

利用の側面からみた国産データベースの特徴
Some Characteristics of Japanese Databases based
on User Survey

細 野 公 男
Kimio Hosono

藤 岡 千 代 子
Chiyoko Fujioka

Résumé

Japanese databases are an indispensable ingredient of information resources. Their importance increases rapidly since they will extensively cover information produced in Japan although the number of the databases is not sufficient enough to do any research without foreign ones.

This paper describes some characteristics of the Japanese databases based on an user survey. 63 academic libraries and 65 corporate libraries are questioned in terms of popular ones merits problems of the databases, government owned information which might be desirable to be included in the databases, and so forth.

Popular databases in academic libraries are only ones produced by JICST. On the other hand, corporate libraries are using wide range of databases such as ones provided by JICST, Japan Patent Information Center, Nihon Keizai Shinbun, and others. Corporate libraries use those more frequently than academic libraries.

Merits to use them are evident since searchers are exempted from the language barriers and able to obtain domestic information efficiently. Diversity of search commands and keyword types among different systems are major problems. Incomplete or biased coverage of information stored in the databases is also indicated.

This paper analyzes reasons and/or backgrounds which resulted in the user responses and finally suggests the way to foster sound growth of the databases.

- I. はじめに
- II. 国産データベースの利用状況
 - A. 概況
 - B. 利用頻度

細野公男：慶應義塾大学文学部 図書館・情報学科教授，東京都港区三田 2-15-45
Kimio Hosono, Professor, School of Library and Information Science, Keio Univ. 2-15-45, Mita, Minato-ku, Tokyo

藤岡千代子：慶應義塾大学文学部 図書館・情報学科，東京都港区三田 2-15-45
Chiyoko Fujioka, School of Library and Information Science, Keio Univ. 2-15-45, Mita, Minato-ku, Tokyo

利用の側面からみた国産データベースの特徴

- III. 国産データベースの利点と問題点
 - A. 利点
 - B. システム利用にあたっての満足度
 - C. 問題点
- IV. 国産データベースの展望
 - A. 需要の高いデータベース
 - B. データベース・サービスの発展
 - C. 国産データベースに対する要望
- V. 国産データベース振興の方向
 - A. 構築面
 - B. 流通・利用面
- VI. おわりに

I. はじめに

わが国におけるオンライン・データベース・サービスは、1976年 JOIS-I サービスの開始によって本格的に始まったといえよう。その後、DIALOG や ORBIT の様な海外のデータベース・サービスが導入され、データベースの利用は隆盛期を迎えたが、国内で生産される情報を対象としたデータベースが脚光を浴びだしたのは、1980年代に入って漢字ワードプロセッサの技術が著しく進歩してからである。

わが国で流通するデータベースには、海外の機関により作成され、日本のディストリビューターによって提供されているものと、日本国内の機関により作成され提供されているものがある。一般に、前者は海外データベース、後者は国産データベースとよばれているが、このうち国産データベースの概念は、以下のように定義することができよう。

国内の文章情報と数値情報、および海外の情報を国内で整理統合し、コンピュータによって処理、検索が可能な形態でまとまった情報群として編集した情報ファイル¹⁾

情報化社会の進展と情報技術の進歩により、最近ではデータベース・サービスに対する関心が高まり、需要も著しく増大し、データベース産業の環境は大きく変化した。また、データベースの有効性の認識も高まりつつある。しかし、依然としてデータベース・サービスは、欧米と比べると10年遅れているといわれている。例えば、国内で流通しているデータベースの約80%は海外製であり²⁾、国産データベースの数はまだ少い。

わが国で流通しているデータベースの数を『データベース台帳総覧』でみると第1表のようになる。第1表は、商用レベルで流通しているデータベースのみを取り上げており、利用可能なデータベースを全て網羅しているわけではないが、大体の傾向はつかむことができよう。そ

第1表 わが国で流通するデータベース数

	収録データベース数				データベース実数			
	57年度	58年度	59年度	60年度	57年度	58年度	59年度	60年度
海外データベース	472 (78.1%)	743 (81.1%)	1,034 (83.3%)	1,401 (82.3%)	334 (73.2%)	522 (76.9%)	725 (78.5%)	1,008 (78.2%)
国産データベース	132 (21.9%)	173 (19.9%)	208 (16.7%)	301 (17.7%)	122 (26.8%)	157 (23.1%)	199 (21.5%)	281 (21.8%)
合計	604	916	1,242	1,702	456	679	924	1,289

