

第2回 SPARC Japan セミナー 2008

「学術出版と XML 対応 - 日本の課題」

事例紹介 2

化学系ジャーナルの場合 -WordからXMLを作る試み-

林 和弘 (社団法人 日本化学会)

国立情報学研究所

講演要旨

化学系の電子ジャーナル作成の事例として日本化学会が検討を進めている XML を利用した出版について、その経緯とともに紹介する。特に SGML、TeX、XML の運用経験からメタデータを作成しながら組版を行う際の諸問題について触れる。

講演者プロフィール



林 和弘

日本化学会学術情報部課長・
SPARC Japan 運営委員

1994年東京大学大学院に大学院生として在籍時、日本化学会の論文誌査読管理データベースを開発したことがきっかけとなって、日本化学会の論文誌の電子ジャーナル化に取り組む。化学の研究者の立場がわかり、かつITスキルを持った氏は日本発の学術情報発信について、電子投稿査読、XML出版、電子ジャーナルプラットフォーム構築、宣伝活動の広いフェーズで実務に基づき考察と改善を加えた。その活動は、結果として日本化学会 Chemistry Letters 誌を一般化学誌として世界最速クラスで発行する電子ジャーナルに整え、化学系学会出版としては世界でもいち早いオープンアクセス対応を開始するなど、数々の実績に反映されることとなった。また、そのノウハウは SPARC JAPAN、J-STAGE 等各種のプロジェクトに反映され、科学技術・学術審議会下の WG 専門委員も務めて日本の学術論文誌のあるべき姿を提言し、現在も日本発の論文誌をより魅力的にするための考察と改善を繰り返している。さらに、文部科学省科学技術政策研究所客員研究官他を兼任し、電子ジャーナルの将来とインターネット時代の科学コミュニケーションのあり方そのものについても興味を持つ。

メタデータ (XML) を利用した出版のキモ

(物理系のジャーナルに引き続いて) 化学系ジャーナルの場合の事例として、Word から XML を作る試みを続けているので、それを紹介します。

まず、ここでいうメタデータ (を利用した) 出版の対象は、学術ジャーナルを指します。そして、メタデータ出版のキモは、設計と運用です。

設計は、簡単に言えばどの程度タグを付けるべきか、また付けたタグが最終的にどのように利用されるかも前提において (DTD やクラスファイルなどの基本設計を行い) タグ付けを行わなければならない、ということです。

また、運用で重要なことは、誰が、どこで、いつ、タグを付けるか、ということと、そのコスト効果も常に気を付けながら考えるということです。

タグ付けには、次のような性質もあります。タグが多いものから少ないものの変換は比較的容易ですが、その逆は難しいということです。

さらに、タグ付けにおいて、学術ジャーナルのような不特定多数の著者にタグ付けをお願いしても、そのクオリティは大概信用できないという問題もあります。

(図2) は組版ソフトを構造化文書と誌面自由度の観点から見て、簡単なマトリックスを書いたものです。縦軸はデータ構造化の高低、横軸が見た目の対応の高低です。

マトリックス的に見ると、この DBside からみた矢印は、SGML という高度に構造化されたデータから、だんだんと誌面の見た目をよくしていく代わりに、多少の構造化を犠牲にする場合もあるという妥協の取り方をします。Publisher side のほうから見た場合は、見た目の対応が高い Word、電算写植、Quark などから、Quark のメタデータアドオン利用など構造化文書を取り扱う方向へ進めながら、妥協点を探すということになります。図には入れ忘れていますが、InDesign などでもそこ (現在の現実的な妥協点) に入ります。

今のところどちらのアプローチも何かを犠牲にしていますので、この取り組みの先に、構造化を高度に維持しつつ見た目の対応も十分に高い「ユートピア」の組版体制が出てくるかどうか今後の展望となりますでしょうか。

