

電子原稿・電子出版・電子図書館
-- 「SGML実験誌」の作成実験を通して --

根岸正光
学術情報センター

電子出版というキーワードを中心に、文献の執筆、編集、印刷、流通の過程の電子化について、その現状を整理して今後の動向を検討する。とくに、電子出版に関する国際規格であるSGML(Standard Generalized Markup Language)に対するわが国での動きに注目し、学術情報センターにおける「SGML実験誌」およびその全文CD-ROM版の作成実験を通じて得られた知見を報告する。また電子図書館について、その諸形態を展望する。出版文化に依拠するこの種の社会システムの場合、技術的な問題というのは、その技術の社会的な受容・浸透の問題であり、この観点から標準化の重要性を論じる。

ELECTRONIC MANUSCRIPTS, ELECTRONIC PUBLISHING AND ELECTRONIC LIBRARY
-- "SGML Experimental Journal": An Experiment at NACSIS --

Masamitsu NEGISHI
National Center for Science Information Systems (NACSIS)
3-29-1 Otsuka, Bunkyo-ku, Tokyo 112, Japan

Current situation on electronic processing in preparation, editing, printing and distribution of documents is summarized and its future trend is discussed, with focus on the concept: "Electronic publishing." Movements in the country concerning an international standard on electronic publishing, SGML(Standard Generalized Markup Language), are assumed to be important, and the results from an experiment at NACSIS to publish "SGML Experimental Journal" and to make its full-text CD-ROM version are reported. Various forms of "Electronic library" are also investigated. The author puts emphasis on standardization, as technological problems for those social systems based on cultural settings of publication of the country, are the problems of acceptance and penetration of the technology in the society.

1. はじめに

文書原稿を電子的な媒体上に作成して、以後の編集から印刷までを一貫して電子的に処理し、最終生産物としての印刷物を得るという方法は、ワープロの普及に従って各所で行われるようになっていくが、おおむね局所的かつ小規模のものであって、この種方式の大規模・組織的な応用はまだ限られている。学術情報センターでは、学術文献の「電子出版」の可能性について検討するべく、国際規格SGML(Standard Generalized Markup Language)¹⁾による学術雑誌の全文データベース化と多種媒体による刊行のための予備的実験を行った。本稿では、この実験の経験を踏まえつつ、電子原稿、全文データベース化、電子出版、さらには電子図書館という一連の流れについて、技術的側面のみならず、技術の受容に関わる社会的側面にも注目して検討することにした。

2. 電子出版

2.1 模擬的学会誌「SGML実験誌」の作成工程

学術情報センターでは、1989年に学会誌全文データベース「学術論文データベース」のサービスを開始し、これには現在、化学系および電子・通信系の5学会6誌が含まれている。検索・表示には端末とファクスを併用し、図表の類は利用者の指示に応じてファクスに送信される。このデータベースの作成方式は、電算写植(CTS)ファイルの再編集方式ともいうべきもので、各学会誌をCTSにより印刷する際に生じる写植機用のファイルに手を加えて、データベースに仕立て直す方法である。この場合、

(1) 貼込みによる最終校正は、CTSファイルに反映されないから、別途の処理を要する、

(2) 雑誌が印刷された後に、データベース化処理に着手するので、データベースの速報性は阻却される、

(3) 何よりも、CTSファイルの再編集には相当の手間がかかる、

といった問題がある。これらの解決には、結局、印刷物とデータベースの主従関係を逆転し、まずデータベースを作成し、これを印刷やオンライン・サービス、CD-ROMなど多様な媒体で供用するという考え方をとる以外にない。