注：昭和60年度版データベース台帳総覧パンフレットから引用

れによると、データベース実数は57年度から60年度までの3年間で、456件から1289件へと3倍近く増大しており、データベース業の順調な発展がうかがわれる。

このうち国産データベースに関しては、122件から281件へと増加はしているが、全体の中で占める割合は26%台から21%台へと低下しており、海外依存度が強い傾向が明らかである。これは、日本では最近まで情報に対する意識が希薄であったこと、日本語という言語の障害があったこと、さらに大学や研究機関と民間のデータベース作成機関との連携が薄いことなどに起因するが、政府のデータベースに対する姿勢も大きな原因となっていたといえよう。

国産データベースの構築・普及は、科学技術の着実な発展、および各種の情報活動の円滑・効果的な推進に不可欠である。また、情報の一方的な輸入によって生じつつある、米国を中心とする先進諸国との情報摩擦を回避するためにも極めて重要である。したがって、国産データベースを発展させその普及をはかることは、急務である。そのためには、国産データベースの利用状況を把握し、今後の国産データベース・サービスの動向を分析することが必要となる。

この点に鑑みて、本稿では、国産データベースの利用の現状および構築に関しての要望を利用者側の観点から明らかにし、今後の国産データベースの在り方、振興の方向をさぐることを試みる。

II. 国産データベースの利用状況

A. 概況

わが国においても情報提供機関を中心に、データベース・サービスの利用に関してさまざまな調査が行われている^{3,4,5)}。しかし、これらの調査はデータベース一般を対象にするか、あるいは限られた種類の国産データベースを対象にしたものであって、まだ、国産データベース

の包括的な調査は行われていない。

そのため、昭和60年8月に大学図書館(77機関)、国立機関の研究機関など(19機関)、企業(84機関)の合計180機関を対象としてアンケート調査を行い、このうち71%の128機関から回答を得た。

国産データベースは第2表のように全体の89%にあたる114機関が利用している。残りの14機関は海外データベースのみを利用しているか、どちらも利用していない機関である。国産データベースを利用していない機関の大部分は、大学図書館である。利用しない理由として、“利用したいデータベースがない”、“端末を持っていない”と答えた機関が多い。その他、“エンド・ユーザからの積極的な要望がない”、“蓄積データの遡及年数が短い”などがあげられている。なお、このような機関でも、国産データベースを利用する具体的な計画を持っていたり、今後利用する希望を示しており、国産データベースに対する関心は強い。

B. 利用頻度

国産データベースの利用頻度を、“非常によく利用する”から“あまり利用しない”までにランク付けして、企業グループ、大学図書館・研究所等のグループ毎に示したのが第3表、第4表である。なお、企業では、JAPICDOC, TEST DRUG, ID-IR 赤外線吸収スペクトル解析, NEEDS-ECONOMYなども利用されている。

表から明らかなように、JICST 科学技術文献ファイルは大学図書館・研究所、企業のいずれのグループでも極めて利用頻度が高く、代表的な国産データベースであることがよくわかる。大学図書館グループでは、次いでJICST 国内医学文献ファイルの利用が続き、その他のファイルはほとんど利用されていないようである。一方企業では、NIKKEI, 日本特許ファイル, 日本実用新案ファイル, JICST 国内医学文献ファイル, JICST 科学技

第2表 国産データベースの利用状況

	どちらも利用している	国産データベースのみ利用している	海外データベースのみ利用している	どちらも利用していない	合計
大学図書館	39	3	2	7	51
研究所など	8	1	2	1	12
企業	57	6	2	0	65
合計	104	10	6	8	128

注：単位は機関数

利用の側面からみた国産データベースの特徴

第3表 企業における国産データベースの利用頻度

データベース名	利用頻度			合計
	非常によく利用する	よく利用する	あまり利用しない	
JICST 科学技術文献ファイル	34	20	4	58
JICST 科学技術研究情報ファイル	2	5	19	26
JICST 科学技術用語シソーラス・ファイル	6	4	6	16
JICST 資料所蔵目録ファイル	5	4	9	18
JICST 国内医学文献ファイル	11	14	4	29
NK-MEDIA	0	5	13	18
日本特許ファイル	27	7	2	36
日本意匠ファイル	2	7	9	18
日本実用新案ファイル	15	10	4	29
日本商標ファイル	3	6	5	14
INPADOC ファイル	6	5	8	19
BRANDY	2	7	3	12
微生物関連特許データベース	1	1	2	4
TSR-BIGS	4	6	4	14
テクノサーチ	3	7	8	18
タイトルサーチ	0	4	11	15
JAPAN/MARC	0	3	18	21
JOINT	3	9	13	25
NIKKEI	10	20	9	39
IEE	0	0	9	9

