

解説



デジタル図書館

1. デジタル図書館とは[†]田 畑 孝 一^{††}

1. 第三の波

The Third Wave!

アルビン・トフラーがその著書「第三の波」[†]を著してから十数年、図書館の分野にいいよ第三の波が押し寄せてきた。コンピュータの小型高性能化、高速大容量通信ネットワーク、マルチメディア、ハイパテキスト、CD-ROM、電子図書、電子出版などの技術の進展により、我々を取り巻く社会情勢は大きく変化し、図書館の分野を直撃しつつある。

アルビン・トフラーの第一の波は数千年にわたってゆるやかに展開された農業革命を指す。それになぞらえてみると、図書館の分野における第一の波は、やはりゆるやかに古代、中世にわたって考案されてきた種々の分類法、そして1870年代に創案された十進分類法(DDC)ということになろう。

アルビン・トフラーの第二の波は産業革命を指すが、図書館の分野における第二の波は1960年代後半に始まるコンピュータ革命ということになる。MARC(機械可読目録)の出現や、オンライン情報検索システムなどがそれである。第二の波は第一の波の時代の古い体制の抵抗に会いながら、ようやく一般に浸透したといえよう。

到来しつつある新たな第三の波!

アルビン・トフラーの第三の波には特定の呼称が与えられていなく「これから」の時代を指したものであるが、図書館の分野における第三の波は「デジタル図書館」革命であろう。

アルビン・トフラーの第一の波の時代は、生産と消費が渾然一体とした手仕事の時代であった

が、何でも屋の個性的な人間の能力が尊ばれた。第二の波の時代になると、生産と消費の決定的な分離ということから派生して、規格化、分業化、同時化、集中化、極大化、中央集権化というどちらかというと非人間的な原則が人間を支配するようになった。生産と消費がふたたび融合されてゆく第三の波の時代では、人々はこれらの非人間的な原則に挑戦し、それを克服していくのだという。

図書館の分野における第一の波の時代はそれぞれの図書館でオールア라운드でかつ個性的な司書が尊ばれたに違いない。第二の波の時代、これは現在も進行中であるが、あまりにも商業的な出版活動によって余儀なくされる著者と読者の隔離、MARCやオンライン情報検索システムによる規格化、分業化、同時化、集中化、極大化、中央集権化が行われており、没人間的な環境となっているかもしれない。

第三の波の時代がやってくると、電子出版・電子の流通により、手作りの、個性豊かな、創造性に富んだ多形態な著述そのものが、時間と空間を超えて、読む人に直接もたらされる。つまり書く人と読む人の一体化、融合化が行われ、分散化、分権化が進む。人間性あふれる第一の波の時代への回帰ともいえるのである。

「デジタル図書館」は第二の波での束縛から人々を解放する担い手であり、エイブラハム・リンカーンのゲティスバーグの演説になぞらえば、それは

‘Library of the people, by the people, for the people’

である。そこでは、だれもが権威や資金にとらわれず自由に出版でき、まただれもがそれらの中から欲するものを自ら探し出し無料を原則に読書することができる。そして、数多くのそれらの中から

[†] What's a Digital Library? by Koichi TABATA (University of Library and Information Science).

^{††} 図書館情報大学図書館情報学部

ら未来にわたる価値を見抜き後世に伝えるものを見い出すには、いまだかつて経験したことがない高度な Librarianship が要求されるのである (文献2) の No.1 の序)。

2. デジタル図書館にアクセスしてみる

デジタル図書館の1つの典型を TULIP で見てみる。パソコンやワークステーションで Web (World Wide Web) に入る。URL (Universal Resource Locator) として、<http://tulipsrvr.engin.umich.edu/tulip/> を指定すると、TULIP のホームページが現れる (米国ミシガン大学にそのサーバがある)。そこで、材料科学 (Material Science) 関係の雑誌論文のデモンストレーション、すなわち量はわずかであるが雑誌へのアクセスを実際に体験できる (雑誌の種類は20)。TULIP Journal Browser には次の検索項目とその各々に入力欄が用意されている。

