

解 説

3. データベースの構築と利用技術の動向

3.1 科学技術情報のデータベース構築,
サービス, 利用の技術の動向†

中 井 浩††

1. データベース技術の発展とデータベース・
サービスの成立

データベース技術は電子計算機の登場以前にその原型が生まれている。19世紀に生まれた写真術は Eastmann Kodak (当時の社名) の写真フィルムの開発によって新しい時代を迎えた。そして 1920 年代に普及した 35 m/m のロール・フィルムを用いて、マイクロ写真が 1930 年代、第二次世界大戦中に実用期に入った。また映画がトーキーとなるに当って、フィルム上に音のトラックを設け、音をフィルム上の光学情報に変換し、再びそれを音に再変換して再現する技術が樹立した。この技術が情報検索に用いられた。米国農務省図書館長であった R. Shaw は、国立標準局 NBS を中心として 1940 年代に Rapid Selector という検索システムを開発した。これは文書を 35 m/m フィルム上に写し、その検索コードをサウンド・トラックと同様にフィルム上に写す。そのコードを光電管で読みとって必要な文献を見つけると未感光のフィルムをスタートさせ、必要な文書のイメージを転写し、後で現像・拡大して利用するものである。米国中央情報局 CIA は IBM のカード統計会計機を用いて検索システムを開発した。Eastmann Kodak はフィルム・チップ上に文書とコードを写し込み、コードによって必要な文書のチップを見つけ拡大・プリントして利用するという mini-card system を開発した。Rapid Selector が sequential file の前身であり、mini-card が random file (access は random ではないが) の前身である。この後 1950 年代にわたって多くの情報検索システムが発表された。しかし本格的な実用システムには発展しなかった。

1960 年代に入って文献の機械編集を電子光学的出

カシステムが普及しはじめた。これを用いて政府刊行物や政府からの刊行助成金を得た科学技術文献目録(抄録誌を含む)その他の学術的出版物の機械編集と、電子光学的なオフセット版下フィルム作成が広く行われるようになった。この印刷用マスタ・テープから、編集・印刷用の情報その他不必要なものを除き、検索に適したフォーマット変換を行えば容易に情報検索用マスタ・ファイルが作成可能である。

一方 1960 年代に電子計算機は第 3 世代に入り、またメモリが急速に安くなった。計算速度が高まり、OS の外部メモリのアクセス機能が高まり、さらにコンパイラの利用によってファイル操作が容易となった。

データの供給、ハードウェアおよびソフトウェアの環境が整うにつれて、1965 年ごろから主としてバッチモードの検索サービスが開始された。しかしこれはまだ商業的サービスではなく、大学の計算機センタのサービス・メニューの中の一つとして行われたものであり、学外への地域サービスの形で一般にもサービスされた。

1963 年米国国防省の Advanced Research Project Agency (ARPA) が MIT に開発委託をした MAC 計画に刺戟をうけて開発された時分割処理技術により、1965 年 General Electric は TSS 計算サービスを開始した。そしてこの分野に CDC, Tymshare 等が参入した。1971 年 Tymshare はその TSS ネットワークに、他のホストコンピュータを接続可能にし、VAN のサービスを開始することを発表した。その最初のユーザとして米国立医学図書館 (NLM) が申込んだ。

NLM は 1962 年、その刊行する医学の文献索引誌の Index Medicus の機械編集を開始し、その全工程と磁気テープを記憶媒体とする検索システムを含むトータルシステム Medical Literature Analysis and Retrieval System (MEDLARS) を開発し稼働させていた。そして米国内および欧州に数カ所、日本に 1 カ所 (日本科学技術情報センター) の地域サービス・センタをおき、国際サービス・ネットワークを形成して

† Current and Future Problems and Trends in Technologies for the Organization, Service and Use of the Scientific and Technological Database, by Hiroshi NAKAI (Department of Communication, Faculty of Human Sciences, Tokiwa University).

†† 常盤大学人間科学部コミュニケーション学科

いた。そして 1965 年に成立した医学図書館援助法に基づき、その全工程をランダム・アクセス・メモリを用いるシステムに改善し、検索システムもオンライン化を進めていた。(この全システムを MEDLARS II と呼び、オンライン・サービスを MEDLARS On-Line (MEDLINE) と呼ぶ。) そのオンライン・サービス回線に公衆電話網とテレックス網(制限された範囲で ARPANET)しかなかった。Tymshare の VAN 計画はきわめて好都合であり、その利用によって 1972 年のサービス開始年においてそれまで 20,000 件/年の利用数が 200,000 を越しその後も急速に増加した。これはデータベース・サービスが商業的に成立しうる可能性を示した。そしていくつかの商業データベース・サービス(現在の DIALOG, SDC 等)が 1972 年よりサービスを開始した。

データベース技術は、1969 年の CODASYL の DBTG の報告書以前にも、インハウスのビジネス計算や科学計算のためのデータ管理のためのデータベース管理システムが開発され、1970 年に Codd のリレーシステム・データベース理論が提唱されたが、データベース・サービスのためのシステム、いわゆる情報検索システムは、それらとは離れて独自の道を歩んできた。その理由はデータベース・サービスが要求するシステムへの要請が次のような特徴をもっていることによる。

