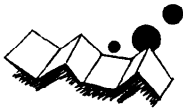


解説



オンライン情報サービスの歴史†

山本 毅 雄††

1. はじめに

情報処理に関する他の多くの概念と同様，“オンライン情報サービス”という概念も定義がむずかしいが、ここでは、

- a. データベースがオンライン記憶にある
- b. オンライン端末からアクセスできる
- c. データベースは継続的に、あるいは定期的に更新されている
- d. 利用は単一機関（一施設、一企業）内ではなく、機関外部からも行われる。

というものをさすでしょう。時に“データベースサービス”といわれるものと、ほぼ同義である。データベースとしては、文献（論文、図書、特許等）、非文献（数値・統計が主、市況情報、経済・産業情報、信用情報、物性・化合物情報等）を含む。主として検索が主目的であるもの（文献情報のほとんど全部と、非文献情報の大部分）と、検索よりもその後の分析処理に重点のあるもの（経済・産業情報や物性情報の一部など）がある。

1948年から稼動した、きわめて初期の計算機のひとつであるマンチェスター大学の計算機は、主メモリである Williams 蓄積管 (5,120 ビット) の他に、120 K ビットの磁気ドラムを備えていた¹⁾。これは、それまで計算処理中心であった電子計算機に、大量データ処理の機能をはじめて可能にしたものといえよう。この後、1950 年中ごろまでに、オンライン記憶としては磁気ドラムが多種類作られたが、 10^6 ビット程度であり、情報検索とくに文献情報など大量の情報検索にはさまざまな工夫をこらさねばならなかった。

最初の磁気ディスク RAMAC は、1956 年 IBM 社によって作られたが、磁気ディスクが普及してきたのは欧米では 1960 年ごろ、わが国では更におくれて 1960 年代末といえよう。そして、オンライン情報サ

ービスへの本格的な動きは、磁気ディスクの普及とともに始まったといってもよい。しかし、ここでは、磁気ディスク以前における情報検索システムの歩みを概観してから、上の a~d の項目に対応したオンライン情報サービスを支える四つの条件、すなわち、

- オンライン記憶（媒体、装置）の発展
- データ通信、ネットワークの発展
- ソフトウェアの発展
- データベース作成の発展

を追い、最後に欧米とわが国それぞれでのオンライン情報サービスそのものの発展を概観する。

2. 磁気ディスク以前

オンライン記憶が未発達であっただけでなく、計算機自体が遅く、高価であったこの時期には、種々の方向への試みがなされた。

2.1 専用検索装置

1951 年、MIT の Phillip R. Bagley は、修士論文で当時の汎用高速計算機 Whirlwind I をモデルとして、5 千万文庫を検索させたらどうなるかをしらべてみた。各文献から 30 語平均の索引語が出ていると、検索には 41,700 時間かかることになり、これでは役に立たない。そこで Bagley は、検索専用機を作るとはかないとの結論に達した²⁾。

これは極端な例であるが、多くの人が同様の結論に達して、種々の装置を作った。パンチカード、紙テープ、磁気テープなど、通常はオフライン記憶媒体であるものを、専用の回路をもつ装置で高速に検索するのが、その一法である。

1957 年発表された Western Reserve 大学の紙テープ検索機³⁾、1958 年発表の IBM 9900 Special Index Analyzer (パンチカードを入力とし、紙テープを中間メモリおよび出力とする)³⁾、その他 GE-250 Information Searching Selector (磁気テープ専用検索機)などが有名であった。わが国では、電気試験所の磁気テープ式情報検索機 (1960 年)⁴⁾、日本科学技術情報センターの JEIPAC 電子式情報処理機 (1961 年)⁵⁾な

† The History of On-Line Information Service by Takeo YAMAMOTO (Computer Centre, The Univ. of Tokyo).

†† 東京大学大型計算機センター

どが作られた。

マイクロ写真によるイメージ情報の検索を、電子計算機によるデジタル情報の検索と組合せ、あるいは単独で用いるのが、第二の方向である。マイクロ写真自動検索装置のアイデアは、1940年始めにMITのV. Bushによって呈示され、これにもとづいた装置Rapid Selectorが1949年Engineering Research Associatesによって製作された。のちこれはYale大学で改良され、各所で使われた。その他、FMA File Searchシステム(1961年)、アパーチャカード(1954年、CIAで使用)、Minicard、IBMのWalnut、その他、多数のシステムが作られた⁶⁾。

これらのシステムは、一部はそのままオフラインの検索機器として特殊用途に使われたり、図書館・情報室等で実用にされてきた例もある。しかし、一般的には、1965年ごろを境として、磁気ディスクによるオンラインデジタル記憶が安価になってきたことと、データ伝送の発達により、同じ磁気ディスク上の記憶を多くの利用者が共用して単位コストを下げることが可能になったこと、などから、磁気ディスク方式に移ったところが多い。