- (a) 全文 (full text)
- (b) 要約 (abstract)
- (c) 表題 (article title)
- (d) 著者 (author)
- (e) 雑誌名 (journal title)

(a) 欄にたとえば 'transition' と入力し検索すると、論文の本文にそのことばが含まれる論文の論文名とその雑誌名がリストアップされる。いわゆる全文検索である (筆者の端末では検索開始4分後に258件リストされた)。リストされた各論文について「要約」および「記事全文」を見ることができる。「要約」を指示すると、その論文の書誌事項と要約が現れる (数秒後に、文字コードとして)。「記事全文」を指示すると、その論文が掲載されている雑誌のページそのもののイメージが1ページごと現れる (最初のページが数十秒後に、ページイメージとして)。(b) 欄に 'transition' と入力し検索すると、要約文にそのことばが含まれる論文が上と同様リストアップされる (検索開始45秒後に51件リストされた)。(c) 欄に 'transition' と入力し検索すると、表題にそのことばを含む論文が上と同様リストアップされる (検索開始15秒後に14件リストされた)。(b)、(c) でも (a) と同様「要約」および「記事全文」を指示できる。

以上のように、机上のパソコンから居ながらに

して他国の学術論文を検索し、検索した論文の記事本文そのものを読むことができる。ついでながら、「TULIP とは何か」もこのホームページをたどることによって知ることができる。TULIP は Elsevier Science Publication の音頭取りによって1991年3月に始められたプロジェクトで、学術論文の電子的配布に関する問題点を調査するものである。これには9大学と約60の材料科学雑誌が関与している。論文記事のデータは次のように作られる。雑誌の背が裁断され、各ページはイメージスキャナでスキャンされページイメージとなる。ページイメージはOCR (Optical Character Recognition) で読み取られる。記事の始めの部分にある書誌データ (表題、著者など)、キーワード、要約は、人手でチェックされ、スペル誤りは正され、それぞれ規定の欄に割り当てられる。未訂正全文、つまりOCRで読み取ったままの記事本文 (文字読み取り誤りが存在するもの) もデータとして蓄えられ、上記の (a) の検索に使用される。ページイメージはファクスG4による圧縮で、1ページ当たり70-80KB (バイト) となる。未訂正全文は1ページ当たり4KB、書誌データなどは1.5KBである。Elsevier は Web サーバを1994年9月に立ちあげた。

3. デジタル図書館とは

「デジタル図書館」(digital library) とは、文章のみならず、画像、映像、音響などを含むマルチメディア図書の収集、蓄積、配布を「デジタル信号」の形態で統合して扱う図書館を指す。国際高速大容量ネットワークで大小、公私の図書館が互いに結合され、各家庭から居ながらにして世界中の図書の内容全文を読むことができる。

「電子図書館」(electronic library) および「仮想図書館」(virtual library) の用語は、デジタル図書館と同義語といわれている。しかしながら、米国では長く使われてきた電子図書館に代わって次第にデジタル図書館の用語が使われるようになり、最近はその定着しつつある。1994年5月に開催された1994 Workshop on Digital Libraries³⁾、続いて6月に開催されたコンファレンス Digital Libraries' 94 (DL' 94)⁴⁾ で初めてその用語が広く認識された。9月には NSF/

ARPA/NASA から総額 2,440 万ドル (4 年間) の Digital Libraries の名が付けられた 6 つの研究プロジェクトへの助成が発表された^{5),6)}。そして 1995 年 4 月号の CACM⁵⁾ において Digital Libraries の特集記事が掲載されるに至ってその用語が一般に認知されたと考えてよからう。ADL'95³⁾ が 5 月に、DL'95⁴⁾ が 6 月に開催され、1996 年 3 月には ACM 主催の国際会議 DL'96¹⁸⁾ が、また NASA/NLM/IEEE 主催で ADL'96 が 5 月に開催され、この用語はますます一般化してきている。

我が国においては、「デジタル図書館」の用語が図書館情報大学で 1994 年 8 月より年 3 回開催されている「デジタル図書館ワークショップ」²⁾、また文部省助成により図書館情報大学主催で 1995 年 8 月に開催された「デジタル図書館国際シンポジウム 1995」⁷⁾ で用いられている。さらにこの用語は、上記「デジタル図書館ワークショップ」の活動を基に 1995 年 8 月に設立された「デジタル図書館学会」⁸⁾ で用いられている。