- (1) 多くのフィールドが可変長である。
- (2) ほとんどのフィールドの属性が文字列である。
- (3) 多数回(実質的には不定回)の繰返し項目が多数存在する。
- (4) 一つのレコードが異質で多数のフィールドから構成される。
- (5) 大容量(数 Gbyte から数千 Gbyte)である。
- (6) 演算は論理演算が主体である。(実質的にはレコード識別記号ないしはポインタのアドレス・データの集合の集合演算である。)
- (7) 一つのフィールドのフィールド値を単位とする演算のみでなく、フィールド値を表現する文字を単位とする演算を必要とする。(文字列探索や距離検索 proximity search。)
- (8) 高トラフィック(数十のアクティブ端末の同時処理が可能であること)。
- (9) 速いレスポンス(いかなる処理であっても送信または carriage return から応答出力開始までに 2 ～ 3 秒)。
- (10) 数多くの回答出力フォーマットに対応した出

力編集機能が必要である。

これに対し、Codd のモデルにしても比較的簡単な構造の、定型的なデータしか扱っていない。またリレーショナル型といわれる DBMS も、データベース・サービスの要求する信頼性(システムの強靱さ)をまで証明していない。このため、現在のところデータベース・サービスに用いられているデータベース・システムは情報検索で十分に使いこなされたアーキテクチャをもっている。

1960 年代米国の目覚ましい技術開発に圧倒されていた欧州諸国は 1970 年代に入り、特に米国の VAN が国際通信回線に乗って欧州にサービスを拡大しようとするのに対して強い警戒心を抱いた。これはかつて 1960 年代にコンピュータのハードウェア産業を IBM に独占され、欧州のハードウェア産業の根が育たなかったことに対する反省でもあった。欧州共同体は 1973 年に科学技術情報活動の強化に関する行動計画を欧州議会に提出し承認を得た。これは欧州のデータベースの構築、情報技術の開発、そしてネットワークの構築を内容とするものであった。このネットワークは、欧州各国の PTT が提供するパケット交換網を相互に結ぶ汎欧州のパケット交換のネットワークであり、各国の PTT の形成するコンソーシアムが EC 委員会と契約して請け負った。これは 1980 年に開通した。そして EC 委員会はこの EURONET の上に欧州のデータベース・サービスが自由に参加(あるいは脱退)できるサービス・ネットワーク Direct Information Access Network in Europe (DIANE) を組織した。そして欧州各国のデータベース・サービスは DIANE を通してそのサービスを展開した。(現在 EURONET は EC の手から離れ EC 各国の PTT の自主運営に切り換えられた。)

またすでに欧州のいくつかのデータベース・サービスは国際通信回線を通して欧州以外にもサービスを拡大した。そして米国、日本、東南アジアにも市場を拡大しつつある。

2. データベース産業の成立に向けて

データベース・サービスの市場の分析はここ十数年の間に数多く行われ、それらはデータベース・サービスが着実に成長しつつあることを示している。しかしこれが産業として安定するにはまだ多くの問題がある。この議論においてまず注意しなければならないことは、データベースに多くの種類があり、そのそれぞれがデータベース化、データベースの構造、市場の大

きさ, 利用ソフトウェア, コスト, 市場特性等の異なるということである。けっして単一視してはならないし, また単純な議論をしてはならない。

データベースには大別して次の二つの種類がある。一つは株価や市況情報のように, 単一あるいは同種のデータのみで十分大きい市場を獲得できるものである。この場合は新聞業や放送事業のような従来のマス・コミュニケーションとよく似た特性をもっている。これに対し一つのデータベースでは十分な市場が確保できず, 他の多くの(必ずしもその個々は採算ベースに乗らないものであっても)データベースと共存することによって全体として大きい市場を確保できるようなデータベース群がある。文献データベースや多くのファクト・データベースはそれに属する。情報を必要とする者はけっして単一のデータベースで要求が充たされることはない。多くのデータベースから必要に応じて必要な情報を入手する。そのためサービス・メニューに多くのデータベースを持つものがより多くの市場をシェアする。この効果を *scope merit* と呼ぶ。

またデータベースの中には常に古いデータを新しいデータで更新し, 古いデータをアクティブ・ファイルから消去してよいものと, 古いデータの上に新しいデータを追加していくものがある。前者には市況データ, ミクロ経済データ, 信用データ, 辞書データ等がある。後者には文献データ, 学術データ, 法令・判例データ等がある。前者はファイル・サイズの極端な増加はなく合理的なサイズで運営が可能である。それに対し後者はファイル・サイズが次第に大きくなり, そのコントロールが政策的にも技術的にも問題である。しかも情報発表量は年々増加し, 問題をより深刻にしている。しかもデータベース間の競争ではより多くの年限についてのデータを保有しているほうがより大きい市場をシェアできるのである。