SPARC JAPAN

メタデータ (XML) 出版のキモ

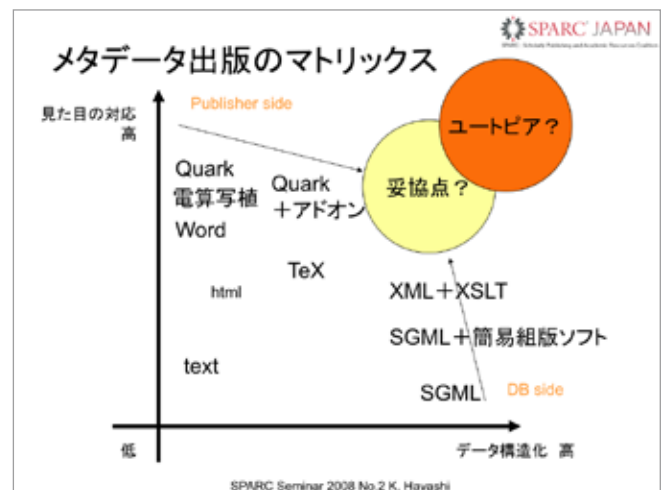
- 設計: どの程度タグをつけるべきか、つけたタグは最終的にどのように利用されるか
- 運用: 誰が、どこで、いつ、タグをつけるか

Tips:

- タグが多いものから少ないものの変換は比較的容易だが、逆は難しい(面倒)
- タグ付けにおいて著者は信用できない(不特定多数の思想のばらつき)

SPARC Seminar 2008 No.2 K. Hayashi

(図1)



(図2)

SGML 出版の取り組み

(図3)は、2002年の『情報の科学と技術』の論文に書いたものですが、まず、日本化学会では、1989年からSGML(を利用した)出版に取り組みました。SGML出版の特徴はワンソースマルチユースで、ひとつのソースから版下や電子ジャーナルファイルなどを作り分けることができることでした。しかし、実際にやってみると、著者校正で修正が入る段階がありますので、最終的にデータがフィックスされるのは、ソースの段階ではなく、誌面に展開し訂正が入ったあくまでプルーフ上です。ここがワンソースマルチユースの理想と現実だったわけです。つまり、SGMLからプルーフを一方に吐き出しているのも、もし版下に訂正があれば、元のSGMLに戻って、TeXのコンパイルと同じように何度も版下を吐き出さなければならず、出版が間近に迫っていると、版下だけを直してもとのソースであるSGMLを直す暇がないケースがでてきました。このため、最終的にできたPDFとSGMLのデータが違うというようなことが煩雑に起こり、その他の要素を含めて非常にコスト高な構成になっていました。慎重に検討し、SGMLを利用した出版は2001年で一旦止めて、IPAPさんと同じシステムを導入しました。

これは太宰さんがTeXと3B2と呼ばれる組版ソフトを利用して開発されたもので、このTeX-3B2は、今はArbotext Advanced Print Publisherと名前が変わりましたが、組版上のシステムの中で版下とメタデータの変換の同期をとることによって、プルーフを直した瞬間に元のメタデータも直る形をとります。ソースに戻って直す必要がありません。このような段階取りで初めていわゆる“One source multiuse”でやりたかったことが実現されたと見ることができると思います。

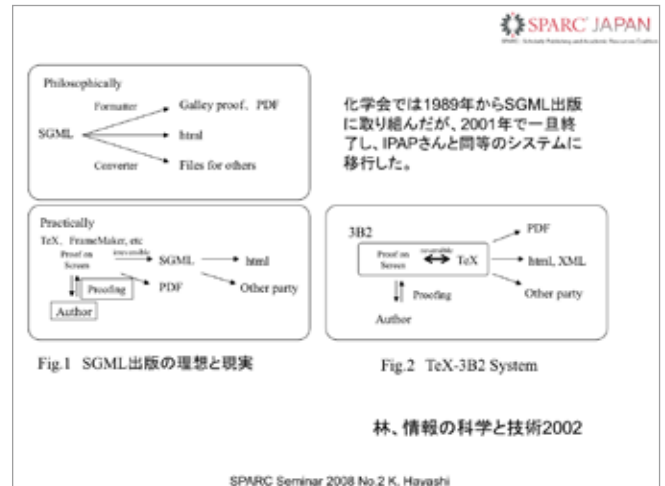
この体制で出版を開始したときに、本当によくあった質問を挙げてみますと、面白いことに、化学系の人はずいぶんTeXでないと駄目なのかと、必ず聞かれました。そして、2番目に多いのは、Wordでは駄目なのか、ということでした。(図4)

情報系の専門家からすれば、TeXとWordと分けて、あるいは同列の扱いで聞いてくるのはおかしいと思われるかもしれませんが、これが(少なくとも化学の)研究者の実態です。

