

W3C Technical Plenary Meeting 参加報告

大野邦夫

株式会社ジャストシステム

W3CのTechnical Plenary Meetingは、標準化を推進するWorking Groupや調査検討を行うInterest Groupが一堂に会して情報交換を行い、最新トピックの紹介を行う年に一度の場である。今年は、2月末から3月はじめにかけて、フランスのカヌ近郊のマンデリューで開催されたので、その模様を簡単に紹介する。Meetingの最大のイベントであるTechnical Plenary Dayにおいては、Webの最新動向のトピックが紹介された。Web2.0の時代を反映してか、Microformatsのセッションが最大の関心を集め、W3Cを含む既存の標準化組織のあり方について考えさせられた。

Report and Impression of W3C Technical Plenary Meeting

Kunio Ohno

Justsystem Corporation

W3C Technical Plenary Meeting, which gives a chance to exchange information within Working Groups and Interest Groups and introduce the latest topics once in a year, was held at Mandelieu near Cannes in France at the end of February and in the beginning of March this year. This paper describes the personal view and impression through the meeting. The latest topics were introduced and discussed on the Technical Meeting Day which was the biggest event of the meeting. Among the sessions, Microformats Session was the most interesting one due to the Web 2.0 era, which was a challenge to existing standardizing organizations as W3C itself.

1. はじめに

今年の2月27日から3月3日まで、フランスはコートダジュール、カヌ近郊のマンドリューでW3CのTechnical Plenary Meetingが開催された。この町は地中海に面した風光明媚な観光地で、付近にラ・ナプール城という中世の古城があることで知られている。Meetingは、この古城からほど遠からぬSofitel Royal Casino Hotel で開催された。

W3CはWebの創始者であるTim Berners Lee博士をディレクターとするWebやXMLに関する標準化団体である[1]。W3Cは2001年から毎年一回ワーキング・グループ(WG)とインタレスト・グループ(IG)が一堂に会して相互の情報交換と情報共有を図るための会議を開催している。これがTechnical Plenary Meetingである。これまで、西暦の奇数年は米国ボストンで、偶数年はフランスのマンドリューで開催されている。

私は、一昨年ジャストシステムに入社して後、AC-Rep (Advisory Committee Representative)と呼ばれるW3Cに関する弊社の代表を担当している。今回は、定例の会議への参加というだけでなく、JAGAT (印刷技術協会) [2]と協力して検討しようとしている日本語組版技術に関する取り組みをW3Cの関係WGに紹介して認知してもらうことを目的として参加した。

日本語文書の組版に関しては、JIS-X-4051として国内的には確立している。この規格は一般的なビジネス文書、文学書等の書籍、雑誌・週刊誌、教科書、学習参考書、官公庁の広報文書、官公庁の条例文書、学術論文集、歴史書等の書籍、技術的な解説書籍、技術的なマニュアルなど、幅広い分野の文書への適用を狙ったものであるが、それをさらに携帯電話などを含む多様な表示系に拡張し国際的に認知されたしっかりした規定にすることは重要な課題である。

以前からJAGAT関係者とこの分野の重要性を議論してきたところであるが、今回のTechnical Plenary Meetingを一つの機会と考え、慶応大学SFCのW3Cメンバーを通じてW3Cの関係者と議論する機会を得たいと考えた。

2. 会議の概要

会議は、2月27日から3月3日まで5日間にわたって開催されたが、中間の3月1日は、WGやIGのミーティングは行われず、Technical Plenary Dayと呼ばれる全体会議が開催される。ここではトピックを持っているWGやIGが中心になって、W3Cの活動の概要やトピックを紹介する。各WGやIGは、この前後の月火、木金のいずれかの2日間にミーティングを行う。表1に、参加組織とそのスケジュール、登録人数を示す。

たいていの参加者は複数のWGやIGに関係しており、前半と後半の両方に出席するためにまる一週間滞在する人が多いようであった。また、Voice Browser WGのように、週の前半と後半の双方にわたってミーティングを行う組織もある。ジャストシステムがメンバーになっているWGは、i18n (Internationalization) とCDF (Compound Document Format) である。i18nは、ITS (International Tag Set) の会議が前半、Coreの会議が後半であるが、両方を兼ねている人が多い。CDFは、参照による複合文書のワーキングドラフトがファイナル・コールの段階であり、それへのコメントの吸収が主たる議題であった。