科学技術データベースは, この横方向の拡大(*scope*の拡大)と縦方向の拡大(時間軸に沿ってのデータ量の増加)に耐えなければならない。そしてこの両方向での機能を持つがゆえに, 社会の記憶機構でありうるのである。しかしこの両方向ともにその拡大はコスト効果を劣化させる方向に働く。情報の利用は時間と共に指数関数的減衰(寿命曲線)を示す。すなわち古いレコードの利用頻度は急速に減少する。また情報需要者の要求を充たすに十分な *scope* を持つためには, 利用度の低いデータベースをも抱える必要がある。このコスト効果の劣化に耐えて, 経営業績を確保するだ

けの市場を確保するには, データベース・サービスのコストを下げる以外にない。

従来, ほとんどのデータベース・サービスは, 印刷物媒体による先行サービスを持っている。そして最初には印刷物媒体のサービスの副産物(印刷方式が機械化するにともなう)とみなされ, コストの大半は印刷物媒体のサービスが負っている。しかし, オンライン・サービスが成長し次第に売上げの中に占めるオンライン・サービスの比率が大きくなり, 印刷物媒体の売上げを越す傾向を示すに至って, オンライン・サービスにおけるコスト構成が従来とは違ったものになってきた。

データベース・サービスには, データベースを供給するデータベース供給者と, それをオンラインで提供するオンライン・サービス(ヴェンダ *Vender*)がある。そしてヴェンダの提供するサービスをエンド・ユーザに代って検索し提供する検索代行者がいる。このデータベース供給とオンライン・サービスは一般には異なるけれども, いくつかのデータベース供給者はみずからオンライン・サービスを兼ねている。データベース供給者においては, 次第に増加する世界の情報発表量を消化し, 漸減する印刷物媒体のサービス市場をオンライン・サービスへ向けての機械可能データベースによって回復し, さらに拡大することが問題である。オンライン・サービスにとっては前述の横方向・縦方向に拡大するコストの重圧に耐えてより大きい市場を確保することが問題である。

オンライン・サービスの市場は着実に拡大している。しかしまだ揺籃期にある。かつて近代産業がまだ揺籃期にあったとき, 製鉄業, 電気事業, 鉄道事業, 道路建設, 通信事業, 航空・輸送事業等, 基盤産業は国営ないしは公的な機関によって運営された。そして産業が成熟するにつれて, これら基盤産業も次第に民間に移行した。情報産業もまだ揺籃期にあり, その産業基盤はまだ国家ないしは公的な整備が必要である。

3. データベース構築における諸問題

データベース・サービスは情報をもつ多くの人たちを情報源とし情報を必要とする多くの人たちを受信体とする“情報の流通過程”という新しい形のコミュニケーションである。しかし情報源から受信体への一方的な情報の流れではなく, 情報を入手した人はそれを用いて, 新しい情報の創造を行う。創り出された情報は新たな情報として流通過程に送り出される。すなわちこの情報の流れはサイクルを形成する。この情報の

生産サイクルの中で科学技術が発展し、文化が形成されてきた。現在、軍事にせよ経済活動にせよ、新しい製品を造り出すためにもまたより合理的な生産技術を生み出すためにも、激しい研究開発の競争の中にある。そして経済発展、特に付加価値生産の大きい部分は、この知的な活動、すなわち情報生産のサイクルの効率に大きく依存する。データベース・サービスは単なる社会の記憶機構であるだけではなく、情報の生産サイクルを通して産業全体のインフラストラクチャであり、またそれと同時にそれ自身一つの産業を形成しようとしている。この二面性は重要である。

データベース構築において問題となるのは次の要素である。

- (1) データの収集、評価、選択の機構の整備。
- (2) データの分析、データベース化のための知的なマンパワーの確保と、それと平行してこの過程を代行する知的システムの開発。
- (3) データベースの品質管理と、データベース間でのデータの互換性のためのデータ表記・表現の統一化・標準化。
- (4) データベース構築のためのコスト低減と、より合理的な機構の整備。(収集・評価・分析・データベース化等の業務の分業体制や協力態勢の推進等。)

データベースの構築においてまず問題になるのは情報源の性格である。文献データベースの場合、研究者は自分の業績を積極的に発表しようとするし、またそれを刊行する出版業者も情報の流通過程に積極的に参加する。この場合、情報源は能動的である。これに対し、みずから積極的に情報流通過程に参加しようしない情報源、特に情報流通に対してネガティブな態度をとる情報源がある。自分の研究のために測定しても他人の利用に供することを全く考えていない学術的データ、公害や風土病に関するデータ、アンケートやインタビューによって入手できるデータ等々である。このような情報源は受動的である。能動的な情報源に対する情報収集操作は受動的でよく、評価と選択が必要な程度である。それに対して受動的な情報源に対しては積極的収集戦略を立て収集のための操作が必要である。これまでのデータベース供給は、収集の容易な積極的情報源からのデータに偏っていた。しかし今後のデータベース・サービスにおいては受動的情報源に対する収集が重要である。