注1 単位は機関数

注2 複数回答

第4表 大学図書館・研究所等における国産データベースの利用頻度

データベース名	利用頻度			合計
	非常によく利用する	よく利用する	あまり利用しない	
JICST 科学技術文献ファイル	5	11	19	35
JICST 科学技術研究情報ファイル	0	2	3	5
JICST 科学技術用語シソーラス・ファイル	2	2	1	5
JICST 資料所蔵目録ファイル	0	1	2	3
JICST 国内医学文献ファイル	8	19	7	34
日本特許ファイル	0	2	1	3
日本意匠ファイル	0	1	0	1
日本実用新案ファイル	0	1	0	1
テクノサーチ	0	0	1	1
JAPAN/MARC	2	1	2	5
JOINT	0	1	0	1
NIKKEI	0	0	1	1

注1 単位は機関数

注2 複数回答

術研究情報ファイル, JOINT などのデータベースの利用度が高い。また, 全体的にみて大学図書館や研究所よりも企業の方が頻繁にデータベースを利用している。

第4表から企業ではデータベースの利用において, 中心的なファイルと補助的なファイルとがあることがわかる。例えば, JICST 科学技術文献ファイル, 日本特許ファイル, 日本実用新案ファイルは, “よく利用する”, “あまり利用しない” に比較して, “非常によく利用する” 機関が多い。したがって, これらのファイルは, 多くの企業にとって中心的なデータベースであるといえよう。それに対して, JICST 科学技術研究情報ファイル, JOINT, JAPAN/MARC, NK-MEDIA, 日本意匠ファイル, テクノサーチは, “あまり利用しない” が相対的に多く, 補助的に使用される傾向があると思われる。NIKKEI, JICST 国内医学文献ファイル, INPADOC ファイル, TSR-BIGS などは, この中間に位置するデータベースということになる。このうち JICST 国内医学文献ファイルが中心的なデータベースとなっていないのは, 外国データベースと併用されることが多いためではないかと予想される。

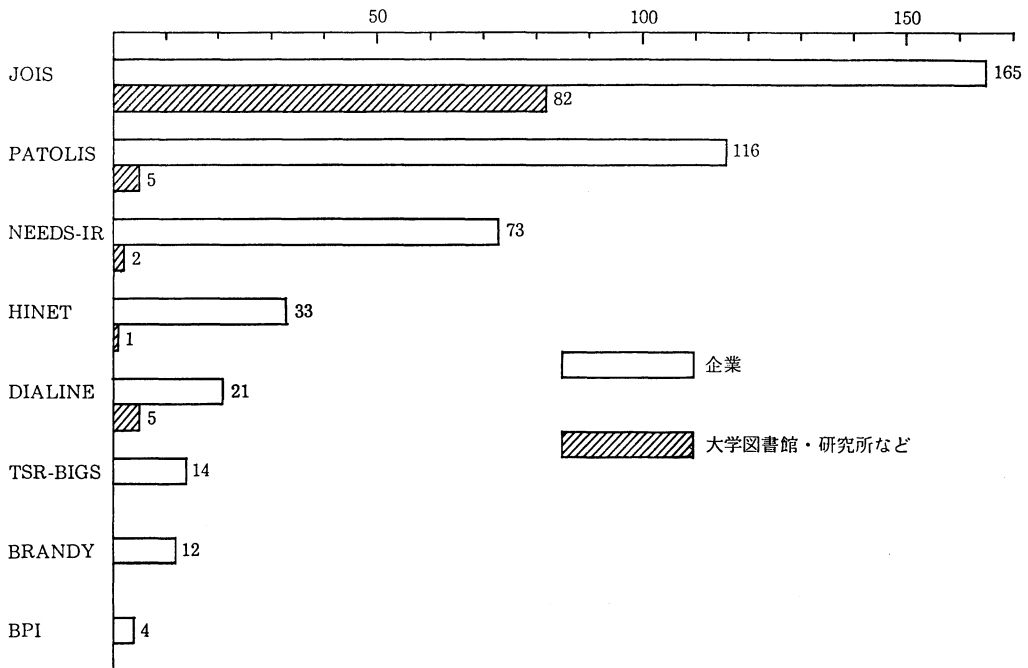
今回の調査では外国データベースとの利用度比較は行っていないこと, また, 利用頻度のレベル分けには個人差が大ききことなどの問題があるにせよ, 第3表, 第4表は国産データベースの利用の全体像を示していると考えられる。