今回の日本語の組版に係るWGは、XSL, CSS, SVGとi18nなどである。これらのWGに個別に参加することを考えていたが、WG会員以外がオブザーバとして参加するためには、事前に議長の手前を得なければならない。場合によってはオブザーバを認めないWGもある。しかしこの時期各WG議長は多忙なのでなかなか時間を取ってもらえそうにない。そのような状況で困っていたら、SFCのフェリックス佐々木さんがこれらのWGの関係者と連絡を取ってくれて、28日の昼食時に関係者を集めてランチミーティングを設定してくれた。

3. 日本語組版ランチミーティング

佐々木さん経由の依頼に基づき、XSL, CSS, i18nの議長または、そのWGに関するW3C責任者がランチミーティングに参加してくれた。SVG-WGは、時間的に調整が取れずに不参加であった。用意した資料を配付して趣旨を説明した。説明のポイントは下記の2点である。

(1) Webや複合ドキュメントの見やすい日本語のレイアウトのために、現状のCSS[3]やXSL-FO[4]の規定では不十分であり、より見やすい標準的なレイアウトを必要としている。従来の印刷・出版や表示画面以外の新しい日本語レイアウト分野として、携帯電話、電子ブック、デジタルTV、ゲーム・アニメーション、カーナビの地図情報などが登場し、そのような新規分野の日本語のレイアウトに対する要求を文書として明確化したい。

(2) 上記検討を行うグループとして、JAGATを中心に出版関係、印刷関係、電子ブック関係、表示デバイス関係の専門家から構成されるタスクフォースを編成し、その結果をW3Cに提案し、関連規定に反映させたい。

上記の提案は概ね了承していただいたが、下記のコメントをいただいた。

(1) W3Cの立場としては、JAGATにW3C会員となっていたいただいた上で標準化活動に参加していただきたい。

(2) 漢字の縦書きのようなレイアウトは、中国、韓国などの東アジア圏で使われており、これらの国々にも適用可能な要求条件としてまとめて欲しい。

(3) 関係するWGが多いので、とりあえずi18nを窓口として活動を進めて欲しい。

表1 参加組織と開催日，登録者数

組織名	組織略称	開催曜日（登録人数）
Advisory Board	AB	月火（14）
Cascading Style Sheet WG	CSS	木金土（14）
Communications Team	Comm	火（9）
Compound Document Format WG	CDF	月火（20）
Device Independence WG	DI	月火（10）
Education and Outreach WG	EO	月火木（18）
Efficient XML Interchange WG	EXI	木金（30）
Evaluation and Repair Tools WG	ERT	月火（10）
HTML WG	HTML	月火（10）
Internationalization (I18N) Core WG	I18N Core	木金（8）
Internationalization Tag Set (ITS) WG	I18N ITS	月火（5）
Mobile Web Best Practice WG	MWBP	木金（20）
Multimodal Interaction WG	MMI	月火（25）
MWI Device Description WG	MWDD	木（10）
Patent and Standards IG	PSIG	月(70) 火(30)
Protocols and Formats WG	PF	木金（12）
Quality Assurance IG	QAIG	月火（12）
RDF Data Access WG	DAWG	木金（12）
Rule Interchange Format WG	RIF	月火（40）
Semantic Web Best Practice WG	SWBPD	金（35）
SVG WG	SVG	月火土（15）
SYMM WG	SYMM	木（15）
Technical Architecture Group	TAG	月(9) 火(15) 金(9)
Voice Browser WG	VB	月火木金（25）
Web Application Formats WG	Application Formats	木金（15）
Web Services Addressing WG	WSA	木金（20）
Web Services Description WG	WSD	月火（18）
XForms WG	XForms	木金（15）
XML Core WG	XML Core	木金（8）
XML Schema Patterns for Databinding WG	XML Schema databinding	月火（10）
XSL FO Task Force	XSL FO	月火（8）

従来であれば，東アジア圏では日本の影響力が圧倒的であったのだが，2番目のコメントからは，W3Cにおける最近の中国の潜在的な影響力がじわじわと感じられる。議論にあたっては，関係者と懇意のi18nのリチャード石田さんが親切に仲介してくれので，非常に助かった。ミーティング終了後