このような全文データベース→多様な公開形態という作成方式に関しては、1987年にISO規格となったSGMLの応用が欧米で進みつつある。こうした状況に鑑み、学術情報センターでは、1990年度に民間等との共同研究として「学術雑誌全文データベースのためのSGML方式による投稿・編集システムの研究」を実施した。この研究は凸版印刷株式会社との共同研究として行われたもので、その目的は、全文データベース作成のための国際規格SGMLを適用して、学術論文の入稿から編集・データベース化・模擬的学会誌印刷の過程をひとわたり実験してみるところにある。データベースの応用としては、論文全文のCD-ROMを作成する。

この実験の工程はつぎのようである。

(1) 各種学会誌における論文の要素、スタイルの分析

(2) 学術論文(和文・英文)のためのSGMLデータ定義(DTD:Data Type Definition)の設計

(3) 学術情報センター「学術論文データベース」からの試験用論文データの複写(一部論文は原学会誌から新たにパンチ入力)

- (4) 試験用論文のSGML形式への変換（上記DTDに即したタグの挿入）
- (5) SGML形式から実験誌印刷形式（Texによる）への変換仕様設計
- (6) スキャナーによる図表・数式等および印刷頁版面の入力
- (7) TeXによる「SGML実験誌」の印刷
- (8) CD-ROMにおける検索仕様の設計およびSGML形式からの変換
- (9) 検索ソフトウェアの開発とCD-ROMの作成

上記(3)において、著者・学会の了解のもとにこの実験データベースに入力された論文は10件であるが、具体的な入力方式はつぎの3種に分かれる。

(A) 学術情報センターの「学術論文データベース」から複写して変換したものの7件：高分子学会“Polymer Journal”（英文）1件、「高分子論文集」（和文）1件、日本農芸化学会“Agricultural and Biological Chemistry”（英文）1件、日本薬学会“Chemical & Pharmaceutical Bulletin”（英文）2件、電子情報通信学会“The Transaction of IEICE”（英文）2件。

(B) 原学会誌から新たにパンチ入力したもの2件：「電子情報通信学会論文誌」（和文）1件、「情報処理学会論文誌」（和文）1件。

(C) 著者保有の原稿ファイル（TeX形式）を借用したもの1件：情報処理学会“Journal of Information Processing”（英文）。

(7)において印刷された模擬的学会誌「SGML実験誌」は、A4版80ページとなった。図版等は、スキャナー入力して電子的に割付けており、いわゆる貼り込みなどのない完全な電子出版になっている。何分とも実験であるので、変換・編集時の単純ミスを完全に除去する余裕がなく、「誤植」が相当に見られるが、体裁、印字品質などには問題がなく、上質な仕上がりの雑誌になったといつてよさそう。

2.2 「SGML実験誌」全文CD-ROM

前記(9)の工程で作成された「SGML実験誌」全文CD-ROMは以下のようなものである。検索ソフトウェアには、凸版印刷で開発されたCD-ROM用汎用検索ソフトウェア「竹簡」を応用した。これは、MS-Windows(Ver. 2)上で稼働する検索ソフトで、各種索引語による検索・表示、イメージ・データの表示など、一般的な機能を備えている。これを応用することにより、学会名、著者名、論文標題中の語などから論文を検索でき（図1）、また該当論文本文中の特定の語の使用箇所を探索できる。

また、本文中で「Ref. 1」などのように参考文献を引用している箇所には、ハイパーテキスト流のボタンを設定し（図2）、これをマウスでクリックすると、論文末尾の文献リストの当該参考文献の記述部分に画面が移行する。またこれとは逆に、文献リストの方にもボタンが設けられており、これをクリックすれば、本文中でのその文献の引用箇所にジャンプする。本文中にある「Fig. 1」といった図表への参照もボタンになっており、これをクリックすると、イメージ・データとして別に保持されているその図表などが表示される。

また、学会誌の印刷物上での仕上がりページ全体もイメージとして保持されており、これを表示すれば雑誌そのものをみるのと同じ結果が得られる（図3）。もっともここでは解像度の問題があり、(1)画面上で印刷物と同程度の大きさになるように表示する、(2)スキャナーの解像度そのままのドット数で画面に大きく表示する、(3)ウインドウ内に頁全体が収まるように小さく表示する、という3種類の表示を用意した。この場合、高解像度ディスプレイを用いた方が、余程鮮明な表示が得られるこ