第3表, 第4表をオンライン・サービス毎にまとめたのが, 第1図である。企業では JOIS, NEEDS-IR, PATOLIS が他のシステムと比べて利用頻度が高いことがわかる。一方大学図書館・研究所では実質上 JOIS だけが利用されているといえよう。

III. 国産データベースの利点と問題点

A. 利点

国産データベースを利用する利点を示したのが第5表である。この表から, “検索結果が日本語なので読みやすい”, “キーワードが日本語なので利用しやすい” ことがそれぞれ1位, 2位を占めているが, これは, 国産データベースの第一の利点が日本語で利用できることを示している。第5表には示されていないが, その他として, “国内情報のきめ細かい検索, 偏りのない



第1図 利用頻度の高いシステム

注1 単位は機関数

注2 複数回答

利用の側面からみた国産データベースの特徴

第5表 国産データベースの利点

	大 学 図 書 館 研 究 所	企 業	合 計
検索結果が日本語なので読みやすい	23	46	69
キーワードが日本語なので利用しやすい	16	34	50
海外データベースよりも利用料金が安い	19	28	47
原報の入手が保証されている	18	23	41
資料として検索結果がそのまま使える	5	23	28
地域ごとの細かい情報が得られる	4	5	9
海外データベースよりもタイム・ラグが少ない	1	1	2
その他	10	13	23

注 1 回答機関数は、大学図書館・研究所46、企業61、合計 107 機関

注 2 複数回答

検索ができる”，“国内情報の網羅性，カバー率が高い”，“日本語文献が入手できる”など，国内情報の利用しやすさと直接結びつく利点があげられている。

日本語データの処理は，データベース構築における労力を増大させる大きな要因である。したがって，国産データベースの効率的な作成をはかるためには，日本語処理技術の一層の進歩・発展と作成者側でのこれまで以上の努力が不可欠である。

B. システム利用にあたっての満足度

国産データベースを提供するオンライン・サービスの利用方法，データの内容，サービス等どの程度満足しているのかを調べてみると，以下のような傾向がみられた。

大学図書館や研究所では，シソーラスの使いやすさ・更新頻度，蓄積データ量・遡及年数に対して否定的な評価が多くみられ，一方，出力形式のわかりやすさ，料金の安さ，サービス時間の長さに対しては，肯定的に評価している。大学図書館・研究所では，JOISが主として使用されているので，この結果は，大学図書館・研究所におけるJOISの評価を表しているともいえよう。

一方，企業グループでは，さまざまなサービスが導入されているため，国産データベース・サービスに対する評価は，肯定的から否定的まで広く分布している。例えば，キーボードの使いやすさ，データの遡及年数などはその一例である。一般に，企業グループは大学図書館・

研究所グループよりも肯定的な反応を示しており，とくに検索方法の容易さ，シソーラスの使いやすさに肯定的な傾向がみられる。また，出力形式のわかりやすさ，料金の安さ，サービス時間の長さに対しても肯定的に評価している。しかし，シソーラスの更新頻度，蓄積データ量・遡及年数に対しては，大学図書館・研究所グループほど顕著ではないが，否定的な評価が多くみられる。

グループによって評価に差異が生じた大きな理由は，経験年数あるいは利用頻度の違いによるものと思われる。これは，オンライン検索は回数をこなせばこなすほど，その技術が向上し，システムが使いやすくなるという経験的な色彩が極めて強い業務だからである。企業グループでは，JOISの利用機関が多いにもかかわらず，シソーラスの使いやすさ，検索方法の容易さに関する否定的な評価が大幅に少なくなっていることは，それを示しているといえよう。

C. 問題点

国産データベースを利用して感じる問題点を示したのが第6表である。この結果から，システム毎に検索コマンドやキーワード，コードが異なること，原情報がデータベースに収録されるまでのタイム・ラグの3点が，利用者にとって大きな問題となっていることがわかる。

“利用したいデータベースの所在，利用法がわからない”のは，国産データベースに関するクリアリング機能が十分整備されていないためであるが，この項目に回答した企業数が相対的に少ないことは，既述したように，両グループ間のデータベース利用度に有意な差があるこ