に、XSL-WG議長であるIBMのシャロン・アドラーさんから、実装依存の部分や、将来の拡張に備えなければならない部分があるので、それらの切り分けが重要とのコメントをいただいた。

4. Technical Plenary Day

3日目のTechnical Plenary Dayは、この会議の最大のイベントである。この日のスケジュールを表2に示す。表から分かるとおり、実に11時間半にわたり講演が行われた。

4.1 ネット社会におけるコンテンツの所有権

最初のセッションは、W3CのCEOであるSteve Bratt氏による歓迎の辞であった。2番目のセッションは“My Data, Your Data, The Web’s Data: Challenges of Data Ownership”というテーマによるモバイルWeb関係の人たちを中心にした講演とパネル討論で、ユビキタスネットワーク社会におけるコンテンツの所有権に関する議論であった。

Webが広範な分野で幅広く応用されるにつれ、様々な人々やコミュニティが、Webのコンテンツの作成や管理に関与するようになってきている。特にWeb2.0というキーワードは、不特定多数の利用者による情報発信が中核的な概念になっており、これらのコンテンツは重要な役割を占めている。しかし、Web上で共有されるこれらのコンテンツの知的所有権、識別される個人情報の扱いなどは必ずしも明確ではない。さらに利用者による情報発信を特徴とするWeb2.0の時代で、コンテンツが多様化している。Daniel Appelquist氏がVodafoneにおけるポリシーを説明してくれた。Vodafoneは、1億8000万人の利用者を対象にグローバルに携帯電話ビジネスを行っているだけあって説得力があった。携帯電話に関する個人情報や位置情報は、基本的には個人のものである。それを活用するためには利用者の許諾が前提になる。そのためのオプト・イン・ポリシーを設定して、企業情報、友人探索、子供の場所確認などに使っているとのことであった。

W3CのDaniel Weitznerは、ネット上でのコミュニケーションのために、将来多くのアプリケーションはFOAF (Friend of a Friend) ファイル[5]を使わざるを得なくなるかもしれないと語った。FOAFは、友達の友達の友達…という連鎖をメタデータとして表現することで、ネットワーク上の興味深い属性や関係を析でけるようにするものであるが、多くの人の連鎖で出来上がるコンテンツの知的所有権は、オプト・インで扱われざるを得ない。従って今後のネット上のコンテンツにとて重要なことは、多くの関係者のオプト・インによる許諾が重要な概念にあると思われた。

4.2 Microformats

3番目のセッションの“Microformats”は、CSS-WGメンバを中心とする議論で極めて興味深かった。Microformatsとは、表示用のHTMLに粒度の小さなXMLデータを埋め込んでサービスに結びつける方式である。既存の仕様を活用するだけで、新サービスが可能になるために、最近は様々な分野で使われつつある。特にhCardと呼ばれるアドレス帳データやhCalendarと呼ばれるスケジュール帳のようなPIM (Personal Information Management) に関するデータが広範に使われ始めている。GoogleのIan Hickson氏が、Googleが集計している統計データを用いて、Webで用いられているコンテンツで使われているタグの多くは、標準的なものではないことを指摘し、利用者は自由にタグを使いたがることを指摘した。Tantek Celik氏は、Technical Plenary DayのアジェンダをhCardでマークアップし、そのデータをアドレス帳にインポートするデモを見せてくれた。hCalendarについても同様に、Web上の案内などを自分のスケジュール帳にインポートすることが可能である。

従来のWebサービスやWebアプリケーションが、データベースとプログラムで利用者を介在させない自動化を推進するのに対して、Microformatsは利用者が参照するWebを仲介して、日常生活で用いるPIMにおけるデータ活用を指向していることが分かった。WebサービスやセマンティックWebといった高度な技術に対する既存技術のしっぺ返しのような感じである。