2.2 情報内容の解析

計算機の記憶容量が絶対的に不足すると、複雑な概念を簡単なコードで表現したいと考える。また、記憶された情報内容と、この内容を知らない利用者の検索質問とのマッチングという、本質的に困難な仕事の手がかりとして、そこにあらわれるすべての概念を定義し、分類して、これをあいまいさのないコードで表現したいと考える。これは、いずれも自然な欲求であり、やはり計算機以前から、図書館学などで古くから追求されてきた一つの方向である。Deweyの十進分類法は1876年に発表されており、これを改良したUDC(統一十進分類法)は1895年の国際会議で採択されている。

計算機による情報処理が盛んになるにつれて、このいわゆる主題分類に関する研究も盛んになった。多種類の分類辞書、いわゆるシソーラスが作られたし、いまでも作られつつある。分類も、単なる階層関係だけでは不十分で、概念の間の“関係”あるいは“役割”が大切であるとの考えから、Ranganathanの“コロソ分類”(1950年)やKentらの“Role Indicator”の考え⁷⁾(1956年)が展開された。しかし、対象分野の変化、対象データ量の増大とともに、柔軟性と普遍妥当性あるいは正確さを兼ね備えた分類を維持すること

は難しく、結局いわゆる“シソーラスによる語彙規制”はデータの作成・検索に補助的な役割を果たすだけであることが明らかになってきた。

関連する問題であるがいわばアプローチの異なるものに、自然語の意味理解がある。これにも、計量言語学的・統計的なアプローチと、構文解析・命題論理演算などを主とする人工知能的アプローチがあり、特に1960年代初めまでの幸福な期待にみちた時代に、計算機による文献情報処理と自動翻訳・質問応答などの人工知能への試みは、互に影響を与えながら進んできた。統計的アプローチに大きな寄与をしたのはLuhnである⁸⁾。彼は、SDI(Selective Dissemination of Information)システム、KWIC(Keyword-in-Context)索引、自動抄録(語の頻度分布にもとづく)などの創始者であった。人工知能研究の歩みとそれに対するBar-Hillelのきびしい批判⁹⁾はよく知られているが、Bar-Hillelは同じころ情報検索の研究の進め方にも批判的な意見を発表している¹⁰⁾のは面白い。

3. オンライン記憶の発展

最初の磁気ディスクとして有名なIBM社のRAMACは、直径約60センチ(24インチ)の円板50枚を組合せた大きな装置で、それにもかかわらず記憶容量は30Mビット(5M字)にすぎなかった。また、読書用ヘッドも1対しかなく、したがってアクセス速度も0.5秒程度であった。しかし、これが製作された1956年ごろは、ドラムでも1Mビット程度までであり、これに比べればやはり飛躍的な増大である。それまでの検索では、外部記憶上のデータを高度な、または高速のアルゴリズムで順に処理するか、あるいは主記憶上の極度に圧縮しコード化した情報に複雑な処理を加えて、わずかの情報からどれだけnon-trivialでない結果をひき出すかという、いわばアルゴリズム主導の考え方が主になっていた場合が多い。磁気ディスクの出現は、大量のデータを比較的圧縮・変換しないでオンライン記憶におき、このデータの記憶のしかた(索引の作り方)アクセスのしかた(質問のくみたて方)あるいは更新のしかたによって、少量のデータから予想されるよりはるかに豊かな情報を得るという、データ主導の新しい考え方にうつるきっかけを与えたものといえよう。

その後、磁気ディスク技術のめざましい発展はよく知られている通りで、例えば記録密度にしても、約25年に 10^8 倍となっている。それに対応してビット当り

コストも低下しており、代替技術（磁気バブル、光メモリ等）をおさえて、オンライン記憶の主座をおさえてきた。

ビデオテープ型の幅広い磁気テープを用いるマストレージシステム (MSS) は、Ampex 社の Terabit Memory (1972 年) が早い、1975 年 IBM 3850 の発表以来、急速に普及している。しかし、これまでのところ、大規模なオンライン情報サービスに MSS が活用されている例はあまりない。これは、MSS がオンライン媒体としてはアクセス時間が長すぎる、通常のオンライン情報システムではファイルアクセスが局在化されておらず、しかも多数使用者の同時使用により MSS の弱点が強調されやすいためと思われる。今後、バッチ出力作業に対し古典的なバッチ検索技術を再生させるなど、MSS の有効な使い方が見出されるのではあるまいか。