第6表 国産データベースの利用における問題点

	大 学 図 書 館 研 究 所	企 業	合 計
システム毎に検索コマンドが異なる	18	40	58
システム毎にキーワード，コードが異なる	14	31	45
タイム・ラグがある	16	24	40
利用料金が低い	10	11	21
ダウンロードができない	3	10	13
データベースの所在，利用法がわからない	5	1	6
導入時に要する経費が高い	4	1	5
その他	7	13	20

注 1 単位は機関数

注 2 複数回答

とを示していると思われる。これは、ダウンロード機能ができないことを問題点としてあげる機関数が、企業に多いことからもうかがわれる。(ダウンロード機能は、検索された情報の二次加工、つまりデータベースの高度利用と結びつく)

その他の問題点として、シソーラスが不備、キーワードが整理されていない、索引方針が不明確、カナ入力なので時間がかかる、蓄積情報の網羅性・収録文献数が不十分、自動切出しなど自然語処理が不十分などがあげられている。国産データベースの強力な武器である国内情報の蓄積に関しては、前述の否定的な評価がここでもみられる。

検索ツールが統一されていないのは、流通しているデータベースが、それぞれ違ったデータ構造やファイル仕様に基づいて構成されているためである。つまり、ある端末から2種類のシステムを利用する場合、それぞれ異なった通信制御手順やコマンド、検索方式でデータベースにアクセスしなければならない。その結果、利用者は利用するシステムやデータベースを変更する度に、そのシソーラスやコマンドなどのツールも変えなければならないのである。

米国では DIALOG や SDC といった膨大な量のファイルを持った大規模なデータベース・サービス提供機関が存在し、利用者は一つのシステムで多くの分野の情報を得ることが可能である。したがって、このような問題は少ないが、わが国では国産データベースを包括的に提供するサービス機関が存在していないため、この種の多様性は極めて問題である。

こうした不便さを解消するためには、各種のツールが統一化されることが理想的であるが、多数のコンピュータ、端末機器メーカーが技術開発を競い、ソフトウェア開発にも個性が発揮されている現在、それらを統一することは非常に難しいと思われる。そこで、現在ではシステム間にインターオペラビリティ(相互運用)と呼ばれるインターフェースを設置し、コマンド、キーワードといった検索ツールを互いに交換可能にする方法が考えられている。

各企業の利害関係や労力と時間の問題などが絡み、簡単には解決できないであろうが、利用者の立場のみならず、国産データベースを広く普及させ、産業として発展させるためにも、早急に標準化を促進していかねばならないことは確かである。

“利用料金が低い”という回答は、国産データベースの

利点として、“海外データベースに比べて国産データベースの利用料金が安い”ことがあげられている、第5表のデータと矛盾するようであるが、これは、蓄積情報の不十分さに対する不満の一表現であると共に、情報を財としてみる意識がまだ定着していない面があることを示しているように思われる。

シソーラスに対する不満は、前述のシステム利用にあたっての満足度の節で言及したように、検索に適当なデスクリプターが存在しないことによるものと考えられる。シソーラスを使用する検索では、その中に自分の検索したい主題を表すデスクリプターが存在していれば問題はないが、適当なデスクリプターがない場合には、最も近い語を使用するしかないからである。

また、シソーラスの更新頻度に対する不満も考えられるが、これは新しい用語の出現にシソーラスが対応しきれず時代遅れになってしまっているためであろう。例えば、『JICST 科学技術用語シソーラス』の最新版は1981年版であり、最新の用語が十分網羅されているとはいえない。

もし、原情報からの自由キーワードの取り出しが精巧になされているのであれば、主題と合致しないシソーラス中の不適当なデスクリプターを使うよりも、この自由キーワードを使った方が、より良い検索を行うことが可能になる。しかし、語の単位が明確でない日本語では、自由キーワードの取り出し、とくに自動切出しに高い精度を期待することは困難であり、したがって、自由キーワードに全面的に頼ることは危険である。これは、シソーラスへの依存度が英語文献よりも高いこと、その結果、シソーラスの不備に対する反応が強くなりがちであることを、それぞれ示している。

内容データの蓄積量と遡及年数についても問題が指摘されている。JICST 国内医学文献ファイルは、1981年4月から1985年時点まで年間約6万件を収録しているが、米国の MEDLINE 医学文献ファイルは、1972年から1985年時点まで年間約30万件と5倍以上も多くの情報を収録している。この差を考えれば、JICST 国内医学文献ファイルの利用者がその蓄積量と遡及年数に対して不満を感じるのはいたしかたなかろう。一方、JICST 科学技術文献ファイルは、1975年4月から現在まで年間約50万件と蓄積量、遡及年数とも少なくはない。それにもかかわらず、このファイルの蓄積データに対する不満があるのは、徹底した遡及的探索を目的としてデータベースを利用する利用者が多かったためと思われる。