表2 Technical Plenary Dayのスケジュール

時間	セッション名	司会者 講演者
8:30-9:00	Welcome	Steve Bratt (W3C)
9:00-10:00	My Data, Your Data, The Web's Data: Challenges of Data Ownership	Mark Nottingham (Yahoo) Daniel Appelquist (Vodafone) Sandro Hawke (W3C) Bruno von Niman (ANEC) Daniel Weitzner (W3C)
10:30-11:30	Microformats	Tantek Celik (Technorati) Ian Hickson (Google) Tantek Celik (Technorati) Hakon Wium Lie (Opera) Rohit Khare (CommerceNet) Dan Connolly (W3C)
11:30-12:00	SQL, XQuery and SPARQL: What's Wrong with this Picture?	Jim Melton (Oracle)
12:00-12:30	Lightning Talks (A)	Paul Downey (BT) Mark Nottingham (Yahoo) Phil Archer (Internet Content Rating Association) Ignacio Marin (Fundacion) Mark Birbeck (x-port.net) Jacek Kopecky (DERI Innsbruck)
14:00-15:00	The Grid: What is the Grid and Why Does it Matter to W3C?	Ian Jacobs (W3C) Dave Snelling (Fujitsu) David De Roure (University of Southampton) Don Deutsch (Oracle)
15:00-16:00	Adventures in Formal Methods	Jonathan Marsh (Microsoft) Jerome Simeon (IBM) Dan Connolly (W3C) Charlton Barreto (Adobe) Arthur Ryman (IBM)
16:30-17:00	Rich Web Application Backplane	Kevin Kelly (IBM) John Boyer (IBM) Steven Pemberton (W3C) Mark Birbeck (x-port) Al Gilman (?)
17:00-18:00	Lightning Talks (B)	Paul Downey (BT) Steven Pemberton (W3C) John Boyer (IBM) Jonathan Marsh (Microsoft) Paul Downey (BT)

Hakon Wium Lie氏が、HTML+CSSの世界で、書籍の製本のためのmicroformatsについて講演してくれた。既存のCSSの枠組みをmicroformatsで拡張するのは良いアイデアかもしれない。Rohit Khare氏は、microformatsとRSSとを組み合わせせて開拓される世界の可能性を指摘した。XMLによりWebは広範な世界に広がったが、その広がりが今日新たな問題を生じさせており、その問題をRSSとmicroformatsが解決してくれるとのことであった。Dan Connollyは、セマンティックWebの構築には、GRDDL (Gleaning Resource Descriptions from Dialects of Languages)とmicroformatsを活用す

べきであると語った。GRDDLは、XHTMLからデータを抽出するための言語で、これとmicroformatsを適用して意味的なデータを取り込むことができ、従来のOWLなどによるセマンティックWebよりも、データ活用などの点で実用的であるとの指摘であった。

4.3 Query言語の比較など

4番目のセッションは、"SQL, XQuery and SPARQL: What's Wrong with this Picture?"というタイトルで、RDB用のSQL、XMLデータ用のXquery、RDFデータ用のSPARQLについて比較検討している。セマンティックWeb、RDF、OWLなどの要請でSPARQLが検討されているが、RDFはタプルの集合であるので、SQLに関連付けることが可能である。従ってSPARQLはSQLに変換可能であり、RDBとの相互運用が考えられるとのことである。さらにSPARQLは、工夫する場合にはXqueryにも変換可能になるとのことで、SPARQLの可能性の講演であった。

5番目のLightning Talkは、中心的なテーマを決めないで短時間に多数のトピックを紹介するセッションである。"Services and the Web" (Mark Nottingham)、"XG-1" (Phil Archer)、"SVG Statistics Demo" (Ignacio Marin)、"SVG and XForms Processing Language" (Mark Birbeck)、"Standard RDF mappings for everything?" (Jacek Kopecky)というタイトルの講演やデモが行われた。その後昼食時間のための休憩となった。

4.4 WebへのGrid適用

6番目のセッションは、"The Grid: What is the Grid and Why Does it Matter to W3C?"というタイトルで、WebにおけるGridの議論があった。Web上での大規模な処理を必要とする分野が顕在化してきたことがその背景である。Dave Snelling氏によると、Grid技術は、並列化、クラスター化といった処理の延長上の技術とのことである。企業においては、DBがらみのコラボレーション処理などでGridの適用が検討されている (Collaboration Grid)。さらに情報の可視化、セキュリティなどの要請からGrid技術の標準化も要請されている。そのためにGlobal Grid ForumやEnterprise Grid Allianceといった組織が立ち上がり、企業向けGridのに関する取り組みがなされつつある。今後のGridの用途は、企業に閉じた世界から、企業横断的な世界に拡張されると考えられ、そのためにはW3Cによる標準化も視野に入れておく必要があるとのことであった。さらに今後は、Semantics, Heuristics, Formal Languageのような取り組みと連携するとのことであった。