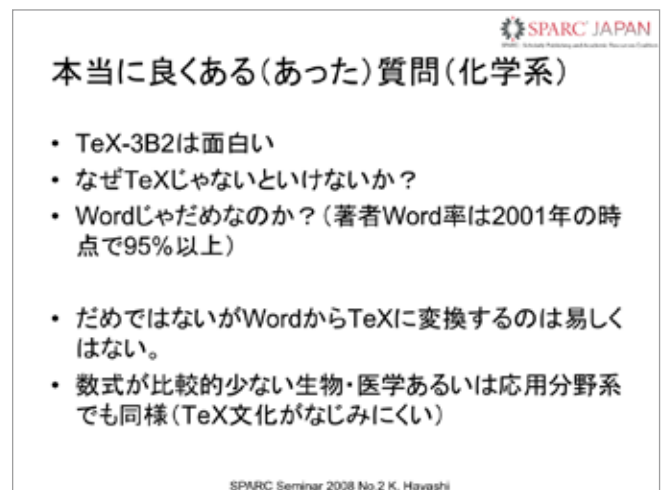
我々のジャーナルも、Word投稿率は2001年時点で95%以上、今ではもう99%以上と言っていいでしょう。もちろん、TeX投稿でも駄目ではありませんが、TeXからWordに変換するのは、そう簡単ではありません。

こういった状況は、化学に限らず、生物や医学、応用分野などでも同様で、TeX文化がなじみにくいようです。(念のため)TeXのことを非難しているわけではなく、こういう研究者がいる分野が、我々のジャーナルパブリッシングに存在しているという現実を紹介しているだけです。

むしろ、TeXが非常に少ないリソースでも高度な組版表現をできる、大変素晴らしい言語だということは間違いのないことです。



(図3)



(図4)

いずれにせよ、Wordから初期TeX、太宰さんが(前の講演で)言われている基本ファイルのTeXまでは、労働集約的処理で、当初はWordからテキストファイルに落としてから、TeXでタグ(コマンド)を打っていました。(図5)

途中からWordからHTMLコンバートのような形で、多少のスタイルを保持してTeXに変換処理を可能にしたのですが、いずれにせよ、人件費の安い地域を利用して労働集約的に行っています。そして、化学会は、TeX-3B2を利用した出版を確立し、版下作成コストを大幅に削減したということ、2003年に掲載された情報管理誌でも報告しています。

これは、非常に現実的なシステムであり、実際にこの基本ファイルのTeXからXMLの出力もLaTeXのコマンドの範囲プラスアルファくらいまでなら、それほど問題なく作成できるということまで来ました。

しかしながら、安定運用を行った上でも、Wordから直接XMLを出力することはできないのか、そうしたニーズがどうしても化学会の中にあるわけです。(図6)

そしてSGML出版として見た場合も、前述したように、2001年で一旦やめてはいますが、理念としては間違っていないわけで、ただ、やるのが早かったわけです。実際、SGMLの欠点を補ったとされるXMLの登場と利用があって世界を含めて今に至っているわけですが、このような動向について、2002年以降も継続調査をしているうちに、一つの解決候補としてeXtylesというツールが登場したことを知りました。WordからXMLに吐き出せるツールができたのです。しかし、当時はなるべく少ないコストで必要なメタデータができて、最終的に必要な電子ジャーナルサービスができるのであれば、それで構わないではないかと考えました。また、先ほど紹介したTeX-3B2の現実的なシステムがちょうど安定して間もないころでしたので、まずは様子を見てみようという状況でした。

そして、そのうちに様相が多少変わってきたと感じさせたのが、NLM - DTDが登場して浸透してきたことです。

XML、SGMLの一つの欠点に、(学術出版系)業界標準のDTD(設計書)が存在しないということがあったのですが、ここへ来て、その候補が出てきたわけです。決定的だと思ったのは、Porticoプロジェクトと呼ばれる、大手出版社のアーカイブデータを共同で保存しようというプロジェクトの標準DTDに採用されたことです。

これは分野、商業出版者、学会出版者を問わず動いているプロジェクトで、これによって、アーカイビングが、NLM - DTDのXMLを使って保存しようということ動いているものでした。この動向を鑑みて、我々のデータも、アーカイブとして後世に残すためには、この業界標準のアーカイブプロジェクトのXMLを作っておけばよいだろうと思いました。何が何でもこのXMLでなければいけないという理由ではありませんが、やれるのであればその当時の標準に合わせておいてもいいだろうということです。

