなところが見受けられる。しかし、セマンティックウェブが実現しようとしているものも広い意味での Web サービス(Web を介したエージェントによるサービス)であり、その質をより高めようという活動に他ならない。そういう意味から、セマンティック Web サービス(S.W.S.)という現状の技術に限定しない Web サービスの将来のあり方が考えられる。

セマンティックウェブの活動の中にも Web サービスのためのものがある。Web サービスのためのオントロジ記述言語 DAML-S である。DAML-S では、プロセス(処理の流れ)記述するオントロジ(プロセス・オントロジ)を備えており、サービスの機能名や入出力の名詞的なオントロジだけでなく、より詳細なサービスの記述を、サービス間連携や、課金・監視等の運用管理のために提供している。

同様な活動は、WSFL さらに BEPLAWS(Business Process Execution Language for Web Services, Ver.1.0; 2002年7月)にも見受けられる。

4. セマンティックウェブの課題

4.1. 意味の抽出と表現

セマンティックウェブでは、エージェントが取り扱うデータは、資源(例えば元データとなるドキュメント)から抽出するものではなく、メタデータとして与えられている。しかし、元データの含まれているような種類のメタデータであれば、エージェントが自動的に抽出することも考えられる。しかし、そのためには、まさにドキュメントの意味を理解することが必要になる(正確に言えば、メタデータを与えられてそれをあつかうエージェントが意味を理解しているとはいいがたい。生データからメタデータに相当するものを抽出することこそ理解と言え

のではないだろうか)。

セマンティックウェブの志向している意味の取り扱い、特に、人間が扱うための自然言語(ドキュメント)から意味を抽出することはこれまで機械翻訳システムなどが内部に持つ自然言語意味理解技術において研究されてきた。この難しさはいままでもない。

自然言語意味理解において意味の表現にはいろいろな提案がある。あくまで自然言語という表層表現から深層的な意味へ変換するということからすれば、RDF とい似た構造を提供できる素性構造(feature structure)をベースに格フレームによって動詞を中心に格構造を表示する方法が対応付けしやすい。RDF のモデルである〈資源、属性、値〉の三つ組は、「述語(主格、目的格)」に対応するので動詞(述語)を中心とした言明(statement)の単純化したものとなっている。とはいえ、本格的に意味を取り扱うためには深層格としてどのようなものを使うかといったことにすら、なかなか解決が得られていないのが現状である。

4.2. データから知識へ

ドキュメントから意味を抽出するのは自然言語理解技術を使って難しいとはいえアプローチしていくとして、他にはデータから意味を抽出するにはどう技術があるだろうか? そうした技術の中で最近注目されているのはデータマイニングである。データから法則性を抽出し、ルールとすることで、「データ→情報→知識」といったもちあげを図るものである。対テロリズム(Counter-terrorism)といった用途には、こういう手法を使った方法が考えられる。

また、不確実情報をあつかうということからすればベイズ推論をベースにしたベイジアンネットワークなどもデータから知識を抽出する方法として有力な手法であるといえる。

4.3. 知識の利用と獲得の問題

4.1と4.2で知識を抽出する手法について簡単に眺めたが、現実には自動抽出以前に、人間が手作業で、オントロジを定義し、資源にメタデータをつけることでセマンティックネットワークに似た知識表現を作成し、その上で知識を補完したり利用していくための推論規則を与え

ることが必要となる。

しかし、そこには大きな課題がある。かつて人工知能ブームだった頃、1970年代後半から1980年代にかけては、知識システム、推論規則に基づいたエキスパートシステム(専門家システム)が多く作られていた。その当時、Feigenbaum は実用指向の AI を目指して知識工学を提唱したが、特定分野の診断や設計に使うためのエキスパートシステムの構築すら、そのドメインの専門家には難しく、知識獲得の専門家である KE (Knowledge Engineer; ソフトウェア技術者 SE に対して) を大量に養成する必要性を説いていた。

今や、まったく同じ問題がセマンティックウェブの実用にあたってはおこるはずである。いかに大量のデータが WWW 空間上に蓄積されていようと、その質はバラバラで統一に欠けている。セマンティックウェブが補完しようとしているところが、そうした統一性のない表現を扱うことにあるとしても、そこから先の利用につなげるためには、やはり一定水準以上のデータ資源を WWW 空間上に載せることが必要である。そして、それを実現するためには、「使う人」ばかりではなく「作る人」を作らねばならない。「作る人」を増やすには動機がある。やはり自分にとって役立つものでなければ他人のためにメタデータを付けてくれる人はいない。また単に「作りたい人」ではなくて一定の水準以上のものが「作れる人」を大量に必要としている。実用を考えるとこうした問題点は大きい。

そうした教育と普及の鍵となるのは動機付けであるが、更に、そもそもの程度の推論能力が与えられるのかという問題もある。【原理 6】「最小の設計」において、「まずは単純な仕掛けで単純なことができればよい」というものがある。しかし自分の要求水準以下のことしかできないのであればいずれ、不満が高じる。たとえば、紀田順一郎「インターネット書齋術」(ちくま新書)では、日本人女性の姓名を調べるのにインターネットを使う例がある。「聖子」と「りえ」のどちらが多いかをしらべるのに、「りえ」の方は「ぬりえ」「後りえ」「にごりえ」といった雑音を拾ってしまうとあるが、仮にこれは、推論機能が文脈から判断して的確にフィルタリングしてくれたとする。また二重にカウントしていないかどうかは URI によってなんとか識別できたとする。仮にそこまでできたとしても、そもそもが、「りえ」という名前を持つ全員がホームページを持ってい

るとは限らない。つまり、網羅性、完全性が必要となる。計算機での推論では、結局、検索に基づいて証明することが多いので、P が成り立たないという証明は、反例を検索し、それが見つかればいえるという方式が中心であって、閉じた有限の世界でないと反例がないことを示すのは難しい。たとえば、日本国内に狂牛病の牛がいないことを、WWW 上に載せられた牛の経歴、検査結果や餌の生産・流通ルートから判断しようとした場合、やはりすべての飼育されている最新の牛の情報を WWW 上に掲載することを義務付けるような強制機構が必要とならざるを得ない。そうした強制がそもそも WWW というコミュニティの中で可能なかどうかということについては、否定的な感触を持つ人も多いと思われる。

知識の信頼性の面でも、そうした一定水準以上のものを保証する機構がほしい。セマンティックウェブでは階層アーキテクチャの上2層で Proof と trust と扱っており、認証に支えられたお墨付き(レーティング)による信頼性の連鎖が、信頼性を考慮した推論を可能にするとしている。しかし、信頼できる情報がある程度以上、集まらないことにはメリットが生まれず動機付けとなりえない。動機付けができるところまで、艇入れをしてうまくリードしていくことが大きな課題といえるだろう。

4.4. メディアータとしてのソフトウェア・エージェント技術

セマンティックウェブにおいては資源が充実すれば、それを使うソフトウェア・エージェントへの期待は当然高まる。特に、人間と資源(サービスを含む)の間で仲立ちをしてくれることと、資源間の仲立ちをして、うまく連携してくれて、ひいては人間の役に立ってくれるという、2つの仲介ポイントがある。前者はパーソナルアシスタントエージェントと呼ばれ、後者は協調エージェントとよばれることがある。主に後者のための規格として、FIPA という団体がエージェントのプラットフォームとエージェント間の通信言語の規格が提案されている。このエージェント間通信言語は、KQML (Knowledge Query and Manipulation Language) という知識ベース間の知識交換を志向した言語の影響を受けている。その通信(対話)モデルは、構文論、意味論(セマンティクス)とならぶ語用論(プラグマティクス)に分類される発話行為理論

(SAT; Speech Act Theory)に基づいてエージェントが影響を及ぼしながら会話するモデルとなっている。

また、エージェント同士の自動的な契約におけるSLA(Service Level Agreement)の調整などは今後の課題といえるだろう。

最後に少し長くなるが Clifford Stoll の著書『インターネットはからっぽの洞窟』 [9]から、何人かの人工知能学者の未来の図書館へのビジョンを引用する。同書では、これらのビジョンは比較的、否定的に扱われているが、セマンティックウェブの将来像を思わせるものといえるかもしれない(以下、この章の終わりまでの斜体部分が引用箇所)。

未来の図書館はどんなものになるのだろうか？ 人工知能の創始者、MITのマービン・ミンスキー教授は、未来の世界から昔の図書館を回顧するかのように語る。「蔵書同士が会話できないような図書館があったなんて想像できますか？」

図書館については、エドワード・ファイゲンバウム、パメラ・マコードック、H. ベニー・ニールといった人工頭脳の権威も同じようなことを述べている。「現在の図書館というのは、蔵書倉庫にすぎない。書物も専門誌も、書架に置かれて利用者がやってくるのを待ち受けているだけである。本を見つけたり、読んだり、理解したりすることは、すべて利用者がやらなければならない。そして内容を理解するという作業も、利用者が自分で考えてしなければならない」

ファイゲンバウム教授の図書館システムは、いろいろな専門家の知識を持った一種のエキスパートシステムで、その知識を使って自分で必要な情報を収集したり要約したりできる。専門知識と判断/推論の能力をあわせもっているから、図書館利用者を積極的に支援できる。専門知識は、システムに依存しない形の知識データベースとして保存され、知識工学技術者(ノレッジエンジニア)によって保守管理される。

つまり本も雑誌も新聞もなければ、職員もいない、知識だけがぎっしり詰まっているのが未来の図書館なのだ。

『ネットワーク、開かれたアクセス、仮想図書館』

の著者、クリフォード・リンチは、この本のなかで1996年の予測を述べている。図書館情報システムの推進者である彼によれば、「どんな情報が必要かを、ユーザーからソフトウェアを介して告げられたワークステーションは、ネットワーク上のいろいろな情報にアクセスし、情報収集に要する費用を計算し、複数の情報を一本化させ、新しい情報を自動学習し、自らの知識データベースを積極的に拡張することができる。このためユーザーが地元の図書館に出かける必要はなくなってしまう」

リンチの図書館システムには、ワークステーションやネットワークはたくさんあるし、知識や情報もあふれているが、本というものは一冊もない。

(引用終わり)

5. おわりに

セマンティックウェブが「空っぽの洞窟」になるのかユダヤの知恵の書「タルムード」のようになるのかは、技術がある程度、固まり、知識の集積が始まってからのことになるかもしれない。しかし、セマンティックウェブは、困難で紆余曲折のある道のりかもしれないが、爆発的な普及発展を遂げたWWWの正当な後継者であることには間違いがない。否定的な見方もあるかもしれないが、最後に引用した未来の図書館を使って、人間の創造的な活動がより推進されることに期待したい。

引用・参考文献

- [1] Tim Berners-Lee: "Semantic Web - XML2000," (2000), <http://www.w3.org/2000/Talks/1206-xml2k-tbl/>
- [2] Marja-Riitta Koivunen and Eric Miller: "W3C Semantic Web Activity," Semantic Web Kick-off Seminar in Finland (2001) <http://www.w3.org/2001/12/semweb-fin/w3csw>
- [3] Tim Berners-Lee, James Hendler, and Ora Lassila: 「自分で推論する未来型ウェブ」, 日経サイエンス, 2001年8月号
- [4] Joe Rockmore: Semantic Web for the Military User Intelligence Breakout Session

- <http://www.daml.org/2001/11/swmu2-intel/Breakout-Intel.ppt>
- [5] Marc Abrams, ed., *World Wide Web - Beyond the Basics*, Prentice Hall, 1998. もしくは,
http://ei.cs.vt.edu/~wwwbtb/book/chap1/htx_history.html
- [6] Vannevar Bush: "As We May Think," *The Atlantic* (1945).
<http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>
- [7] <http://www.xanadu.com/xuhistory.html>
- [8] Tim Berners-Lee: "Information Management: A Proposal," (1989).
<http://www.w3.org/History/1989/proposal.html>
- [9] クリフォード・ストール: 「インターネットはからっぽの洞窟」, 草思社(1997).

eCampus

- SGML/XML の教育システムへの応用 -

ネクストソリューション株式会社
ドキュメント・ソリューション事業部/研究開発課
次長役/主任研究員
海田 茂

概要

SGML/XML の教育システムへの応用としての eCampus を紹介する。将来型統合文書システムのアプローチ対象としての eCampus モデルを提示し、論文出版支援システムと教務支援システムとしての応用事例、対話型文書とマルチメディア統合言語の応用事例などを合わせて紹介する。SGML/XML の応用課題を提起する。

キーワード

eCampus, 将来型統合文書システム, 論文出版, 教務支援, 対話型文書(ISMID), 電子運用支援システム(EPSS), 対話型コースウェア, 対話型電子技術マニュアル(IETM), 同期化マルチメディア統合言語(SMIL), 文書スタイル意味指定言語(DSSSL), eラーニング, Web Based Training(WBT)

1. はじめに
インターネット利用者が増加するにつれて、インターネットを活用した新しい教育システムが開発され、サービスが利用されるようになってきた。インターネット上の教育サービスは、いつでも、どこでも、誰でもが各々の掲げる学習目的のために個別に組上げたカリキュラムに従い学習を進めることが可能である。このような教育サービスでは、利用者個々の学習目標を達成するために必要な時間かつ最低限のアクセスによって高い教育効果をあげることが期待されている。

旧来から大規模な教育サービスを提供してきた受験学習、資格試験講座、語学

教育などに加えて、改革の波に洗われている大学教育や専門技術教育の場でも、社会人の生涯学習及び専門技術の再教育を対象にしたサービスの拡大を強く望むようになってきている。また、情報社会への構造転換が進んだことで、企業内教育を含む企業資源の共有化が進められ、それに伴って文書管理の標準化が求められるようになってきている。

このような社会的な要請を受け、将来型統合文書システムの標準化の実証実験の場の一つとしての eCampus が構想され、推進されている。この eCampus を SGML/XML 技術を前提とした教育システムへの応用事例の一つとして紹介する。

2. eCampus の目的

ここで紹介する eCampus は、SGML/XML 技術の利用を前提とした標準文書によって、いつでも、どこでも、誰でもが、必要な時間に、最低限のアクセスで、十分な教育効果をあげ得るために、大学教育及び専門技術教育を提供する教育機関の教務支援、研究機関の論文出版支援、マルチメディア表現の記述及び交換性の確保された対話型のコースウェアなどの開発サイクル及び管理システムの標準化を達成することを目的としている。

eCampus という呼称を広義に解釈すれば、インターネットを活用した学習教育システム全般を指しているといえる。インターネットを利用した教育は、今日では、Web Based Training(WBT)又は、eラーニング[1]と呼ぶ方が一般的で、理解の入り口に立ち易いかもしい。ここでわざわざ、eCampus と呼ぶことの意味は、前提としている技術が SGML/XML であること及び SGML/XML 技術を応用する実証の場を形成することを意味すること、その上での結果を国際標準の開発にフィードバックすることを目的にしているということである。

3. eCampus の要件

eCampus は、次の要件をもつことが望ましい。

- ・ 図書館、博物館、美術館の書籍、資料、所蔵品を横断検索可能であること。
- ・ 特許情報の検索、特許申請に対しての手続きが確立されていること。
- ・ 論文の検索、閲覧、執筆、提出、審査、出版が、支援されていること。
- ・ 学校等の教務、事務に必要な文書情報が共有化され、手続き処理などが支援されていること。
- ・ 教材(コースウェア)の企画、開発、教

材部品の再利用、配布、教材開発/改訂のサイクル、教材利用実績の統計情報、利用者個別又はグループの成績管理が支援されていること。

現時点で、これらの要件のすべてを十分に備えることは容易であるはずがないが、これらの個々の要件については、すでに取り組みが本格化していて、身の回りでも容易に事例を探すことができる。今後、これらの要件を統合し、情報技術を十分に活用するシステム化が社会から要求され、それに伴う標準化が進展するのにもまず間違いのないことである。

4. 将来型統合文書システムのアプローチ対象としての eCampus モデル

先に述べた、eCampus の要件を将来型統合文書システムの目指す標準化に照らしてモデル化したものを次に順に紹介する。

(1) 図書館、博物館、美術館

電子図書館、電子博物館、電子美術館の知的横断検索[2][3]が可能であること。

ここでは、ダブリンコア(Dublin Core)のメタ要素化の取り組み及び Z39.50 のポータルサイトによる横断検索を紹介する。

(2) 特許情報

特許情報の検索、有益な特許情報の利用、新たな特許申請が速やかにかつ確実に行えること。

ここでは、特許庁が推進している特許情報の SGML/XML 利用のシステムなどを紹介する。

(3) 論文出版

論文の提出、審査、編集、出版が速やかにかつ安価に行えること。

ここでは SGML/XML Conference Japan 1998 の際に利用された、文書のワンソース、マルチユースの事例を中心に、SGML/XML の既存技術構成及び今後必要となる技術又は標準について紹介する。

5. 論文出版支援システムとしての eCampus

論文の著作から出版までの論文出版支援システムの特徴の説明及び処理フローの解説など。ごく簡単に述べる。

6. 教務支援システムとしての eCampus

大学教務の文書情報共有化システムの特徴及びその他の解説, 紹介など。ごくごく簡単に述べる。

7. 対話型文書 (運用支援, 教材, マニュアル)

対話型文書の例として, ISMID の EPSS, コースウェア, IETM などを紹介する。何を, どのように対話型文書として記述可能であるかを事例とともに解説する。

8. マルチメディア統合言語

同期化マルチメディア統合言語 SMIL の記述例を動作させながら解説する。会場の要望に応じながら記述の説明か, 機能の説明かどちらかに重点を置いて述べる。

9. eCampus の課題と将来型文書統合の動向

MIT の情報教育カリキュラムと日本の情報教育カリキュラムとの決定的な相違, その帰結の意味するものなどを述べる。

謝辞

この発表をわたしに勧めてくれたシナジーインキュベートの菊田氏に感謝する。

付録

発表の内容及び関連資料は, 次の URL でダウンロード可能とした。

<http://homepage.mac.com/skaida/eCampus/>
解凍キーは, 20021022。

参考文献

- [1] eラーニング白書 2002/2003 年版, 先進学習基盤協議会(ALIC)編著, オーム社, 2002
- [2] 博物館情報の知的横断検索のためのフレームワーク, 山田篤ほか, 画像電子学会 VMA 研究会, 2002 年
- [3] 博物館情報の知的横断検索の試み, 今門政記ほか, 画像電子学会 VMA 研究会, 2002 年

発表者紹介

海田 茂 (かいだ しげる)
教育システム研究所, コースウェアプランナー, などを経て, 現在, ネクストソリューション株式会社の標準化活動を, 総括中。ISO/IEC JTC1/SC34 では, 文書スタイル意味指定言語(DSSSL)のエディタを務めた。最近は特に, 同期化マルチメディア統合言語(SMIL) 2, 対話型マルチメディア文書の交換規格(ISMID)などの規格の検査と実装を推進中。過去に, 電力会社向けのデジタルネットワークアナライザのシミュレーション教材, 石油会社向けの CAI システムなどの開発にもプロジェクトリーダーとして携わっていた。その開発で知り合い結婚した夫人もまた, 遠隔教育及びハイパマルチメディア教材に興味をもつとのこと。

XMLの標準化

小町 祐史/松下電送システム(株)
JTC1/SC34専門委員会委員長
ISO/IEC JTC1/SC34/WG2 Convener

2002-09-25

1. はじめに

XML(Extensible Markup Language)は、ISO/IEC JTC1におけるSGML(Standard Generalized Markup Language)を中核とする文書記述言語に関する地道な標準化活動^{[1], [2]}を基盤として、W3C(World Wide Web Consortium)のエキスパート達によって開発された。

ここではXMLの標準化活動の中でも特に国内における活動を中心として紹介する(規格等の内容には言及しない。)。XMLの国内標準化活動は、W3CにおけるXMLの検討とほぼ並行して行われたため、その検討結果は直ちにW3Cでの議論に組み入れられ、W3Cの勧告(Recommendation)に色濃く反映された。その意味では、XMLの国内標準化活動は、他の多くの国内標準化活動とは幾分趣きを異にする。

そこでその背景と経緯をここに纏めて紹介する。

2. XMLの標準情報(TR)化

W3CにおいてXMLは、次のドラフトを経て今日のExtensible Markup Language (XML) 1.0 (Second Edition)に至っている。

- <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、XSLソースに対してDSSSL^[10]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、直ちに国内のエキスパート達(村田, 小町, 他)によって独立に翻訳されて、それぞれのWeb等で国内に紹介された。

XMLに関する国内エキスパート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、国内エキスパートの参加をお願いし、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^[3]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]、その後、5.に示すとおりW3CのXML関連規定が次々とTR化された。

[注] 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。

2.1 PR(1997-12)のTR化

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)
- 97年12月版(PR)

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97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98年2月に当時の通産省工技院に提出した[4]。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月に発行された。

2.2 REC(1998-02)のTR化

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[5]。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月に発行された。

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。しかしその後に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。

この標準情報(TR)で採用した主な訳語の例を表2.1に示す。

表2.1 TR X 0008:1999における訳語

原語	訳語
well-formed	整形形式
valid	妥当
validity	妥当性
parsed entity	解析対象実体
unparsed entity	解析対象外実体
escape	別扱い
surrogate blocks	サロゲートブロック
ideographic	統合漢字
base character	基底文字
composed form	合成形式
validating	妥当性を検証する
content particle	内容素子
mixed content	混合内容
match	マッチ
character value	文字番号
byte order mark	バイト順マーク
extender	エクステンダ

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、章・節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、次に示すだけの修正(章・節番号の変更なし)を施して、TR原案とした。

- Versions/Editors, Status of this document, Abstractなどをまとめて、まえがき(目次の前)とする。
- "1. Introduction"を、"1. 一般"とする。
- "1.に続く3パラグラフ"に、見出し"1.0 適用範囲"を付加。

- “解説”の追加

3. XMLのJIS化

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が公表されていたため、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[6]。省内の都合によって、02年3月の審議にはかけられず、02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、02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ている。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JSA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った。

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、使用、複製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製物が次を含む場合に制限されている。

1. W3C文書(原規定)へのリンク又はURI。
2. 原著者の既存の著作権表示。それが無い場合は、次の形式の表示。"Copyright © [date-of-document] 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、テキスト表現は許可する。)
3. 存在する場合には、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担当者との打ち合わせもたれた。

4. XML日本語プロファイル

4.1 TR X 0015:1999

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ている。

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[5]、TR X 0015:1999、XML日本語プロファイルとして99年5月に公表された。

4.2 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 0015:1999を英訳し、W3Cに対してNoteとして提案を行った[7]。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 Requestは99年12月に送付され、AC Repからのconfirmを受けて、このTRの英訳は、XML Japanese Profile, W3C Note 22-12-1999 として公表された。

4.3 TR X 0015:2002

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[6]。

これは、TR X 0015:2002, XML日本語プロフィール として02年6月に公表されている。

5. W3Cの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として公表されている。2002年度末までのこれらの活動により、accessibilityとsecurityを除くW3Cの主要な勧告のTRは、概ね完了する予定である。

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(W3C Rec. 等)の問題点は、W3Cにフィードバックされている。