4. データ通信ネットワークの発展

1973 年の Telenet, 1976 年の TYMNET の発足が米国におけるオンライン情報サービスの普及にきわめて大きい役割を果たしたことは、よく知られている。これらのデータ通信ネットワークは、米国内主要地間の専用回線ネットワークを多数の利用者とホスト・コンピュータが共用することにより、地域差のない安価なデータベース・アクセスを可能にし、オンライン情報サービスの発達を促した。しかし、これよりはるか以前から、データ通信の着実な利用がなされていたことに注意が必要である。

オンライン・リアルタイム・システムとしては、早く 1958 年に運転を開始した米空軍の SAGE 半自動防空システム¹¹⁾、座席予約業務のみならず予約情報・飛行情報のサービスも行い、経営情報システムとしての性格もある American Airline 社の SABRE (1963 年稼働) 等、国土の地理的条件もあるが、データ通信機能とデータ記憶 (データベース) 機能を総合した大規模システムであった。

1960 年代前半から盛んになった汎用 TSS の開発においても、安価な端末の存在や、早くから音響カップラーを用いて通常の電話回線による遠隔アクセスが行われていたことなどから、手軽で実用的なデータベース・アクセスへの素地が米国にあったといえる。

わが国では、1960 年の MARS-1, 1963 年の MARS-101¹²⁾にはじまる国鉄の座席予約システム、1965 年ごろからの各銀行のハンキング・システム¹³⁾など、業務

オンライン・システムにおいては世界をリードするシステムが作られた。このような活動をオンライン情報サービスに拡張したものとして、市況情報センターの QUICK システム (1974 年)、農林省の生鮮食糧品流通情報サービス (1976 年) など、活発に使われているサービスがある。これらは大部分が専用回線と専用端末を用いるオンライン・リアルタイム型のシステムであり、公衆回線の利用による安価・大衆的なサービスあるいは汎用 TSS を用いた柔軟・多目的なシステムという面では、わが国は未だ欧米におよばない点が多いようであるが、今後の発展が期待される。

5. ソフトウェアの発展

オンライン情報サービスをソフトウェア的にみればほぼ 3 種に分類できよう。

a. TSS またはオンライン・リアルタイムで、文献情報等に対し比較的複雑な問合せをするシステム

b. 主としてオンライン・リアルタイムで、ごく単純な問合せをきわめて多数受け付けるシステム

c. モデル計算、グラフ作成、シミュレーション等データ内容に演算処理を加えて出力することに重点をおくシステム

a は文献情報システム、化合物情報システム等が主であって、大量の変長データ、不定回線返しデータ等の管理や、転置索引リスト長の質の悪い分布に対しいかにパフォーマンスを維持するかが問題である。また、質問者とシステムとのマン・マシン・インタフェースが重要な問題となる。b では、主としてトラフィック処理と、もしデータベースの同時更新を行う場合は、データの保全が重要となる。c では、分析用のすぐれたプログラムを作ることと、多様な機能をいかに利用者に指定させるかの工夫が大切となる。

文献検索型のシステム^{14), 15)}は、Lockheed 社の DIALOG と SDC 社の ORBIT がそれぞれ 1965 年前後に発足した。その他にも多くのシステムがあるが、原理的にはいずれも転置ファイル検索方式であり、上の a. にあげたような問題をいかに解決するかで異なっているといえる。問合せシステム型は、いわゆるオンライン・リアルタイムシステムの問題で、システムの規模とトラフィックによるが、最近ではいわゆる DBMS を用いる場合もある。演算処理型 (プログラム分析型) では、シミュレーション、統計、推測、図形出力などのプログラム機能が最重要となる。この分野では、計量経済モデルが社会的に認められ、広範に実

用されるようになったこと(1970年代)が特筆される。石油ショックは計量経済モデルの限界を明らかにしたが、またこれに関する知識(どれだけのことが可能であり、現実にどれだけ使われているか)を普及させた。

6. データベース作成の発展

米国におけるデータベースの作成¹⁶⁾、特に文献データベースの作成は、まず出版物の電算機による作成の副産物として始まった場合が多い。特に抄録・索引誌出版の電算化の際、材料となるデータの一部をデータベースとして販売することが広く行われるようになったのは、1960年代末からである。1970年代にデータベースの数は増加し、現在ますます増加しつつあるが、一方では世界的なオンライン情報サービス網の存在により、同一分野でのデータベースの整理が自然に発生し、少なくとも重なりのある分野での新規作成は困難となりつつある。

欧州においては、特に1970年代後半、米国のオンライン情報サービスのヨーロッパへの進出に対抗するため、データベースの作成に力を入れている。このため、データベースの数・カバーする分野等は急速に増加しつつある¹⁷⁾。データベースを作成して、これが世界的にその分野を制するまでには、長期にわたる実績と継続的な努力が必要であるが、欧州には既に世界的なデータ集・文献集等で伝統あるものがあり、この点で有利であろう。

