

一次情報発表技術の展開
——特に専門情報の電子刊行に関して——

Coming Publication Technology for Primary Information:
Especially of Electronic Publication of Specialized Information

中 村 幸 雄
Yukio Nakamura

Résumé

Until recently, the technology of information handling has been concentrated on secondary information but new developments in electronic technology have relieved us from limitations in the text length for handling. Existing full-text databases are now used for retrieval purposes rather than for reading. To convert these full-text databases into ones that provide easiness in reading, a new technology for skip-reading is pointed out as necessary. Also an important factor is the way of thinking of human readers. Publications in specialized fields are going to be more oriented to small quantity publishing and databases can be used for such purposes under certain conditions. In this case, however, some new concepts relating to databases, such as "original edition," meaning of rights of publication and the liberty of processing by users, must be examined and established.

- I. 今までの習慣
- II. 新技術のもたらすもの
- III. 少数部数刊行とデータベース
- IV. 一次資料の多段階刊行
- V. テキスト加工
- VI. 専門的一次情報の変貌
 - A. 第二種刊行物のためのデータベース
 - B. 第二種刊行物の利用
 - C. 付随して起こる問題
 - D. 情報送り出しの経費
- VII. むすび

本論文の内容は、1989年11月11日に三田図書館・情報学会大会で行なった橋本記念講演に若干の内容訂正をおこなったものである。
中村幸雄：(社)情報科学技術協会会長，東京都文京区小石川 2-5-7。
Yukio Nakamura: President, Information Science and Technology Association, INFOSTA. 2-5-7, Koishikawa, Bunkyo-ku, Tokyo.
1990年12月7日受付

I. 今までの習慣

図書館人をつかまえて聞くと、図書館職員の中では「目録のとれる人」が一番大切なのであり、中でも洋書の目録をとれる人が最高なのだそうである。つまり図書館を構成する諸技術のなかでは、目録、即ち二次情報を作り出す技術が最もむつかしいし、それだけに、その技術を扱う人のランクが上ということであろう。

また少し河岸を変えて、情報検索をする人々をみると、データベース（以下 DB と略す）というものは、大抵が二次情報の DB であり、それらから、キーワード等を扱ってユーザの要求する記事を取り出すことが自分の仕事だと考えている。従ってヒットした記事を丁寧に調べて、内容をチェックする等は、メッタにしない。つまりキーワード・レベルに関心があり、記事・資料の本体（即ち中味）には関心は低いのである。

情報専門家という人々でも、二次情報データベースを中心に考える。最近の全文データベース（例えば CJACS）でも、大部分のユーザはこれを高級な二次情報データベースと見做して、新しい近接演算子等が駆使できることに満足しているらしい。

以上を見渡すと、情報の世界では、二次情報こそ仕事の中心なのであり、一次情報については（殊に技術的な）関心はほとんどない、といって差し支えなさそうである。要するに「二次情報崇拜」が常識なのである。

この一文は、現在までそうであったのには、相応の理由があるのだが、いつまでも、そうなのか、という観点に立っている。

II. 新技術のもたらすもの

電子技術の進歩は著しい影響をわれわれに与えている。誰の目にも明らかなのは、記憶容量の著しい増大である。半導体の側では1個のチップの記憶容量は3年ではほぼ4倍になるというのが今までの事実である。

また、いわゆるスタンドアローンの記憶媒体としての CD 系から光ディスクへの発展により、ギガバイト程度のものが容易に使えるようになってきた。

この技術面での変化が、情報業務のいろいろな面で量の拡大をもたらしていることは、誰の目にも疑いない。しかし量の著しい変化は質の変化を生み出す、という法則が働けば、どうなるのであろうか。これが本文の第二の目的である。

新技術の中心はテキストの電子化である。その要因と

しては記憶容量の著しい増大と処理速度の向上により、長文を扱うことに、データベース関係では苦勞を感じなくなつた。そこで全文データベースに何を入れるかが問題となる。しかしここではその議論は止めて、専門論文をのせる専門（学術）雑誌が近い将来大きな変化をはらむことに注意しておけば十分である。既にいくつかの分野でこれが実行にはいっている。その好例がアメリカ化学会の論文を収容する CJACS であるというに止めよう。