5.1 文書構造等

5.1.1 XML関連

TR X 0023:1999, XML名前空間, Namespaces in XML, 1999-11

TR X 0076:2002, XMLリンク付け言語 XLink 1.0, XML Linking Language (XLink) Version 1.0, (2002-10のMETI審議を予定)

5.1.2 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, 2000-09

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 Specification, 2002-06

TR X 0037:2001, 拡張可能なハイパテキストマーク付け言語 XHTML 1.0 XHTML 1.0: The Extensible HyperText Markup Language, 2001-02

TR X 0051:2001, XHTML基本, XHTML Basic, 2001-12

TR X 0056:2002, XHTMLのモジュール化, Modularization of XHTML, 2002-06

5.1.3 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, 1999-11

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, 1999-09

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, 2002-09

TR X 0060:2002, 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, (2002-10のMETI審議を予定)

TR X 0078:2002, 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, (2002-10のMETI審議を予定)

TR X 00**:2002, 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, (2002-12のMETI審議を予定)

5.1.4 Schema関連

TR X 0054:2002, XMLスキーマ 第0部 基本, XML Schema Part 0: Primer, 2002-06

TR X 0063:2002, XMLスキーマ 第1部 構造, XML Schema Part 1: Structures, 2002-09

TR X 0064:2002, XMLスキーマ 第2部 データ型, XML Schema Part 2: Datatypes, 2002-09

5.1.5 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, 1999-05

TR X 00xx:2002, 同期化マルチメディア統合言語(SMIL 2.0), Synchronized Multimedia Integration Language (SMIL 2.0), (2002-12のMETI審議を予定)

TR X 0077:2002, Xフォーム1.0, XForms 1.0, (2002-10のMETI審議を予定)

5.2 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), 1998-10

TR X 0032:2000, 段階スタイルシート 水準2(CSS2), Cascading Style Sheets, level 2 CSS2 Specification, 2000-09

TR X 0048:2001, XSL変換(XSLT) 1.0, XSL Transformations (XSLT) Version 1.0, 2001-07

JIS/TR X 00**:2001, 拡張可能なスタイルシート言語(XSL) 1.0, Extensible Stylesheet Language (XSL) Version 1.0, (2002-12のMETI審議を予定)

5.3 W3Cの翻訳でない関連TR

次のTRは、INSTACの委員会で独自に開発された関連技術の規定である。TR X 0047は、英訳されてW3CにNoteとして提案されている。

TR X 0059:2002, XSLTライブラリ, XSLT Library, 2002-09

TR X 0047:2001, XMLによる画像参照交換方式, Picture Reference Exchange by XML, 2001-07

6. XMLのISO/IECへの影響

6.1 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Technical Corrigendum 2の編集作業に着手した。SGMLのEditorからのemailによる呼びかけに応じて各国のエキスパートが議論に参加した。

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 8879:1986/Cor.2:1999として発行されている。

ISO 8879:1986/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。

6.2 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, ITUの規格)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 8879:1986/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

7. スキーマ言語

7.1 背景

文書型定義(DTD)はパーサで構文解析できないため、DTDを処理するツールを作りやすく、しかも基本的なデータ型を扱うことができない。そこで、この問題を解決できる言語が強く望まれていた。既にいくつかの提案が出されていたが、それらの共通機能を実現する標準的な規定として、XML正規言語記述(Regular Language Description for XML, RELAX)が考案された。

この活動は、村田らによる小グループのものであったが、その成果を国際的な規定に位置付けるために、まず国内の標準情報(TR)としてオーソライズしてから、それをFast-track手続きによってISO/IEC JTC1に提出するという戦略が採用された。TR化の活動は、日本規格協会のINSTACにおいて99年から開始された。

7.2 RELAXコア

最初の規定は、単一の名前空間だけを扱うRELAXコアであり、TR X 0029:2000, XML正規言語記述 RELAX コア として2000年5月に公表された。このTRは、XML文書の正規集合を記述するための言語を規定し、その文法は、XML文書で使用可能なタグ名、属性名、許される文字列、それらの可能な組合せを扱う。

TR X 0029の英語版は、通産省工業技術院(当時)からISO中央事務局にFast-track提案され、ISO中央事務局がJTC1に対して01年5月を期限とするDTR投票(DTR 22250-1)を指示した。

この投票に触発されて、それまで進捗がはかばかしくなかったW3CのXML Schemaの活動に大きな進展があり、急遽、勧告が公表されるに至った。W3CのXML Schemaの公表を著しく加速化したことは、このDTR投票の副作用であったかも知れない。同時期に、J. ClarkがTREG(Tree Regular Expressions for XML)を発表したため、RELAXとTREGとの整合作業が進められて、その成果はRELAX NG(RELAX New Generation)としてOASISから発表された。

DTRは賛成多数で可決されたが、各国からの投票コメントに対処するため、01年8月にBallot resolution会議を新潟の国際大学で開催し、Disposition paperを作成すると共に、最終テキストの準備を行った。最終テキストは01年9月に提出され、ISO/IEC TR 22250-1として02年2月に発行された[8]。

7.3 RELAX名前空間

RELAX名前空間による記述は、幾つものRELAXコアによる記述を組み合わせることによって、複数の名前空間を扱う。一部の名前空間について、RELAXコア以外の言語によって記述することもできる。

この規定は、TR X 0044:2001, XML正規言語記述 RELAX 名前空間 として、01年7月に公表された。その英語版は、情報規格調査会からJTC1にFast-track提案され、02年5月を期限とするDTR投票が開始された(JTC1 N6616; N6643)[9]。このDTR 22250-2は、賛成多数を得て可決された。Ballot resolutionは02年12月を予定している。

7.4 DSDL(文書スキーマ定義言語)への組込み

日本提案のRELAXに対する対抗策としてUKは、文書スキーマ定義言語(DSDL)のNP提案をSC34に提出してきた。NP提案

は承認されたが、提案の後にUKから提出された作業ドラフトは、J. Clark, 村田らのエキスパートのレビューに耐えるものでなかったため、01年12月のSC34会議において、DSDLをマルチパート化して、XSL Schema, RELAX NGなどの既存の複数のスキーマ言語を含む規格とする方針が承認された。

DSDLのパート2としてRELAX NGが位置付けられ、OASISの仕様書の表紙だけをISOの体裁にしたCDテキストが、02年3月を期限とする投票にかけられた。これは反対なしで承認され、02年5月のSC34会議においてCD投票コメントの反映が行われて、改訂テキストが配布された。

02年5月のSC34会議では、DSDLのパート構成がさらに見直されて次のリストのとおりとなり、RELAX名前空間がDSDLの一部となることがほぼ決まった。この作業が早く進めば、ISO/IEC TR 22250-2の発行に代えて、DSDL(ISO/IEC 19757)のそのパートの発行を急ぐ方が適切かもしれない。

- Part 0: Overview, G. Ken Holman (Canada)
- Part 1: Interoperability framework, Eric van der Vlist (ISUG)
- Part 2: Grammar-based validation - RELAX-NG, James Clark (UK) and Makoto Murata (Japan)
- Part 3: Rule-based validation - Schematron, Rick Jelliffe (Australia)
- Part 4: Selection of validation candidates, Makoto Murata (Japan)
- Part 5: Datatypes, Martin Bryan (UK)
- Part 6: Path-based integrity constraints, James Clark (UK)
- Part 7: Character repertoire validation, Diederik Gerth van Wijk (Netherlands)
- Part 8: Declarative document manipulation, G. Ken Holman (Canada)
- Part 9: Datatype- and namespace-aware DTDs

8. むすび

国際標準化と国内標準化との間には、英語-日本語変換という大きな障壁が存在している。この障壁にもかかわらず、XML関連技術については国際と国内とがほぼ同時進行で標準化が推移し、互いに影響を及ぼし合い、相互に寄与を与えてきた。

これらの経緯をここで整理したことが、今後の標準化活動を推進する際の一つのReferenceとなり得るであろうことを期待する。

寝食の時間を割き、国際会議参加等にも休暇と私費を遣うことも可としてこれらの標準化活動に参加されたXML標準化のメンバに感謝する。

文献

- [1] 小町: SGML/XMLの入門から応用まで, 1. SGML/XML開発の経緯とその関連規格概要, Computer Today, No.83, 1998-01
- [2] 小町: SGML/XMLの入門から応用まで, 7. SGML関連規格 - SGML Support Facility -, Computer Today, No.86, 1998-07
- [3] 通商産業省: 標準情報(TR)制度実施要綱, 1996-08-01
- [4] 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
- [5] 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
- [6] 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
- [7] 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
- [8] ISO/IEC TR 22250-1, Information technology -- Document description and processing languages -- Regular Language Description for XML (RELAX) -- Part 1: RELAX Core, 2002-02
- [9] ISO/IEC DTR 22250-2, Information technology -- Document description and processing languages -- Regular Language Description for XML (RELAX) -- Part 2: RELAX Namespace, 2002-01
- [10] JIS X 4153, 文書スタイル意味指定言語(DSSSL), 1998-03

トピックマップ

株式会社シナジー・インキュベート
内藤 求 (motom@synergy.co.jp)

Ontopia AS
Lars Marius Garshol (larsga@garshol.priv.no)

1. はじめに

Web の発展は驚異的であり、この 10 年の間に、我々の情報へのアクセス方法を決定的に変化させたといっても過言でないように思える。今や、Web なしでの日々の業務は考えられないほどである。ただ、現在、情報の発見方法の主流は、文字列の一致によるものであり、ときには、山のような情報の中から、ほんとうに必要な情報を探し出すのに大変な労力を必要とすることがある。これに対処するため、情報にメタデータを付与したり、検索エンジンを工夫したりと、様々な試みがなされている。

そのうちのひとつで、最近の XML カンファレンスで最も発表件数の多いテーマの一つであり、2000 年には ISO の標準にもなっているトピックマップについて、その概要と応用例について述べる。トピックマップは、利用者の持つ概念体系に合わせて情報を分類、整理する手法であり、情報洪水、情報過多に対する対応方法の一つとして利用することができる。

2. トピックマップ

2. 1 トピックマップとは

トピックマップは、ISO/IEC JTC1 SC34 WG3 (以降、SC34 WG3 とよぶ) で策定された標準 (ISO/IEC 13250 Topic Maps) である。以降 ISO 13250 と呼ぶ。現在も、SC34 WG3 と非営利のコソソーシアムである OASIS (Organization for the Advancement of Structured Information Standards) において、ISO 13250 の改定と、関連規格の策定について検討が続けられている。

トピックマップは、情報リソースとは独立した上位層 (メタレイア) に位置付けられ、トピックマップの作成者が認識している概念体系に則って、索引を作成し、情報を分類、整理し、管理、検索、ナビゲートを可能にするための新しいパラダイムである。トピックマップは、情報リソースが持つ主題 (概念) と主題間の関係を、情報リソースとは独立にコンピュータ上でモデル化 (reify:具体化) し、主題に関連する情報リソースに対しては、リンクを張ることにより関係を明示する手法である。

トピックマップを表現するためのシンタックスは、マークアップ言語の形式をとっており、現在、SGML/HyTime に則った HyTM シンタックスと、XML に則った XTM シンタックスの 2 種類があり、その構造は、DTD によって定義されている。現在、よく使用されている XTM (XML Topic Maps) の主な構成要素を以下に示す。

(1) 三大要素

- ・ topic (トピック) : 人間が認識する具体的または抽象的な主題/概念を表す要素
- ・ association (アソシエーション) : トピック間の関連を表す要素
- ・ occurrence (オカレンス) : 実際の情報リソース

(2) 3種類のポインタ

- ・ topicRef: トピックへのポインタ
- ・ subjectIndicatorRef: 主題の所在を示す情報リソースへのポインタ
- ・ resourceRef: 主題を構成する情報リソースへのポインタ

(3) mergeMap 要素

他のトピックマップの併合を指定するための要素

(4) subjectIdentity 要素

subjectIdentity 要素の親要素である Topic 要素によって具体化 (reify) される主題を指定するための要素

(5) baseName 要素

トピック名を指定するための要素

(6) scope 要素

トピック特質の割り当ての有効範囲を指定するための要素

トピック特質は、baseName、オカレンス、アソシエーションの集まり

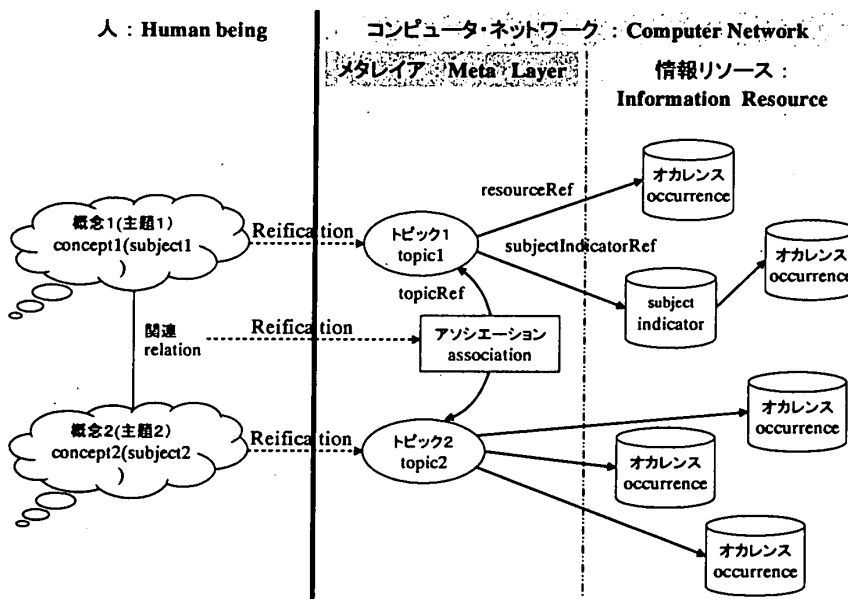


図1. 主題、トピック、アソシエーション及びオカレンス

図1. は、主題、トピック、アソシエーション及びオカレンス（情報リソース）の関係を表現している。トピックマップでは、概念として認識できるすべてのものをトピックとすることができる。アソシエーションについては、XTM (XML シンタクスのトピックマップ) では、以下の2種類が標準の中で定義されている。

- ・ class-instance 関係 (トピック間のクラスとインスタンス関係を表現するアソシエーションのクラス)

- ・ superclass-subclass 関係 (トピック間の上位クラスと下位クラスとの関係を表現するアソシエーションのクラス)

それ以外、アソシエーションは、例えば以下のような、概念 (主題) 間の任意の関係を自由に定義することができる。

- ・ 等価関係 (同義、類似・対比)
- ・ 階層関係 (上位-下位、全体-部分)
- ・ 連想関係 (並列、用途・環境、因果)
- ・ その他任意の関係

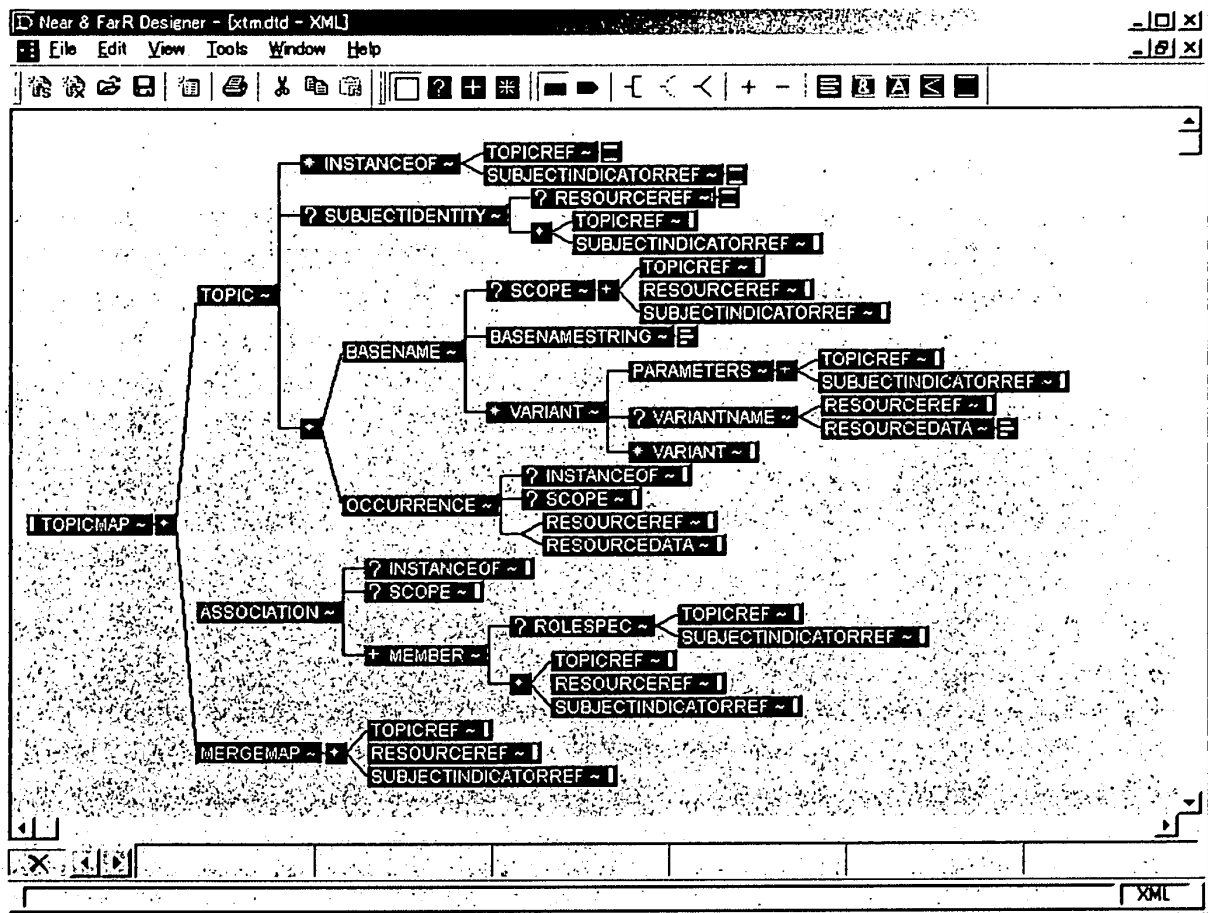


図 2. XTM DTD の構造

図 2. は、XTM DTD の構造を表している。XTM DTD は、19 の要素からなる非常にシンプルなものである。木構造の最も末端、すなわち、葉要素のほとんどは、上記の 3 種類のポインタである。このように、主題、主題間、情報リソースへのリンク付けのための豊富な手段をもっている。

例えば、トピック”論文”及び”著者”、アソシエーション”論文の著者”を XTM で表現すると以下のようになる。

```

<topicMap xmlns="http://www.topicmaps.org/xtm/1.0/"
  xmlns:xlink="http://www.w3.org/1999/xlink">
  <topic id="authorship">
    <baseName>
      <baseNameString>論文の著者</baseNameString>
    </baseName>
  </topic>
  <topic id="author">
    <baseName>
      <baseNameString>著者</baseNameString>
    </baseName>
  </topic>
  <topic id="article">
    <baseName>
      <baseNameString>論文</baseNameString>
    </baseName>
  </topic>
  <association>
    <instanceOf>
      <topicRef xlink:href="#authorship"/>
    </instanceOf>
    <member>
      <roleSpec>
        <topicRef xlink:href="#author"/>
      </roleSpec>
    </member>
    <member>
      <roleSpec>
        <topicRef xlink:href="#article"/>
      </roleSpec>
    </member>
  </association>
</topicMap>

```

2. 2 トピックマップの歴史

現在トピックマップコミュニティで思想的リーダーの一人である Steven R. Newcomb 氏によると、1991年にUNIXのシステムベンダーによって Davenport グループが設立され、そこでトピックマップの作業が開始された、ということである。その作業の中で、別々に保守され常に進化する技術文書を集めてシステムマニュアルにするためのマスターインデックスの作成が大きな問題になった。その解決策としての最初の試みが、SOFABED (Standard Open Formal Architecture for Browseable Electronic Documents) であった。

1993年には、GCARI (Graphic Communications Association Research Institute、現在の IDEAlliance) 主催による CAPH (Conventions for the Application of HyTime) グループが結成され、1992年にISOの標準になった HyTime を使用して SOFABED を Topic Maps として洗練させた。

Topic Maps は、1995年に、ISO/IEC JTC1 SC18/WG8において、“new work item”となり、ISOの組織再編を経て ISO/JTC1 SC34/WG3の担当になり、最終的には2000年に、ISO/IEC 13250:2000として国際標準になった。

ISO 13250 が出版されるとすぐに、XTM 活動が開始され、独立の組織である TopicMaps.Org が結成され、そこから、XTM 1.0 規格が2001年3月に出版された。XTM は、2001年のうちに、ISO 13250 に組み入れられることが決定し、2002年には、ISO 13250 の第二版が出版された。現在も、引き続き、ISO 13250 の改善、および関連標準の検討が進められているが、それについては、項番 2. 4 トピックマップの標準化動向において述べる。

2. 3 活動組織

(1) ISO/IEC JTC1 SC34

トピックマップの標準策定作業は、ISO/IEC JTC1 SC34（文書の処理と記述の言語）の Working Group 3（情報の関連付けを担当）で行われている。国際会議は、アメリカやヨーロッパでの XML カンファレンスなどと併設で、大体年 4 回開催されている。ISO の国際会議は、W3C、IETF や OASIS と異なり参加できるのは、National Body のメンバだけである。WG3 の convener は、ノルウェーの Steve Pepper 氏であり、投票権のある参加国は、アメリカ、フランス、イギリス、ノルウェー、日本である。その他に、ドイツはこの WG では投票権を持っていないが、積極的に参加している。

その他、SC34 では、WG1 で、マーク付け言語関係の標準化を、WG2 で、情報表示関係の標準化が進められている。WG1 の convener は、Charles Coldfarb 氏であり、Martin Bryan 氏や James Clark 氏などのビッグネームもメンバである。WG2 の convener は、小町祐史であり、日本中心に活動が行われている。

（2）OASIS

OASIS のトピックマップ関連の活動としては、Published Subject に関する以下の 3 つの委員会がある。Published Subject は、多くの人が共通に使用できると思われる主題（トピック）を定義し、Web 上で永続的に公開していこうというもので、それにより、トピックマップのみならず、他の知識表現言語も含めて、知識の共有／交換を容易にしようとするものである。国際会議は、SC34 WG3 の会議と同じように、アメリカやヨーロッパでの XML カンファレンスなどと併設で、大体年 4 回開催されている。メンバの多くが、SC34 WG3 のメンバでもある。

・ Topic Maps Published Subjects Technical Committee

この委員会の Chair は、フランスの Bernard Vatant 氏が務める。この TC の目的は、Published Subject についての定義、管理、使用方法について、要求、勧告、及び最優良事例を指定することにより、その利用を促進することである。

・ Vocabulary for XML Standards and Technologies TC (XMLvoc TC)