映像や通信ネットワークの技術がアナログ技術であったかなり以前から用いられてきた電子図書館の用語は、その当時としてはいわば「非現実な」理想の図書館を意味することになってしまい、その後今日まで長らく使われてきた。最近になって、デジタル高速大容量通信網が敷設され、デジタル映像技術によって映像を含むマルチメディア情報がすべてデジタル形態に統合され、その処理、蓄積、伝送がデジタル技術で統一的行われるようになり、そのことによってようやく「現実」に可能となった理想の図書館を表す用語としてデジタル図書館が用いられるようになったといえよう。とはいえ、1994 年 9 月に公開実験された電子図書館 Ariadne⁹⁾、1995 年 2 月から試行実験サービスを始めた学術情報センターの電子図書館サービス、1995 年 9 月に公開実験を始めた通商産業省のパイロット電子図書館システム事業など、我が国では電子図書館のことがばむしろ通用している。

3 つ目のことば、仮想図書館とは自分の机においてあるパソコンやワークステーションが実質上の (virtual) 図書館であるとの意で、それを用いて世界中の図書館の図書を直接読むことができ

る。デジタル図書館の特性を別の観点から言い表している。

4. もう 1 つのデジタル図書館

灰色文献 (Grey Literature) と呼ばれるカテゴリの文献がある。灰色文献とは通常の出版・流通の経路で扱われていなく、またそれについての検索手段が整備されていないので、入手が困難な文献を指す。これにはテクニカルレポート、会議資料、学位論文、大学出版物、官公庁資料などが含まれる。これらのうちいくつかがたとえ正規の流通経路にある学術雑誌に載ったとしても、研究成果が得られてから出版までに時間がかかり、現場の研究者にとってはテクニカルレポート、会議資料などがより重要となる。また官公庁資料は民間の企業活動にとって貴重であるがなかなか入手できない。灰色文献の収集と蓄積、検索、配布については、従来、英国の BL (British Library)、欧州の EAGLE (European Association for Grey Literature Exploitation)、米国の NTIS (National Technical Information Service)、我が国の JICST (日本科学技術情報センター) などがその事業の一環としてそれに努めている。灰色文献の 2 次情報をオンライン情報検索システムや CD-ROM で提供し、ものによってはその 1 次資料のコピーを郵送あるいはファクスで配布している。しかしながら対象とする文献は限られ、またその事業の採算性に難点があつて、これまでの体制では問題の解決にほど遠いものであった。

最近の Web の状況を見ると事態の解決の糸口が見えてきたように思われる。文献の発生源においてそれぞれが競って HTML (HyperText Markup Language) を用いて Web 上に発信するようになってきており、机上のパソコンから直接全文を入手できる。上述の「TULIP とは何か」の記事はその例で、「デジタル図書館ワークショップ」²⁾、「デジタル図書館国際シンポジウム 1995」⁷⁾ の論文集もその全文が Web で読める。1995 年 7 月創刊の D-lib Magazine¹⁰⁾ は当初から紙媒体ではなく Web で発行されている。灰色文献はデジタル図書館の出現によって、もはや灰色でなくなりつつあるといえる。けれども、無数に発生する文献のうち、どこにどんな有

用なものがあるかの案内情報や検索情報を提供するための手段の整備がこれからの課題となる。

Webのすべてのページ全体をクロール(crawl:泳ぎまわる)し、ページの内容(文献や記事)を解析し、それらに索引付けすることを、一貫して自動的に行うWeb検索サーバが出現しており、使ってみると結構役に立つ。これらの検索サーバは日々Webをクロールし索引付けを更新している。1つの例として、Webにて<http://www.opentext.com/>のホームページにて“Search the Web”をクリックすると検索のページとなる(無料)。ここの入力領域にいくつかの単語を入れ、それらのANDあるいはORを指定して、検索実行ボタンをクリックすると、数秒後、本文中にそれらの単語のすべてを含むあるいはいずれかを含む文献や記事の表題の一覧が出る。いずれかをクリックすると、URLが埋め込んであるのでWebによってその本文が元来の所在から取り寄せられて表示される。単語として適切な専門用語をいくつか指定すればノイズ(関係ないものが出てくること)も少なくなり満足できる結果となる。試しに自分の名前といくつかの専門用語を入れて検索すると自分自身のWeb文献や記事を発見できる!このような自動的、機械的なものが十分であるわけではなく、分野ごと人知を尽くし組織化し、再現性、適合性を高めた情報検索手段が必須であることは論をまたない。