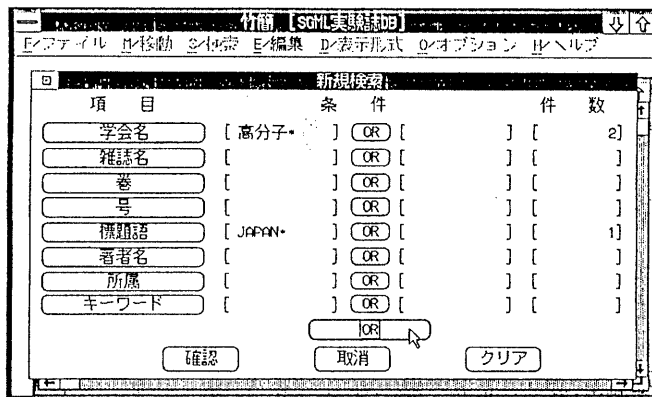


図1 「SGML実験誌」全文CD-ROM: 検索条件入力画面例

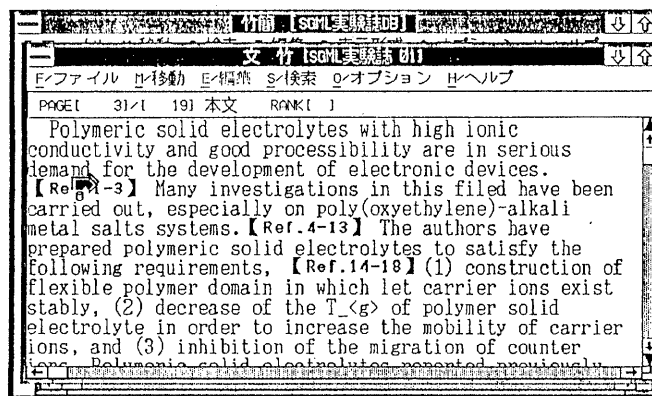


図2 参考文献引用個所のボタン表示例

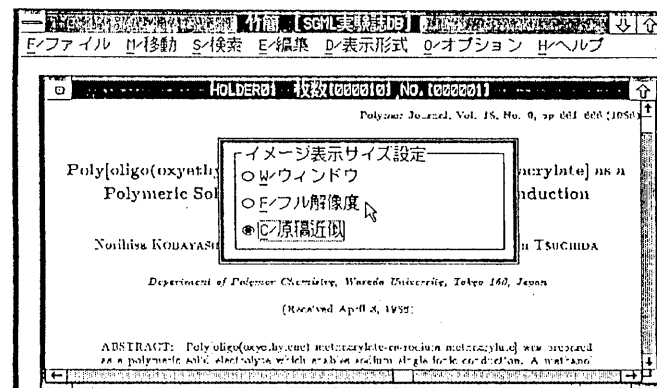


図3 印刷頁イメージの表示例

とはいうまでもない。また、ウィンドウ・システムで稼働しているので、本文の一部のカット・アンド・ペーストによる取り込みなどももちろん可能である。

2.3 版面イメージCD-ROM

全文データベースと版面イメージ・データベース ー ー ところで一般商用のオンライン・データベース・サービスでも、米国などでは全文データベース化が急速に進行している。たとえばDialogでは、160種の全文データベース（雑誌など原資料の種類で勘定すると実に1,900種）がサービスされている。²⁾ もっとも、これらはテキストのみををコード化データとして保持するもので、図表類は提供されない。