化学物質、新素材、合金等の種々の物質に関する物理的、化学的、生物学的性質や構造、スペクトルに関

するデータは最近注目されているが、これらのデータの測定や評価を専門に行う機関はなく、多くのデータ発生は研究や生産活動の中でそれぞれの目的をもって行われる。これを集め、評価し、データベース化する機構が必要である。かつて(現在も)このようなデータ供給の重要性を認識した研究者が、自己の研究を超えてデータ集積を造り提供システムを開発し、みずからサービスを試みたいいくつかの事例がある。目下のところ、学術データの提供はこのような研究者の自発的努力に負うところが大きい。しかしこのような状態は永続的なものではない。

マンパワーの確保は深刻な問題である。情報の発表量は科学技術文献においては年率 7~8% で急速に増加し、データベース化のためのマンパワーもそれに比例して拡大しなければならない。従来、科学技術論文のデータベース化は、研究者のサイド・ビジネスであった。しかし研究活動が大規模になり研究者の間の競争が従来よりも激しくなるにつれて、データベース化の要員を確保することは従来よりも困難となりつつある。このマンパワーは、それぞれの科学技術の専門分野の知識を持つ必要があるから、その人材を安定して確保するためには、財政的にも組織的にも安定した運営環境がなければならない。

この問題の解決のために次のような措置が必要である。

- (1) 従来の現役研究者のサイド・ビジネスから、研究開発に従事していた停年退職者や家庭に入った婦人等の、データベース化を半ば専門とするマンパワーの養成と組織化。
- (2) データベース化の生産性を高めるための支援システムの開発。その延長としての半自動化システムの開発。

(3) データベース化の効率化、生産性の向上による、コスト低減の努力と併行して、データベース化のための人件費を、半ば専門として生計が立てられるレベルに引上げること。

世界中のデータベース供給者は、これらの問題に対してそれぞれの努力を行っている。

4. データベース・サービスにおける技術的諸問題

4.1 down loading, easy system, intelligent system

科学技術データベースのサービス、特にオンライ

ン・サービスにおいて端末利用者が検索結果を手許のFD その他のメモリ内に保存する down loading は、もはや避け得ないとの認識が深まってきている。このため多くのデータベース・サービスは down loading を前提として価格政策をとっている。down loading は従来の著作権や公正使用の概念では捉えられない新しい情報利用形態であり、法制度上の研究課題である。

この down loading の普及する傾向の中で、いくつかの技術的課題がクローズ・アップされる。まず第一に、down load する最大の目的は、一つの研究のためには多くの文献やデータが必要であり、特に学術データの場合は、それらのデータを組み合わせたり、なんらかのソフトウェアによって処理するための原データとする。このときいくつかのデータベースから得るデータの単位系や統計におけるサンプリング方式等、データの表現や特性に互換性がないとき、利用者の要求を充たすことはできない。この解決のためには

(1) データの表記・表現の標準化。

(2) down load したデータ間の調整を行い、統合利用が可能とする支援システム。

が考えられる。前者については最近工業技術院においてもデータベースの標準化に着手したがその実現にはまだ時間を要する。現実的な解決として後者についての多くの試みが期待される。

マイクロ・コンピュータの普及は、オフィスや研究所での新しい tool となり、種々の user friendly なシステムが造られて新しい環境を造り出しつつある。オンライン・サービスの黎明期においては、オンライン・サービスはエンド・ユーザみずからが利用するため、バッチ・モード検索のときのような質問分析と質問式作成の要員や図書館員のような中間介在者が不要になるのではないかとこの憶測があった。しかし現実とは逆であった。オンライン・サービスを使いこなすとは、

(1) データベースの特性：どのデータベースにどのようなデータが、どのフィールドに、いかなる基準でフィールド値を与えられて、どのような記号表現をもって格納されているか。

(2) データベース・サービスの特性：どのような手法で Log-on, Log-off し、どのようなコマンド言語を用いて、アクセスし会話するか。

(3) どのような質問にはどのデータベースにアクセスすればよいかの知識。

これらの知識を数多くのデータベースの内容とオン

ライン・サービス、サービス・システム（たとえばコマンド言語等）について熟知しないと効率的で効果的な検索は不可能である。このため“searcher”と呼ばれる職種や、“検索代行業”と呼ばれるヴェンチャー・ビジネスが生まれた。このこと自体はデータベース・サービス産業の形成の上で、それを業とする人々の増加、分業体制の形成として有意義である。しかし一方、机の傍らにキーボードとグラフィック端末を置いて、従来多くの文献やデータ集を使って研究者自身が自分で種々のデータを扱っていたように、多くのデータベースから種々のデータをとって即座に自分のソフトウェアで処理できる環境を与えることは、これもまたきわめて重要である。

高機能の検索やデータ操作を、エンド・ユーザ自身が容易に行える easy system の開発も行われつつある。ユーザ支援システムや、データベースを効率的に利用するためのデータベースの構築も含めてこの面でのより多くの努力が望まれる。