また、組版システムのバージョンが上がってXMLに対応しやすくなったのも再検討の理由の一つでした。

SPARC JAPAN

Word-(Html)-TeX-3B2システム (WHT3)

- Wordから初期TeXまでは労働集約的処理
- 最初はテキストに落としてからコマンド(タグ)を直接打っていた
- 途中からhtmlコンバートで多少のスタイルからTeXへの変換処理を可能に
- いずれにせよ人件費の安い労働を利用
- 化学会としてはメタデータ出版を確立しつつ、版下作成コストを大幅に削減した(林、情報管理2003)
- 非常に現実的なシステム XML出力も可能

SPARC Seminar 2008 No.2 K. Hayashi

(図5)

SPARC JAPAN

Wordから直接XMLを出力できるか

- SGML出版も理念としては正しかったので、SGMLの欠点を補ったとされるXMLの利用については2002年以降も継続調査
- 一つの解決案としてeXtyles(2004登場)
- 当時は前述WHT3のシステムが安定して間もない頃のために、存在を認識しながらあくまで様子を見ていた。
- NLM-DTDの登場と浸透
- PorticoプロジェクトによるNLM-DTDのデファクトスタンダード化
- 組版システムのバージョンアップによるXMLの親和性アップ
- 再検討

SPARC Seminar 2008 No.2 K. Hayashi

(図6)

eXtyles とは

ここで eXtyles について説明します。(図 7)

eXtyles は、ボストンにある Inera 社が開発した、MS - Word のプラグインソフトで、単独のソフトではありません。任意の著者ファイル (Word) から始めて、組版やデータベース作成に要らないデータを取って、パラグラフ単位でスタイル付けをすることで、そのスタイルと基本的なメタデータ、XML タグが連動しています。

また、パターンマッチングによって、従来、人の手で同じように何度も直していた編集作業を、完全自動ではありませんが、半自動編集できるようになりました。

さらに、ここから先はツールのオプション機能も入りますが、まず、リファレンスの構造を見て、どれが著者で、どれがジャーナルタイトルで、どれがページで、というものを解析してタグ付けをします。それを基に、リファレンスチェックを行います。我々が導入しているものは PubMed と CrossRef ですが、このデータベースに当てて、引用文献のリファレンスが存在するかということ電子ジャーナルリンクの有無で見るわけです。これは地味ですが重要なことです。

他にも、例えば、図表引用が Figure.1 と書いてあれば、本当に Figure.1 があるか、Reference.3 があれば、後方に Reference.3 があるかということも確認します。さらに同時に、順番が逆になっていたら勝手に直してくれます。1、2、4、3、5 であれば、1、2、3、4、5 に直してくれるわけです。このような整形を行いながら、最終的なメタデータ生成 (要素を入れることとその中のコンテンツ作成) に必要な工程を、一つの Word の中で行います。NLM - DTD の XML が生成できるソフトウェアということで、この導入を進めています。

では、eXtyles 導入のメリットは何かというと、まず、ある時点で NLM - DTD に基づいた質の高い XML の出力が可能と


なるということです。(図 8) その XML を出すタイミングが重要なのですが、そこは今後検討課題だと思います。また、このツールが NLM-DTD に準拠した XML を吐き出せる保証に関しては、Inera 社 CEO の Bruce Rosenblum 氏が、実は NLM - DTD の共著者であり、その彼が作ったツールという事情もあります。NLM-DTD の仕様は公開されておりますが、その NLM - DTD の仕様書を読んで一から自分の雑誌のデータを作るのは、なかなか大変なことなのです。その作業を省くことができます。

次に、手入力で行うと非常に非効率な XML タグ付けが、Word 上で簡単にできるというメリットがあります。

TeX に比べ、XML の欠点は、手入力する際に、例えば、後ろのコマンドはいちいち要素名を入れなければならないなど、手間がかかり、それを労働集約的な作業によって行えば、確実にコスト (人件費) の負担になるところでした。それが解消されたこととなります。