David De Rour氏等によると、Grid技術は、データGrid、情報Grid、知識Gridの過程を経て進歩しているとのことである。知識Gridの用途に関しては、セマンティックGridとしての意味的な処理への取り組みが期待され、WebサービスとセマンティックWebとにGrid技術が貢献していくとの予想であった。この分野に関しては、今のところ研究開発段階で、ベストプラクティスのような事例の出現が望まれている状況とのことであった。

Googleの台頭からも明らかなように、Web上のコンテンツの検索サービスやそのための系統的かつ有機的なインデックス作成などが今後のニーズとして考えられるであろう。それをさらに活用するためのオントロジの作成や管理などにセマンティック Gridは活用されると考えられる。

4.5 XMLへのデータ型適用

7番目のセッションは"Adventures in Formal Methods"という数学や論理学の素養のある言語屋さん以外には馴染みの薄い分野に関する講演であった。Jerome Simeon氏による "The XQuery1.0 and XPath2.0 Formal Semantics" では、昨年末に勧告候補となったタイトル名の規定に関する説明であった。XQueryとXPathの組み合わせを、関数型言語とみなしXML Schemaで定義される型を拡張して定式化している。XQuery自体はXMLではなく独立した言語なので、このような定式化が必要

なことは分かるが難解である。Dan Connolly氏は、"DTD/UML/OWL"とい講演で、HTML/XML関連の情報構造の記述方法に関する考察を述べたものであった。SGMLのDTD記述は効果的かという問題設定から出発し、情報構造の記述が項目と矢印で関係付けられるというプリミティブな議論に掘り下げて、UML記述、さらにOWLの情報構造を見直すというアプローチであった。

Charlton Barreto氏の"The π of Choreography"は、Webサービスにおけるフォーマル・メソッドの適用に関するものである。W3CにWeb Services Choreography WGという組織があることを知らなかったが、そこでの活動の紹介も含めた講演であった。WS-CDL (Web Service - Choreography Description Language) という、Webサービスにおける処理階層やワークフロー記述を定式化するもので、そのモデルに π -calculusを適用しようとするものであった。Arthur Ryman氏とHugo Haas氏による"Using Z Notation to Specify WSDL 2.0"は、WebサービスのWSDL2.0の規定のモデルをZ記法を基に構築する試みを解説したものである。Z記法自体は、ラッセルのパラドックスを回避可能な型集合理論に基づくもので、身近な例ではLaTeXの仕様記述で用いられているとのことである。Z記法をWSDL2.0に導入するメリットは、型検査が可能になり、実装以前にエラーチェックが出来ることであるとのことであった。

4.6 Rich Webなど

8番目のセッションは"Rich Web Application Backplane"ということで、リッチクライアント化を指向するWebアプリケーションに関するものであった。John Boyer氏による"Intelligent XML Data Model"という講演は、クライアントとクライアントに近い処理を行うサーバ・ミドルウェアを知的なXMLオブジェクトと見なし、データベースのような単純なサーバ機能に対する複雑な処理を実行するサブシステムと考えるモデルである。このモデルは複合文書だけでなくWeb2.0におけるAJAXにも適用可能とのことである。Steven Pembertonは"Events"というタイトルの講演で、DOM3をはじめとするイベントが、リッチクライアントに対するバックプレーンの役割を果たすことを述べ、種々のイベントの体系付けを提案した。

最後のセッションは、再びLightning Talkで、"What if a rose didn't have a name?" (Steven Pemberton), "Signatures and Namespaces" (John Boyer), "Single Transferable Vote" (Jonathan Marsh), "An Audience with Chad" (Paul Downey) というタイトルの講演があった。

以上のように、W2.0に関係する世の注目を集めるトピックから、型に関する理論的な検討に至るW3Cの幅広い分野における活動の紹介が行われた。全てのセッションの終了後に、最も興味深かったセッションの投票が行われたが、それは3番目のセッションの"Microformats"であった。Microformatsは、既存の規格の膨大さや標準化プロセスへの批判や反省から出てきた産物である。そのセッションがW3Cという代表的な標準化団体で最も関心を持たれたのは一つの事件であった。