4.2 DBMS と情報検索, 文字列・数値・図形データの統合利用システム

現在の情報検索システムから派生している一連のデータベース・システムは、対象とするデータベースを指定し、そのデータベースに対して検索操作が行われる。かつての文献を対象とする検索において、検索操作を単純にするためにとられた措置である。しかし前述の down loading で述べたように、情報を必要とする者はいくつかのデータベースから異なるデータを入手する必要があるし、また一つのデータベースから得た結果に基づいて他のデータベースへのアクセスが時には連鎖的に、時には試行錯誤的に行えることが必要である。このため多くのデータベースが同時に open され、リレーショナル・データベース操作の結合その他の操作が可能でなければならない。しかし一方現在のリレーショナル・モデルではデータモデルが定型的であり現実のデータベース・サービスにおいて扱われるような複雑で不定型のデータ構造を扱うことが困難である。この問題の解決は早急に検討される必要がある。

研究開発においてデータを扱う表にとっては、単に文献のみでなく、種々の事実 (factual) データ、数値データ、図形データ、映像データ、さらに音声データ等を自由に扱っていくことが必要である。データベース・サービスはこれの中で、やっと文献データについてやや満足ができる規模のサービスが行われているが、

数値や予定データについてはまだ不十分である。図形・画像・音声についてはまだ開発段階である。

Chemical Abstracts Service の化学物質の構造ファイルとその全構造および部分構造検索システムでは、物体の化学構造式の図形表現とそのパターン・マッチングを可能にしている。フランスの Télésystem Questel の DARK system でも同様の機能およびその立体表現を可能にしている。日本科学技術情報センター (JICST) においても化学物質の立体構造の表現を可能にしている。また質量スペクトル等においても、スペクトルのピークの位置の図形表現を行っている。しかしまだ線表現の図形である。図形表現としては初歩的なものにすぎない。

文献データベースにおいては必要文献の存在のみを回答するにすぎない。そのオリジナル文献自体のオンライン伝送がいくつか試みられている。EC は DIANE と並んで1次文献のオンライン・システム ARTEMIS の開発を計画した。これは文章部分を文字コードで、図形部分をファクシミリを用いる複合システムとして提案されたがまだ実現していない。国際的出版社である Elsevier Science Publishing, Co. はオリジナルのコピーでの利用があまりにも簡便に行われ、本来出版物として購読されるべき利益が損われているとして、出版社みずから1次文献の伝送サービス ADONIS を計画した。これには Springer Verlag, Pergamon Press 等欧州の出版社および米国のいくつかの出版社が協力することとなっていた。これはすべての出版物を光ディスクに蓄積しコンピュータ操作で必要文献を検索しオンライン伝送しようとするものであるが、技術的障害が克服できず中断されている。

米国議会図書館においても資料の光ディスクによる蓄積・保管を計画しており、プロジェクトが進行中である。この光ディスクによる静止画像の蓄積は多方面から関心が示され、数多くの試みがある。しかし1次文献全体の蓄積・伝送には、蓄積媒体としてもまた通信回線（速度と容量）からみても、まだ少し問題が残されている。

最近、地図のデータベース化と、その地図に示される地点や地域に関する種々のデータの複合利用が種々検討されている。これはベクトル化された白地図のデータベースの各点、あるいはエリアに人名や国地名や建物名、住民や行政やマーケティングのデータ、地下埋没物に関するデータ等のデータベースがそれぞれリンクされ、ユーザからみて一つのデータベースのよ

うに利用しようとするものである。これはデータベース技術として新しい問題の提起であり、目下のところ、対症的にシステムが造られている。これらを統合したマルチ・メディア・データベースについてはまだ模索の状態である。

4.3 full-text database と電子出版

すでに述べたようにオンライン・サービスの前身には何かの形での先行サービスがある。文献や数値データの場合は印刷物媒体でのサービスである。この印刷物版下の作成工程が、従来の人間による植字、グラ出し、活版印刷の工程がキーボード入力、コンピュータ編集、オフセット用フィルムへの電子光学的出力に変化し、その編集用マスタ・ファイルがオンライン・サービス用にコンバートされる。

しかしこの工程をとるにしても印刷物媒体によるサービスは、印刷、製本、装訂、発送、郵送の期間（約1カ月程度）を必要とする。しかしオンライン媒体の場合、マスタ・ファイルのコンバージョンとオンライン・ファイルの更新が行われれば即座にサービス可能になる。

物理的なファイル構造がどうあれ、ユーザからみてキーワードその他の書誌的事項によって必要な文献を見つければ、即座に同じキーボードでその文献（全体または一部分）の伝送を要求しリアルタイムで入手できることを望んでいる。今後の文献データベース・サービスはこのような方向に進むと考えられる。Chemical Abstracts Service は、米国化学会の出版物についても full-text database を造り、オンライン・サービスのシステムを開発した。前述の図形・映像のデジタル処理が容易になれば百科事典や多くの出版物がオンライン伝送（おそらく資料全体でなく、ユーザの必要とする項目あるいはパラグラフ）でサービスされるようになると考えられる。

電子出版とは、目下のところ印刷物形態のサービスを持たず FD または磁気テープの形、あるいはオンラインによって提供される出版形態を指す。しかし現実には、印刷物形式と併行して（あるいは印刷物に付随して）FD かカセット・テープの形で印刷物と同内容で頒布されるケースが現れている。しかしソフトウェアで種々の処理を行うことによって実際の利用が可能になる原データを、ソフトウェアつきで売り出す形も次第に現れている。電子出版の将来の形は、従来の印刷物が果し得なかった利用を可能にする方向で成長するであろう。