ファクスの普及が著しいわが国では、原資料の版面そのままをデータベース化しておき、これをファクスで送信するオンライン・サービスが行われている。エレクトロニック・ライブラリー社のELNETと称する新聞記事データベースがこれで、昔からの記事切抜きサービスの電子化版といえる。技術的には先述の学術情報センター「学術論文データベース」と同等のものであるが、コード化データとイメージ・データの主従が入れ替わっており、ELNETでは、コード化データは別途入力されたキーワードだけに限定され、これはもっぱら検索のための索引ファイルという位置付けになる。

版面イメージCD-ROM ー ー 上記のような版面イメージ・データベースは、これをCD-ROM版で頒布する場合、専用の検索ソフトウェアを付属できるので、その分、標準規格化との関連はゆるくなり、各社独自の商品展開が可能となる。欧米ではすでにADONIS、UMI、IEEEなどがこれを手掛けている。³⁾ 版面イメージCD-ROMは、その作成面においても、既製の印刷物版面のスキャナー入力を始点とするから、その印刷物の編集工程とはまったく無関係に作成でき、過去分へ遡ってのデータベース化など、適用範囲は広い。従って、今後多くのものがこの形態で頒布されるようになるものと思われる。

2.4 電子出版における標準化

このように版面イメージCD-ROMは比較的安易に製造可能であるが、これに対して、本来的な電子出版は、全文データベースを基礎とした多種の媒体・形態による資料の刊行であると解すれば、この方については依然大きな問題がある。すなわち、標準規格の普及という問題である。

例えば、前述の「SGML実験誌」の実験を通じてみると、まずSGML回りのソフトウェアが、わが国では未だ皆無に等しいということがあげられる。実験では、SGMLのパパーとして「MARK-IT」、フォーマッターとしてTeXを用いた。MARK-ITはECの支援により欧州で開発されたSGMLパーザーであり、欧文対応のものであるが、漢字コードのデータも一応そのまま通過するので、日本語文書にも適用可能である（ただし、各種表示では漢字は英文字に字化けするので、実務的には非常に使いにくい）。わが国独自の、あるいは日本語対応のSGML関係ソフトウェアとしては、松下電送の日本語SGMLエディター「MJSE-90」が最初でかつ現時点では唯一の市販ソフトウェアであると思われる。

標準規格制定の重要な利点は、その規格に対応した高品質なソフトウェアが廉価に供給されるという状況の現出にあるわけだが、わが国でのSGML回りに関する限り、こうした状態には至っていない。米国では国防総省、欧州ではECでのSGML採用といった刺激もあって、SGMLは欧米ではかなりの普及を見つつあるようである。わが国でも追っ付けこうした状態になるであろうと期待するものの、被我における出版文化と業態の相違に鑑みて、予断を許さないという点は、別に論じたことがある。^{4) 5)}

3. 電子原稿 ー 編集業務の電算化システム

3.1 ワープロ原稿

ところで、上で注目した全文データベースとか印刷物、CD-ROMなどの製品は、原稿から公刊に至る全過程のいわば川下部分である。そこで川上側に目を転じると、その電子化にはとりあえず目ざましいものがある。ワープロの普及がすなわちこれである。ワープロを用いれば、原稿は著者から電子的媒体（フロッピー・ディスク）で供給できる。すでに、主なワープロまたはワープロ・ソフトのファイルを写植機向きに変換するソフトが開発されており、これに依拠した原稿のフロッピー渡しは、出版・印刷実務において随分普及している。この場合、印刷所における新規入力とは異なり、誤字は原則として生じないから、校正量は随分少なくなる。ただし、上付き、下付き、倍角文字、罫線など、特殊な字体は「正しく」変換されないことも多く、これらには写植システム上での校正を要するから、自動的な原稿入力が可能になっているわけではない。

ワープロには、文書の蓄積、再利用等々の機能があるが、基本的には簡易印刷機であるとの認識が一般であろう。つまり、ワープロは最も簡略な電子出版機器であって、それなりの完結性を持っており、従って本格出版との関係性などは当面その守備範囲外にある。ワークステーションによる「高級な」DTPシステムにしても、事態は基本的に同様である。従って、これらは、ファイルの流通性、互換性の要件を満たして、本格的な電子出版の川上部分に配置されるというようなものにはなっていない。電子出版の標準化が未発達である現状からすれば、これは当然のことであり、やむを得まい。