また、初期 XML 出力までが非常に早いというメリットがあります。Word 上で全て処理して最後に出せますので、我々のジャーナルのものでは、3 カ月分くらい流してみてもわかったのですが、図表が少なければ、著者の Word から数時間以内に出力することが可能です。場合によっては、そのまま XML から XHTML で、超早期公開という展開もあり得ます。そうしたことも含めて、全文データ作成が容易ということです。

そのほか、XML を作ることは直接関係ありませんが、Word 上の編集なのでヒューマンリソースの確保がしやすいことがあり、また、中のコンテンツを学会編集側で正しくすることもできるので、校正が全く必要なくなるわけではありませんが赤入れ担当者の負担を軽減できる可能性があります。リファレンスの多重チェックも可能です。




eXtylesとは

- Inera社開発のMS-Wordのプラグインソフト
- 任意の著者ファイルから始めて
 - 不要なデータの削除
 - パラグラフ単位でのスタイル付けとタグ付け
 - パターンマッチングによる半自動編集
 - リファレンスの解析とタグ付け
 - リファレンスチェック
 - 図表引用、リファレンス番号の整合性チェック

→最終的なメタデータ(要素とコンテンツ作成)生成に必要な工程をWord内で行い→NLM-DTD XML

SPARC Seminar 2008 No.2 K. Hayashi

(図 7)



eXtyles導入のメリット

- (ある時点で)NLM-DTDに基づいた質の高いXMLの出力が可能 (Inera社CEOのRosenblum氏はNLM-DTDの共著者)←一からDTDに合わせてXMLを作成するのは大変
- 手入力では非常に非効率なXMLタグ付けが容易に(閉じタグの入力の手間 TeX] XML</要素名>)
- 初期XML出力までが早い(図表が少なければ著者Wordから数時間以内に出力することも十分に可能)→超早期公開
- 全文データ(html, xhtml)作成が容易
- ヒューマンリソースの確保がしやすい(Word処理)
- 学会編集側で処理できる可能性もある
- 赤入れ担当者の負担軽減(なくなりはない)
- リファレンスリンクの多重チェック(PubMed, CrossRef)

SPARC Seminar 2008 No.2 K. Hayashi

(図 8)

では、デメリットは何かというと、当たり前ですが、著者の Word 利用が多くなると力を発揮できないということです。(図 9)

また、一つの OS ツールに依存することによるリスクが当然あります。この辺りは、著者や業界の動向を常にウォッチしながら、柔軟に対応していく必要があります。

一方、著者原稿の当たり外れによって、前処理の負担が大きい場合があり、ツールの入れ甲斐がないと感じる原稿も出てくる場合があります。また、TeX ファイルを労働集約で人海戦術的に行ってやっていたものを、このツールを導入することにより、安い人件費を使う作業にしにくいところができて、導入コストと併せて、これらコストの問題もあります。さらに、カスタマイズに一時的に負荷がかかるという問題もあります。例えば、便利にしたい、特に編集効率を上げようとして、コンピュータになるべく赤入れさせようとするると負荷がかかり、他の作業に影響を与えることもあります。しかし、これらはどのようなツールを導入しても、一部を除けば、ある程度は想定されることなので、とにかく実施して知見を得ることにしています。

また、このツールをこの段階で導入しようと思った一つの理由は、コンピューターリソースの変化にあります。(図 10)

PC の性能が各段に向上してツール類が充実してくると、従来、面倒でかつ、マシンパワーとマンパワーを必要としていた作業が手軽に行えるようになったことが XML 化を進めました。TeX に比較して、煩雑ともいえる XML のタグ付けも可能になったわけです。

また、前述のように、Word 上で全てできるので、印刷会社のオペレーターも含めてユーザーサイドが、タグ付けをあまり意識しなくとも操作できるようなインターフェース作りが可能になってきました。

SGML 出版の検討段階で理想的な組版とデータ作成が実現できると言われてから 20 年以上を経て、ようやく「素地ができたと言えるかどうか」ぐらいのところに来ていると思います。しかし残念ながら、それはまだ、断言できません。



eXtypes導入のデメリット

- 著者のWord利用が多くなると力を発揮できない
- 1OSの1ツールに依存することによるリスク
→著者や業界の動向は常にウォッチ
- 著者原稿の当たり外れによる前処理の負荷
- 労働集約的な作業にしにくい(今のところ)
- カスタマイズに一時的に負荷がかかる(便利にしたいなら)
- 導入コストの償却
→これはどのようなツール導入でも同じ