4.4 分散データベースとネットワーク

VAN の誕生はデータベース・サービスの成立を支えた。しかし現状のデータベース・サービスの VAN 利用はまだ限られた形にすぎない。米国のデータベースサービスを行うオンライン・サービスは寡占の状態にあって各サービスは独立している。それぞれ VAN を自分の通信回線網として利用しているのみである。欧州の EURONET 上のオンライン・サービスの協同体である DIANE の場合でも、各オンライン・サービスが相互に独立であることは変りがないが、一つの規約の中で結ばれており、ユーザに対して独自のコマンドのみでなく共通コマンド言語でも利用できるような協力関係をもっている。しかしこの形では多くのオンライン・サービスが併立しているのであって分散データベース・システムとは言えない。

これに対し Chemical Abstracts Service (CAS) が主唱して現在実現の計画が進行中の Scientific & Technical Information Network International (STN International) は分散データベース・システムを国際規模で実現しようとするものである。STN は CAS の提供するソフトウェア・システムを使用して行うデータベース・サービスの連合体である。STN は、ソフトウェア供給者とデータベース供給者とオンライン・サービスとユーザからなる。ソフトウェア供給者は CAS である。データベース供給者は目下のところ米国は CAS、欧州は西独の Fachinformationszentrum Karlsruhe (FIZ-K, 旧称 INKA) と東洋は日本の JICST である。オンライン・サービスは上記 3 機関がそれぞれみずから行っているオンライン・サービスである。STN はこの構成員同士の二者間契約 (ソフトウェア供給者とオンライン・サービス, データベース供給者とオンライン・サービス, ユーザとオンライン・サービス) によって構成される。

同一ソフトウェアによってそれぞれ特色あるデータベースを持つ米国、西独、日本にあるサービスノードがオンライン・サービスを行いユーザに各ノードを通して他のノードの提供するデータベースを利用できるネットワークである。

これは CAS が、単なるデータベース供給者であるだけでは印刷物媒体の売上げが漸減化方向にある中で将来の経営戦略を立てるに当り、

○すべての印刷媒体のサービスをオンライン媒体によってサービス可能にする。(これはすべて CAS Online というサービスで実現。)

○データベース供給者もみずからオンライン・サービスを行うことによって、今まで他のオンライン・サービスに奪われている利益を回復する。

○化学のみでなく科学技術全般についてのサービスを連携して行うことによりサービスを強化する。(すなわち scope merit を生み出す。) という方向を打出した。

この中で問題は「同一ソフトウェア」という点である。FIZ-K はその方針を受け入れ、それまで稼働していたデータベース・ソフトウェアを CAS のものに替え、その持つすべてのデータベースを CAS ソフトウェアに合わせてコンバートした。1983 年に CAS より STN 参加を要請された JICST はこの条件を受け入れることを拒否し、CAS と交渉を続けてきた。1985 年の秋に、

○日本国内について現在 JICST の行っている JOIS のサービスはそのまま行う。(ただし現在の JOIS・II システムの JOIS・III へのレベルアップも行う。)

○JICST の作成する英文データベースを、STN 専用のコンピュータにより、CAS のソフトウェアを用いてサービスする。

ということで了解が成立し、目下その準備を行っている。

STN International の計画は広い範囲に刺戟を与えた。欧州のいくつかのデータベース・サービスは、STN に対抗するネットワークの形成について協議をした。しかし STN もまだ本格的稼働はしておらず、DIALOG, SDC, BRS 等の米国系オンライン・ヴェンダが圧倒的に市場をシェアしている現在、種々の試みがなされるべきである。

すでに述べたように、目下のところデータベース市場は大きい資本力をもつヴェンダの寡占状態が続いており小さいサービスが参入することはむずかしい。この場合、巨大な集積をもつ化学物質のデータベースや文献データベースのバック・アップを受けて、比較的小さいデータベース・サービスが相互に協力して付加価値と scope merit を生み出すことにより商業的に成り立つ状況を造り出す可能性がある。これを意図して、科学技術庁の科学技術振興調整委に基づいて「ネットワーク共用による化合物情報等の高次利用の研究」が、日本科学技術情報センター (JICST)、日本電信電話公社、をはじめとして約 10 機関の協力の下で 1982 年より進められている。このネットワークの規模や機

能は当初の構想よりは縮小されているが、このようなネットワークが公的な力によって建設され実用されることは情報産業の育成の上できわめて有意義である。たとえば文部省の学術情報システムのネットワークとJICSTを中心とする文献/ファクトのデータベース・サービス・ネットワークが相互接続して民間のデータベース・サービスを支援するような国家的構想もあってよいと思われる。

5. おわりに

人間の知的な活動の成果を保管し供給する仕事は、人間の文化の発展とともに始まり、常に最も新しい技術を利用し、常に次の社会への胎動をその中に噛みこんできた。そして現在も最も新しい情報技術を利用し多様なサービスを展開しつつある。社会が知的に成熟し、競争が研究開発やオフィス事務やシステム設計や生産工程の改革、効率化になっていくにつれて、データベース・サービスの要求はより強く、より高次で高機能なものになる。そしてこのデータベース・サービスと情報創造の活動が結んで造る情報生産のサイクルの効率が経済活動全体の発展に重要な役割を果たす。