IV. 国産データベースの展望

A. 需要の高いデータベース

現在提供されているデータベースの他に、これからどのようなデータベースが必要かを示したものが、第7表である。“学会会議開催予定情報”と“各種統計”のデータベースに対する需要は、大学・研究所、企業を問わず高いことがわかる。

わが国の統計制度はよく整備されており、各種の統計データが政府に集積されているが、そのほとんどがまだ未公開である。そのうちわずかながら磁気テープの形で民間に提供されているものもあるが、それを利用している機関はこの調査では皆無に等しい。これらの磁気テープは外郭団体経由で民間に提供される慣行になっており、民間業者による再加工、再編集、再販売は禁止されていることが多い。さらに、民間に提供されるテープは素材のものに近く、これを活用できるのは、大企業やシンクタンクなどコンピューター利用能力が高い機関のみであるといわれている。

したがって、政府保有の情報がデータベースとして、広く利用されることは、極めて意義があるといえよう。

“先端技術の研究・開発情報”は現在、誰が(どこが)、どんな研究を行っているかを知るための情報ということが出来る。このような情報を提供するデータベースとして、JICST 科学技術研究情報ファイルがあるが、企業におけるこれら最新情報のデータベースの重要度は、ますます高まることが予想される。

一方、“新聞、雑誌記事の全文”、“判例・法律の全文”、など全文データベースへの期待も大きい。新聞記事を対象にした全文データベースの提供が開始されている

第7表 新たに必要なデータベースの種類

データベースの種類	大学 図書館 研究所	企業	合計
学会会議開催予定情報	27	30	57
各種統計	22	30	52
先端技術の研究・開発情報	14	31	45
新聞、雑誌記事の全文情報	18	20	38
化学物質、化合物	17	18	35
化合物辞書	8	12	20
判例、法律の全文情報	10	6	16

注1 単位は機関数

注2 複数回答

第8表 公開が望まれる政府保有の情報

種類	大学図書 館研究所 等	企業	合計
工業統計調査	16	40	56
消費者動向調査	10	26	36
貿易統計	11	20	31
人口動態調査	22	8	30
国勢調査	20	6	26
商業統計調査	8	17	25
法人企業投資動向調査	2	19	21
景気動向調査	5	15	20
全国消費実態調査	3	15	18
事業所統計調査	2	15	17
工業実態基本調査	4	10	14
学校基本調査	12	1	13

注1 単位は機関数

注2 複数回答

が、一般に全文データベースの提供はまだ緒についたばかりである。なお、その他にこれから必要なデータベースとして、社会科学・人文科学分野の雑誌記事索引のデータベース、各種の物性データベースなどがあげられている。

第8表は、日本データベース協会が、政府保有データ流通方式整備への提言書⁹⁾の中で特に利用を希望するものとして選んだ24ファイルのうち、今回の調査で希望の高かったファイルを示したものである。企業グループでは、工業統計調査、消費者動向調査、貿易統計、法人企業投資動向調査、商業統計調査など企業活動と密接に結びつく情報に関心があるのに対し、大学図書館・研究所等は、人口動態調査、国勢調査、学校基本調査などの情報に関心があり、際だった対照をなしている。

上述のような新たなデータベースの提供を望む声がある一方、国産データベースを導入している大学図書館・研究所等51機関のうち20、63企業のうち13機関、全体で33機関(約30%)はこの質問に無回答であった。これは、統計データへの要望は高いが、政府保有データの中からどんな内容のものを何に利用したらよいかという具体的なニーズは、まだ持っていない機関が多いことを示している。さらに政府保有データの所在に関する情報不足のため、この種の磁気テープの存在すら知らない利用者も多いのではないかと考えられる。利用者の要望に応えるためにも、また国産データベース振興のためにも、政府保有データの公開は急がれるが、同時に利用者へのアピ

ールを積極的に行い、これらのデータが有効に利用されるよう働きかける必要もあろう。

B. データベース・サービスの発展

データベース・サービスの発展に必要とされる要件をまとめたのが第9表である。利用者側と提供者側の意識の違いを示すため、ここでは昭和60年に郵政省がデータ通信業者を対象にして行った調査⁷⁾と同じ項目について利用者の意見を求めている。第9表から明らかなように、2つの調査共通に、“情報収集体制の整備”の必要性があげられている。とくに利用機関では82%と極めて高い。これは、データの蓄積量や遡及年数に問題を感じている利用者が多いためであろう。実際、原情報の収集体制が十分整備されていないことは事実である。大学や政府など公的機関との協力体制を確立し、原情報収集段階での専門家による情報の評価、加工体制を充実させることが利用者側からも提供者側からも望まれているといえよう。