3.2 電子投稿・電子査読 ー 投稿・編集・公刊の一貫システム

通常の出版では著者と編集者あるいは印刷所との往復で編集・校正が進行するが、学会誌を考えると、そこに特有の業務として査読がある。つまり、編集者と査読者間での往復が付け加わるので、その分事務運行は複雑になり、その全体を電子化するには余程の準備を要する。電子メールによる投稿と査読というアイデアは、当面誰も思い付くところであるが、実際の事務システムとしてこれを運用するには周到な検討を要する。

技術的観点からは、問題はやはり論文中の図表や写真、また仕上がり頁イメージの転送であろう。学会では、不特定多数という程ではないが、多くの会員が投稿と査読に関与するから、学界全体でのこの種の標準互換設備の普及状況が問題になる。ここでも問題は、技術それ自体の可否ではなく、もっぱらある種標準の普及・浸透という点にある。

この点は、送稿・審査・公刊の一貫システムの先行事例である特許庁の「電子出願システム」をみるとよくわかる。すなわち、このシステムは強力な制度的背景に裏付けられた業界システムであり、専用端末の開発・普及などもこうした背景において実現されている。学会誌や一般の出版物のようないわば開放型のシステム、あるいは社会システムでは、このような統括的機制は働かない。ここでは、当初、模様ながめ、なりゆき注目的な発想が支配的であるが、ある時ある種の臨界に到達して、にわかにはバスに乗り遅れるな論が全体を制するといった現象が、観察されるのが普通である。このように見てくると、電子原稿からの一貫システムを見通すのは相当に困難であり、当面、川下部分に限定した狭義の電子出版という観点での取り組みの方が現実的のように思われる。

4. 電子図書館 —— 文献流通の電子化

4.1 図書館機械化・資料の電子化・電子化ILL

文献流通の最も川下に位置する重要な機関として図書館がある。図書館の電子化はこれまで図書館機械化 (Library Automation) として語られてきた。これは要するに図書館業務の電算化であり、たとえば受入、閲覧、目録等々の業務の電算化である。しかし一方、図書館で収集する資料自体の電子化も急速に進みつつある。抄録誌のオンライン・データベース化は以前からことで、図書館では抄録誌を購入するのに替えて、外部のオンライン・サービスを利用して参考業務を行うようになった。昨今は各種のCD-ROMが市販されるようになり、これを図書館で購入して、利用に供することが広まりつつある。大学図書館では学内LANを介してCD-ROMを検索するシステムなども導入されているが、これなどはいかにも電子図書館的雰囲気を持たせようとするものといえよう。

つぎに、電子図書館的イメージをもたらすものとして、文献複写サービスの電子化が考えられる。すなわち、複写のオンライン発注とファクスによるコピー受信といった体裁のものである。実際、文献抄録データベースや雑誌総合目録データベースと連動する形での文献複写のオンライン発注は、学術情報センターにおいて、大学の研究者・図書館を対象として、1992年度から運用が開始される。そこでつぎに電子的な配布手段の方であるが、これに向けて当面検討に値するのはファクシミリであり、学術情報ネットワークを介する大学図書館間ファクシミリ・ネットワークはすでに稼働している。安達によれば、ファクシミリ・ネットワークは電子図書館システムの足回りとして将来とも有望で、とくにG4ファクスは電子メールの標準プロトコルであるMHSに包含されている点からも、十分注目するべしとされる。³⁾

