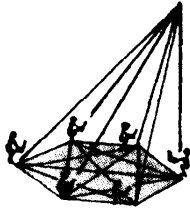


展 望

通信網の変革と情報処理

電子情報サービスの展望†

瀬 谷 重 信††



1. はじめに

グーデンベルグが印刷技術を発明して以来、活字文明は経済、技術、文化などあらゆる活動をはぐくんできたといえる。しかしコンピュータの出現によって、伝統的な活字印刷媒体による情報表現は大きく変革しようとしている。

電子情報サービスの概念は明確に規定されたものではない。本編では新聞情報や雑誌、文献などの文献情報のように活字印刷媒体に表現されている情報がコンピュータ技術や通信技術の適用によって展開されるサービスを電子情報サービスと呼んでいる。電子情報サービスは単に印刷業務が電算写植に置換され、経済化、迅速化の効果が生まれるのみではない。機械可読化された一次情報をもとにデータベースを構築したり、また文章情報データベースの一般的な検索はもちろんのこと文章中の単語頻度分析や定量化利用などの広範な領域にまで電子情報サービスが拡大されて行くものと考えられる。

以下には電子サービスの今後の展開について、利用ニーズとサービス高度化の基盤あるいは促進となる技術について展望を試みる。

2. 電子情報サービスの芽と今後の成長
—傾向分析—

2.1 印刷メディアの電子化

長い間印刷に欠かせなかったのが鉛の活字であった。今やワードプロセッサやコンピュータが鉛に取ってかわろうとしている。

最近のシステム例をみるとワードプロセッサと電算写植機が連動し、スピードアップや省力化の効果を発揮している。対象も雑誌から百科事典づくりにまで電

子化の波が広がって行きつつある。このような編集、版下作成等の印刷業務のシステム化は単に印刷工程のスピードアップや省力化効果をもたらすだけではない。情報を従来のように、紙という媒体で一過性のものとして生産する方法から電子的に蓄積することによって、多面的な情報の付加価値が生まれてくる。印刷メディアの電子化のねらいはこの一次情報を電子化することによって生まれる新たな情報価値を高めることにあると筆者は考えている。

新たに生まれる情報付加価値とは、次の2つの傾向を持つと思われる。詳細は後述するが、1つは印刷メディア以外にコンピュータ端末、CATV、ビデオテックス、音声サービス等のマルチメディア媒体による情報表現を可能とすることであろう。さらに、第2の傾向はデータベースに代表されるように一次情報をもとに索引語の付与、抄録作成の自動化、文章情報の定量化分析など電子化された情報内容に高度な付加価値が生まれてくる。

このような傾向が促進されるための基本的前提条件となるのが印刷メディアの電子化である。

米国印刷業界のコンピュータリゼーションの伸展は我が国の展望に有用である。その一例を米国クライン社の調査レポート「コンピュータ編集技術の印刷産業に対する影響」でみると、米国の印刷産業界は1986年までの5か年間に、約11億2500万ドルのコンピュータ写植・編集システム機器を導入する見通しである。市場全体の78%（8億7500万ドル）がコンピュータ編集システムであり、他にコンピュータ写植機（1億5000万ドル）などの導入がはかられ、一層印刷メディアの電子化に拍車がかけられる傾向にある。

このような印刷業界のコンピュータシステムの導入は文章情報における機械可読可能な一次情報の生産性を飛躍的に高めることになり、前述のような2つの利用傾向が進むことは必要と考えられる。

紙に印刷するかわりに電子メディアを使用して読む

† Future Trend of Electronic Information Services by Shigenobu SEYA (Data Communication bureau, N.T.T.).

†† 日本電信電話公社データ通信本部

ペーパーレス雑誌、文献はさらに革新的な変化をもたらすことになろう。

電子雑誌（エレクトロニック・ジャーナル）の“創刊”が米国で進められているという報道がある。これは米国コムテック社の F. プロトキン社長（前コロンビア大教授）によって推進されているシステムであり、天体物理学、原子力工学、遺伝子工学などの 37 分野で進展中の実験・研究レポートを蓄積しオンラインによる検索を可能とするシステムである。レポート投稿者はワードプロセッサによりレポートを作成し、フレキシブル・ディスク（FD）を編集者に送る。編集者によって評価採択されたレポートがシステムに入力され、購読者が端末装置から検索する仕組みである。本システムによれば従来、原稿が出版されるまで 3 カ月から 1 年程度要していた期間が 2 週間以内に短縮されるメリットがある。一方では入力されるレポートの内容について質の低下が避けられないと既存雑誌の編集者たちが批判しているという。このような新しいシステム化の試みは単に技術的可能性や経済的メリットの観点からのみならず、情報内容あるいは社会的な受容性の双方から評価され導入されて行くものである。