SPARC Seminar 2008 No.2 K. Hayashi

(図 9)



コンピューターリソースの変化

- 昔に比べてPCの性能が格段に向上
- ツール類の充実
- 従来面倒でかつマシンパワー&マンパワーを要していた作業が手軽に=TeXに比較すると煩雑ともいえるXMLのタグ付けも可能に
- また、そもそもユーザーがタグ付けを意識しないインターフェース作りが可能に
→ようやくXMLを実際的に効率よく作成する素地ができたと言えるか？

SPARC Seminar 2008 No.2 K. Hayashi

(図 10)

Word (→ eXtyle) → XML → 3B2

以上のような検討を重ね、Word のプラグインとしての eXtyle を利用して、そこから XML を吐き出し、組版ソフトに流すというオペレーションを、試験フローとして確立しています (図 11)。

その中で、新しい赤入れと入稿、新しいメタデータ生成など、システムフローに応じて新しいワークフローを作って、それぞれのオペレーターの役割の再確認を行いました。その各オペレーターの位置づけを示した簡易フローが (図 12) です。

化学会の場合、著者から事務局へ原稿が来ると、化学会事務局の赤入れ担当が赤入れを行います。赤入れされたファイルと図ファイルなどを、FTP Server を介して印刷会社に渡します。本文に関しては Word で処理されるので、その eXtyle を流す担当者がいます。

一方、画像ファイルの処理もノウハウのいる非常に重要なところで、担当者がそれを行います。画像とテキストファイル、メタデータとテキスト本文ファイルを合わせて組版ソフトに流し込み、そのときに数式処理をして、最終的に XML を吐きつつ、今は J - STAGE を使っているので、XML ファイルから BIB - CIT に変換しています。

冊子のほうは今まで通りに出し、それを工程管理が担当しています。

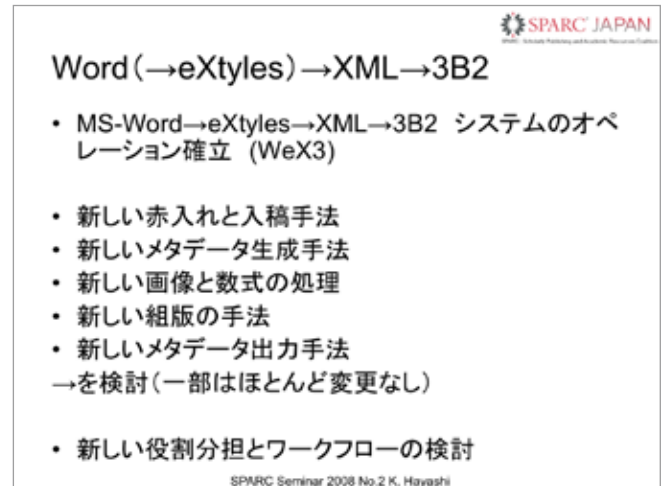
点線の矢印は、データ作成やリファレンスのチェック、版下作成の時の赤入れ担当とやり取りしながら行っていることを示しています。

(図 13) は、XML を出力するところにもう少し注力したものです。

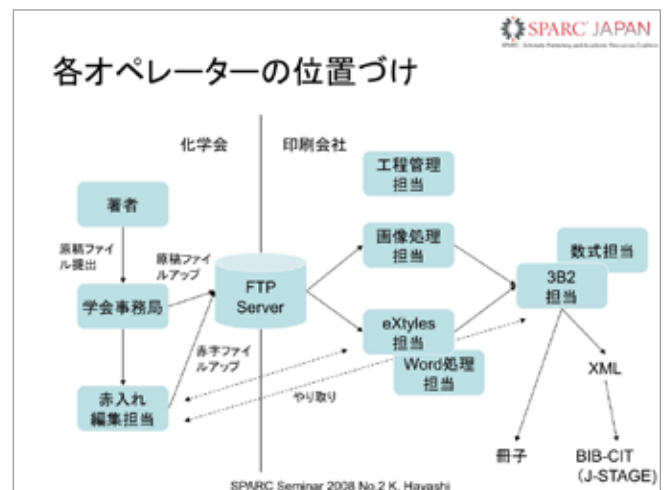
著者の原稿だけ使って Word with eXtyle で XML を作ったとき、実は足りない情報がいくつかあります。例えば、受付番号や、場合によってはページ、あるいは DOI など、学会側 (パブリッシャー側) で発行のために必要な情報があるので、それとマージさせることによって、最初の XML ができます。また、数式処理に関しては特別な処理を施しているのので、それを合わせて、初めて組版用の基本ファイルの XML ができます。