ここでは全文データベース（FTDB）が使われるようになった場合に生ずる問題点を指摘しておこう。

(1) 全文データベースの検索

在来型の DB では、孤立したキーワードが多数存在し、それら複数のキーワードの間にブール代数的演算を認めるだけである。ソソーラスにおける BT, NT, RT の関係は検索の中に直接取り上げられてはいない。これらの関係は人間が検索の前に、検索式を考えるとときに使うのだと従来は考えてきた。

そこに現れてきたのが位置演算子である。現象的には、確かにワードの位置に関する問題であるが、その主な狙いは syntax を検索に取り入れることにあり、と考えるてはならない。Syntax は機械翻訳の中で、あれほど使われているのに、検索では取り入れられることが従来少なかった。それがようやく利用され始めたのが、位置演算子という名前を借りているといえよう。

一方 semantics はどうであろうか。検索演算の中に取り入れられているのは、BT の操作を up-posting という形で採用している JICST-File くらいではないか。

従って検索の技法において、まだまだ多くの新しい工夫が生れて然るべきである。位置演算で止まっている、いけないのではないか。

しかし FTDB を検索の立場から見る議論の対象に止めておくのでは不十分という他ない。もっと別の眼で見る必要がある。

(2) 全文データベースはどう読むか

全文データベース（FTDB）は読むためのものでなければ、なぜ全文を収容しているのか意味がなくなる。ただし、全てのユーザが FTDB をいきなり読むか、となると大いに疑問がある。

DB の全文を読もうとしているユーザ（彼こそ読者である）が存在するとしたら、彼は読み始める前に、色々

なことを既に行っているはずである。この点では従来の印刷物についても同じである。

読者は、対象となるテキストについて何等かの情報を手に入れて、それを読むことが自分にとって得であるかを検討しているはずである。それを既にしたとしても、いきなり第1ページから読むことは普通しない。全テキストをパラパラと読み、大体どの辺には何が書いてあるかを偵察し、場合によっては、頭の部分を飛ばして、先の方のどこかを読み始める。それで理解が得られることもあるが、前提条件や予備知識の不足を感じて、頭の部分から読むべきだと考えたり、他にもっと分かる部分がないか捜したりする。

相当の迂回をした後に、では全文を読もうと決心する。また頭から読み始めても、一本調子に進むことは少く、ときどき前に戻って既読の部分を確認めたりする。要するに、行ったり来たりを繰り返しながら逐次進んでゆくものである。読むところをあちこちと動かす(変える)ことに、読み取る機構が適切に動かないといけぬ。この見地からすると、現在の端末機はあまり工合はよくない。スクロールができればよい、という程度では不完全なのである。

またアンダーラインを引いたり、淡い色のペンでテキストを塗りつぶすことも、読者はしばしば行う。これに相当することが端末機上で、できないといけぬ。それは現行の技術で、できることなのだが、この動作は端末機が受け入れたデータを、読者が加工することになる。この読みながら行う加工の可能性が従来は無視されている。

(3) 端末機における文字表示法

現在のところ、発光型の端末機が多い(CRT型)。これが目を疲れさせることは大抵の人が経験している。液晶型はコントラストが不足であり、見える角度が限定され過ぎていて、バックライト型でもまだ不十分である。つまり端末機の表示メカニズムはまだ改良すべき点が多いのである。

一方で発光型の特長として、多色化があげられる。これは従来の印刷における多色刷より、はるかに簡単に実現できる。またハイライト表示も有用であるが、これは字体の不足(例えばイタリック体の欠除)の対策にもなる。それにも拘らず、基本的な問題としてキャラクタ・セットの不十分さ(欧文、邦文を通じて)は、いずれは解決しなければならない。これは端末機だけの問題ではなく、電子化文字表示システムの共通の問題である。

これはコンピュータが数値の計算機として出発し、次いで事務用機器として生きてきたために、長文を読む道具としての意識が弱いために起ったことである。この解決には技術的に見れば基本的な困難はないのだが、使う方の人間の意識の問題が大きいのであり、ユーザの意見に基づき改善が必要な部面である。ハードウェアの専門家やシステム設計者の注意が欠けているのは、このようなユーザインターフェースの問題である。