5. 技術動向と標準化状況

以上、いくつかのWGとTechnical Plenary Dayに参加して感じた印象を述べたい。まず、ユビキタスネットワーク化でXML関連技術が企業情報システムからモバイル環境や家庭環境に移行しているがW3Cにおいてもそれは顕在化している。MWIが昨年設立され、携帯電話などにおけるWebの世界にもXMLが導入されつつある。この状況はCDF-WGにおいて特に顕著で、携帯電話画面の複合ドキュメント標準化が推進されている。複合ドキュメントは、XHTMLにSVGを統合しリッチマルチメディア・コンテンツとしてユーザインタフェースの向上を目指している。この動向はリッチクライアント化を推進するもので、先の8番目のセッションで取り上げられたテーマに関係する。リッチクライアント化の動向は、Web2.0におけるAJAX活用なども軌を一にするもので、今後のクライアント環境の一つの流れを形成するであろう。他方、企業に於いては、端末におけるデータの保存は情報漏洩のリスクとなるので、常時接続環境に於いてはリッチクライアントよりはシンクライアントを指向する動きもある。ただしWeb2.0の動きを見ている限りは、リッチクライアント化が今後の動向であろう。

他方、セマンティックWebと連携したWebサービスの高度化も一つの動向である。後半のセッションで話題になったGRID技術やフォーマル・メソッドの導入はそのあたりの状況を物語るものであろう。これらの技術は、専門家にとっては興味ある技術分野ではあるが、すぐに実用的に使える技術にはなりそうもない。XML Schemaのデータ型を確立し、XMLの世界を型で体系化することは興味深い、それが実用的に有効になるためにはかなりの時間を必要とすると思われる。ハードウェアをベースとする技術進歩は、それを待つてはくれそうもない。

最近Webサービスで用いられているRESTは、標準化団体が決めたものではなく、大学院生が修士論文のために開発した技術であるとのことだが、単純で使いやすいためにあっという間に普及してしまった。単純な技術やその組み合わせ技術は、標準化プロセスなど経ないでWeb上の情報交換で迅速に普及してしまうのがWeb2.0時代の特徴である。そのような意味では、標準化プロセスの意義を問うようなMicroformatsに関心が集まったのは、自然なことでは驚くには当たらない。標準化団体はより迅速な標準化プロセスを確立する必要がある。しかし、それが容易ではないことも事実である。

6. おわりに

以上、Technical Plenary Meetingにおける、日本語組版に関するミーティングと、Technical Plenary Dayに関する状況を紹介した。GoogleがWeb上の検索ツールの事実上の標準として君臨し、マイクロソフトに次ぐ大企業として成長しつつある現在、Webは便利なネットワークといった位置づけから、企業のビジネスを制する基幹インフラとして変貌しつつある[6]。Web進化論[7]の著者である梅田望夫氏は、Web上で衆知を集めるビジネスのみが今後成長するとまで述べている。そのWebに関する標準化団体であるW3Cが、現在何を重視しているかを知るためにはTechnical Plenary Meetingは良い機会であった。

Web2.0というキーワードが飛び交い、Googleが提供するAPIとAjaxでアプリケーションを開発する動きが顕在化しているが、これらは今後のWebの世界の変貌の発端に過ぎないであろう。他方、上記のような電子メディアの変貌を経過してもなお既存の印刷メディアも存続し続けることであろう。日本語組版というような地味な分野のノウハウを記録に残し標準化して電子的メディアに継承してゆくことも重要な仕事であると考え次第である。

参照情報および文献

- [1] <http://www.w3.org/>
- [2] <http://www.jagat.or.jp/>
- [3] <http://www.w3.org/Style/CSS/>
- [4] <http://www.w3.org/TR/xsl/>
- [5] <http://www.foaf-project.org/>
- [6] 佐々木俊尚; “グーグルGoogle一既存のビジネスを破壊する”, 文芸春秋新書 (2006)
- [7] 梅田望夫; “ウェブ進化論 本当の大変化はこれから始まる”, ちくま新書 (2006)