データベース・サービスは社会全体が情報化していく中で新しい社会・行政・経済のインフラストラクチャの役割をもつ。そしてそれと同時に商業的に成立する部分から次第に商業化しつつある。すなわち、インフラストラクチャであると同時にそれ自身が商行為の対象となる財であるという二重性によって従来の商習慣や法制度の変革が必要となってきている。

データベース・サービスを論ずるに当たり、その技術的側面（これだけでも非常に多くの誤解があるが）のみならず、それを必要とし、その利用を通して変革を加速しつつある社会・経済の面の分析も重要である。そしてデータベース・サービス産業の成立の要件とその実現過程の中で、データベース・サービスのもつ上記の二重性をよく分析する必要がある。

参 考 文 献

(データベース技術)

- Amano, K. and Mochida, A.: A Supporting System for Effective Construction and Sharing of Scientific Databases by General Researcher. *Inf. Proc. Manage.*, Vol. 21, No. 6 (1985).
- Bäszczyk, R.: Information Systems with Attribute Decomposition, *Inf. Syst.*, Vol. 10, No. 4 (1985).
- Blair, D.C. and Maron, M.E.: An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System, *Commun. ACM*, Vol. 28, No. 3 (1985).
- Bhatt, R.V., Fisher, E.L. and Rasdorf, W.J.: Information Retrieval Architectures for Expert System/DBMS Communication, *Proc. Fall, Ind. Eng. Conf.* (1985).
- Boisvert, R.F. et al.: GAMS—a Framework for the Management of Scientific Software, *ACM Trans. Math. Software*, Vol. 11, No. 4 (1985).
- Chang, C.C.: Optimal Information Retrieval When Queries are not Random, *Inf. Sci.*, Vol. 34, No. 3 (1984).
- Croft, W.B. and Parentv, T.J.: A Comparison of a Network Structure and a Database System Used for Document Retrieval, *Inf. Syst.*, Vol. 10, No. 4 (1985).
- Desalvo, D.A. and Liebowitz, J.: The Application of an Expert System for Information Retrieval at the National Archives, 1985 Expert Syst. Gou. Symp. (1985).
- Fritz, F. and Granitzer, W.: Dater/Fakter Systeme anfair Basis Relationaler Daterbanksysteme, *Machr. Dok.*, Vol. 36, No. 4/5 (1985).
- Green, M.R.: Providing Online Access to Scientific Data, *Proc. Natl. Online Meet* (1985).
- Hake, J.F.: PROC—a Database on Parallel Processing Literature, *Comput. Archit. News*, Vol. 13, No. 4 (1985).
- Hampel, V.E.: TIS—the Intelligent Gateway Processor, *Proc. Hawaii Int. Cont. Syst. Sci.*, Vol. 18 (1985).
- Helbig, H.: Natural Language Access to the Database of the AIDOS/VS Information Retrieval System, *Artif. Intell. Inf. Control, Syst. Rob.* (1984).
- 洪 政国, 六川修一: カラー画像検索システムへのアプローチ, 情報処理学会, 全国大会講演論文集, Vol. 31, No. 2 (1985).
- Kowalski, A. and Strazalkowski, T.: A Natural Language Information Retrieval System with Extensions towards Fuzzy Reasoning, *Int. J. Man-Mach Stud.*, Vol. 23, No. 4 (1985).
- Mistrik, I. and Graz, G.K.: Das (IV + V) Information. on System Paket, *Nachr. Dok.*, Vol. 36, No. 4/5 (1985).
- Newton, U.A.: Intelligence in Database Management Systems, *Informatics*, Vol. 7 (1983).
- 二村祥二他: 2階層記憶による索引語転写索引, 情報処理学会, 全国大会講演論文集, Vol. 30, No. 2 (1985).
- 二村祥二他: 情報検索システム AIR の性能評価, 情報処理学会, 全国大会講演論文集, Vol. 29, No. 2 (1984).
- Ohtomo, M., et al.: Extension of Information Retrieval System on Multi-Files, 情報処理学会,