III. 少数数刊行とデータベース

読者の数は少ないが、刊行の必要はある、という刊行物のことを「少数数刊行物」と呼ぶ。この種の刊行物は以前から必要があるのだが、適切な発行形式がないため、必要もない大部数(普通1,000部以上)を印刷し経費高に泣くか、または発行をあきらめるしかなかった。その中でも必要で欠くことのできないものが、専門的内容を持つ刊行物である。

ところで学問の細分化は避けられない傾向である。このため、教科書は別として、一つの論文・記事をほんとうに丁寧に読む人は、全国で数十人程度でしかあるまい。例えば専門的雑誌を買う人は(もちろん大衆雑誌とは事情が全く異なる)、その点は心得ていて、一号の雑誌に自分の読みたい論文・記事は精々一つあるかないかでも、専門家は雑誌を買うことをあきらめない。また何千部も発行する専門雑誌でも、原著記事では、一人の読者に喜ばれる記事は全体の2~3%もないとしても、それは当然だと考えている刊行者もあれば、またそんな雑誌を買う方の人もある(これが現状)。技術の進歩を考えると、これは人がよすぎるのではあるまいか。

一つの専門的論文・記事とは、熟読するごく少数の読者の手に渡るようになっていけばよいのであって、他の読まない読者にまで配布する必要はない。しかるに大量に印刷しなければ単価が下らないような現在の印刷法+配布法を継続しようとするから専門雑誌や専門図書の発行は困難になり、一方出版社はコピーは不当だと主張する。

従って専門論文・記事には、別種の刊行形式を考え出し、従来型の印刷はしないでも、必要な手に渡る新しい刊行形式を考え出すことが、専門情報流通の専門家の大きな仕事であるのに、それを真面目に考える人は少ない。考えないのはそれら専門家の怠慢としかいえない。

そこで当面考えられる方法をいくつか挙げよう。

一次情報発表技術の展開

(1) CTS 紙印刷の資料を作る、コンピュータ利用の印刷法であり、この方法の特長は、印刷した後に電子記録が必ず残るから、この記録を活用できる形式の刊行物に向いている。目下のところ、部数の多い方の印刷に使われることが多いが、①内容の修正をしばしば行う必要がある場合、②一旦印刷発行後、内容は余り変えずに別種の刊行物を作る場合、③作った記録はデータベースに転換が容易、という特長を活かすことを考えるのにむいている。

(2) DTP DTP は広義の CTS といえなくもないが、版下の作成に印刷業者を頼む必要がなく、入力にワークステーションを使い WYSIWYG の原則を生かせば、ユーザー自身の手で少数の印刷物の刊行が行える。かつての軽印刷にとって代る勢である。印刷仕上りは、活版より劣るが、内容本位の刊行物には、少しも差し支えない。

学術的刊行物(単行書を含めて)は今後この方法で作られる傾向にある。こうなれば 500 部程度で採算がとれる。またもし、著者の原稿が、著者の手でワープロ等により入力されていけば、コストは更に下がるので、もっと低部数でも引き合うようになる。

(3) オンラインデータベース

オンラインデータベースは印刷物ではないと、誰しも考えるが、公開されるものであるから、一種の刊行形式である。商業ベースのデータベース(DB)は作成経費とからめて、必ずしも安い刊行物ではないが、提供方式によってはユーザにとっては、希望のものだけを容易に入手し得る方法となる。ディスプレイ上での閲覧のほかプリントアウトもできるので(ダウンロードを禁止するベンダは近頃少いので)、立派な刊行形式といえる。

(4) 可搬性データベース

CD-ROM およびフロッピディスク(FD)で全文データベース(FTDB)を提供すれば、これも刊行形式の一つと見なせる。事実 CD-ROM の FTDB は最近方々のベンダまたは発行者が提供している。

媒体は上のどちらでも可搬性が高いので、オンラインによるよりも便利ともいえる。従って、ときどき改版の必要があり、少数しか作らない(売れない)刊行物はこれらの形式で発行することを考える人がふえている。数百部程度なら FTDB を CD-ROM で、またその一部