2.2 新聞情報システムのマルチメディア化

新聞情報システムが先進エレクトロニクス技術のシステムの高度化に組み込もうとする視点は 2 つある。1 つは他のシステムと同様に効率性、経済性追求のためのシステムづくりがねらいである。新聞の製作工程にコンピュータを導入して工程の短縮、効率化、経済化をはかることが目的である。第 2 のねらいは、新聞ニュースの商品性に注目して、従来紙面で提供・販売していた情報を印刷メディア以外の各種のメディアで付加価値をつけて提供する、すなわちマルチメディアの方向である。マルチメディア化は第 1 のねらいである新聞制作のコンピュータ化が前提となるが、データベース、ビデオテックス、CATV、文字多重放送など多様な形態、媒体でのニュース情報サービスが進行あるいは企画されている。

新聞製作工程へのコンピュータの導入は以前から行われており、日本経済新聞社の ANNEX システムや朝日新聞社の NELSON システムなどがよく知られている。大手新聞社のみならず地方新聞社においても最近コンピュータ導入が積極的に検討されている。新聞製作工程のコンピュータ化とは従来人手作業で行われていた校正・組版の過程をコンピュータ処理に置換えて活字を追放し、記事や写真などの紙面情報をすべて

符号とイメージデータで取り扱うことにしたことである。新聞記者の書いた原稿がキーボード端末から入力されシステムに蓄積される。記事 ID をもとに入力記事の訂正（削除、挿入、差換えなど）をする。次に記事の見出しを付与した後、紙面のレイアウトを決定する（前組み、大組み）。紙面が完成すれば各地の印刷工場に紙面フィルムがファクシミリ伝送されて印刷され、各家庭に配達される仕組みである。

コンピュータ技術や通信衛星技術の進歩は新聞製作工程の上流工程である紙面編集のシステム化をベースにして、今後下流工程さらには販売の仕組みにまでも大きなインパクトをもたらす可能性がある。米国では経済紙「ウォールストリート・ジャーナル」（ダウジョーンズ社）や「ニューヨーク・タイムズ」紙以外にも「USA トゥデイ」が通信衛星を使って送信・発行している。我が国でのインパクトがどのように出てるか注目したい。

新聞情報を紙面以外の形態で提供するサービスとして大いに注目したいのがニュース・データベースサービスである。新聞が一過性のサービスであるのに対して、ニュース情報を蓄積して、索引語を付与することによって利用者ニーズに合致した情報を選択して検索可能とした点が特徴である。国内では日本経済新聞社がオンライン記事情報サービス NEEDS-IR を提供しており、新聞・雑誌記事の蓄積件数は 73 万件に達している。このような記事情報のデータベースは、新聞製作のコンピュータ化が促進されるにしたがって必然的に増加していくと予想される。現状のデータベースの利用者は主として企業中心の、いわゆるビジネス・ユースが主流である。しかし、今後我が国でも CAPTAIN サービスの商用化をきっかけに家庭向けの情報サービスの実現が予想される。ホーム・ユースのニュース・データベースサービス企業例として Compu-serve 社をあげることができる。この企業はワシントン・ポストやニューヨーク・タイムズなど主要 10 紙の新聞記事や AP 通信サービスの最新ニュースさらには株式、債券、商品ニュースなどおよそ 4 万件以上の情報を家庭に提供しているという。このように、伝統的な新聞という印刷メディアで提供していた情報を、ユーザーニーズにマッチした形態に加工・編集して、各種のメディアを通して提供することが新しいビジネス・チャンスとなるものと予想される。

2.3 文章情報利用の高度化

新聞、雑誌にかぎらず現在印刷メディアで流通して

いる文章情報が電子化（コンピュータ可読化）されることによって、多角的な利用方法が期待できる。

まず、前述のように、第1の利用高度化は一つの文章情報を各種のメディアに変換して提供するマルチメディアサービスの形態である。CATV、文字多重放送、ビデオテックス、音声出力、FAX 同報サービス等による情報提供形態である。

第2の利用形態の高度化は文章情報を蓄積し、検索に供するデータベースサービスである。現状におけるデータベースサービスの最大の問題点は原文献からのデータベースづくりをいかに経済的に行うかであり、また原文献の発生から、オンライン検索サービス提供までの期間をいかに短縮するかである。