この委員会の Chair は、ドイツの H. Holger Rath 氏が務める。この TC の目的は、XML 標準及び技術の領域についての語彙を定義することである。それらの語彙は、トピックの参照セット、トピックの型、アソシエーションの型などになる。それにより、XML、関連する標準、及び XML コミュニティに関する情報の発見を容易にする。

また、それらの語彙を、OASIS Topic Maps Published Subjects Technical Committee の勧告に則って、Published Subjects として定義する。

・ Topic Maps Published Subjects for Geography and Languages (GeoLang) TC

この委員会の Chair は、ノルウェーの Lars Marius Garshol 氏が務める。この TC の目的は、Published Subjects TC によって作成される Published Subjects の案内に従って、言語、国、地域の Published Subjects セットを定義することである。

言語、国、地域は、広い範囲のトピックマップでしばしば使用される主題である。最大限の再利用性、交換可能性、併合可能性を促進するために、これらの領域をカバーする Published Subjects の標準セットが要求されている。

2. 4 トピックマップの標準化動向

（1）ISO

今後のトピックマップの標準活動については、策定または検討中の標準および標準の位置付け、関連等、2001 年 12 月に米国フロリダ州オーランドで開かれた SC34 WG3 の会議で作成されたロードマップで明確に示され、それに沿って作業が進められている。

現在は、2 種類のデータモデル、すなわち、標準応用モデル：Standard Application Model (SAM) と、参照モデル：Reference Model (RM) が優先的に検討されている。2002 年 5 月のスペインのバルセロナの会議では、Reference Model について、新しいドラフトの提案があり、Standard

Application Model のドラフトについて、前回からの審議が継続して行われた。また、2002年8月の Montreal の会議では、再度、Standard Application Model の審議が継続して行われ、今後も継続して審議される予定である。

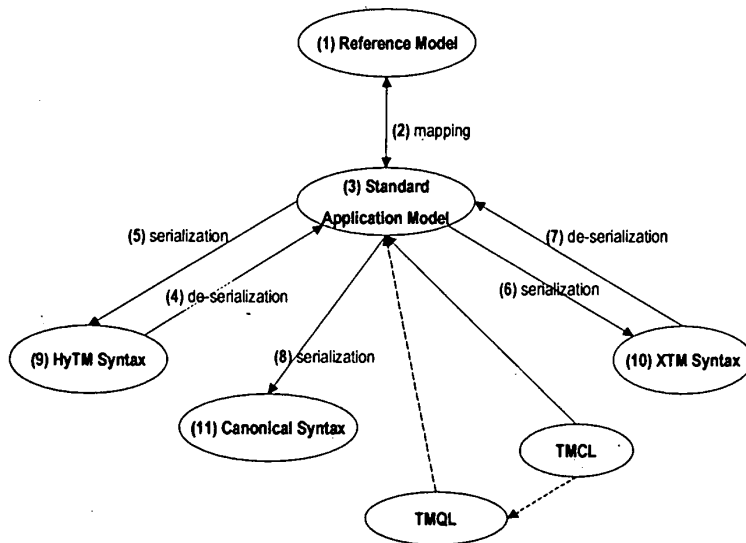


図 3. トピックマップ標準策定ロードマップ

将来、ISO 13250 は、マルチパート標準として、これらを束ねたものになる予定である。例えば、以下のような構成になることが予想される。

- ・ Part 0 : この規格の案内
- ・ Part X : 標準応用モデル (Standard Application Model)
- ・ Part X : 参照モデル (Reference Model)
- ・ Part X : XML Topic Maps (XTM) シンタックス
- ・ Part X : HyTime Topic Maps (HyTM) シンタックス
- ・ Part X : トピックマップの規範化 (Canonicalization of topic maps)

標準応用モデルの標準化の完了後に、以下の関連規格の検討が予定させている。

- ・ ISO/IEC 18048 : Topic Maps Query Language(TMQL) : トピックマップ用の検索言語
- ・ ISO/IEC 19756 : Topic Maps Constraint Language(TMCL) : トピックマップ用のスキーマ言語

日本では、ISO 13250 の翻訳版、及び、XTM 1.0 の翻訳版が、それぞれ JIS X 4157、及び、TR X 0057 として経済産業省の技術委員会で承認された。また、現在 ISO で策定中の標準も、適宜、JIS 化を検討していく予定になっている。

(2) OASIS

OASIS では、項番 2. 3 に記述したように、3つの委員会で、Published Subject に関する標準化およびその普及が進められている。

3. トピックマップの適用例

トピックマップの適用例として、研究者と論文のトピックマップを考えてみる。

(1) “研究者”トピックマップ

“研究者”トピックマップとしては、例えば、以下のトピック、アソシエーションが考えられる。

- ・トピック：研究者、著作論文、所属組織、役職、専門分野、学歴、経歴
- ・アソシエーション：研究者－著作論文、研究者－所属組織、研究者－役職、・・・

(2) “論文”トピックマップ

“論文”トピックマップとしては、例えば、以下のトピック、アソシエーションが考えられる。

- ・トピック：論文、雑誌、出版社、タイトル、ジャンル、キーワード
- ・アソシエーション：論文－雑誌、論文－出版社、雑誌－出版社、・・・

“研究者”トピックマップと“論文”トピックマップのオントロジを図4.、図5. に示す。

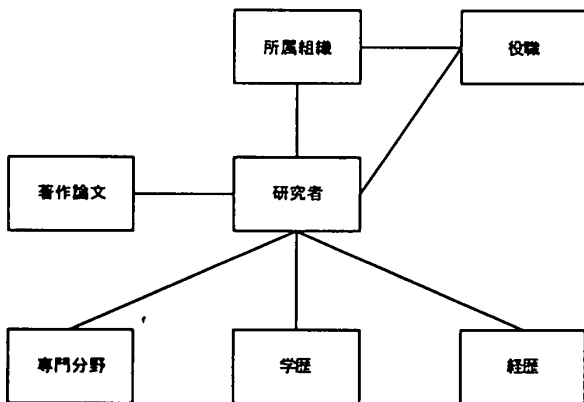


図4. “研究者”オントロジ

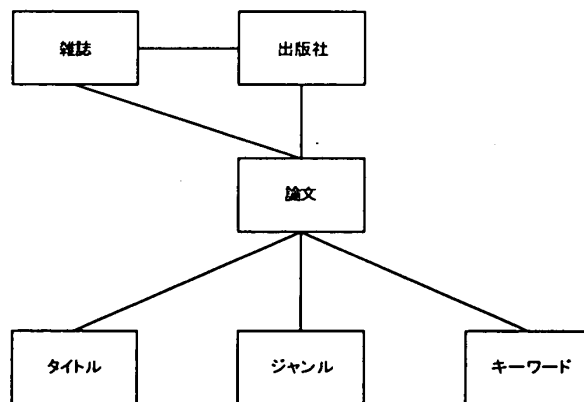


図5. ”論文”オントロジ

(3) “研究者”トピックマップと“論文”トピックマップの併合

別々に作成した“研究者”トピックマップと“論文”トピックマップを併合すると、“研究者論文”トピックマップが作成される。

“研究者論文”トピックマップのオントロジを図6に示す。

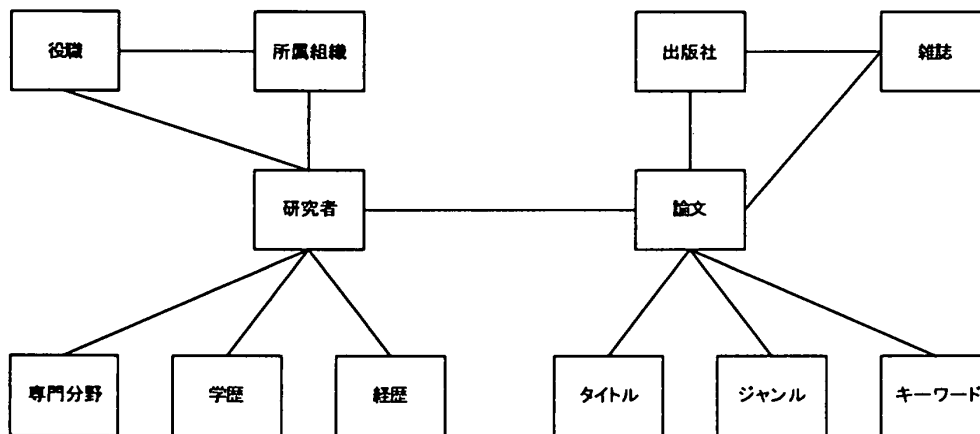


図6. “研究者論文”のオントロジ

この併合により、研究者、論文、それぞれ別々に分類・整理した情報を、元の情報を失うことなく統合することができる。

4. むすび

トピックマップは、非常にシンプルである。トピック、アソシエーションを自由に定義することができ、情報リソースと関連付けることができる。それらにより、任意の問題領域でのオントロジを容易に作成することができる。トピックマップ作成者自身による視点の設定、それに基づく情報の分類、整理が可能になる。また、併合機能により、別々に作成されたトピックマップを容易に併合することができる。そのときの、世界共通の結合点として、Published Subject を使用することができる。いろいろなアプリケーションにトピックマップを適用することで、直観的、ユーザフレンドリなユーザインターフェースの構築が可能になり、その付加価値を高めることができる。

オントロジ及びその構成要素であるトピック、アソシエーションに対する計算機処理も容易で、データ処理から情報処理へと進化してきた計算機利用を、知識処理へと進化させる可能性を持った方法(道具)の一つとして多様な利用方法が考えられる。

<Lars Marius Garshol 氏のプロフィール>

Lars Marius Garshol 氏は、トピックマップソフトウェアのリーディングベンダーである Ontopia 社の開発マネージャです。彼は、オープンソースの XML パーサの開発者であり、フリーXML ツール Web サイトの管理者として、長年にわたり XML コミュニティで積極的に活動しています。また、Prentice-Hall 社から出版された"Definitive XML application development"の著者でもあります。現在は、ISO のトピックマップの標準化活動において、SAM (Standard Application Model) 及び TMQL (Topic Map Constraint Language) の編集者の一人であり、OASIS のトピックマップ関連の委員会の1つである "GeoLang TC" の委員長でもあります。

参照情報と文献

- [1] Steve Pepper, "The TAO of Topic Maps." Conference Paper at ACM CIKM'99 .
- [2] International Standard Organization, "ISO/IEC 13250 Information Technology - SGML Applications - Topic Maps", ISO/IEC 13250:2000.
- [3] John F. Sowa, "Knowledge Representation Logical, Philosophical, and Computational Foundations", Brooks/Cole, ISBN: 0-534-94965-7.
- [4] TopicMaps.org, "XML Topic Maps (XTM) 1.0" .
- [5] Topicmaps.net, "Topicmaps.net's Processing Model for XTM 1.0, Version 1.0.1" .
- [6] Steve Pepper, "Modelling Topic Maps", Tutorial Handout, XML 2001, Orlando, Florida .
- [7] ISO/IEC JTC1 SC34, "Topic maps, roadmap for further work", ISO/IEC JTC1 SC34 N0278 .
- [8] Michel Biezunski, Steven R. Newcomb, "The Topic Maps Technical Workshop", KT 2001, Austin, Texas .
- [9] Martin S Lacher, Stefan Decker, "On the Integration of Topic Maps and RDF Data", Proceedings of SWWS' 01, pp 331-344, 2001 .
- [10] Graham Moore, "RDF and TopicMaps An Exercise in Convergence", XML Europe 2001, Berlin .
- [11] Mari Nagase, Motomu Naito, "Application and Evaluation of Topic Maps for the Cultural Resource Data --- Experiment with the Graphic Data of "Genjji Monogatari (the Tale of Genji)" --- ", KT 2002, Seattle, 2002.
- [12] ISO/IEC JTC1 SC34, "A Draft Reference Model for ISO 13250 Topic Maps", ISO/IEC JTC1 SC34 N0298 Rev.1
- [13] ISO/IEC JTC1 SC34, "The Standard Applicatoin Model for Topic Maps", ISO/IEC JTC1 SC34 N0299
- [14] Kal Ahmed, "Topic Maps - A practical Introduction With Case Studies", XML Europe 2002, Barcelona, Spain .
- [15] Jack Park, Sam Hunting "XML Topic Maps - creating and Using Topic Maps for the Web", Addison-Wesley, ISBN: 0-201-74960-2.

XML署名

NEC インターネット基盤開発本部
マネージャー 杉山高弘

政府主導によるスーパー電子政府では、申請届出手続きや政府調達など行政手続きの電子化を進めている。また、インターネットにおける電子商取引が普及してきた。このとき、取引相手の本人確認による電子商取引の安全性の確保、および、省庁間で交換される電子化公文書の改竄検出を保証する電子署名技術が重要になってきている。

本稿では、W3C XML署名標準規約に基づきXML署名技術を解説し、XML署名基盤ソフト技術開発への取り組みを紹介する。また、民間を含めた電子商取引への適用システム/サービスイメージ等をご紹介します。

1. はじめに

政府は、2003年度までに行政の効率化や国民負担の軽減を目標に行政手続きを電子化する電子政府の基盤を構築することを目指している。

電子政府の構築は、デジタル経済・社会の一つのモデルであり、その中で実施される情報セキュリティ確保のための対策もまた、広く民間の範となるものとなり、それによって、我が国のネットワーク全体の安全性・信頼性を高め、更に、具体的に進みつつある同様な取り組みと連携していくことにより国際的な貢献につながることも期待される。このため、デジタル経済の1つのモデルである電子政府の構築に向けて、積極的な貢献を行っていく必要がある。

2. 背景

XMLは、電子政府や電子商取引など多くのシステムにおいて、電子文書を交換するための標準フォーマットとして使用されているのが現状である。今後も、XMLがデファクト標準として普及していくことが予想される。一方、電子商取引システムなどで使用される電子署名を付与した文書の標準フォーマットについても、XMLの標準化団体であるW3Cが、IETFと共同で、XML署名と呼ばれる標準フォーマットを検討してきており、2002年2月に正式勧告された。このXML署名は、デジタル経済の進展とXMLの普及に伴い、電子署名付文書のデファクト標準

となることが極めて有望とされている。また、政府が推進する電子政府アプリケーションにおいても、電子署名として国際標準のXML署名フォーマットが採用されることが期待されている。

電子署名付文書フォーマットが標準化されていない場合、電子署名付文書の構造は各アプリケーション固有の構造となり、電子署名ライブラリも個別に作られることになり、電子商取引が普及する際に大きな問題となる。これを解決するために、以下を推進する必要がある。

- ・電子政府における公文書交換への電子署名付文書構造の標準化
- ・電子申請、電子商取引における利用者の利便性を考慮した電子署名付文書構造の標準化、および、電子署名を処理するライブラリの標準化

また、電子政府においては、今後の申請業務の多様化と民間電子商取引との連携、および、海外政府との公文書交換等の適用場面の広がりを見ると、マルチベンダ環境での電子署名の使用は避けられない。マルチベンダ間での、XML署名の真の相互運用性を確保することが重要となる。

3. 目的

本講演のベースとなった技術開発は、電子政府における情報セキュリティ確保の一環として、以下を目指した。

1) 標準準拠のXML署名基盤ソフトの開発

電子文書の交換フォーマットにXMLがデファクト標準となっている現状に鑑み、XMLで整合のとれた世界標準の電子署名フォーマットが望まれている。本技術開発では、W3C XML-Signature標準に準拠した汎用性のあるコンポーネントとしてAPIを策定し標準電子署名ライブラリを開発する。本ライブラリは、署名付与機能、検証機能、XML署名生成機能等から成る。本成果は、政府の承諾を得てW3CやIETFへリファレンスソフトとして公開し、デファクト標準化を促進する。

2) XML署名における署名付与・検証時の相互運用性確保

電子政府においては、今後の申請業務の多様化と民間電子商取引との連携、および、海外政府との公文書交換等の適用場面の広がりを見ると、マルチベンダ環境での電子署名の使用は避けられない。現状では、XML処理ツール等の非互換により各社の署名付与、検証システム間の相互運用性が保証されない場合がある。これを解決する手段の研究、および、相互運用性を確保するためのXML署名ガイドラインをまとめる。

4. XML署名基盤ソフトの機能説明

W3C XML Signature Syntax and Processing標準規約に基づきXML署名基盤ライブラリの各機能を以下で説明し、機能間の関係を図1に示す。

① XML署名文書生成機能

XML署名文書の生成を行うAPI。アプリケーションに対するXML署名機能の基本的なインタフェースも提供する。アプリケーションはこの機能を通じて署名対象の文書を指定して、署名計算に必要なパラメータを設定し署名を作成する。

- W3C XML-Signature Syntax

and Processing仕様に準拠したXML署名文書の生成

- 電子文書の一部を指定する部分署名や、既に署名付与された電子文書に自分の署名を付与する多重署名（証明書更新等にも有効）を容易に表現、処理可能な機能をサポート
- 署名対象文書として、XML文書だけでなく、MS-Word等のバイナリの文書もサポート
- 署名付与日時、署名を検証する公開鍵証明書情報などXML署名仕様でオプションであるがよく利用される要素をサポート
- 証明書情報など署名関連要素でユーザ定義可能な要素の操作のためのユーザ拡張用インタフェースの提供

② XML文書正規化機能

XML文書に署名計算を行う際に、表現形式は異なるが意味的に同一のものとして扱えるように正規化する機能。以下のアルゴリズムを標準でサポートする。

- Canonical XML
(<http://www.w3.org/TR/2000/WD-xml-c14n-20000119>)

また、XML署名機能の相互運用性試験や評価を通じてXMLパーサの実装の非互換性等に起因する相互運用上の問題点を解決するために、任意の正規化機能をアドオン可能とする。

③ XMLトランスフォーム機能

署名対象物に対して署名計算を行う前に、ユーザが指定する前処理として実行するXML文書変換機能。XML電子署名仕様に規定されている以下の機能を実装する。

- base64 エンコード、デコード
- XML フィルタリング機能

(XPath フィルタリング)
署名要素内に署名対象データを含む場合(enveloped signature)、署名値など署名計算に当たって除外すべき要素を指定することなどに利用
XML 変換エンジン(XSLT 等)により XML 文書変換処理記述を設定するためのインタフェース

④ 署名付与機能

XML 署名文書に対して署名付与する機能。署名対象電子文書に対して XML 文書正規化機能、XML 文書トランスフォーム機能を用いて得られた電子文書のハッシュ値を求める。ハッシュ値を署名付与依頼者の秘密鍵で暗号化することによって署名値を求める。XML 署名文書生成機能を用いて生成された XML 署名

文書に対して、上記ハッシュ値、署名値を格納する。

⑤ 署名検証機能

XML 署名文書に対して署名検証する機能。署名対象電子文書に対して XML 文書正規化機能、XML 文書トランスフォーム機能を用いて得られた電子文書のハッシュ値を求める。署名データを署名者の公開鍵で複合化する。上記のハッシュ値と複合化した値が一致するかどうかを検証し、その結果を通知する。

⑥ セキュリティ機能設定インタフェース

ユーザが XML 署名の計算を行う際に、セキュリティ機能に各種の設定を行うためのインタフェースを提供する。

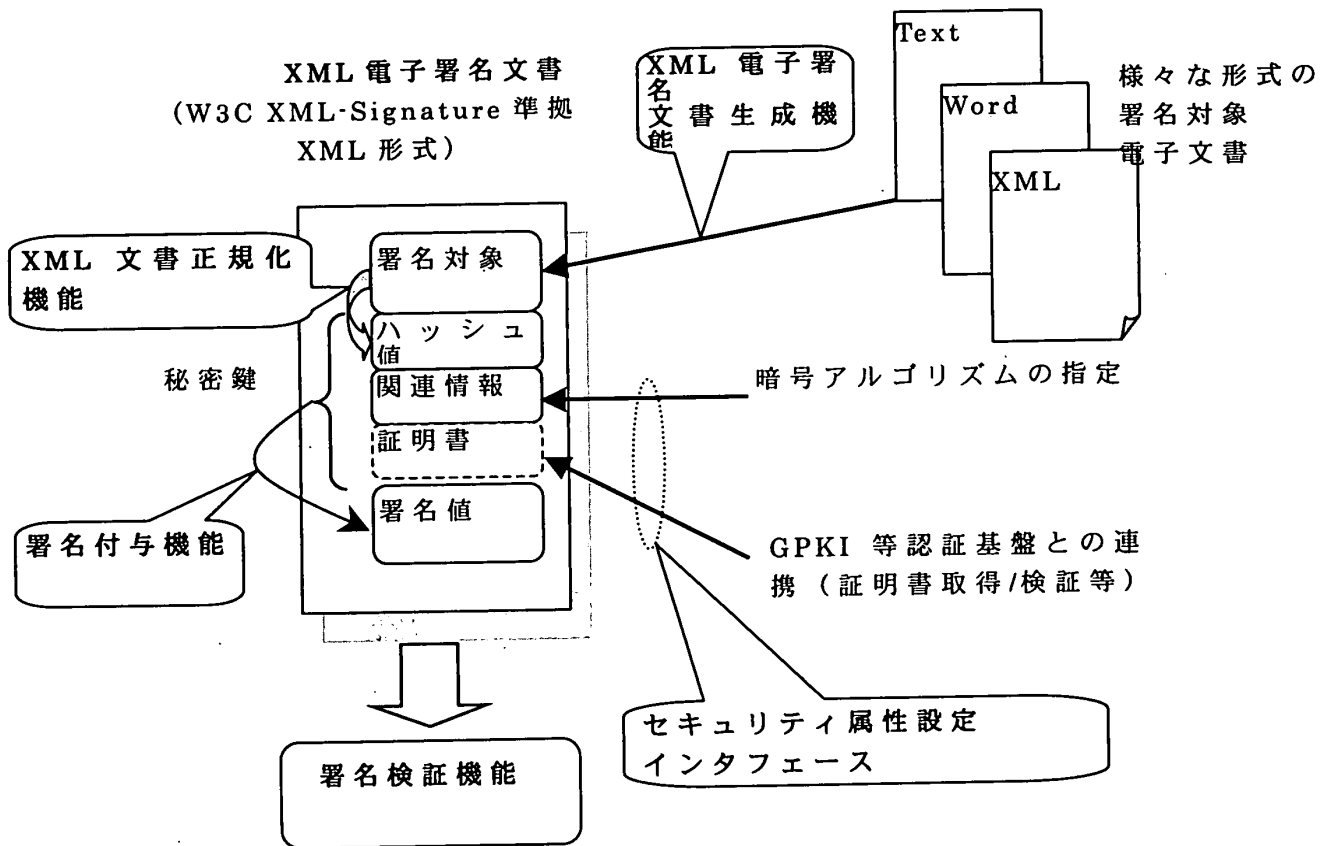


図 1. 署名文書と各機能の関係図

- XML 署名文書内で指定されている公開鍵証明書取得し利用する機能
 - CRL を取得し利用する機能
 - アプリケーション固有の署名アルゴリズムやハッシュ計算機能などを新たに付加する機能
- またセキュリティ機能としては、署名アルゴリズムとして RSA-SHA1、DSA-SHA1 を、ハッシュアルゴリズムは SHA-1 を標準でサポートする。