分を FD で、さらに小さな DB なら始めから FD で頒布することができる。今後さらに新しい類似媒体も現われよう。

IV. 一次資料の多段階刊行

読者の「きままな」読み方に、発行側はとでも対応できないというのが、従来の考え方である。読者のいろいろな読み方に対応することは

- ① それは著者の責任ではない。
- ② やればよいとしても、著者にはその努力は払えない。
- ③ 著者としてはなるべくやってみようとは思うが時間がかかり過ぎるのでできない。
- ④ 読者の要求に対応する「巧妙な手段」があり、著者の努力が過度に渡らないなら、著者はやってもよい。

すると問題は「巧妙な手段」があるか、ということになる。

ところで現実はどうなっているのであろうか。

(1) 本文だけあった時代

20 世紀の初頭をみると、発行される論文・記事には本文(全文)しかなかった。つまり著者が十分と思うテキストを読者に提供すれば、それで著者は責任を果たすと認められた。

二次資料は、それ以前から勿論あったが、それは各著者とは無関係に、二次資料の作成者がおり、この人々の考えに従って自分の思い通りに作られていた。抄録も二次資料の一つであり、著者にこのような抄録を作らせても満足なものではない。練達の抄録者には所詮かなわないと二次資料の発行者は考えていた。本著者の記憶によれば、1960 年においても、Chemical Abstracts の編集部はそう主張していた。

(2) 著者抄録を付ける時代

戦争後、刊行物は異常な増加が始まった。これは戦時中の秘密の解除のほか、科学技術者の増加による。このため、二次資料、特に抄録雑誌の発行者は抄録作成の人手不足と、作成原価の高騰に直面した。かりに専門的抄録者は得られても、対象となる一次資料の数の急増には追い付けない。抄録雑誌の発行原価の高騰は、予約価の増大を呼び、購読者の減少につながった。つまり二次資料発行の経営の危機である。

従って発行者は原価の上昇を少しでも和らげるため、著者の抄録（しかも良質の）の提供を求めようになった。そこで各著者の全文に心ず良質の抄録をつけることを呼びかけた。これは例えばイギリスの Royal Society の勧告等に始まり、ユネスコの著者抄録 (synopsis) の向上の呼びかけなどを契機として、国際的な運動が生じ、現在のように専門雑誌には全て著者抄録がのり、編集者は投稿される著者抄録の質に目を光らせる時代がきた。これにより、二次資料の作成者は危機を乗り越え、今日の DB 時代まで持ちこたえられたのであった。

勿論抄録作成の自動化は当然試みられ、数多くの工夫や研究が行われた。イギリスの Cambridge を中心とする automatic abstracting の研究は代表的な例であるが、1980年代の中頃にはついに難関に打ち勝つことができず、衰微してしまった。

今では各著者は抄録を作ることを（少なくとも公式的には）いやがらないが、これは今世紀中の大きな習慣の改革だったのである。

(3) 学協会雑誌の経営危機と分割・専門化

ところが一次資料、特に専門雑誌の慢性的危機は解消していない。研究発表及び発表論文記事を読む人口の増加は止まない。このため学協会会員の増加にも拘らず、投稿論文の全文を、会員全員に配布する前通りの方式で刊行することが困難になってきた。そのため少なくとも二通りの解決法が工夫された。

① 今日、日本の大学会で、1種類の雑誌だけで間に合うところはない。まず一般会員誌と論文誌の分裂が戦争後起り、論文誌も1点だけでなく、細かく分冊化が要求される。あるいは、ごく狭い分科の雑誌が中心部分から分裂し独立してゆく。例えば会員数が3万6千を算える某学会は、会誌の他に実に8種の論文誌を刊行し、合計1千件強の論文をそれらは含んでいる。

そのような学会では、会員の負担は論文誌を何種類受け取るかによって異なる。広い範囲の論文・記事を読み取れば、高い会費を払うという、受益者負担制となる。個人会員で全論文誌を買う人はまずない。全部を買うのは機関の図書室だけである。自分で購入していない論文誌中の記事を利用するには、図書室にある論文誌を、その機関の構成員はコピーして利用する。出版業者がというようなコピーの禁止は実際上できない。結局コピーの可能性のあるところからは、可能性に見合う超過料金を発行者（またはその代行機関）が受取るしかない。