利用機関では、“情報入力コストの低減”、“情報更新の迅速化”が次に続いており、料金の低減化、情報の質

の向上に対する期待がうかがえる。一方、提供者側では、“データベース関連技術の開発”が次にあがっている。これは、近年のニューメディアの普及に伴って生じた、データベース利用形態の多様化に対応する技術開発が必要となっていることを示すものである。また、利用者にとってより使いやすい検索方法やコマンドの開発、インターオペラビリティの技術開発の必要性も意識されているといえよう。

“データベース・ネットワークの充実”については、提供機関の53%が回答しているのに対し、利用機関では25%しかなく、データベースの流通に対する問題意識は作成者側の方がずっと強いことがわかる。

以上の結果から、立場の違いにより解決すべき問題をとらえる観点が異なっていることがわかる。総じて利用機関は、データベースの収録内容に関連する事項に主として関心を示し、提供機関は、それに加えて処理技術にも関心が強いといえよう。わが国のデータベース発展のためには、それぞれの立場から問題を解決していくように努力していかなければならない。

C. 国産データベースに対する要望

国産データベースに対する利用機関からの要望は以下のようにまとめることができる。

a) 情報量の増大、網羅性の向上、タイム・ラグの縮小

情報量と網羅性については、情報収集体制の整備が今後の課題となるだろう。タイム・ラグは、特に科学技術分野において深刻な問題である。科学技術文献がデータベース化されるまでには、検索を可能にするための加工に4-6ヶ月の作成期間が必要といわれる⁸⁾。タイム・ラグの短縮は、日進月歩する科学技術分野において非常に重要である。将来、自動分類、自動キーワード付与、自動抄録といった情報処理技術の開発が軌道にのれば、その作成期間を大幅に短縮することができるだろう。

b) デスクリプターの絶対数の増大、シソーラスの充実、自由キーワードによる検索の高度化

わが国では、従来索引誌、抄録誌の作成・刊行に伴う索引作業に関する知識、技術に対する評価が高いとはいいがたく、データベースの構築・利用においても、収録内容の充実、索引の充実よりもコンピュータ技術を駆使した情報処理に力を注ぐきらいがあった。このような傾向は、デスクリプターの付与、自由キーワードの抽出、シソーラスの整備などに関する作業を軽んじることにな

第9表 データベース・サービス発展のための要件

	回答機関数	
	利用機関	提供機関
情報収集体制の整備	97(82.2)	93(73.8)
データベース関連技術の開発	49(41.5)	82(65.1)
データベース関連技術の高度化	49(41.5)	48(38.1)
情報入力コストの低減	57(48.3)	77(61.1)
情報更新の迅速化	56(47.5)	68(54.2)
データベース・ネットワークの充実	30(25.4)	67(53.2)
クリアリング機能の強化	8(6.8)	16(12.7)
データベースにかかわる諸権利の明確化	14(11.9)	23(18.3)
データベース検索技術の開発	20(16.9)	27(21.4)
データベース検索者の養成	31(26.3)	19(15.1)
政府機関等保有のデータの公開	27(22.9)	37(29.4)
公的資金による育成強化	42(35.6)	31(24.6)
英語による展開(海外販売)	19(16.1)	11(8.7)
その他	5(4.2)	1(0.8)

注 1 有効回答機関数 118 126
 注 2 複数回答あり

利用の側面からみた国産データベースの特徴

り、今回のアンケート調査でも指摘された、シソーラスやキーワードに対する問題点を生じさせる遠因となっている。したがって、検索キーの充実にはこの面での意識の改革も重要である。

- c) 入力コスト低減による利用料金の低下，料金支払い先の一元化，一次情報の提供

検索結果に基づいて一次情報をも提供するためには、全文データベースの作成を進めなければならない。近年、光ディスク等の蓄積媒体やワードプロセッサなどの進歩により、原情報そのものをデータベース化することが可能となってきた。近い将来全文データベースは確実に増加すると思われる。