5. XML 署名文書の処理プロセス

XML 署名文書の生成処理と検証処理のプロセスを図 2 に示す。処理内容は以下の通りである。

[XML 署名文書の生成処理]

- ① 署名対象文書、XML 署名文書に対して実行する変換処理に関する情報、検証時に用いる公開鍵に関する情報を、それぞれ Transforms 要素、KeyInfo 要素に設定する。
- ② 変換処理に関する情報を元に、署名対象文書の変換処理を行う。
- ③ ダイジェスト値計算で用いるハッシュ関数のアルゴリズムを、DigestMethod 要素に設定する。さらに、変換した署名対象文書をハッシュ関数で圧縮してダイジェスト値を取得し、DigestValue 要素に設定する。
- ④ 得られたダイジェスト値を秘密鍵を用いて暗号化して署名値を生成し、SignatureValue 要素に設定する。

以上のプロセスをもとに、XML 署名文書を生成する。

[XML 署名文書の検証処理]

- ① XML 署名文書の SignatureMethod 要素と KeyInfo 要素から、署名値と公開鍵に関する情報を取得する。
- ② 公開鍵値で署名値を復号し、ダイジェスト値を取得する。
- ③ XML 署名文書の Transforms 要素から、変換処理に関する情報を取得して変換処理を実行し、署名対象文書を取得する。
- ④ 得られた署名対象文書を、XML 署名文書の DigestMethod 要素から取得したハッシュ関数で圧縮し、ダイジェスト値を取得する。
- ⑤ ② で得られたダイジェスト値と、④ で得られたダイジェスト値を比較して、一致すれば検証成功、一致しなければ検証失敗とする。

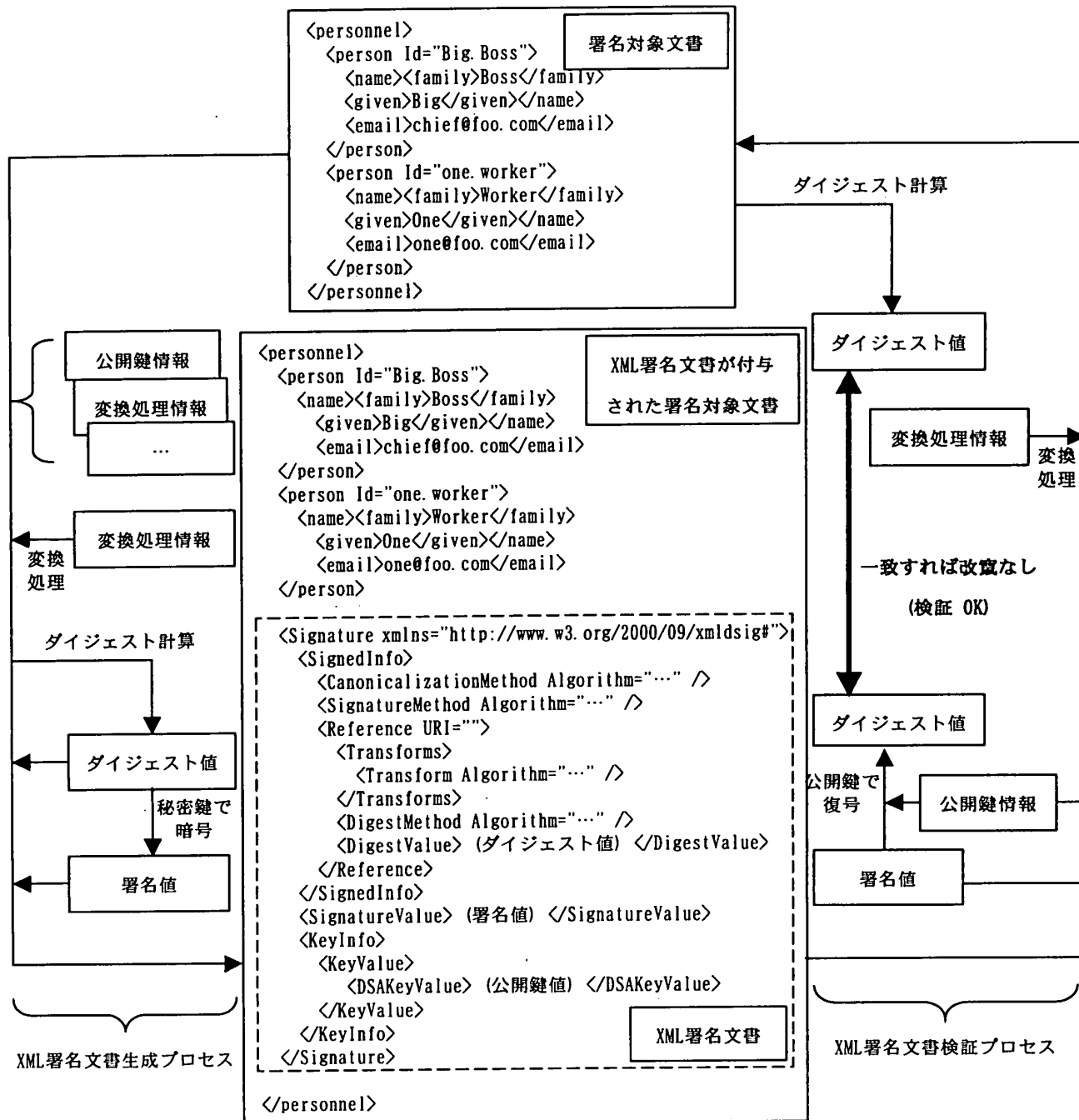


図2. XML 署名文書の生成処理と検証処理のプロセス

6. XML 署名技術の応用例

本章では、電子政府を構成するシステムのうち重要な位置を占める電子申請システムや、民間の電子商取引システム等への XML 署名技術の応用について具体的に示す。

[電子申請システム]

電子申請システムは、企業・個人が申請書を手入力してから提出まで、従来窓口で行っていた申請業務をインターネットで、行うことができるシステムである。データ記述方式として XML を採用することにより、データの再利用やデータベースと連携したシステム開発によるデータの効率的な管理、活用

が可能である。不特定多数の申請者に統一したフォーマットを適用し、大量の申請データを効率的に処理可能になる。

XML署名基盤技術を適用すれば、上記に示したXML文書の利点を損なわずに申請書に電子署名を付与できる。XML署名によって、電子申請文書に対して以下の三つのセキュリティ機能を実現する。

- 1) 申請内容の第三者による改竄防止
- 2) 第三者による申請の成りすまし防止
- 3) 申請者の申請内容の否認防止

また、現行手続きで認められている代理申請を電子申請にて実現する手段として、申請者の電子署名付き委任状を添付した申請書に対して代理申請者が電子署名を付与することが考えられる。XML署名基盤技術を用いれば、このように複数人数の署名を扱う電子申請書でも、署名付き文書モデル、および、システム化が簡単に実現できる。

[EDIシステムへの適用]

EDI関連分野でのXML署名基盤技術の適用例として、金融取引システム(手形決済、証券取引、保険業務等)が考えられる。これらの企業間の取引をオープンなネットワーク上で安全、かつ、確実におこなわなくてはならない。手形を始めとする金融券面を電子文書化する際の重大な問題点の一つである改竄、成りすまし、発行者の否認等をXML署名基盤技術を適用することにより解決できる。

特に、手形決済においては、手形の権利所有者の移行を裏書として記録していく現行方式を電子手形においても実現しなくてはならない。XML署名を用いれば、XMLが持つ文書の構造的処理の容易さから、裏書の署名を多重署名形式による電子手形として実現できる。

7. 課題

XML署名標準を電子政府システム、または、民間システムにて安心、確実に利用できるようにするには、以下の課題に取り組むことが重要である。

① 電子政府システムにおける公文書交換へのXML署名付文書構造の標準化

XML署名標準に基づく電子署名付文書構造は、Detach型、Enveloping型、Enveloped型の3種類をもち、鍵アルゴリズムや証明書情報の格納や取得の仕方も自由度が高く、仕様そのものはどのようなAPの要件も満足できるように設計されている。電子政府システムにおける電子署名付文書構造の標準化、処理手順を含んだ作業フロー、政府認証基盤GPKIとの連携等をさらに規定していく必要がある。

② XML署名ライブラリAPIの標準化

XML署名ライブラリAPIが標準化されていない場合、ベンダが提供するXML署名ライブラリ毎に各アプリケーションが個別に処理記述を記述することになる。Webブラウザ等を用いたクライアントシステムではクライアントにインストールされているXML署名ライブラリに従って、最悪、呼び出し記述を取り替えなくてはならなくなり、電子政府システムや電子商取引が普及する際に大きな問題となる。これを解決するために、XML署名を処理するライブラリの標準化をおこなう必要がある。

③ XML処理ツールのコンフォーマンスチェックによる相互運用性確保

電子政府においては、今後の申請業務の多様化と民間電子商取引との連携、および、海外政府との公文書交換等の適用場面の広がりを考えると、マルチベンダ環境での電子署名の使用は避けられない。XML署名は、XML文書のシンタックスに対して直接署名値を計算するのではなく、XMLが仕様で許されている冗長性を正規化した意味に対して署名値を計算する。このため、現状のXML処理ツール(XMLパーサ、XSLTエンジン、XPath

エンジン)の非互換により署名値が正しく計算されない障害が発生し、各社の署名付与、検証システム間の相互運用性がないことが問題となっている。運用システムで実際に使用するXML処理ツールと環境に対して、運用に先立ち事前に、XML文書によるあらゆる形式を用意したパフォーマンスデータにより検査する。発見された非互換部分のXML仕様を用いないガイドラインを作成し、システム開発者に徹底させることが考えられる。

また、マルチベンダ環境にて確実にセキュアな電子政府システムの早期立ち上げのためには、XML署名に確実に利用できるXML処理ツールであることを保証してくれる信頼に足る認定機関の設立等も望まれる。例えば、その機関は、W3C XML1.0完全準拠なXML処理ツールであることをパフォーマンスデータにより検査し、合格した際に完全準拠の認定を行うものとする。これに従って、W3C XML1.0完全準拠と認定されたXML処理ツールを使用していれば、マルチプラットフォーム環境でのXML処理ツールに起因したXML署名の非互換問題は解消される。

④ 日本語文字コード変換表の非互換吸収による相互運用性確保

XML署名は、処理時に、Unicode値に変換され、署名値が計算される。現状のXML文書は、Shift-JISやEUC-JPで作成・保存されることが多くコード変換表により実際にUnicodeへ変換している。このとき、ベンダ毎にコード変換表が実装されており、変換の相違点が「XML日本語プロファイル解説」(http://www.y-adagio.com/public/standards/tr_xml_jpf/kaisetsu.htm)等に報告されている。この非互換に起因して、A実行環境で署名付与したShift-JISのXML署名文書をAプラットフォームと異なる変換表を

もつB実行環境で署名検証すると、改竄等はしていなくても署名検証に失敗することがあり、非常に問題となっている。さしあたっては、以下のガイドラインをXML署名関連システム開発者に徹底させることにより回避は可能である。

- ?? Unicode変換表をシステム内で統一する。
- ?? システム内で利用するXML署名関連XML文書は全てUTF-8、または、UTF-16形式で保存する。
- ?? 上記のいずれもシステム側の制約により採用不可の場合は、署名対象の電子文書をBase64エンコーディングしたバイト列に対してXML署名を付与する。

早急な日本語コードの根本的解決は無理にしても今後UTF-8形式の処理・表示ツールの普及等を強力に推進する必要がある。

⑤ XMLパーサ非妥当性検証モードでの外部実体参照の解決

XML1.0仕様では、実装に任せられた仕様部がいくつかあるXML署名でXMLパーサを用いるときの注意として、XMLパーサを非妥当性検証モードで使用する際、XML文書に外部実体参照指定が存在するとき、外部実体参照している部分を展開するか、しないかは、パーサの実装にゆだねられている。XML署名値がXMLパーサの実装に依存して異なるという望ましくない状況が発生する。

8. まとめ

電子政府、電子商取引の情報セキュリティ確保のための基盤技術開発の一環としてXML署名基盤技術の研究開発を行ってきた。研究開発成果を関連する電子政府システム、電子商取引システムへ適用すべく連携を持っていく。同時に、他の先行電子申請プロジェクトへ適用可能性の評価を依頼することも考え

る。そのために、XML署名基盤ソフトのAPI仕様等の標準化も推進していく。

開発したXML署名基盤ソフトによりマルチプラットフォームでも間違いなく署名、検証ができることを確認するために、相互運用性評価検証実験を行った。この実験結果から、XML署名技術により、海外政府との公式電子文書のやり取りや民間システムとの連携に代表されるように、電子政府、電子商取引業務が多様化した際にも、問題なく署名付与・検証が確実に行えることを保証していく。

謝辞

本講演のWebサービスにおけるXML署名相互運用性評価に関する報告は、通信・放送機構(TAO)プロジェクトの研究開発成果を活用しています。

地質情報とXML

- 国土交通省の電子納品における地質情報に係わるXML-

(社)全国地質調査業協会連合会 情報化委員会 委員
川崎地質株式会社 中田 文雄

1. はじめに

建設業界(土木工事)の衛星圏には、測量、地質調査や建設コンサルタント(土木設計)などの業界があり、国土交通省などの官庁がこれらの業界に仕事を発注する場合には、工事発注とは異なる「業務発注」という契約形態を取ることが一般的である。

これらの業務における成果品は、土木や建築工事によって建設された構築物のように、目に見える形をしている「モノ」ではなく、測量図面、地質断面図や設計図などを含む「報告書」であることが一般的であり、言葉を換えて述べるならば「情報」を納品することにはならない。

報告書は、従来から用紙に印刷・製本されて提出されているが、基本的なことは報告書に印刷されている情報(内容)そのものが重要であることから、何らかの技術により報告書に記載された情報が電子化されても、価値が損なわれてしまうといたことではない。

一方、国土交通省が推進している「CALIS/EC直轄アクションプログラム」(国土交通省、2001、2002)の一つの柱である電子納品では、直轄業務の場合は2001年度から、成果品である報告書類はすべて「印刷媒体=本」から「電子媒体」に変更された。地質情報においても例外ではなく、従来は紙に印刷し製本して提出したボーリングデータや報告書本文などをXMLやPDF等で保存し、電子媒体で納品・提出するよう義務づけされた。

印刷・製本するようにシステムが構築されている情報をデジタル化するためには、ある一定の標準(基準や要領)が必要となるが、国交省では建設省時代の1998年度からその策定作業を開始し、2000年度から徐々にその標準を公表・適用してきている。

筆者が属している地質調査業界では、こうした国交省の動きに呼応して、地質に関する標準化の基本的な作業を担ってきた。

本文は公表された国交省の標準と、その策定経緯の概略を述べるものである。

2. 国土交通省の規定する標準

2.1 標準の種類と公表年月

国交省の規定によると、電子納品に関する定義は以下のとおりである。

- ① 電子納品とは、調査、設計、工事などの各業務段階の「最終成果を電子データで納品」することをいう。
- ② 電子データとは、各電子納品要領(案)等を示されたファイルフォーマットに基づいて電子化された資料・情報をいう。

表-1は、国交省が規定した電子納品に関する標準の種類と、それに使用されている全てのXMLファイルであって、管理(情報)と呼ばれるインデックスファイルの全てと、地質データの一部がXMLで提出するよう規定されている。

2.2 各標準の概要

土木設計業務等の電子納品要領(案)と工事完成図書の電子納品要領(案)は、基本的な標準の2つの柱であるが、その相互関係を図-1に示した。以下に、各標準を略記する。

表-1 国土交通省の電子納品の標準と全XMLファイル

要領(案)・基準(案)の名称	フォルダ	管理ファイルの名称	ファイル名	同左のDTD	最新年月
土木設計業務等の電子納品要領(案)	ルート	業務管理ファイル	INDEX_D.XML	INDE_D02.DTD	2001/08
	¥REPORT	報告書管理ファイル	REPORT.XML	REPO2.DTD	2001/08
CAD製図基準(案)	¥DRAWING	図面管理ファイル	DRAWING.XML	DRAW02.DTD	2002/07
デジタル写真管理情報基準(案)	¥PHOTO	写真管理ファイル	PHOTO.XML	PHOTO02.DTD	2002/07
測量成果電子納品要領(案)	¥SURVEY	測量情報管理ファイル	SURVEY.XML	SURVEY01.DTD	2002/07
	¥SURVEY¥KITEN	基準点測量成果管理ファイル	SURV_KTN.XML	SUR_KT01.DTD	2002/07
	¥SURVEY¥SUJUN	水準測量成果管理ファイル	SURV_SJN.XML	SUR_SJ01.DTD	2002/07
	¥SURVEY¥CHIKAI	地形測量成果管理ファイル	SURV_CHI.XML	SUR_CH01.DTD	2002/07
地質調査資料整理要領(案)	¥BORING	地質情報管理ファイル	BORING.XML	BRG0130.DTD	2002/07
	¥BORING¥DATA	(ボーリング交換用データファイル)	BEDNNN.XML	BED0200.DTD	2002/07
	¥BORING¥PIC	コア写真管理ファイル	COREPIC.XML	CPIC0100.DTD	2002/07
	¥BORING¥TEST	土質試験・地盤調査管理ファイル (土質試験結果一覧表データファイル)	GRNDTST.XML	GTST0100.DTD	2002/07
	¥BORING¥OTHR	その他の地質資料管理ファイル	OTHRFLS.XML	OTHR0100.DTD	2002/07

注：2002年9月現在、測量成果電子納品要領(案)で未公表の内容あり

注：CAD製図基準のうち、工事に関する規定は省略した。注：(****データファイル)付きのファイルはデータファイルである

- ① 土木設計業務等の電子納品要領(案)：地質調査報告書や設計書などを電子納品するために遵守すべき基本的な要領である。
- ② CAD製図基準(案)：CADソフトを使用して設計図面を作成する際に使用する基準であって、道路、橋梁、樋門と山岳トンネルなどがその対象となっている。主として、図面や線のサイズ、レイヤの命名規則から構成されている。
- ③ デジタル写真管理情報基準(案)：工事(現場)写真と参考図面のフォーマットや保存方式に関する基準である。
- ④ 地質調査資料整理要領(案)：ボーリング柱状図、土質断面図や地質断面図の記載方法に関する要領である。本文発表時(2002年10月)においては、岩盤柱状図や地質図に関する要領案は公表されているが、適用は2003年4月からとなっている。それ以前では、CAD製図基準(案)に準拠する(具体的には、業務別の協議結果による)。
- ⑤ 工事完成図書の電子納品要領(案)：土木工事を遂行する過程で発生する様々な書類や図面の電子化に関する基本的な要領である。

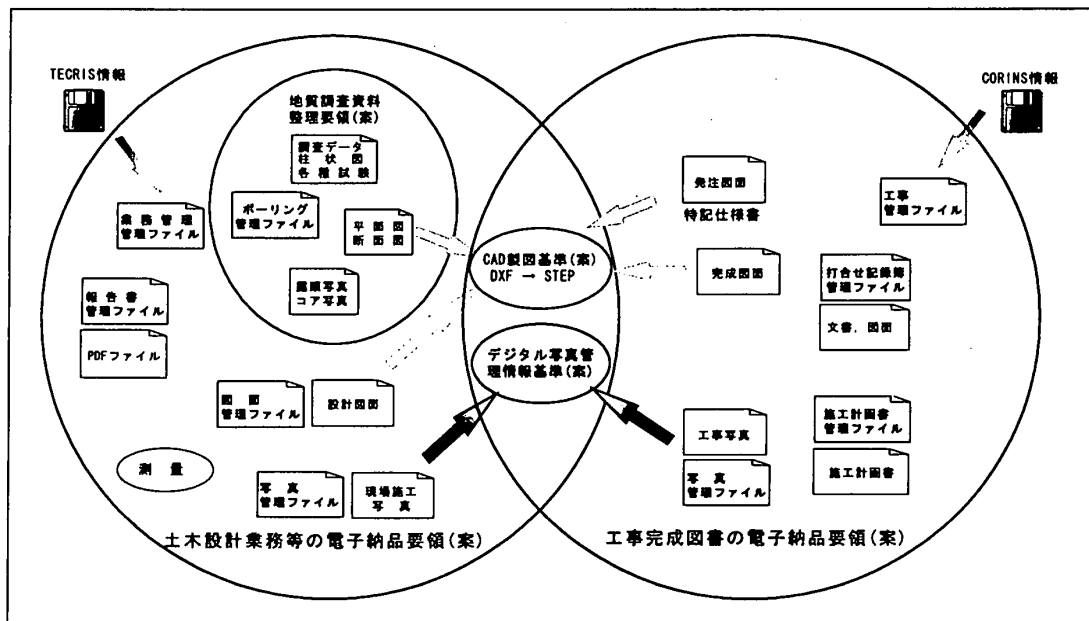


図-1 電子納品に関する各標準の相互関係

2.3 各標準の管理情報ファイル

それぞれの標準(要領や規準)には、業務内容や工事内容などをデータベース化する目的で、XML形式の管理(情報)ファイルが使用されている。DTDは外部参照ファイル形式となっており、XMLとDTDの文字コードは以下のようにになっている。

- ・XML : Shift_JIS(TRX0008:1999, TRX0015:1999)
- ・DTD : Unicode (UTF-16)

XMLにShift_JISが使用されているのは、国内で多く使用されているコンピュータの実情からとされている。

3. 土木設計業務等の電子納品要領(案)

同電子納品要領(案)は、委託業務に関する電子納品の基本的な規定を定めたものである。その骨子を以下に示す。

- ① 報告書本文の保存形式はPDFとするが、オリジナルファイルも提出する。
- ② ファイル管理を確実にするために、それぞれの電子データを格納するフォルダとファイル名は予め定められたものを使用する。
- ③ 報告書ファイル(PDF)とのリンクを確保するために、XMLで記載された管理情報(インデックス)ファイルを使用する。
- ④ 提出する電子媒体は、「650MBのCD-R」もしくは「230MBのMOディスク」となっているが、当面は改ざん防止のために650MBのCD-Rを使用するように規定されている。

3.1 成果品(報告書)のフォルダ構成

電子納品要領に規定された電子成果品(報告書)のフォルダ構成を図-2に示す

3.2 業務管理ファイル

業務管理ファイル(INDEX_D.XML)は、業務全体の概要を記述した管理ファイルであるが、図-3に示すDTDに従ってXML(Ver. 1.0)で記述したファイルである。

3.3 報告書ファイルと同管理ファイル

報告書管理ファイル(REPORT.XML)は、報告書名、同副題、作成したソフト名、オリジナルファイル名などの報告書管理項目がXML(Ver. 1.0)で記述されている。

なお、紙面の都合でXMLとDTDは省略した。

報告書ファイルは、設計図書に規定された成果品のうち、報告書、数量設計書、設計計算書などの電子データファイルであって、保存形式はPDF、ファイル名はREPORTnn.PDFである。なお、nnは原則として01~99の連続番号である。