5. SGMLとページイメージ

我々が原稿を書いて出版社に渡すと、編集者はその原稿に赤インクでいっぱい印を付ける。この部分は「表題」で中央に配置し活字の大きさは18ポ、ここは「見出し」で14ポのゴチ、文章のここでは改行せよなどと、印刷業者に指示する印を付ける(markup)。植字工が見やすいように赤インクで入れていたこれらの印の部分電子化するにあたって、それを特殊な制御コードではなく、本文に用いられる通常の文字で表し、原稿本文の文字列の中に埋め込むことにした。その国際標準がSGML(Standard Generalized Markup Language)である。

記事の原文に、原文のどの部分が、「表題」、「著者」、「要約」、「章見出し」、「節見出し」、「文章段落」であるかを示すため、それぞれに対

応した一対の印でそれらの部分を両側から囲む(開き括弧、閉じ括弧の対のように)。これらの印(タグと呼ばれる)は印刷に役立つのみならず、この形式で収集された記事のデータベースを検索する際にも役立つ。表題のみを対象とした文字列検索、要約のみを対象とした文字列検索、章見出しあるいは節見出しを対象とした文字列検索、本文全文を対象とした文字列検索などが可能となる。

SGMLは、記事を蓄積する方法として、ページイメージとならんで最も利用されているものである。文字コードとして蓄積されるので必要な記憶容量は極小となるが、文字の入力やタグの挿入のためページ当たりの単価は高くなる。OCLCの経験によると、SGML(技術資料)では\$15/page、8-10 KB/pageであり、一方、ページイメージでは\$0.50/page、80-100 KB/pageである¹¹⁾。SGMLはそれ自身には図面の記述能力がないので、図面を参照する場合にはそれらが格納されている外部ファイルの識別子によって指し示すことになる。なお、WebのMosaicやNetscapeのビューアで利用されているHTMLはSGMLで定義された文書形式の1つである。

図面を持つ科学技術論文はそれらを数多く蓄積するとなると、SGMLではテキストと図面がリンクという弱い結びつきなので保守性に難があり、それらが一体化しているページイメージ方式が有利となる。一方、文学など、文脈上のことば1つ1つが研究対象となる分野では必然的にSGML方式となる。学術情報センターの電子図書館サービスや奈良先端科学技術大学院大学附属図書館の電子図書館サービス¹²⁾はページイメージによる。

6. 多様化する電子出版

いま私達が手にする図書の形態は、グーテンベルクによる印刷術が発明されてすぐにそうだったのではない。当初の印刷本はそれまでの写本の形態を踏襲しており、その後、数十年かかって文字のデザインをはじめ、表題紙、目次、ページ、巻末索引、通し見出しなどが考案されていったという。ネットワークで流通するマルチメディア電子図書の出版はいま始まったばかりである。ページイメージはいうまでもなく冊子形態の踏襲であ

る。SGML も文書の形式化の技法で基本的にはやはりこれまでの延長である。

HTML はハイパテキストの技法を取り入れ、確かに 1 次元のテキストを超えているが、図面の表示、音響の再生は Web ビューアの機能に任せられ、かつ限られた能力しかない。いまのままの機能では「とび出す絵本」という紙媒体の絵本さえ模倣できない。レシプロ型の自動車エンジンのピストンの動きの説明図は子供ごろに理解し難かった。ピストンの行程対応にいくつかに分けた図がありそれをあれこれ頭の中でつないでみるが、もう 1 つピンとこなかった。これは 1 つの図の中でピストンがゆっくり動いてくれればただちに理解できるのであるが、いまの HTML とそのビューアの機能では表現し難い。