(図 13) の点線で囲んだ右側の青い囲み、3B2 をコンピュータの画面だと思ってください。ここを直しますとその左側の元のファイル (2nd XML) も直りますので、そこから最終的にフィックスした対外的に回せるような XML ができます。

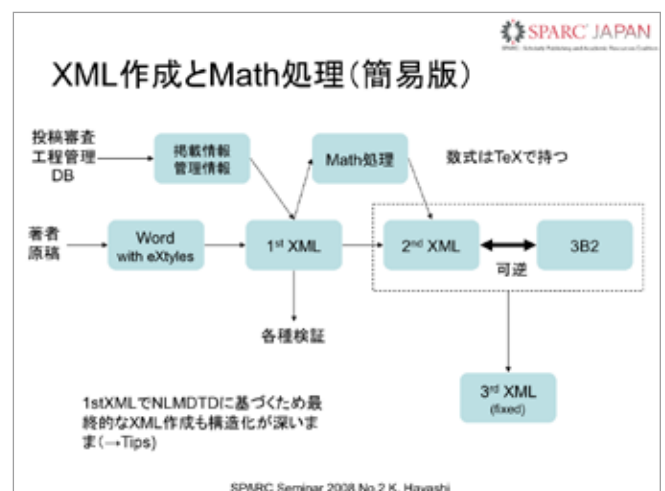
ここで重要なことは、1stXML が NLM - DTD に基づくということです。1stXML の時点で比較的高度に構造化された XML ができるので、そこから出力する分にはそれほど難しくありません。1stXML が簡単なままだと、例えば、タグ要素を増やす、構造を階層化するなどの処理が必要になってくるため、最初の段階で 1 度きっちりと行っておくのが良いと思います。ただ、この場合、ここで出力されるものが必ずしも NLM-DTD に厳密に基づいている必要はなく、高度に構造化されていれば良いので、(NLM-DTD に基づくというのは) 後で付いてきたメリットと言ったほうが正しいかもしれません。



(図 11)



(図 12)



(図 13)

課題と展望

とにかく、XML ベースで進めることがどこまで有効かを見極める (XML ありきにならない) ことが大切です。

特に、英文誌の学術出版に関しては、欧米と比較してコスト対効果をどこまで出せるかといった事業効率の問題もあります。というのも、他の大きな出版社が、どのように SGML、XML 出版をしてきたかという、学術プランテーションという言葉も先の講演で出ていましたが、煩雑なタグでも何でも、手間暇かかるものは全て、時間がかかっても人件費の安い国、地域で行って (オフショアビジネス)、ビジネス効率を出す、そういうやり方で SGML を組んでいたわけですから。それを、このようなツールを利用して国内、インハウスで作った場合に、どこまでコスト対効果を出せるかということが課題になります。

また、まずは全文データの XHTML をつくることで、XML の実際的な利用にしたいと思っています。

よく日本語のニュースサイトなどで、冒頭の 4 行くらいが書かれたあとに「続きを読む」や「記事全文」などある場合があります。その「続きを読む」のファイルが PDF であつたらどうでしょうか。日本人が日本語のニュースなど新着の情報に関しては、全文 HTML のまま読み切ってしまうと思うのが普通でしょう。アナロジーを説明するのが分かりやすいと思ひ、ニュースサイトの例を出しましたが、今の日本の学術ジャーナルは、欧米の英語圏のリーダーから電子ジャーナルを見た場合、先のアナロジーのような状態なのです。こうしたビズビリティの差を埋めなければならないと思います。分野やそれぞれの雑誌が置かれた状況にもよりますが、私個人としてはこのように考えています。

展望としては、とにかく NLM - DTD で本格的な XML を作る、恐らく日本で最初の例だと思うので、その浸透具合を見てみたいと思います。また、どうしても TeX 文化になじまない分野でジャーナルへの展開ができるかどうかということも見ることにあります。あるいは和文誌の対応の検証も同様です。