報告書オリジナルファイルの形式については、特に制限は設けられていない。

3.4 CADデータと工事写真

本文では省略したので、詳細が必要な場合は巻末の参考文献を参照されたい。

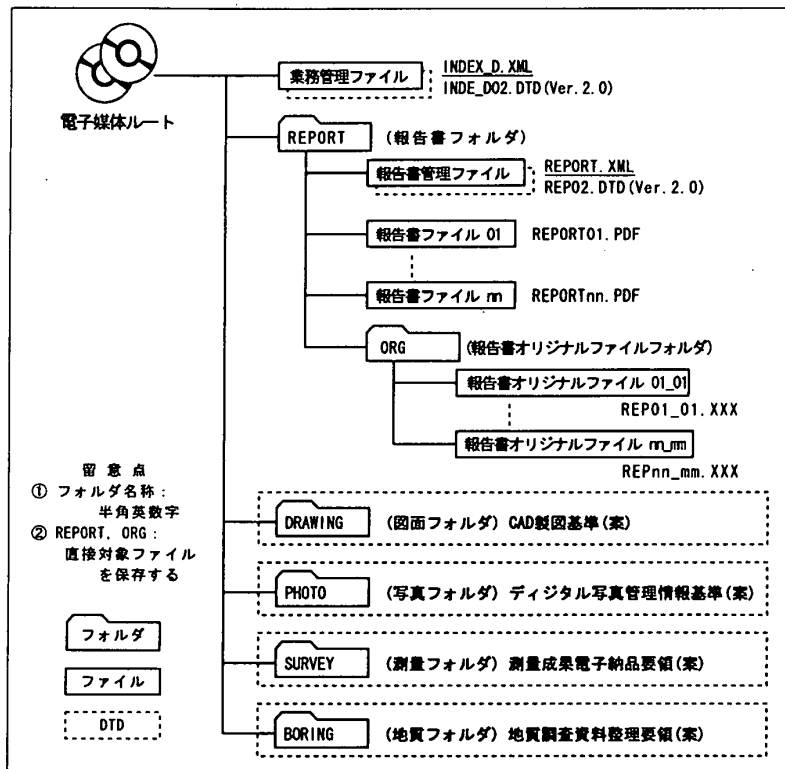


図-2 電子納品要領(案)のフォルダ構成

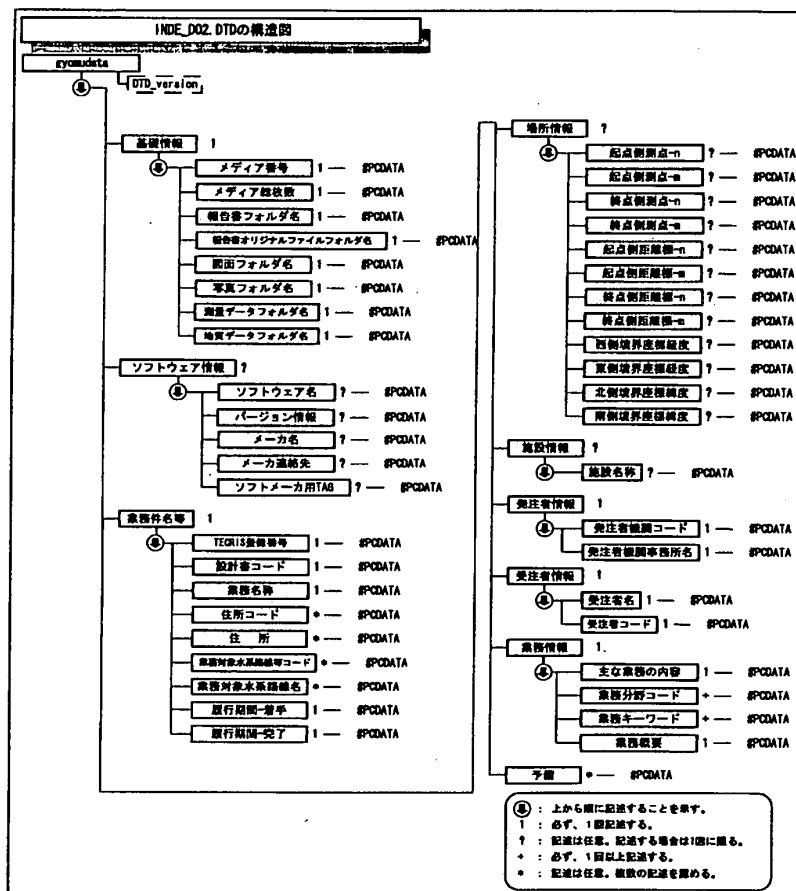


図-3 業務管理ファイルのDTD構造(国交省, 2001)

4. 地質調査資料整理要領(案)

ボーリングなど地質調査業務によって得られたデータ類は、全て本要領(案)に従って整理することになっている。ただし、2002年10月現在では、土質ボーリング柱状図と土質断面図だけが適用対象であり、その他の資料については2003年4月に追加適用されることになっている。

図-4 と表-2 は、地質データを格納するフォルダ・ファイル群と登録項目等である。

地質情報管理ファイルなど4種類のXML は、それぞれのデータ群を管理するインデックスファイルであり、「ボーリングファイルnnnn」と「土質試験結果一覧表ファイル」は、XML で記述されるデータファイルである。

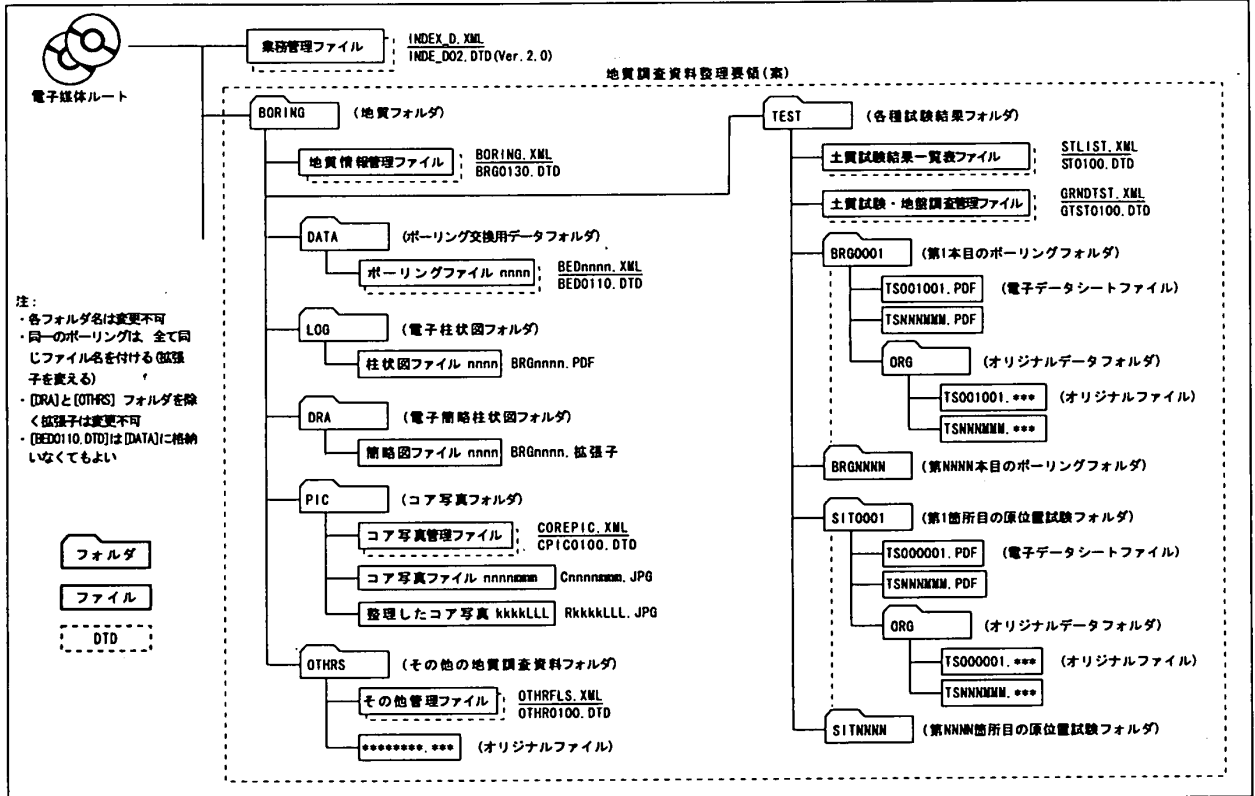


図-4 地質調査資料整理要領(案)のフォルダ構造

表-2 地質調査に関するデータの登録項目とフォルダ等(国交省, 2002に加筆修正)

地質調査資料の種類	フォルダ	サブフォルダ	ファイル形式	電子データの目的・内容	備考	
(1) 報告文	REPORT	-	PDF	閲覧用データ		
(2) 柱状図	BORING	① ボーリング交換用データ	DATA	XML	データ交換	専用ソフトの使用
		② 電子柱状図	LOG	PDF	イメージ出力用データ	同上
		③ 電子簡略柱状図	DRA	SXF (P21)*	設計過程へのデータ交換	同上
(3) コア写真	PIC	-	JPEG	200万画素以上		
(4) 土質試験・地盤調査	TEST	-	XML, PDF, 協議	データ交換, イメージ図	専用ソフトの使用	
(5) その他の地質調査資料	OTHR	-	協議	データ交換		
(6) 地質平面図	DRAWING	-	SXF (P21)*	設計過程へのデータ交換	専用ソフトの使用	
(7) 地質断面図		-	SXF (P21)*	設計過程へのデータ交換	専用ソフトの使用	
(8) 現場写真	PHOTO	PIC	JPEG	100万画素程度		

*: 2002年度のファイル形式は協議可

主なデータ項目の特徴を以下に略記する。

(1) ボーリング柱状図

ボーリング装置によって地盤を掘削して得られる地質情報は、従来ボーリング柱状図と呼ばれる印刷図面で提出することが一般的であったが、電子納品制度によって下記の①～③のように、3種類のデータを提出することになった(表-2, (2)①～③参照)。

- ① ボーリング交換用データ : データベース登録と受発注者間などのデータ交換に使用する XML である。参考資料を表-3 と図-5 に示す。
- ② 電子柱状図 : 従来から使用されてきたボーリング柱状図の印刷イメージファイルである。参考資料を図-6 に示す。
- ③ 電子簡略柱状図 : 設計・施行段階の構造物断面図などに使用されるCADファイルであって、通常CADソフトで再加工される。

表-3 ボーリング交換用データの登録項目(国交省, 2001)

登 録 項 目
表題情報, 地質区分, 色調区分, 観察記事, 観察記事枠線, 標準貫入試験, 相対密度・相対稠度, 原位置試験実施の有無, 孔内水平載荷試験, ボーリング孔を利用した透水試験, PS検層, その他の原位置試験, 試料採取情報, 土質試験結果, 地盤材料の工学的分類, 地質時代区分, 地層区分, 孔内水位, 掘削孔径, 孔径・孔壁保護, 断層情報, フリー情報

```

<?xml version="1.0" encoding="Shift_JIS"?>
<!DOCTYPE ボーリング情報 SYSTEM "BED0110.dtd">
<ボーリング情報 DTD_version="1.10">
<標題情報>
<調査基本情報>
<事業工事名>国道〇〇号建設事業</事業工事名>
<調査名>〇〇共同溝土質調査(その2)</調査名>
<調査目的>01</調査目的>
<調査対象>04</調査対象>
<ボーリング名>B-2</ボーリング名>
<ボーリング総数>10</ボーリング総数>
<ボーリング連番>0001</ボーリング連番>
</調査基本情報>
.....
<コア情報>
<地質区分>
<地質区分_深度>1.80</地質区分_深度>
<地質区分_地質名称1>埋土</地質区分_地質名称1>
<地質区分_地質コード1 >00001</地質区分_地質コード1>
<地質区分_地質名称2></地質区分_地質名称2>
<地質区分_地質コード2 ></地質区分_地質コード2 >
</地質区分>
.....
    
```

図-5 ボーリング交換用データ見本(国交省, 2001)

(2) 土質試験・地盤調査

ボーリングによってサンプリングされた土質試料は、必要に応じて土質試験によって密度などの物理的情報や圧縮強度のような力学的情報が求められる。本情報も、従来は用紙に印刷された情報であったが、電子納品では表-2 のように、交換用データ(XML, 他)と、印刷イメージ(PDF)ファイル等で提出することになった。

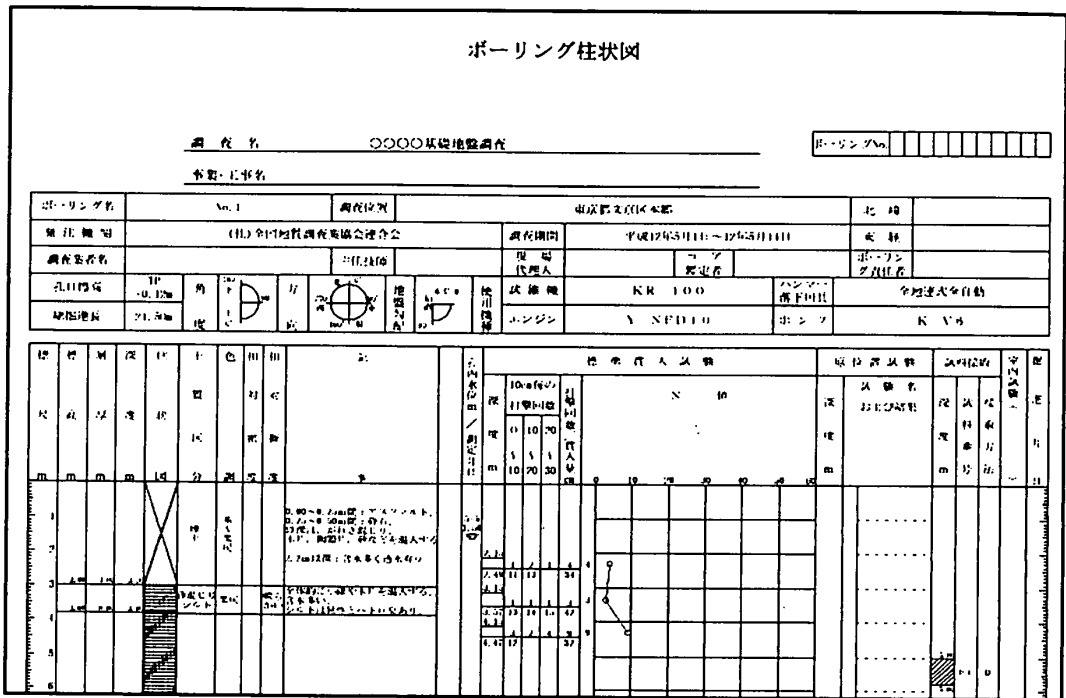


図-6 地質調査で得られるボーリング柱状図の例

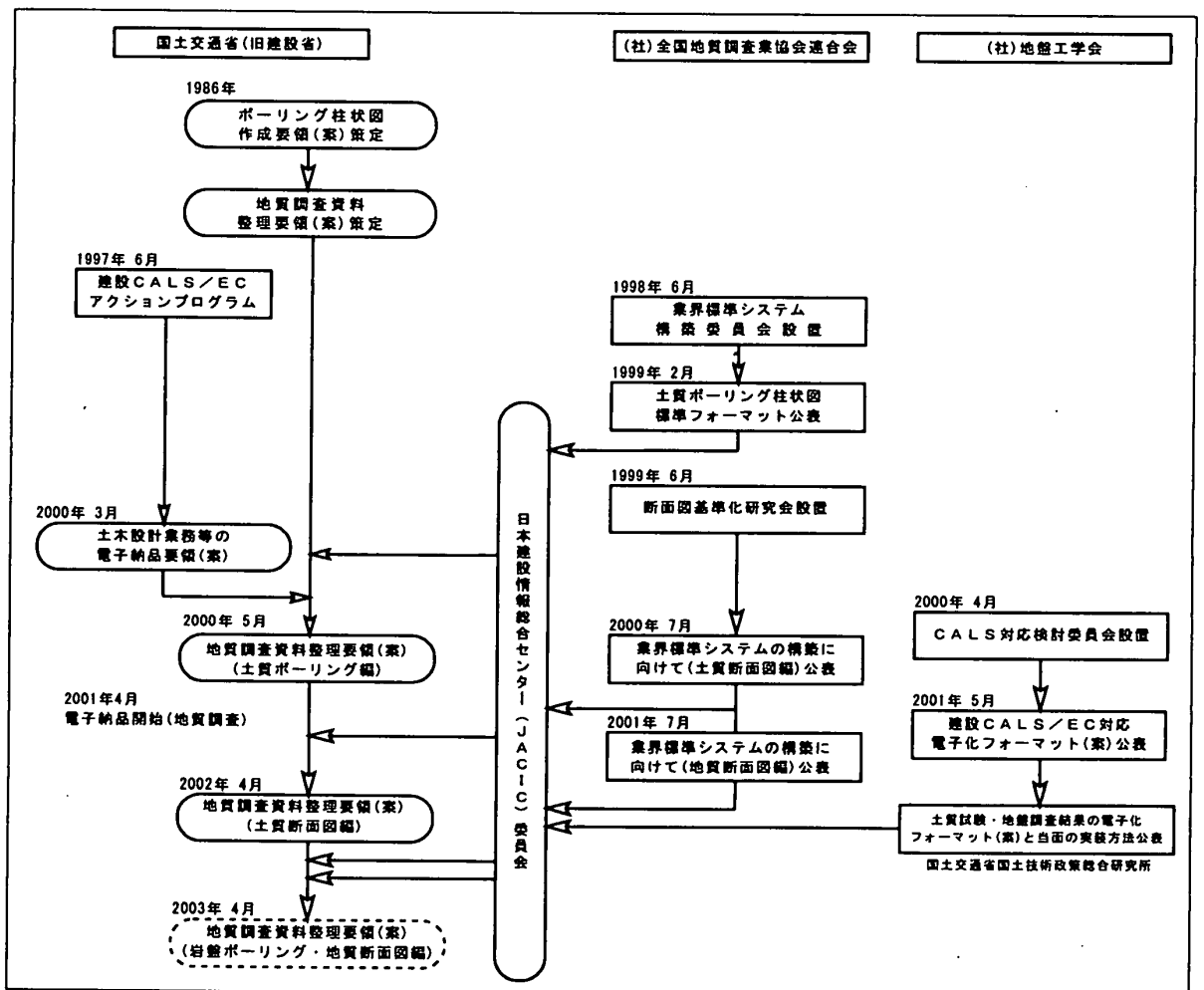


図-7 地質情報に関する国土交通省の標準と関連団体の関わり

5. 地質情報に関する国土交通省の標準と関連団体の動き

地質調査の電子納品を行う際に、遵守すべき地質調査資料整理要領(案)が制定された過程の概略関係を 図-7 に示した。

同要領(案)は、(財)日本建設情報総合センター(JACIC)内に設置された専門委員会で検討されているが、(社)全国地質調査業協会連合会(以後、全地連)と(社)地盤工学会の担当委員会が策定した提案がベースとなっている(全地連, 1999~2001)(地盤工学会, 2001)。

全地連では、建設CALS/ECアクションプログラムが公表になった1年後の1998年6月に標準化に向けての自主活動を開始し、その成果を当時の建設省標準として採用するように行政側に働きかけを行ったことが契機となり、合わせて3種類の原案が採用された。

地盤工学会の場合は、建設省土木研究所(現、国土交通省国土技術政策総合研究所)との共同研究として、学会制定の土質試験や地盤調査結果のデータシートを基にして策定した「土質試験・地盤調査結果の電子化フォーマット(案)と当面の実装方法」が原案である。

ただし、これらの原案はそのまま採用されたわけではなく、行政の施政方針に基づいてJACICの委員会や国土交通省(旧建設省)の担当部署で他の標準類との整合性が行われた後、国民に対するパブリックコメントを経て決定されている。

6. おわりに

2002年10月現在、国土交通省以外にも、農林水産省、日本道路公団、水資源開発公団など、最終成果を電子納品で提出を求めている発注者がある。一方、2~3年後には大多数の都道府県でも同様の電子納品が開始されることになっている。

成果品の電子化によって、保管場所の削減やデータベースの構築・利用といった、多くの利点が生まれるものと思われる。しかし、大きな図面を総括的に検討するといった場合などでは、ディスプレイでスクロールしながら閲覧するよりも、従来通り用紙に再印刷した方が見やすいなど、情報の受け手ではかえって不便や手間が掛かるような場合があり得るため、今後も標準の見直しが行われるものと考えられる。

筆者ら地質調査に関する企業活動に携わっている者としては、ISO9000sの規定もあって発注者側の意向に添う必要があり、地質調査報告書やそれに関連して提出する諸資料類の全てを電子化して提出するべく、社内業務遂行のプロセスの改変に取り組んでいる。

参 考 文 献

- 国土交通省(2001) 土木設計業務等の電子納品要領(案)
- 国土交通省(2002) CAD製図基準(案)
- 国土交通省(2002) デジタル写真情報基準(案)
- 国土交通省(2002) 地質調査資料整理要領(案)
- 国土交通省(2001) 電子納品運用ガイドライン(案)
- (社)全地連(1999) 建設CALS/ECに対応する業界標準システムの構築に向けて
- (社)全地連(2000) 建設CALS/ECに対応する業界標準システムの構築に向けて—土質断面図編—
- (社)全地連(2001) 建設CALS/ECに対応する業界標準システムの構築に向けて—地質断面図編—
- (社)地盤工学会(2001), 建設CALS/EC対応電子化フォーマット(案)
- 原 弘, 中田 文雄(2001), 地質調査業界の提案する地質図標準, 地質企業活動を巡る地質データ標準。
- 日本情報地質学会, インターネット時代の地質図標準
- 中田文雄, 金澤直人, 和田弘, 丸山昌則, 中川継一, 全地連(2002), IT対応地質調査報告書のまとめ方,
- (株)オーム社

XML 電子カタログと電子商取引の考察

沖電気工業株式会社

エンタープライズソリューションカンパニー (ESC)

技術情報ソリューション部シニアコンサルタント

XML コンソーシアム XML エバンジェリスト

藤岡慎弥

住所：東京都港区芝浦4丁目10番3号

E-mail：fujiooka563@oki.com

1. はじめに

筆者は電子部品情報について過去 8-10 年、データブックの電子化から始めてデータの一元管理多目的出版のために SGML や XML の活用も検討して来たが現在は、XML を応用した電子カタログと電子商取引システムを活用した RosettaNet 支援ビジネスを担当している。その経験で言うと日本だけかもしれないが、XML の一般ドキュメントへの応用より、XML/EDI システムの実装化の実現のほうが早まりそうな気が濃厚となっている。ECALS-2 補正予算プロジェクトでも「半導体と一般電子部品グローバル SCM インフラ整備」というテーマに挑戦したおかげで世界に誇れる ECALS 辞書が生まれ、部品情報を登録して検索活用する技術については各専門家がその要素技術の開発に専念され、それを引き継いだ JEITA でのポスト ECALS の活動としてボランティアベースで辞書の修正追加、維持管理活動で確実に進められている。しかし、なぜか日本の半導体業界では電子部品情報の準備では上流工程から一貫した合理化を考える標準化が進行しているようには見えない。XML 電子カタログや XML/EDI を活用した日本の SCM や電子商取引への取り組みも米国の RosettaNet プロジェクトに見られるような実用的でありながらいくつかの業界にまたがる流通プロセスの標準化や改革を迅速に推進しているようには見えない。何とか米国に太刀打ちできる実践的な仕組みを日本でも実現できないのか個人的に検討してきた結果をまとめ紹介するので、今後の取り組みの参考になれば幸いである。