4.2 エレクトロニック・キャンパスと「壁なし図書館」

ともあれ、電子図書館をどのようなものと想定するかについては、未だ議論の分かれるところである。従来型の図書館の単に高度に電算化されたものをいうとする立場ももちろん可能であるし、一方、これまでの図書館の概念を根本的に再定義するような発想もありうる。ここで、現在米国で進行中のふたつのプロジェクト、Cornell大学、Bellcore、CAS (Chemical Abstract Service)、OCLC (Online Computer Library Center) によるCORE (Chemical Online Retrieval Experiment) 計画と、Carnegie Mellon大学、OCLCによるMercury計画は、今後における電子図書館システムのイメージを与えるものとして、注目してゆく必要がある。^{6, 7, 8)}

松村によれば、Lancasterらによって1980年代に想定されたペーパーレス・ライブラリーは現実のものとはならなかった。⁹⁾ 実際、むしろOA化による紙多消費現象の方が顕在化しているのは周知のとおりである。そこで、今後予測されるのは、紙なし図書館ではなく、「壁なし図書館」であり、これこそが電子図書館の基本概念であるという。これは、大学図書館の場合、LANにより結合されたエレクトロニック・キャンパスの中で実現されてゆく。この場合、これまでの個々の図書館が解体・解消されるというような単純なものではなく、たとえば外部情報源とのゲートウェイとして機能展開などが考えられ、松村はこうした新しい図書館システムのモデルを提示している。いずれにしても、図書館は、物理的存在である従来風の図書『館』からはいよいよ離れて、より機能的な存在に衣替えしてゆくこ

とはなるのであろう。

5. おわりに

これまで、学術情報センターにおける全文データベース・電子出版関連の経験および実験と、電子図書館システムに関する検討結果を基礎として、文献情報電子化の諸側面について概観した。同センターで現在進行中である電子図書館システムの実験計画については、機会をみて別途報告することにした。それにしても、紙という人類史上稀有の発明品を、われわれはにわかには放擲できそうもないようである。文献の電子化といっても、結局はその内容を紙の上に定着させたところから、情報内容の本格的な受容が開始されるように思われる。従って、文献の電子化にあたっては、必要な文献の探索・選択・送受信・保存などにおける自由度を飛躍的に向上させるのみならず、文献内容の紙の上への定着過程を、個々人の需要に応じて適時・適切に制御できるようなシステムが目標とされなければならないということ、どうやら確かなようである。

なお、「SGML実験誌」関係研究は、学術情報センター・スタッフの他、図書館情報大学・石塚英弘、長谷部紀元両助教授、凸版印刷・田中洋一、平澤道彦、坂田英俊の各氏の協力によるものである。

《参考》

- 1) Charles F. Goldfarb, "The SGML Handbook," Oxford University Press, 1990, ISBN 0-19-853737-9. 664p.
- 2) "Dialog Full-Text Sources," Dialog Information Services, Inc., 1991. 12p.
- 3) 安達 淳「新技術によるILLシステムの構想」、根岸・猪瀬編「図書館システムの将来像——密結合型図書館ネットワークと電子図書館」、紀伊國屋書店、1991、p.152-169.
- 4) 根岸正光「フル・テキスト・データベースの応用動向」、情報処理学会誌、掲載予定.
- 5) 根岸正光「全文データベースの課題と展望」、根岸・猪瀬編「図書館システムの将来像——密結合型図書館ネットワークと電子図書館」、紀伊國屋書店、1991、p.180-204.
- 6) Erick Jul, "CORE project enters phase III," OCLC Newsletter, May/June 1991, p.12.
- 7) William Y. Arms and Lisa D. Holzbauser, "Mercury: An electronic library," OCLC Newsletter, September/October 1988, p.15-18.
- 8) Mark Kibbey and Nancy H. Evans, "The network is the library," EDUCOM Review, Fall 1989, p.15-20.
- 9) 松村多美子「海外動向にみる今後の図書館システム」、根岸・猪瀬編「図書館システムの将来像——密結合型図書館ネットワークと電子図書館」、紀伊國屋書店、1991、p.111-121.