7. オンライン情報サービスの発展

7.1 米 国^{15),16)}

文献検索においては、Lockheed社のDIALOGとSDC社のORBIT、およびそれらの応用であるNASA/RECONとNLMのTOXLINEがそれぞれ重要な時期を画した。DIALOG、ORBITに対抗するものとしてBRSがあるが、まだその勢力は小さいようである。

数値検索の面では、DIALOG、ORBIT等に含まれる経済統計の類もさることながら、TRW社の信用情報、PREDICASTS社やChase Economics社の経済予測等の、企業活動に密着した大規模なサービスがある。

7.2 欧 州^{15),17)}

NASA/RECONを輸入したESAのサービスが欧州でのオンライン情報サービスの始まりであったといえる。最近では、英国のBLAISEその他多くのサー

ビスがあり、EURONETに統合されようとしている。PRESTELその他画像データのサービスの今後は興味深い。

7.3 日 本¹⁹⁾

市況情報センターのQUICK(1974年)は大きな成功をおさめた。文献では、利用者が学術研究者に限られるが、東大のTOOL-IR(1975年)が公開外部利用の最初と思われる。JICSTのJOIS(1976年)、JAPATICのPATOLIS(1979年)等は本格的に利用されてきており、今後の発展が期待される。その他、筑波大学(IDEAS/77)、広島大学(HUNDRED)、京都大学、東北大学、名古屋大学等の各大学、アジア経済研究所等のシステムがある。特に、日本語端末の普及により、今後国産データベースの蓄積と利用は大きく進むであろう。

経済・統計関係では、日経NEEDS、電通MARK-IIIの各種データベース(NRI経済情報、TSR企業財務データ等)が主なものである。この分野も、発展が期待される。

画像情報のCAPTAINシステムは、巨大な実験であり、興味深いシステムといえる。

8. おわりに

手持ちの不十分な資料にもとづき、簡単にまとめたため、十分な紹介ができなかったことをおわびしたい。種々御教示頂いたJICSTの中井浩氏に深謝する。

参 考 文 献

- 1) Matick, R. E.: Computer Storage Systems and Technology, pp. 12-13, John Wiley and Sons, New York (1977).
- 2) Becker, J. and Hayes, R. M.: Information Storage and Retrieval: Tools, Elements, Theories, p. 146, John Wiley and Sons, New York (1963).
- 3) 同上書, pp. 147-154.
- 4) 木澤 誠, 実川卓次, 大岸 洋, 田島智平, 木村 啓: 磁気テープを用いた情報検索機, 電気通信学会誌Vol. 44, No. 11, pp. 1788-1795 (1961).
- 5) 日本科学技術センター十年史, pp. 235-240, 日本科学技術情報センター, 東京 (1967).
- 6) Becker, J. and Hayes, R. M.: 上掲書, pp. 195-217.
- 7) Perry, J. W., Kent, A. and Berry, M. M.: Machine Literature Searching, Interscience, New York (1956).
- 8) Schultz, C. K. 編: Luhn, H. P.: Pioneer of Information Science, Spartan Books, New York

- (1968).
- 9) Bar-Hillel, Y.: Report on the State of Machine Translation in the United States and Great Britain, Report on NONR-2758 (00) (1959).
 - 10) Bar-Hillel, Y.: A Logician's Reaction to Recent Theorizing on Information Search Systems, American Documentation, 8, pp. 103-113 (1957).
 - 11) Everett, R. R., Zraket, C. A. and Benington, H. D.: SAGE-A Data-Processing System for Air Defence, Proc. EJCC, pp. 148-155 (1957).
 - 12) 穂坂 衛, 大須賀節雄: 実時間システム開発, 共立出版, 東京 (1970).
 - 13) 石崎純夫, 藤田 献: トータル・バンキング・システム, 産業図書, 東京 (1972).
 - 14) Walker, D. E. Ed.: Interactive Bibliographic Search: The User/Computer Interface, AFIPS Press, Montvale, N. J. (1971).
 - 15) Hall, J. L.: On-line Information Retrieval Sourcebook, Aslib, London (1977).
 - 16) Williams, M. E. and Rouse, S. H.: Computer-Readable Bibliographic Data Bases, A Directory and Data Sourcebook, ASIS, Washington, D. C. (1976-).
 - 17) EUSIDIC Database Guide, Eusidic, London (1978-).
 - 18) Walker, D. E. 編: Interactive Bibliographic Search: The User/Computer Interface, AFIPS Press, Montvale, N. J. (1971).
 - 19) シンポジウム「オンライン情報サービス」予稿集, 情報処理学会 (1979).

(昭和55年1月11日受付)