この問題解決には最近技術進歩が著しいコンピュータによる日本語文章処理技術の導入が期待できる。もちろん、この前提として文章情報がコンピュータ可読化されている必要がある。現在人手に依存しているデータベースの索引語付けの作業は日本語文章処理技術によって省力化、迅速化の効果が期待できる。その処理の流れは次のとおりである。文章情報を入力して、単語認定、品詞判定、複合語の分割等の形態素解析にもとづいて、ソーラス、ストップワード（不要語）の辞書を参照しながら該当のキーワードを抽出する。米国においては国防省ドキュメンテーションセンター（DDC）で採用されているシステムのように、年間1,000万語の技術文献の標題および抄録から術語辞書を用いて索引語となり得る単語列を抽出するシステムが1969年から稼働している。さらに、科学情報協会（ISI）のシステムは1972年より実用に供されている。論文標題の中の語から不要語を除き、標準の綴りや表記法に統一し、複合語の場合は語順を統一するなどの処理により索引語を作成している。

国内でも日本語解析技術の進歩を背景に索引語抽出システムの検討が進められている。日本語ソーラスも各分野で充実されつつあるので実用システムの可能性が現実のものとなってきている。

このようにデータベースサービスの中でも特に日本語データベース構築の経済化という障害が克服されてデータベースの量の拡大がはかられば多くの分野でのオンラインによるデータベース利用が活発になると予想される。

第3の文章情報の利用高度化の方向は文章を構文的、意味論的に解析して、コンテンツ・アナリシス等の定量分析手法を使って文章情報から潜在的な事象や

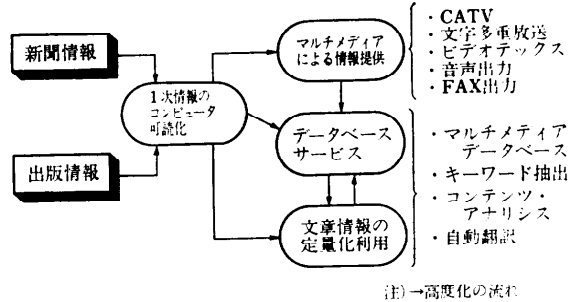


図-1 電子情報サービスの利用高度化のプロセス

傾向のは握に役立てようとするものである。現在、通産省が「文章情報データベースの総合利用に関する調査研究」の中で文章情報の定量化利用の調査研究、システム開発に取り組んでいる。

コンテンツ・アナリシスの手法には、文章中から主題や重要語（キーワード）を抽出して頻度分析を行う単純な方法から政策決定のプロセスをモデル化して数量化された文章情報をそのモデルに組み込んで使用するなど利用形態は多様である。この利用分野はカントリーリスク分析、エネルギー問題分析、通商問題分析、経済安全保障分析などの政策決定のための支援システムとして期待できる。また企業経営の支援システムとして、マーケティング、新製品開発計画、広告等の分野への適用が考えられる。

すなわち、文章情報が数値統計データを補完して各分野の予測、計画、意志決定の基礎データとして有用性が増していくということである。

以上、考察してきた新聞、出版情報を中心とした情報サービスの今後の展開を構造化すると図-1となる。

3. 電子情報サービスの高度化のための技術課題

新聞、雑誌、書籍など伝統的な印刷メディアによる情報が制作工程で電子化され、多様なメディアによる情報提供が促進することは必須である。この電子情報サービスをさらに活性化するための技術課題を以下に考察してみる。電子情報サービスはその性格から相互に密接にかかわり合う、電子情報生産サブシステムと電子情報流通サブシステムに分けられる。一方サービスを実現するためのシステム構成技術を次の4種の機能に区分することができる。それは、利用者との間で情報の入出力を効率的に行う端末機能、高速・経済的

表-1 電子情報サービス高度化のための技術開発課題

ネットワークシステムの構成機能	端末機能	網機能	通信処理機能	情報処理機能
サービスシステム				
電子情報生産サブシステム	日本語情報の入力 漢字 OCR 音声入力 図形・画像情報入力			日本語知能処理 キーワード自動抽出 シソーラス作成 自動翻訳 認識処理 音声 文章 画像
電子情報流通サブシステム	端末装置の多様化 電話(音声入出力) FAX(パターン入出力) 図形・画像 端末のインテリジェント化	ネットワークのデジタル化による高速、高品質化 網間接続 情報蓄積 同報通信 代行受信	コード変換 速度変換 プロトコル変換 メディア変換 漢字カナ変換 カナ音声変換 文字コード・パターン変換 情報蓄積 クリアリング機能	DBMS のファームウェア化 データベースマシン 分散型 DB アクセス プロトコル DB の結合処理 コンテンツ・アナリシス