2. ECIX(Electronic Component Information Exchange)

米国の半導体業界では部品開発・供給者として供給部品に関するデータを可能な限り、すば

やく、安価に、能率的に提供することを要求されている。そのためにこれまで時間と労力と資源と資金の面で信じられないほどの量が費やされてきた。この大量のデータブックの情報を作成提供するために抱えている共通の問題を解決するために 1993 年に日立アメリカ、Intel、National Semiconductors、Philips Semiconductors、Texas Instruments が PCIS(Pinnacles Component Information Standard)と呼ばれる共通の標準を定義するために Pinnacles Initiative を結成した。

PCIS の核心は SGML DTD である。ベンダーとプラットフォームに依存しない SGML の選択には時間はそれほどかからなかったが DTD の開発には数年をかけた。PCIS DTD が完成してもそれを実装するアプリケーションの開発にはいくつかの困難をともしない、情報交換の標準として利用するにはさらに時間がかかるかと思われていた。幸い、選ばれた SGML の構造と内容の表現の豊かさが業界共通のデータ構造と内容を示すタグをつけたデータで管理状態に置くことによってデータの一元管理や多目的出版、それから設計技術者が特定のパラメータやキーワードで検索することを可能にしていた。

Pinnacles Group は CAD フレームワーク・イニシアチブ (CFI) の組織下で活動をしていたがその後、Silicon Integration Initiative(Si2)のもとで電子コンポーネント情報交換プロジェクト (ECIX) に継続され、半導体業界のためだけでなく、頓挫していたかもしれないが、結局、すべての電子部品業界や電子機器業界のために活動の重点を移しながら、最終的には電子設計アプリケーションや振る舞いと機能モデルを作成するためのシミュレーション・ソフトで直接使えるデータが提供できる仕組みへと発展してきた。

その実装化の前にインターネットが普及し、この SGML 化のおかげで SGML から HTML への自動変換も可能となり当初から予定していた

CD-ROM と平行して流通の手段として、Web で HTML 電子カタログが実現されたのは自然の成り行きであった。しかし、HTML で流通できるようになったとはいえ、データの再利用には制限が多すぎて役に立たないことも判明、まもなく XML が検討されることになったのである。

一方、SGML や HTML と Java はいっしょに実行可能であった。Norbert Mikula という大学生が Philips Semiconductors のために動作可能な試験概念を開発している。この電子データブックは PSC-EDB と呼ばれているが SGML、DSSSL、Java を使っている。但し、PSC-EDB は Java のモバイル機能を一部実現するのに成功しただけである。もちろん、Java は安全を意識して設計されているので Web ブラウザでアプレットが動作している時にコンピュータにダメージを与えることはない。少なくとも、XML と Java の組み合わせでもそれ以上のことができそうであることを予感させるものであった。しかも Web の世界で使いやすい XML の標準が決められてから、SGML の経験が生かされ、辞書の保守に活用される辞書ビルダーから部品情報検索ツールまであらゆるツールが XML 対応化されて来ている。

電子データブックの大部分はプロパティの値やパラメータ名やパラメータ・シンボルとパラメータ記述であり、これらが設計技術者のユーザにとっても非常に重要な要素となる。これらの機能特徴は各々電氣的、機械的情報で、適用される製品の種類に従って、さらに識別できる必要がある。情報が十分に細分化され識別ができるようになるとソース&リフレクションという概念を使い、データシートのパラメータ値や記述などが別途保管され、必要なときに共通のデータとして Id で参照されることになる。データがどのように現れ、データの参照方法をどうするか判断する文脈を著者は提供することになる。そのデータを人間が読めるようにスタイル・レイアウトを提供する仕組みに加えて同じデータを機械可読にすることも可能となる。^[1]

部品情報辞書システム (CIDS)^[2] も XML 対応となり、IEC や JEDEC のような各種辞書標準データを取り込んで活用する仕組みはできているので表現規約や流通規約をお互いに理解した上で各々得意の分野の辞書データを交換することで辞書内容の充実や保守が容易になることが期待されている。

実際、ECALS-2 の標準化活動の成果となる ECALS 辞書コンテンツもロゼッタネット辞書 ECTD に 7-8 割取り込まれ実際に電子部品情報の照会応答システムで使われようとしている。ECTD は現在では IT 辞書と一緒に RNTD (RosettaNet Technical Dictionary) となっ

ているがその辞書の情報モデルは CIDS である。

従って、当初の SGML DTD ではその設計思想からすれば、かなり高邁な理想を実現しようとしたためにその実装化に手間取り、ECIX プロジェクトは実用化への道のりは長いという人もいたが、必要なクラスをプロパティとともにどんどん追加定義していき、その辞書に準拠して電子カタログコンテンツを準備することにより、半導体技術情報の XML 化を検討している現在では意外と早く実用化が実現される可能性が出てきている。

TDML (Timing Diagram Markup Language) は、デジタルコンポーネントのタイミングダイアグラム情報を取り扱う為の新しい標準規格である。Si2 の ECIX (Electronic Component Information eXchange) プログラムを通して Chronology 社と共に現在開発が行われている。これは人とコンピュータの両方に理解し易いタイミングダイアグラムのフォーマットを提供する。TDML はコンポーネント情報の取り扱い方式である ECIX の中で重要な役割を果たすと同時に、ECIX が EDA やシミュレーションを意識したプロジェクトであることのもう一つの証左である。

今日、流通しているタイミングダイアグラムの一般的なフォーマット (GIF、JPG、PDF) は、タイミングダイアグラムの“絵”を取り込んだだけのものでしかなく、従ってコンピューターベースの解析に使用する事が出来ない。コンポーネントベンダーは、SGML (Standard Generalized Markup Language : ISO 8879:1986) をベースにした TDML により、タイミングダイアグラム仕様と合わせてモデルの基礎となる情報の取り込みが可能になり、これらを直接デザインプロセスに使用する事が出来るようになる。もちろん、今ではこのツールも XML 対応となっていることは言うまでもない。そして設計者は、TimingDesigner のようなタイミングダイアグラムツールを入手して表示、編集、そしてベンダーが供給するコンポーネント仕様を直接設計に反映させる事が出来るのである。^[3]

3. ECIX の QD 仕様から PIP2A9 へ

当初、QuickDataV.1.01 では PCIS-DTD のために用意されていた CIDS 辞書に基づく ECIX-I 情報コンテンツを用意して ECIX レジストリに登録されたサプライヤと顧客の間で必要な電子部品情報を照会すると応答が得られる仕組みを立ち上げようとしたが、その後、QuickDataV.2 ではロゼッタネット実装仕様書 RNIF1.1 への対応を目指して EC 辞書の充実と ECIX-II の情報コンテンツを用意して SSL セキュリティのインターネット環境で PIP2A9 のドラフトに基づく電子部品情報の照会応答の仕組みを実現している。

そしてこれがインダストリ・レビューや投票を経て、現在の PIP2A9 の仕様書 V.2 として規格が制定され、PIE-MP(製品情報交換参照ストーンプログラム) で実装化の努力が行われ NEC、東芝、三菱電機、沖電気の主導する活動で日本の半田対応メーカーもグローバルの活動でも実証実験の成果を標準仕様の改定にフィードバックする形で貢献している。

(1) RosettaNet プロジェクト

1998 年の 2 月、XML の標準が制定されるや否や、それを待っていたかのように RosettaNet プロジェクトが開始され、パソコン流通の改革が始まった。流通大手のイングラムマイクロやパソコンメーカーのコンパックなどの業界の主だったメンバーが集まり、ビジネス・プロセスの見直しから、あるべきプロセスのモデリングまで検討し、Partner Interface Processes (PIP) という実際のデータの取り扱いまで含めた企業間データ交換の定義と辞書の定義を行い、PIP 対応の XML DTD や技術プロパティとビジネス・プロパティを含むマスタ辞書の制定も行なわれた。

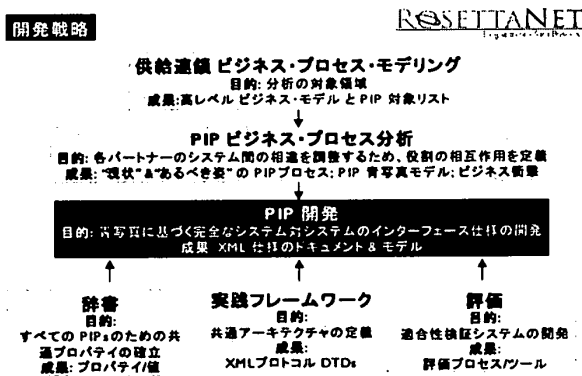


図 1: 開発戦略

標準化作業の進め方としてはクラスター・ワークショップではビジネス・プロセスの専門家が集まり、ビジネス・プロセスの現状の確認とあるべき姿について上位モデルでの合意を全体的視点から行なわれた。また、セグメント・ワークショップではシステム実装を担当する技術者が集まり、特定のデータの相互関係と技術的な振る舞いを定義した。しかもその検討結果を UML で表現してビジネスの専門家とソフト開発の専門家が共通の理解が得られる努力した結果がすべて標準仕様書としてまとめられている。同時に物流まで考慮した各種バーコードの確認と取り決めまで検討され、自動受発注から在庫管理や物流まで連動する仕組みが検討された。

[4]

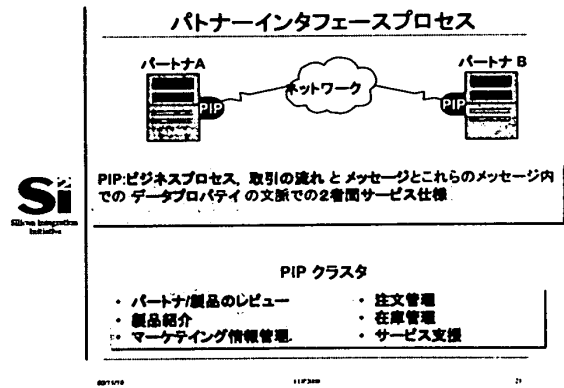


図 2: PIP 概念図

(2) ECIX/RosettaNet 戦略

各々、コア・コンピタンスに投資、RossetaNet は電子的ビジネス・プロセスの標準仕様の制定、ECIX は電子部品情報交換の標準制定に注力する。下記のような関係で協調作業を進め、Si2の専門家の協力を得て PIP2A9 として標準仕様書が定められている。

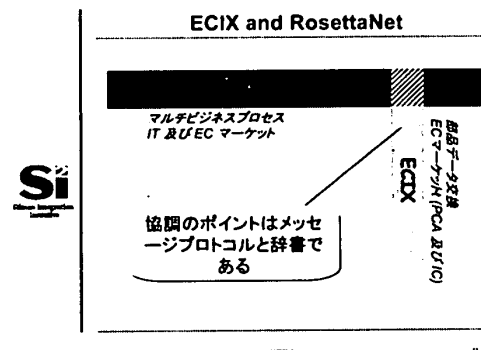


図 3: ロゼッタネットと ECIX の関係

4. PIP2A9 とは

(1) ECIX/QuickData の位置付け

PIP2A9 の EC 技術情報の照会は QuickData のメッセージがコアとして使われており、それがロゼッタネット実装フレームワーク (RNIF) の仕様に基づいて取り扱われ、部品供給者と顧客の間はレジストリで仲介されて効率の良い部品情報の照会応答システムを構築しようとしている。

(2) PIP2A9 の目的

電子部品のサプライチェーンではさまざまなレベルの顧客が部品の技術情報を必要とする。

次の 5 段階のプロセスに応じてそれぞれの情報を必要とするが、PIP2A9 では下記の図に示す仕組みで一貫性のある実装方法によらない技術情報要求応答手順の提供を目的としている。5 段階のプロセスとは発見、試作、購買、設計、製造のことであるが、その各段階で PIP 2 A9 の活用目的と考えられるのは次の通りである。

1. 発見：製品 ID を把握していない OEM 企業が本質的な必要性に基づいてベンダや販売流通業者を検索して、部品を特定する。
2. 試作：OEM 企業が設計課題に合わせて選択した情報(モデル、テスト・ベンチ、パッケージ・オプションなど)を評価する。
3. 購買：この段階では、価格決定、入手可能性、危険度の予測情報、ライフサイクル・データ、サンプル部品の入手可能性が重要となる。
4. 設計：OEM 企業が選択した部品を使用して設計する。この段階では、EDA のライブラリとモデル・ファイルが特に重要となる。
5. 製造：OEM 企業がプリント基板組立品を製造してテストする。この段階では、テスト・パターンと設定条件、部品を選択して配置するための物理的な情報、2次ソース情報源などが必要となる。

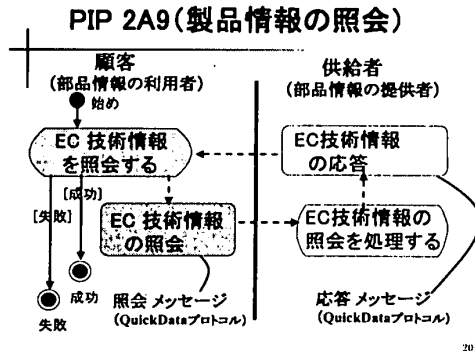


図 5 : PIP 2 A9 の照会/応答アクティビティ図

情報は、部品の諸特性 (characteristics)、諸特性セット群 (CharacteristicsSets)、または製品情報オブジェクト (ProductInformationObjects) のセットとして提供される。パートナーは、オンラインによる電子カタログや販売構成システムの運用、実際のプリント基板アセンブリ (PCA) 設計、PCA の製造とテストなどの電子部品のビジネスアクティビティに必要な情報の要求に応じて返す。この PIP で規定するのは、実際の情報交換プロトコルだけである。照会を作成するメカニズム(デーモン・プロセスを使用する方法やユーザが GUI を操作する方法など)と送信された情報を実際に使用するメカニズムは、パートナー固有のメカニズムとみなし、この PIP では規定しない。

この PIP は、実装状態に関係なく、恒常的に EC 情報の利用者がビジネス処理に必要な情報を必要ときに要求してコンピュータで認識可能な形式で受信できるという点で EC のサプライ・チェーン全体で重要な意味を持つ。また、情報提供者のさまざまなウェブ・サイトから人間が読み取り可能な形式で作成された多様な情報を手作業で取得するよりもコストを大きく削減できる。コンピュータで認識可能な形式(粒度の細かいレベルで定義される意味表現付き)で情報を受信すると、ソフトウェアを使用して、データを内部データベース、EDA ライブラリ、およびユーザが読み取り可能な表示形式に自動変換できる。そのため、データを再入力する必要がなく、エラー、時間、および費用を削減することができる。

また、情報提供者は、この PIP を使用すると、情報を要求するすべての顧客に対して一つの形式でデータを提供できる。情報を複数の形式で提示する必要がないため、費用を大幅に削減できる。

PIP2A9 による検索方法の特長

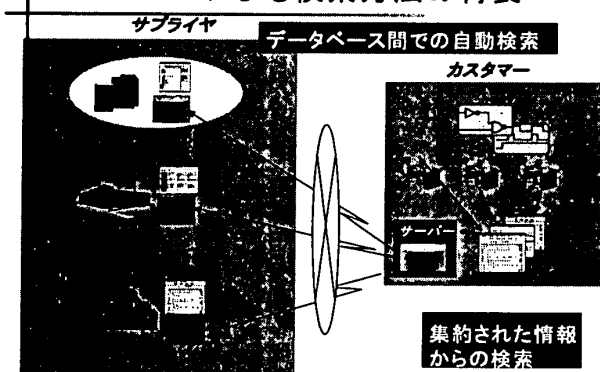


図 4 : PIP2A9 の目的と仕組み概要

(3) P I P のビジネス処理フロー図

下図の「電子部品の技術情報に関する照会 PIP 2A9」は、要求に応じて電子部品(EC)の製品情報を送信することを目的とする。

5. XML データの作成提供管理システム

上記のアイデアに基づき、半導体業界の関係者と相談結果、情報共有と流通のためのXML/DTD が作れたと仮定して、実際にデータを作成提供するにあたっての要望もしくはシステム要件をここにまとめ、紹介したい。

下記システム要件をみたすシステムを納入できるベンダーの協力を得て、半導体情報サービス提供を希望する各社に必要な設備を設置する必要がある。XML DTD の見直しや既存のXMLデータの活用まで考えながら、作業を進め、導入したシステム用のインスタンス処理テンプレートまで含めたスターキットやアプリケーションの開発を目指す。そのためにこの目的に関心のあるシステム・ベンダーなどに協力してもらって、実際に検索ロボットやエージェント機能に対応できる情報の作成提供に貢献できるシステム開発を提案する。

(1) 検討システム

まず、ロゼッタネットのSCM構築実証実験やebXMLの実証実験に参加する各社は協力希望のメーカーからXML電子カタログ対応のコンテンツマネジメントシステムやB2Bサーバーを設置してもらい、現在ある電子部品情報をXMLの構造タグや内容タグをつけて機械可読の状態にできるようにする。そのためにはスターキットやインスタンス処理テンプレートの作成まで協力してくれるメーカーを優先的にこのプロジェクトに参加させる。XMLで一元管理されたデータを必要に応じ、XMLもしくはHTMLに変換をかけてインターネット上で公開する。

もし、電子カタログのレジストリ&レポジトリ(R&R)センターのようなものを立ち上げるのであれば、そこで「御酒の筆筒」のようなXML対応の電子カタログ検索システムを立ち上げ、ディレクトリサービスを可能とする。ここでは検索した結果、各社にあるホームページにXMLデータとして混在させておけば、そのリソースをそのまま流用して、ある情報や画像などを使ってダイナミックに一覧表にして閲覧ができるようにする。^[5]

Si2では図6に示すようなレジストリ構想があり、実証実験も行われている。RosettaNetではWebサービス実装標準技術であるUDDIを利用した横断検索の可能性や見積もりや与信、受発注サービスに連動するSCMの可能性も検討されている。これからはWebサービスも意識したUDDIの構築が現実的である。ロゼッタネットも将来的にはebXMLのレジストリ&レポジ

トリ構想に従う予定であるが今のところはUDDIの採用もすでに宣言して実際にRosettaNetの各種標準であるPIPや辞書もUDDIに登録されている。

レジストリ

- ・参加者(プレイヤー)を特定する
 - 例: 会社名、URL、連絡窓口
- ・役割を特定する
- ・使用領域を特定する
- ・参加者を利用モデルや辞書に結びつける

・レジストリデータ例

Supplier	Part #	Quantity	Unit	Part #	Quantity	Unit	Part #	Quantity	Unit
000000	000000	000000	000000	000000	000000	000000	000000	000000	000000
000000	000000	000000	000000	000000	000000	000000	000000	000000	000000
000000	000000	000000	000000	000000	000000	000000	000000	000000	000000

図6: レジストリデータ例

実際、RosettaNetもUDDIの活用を前提にRosettaNet.orgをレポジトリとしてグローバルUDDIに公開するためにさらにレジストラーの必要性も検討しており、そのイメージを下記の図7に示す。

UDDIにRosettaNet®を登録

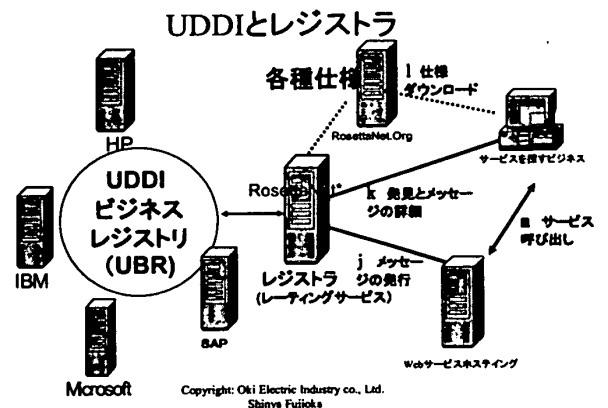


図7: レジストラー

もちろん、著作権の所在が明らかな素材データは直接レポジトリに登録しておいて有償の場合は利用状況に応じて決済・サービスにも連動する仕組みが必要となってくる。当面は銀行のファクタリング・サービスや宅配便の代引きサービスを利用して公開各社が小口ユーザーの代金も容易に回収できるようにする。必要に応じ、決済システムやXMLを応用したビルディングシステムと連動する自動受発注システムやXML/EDIシステムと連動させる必要がある。

そのために必要なら、PIP2A9対応の電子カタログポータルサーバーなどをレポジトリ構築のために導入し、各社から提供を受けたXMLデータを使って、見栄えの良い状態で対話型の電

子カタログを作成する。必要に応じ、デジタル衛星やインターネットを活用して各社の最新データにアクセスできるようにする。技術的に可能であり、必要なら、Verisign 等に登録して認証サービスや代金回収サービスを受けられる仕組みを構築する。

因みに最近では Web サービスが RosettaNet はもちろん、ebXML の実装標準技術として期待されているが、メインフレームとインターネット上のシステムを接続するために記述される WSDL は IBM やマイクロソフトで改良を加え、WWW の標準としても提案され、webMethods の B2B 統合サーバーの実装標準の一つとして活用され実用化され始めている。

この他、CommerceOne で採用している XML 応用規格である xCBL や Ariba 社が採用している cXML に各々対応する eMarketplace システムはもちろん、最近ではロゼッタネット対応の B2B サーバーもいろいろなソフトベンダーからソリューションが提供され実稼働を始めている。将来は ebXML 対応のメッセージ交換仕様 ebMS 対応のシステムを活用することも考えられる。

(2) 日本の半導体情報作成側の問題点と解決案

先行事例があるにもかかわらず、日本の活動を見ているとどうも日本で独自に標準を検討して世界に情報発信をしたいという声はどこからか聞こえてきて、米国の参考になる事例やシステムがすでにあるにもかかわらず、十分検討もされないまま、補正予算プロジェクトでは日本の独自性や日本独自のシステム開発になりがちであった。欧米の既存の技術や考え方は参考になるものはグローバルな視点で取り入れるべきものは取り入れて改良を加え、世界に対して逆提案できるようになればそれも世界の電子商取引の実現に向けての立派な貢献につながることになるのではないか。

どうしても日本の既存の技術に拘ると実績のある RDBMS の公開情報のデータベースや各社固有の Web アプリケーションサーバーシステムを構築することで留まってしまい、あまり斬新なことができないのが今までの先例から言える。RDB を中心とする過去のデータベース技術だけではメンテナンスも含め各種公開データをいちいち登録するのが面倒である。そのためには実用化の段階に入ってきた ORDB やさらに実用化の期待される XML ネーティブの OODB システムの構築が必要となってくる。