NLM - DTD は非常に優れていて、マルチ言語対応になっています。例えば、ルビ対応などが容易にできるかなど、日本のローカリズムにどこまで対応できているのかをちゃんと調べていませんが、学術ジャーナルレベルであれば、和文誌対応も可能ではないかと思ひます。

さらに、実はこれが XML を使う大きな理由なのですが、投稿、査読システムなどの本格的な連携です。前述のリファレンスチェックなども、本来投稿段階で行いたいところなのです。XML のタグ付けも、著者にやってもらうのであれば、最初の投稿段階で行ってもらいたいと思ひていたのですが、Editorial Manager と呼ばれる世界最大の電子投稿査読プラットフォームが、すでに eXtyle のエンジンを導入したので、実はもう裏では、投稿の段階でリファレンスチェックが進むようになっていきます。

同様に Editorial Manager 派生の電子投稿査読システム

(ESS) を使っているエルエビア社のジャーナルでは、裏で 2 次情報データベースの Scopus が動いていて、リファレンスとの連携を取っています。このような形で、投稿査読システムとも XML を通じて連携していくことを今後は考えなければいけません。J - STAGE の方にはもう既にお願ひしているのです、その対応に期待しているところです。

最後に、私は今、文科省の科学技術政策研究所の客員もやらせていただひいて、そのリサーチインタレストですが、Web-oriented publishing の世界に今後は向かひていきます。(図 15)

どうということかという、Web による研究者コミュニケーションの多様化が進んでいき、既にブログなどで行われているように、論文を書く前からのタグ入れを行う可能性があるということです。

課題と展望

- XMLベースで進めることがどこまで有効か見極める(XMLありきにならないように)
- 「欧米に比較して」コスト対効果をどこまで出せるか(事業効率)
- 全文データ(xhtml)の実際的な利用
- NLM-DTD XMLの国内での浸透具合を見る(人柱)
- 生物、医学系ジャーナルへの展開
- 和文誌対応
- 投稿、査読システムなどとの本格的な連携
 - Editorial Managerの例
 - J-STAGE

SPARC Seminar 2008 No.2 K. Hayashi

(図 14)

Web-oriented publishing

- Webによる研究者コミュニケーションの多様化
- 論文として書く前からのタグ入れの可能性(blog)
- 発行した後のコミュニケーション
- PCリソースとツールの発達および研究者の行動習慣の変化が、審査済みの情報からのlegacy publishingではなく研究情報が生まれる段階からのpublishingを加速する可能性が高い
- ただし、研究評価も変わらないといけなが(これが重要)
- いずれにせよ新しい時代の情報流通ソリューションが構築される可能性は十分にあり、そのときの標準(流通)メタデータの有候補はXMLであろう


SPARC Seminar 2008 No.2 K. Hayashi

(図 15)

発行した後のコミュニケーションもできます。例えば legacy な枠組みでネイチャーが行ったオープンピアレビューがあり、これはうまくいきませんでした、これまでもいろいろな例がありました。しかし、PC リソースとツールの発達に伴い研究者の行動習慣が変化することによって、結果的に、審査済みの情報（原稿）から始める今までの出版スタイルではなく、研究情報が生まれる段階からの出版に加速する可能性が高いと思います。ただし、研究評価を同時に変えないなりません。ここが非常に重要ですが、これを踏まえて変わる必要があると思います。（いくら新しい便利な情報流通手段が生まれてもその利用が研究者のキャリアパスに影響を与えるものでなければ、受け入れられないと考えています。）

いずれにせよ、新しい時代の情報ソリューションが構築される可能性は十分にあります。そして、その際に、新たな研究情報が生まれるときから世に出るまでの標準メタデータの最有力候補は、やはり XML だろうと考えています。

しかし、そういう時代がいつ来るかは、まだ分かりません。（図 16）したがって（単なる技術的な興味だけで見ずに）事業運営とビジネスの点を考えながら現実的な手法を下すことが必要であり、当然、開発と安定運用のバランスを取ることが重要です。ですから、XML を作るために XML を作ることは避けたいわけで、実は、それが皆さんに最も伝えたいことでもあります。



ただし、

- いつその時代が来るかはわからない。
- 事業運営、ビジネスの点からは現実的な手法とが必要。
- また、開発と安定運用のバランスをとることが望ましい。
- XMLを作るためにXMLを作ることは避けたい。

SPARC Seminar 2008 No.2 K. Hayashi

(図 16)