「青い鳥」の読者は主人公のチルチルとミチルになりきって Virtual Reality の森の中をさましてもよい。森の分かれ道をどちらに行くかで話の展開が異なってもよいのである。そこにはもはやページの概念はまったくない。紙媒体から脱却したマルチメディア電子媒体に固有な表現法は人々のこれからの工夫によって出現してくる（本特集：Informedia 参照）。

ビューアの上で動作できる Java 言語が注目されているが、当面はそのような言語を利用した手作りのネットワーク流通電子出版物があれこれ試されるであろう。

7. 新たな情報収集・蓄積・サービス

出版社や学会はこれまで「価値ある」著作物を生み出す媒体として機能してきた。新人文賞によって才能ある作家の発掘を行ったり、社会的に活動している人々に著作を促したり、質の高い学術論文の学会誌への投稿を促してきている。図書館はそれら著作物の中から予算の限度内で選択するという作業を行ってきた。

いまこうして、だれもが自由に出版できる Web という媒体が出現すると、数多くの著作物の中から未来にわたる価値を見抜き後世に伝えるものを見出す収集・蓄積の作業は、出版社や学会の仲介なく、図書館員が直接自ら負わねばなくなる。一方、その作業は図書館員に限らずネットワークの利用者のだれもが行える。図書館員をはじめ一般のだれもが自身の関心事について全

世界を対象に収集を行い、また参考奉仕に応じるという、それぞれがその道の第一人者となり得るのである。

望みの情報にたどり着くには、その道の専門家から助けを得ることができればありがたい。図書館の分野ではそのようなサービスを参考奉仕 (reference service) という。従来から図書館の窓口で行われている参考奉仕は、ネットワークを介したワークステーション上のディスプレイを舞台にした電子的協調作業 (collaboration) に発展していくに違いない¹³⁾。利用者は白室に居ながらにして電子図書の本文そのものを対象とした専門家の支援を受け得る。さらに専門家の参考奉仕の知識やノウハウを情報検索システムやエキスパートシステムに作り込み、ネットワークを介してだれもが利用できるようにするとその効果は大きい。たとえば、心理学、教育学的観点から子どもの発達状況に合わせて読ませるべき本を母親にアドバイスすることができる¹⁴⁾。

ことばをいまだ十分に使えない幼児が感性によって自ら本を探して読むことができるように、マルチメディアによる絵本の紹介システムを用意し、ネットワークを介して選んだ本を読んだり読み聞かせてもらったりすることも可能となる¹⁵⁾。何を読みたいか自分自身の心情をつかみかね、感性に訴えて本を求めることは、こどもだけではなく大人にとっても必要なときがある。

それにつけても、デジタル図書館が社会においてより一層機能するには新たな社会の仕組みが求められる。ネットワークを介した情報資源の提供者として中小組織のみならず個人も可能で、かつ 1 回に 1 ページのみ 5 円というような小口の課金も可能な料金徴収システムが出現しつつある¹⁶⁾。また、これまでの紙媒体出版物のための著作権制度はもはや適用が困難として、ネットワーク情報流通のための新しい著作権制度の模索が行われている¹⁷⁾。

8. コミュニティ発展のために

思いを同じくする同好の人々によるコミュニティの活性化にとって、デジタル図書館の役割は計り知れない。子ども達の文集、主婦の文化サークル誌、学会誌など、ありとあらゆるコミュニティにおいて、だれもが自由に自己を発現できかつ

互いに情報共有できるこれほど有意義なものはない。ネットワークで流通している電子文書や電子図書の自動点訳、自動朗読によりコミュニティのひろがりやさらに進む。

コミュニティのひろがりや身近なことに限らない。ガリ国連事務総長の国連創設 50 周年総会演説に「Globalization と Fragmentation の現代」の一節がある。地球規模の連帯化の進展の一方で民族の限りなき細分化という、相反する価値感が支配する現代にいかに対応していくかが課題というのである。民族固有の言語と文化を尊重する一方、相互理解を深め、知性と感性を地球規模で共有するには、デジタル図書館による寄与が大きい。それは世界のあちこちに離ればなれに住む少数民族のコミュニケーションと文化の伝承に欠かせないものとなるに違いない¹⁸⁾。