- 全国大会講演論文集, Vol. 31, No. 2 (1985).
- Ozkarahan, E. A. and Can, F.: *An Integrated Fact/Document Information System for Office Automation*, Inf. Technol., Vol. 3, No. 3 (1984).
- Page-Castell, J. A. and Hollisten, C.: *The Chemical Substances Information Network-User Service Office Evaluation and Feedback*, J. Chem. Inf. Comput. Sci., Vol. 25, No. 4 (1985).
- Salton, G.: *Some Characteristics of Future Information Systems*, ACM SIGIR Forum, Vol. 18, No. 2/4 (1985).
- Schek, H. J.: *Nested Transactions in a Combined IRS-DBMS Architecture*, Res. Dev. Inf. Retr. (1984).
- Shepherd, M. A. and Watters, C.: *A Common Interface for Accessing Document Retrieval Systems and DBMS Retrieval of Bibliographic Data*, Inf. Process. Manage., Vol. 21, No. 3 (1985).
- Tidwell, D. et al.: *Chemical Substances Information Network(CSW)-an Overview*, P.B. Report, No. PB-85-135812 (1984).
- 浦谷則好他: 静止画像検索システム FORKS, 情報処理学会, 全国大会講演論文集 Vol. 31, No. 2 (1985).
- Urban, S. D. et al.: *Utilizing an Executable Specification Language for an Information System*, IEEE Trans. Software Eng., Vol. 11, No. 7 (1985).
- Wald, J. A., and Sorenson, P. G.: *Resolving the Query Inference Problem Using Steiner Trees*, ACM Trans. Database Syst., Vol. 9, No. 3 (1984).
- (Full-Text Searching, Full-Text Database, 電子出版)
- Abbott, J. P. and Smith, C. R.: *Full-text and Bibliographic ACS Databases-Rivals or Companions*, Proc. Natl. Online Meet. (1985).
- Aitchison, T. M.: *Online and the Database Producer*, J. Inf. Sci., Vol. 9, No. 2 (1984).
- Benson, D. A. et al.: *A Microprocessor-Based System for the Delivery of Full-Text Encyclopedic Information*, Proc. ASIS Annu. Meet., Vol. 18 (1981).
- Brunelle, B. S.: *The Production of a full-text Database*, Proc. Natl. Online Meet. (1985).
- Buttler, B. and Slade, R.: *Economics of Full-Text Information*, Proc. Natl. Online Meet. (1985).
- Cannon, C.: *EMIS-The Electronic Publication of Material Data*, Inf. Serv. Use, Vol. 5, No. 2 (1985).
- Callen, U. F. and Flagle, C. D.: *Full-Text Medical Literature Retrieval by Computer-a Pilot Test*, J. Am. Med. Assoc., Vol. 254, No. 13 (1985).
- Goebel, J. W.: *Zur Kurztext-Langtext Problematik Bei Juristischen Informations Systemen*, Nachr., Pok. Vol. 36, No. 2 (1985).
- Goodstein, D. H.: *Electronic Publishing*, TAGA Proc. (1982).
- Love, R. A. and Garson, L. R.: *Precision in Searching the full-Text Database-ACS Journal Online*, Proc. Natl. Online Meet (1985).
- Niehoff, R. T. and Kutz, M.: *Electronic Publishing from the Perspective of a Large Engineering Society*, Proc. ASIS Annu. Meet., Vol. 19 (1982).
- Norris, C.: *Full-text Searching-or You Find the Needle in the Bayotack*, Online 85 Conf. Proc. (1985).
- 白井英子: オンラインと冊子体の関係およびその将来, 薬学図書館 Vol. 30, No. 3 (1985).
- Spencer, M. and Rothenberg, D.: *Building an Online Full-Text Files*, Proc. Natl. Online Meet. (1985).
- Tenopir C.: *Contributions of Value Added Fields and Full-Text Searching in Full-Text Databases*, Proc. Natl. Online Meet. (1985).
- Tenopir, C.: *Full-Text Database Retrieval Performance*, Online Rev., Vol. 9, No. 2 (1985).
- Tittlbach, G.: *Electronic Publishing and Document Delivery of German Patent Information*, J. Chem. Inf. Comput. Sci., Vol. 26, No. 1 (1986).
- (downloading)
- Broering, N. C.: *The Mini MEDLINE SYSTEM-TM-a Library-Based End-User Search System*, Bull. Med. Libr. Assoc., Vol. 73, No. 2 (1985).
- Cotter, G. A. and Hartt, R. W.: *Integrated Bibliographic Information System-Integrating Resources by Integrating Information Technologies*, Proc. Natl. Online Meet. (1985).
- Feeney, N. J., Miller, R.: *Downloading-Piracy or Panacea?*, J. Inf. Sci., Vol. 8, No. 1 (1984).
- Gray, R.: *Databases, Downloading and Proprietary Rights*, Program, Vol. 19, No. 4 (1985).
- Lardy, J. P.: *Connecter un Micro-Ordinateur-un Serveur, Documentaliste*, Vol. 23, No. 2 (1986).
- Maxwell, R. J. and Nardi, B. J.: *Downloading-a Solution for Providing Technical Information in a Corporation*, Proc. ASIS Annu. Meet., Vol. 22 (1985).
- Mortensen, E.: *Downloading-Potentials and Restrictions in Online Searching*, Proc. ASIS, Annu. Meet., Vol. 21 (1984).
- Osborne, L. N.: *Downloading Overviews*, J. Libr. Adm., Vol. 6, No. 2 (1985).
- Ralph, R. D.: *Automating Data Capture from CAS Online*, Proc. Natl. Online Meet (1985).
- Ralph, R. D.: *Down Loading and Data Conversion Doing it Yourself*, Online 85 Conf. Proc. (1985).
- Wanger, J.: *Downloading Data for Use in a Local Database*, Proc. Natl. Online Meet (1985).

(昭和61年8月19日受付)