SGML でも XML でもやはりデータの構造内容は一応決めておく必要がある。これからは DTD が XML スキーマとなり、さらに便利にな

る可能性はあるが、いずれにしても DTD 的なものを検討する必要はある。はたして、各業界で必要な DTD を決めるための標準化が日本でも可能なのだろうか。ECIX で提案された PCIS-DTD を日本の業界でそのまま利用して良いのかあるいはそれを取り込んで ECIX の QD 仕様が開発され、ロゼッタネットの PIP2A9 仕様へと発展してきた仕様をさらにカスタマイズする必要はないのか検討する必要がある。ECALS-2 の標準化活動の一つとして辞書の作成に参加したが各社間の用語の不統一性は想像以上のものがあり、同じ会社の製品群の中でも用語が統一されておらず、この機会に統一できるものは統一すると良いのだがこのまま放置しておけば、いたちごっこの様相も呈してくる恐れがあった。この現状に対して日本ではコンテンツの準備と補正予算を使って開発された ECALS システムはともかく、幸い、ECALS や RosettaNet の技術辞書の作成と追加修正については業界の関係者の涙ぐましいボランティア活動でこの辞書の標準化はかなり進められている。この際、全製品で用語の辞書を作って統一することは至難の技なので少なくとも各製品群ごとに最低限度、統一したい用語を辞書として登録し、同じクラス名やプロパティ名が製品によっては微妙に違う内容で使われていることが多いのであれば、XML のネームスペース機能を活用して各製品群ごとの DTD を作成し、その要素名に従ったプロパティ情報の作成、登録が自由にできるようにすることも考える必要がある。幸い ECALS 辞書作成ではその問題を認識している JEITA の関係者が一致協力して作業した結果が実を結んでおり、ロゼッタネットに大部分取り込まれた ECALS 辞書も JEITA による地道な今後の維持管理はその重要な足がかりとなるであろう。

マルチメディア化を迎えるに当たり、今までのように静止画や CAD データは DTP 用の EPSF の保存だけではなく、IGES、EDIF、STEP のどのフォーマットが良いのかデータフォーマットの標準化についても検討を要する。また、現状では実用的観点から、素材データとして CAD、CG などの各種属性フォーマットのままで保存しておく必要がある。XML 対応の MPEG などの動画データはこれからの問題としても、静止画像も PICT、TIFF、GIF、BMP、PNG、G4 などの各種フォーマットに対応できる必要がある。あるいは XHTMLbasic や SVG ならびに VoiceXML や X-BML などブロードバンド時代の携帯電話やデジタル衛星の活用も意識した対応を迫られている。

6. システム構築の今後の課題と期待

(1) XML 関連ツール

過去 10 年以上も実績のある SGML 対応のアプリケーションを Web テクノロジーとして注目を集めている XML にどんどん対応させる必要がある。さらに SGML/XML 関連のシステムは当初 UNIX 対応が多かったため Mac や Windows クライアント対応の操作性を実現する必要がある。

Web ブラウザ上で動作する SGML/XML 対応のレポジトリ・ブラウザを提供する必要もある。SGML/XML エディタとして ADEPT シリーズや FrameMaker + SGML など直接対応できるシステムが望まれる。Office2000 やその後継ソフトも XML に対応するようになってきているので今まで以上にデータシート情報が作成管理しやすくなることも期待されるが、マルチメディア対応で電子部品情報を提供するなら、JSP のようにサーバ側ですべてのコンテンツを提供処理するにも限界があるので、3次元 CAD や動画によるシュミレーションのためにも curl^[6] などの XML にも対応しておかつクライアント側でも分散処理可能なアプリケーション用オーサリング・ツールの開発も SMIL や SVG 対応のメディア・ブラウザのインテリジェント化と同時に必要となってくるかもしれない。^[7]

7. 結論

ECALS-2 と RosettaNet の PIP2A9 のたたき台となった ECIX 標準を比較するとどちらも電子カタログの標準化と検索性をよくして設計者の便宜を図ろうとしているようだが若干異なるところがあると思われる。

その一つは ECIX ではデータシート全体の情報を含む情報の構造と内容を検討した過去の資産を活用した XML の PCIS-DTD があり、TI や National Semiconductors はそのサブセットの DTD で半導体情報を管理状態に置いてあるのでいつでも RNTD の辞書にしたがってロゼッタネットの PIP2A9 で標準化された照会応答の DTD に対応して必要とあれば検索サービス用に項目をしぼって公開データベースにデータの登録をする際にも人間が介入することなく自動的に必要な項目を選択し、登録できる可能性がある。一方、ECALS-2 では実証実験のための登録公開用 XML-DTD はあるかもしれないが、各社の情報作成提供側は上流工程のワークフローは各社各様となっており、ECIX のような流通と交換を意識した各社の標準となる SGML はもちろん、XML の DTD の議論は進んでいない。結局、公開のためのデータの登録はテンプレートを活用

して人間が登録に介入しないとデータの公開や検索サービスに利用してもらえないことになるが、その手間を惜しんでそのうち誰もデータの登録に協力してくれなくなる恐れがある。一方、2-3 年先行している欧米は余り大きな問題もなく、新しい E-ビジネス対応の情報提供が可能である。ロゼッタネットへの対応もそんな工数をかけなくとも同様のサービスやそれ以上のことが実現できるようになってきているからである。

また、ECALS-2 ではまだデータブックに記載されていた情報しか、実際の検討の対象となっていないが、ECIX では当初から EDA データをどうするか、CAD 設計のための仕組みを検討してきている。その証拠に TDML というタイミング・ダイアグラムのマークアップ言語の検討も進められてきており、ECIX の活動も Si2 で継続する条件に Express モデルの検討が入っていた。また、最近では E-コマースや IP 流通の時代も予測して ECIX の手法を拡大利用できる仕組みづくりも検討していることが論文から推測できる。^[7]

従って、欧米の演繹的なやり方には日本の企業はなじまないところがあるが、日本でも ECIX や RosettaNet の活動と歩調を合わせながら、学ぶべきところは学び、仮説と検証を繰り返しながら、実用化の面では日本も得意とするところがあるのでお互いに得意な分野の成果を持ち寄り、同じ土俵の上で、正々堂々と戦える環境作りのためのビジネス協調も必要である。

こうして、上流工程まで標準化を意識した電子部品情報の作成提供の環境ができてくれば、そのデータをユーザの要望で構築した検索サービスへの自動登録も可能となり、最新のデータを検索して評価検討結果、そのまま見積り用のフォームに必要なデータを入れて、見積り依頼ボタンをクリックするとロットや各種条件に応じた見積りも瞬時に行なわれるようになる。そのまま OK であれば、XML/EDI による自動受発注はもちろん、発注ボタンを押すと WebEDI システムを経由してもサプライヤの ERP に連動し、在庫の確認を行うとともに出荷の指示がでて、FedEx などの 3PL アウトソーシング・サービスで各国の港の倉庫で管理されている商品が即座にトラックで出荷される。一方、受注の指示が出た瞬間、決済引き落としサービスに連動しており、キャッシュフローも良くなり、経営状況も改善できる。昔のように銀行から不必要に借りた金で大量生産した結果、在庫を抱え、代金回収まで長期間かかることもなくなるのである。これが可能となるにはセットメーカーからの受注情報がリアルタイムに流れ、毎週のように半導体メーカーが生産計画を見直し、適正在庫に自動的に収束するようないわゆる SCM

の仕組みを各社が作る必要がある。そのためにも XML 電子カタログと XML/EDI システムが連動する新しい電子商取引システムを構築する必要があり、XML の応用がその鍵を握っているといっても過言ではない。^[9]

謝辞：ED-CALS 研究会や ECALS-2 の活動に参加されている JEITA 関係者の方々のご意見を聞き、ECALS 辞書づくりの地道な活動にアドバイザーとして参加させていただいたおかげで、いろいろ考えさせられました。また、その後、ロゼッタネットジャパンの活動に参加して調査研究した結果を参考にまとめてみました。関係者の方のご指導御鞭撻に対し、この場を借りて感謝申し上げます。

参考文献

- [1] XML を知る(株)プレントイスホール出版 18.5 電子データシート pp340 - 343
- [2] <http://www.si2.org/ecix/CIDS/OVERVIEW/architecture.html>
- [3] <http://www.si2.org/ecix/wml/>
- [4] http://www.rosettanet.org/general/index_general.html
- [5] <http://www.watch.impress.co.jp/internet/ww/article/980522/mallof.htm>
- [6] <http://www.cag.lcs.mit.edu/curl/>
- [7] 日経マルチメディア 1998 年 7 月号 pp190-191
- [8] Paper 'Sensible Exchange of IP Information' Mr. Donald. R. Cottrell
- [9] ロゼッタネット完全解説オーム社出版

お知らせ：新しい編集規程が整備，制定されました

2002年9月20日

情報知識学会誌編集委員会
委員長 安永尚志

8月27日の本会理事会におきまして，長年の懸案でありました本会誌の編集委員会規程，編集規程が制定されました。今まで，このような本会誌の規程類は整備されていませんでしたが，ここに新たに整備，制定されました。これらは本会の定款に基づいて制定されました。

また，これに基づき本会誌への論文などの投稿規定および執筆要領も改訂し，新たに制定しました。したがって，従来の投稿規定は廃止されましたので，ご注意下さい。今後はこれらの新しい投稿規定，執筆要領などにより，ご投稿をお願いします。

新規程類は本号（3号）でご案内していますが，本学会のホームページにも掲載しています。ご参照下さい。

学会ホームページ URL： <http://www.jsik.jp>

なお，同時に論文の査読規定も新たに整備しました。これにより，本会誌のレベルも一層高まるものと期待しています。

今回整備された規程類は以下の通りです。

1. 編集委員会規程 (定款に基づく編集委員会に関する規程)
2. 編集規程 (定款に基づく本会誌編集に関する規程)
3. 投稿規定 (本会誌に投稿するための規定)
4. 執筆要領 (同上のための原稿の書き方など)
5. 様式1 (投稿に当たっての投稿整理カード)
- 様式2 (採択原稿の掲載整理カード)
6. 査読規定 (投稿原稿の査読の方法などの規定 (委員会内規))

以上，6. 査読規定を除き，ホームページ，会誌で公開しています。

なお，編集委員会も，今年5月の総会において新たに発足しました。現在までに上記の規程類の整備を精力的に進めてきました。新たな編集方針として，編集と印刷出版の分離をはかりました。編集作業は新編集委員会が当たり，印刷作業は本会事務局が当たります。新しい体制は，9月からスタートし，今年度第3号特別号から作業を開始しました。

投稿論文は4-6ヶ月で掲載できるよう査読などの迅速化に努めます。投稿原稿の種類も研究論文は基より，事例／調査報告，解説／展望，論談，討論，研究速報，講座などを設け，これらも重視します。

会員のみなさまのご投稿を歓迎いたします。よろしくをお願いします。

「情報知識学会」編集委員会規程

2002年8月27日

制 定

(設置)

第1条 情報知識学会定款第7章に基づき、本会に編集委員会（以下、「委員会」という）を置く。委員会に関する事項は、この規程に定めるところによる。

(任務)

第2条 委員会は、会長の諮問に応じ、情報知識学会誌（以下、「会誌」という）の編集に関する事項について審議し、会誌の発行に当たる。

(組織)

第3条 委員会に委員長を置く。

2 委員長は定款第6章による編集長が担当する。

3 委員長は情報知識学会正会員（個人会員）の中から若干名を委員として任命する。

(委員会)

第4条 委員会は委員長が招集し、その議長となる。

2 委員会に必要に応じて役職委員を置くことができる。

3 委員長に事故あるときは、あらかじめ委員長が指名する委員がその職務を行う。

(任期)

第5条 委員の任期は2年とし、再任を妨げない。

2 委員長が交代したときは前項に関わらず任期は終了する。

(委員以外の者の出席)

第6条 委員長が必要と認めるときは、委員以外の者を委員会に出席させることができる。

(小委員会)

第7条 委員会に具体的、専門的な事項を検討するために小委員会を置くことができる。

2 小委員会に関し必要な事項は、委員会が別に定める。

(庶務)

第8条 委員会の庶務は情報知識学会事務局において処理する。

(雑則)

第9条 この規程の定めるもののほか、委員会の運営などに関し必要な事項は、委員会において別に定める。

第10条 この規程の改定は理事会の議決を経て行う。

付則

1 この規程は2002年8月27日から施行する。

「情報知識学会誌」編集規程

2002年8月27日

制 定

0. 情報知識学会定款第6章による機関誌「情報知識学会誌（以下、会誌という）」の編集に関する事項は、この規程の定めるところによる。
1. 原稿の提出と受付および保管
 - 1.1 本学会会員は本学会が発行する会誌に投稿することができる。ただし、その内容は著者の責任とする。
 - 1.2 原稿の書き方ならびに投稿の手続きは別に定める「投稿規定」「執筆要領」による。
 - 1.3 投稿時の原稿はすべて事務局に提出する。
 - 1.4 事務局は受けとった投稿原稿の受付年月日を記録し、保管する。
 - 1.5 編集委員会は会員または非会員への依頼によって原稿を集めることができる。
 - 1.6 討論を主な内容とする原稿のときは、編集委員会はそれを討論相手に見せ、なるべくその回答の原稿を求める。
2. 原稿の審査と採否
 - 2.1 編集委員会は原稿を審査し、掲載するかどうかを決める。
 - 2.2 編集委員会は、原則として原稿の査読を会員または非会員に依頼する。ただし、この場合も最終的な審査は編集委員会の責任において行うこととする。査読結果は原則として3ヶ月以内に著者に通知する。
 - 2.3 編集委員会は投稿規定と査読結果に基づいて、著者に修正を求めることがある。修正の求められた原稿は修正後すみやかに返送するものとするが、原稿が3ヶ月以上たつて編集委員会に返送されてきたときには原則として新規の投稿原稿として取り扱う。
 - 2.4 編集委員会が掲載適当と認めたとき（以下、これを「受理」とよぶ）、編集委員会は受理年月日を記録し、著者に通知する。
 - 2.5 編集委員会が掲載不適当と認めたとき、その理由を明らかにした文書を著者に送り、その了解を得たうえで原稿を返却する。
 - 2.6 掲載された原稿は原則として返却しない。ただし、図についてはあらかじめ申し出があれば、返却する。
3. 原稿の掲載
 - 3.1 投稿原稿の掲載順は原則として受理の順とする。ただし、編集上の都合により順序を変更することがある。
 - 3.2 学会費など、本会に納入すべきものを著しく滞納している会員の投稿原稿は、それが納入されるまで、掲載を延期することがある。

4. 校正

- 4.1 初校正は原則として著者が行う。第2校正以降は編集委員会が行うが、必要に応じて著者に依頼することがある。
- 4.2 著者は、校正した原稿を指定期日までに編集委員会に返送する。返送が著しく遅れた場合には編集委員会で校正するか、次号に回すこともある。
- 4.3 雑誌発行の日時が迫っているときなど、時間的な制約を受けるときには、著者による校正を略し、編集委員会が校正を行う場合がある。
- 4.4 校正のときに、著者が原稿と著しく異なるように書換えを行ったときは、その掲載を次号以降に延期することがある。

5. 別刷

- 5.1 別刷（抜刷）は著者の申し出により可能であるが、実費負担とする。

6. 会誌の内容

- 6.1 広い意味での情報知識学に関連し、またその発展に貢献するもの（情報／知識の収集、整理、蓄積、検索および各種解析、利用などに関するもの）とする。
 - 6.1.1 研究論文（Research Paper）：オリジナルな研究論文で、内容の主要な部分が学術論文として他に公表されていないもの。
 - 6.1.2 事例／調査報告（Report）：情報知識学に関連したシステムなどの開発、利用、調査に関するもの。資料も含む。
 - 6.1.3 解説／展望（Review）：情報知識に関連した特定分野の論文や学説などを総括、解説、紹介、あるいは技術動向などを展望したもの。技術、研究上の処理、解析方法などに関する解説。
 - 6.1.4 論談（Proposal Paper）：情報知識学に関連した新たな知見の表明、提案など。
 - 6.1.5 討論（Discussion）：本会誌に掲載された論文についての学術的な討論。
 - 6.1.6 研究速報（Notes）：技術、手法、新事実などの簡単な報告。
 - 6.1.7 講座（Lecture）：情報知識学の各分野に関する基礎理論、技術の適用などについて、テーマを定めて系統的に説明するもの。
- 6.2 学会記事（News）
 - 6.2.1 本会の事業、運営などに関する報告、記事、資料。
 - 6.2.2 ニュース、お知らせ。最近出版された単行本やモノグラフの紹介。
- 6.3 講演（Lecture）
 - 6.3.1 特別号などにおける講演資料。
- 6.4 その他
 - 6.4.1 上にあげたもののほか、編集委員会が適当と判断した事項。

7. 規約の改訂

- 7.1 本規約の改訂は、編集委員会の議を経て、理事会の承認を得なければならない。

8. 施行

- 8.1 本規約は2002年8月27日から施行する。

「情報知識学会誌」投稿規定

2002年8月27日

制 定

0. 情報知識学会誌編集規程による本会機関誌「情報知識学会誌（以下、会誌という）」への投稿に関する事項は、この規定の定めるところによる。

1. 投稿資格

投稿者の少なくとも1人は本会員でなければならない。ただし、編集委員会による依頼原稿の場合にはこの限りではない。

2. 投稿原稿

2.1 広い意味での情報知識学に関連し、またその発展に貢献するもの（情報／知識の収集、整理、蓄積、検索および各種解析、利用などに関するもの）とする。刊行時において未発表の原著でなければならない。本会誌の記事の種類を以下に示す。

2.2 投稿者は会誌記事の種類を明記して投稿しなければならない。ただし、編集委員会で変更することがある。

- (1) 研究論文 (Research Paper)：オリジナルな研究論文で、内容の主要な部分が学術論文として他に公表されていないもの。
- (2) 事例／調査報告 (Report)：情報知識学に関連したシステムなどの開発、利用、調査に関するもの。資料も含む。
- (3) 解説／展望 (Review)：情報知識に関連した特定分野の論文や学説などを総括、解説、紹介、あるいは技術動向などを展望したもの。技術、研究上の処理、解析方法などに関する解説。
- (4) 論談 (Proposal Paper)：情報知識学に関連した新たな意見の表明、提案など。
- (5) 討論 (Discussion)：本会誌に掲載された論文についての学術的な討論。
- (6) 研究速報 (Notes)：技術、手法、新事実などの簡単な報告。
- (7) 講座 (Lecture)：情報知識学の各分野に関する基礎理論、技術の適用などについて、テーマを定めて系統的に説明するもの。
- (8) 学会記事 (News)：本会の事業、運営などの報告、記事、資料など。
- (9) ニュース、お知らせ (News)：ニュース、お知らせ。最近刊行された単行本やモノグラフの紹介。
- (10) 講演 (Lecture)：特別号などにおける講演資料。
- (11) その他：編集委員会が適当と判断したもの。

2.3 会誌記事の種類のうち、(1)から(6)までは査読を行う。その他については編集委員会で編集を行う。

3. 投稿原稿

3.1 原稿の形式

(1) 投稿時の原稿

以下の A, B のいずれかの体裁でプリントされたワープロ原稿（横書き）4部の提出とする。その他、執筆に関する詳細は「執筆要領」を参照のこと。

A 刷り上り原稿を想定したレイアウト (A4 判, 2 段組, 20 字×46 行×2 段).

図, 表は希望の位置に配置すること.

B ベタ打ち原稿 (A4 判, 40 文字×40 行).

図, 表は, 1 枚ずつ別の用紙に印刷すること. ベタ打ち原稿右余白に図表の挿入位置を朱書きすること.

(2) 採択決定後の原稿

以下の C, D の両方の形式で記録された電子媒体 1 部の提出とする.

原稿の送付にあたってはフロッピーディスクなどの適当な電子媒体とする. 詳細は提出時に事務局に相談のこと.

C Microsoft Word, 一太郎, DVI, PDF などの代表的なフォーマット.

D プレインテキスト (文字コードは Shift-JIS).

図, 表は十分な品質で印刷できる形式 (JPEG, GIF など).

3.2 原稿の制限

(1) 原稿の長さを原則として次のように制限する.

研究論文, 事例/調査報告, 解説/展望, 論談	:	刷り上がり	20 ページ以内
討論, 研究速報, 講座	:	刷り上がり	6 ページ以内
ニュース他	:	刷り上がり	2 ページ以内

(2) 図原稿 (原図) の大きさは A3 判を越えないものとする.

(3) 原則として, 図版も含めてモノクロ印刷とする. ただし, カラーでなければならない図版を使用する場合は, 別途編集委員会と相談する. なお, カラーページやページを超過する分については, 印刷費を著者の全額負担とする.

(4) 使用言語は日本語または英語とする.

4. 原稿の採否

投稿原稿の採否は, 専門家による査読の後, 編集委員会において決定する. 不採択となった原稿は, 編集委員長より理由を付して通知する.

5. 査読のプロセス

学会員の中から編集委員会が指名した査読者 2 名によって査読を行う. 内容によっては, 編集委員会は著者に照会し, 原稿の修正を求めたうえで, 再査読を行うことがある.

6. 校正のプロセス

採択が決定した投稿原稿は, 掲載原稿として著者に校正を依頼する. 著者による校正は原則として 1 回とする. その際, 字句の修正以外は原則として認めない.

7. 別刷

別刷 (抜刷) は著者の実費負担とする. 希望部数を事務局に申し出ること.