それが各著者の利益になるのではなく、発行機関の収入増にしかならない。

この方法はある程度の成功を収めてきたが、専門の交錯と学際化により、一人々々の人が参照すべき雑誌等の点数はへらせない。従って各人の手持ち雑誌のタイトル数を再び増加させるか、または図書室等によるコピー数を増すか、或は両者の組合せによるかしかない。そこで再びコピーの問題が出てくる。

② 別の方法は、従来の方法を全く裏返しする方法で、ユーザの各個人は雑誌等を全く予約購入せず、あるところ（配布点と仮称し、その場所は別に考える）から、自分に必要な論文・記事のコピーをもらう（対価支払は必要で、その方法は別に考える）こととする。ADONIS方式はこの一つの例である。

配布点は、i) 現在の雑誌の発行元、ii) 著者個人、iii) 全国的あるいは地域的情報センタ、iv) もう少し細かい単位センタ（例えば各機関の情報資料室）と考えることができる。

i) の方法は会員数1万人程度の学協会では可能だが、1千人以下の小学協会では困難であろう。従って連合して配布センタを作ることになる。

ii) 現行の各著者に対する別刷請求制度を拡大したようなもので、不可能ではない。しかし評判の論文を書いた著者は到底処理しきれない。従って、各著者中、希望する人々から委任された配布代行センタが発生しよう。

iii) 役人的発想からすれば、全国统一センタはムダがなく（形式上）見事に運営されるから理想的だということになろう。しかしこれには大機構や大規模な機械化システムが必要となろう。

しかし英国図書館 (BL) の資料配布センタ (DSC) は年間330万件（そのうち23%は外国あて）のコピーを個別ユーザに直接提供し、その必要経費の約70%を売上で賄っているのだから、方法を誤まらなければ可能だといえよう。それにしても独占は危険で、競争のある方がよい。

iv) この方法は「単位センタ」の総数と、受持つ平均ユーザ数により可能性が異なる。一専門の資料を受取る単位センタが500程度以上あれば、源資料全体を例えばCD-ROMに収容して経済的な価格で配布できよう。

技術的問題は、各著者から配布点までの情報メディアに何を選ぶか、である。従来の印刷法の可能性を考えるのは行き詰るが、現存の手段で解決できよう。

即ち各著者が電子媒体（現在なら恐らくFD）で提供

し、これを簡単な編集の後、適切な媒体（例えば CD-ROM）に移し、一定量になる毎に配送すればよからう。

(4) 提案する配布法と読み方

この章で述べてきたことを、別の目でまとめてみよう。雑誌の全購読者（全会員）に全文を配布することは、まずあきらめよう。その代り要約論文を配布しよう。そして（ふさわしい技術手段により）全文は限定少数部数発刊に止めよう。1,800字程度のページを持つ雑誌においては要約は2ページ程度とする。

この方法は、例えば化学全体、機械工学全体というような広い専門分野ではなく、それらを分割した小専門分野で、全会員が一専門の全分野にわたって、とにかく目を通せるような場合に適している。大専門分野は適宜にこのような小専門分野に分割すればよい。

要約の量として2ページというのは、一論文の内容をこの程度の量に圧縮できないのは、各著者の腕が悪いということになる。また既に述べたように、読者は数多くの論文を通覧するときには、この程度の分量しか、まともには読んでくれない、ということでもある。

同じような事情から、部厚い単行報告書についても、別冊として「幹部むけ要約版」Executive Summary を発行することが常識となりつつある。また専門的記事ではない一般的雑誌についても、ダイジェストを別途発行して成功している例も見受けられる。

要約版だけで終る読者も多いが、多数の要約中で、これと思う記事については、精読のため、改めて全文を読むことになる。そのため全文の請求者は、全文雑誌を保有するところ（前の IV の (3) の iv) にいう「単位センタ」に連絡してコピーを入手する。