参考文献

- 1) Toffler, A. : 第三の波, 徳山二郎ほか訳, 日本放送出版協会 (1980).
- 2) 「デジタル図書館」, デジタル図書館編集委員会, 図書館情報大学内, No. 1, 1994 年 8 月, ISSN 1340 7287, <http://www.DL.ulis.ac.jp/> (1994) 以下続巻 (年 3 回).
- 3) Adam, R. N., et al. : Digital Libraries, Lecture Notes in Computer Science 916 and 1082, Springer-Verlag (1995, 1996).
- 4) Proceedings of Digital Libraries '94 and '95, Texas A & M University, <http://www.csdl.tamu.edu/DL94> and <http://www.csdl.tamu.edu/DL95> (1994, 1995).
- 5) Communications of the ACM, Vol. 38, No. 4 (1995).
- 6) Computer, IEEE Computer Society, Vol. 29, No. 5 (1996).
- 7) ISDL 95 : Proceedings of International Symposium on Digital Libraries 1995, University of Library and Information Science, <http://www.DL.ulis.ac.jp/ISDL95/> (1995).
- 8) <http://www.DL.ulis.ac.jp/SDL/>.
- 9) 長尾 真, 原田 勝, 石川徹也, 谷口敏夫, 澤田芳郎, 吉田哲三, 柿元俊博 : 電子図書館 Ariadne の開発 (1), 情報管理, Vol. 38, No. 3, pp. 191-206 (1995) 以下続 4 編.
- 10) D-lib Magazine : Magazine of Digital Library Forum, July 1995, ISSN 1082-9873, <http://www.dlib.org/> (1995) 以下続巻 (月

刊).

- 11) Noreault, R. T. and Crook, A. M. : Page Image and SGML : Alternatives for the Digital Library, ISDL 95, pp. 145-150 (1995).
- 12) Imai, M., Horii, C., Hada, H., Yokota, N. and Chihara, K. : Design of a Digital University Library : Mandala Library, ISDL 95, pp. 119-124 (1995).
- 13) Sugimoto, S., Gotou, S., Zhao, Y., Sakaguchi, T. and Tabata, K. : Enhancing Usability of Network-based Library Information System—Experimental Studies of a User Interface for OPAC and of a Collaboration Tool for Library Services, Proceedings of Digital Libraries '95, Texas A & M University, pp. 115-122 (1995).
- 14) 田畑孝一, 杉本重雄, 阪口哲男 : 児童の発達課題からみた児童図書の選択とその選択システム, 「デジタル図書館」, No. 7, ISSN 1340-7287, pp. 51-59 (1996).
- 15) 杉本重雄, 藤田岳久, 阪口哲男, 田畑孝一 : マルチメディア知識ベースとその児童読書相談システムへの適用, 情報処理, Vol. 33, No. 7, pp. 838-845 (July 1992).
- 16) Sirbu, M. and Tygar, D. : NetBill : An Electronic Commerce System Optimized for Network Delivered Services, Proceedings of IEEE CompCon, 1995, pp. 20-25, <http://www.ini.cmu.edu/netbill/> (1995).
- 17) 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, ISSN 1340-7287, pp. 8-12 (1995).
- 18) Sakaguchi, T., Maeda, A., Fujita, T., Sugimoto, S. and Tabata, K. : A Browsing Tool for Multi-lingual Documents for Users without Multi-lingual Fonts, Proceedings of the 1st ACM International Conference on Digital Libraries, pp. 63-71 (1996).

(平成 8 年 2 月 14 日受付)



田畑 孝一 (正会員)

1941 年生。1963 年京都大学工学部電気工学科卒業。京都大学工学博士。1973 年同大学助教授。1982 年より図書館情報大学教授。

現在、同大学総合情報処理センター長。同大学で開催されている「デジタル図書館ワークショップ」の主催者代表 (<http://www.DL.ulis.ac.jp/>)。デジタル図書館学会会長 (<http://www.DL.ulis.ac.jp/SDL/>)。