8. 投稿の手続き

原稿投稿時には下記の書類を添え、原稿送付先に郵送する。

8.1 必要書類

(1) 最初の投稿時

- a. 投稿原稿整理カード：ホームページからコピーして、必要事項を記入し、印刷したものを1部。掲載原稿整理カードと兼ねるので、コピーを保存しておくこと。
- b. 紙媒体の原稿（図、表を含む）：4部。
なお、投稿者は著者校正用に原稿のコピーを保存しておくこと。
- c. E-mailによる連絡票
 - ・ 標題（和文、英文）、著者名（和文、ローマ字）、所属機関／住所（和文、英文）、要旨（和文、英文）、キーワード（和文、英文）、刷り上り予定ページ数
 - ・ 連絡先：著者1名の連絡先（氏名、所属機関／部局、所属機関住所、電話番号、Fax番号、E-mailアドレス）なお、投稿後の連絡は主としてE-mailで行う。

(2) 採択決定後の投稿

- a. 掲載原稿整理カード：投稿時のカードに追加事項を記述し、印刷したものを1部。
- b. 3.1(2)に指定した電子媒体：1部。
- c. 3.1(2)のCの印刷出力（プリントアウト）：1部。

8.2 原稿の送付先

〒110-8560 東京都台東区台東1-5-1 凸版印刷(株)内
情報知識学会事務局
電話：03-3835-5692 Fax：03-3837-0368
E-mail：LDE01013@nifty.ne.jp

8.3 原稿の受付

事務局が原稿を受け取った日を受付日とする。受付の確認を1週間以内に投稿者の連絡先にE-mailで通知する。不備のある投稿原稿は返送し、再提出するものとする。

9. 原稿提出期日

投稿は随時とする。ただし、特集号などは除く。

10. 著作権

- 10.1 機関誌『情報知識学会誌』に掲載された論文（電子版を含む）の著作権（著作財産権、copyright）は情報知識学会に帰属する。
- 10.2 掲載論文は冊子による出版の他、電子的に蓄積し、本会が行う情報提供サービスなどを通じて公開する。

11. 規定の改訂

- 11.1 本規定の改訂は、編集委員会の議を経て、理事会の承認を得なければならない。

12. 施行

- 12.1 本規定は2002年8月27日より施行する。
- 12.2 本規定の施行により、現行規定（第4版（暫定版）2002年3月）は廃止する。

「情報知識学会誌」執筆要領

2002年8月27日

制 定

1. 一般的な事項

本会誌への投稿は、「投稿規定」に従い、投稿原稿は本執筆要領に従って作成されなければならない。

本会誌の投稿原稿の種類には、研究論文、事例／調査報告、解説／展望、論談、討論、研究速報、講座、本会記事、講演、ニュース、その他がある。

2. 日本語原稿の構成

2.1 全体構成

(1) 第1ページ（査読者には見せない）

- ・ 標題（和文および英文）
- ・ 著者名（和文およびローマ字、ローマ字による著者名は、名、姓の順で、姓は全て大文字を使用する。）
- ・ 所属（和文および英文による所属機関名）
- ・ 住所（和文による所属機関の住所、E-mail、脚注とする。）
- ・ 見出し用原稿（研究論文、事例／調査報告、解説／展望、論談の原稿には、刷り上がりページ上部欄外につける著者名および標題を30字以内で書く。）

(2) 第2ページ目以降（査読者に見せる）

- ・ 要旨（研究論文、事例／調査報告、解説／展望、論談の原稿には、和文および英文で要旨をつける。和文要旨の長さは400字以内とする。英文要旨の長さは500語以内とする。要旨中には、図、表、数式などを用いない。本文中の図、表、数式、文献などを番号で引用しない。）
- ・ キーワード（研究論文、事例／調査報告、解説／展望、論談、討論、研究速報、講座にはキーワードをつける。和文および英文でそれぞれ5個程度、和文と英文のキーワードは、対応することが望ましい。キーワードはカンマ（,）で区切る。）
- ・ 本文（和文または英文）
- ・ 文献、付録など（和文または英文）
- ・ その他（とくに長い論文の場合、読者の便宜を考えて内容目次を付してもよい。ただし、章、節の見出し程度とする。）

2.2 本文 (Body)

(1) 構成

章、節などの構成は、第1レベルは1, 2, …, 第2レベルは1.1, 1.2, …, 第3レベルは1.1.1, 1.1.2, …のようにする。

(2) 脚注

脚注はできるだけ避ける。止む無く使用する場合は簡潔な文とする。

(3) 図および表

- a. 図、表にはそれぞれ通し番号をつける。図1 (Fig.1), 図2 (Fig.2),… 表1 (Table 1), 表2 (Table 1), …など。

- b. 通し番号とともに説明文（キャプション）をつける。キャプションの位置は図は上部に，表は下部とする。
- (4) 数式，化学式
 - a. 数式（独立式），化学式は，段落外で記述されているものも本文中で一回は参照する。
 - b. 数式には，通し番号を振る。
- (5) リスト（または箇条書き）
 - a. 記号なしリスト。
 - b. 記号つきリスト。リストの記号は，数字，アルファベット，記号を用いることができる。ただし，これらの混在した使用は避ける。アルファベットは 1 論文中では大文字，小文字の使い分けをしない。
 - c. 複雑化を避け，せいぜい 2 段（親子関係）のリストとし，ネストを跨ぐ順序づけを用いない。
- (6) 注記および参考文献

本文中で少なくとも一回は参照すること。通し番号で参照し，タイトルなどでの参照は避ける。

2.3 後付け (End)

(1) 謝辞

本文の最後に続けて記述する。章番号は用いない。章題は「謝辞」とする。最終原稿時に記述することが望ましい。

(2) 注記および参考文献

- a. 注記または参考文献には，参照順に通し番号を付し，本文の最後に番号順にまとめて記述する。章題は「参考文献」とし，章番号は付さないこと。
- b. 1つの番号には1つの注記または参考文献を対応させる。
- c. 注記中には参考文献を含めない。注記はできる限り簡潔に表現すること。
- d. 参考文献の記述形式は，以下の形式を満たさなければならない。
- e. URL を参照してもよいが，移動または削除される可能性があるため，極力避ける。原著が URL のみしか参照できない場合など，やむをえない場合は用いてもよい。その場合，参照時点でのハードコピーを保管しておくなど，参考文献へのアクセス手段を確保するよう努力しなければならない。

【参考文献の形式】

1. 雑誌中の 1 論文

[引用通し番号] 著者名：論文名，雑誌名，巻号，掲載ページ，出版年，その他。
2. 図書 1 冊

[引用通し番号] 著者名：書名，版表示，出版地，出版社，総ページ数，出版年，その他。
3. 図書の 1 部

[引用通し番号] 著者名：論文名，書名，版表示，出版地，出版社，掲載ページ，出版年，その他。
4. 会議報告

[引用通し番号] 著者名：論文名，書名（会議名），版表示，編集者名，会議開催地，会議開催年，会議開催機関，出版地，出版社，掲載ページ，出版年，その他。
5. インターネット上の論文

[引用通し番号] 著者名や標題など可能な限り詳細な書誌事項, URL, 参照年月日. (単なるホームページなどは参考文献にしないこと).

【参考文献の記述】

1. 著者名, 編集者名の記述

- (1) 個人著者名は, 姓, 名の順に記述する. 欧文著者名は, カンマ (,) で姓, 名を区切る.
- (2) 複数著者の場合は, 各著者をセミコロン (;) で区切る.
- (3) 翻訳図書などの翻訳者名の場合は, 著者名の後に括弧 () に入れて記述する.

2. 論文名, 書名の記述

- (1) 論文名, 書名は, 和文の場合はかぎ括弧 (「 」), 欧文の場合はダブルクォーテーション (“ ”) に入れて記述する.
- (2) 図書中の一部を引用した場合の書名は, 和書の場合は二重かぎ括弧 (『 』) に入れ, 欧文の場合はイタリック体で記述する.

3. 掲載ページの記述

- (1) 論文の場合は, 開始ページと終了ページを記述する. 「pp. 開始ページ-終了ページ」とする.
- (2) 図書の場合は, 総ページ数とする. 「総ページ数 p.」とする.

【参考文献の記述例】

- [1] 藤原譲: 「情報知識学試論」, 情報知識学会, Vol.1, No.1, pp.3-10, 1990.
- [2] 原正一郎; 安永尚志: 「国文学研究支援のための SGML/XML データシステム」, 情報知識学会, Vol.11, No.4, pp.17-35, 2002.
- [3] Fujiwara, Shizuo: “East-West Communication and Information Transfer -- Coordination of Specificity”, Journal of Japan Society of Information and Knowledge, Vol.4, No.2, pp.11-18, 1994.
- [4] Ellis, David (細野公男監訳, 斎藤泰則, 鈴木志元, 村上泰子訳): 「情報検索論」, 丸善, 180p., 1994.
- [5] 根岸正光: 「学術情報の流通と利用」, 『情報学とは何か』情報学シリーズ 3, 丸善, pp.43-69, 2002.
- [6] 名和 小太郎: 「デジタル図書館と著作権」, デジタル図書館, No.4, http://www.dl.ulis.ac.jp/DLjournal/No_4/nawa/nawa.html (2002年8月27日参照)

3. 文章と文体

- 3.1 文体はひらがなと漢字による口語常態 (である調) とし, 現代かなづかいを用いる.
- 3.2 漢字は当用漢字とする. ただし, 固有名詞や学界で広く用いられている慣用の術語はこの限りではない.
- 3.3 句読点その他には 「,」 「.」 を用いる.
- 3.4 本文中の人名には敬称をつけない. ただし, 謝辞の人名はこの限りではない.
- 3.5 数量を表す数字はアラビア数字とする.
- 3.6 数式は印刷に便利のように十分注意して記号を記すこと. 原則として数量 (変化量) を表す記号はイタリックとする.

- 3.7 ローマ字の人名の姓は大文字体とする。
- 3.8 固有名詞で読み誤るおそれのあるものにはふりがなをつける。
- 3.9 英数字は原則として半角英数文字で記述する。

4. 英文原稿

英文による投稿原稿の場合も、原則として和文による投稿原稿の諸規定に従う。英語圏以外の著者の場合、著者名表記にその国語による表記を認めるが、可能な限り英文表記とする。

- 4.1 研究論文、事例／調査報告、解説／展望、論談、討論、研究速報などの原稿は英文でもよい。
- 4.2 英文原稿は語学的に難点の少ないものであることを必要とし、著者の責任において完全を期する。
- 4.3 英文原稿には、英文による要旨 500 語程度、ならびに日本語による 400 字以内の要旨をつける。ただし、著者が日本語を理解できない場合は日本語要旨を省略できる。

5. その他

原稿は和文または英文によるものとする。文章は語学的に難点の少ないものであることとし、著者の責任において完全を期する。編集委員会は語学的校正を行わない。

6. 要領の改訂

- 6.1 本要領の改訂は、編集委員会の承認を得なければならない。

7. 施行

- 7.1 本規定は 2002 年 8 月 27 日より施行する。

様式 1 / 様式 2

「情報知識学会誌」投稿原稿整理カード／掲載原稿整理カード

1. 論文種別
 2. 標題（和文）
 標題（英文）
 3. 著者名（和文，ローマ字）
 4. 所属機関名（和文，英文）
 5. キーワード（和文，英文）
 6. 連絡責任者（1名）
 氏名，所属機関／部局，同住所，電話番号，Fax 番号，E-mail アドレス
 7. 送付投稿原稿
 テキスト部分の枚数
 図の枚数
 表の枚数
 付録の枚数
 8. 図の返却希望（YES，NO）
 9. カラー図の有無（原則として認めていないが，カラーでなければならない場合など，全額実費著者負担）
 10. 投稿日
-
11. 登録番号
 12. 受付日（再受付日）
 13. 受理日
 14. 送付掲載原稿
 フロッピーディスクなど 枚数
 プレインテキストプリント 枚数（刷り上がり見本，図表なども貼り込んだもの）
 15. 別刷り（抜刷り）の希望部数（全額実費負担）

【投稿原稿整理カード】

1. 標題を「様式 1 投稿原稿整理カード」とし，1 から 10 項目（11 項目以降は採択後）をもれなく A4 判横書き 2 枚程度に，ワープロでお作りください。
2. 投稿原稿と一緒に送ってください。
3. なお，投稿時には「投稿規定」にある E-mail による連絡票もお忘れなく，送ってください。

【掲載原稿整理カード】

1. 標題を「様式 2 掲載原稿整理カード」とし，全項目をもれなく A4 判横書き 2 枚程度に，ワープロでお作り下さい。掲載原稿と一緒に送ってください。

お知らせ [論文募集]

情報知識学会第11回(2003年度)研究報告会 発表論文募集について

実行委員長 国沢 隆 (東京理科大学理工学部)

情報知識学会では平成15年5月24日(土)に、前回と同じ学術総合センター(東京・一ツ橋)の会場で総会とともに研究報告会を開催する予定です。この研究報告会の発表論文を下記要領で募集いたしますので、学会員の皆様どうぞ奮ってご応募ください。

1. 募集分野

- (1) 情報知識の構造解析、モデル化、可視化、知識発見
- (2) 情報・知識の表現、生産、組織化、検索、提供
- (3) 電子出版、電子図書館
- (4) マルチメディア、電子ミュージアム
- (5) 用語、シソーラス
- (6) 知識情報の流通と知的所有権
- (7) 専門分野における品質管理、基準化
- (8) その他情報知識学に関連する諸研究・開発

2. 応募方法

発表論文題目、著者名(連名の場合、登段発表者に○印)、所属、論文概要(200字以内)、連絡代表者の氏名、住所、電話・FAX番号、電子メールアドレスを明記の上、下記宛になるべく電子メールにてお申し込みください。

〒278-8510 千葉県野田市山崎2641

東京理科大学理工学部 国沢 隆

E-mail: kunisawa@rs.noda.tus.ac.jp FAX: 0471 23 9767

応募期限: 2003年3月31日 採択可否通知: 2003年4月7日

3. 論文執筆・発表について

- (1) 採択されたものについては論文執筆要領の詳細を別途お知らせいたします。発表論文はA4版で4ページを目処とお考えください。4ページまでは無料ですが、それを超えるると有料(1ページ1,000円)になります。
- (2) 発表時間は、質疑応答を含めて30分を想定しています。なお、論文提出がないと発表はできません。
- (3) 登段発表者は当学会員に限ります。当日入会も可能です。

情報知識関連新刊図書一覧 2002年6～8月期

- 知識と推論 新田克巳著 サイエンス社 2002・6 1,500円
- I T社会総合データブック 2002 生活情報センター編集 生活情報センター 2002・6
14,800円
- 情報システム監査 吉田洋著 税務経理協会 2002・6 3,000円
- デジタル・エコノミー 2002/03 米国商務省著 室田泰弘編訳 東洋経済新報社 2002・6
1,800円
- 文化情報学 安沢秀一・原田三朗編著 北樹出版 2002・6 2,700円
- 電子署名・認証 渡邊新矢・小林覚共著 青林書院 2002・6 2,100円
- 情報のエコロジー 吉井博明著 北樹出版 2002・6 2,900円
- 知識の哲学 戸田山和久著 産業図書 2002・6 2,600円
- 知的財産権用語辞典 知的財産権用語辞典編集委員会編 日刊工業新聞社 2002・6 2,400円
- システムの思想 河本英夫著 東京書籍 2002・7 2,000円
- 電子自治体 榎並利博著 東洋経済新報社 2002・7 1,800円
- よくわかる.NETテクノロジーのすべて David Chappell 著 インプレス 2002・7 2,500円
- 情報リテラシー入門 室賀進也・山下巖共著 コロナ社 2002・7 2,200円
- 情報デザインケーススタディ Jules Yoshiyuki Tajima 著 翔泳社 2002・7 2,800円
- 情報社会とデジタルコミュニティ 大沢幸生ほか著 山田誠二編 東京電機大学出版局
2002・7 1,900円
- 「IT革命」はどこへ消えた 三石玲子著 主婦の友社 2002・7 1,300円
- ユビキタス・ネットワークと新社会システム 野村総合研究所著 野村総合研究所広報部
2002・7 1,800円
- eラーニング白書 2002/2003年版 先進学習基盤協議会編著 オーム社 2002・7 2,800円
- デジタル・デバインド C&C 振興財団編著 NTT出版 2002・8 4,400円
- 知能システム工学入門 松本啓之亮・黄瀬浩一共著 コロナ社 2002・8 2,200円
- 学習とそのアルゴリズム 電気学会GA・ニューロを用いた学習法とその応用調査専門委員会編
森北出版 2002・8 3,400円
- IT法大全 西村総合法律事務所ネット・メディア・プラクティスチーム編著 日経BP社
2002・8 4,800円

(注) 図書館流通センター刊「週刊新刊全点案内」を参照

コンピュータ関連の技術・操作関連のものは除外した。また、初心者入門的なもの、
ビジネス色の強いもので除外したものもある。

制作者 平田周 2002/9-30

会告 平成15～16年度役員選出について

来春は情報知識学会役員の改選期に当たります。前回（2001年春）の選出方法を踏襲し、さらに全会員の意向を反映させるべく、以下の要領で役員選出を実施します。

[1] 役員候補者の推薦

後掲の正会員名簿の中から役員候補者を推薦してください。

- a. 推薦者は正会員3名以上で、会長・理事・監事各1名を推薦する。
いずれか1名のみ（例：会長1名）を推薦することも可。
- b. 原則として候補者本人の了解を得、推薦者代表は平成14年11月末日までに文書で情報知識学会事務局へ提出。FAX、電子メール可。郵送の場合、当日消印有効。書式自由。
- c. 次期副会長は次期会長が指名する。
- d. 理事会も独自に役員候補者を推薦する。事業の継続性から理事は約半数留任。

[2] 役員候補者の確定

推薦受付終了後、本年12月4日に理事会を開き、理事会が推薦する者を含めた候補者全員を確定。情報知識学会誌12巻4号で候補者リストを公表し、投票方法をお知らせします。会長候補者は学会運営について、総会までに所信を表明します。

[3] 投票および総会での承認

平成15年度総会（2003年5月下旬）に投票結果を報告し、新役員について承認を求めます。

[4] 推薦書送付先（平成14年11日末日締切）

情報知識学会事務局

〒110-8560 東京都台東区台東1-5 凸版印刷(株)内

FAX:03-3837-0368 E-mail:LDE01013@nifty.com

情報知識学会／正会員名簿（平成14年10月1日現在、五十音順、敬称略）

相田満, 青木俊樹, 青木仕, 青戸邦夫, 赤井寧子, 茜ヶ久保幸一, 赤星哲也, 秋元学, 朝尾正昭, 浅見真一, 芦沢実, 東英弥, 安宅彰隆, 阿部正信, 天野力, 天野久栄, 荒井元明, 有川節夫, 安斎裕臣, 安東孝二, 安藤正人, Andrew J. L. ARM, 池田聡史, 池田祥子, 池村匡哉, 石井啓豊, 石井正彦, 石川徹也, 石島優子, 石塚英弘, 泉文男, 伊東千夏, 伊藤鉄也, 伊藤勝, 井上孝, 今坂直子, 今沢真, 岩崎政彦, 岩澤まり子, 岩田修一, 岩淵幸雄, 上田昇, 上田英代, 上村隆一, 牛島利明, 薄葉威士, 宇陀則彦, 内田尚子, 内田美由紀, 内田保廣, 内山哲夫, 宇都宮啓吾, 江見伸夫, 近江晶, 大内英範, 大江和彦, 大久保恒治, 大島輝夫, 太田泰弘, 大野照夫, 岡部建次, 岡本和彦, 岡本由起子, 岡谷大, 奥雅博, 小口雅史, 長田孝治, 尾身朝子, 貝島良太, 海田茂, 香川靖雄, 影山千花, 梶谷二郎, 加藤愛二, 加藤恭子, 加藤信哉, 神尾達夫, 河合勝彦, 川添良幸, 川俣修寿, 川村敬一, 神立孝一, 菊田昌弘, 木澤誠, 岸田和明, 岸永康延, 岸本吉浩, 喜多村政美, 衣川純一, 木村三郎, 木本幸子, 工藤喜弘, 国沢隆, 久野高志, 久保田, 久保正敏, 倉島節尚, Christian Galinski, 小池恵子, 小池澄男, 児島秀樹, 小島正美, 小林昭夫, 小林康隆, 小松功治, 小宮千鶴子, 小山照夫, 惟村直公, 後藤智範, 五島敏芳, 後藤将之, 権藤卓也, 斎藤尚夫, 阪口哲男, 阪田貞宜, 坂本光, 相良佳弘, 桜井仁, 笹井実, 佐々木耕侖, 佐々木由香, 笹森勝之助, 佐竹一夫, 佐藤成子, 佐藤純一, 佐村敏治, 沢恒雄, 澤伸恭, STEFAN KAISER, 重元康昌, 宍戸駿太郎, 柴田武, 柴山守, 斯波義信, 島袋徹, 志村純子, 下垣弘行, 仲宗根, 周防節雄, 菅原秀明, 杉尾俊之, 杉谷嘉則, 杉山真澄, 須崎琢也, 鈴木功, 鈴木尚志, 鈴木英夫, 鈴木博実, 瀬尾雄三, 副島博彦, 孫媛, 高橋正美, 高橋靖明, 高山隆裕, 田窪直規, 竹内寿, 竹下精一, 竹森利清, 田隅三生, 館野義男, 田中一雄, 田中成典, 田中猛彦, 田中久徳, 田中康仁, 谷口祥一, 谷本玲大, 田畑孝一, 田村貴代子, 田良島哲, 大師堂清美, 崔杞鮮, 千速敏男, 千原秀昭, 次田皓, 辻宏子, 土本茂, 土屋達, 津原重久, 當山日出夫, 時実象一, 戸田慎一, 富安寛, 豊島正之, 中尾浩, 中川優, 中条宗彦, 中谷登志雄, 中根康雄, 中挟知延子, 中挟義夫, 中林和典, 中道琢郎, 中村勝己, 中村隆志, 中村英明, 仲本秀四郎, 中山堯, 中山幹敏, 中山亮一, 長尾真, 長塚隆, 長山泰介, 名和小太郎, 二階堂善弘, 西澤正己, 西村健, 西脇二一, 二村健, 根岸正光, 野須潔, 野添篤毅, 野々村敏, 畑口冬彦, 花田岳美, 原正一郎, 原田隆史, 原田勝, 原芳人, 春山晓美, PALVADEAU SOPHIE, 樋口忠治, 平井歩実, 平田周, 蛭田尚子, 廣木守雄, 深見拓史, 福島俊一, 福田滋, 福室淳宏, 藤枝真, 藤田宏明, 藤田政徳, 藤田充苗, 藤則幸男, 藤本一男, 藤原鎮男, 藤原譲, 船渡川清, 船橋正美, 蓬萊尚幸, 星野聰, 細井克臣, 細野公男, 堀井健一, 真栄城哲也, 増田修一, 増田弘夫, 松井幸子, 松倉利通, 松下巖, 松田潤, 松田純一, 松田芳郎, 松本勝久, 松本紳, 的場尚雅, 間所一彦, 三木邦弘, 三品誠, 水田英実, 御園生純, 南直樹, 宮崎智, 美代賢吾, 三輪恵輔, 牟田昌平, 村上弘幸, 村上征勝, 村越一哲, 村主朋英, 村山優子, 森川弘信, 森宗明, 森田歌子, 森本敬三, 森本貴之, 師茂樹, 諸谷小四郎, 門條司, 門馬幸夫, 門馬義雄, 八重樫純樹, 安井至, 安澤秀一, 安永尚志, 安野一之, 矢野正晴, 山内繁, 山口倫子, 山崎清巳, 山崎暢也, 山崎久道, 山崎正人, 山下泰弘, 山島一浩, 山田日出夫, 月見里禮次郎, 山本昭, 山本順一, 山本隆彦, 山本毅雄, 山本晴彦, 山本昌弘, 湯浅吉美, 横井俊夫, 横田節子, 吉岡康, 吉田稔, 吉野敬子, 鷺谷正史, 渡辺辰郎, 渡邊智山, 渡辺雅仁, 弘原海清。

情報知識学会誌 12 卷 3 号 (第 7 回 SGML/XML 研修フォーラム特別号)

2002 年 10 月 22 日 発行 頒布価格 3,000 円

発行者 藤原 鎮男

編集者 根岸 正光 白鳥 裕

発行所 情報知識学会 〒110 - 8560 東京都台東区台東 1-5-1 凸版印刷 (株) 内

©2002, Japan Society of Information and Knowledge

TEL : 03 (3835) 5692 FAX : 03 (3837) 0368

URL : <http://www.jsik.jp>

この特別号は、情報知識学会が平成 14 年 10 月 22、23 日にグランドヒル市ヶ谷にて開催した第 7 回 SGML/XML 研修フォーラムのテキストとして使用されました。