従って一つの論文・記事については

- i) 抄録, ii) 要約, iii) 全文

という三段階を各著者が用意することになる。各著者にとって負担は増すが、多数の読者の「読む」過程を改善し総体的な労力節約が可能となる。

V. テキスト加工

各著者が提供する全文が電子記録になっていると、加工がしやすい。その加工は著者の手でも、他人の手でも行える。

全文中のワードを依りどころとして検索を行うことは、現在全文検索の諸手法により行っている。これは特定のワードが文中に「存在する」ことを根拠として行な

われ原文を加工し変形するのではない。

一方で原文をワード単位で加工することは容易にできる。機械翻訳はこの原文を徹底的にいじって、別のワード集に直すことである。

そこでテキストを読む立場で、テキストに変形を与えることはできないか。つまり重要性の低い単語は飛ばして、あるレベル以上の重要度を持つワード、フレーズ、パラグラフだけを読むことはできないだろうか。

この場合問題点は、読者一人々々の持つ重要度による判断を何によって行うかということである。読者が眼と頭で読むときには、この種の判断を日常実施している。従って読者の読み方を分析すれば、その方法が見出せるはずである。今まで莫然とブラウジングと呼んできた動作をもう少し細かく観察することである。

本著者は、このことは比較的簡単に実施できると考えている。①最も粗い方法は、パラグラフ毎に、重要度の指標（例えば5段階の）を各著者がフラグの形で与えておく。読者はこの指標で希望するグレードを選択入力して、その値以上のフラグを持つパラグラフだけを読者に読みやすく示す。

②精度を上げるには、パラグラフ単位をセンテンス単位にしてもよい。また③特定の用語の出現と、重要度の指標との論理積を使うことも、また重要単語の前後 n語の部分拾い出すこともできる。

読みやすくするには、不要部分をマスクしたり、重要部分だけ色を変えたり、ハイライトしたり、採るべき方法はいくらかでも工夫できる。

さらに読者は、特定のアルゴリズムを与えて、全文を処理し、その結果を読み取ることもできる。

要するに、選択的に読むことに、情報処理の能力を活用することを実行するのである。

VI. 専門的一次情報の変貌

以上述べてきたことの結論として、各専門における資料的形式の一次情報は、今後その流通・配布形式において二つに別れるであろう。これを第一種および第二種と仮称しよう。

第一種刊行物とは、従来通りの印刷配布によるもので、多くの読者があることを前提とする。そして読者自身がその刊行物を長期に保存する可能性は低い。つまり一過性の強い資料である。社会全体としては保存の機能を持つ（図書館等の利用）であろうから、これに頼ることとなる。

第二種刊行物においては、読者の数は一刊行物については限定されるが、各読者は（業務上）読む必要を強く感じるところの、ある量の資料或は刊行物が常にあると考えている。従って小部数刊行の手段に依らざるを得ない。

もちろん両者の中間はあろうが、この両種の刊行形式がはっきりと違うので、意識して区別する必要がある。

A. 第二種刊行物のためのデータベース

第二種刊行物は、少部数印刷（複製）発行によることは当然できるが、データベース（DB）も一つの刊行形式として利用されよう。ただ少数者利用のDBが成立するためには、現在のように本文の入力経費をDB作成側（Producer）が負担せず、原則的に各著者が入力作業を実施する（もちろん他人に代行を頼むことはある）ものとしなければならない。そのためには、その種の負担にも拘らず、著者がDBへの編入を希望しなければならない。さらにその編入のためには、ある程度の費用を支払う（現行のページチャージに相当）ことも生じるかもしれない。

従ってだれもが、やたらにDBへの投稿を行うのではなく、投稿して発表するのは、それが著者の「業務」である場合である。研究者は確かに、そういう人達であり、ほとんどの場合、研究者は給与をどこかから受けているはずである。いわゆる作家は、給与を受けることなく、作成原稿から収入を得ることを前提としている。従ってこの作家たちは、上にいうようなDBを刊行形態とはしないであろうし、第一種の刊行に頼るはずである。