な情報伝達機能を実現する網機能、情報蓄積・形式変換等の通信処理機能、さらにコンピュータセンタによる情報処理機能である。

表-1 は電子情報の生産および流通の各サブシステムのサービスの高度化をはかるうえで4種のシステム構成技術に求められる開発課題を示したものである。

3.1 電子情報生産システム高度化のための開発課題

コンピュータ可読な一次情報の量的増加を促進することが電子情報サービス高度化の必須条件である。それには新聞、雑誌や各種文献の製作工程を電算化することである。しかし、現状ではコンピュータ可読な形態となっていない文献が多いので文献、文書、資料等を直接機械で読み取りコンピュータに入力する技術が情報生産量を拡大するには必須となる。

漢字 OCR の商品化が現在進められているが、印刷された漢字、かな混り日本語文書についてみると、漢字認識技術および認識率を向上させる日本語処理技術の進歩により、正読率も 99% を確保することが可能となり、実用レベルに達してきたといえる。

漢字 OCR による入力技術とキーワード抽出技術とを組み合わせることによって単に一次情報の機械可読化のみならず、データベースの作成をも可能とする。

音声入力の実用化はマンマシンインタフェースの向上をもたらす。さらに図形・画像情報の入力技術の開発や蓄積、認識処理技術の開発は本格的な図形・画像データベースの実現を可能とし電子情報サービスの領域を広げることになる。

3.2 電子情報流通システム高度化のための開発課題

情報生産サブシステムによって生成された情報を通

信ネットワークを介して利用者のニーズに合った情報を、利用者が求める形態で選択あるいは検索可能とすることが電子情報流通システムの課題である。

この課題実現には次の観点からの技術開発が必要となるであろう。

(1) マルチメディアサービスの実現——端末機能の多様化——

端末装置が多様化していくと従来の出版物をディスプレイで見るとともに、液晶、エレクトロルミネセンスなどを利用した高機能端末や、自然画表示可能な高機能端末等の開発が必要となるであろう。

また情報の検索結果を従来のプリンタやディスプレイに出力するだけでなく CATV やビデオテックスサービスのように TV スクリーンあるいは音声情報や FAX パターン情報として出力するなど、同一のデータを複数のメディアに出力するマルチメディアサービスへのニーズが高まってくる。これを実現するには漢字混じりの文章情報を解析し、漢字カナ変換、カナ音声変換、アクセント付与などの日本語知能処理技術が必要になってくる。

(2) 検索システムの高性能化

現行の汎用計算機のもとでのデータベースマネジメントシステムでは検索性能の向上に限界がある。データベースの特性である多量データの蓄積・追加更新、検索処理の高速化、効率的実行を図るためにデータベース処理に適した専用計算機、すなわちデータベースマシンの開発実用化が必要とされてきている。

(3) 分散型ネットワークシステムの実現

1つの情報センタですべての領域の情報を蓄積し、提供することは不可能である。したがって計算機間通信機能を利用して複数の情報センタを結合し、利用者

には関連する複数のデータベース検索の便を供する分散型ネットワークシステムが求められてくる。分散型ネットワークシステムは端末機能、総機能、通信処理機能さらに情報処理機能が最適化して構築されるものである。

情報処理機能には異機種の間通信および分散データベース相互の共用を可能とするための各種の標準通信規約（プロトコル）が必要となる。しかし、標準的プロトコルを採用していないシステムの相互接続には、それぞれのシステムで用いられているプロトコルを相手システムのプロトコルに変換する機能が必要となる。これは通信処理機能で実現する。

電子情報サービスは新聞、出版、雑誌などの多量の一次情報を流通の対象とすることからネットワークに広帯域あるいは高速の加入者ケーブルの導入と拡大が流通促進のために必要となってくる。

4. あとがき

いかにコンピュータリゼーションが進もうとも、新聞、雑誌、書籍が消える世の中にはならない。多様な媒体に情報が表現され、相互に補完関係がつけられると筆者は考える。まさにメディア選択の時代ではなからうか。確かな目を持つ消費者とならねばならぬ昨今ではあるが、情報を見る目も同じ。電子情報サービスの到来は情報消費者の主体性をますます必要とするのではないだろうか。

参 考 文 献

- 1) 文章情報データベース総合利用調査研究報告書。(財)日本情報処理開発協会(昭和58年3月)。
- 2) 瀬谷重信: INS(高度情報通信システム)におけるデータベースサービスの展望, 情報管理, Vol. 26, No. 3 (1983)。

(昭和58年7月12日受付)