このような第二種刊行物については、著作権の考え方を変えなければならない。著者には知的所有権はあるが、刊行による収入はなく、むしろ刊行費用の一部を負担する者である。現行の著作権法には条文の中に、明らかに著作者の収入のことは現われていないが、財産権に近いものとして著作権を考えている。従って第二種の刊行は著作権法が全く予想しない存在なのである。従って著作権のあり方を、この種の刊行に対しては、全く変えないといけない。

B. 第二種刊行物の利用

この種の刊行物は、代表的なものが論文であり、論文は（理工系においては）一般に短い刊行物である。この点で人文分野は全く性質が違い、刊行の数も多くないので、ここでは検討の対象としない。

大抵の論文は、平均して2万字（1万語）を越えることはない、従ってデータベースも、それほど大きくはない。例えば1万件を収容しても500 M Byteの程度である。

さてこのようなDBは、記事の件数が多いので、アクセスには良い二次資料（これもまたDBとなる）がないと使いにくい。従ってこれは抄録まで含むもので現行の二次情報DBとほとんど変わらない。従って必要な文献資料を捜す人は次のようなプロセスを通るであろう。

- ① 専門別二次資料DBの利用
- ② 候補資料の書誌アドレスを知る。
- ③ 全文DBから概当記事を引出し、ダウンロードする。
- ④ 目的とする記事だけの個人用小規模DBを作る。
- ⑤ 各記事の概要を飛ばし読みにより知る。要約版DBがあれば、これを利用し、なければ自分の小DB中の全文を処理して作る。
- ⑥ ほんとうに必要な少数の全文を読む。

この⑥までのプロセスにおいて、どの道、何分の一かの要約が入用となる。最終的に全文を読む人は、例えば平均して数十人程度となろう。専門的刊行物とは、このような「非能率」の上に成り立つものなのである。

C. 付随して起る問題

(1) 「原版」の概念. テキスト加工が容易になると、一つの資料について、いろいろな程度に加工・変形したテキストが世の中に同時に存在することになる。これでは、どれが「ほんとうの」テキストであるか疑わしくなる。従って変形後の各テキストには、それと原テキストとの関係を明記しなければならないし、その記載法を標準化しなければならない。

そこで各種のテキストを、仮に「版」と呼ぶなら、各版の出自を明示し、「原版」が何であるかを常に示す方法を確立しなければならない。このような処置の下で原著者の方では、自分が責任を負う内容を明らかにすることは、知的所有権確立のプロセスとして大切なこととなる。

(2) 「公刊」の概念. 公刊 (publish) するとは、このような来るべき事態において、どんなことを指すのかを明瞭にする必要が生まれよう。

従来公刊とは、多くの人がアクセスできる形で世に送り出し、誰でも見得る状態にすることだと考えてきた。その大きな根拠となる技術的手法が「印制発行」であり、

その投稿の日時が論文の priority 主張の根拠となっている。しかも priority の問題がやかましいのが専門の原著論文であるが、将来の第二種刊行物においては何を以て投稿及び刊行の日付とすべきであろうか。この日付は一方で特許の審査における「いつ公知になったか」の問題に密接に関係してくる。

特許審査あるいは特許訴訟においては、「公知となった日付」が常に問題となる。この場合、印刷物の（例えば奥付の）「発行日」は確実な証拠とはならない。そして公然と閲覧できる場所に受入れられた日付（たとえば公開されている図書館（室）における閲覧業務担当者の押す受入日付印）が最も確実な証拠と一般に認められている。

データベースの世界においては、このようなデータベース公開の日付を明示できる習慣がまだに確立されていないが、いずれは上記の受入日付に相当するアップデート公開の日付を、公然と記録する制度を実施しなければならぬ。論文については編集者の受付日は示す習慣がほぼ確立しており、priority はこの日付により判断するとしても、この日付を直ちに公知の日付とは見なされない。

ところで専門資料・論文は上に述べたように少数刊行の方向に進むし、その形式としてデータベースによる全文刊行を行う場合には、特に「原版」の公開日付を自動的に記入する手段を取らなければならない。このような観点は従来データベース関係者の注意からされているのでここで明確にしておくべきだと考える。

少数発行においては、配布先が限定されやすい。しかし各著者の priority 主張を傷つけないためには、その刊行物の登録場所（いわば刊行物の登記所）を指定すべきであろう。それは国立国会図書館等でよいと、理論的には一言言えそうに見えて、実はその機関において閲覧可能になるまでの遅れの時間は相当なものであるから、けっして工合よくはない。このため、各専門分野毎に、公開の条件を満たし、「不特定多数」に対してではなくともその分野で「特定多数」に開示される機関あるいは機構を確定しなければならない。

公開のデータベースによる発刊の場合には、上に述べたように、各論文・記事の公開日を自動的に記録することができれば、公知の日付を明瞭にできるという点においても、データベースによる刊行は大きな特長を持つこととなり、そのような必要性の高い人達の間では評価されることとなる。

以上の点については「査読」の問題がからむが、本稿ではこの問題には触れない。

(3) 読者には加工の自由がある

今までのところ、読者がある刊行物を購入（入手）した後、これをどう取扱おうと（即ちアンダーラインを引こうと、書込みをしよう）読者の自由である。電子的資料を読者が入手した場合も全く同じでなければならない。オンラインでユーザの端末機の中に信号がはいってしまえば、これをユーザが自己の利用目的のためにどう加工しようと、提供側は文句をいうべきではない。加工したものを業として他人に提供した場合には対抗策を講じればよいのである。

電子テキストを読むためには、加工して読むのが最も工合がよいことは上に述べた通りであるから、加工を一切させない、という考え方は今は成り立たない。これは現在において、読者が購買した刊行物（印刷物）を自家用に加工（たとえば複写すると）しても、出版者は文句をつけるのは不条理である。複製再版して販売することは禁止あるいは原出版者の許諾がいるとするのは当然である。このような現在の状況を拡張して、ダウンロードを危険視するのは、どうみても不当であろう。

D. 情報送り出しの経費

以上述べてきた諸事情から読者は気付くであろうが、情報を世に送り出すには、今後経費がますますかかるであろうという問題が生じる。しからば、このような経費増大を当然であると認める理由はどこにあるのか。

従来からの狭い経済的観点からすれば、知識の獲得には経費のかかることを認めはするが、自から獲得した知識を開示することに多大の経費を費すことは、自殺的行為とする見解が日本では多い。

これに関して、特許権の確立のための経費は既に当然のことと認められている。これは特許権という経済的利得を生じ得る権利のために是認されている。しかし論文発表というものは何等の利得を生まないから、これに多大の経費を投じることは不当であると考えられるためである。

一方で広義の「情報化」の進むにつれて、「情報」を持つこと自身にある価値を感じるようになり、それにつれて、情報を豊富に持つ者は、少なくともその情報の一部を公開すべきである、という観念が徐々に形成されつつある。

現に多くの製造業者は、自身が優秀な技術を持ってい

ると広告宣伝に相当な経費を投じている。こうして消費者あるいは直接の購買者に信頼感を与えることに、出費是認の理由を求めている。また関心を持つ「特定多数」の人々に技術を進んで公表する催物を開くことも増加しつつある。また多くの企業は、その職員が論文を発表するに当って、ページチャージを支払うことを嫌ってはいない。

このように企業は、市場に与える好影響と、社会に対する還元とのため、情報に対する支出を認め始めてきた。研究投資は直接的には将来への投資であるが、一番手の諸国は、多くの試行錯誤を繰返し、自身の成果を周囲に与えながら、全般的レベルの向上を図ることによって、自身の企業目的を追求しようとしている。

ここに述べている情報の送り出しは、その行為の一環

として考えるべきものであり、それに相応な出費を行うことは十分な理由のあることである。専門的刊行物の発行は研究経費の重要な一部分と考えなければならない。

VII. む す び

将来の方向を洞察すると、情報化の奔流の中に一次情報の技術が大きく現われてくる傾向を指摘し、殊に科学技術研究の進展が生み出す一次情報問題を具体的な表現で検討した。その中から情報学の諸技術のうち、今後は（二次情報ではなく）一次情報に関して、いろいろな発展の期待が大きくその方面での発展に務めるべきことを論じた。その意味において、著者は「一次情報の復権」を一つの大切な観点であると考えている。