V. 国産データベース振興の方向

以上の調査結果を踏まえて、データベース・サービス振興のための課題とそのために行うべき方策をまとめると、次のようになる。

A. 構築面

1. 情報源の確保

現在、広範で質の高い大量の情報を効率的に低コストで収集できる体制は、十分整備されているとはいえない。この体制を整備するためには、政府による公的データ収集の拡大と、政府保有データの公開、磁気テープによる提供が急務である。政府保有データは、社会経済活動のあらゆる側面を反映し、質量ともにその価値は極めて高い。とくに統計類は公的権力による強制力と巨額の費用をかけて整備されており、規模の面および正確性の面でこれに及ぶものはない。また、従来は手作業で集計され印刷物にまとめられていたものが、近年のコンピュータ技術の発展に伴って、機械集計され磁気テープに記録されるようになってきた。

しかし、総務庁が旭リサーチ・センターに委託した調査⁹⁾によれば、磁気ファイル化されている805のデータのうち民間への提供が可能なのは41.6%の336件しかなく、この中でも無条件で提供されるものは99しかない。このように外部への提供が制限されている理由として、

- i) 法令、内部規程などにより部外秘にしている
- ii) 部内データであり提供の対象としていない

などがあげられる。さらに、冊子体としては公開されているが磁気テープ化されていないものや、民間への提供の条件にデータ管理者の承認を必要とするものも多い。こうした状況では、利用したいデータがあるとして

も容易には入手できないだろうし、円滑な利用も期待できない。また、この種の磁気テープの存在すら知られていない可能性も大きい。

最近この問題に対する政府、民間の動きは活発で、統計審議会では磁気ファイルについて制度上の検討を行っているし、民間でも日本データベース協会が3回にわたって公開の提言を行っている。したがって、近い将来政府保有データの公開に向けての急激な展開が期待される。

2. 政府の助成

データベースの構築には多大の初期投資を必要とし、その回収期間も長い。構築後もデータの更新、追加、システムの維持管理など莫大な資金が必要である。このため、利用者数が少ないと採算はとれず、重要度は高くても利用頻度の低いデータベースの構築・維持は困難になる。したがって、わが国にとって必要なデータベースに関しては、政府の思い切った公的投資が必要である。

データベース振興センターの設立は、国産データベースの振興をはかるための一施策である。このセンターの今後の活動実績は、国産データベースおよびデータベース・サービスの発展に大きな影響力を与えられると思われる。

3. 新しいデータベース

データベースに収録される情報は、文字・数値情報だけではない。画像情報の中にもデータベース化の意義があるものも多い。したがって、この種のデータベースの研究・開発、および従来のデータベースとの複合利用などを考える必要がある。そのためには、画像情報の入力、処理などに関連する技術の開発が極めて重要である。

4. 英語によるデータベース

欧米諸国では、日本の情報は入手しにくく、情報の流れが一方に偏っているという批判の声が高まっている。とくに米国ではこの問題に対する関心が強く、今後海外への提供を積極的に進めていかなければ、情報摩擦がますます拡大する恐れがあるとさえいわれている。全体の80%を海外のデータベースに依存している現在、海外流通の促進をはかり、日本に流入する情報量と日本から流出する情報量の間にみられる不均衡を是正することは、極めて重要である。

わが国のデータベースが海外で流通するようになるには、多くの障害を乗り越えねばならない。なかでもデータベースの収録内容の翻訳は必須である。現在、自動翻訳システムが開発されつつあるが、まだコスト、品質、

迅速性の面で解決すべき問題が多く残されている。今後一層研究・開発を推進していかねばならないだろう。さらに、海外での需要が高いわが国の情報、例えば、経済、先端技術、産業に関する情報、特許情報等の内容を充実させ、英語化していくことが望ましい。日本科学技術情報センターが現在開発中の英文データベースはその第一歩であり、成果が期待される。

B. 流通, 利用面

1. 通信回線の自由化

わが国における通信利用は共同利用制限, 他人使用制限, 相互接続制限など制度上制約があり, 複数のデータベース・サービスを接続した大規模なシステムの構築が妨げられていた。オンライン・データベース・サービスの発展のためには, このような回線使用制限の根本的な見直しが行われ, 完全に自由化される必要がある。

2. クリアリング機能の強化

利用者が必要としているデータベースの所在案内, 内容案内, 利用案内などを行うクリアリング機能は, 利用者がデータベースを活用する上で必須の機能である。国産データベースの利用を高め, 流通を促進させるためには, このようなクリアリング機能を強化しなければならない。現在データベース振興センターは、『データベース台帳総覧』に基づいたデータベース・クリアリング・システムを開発中である。この開発が順調に推移すれば, 近い将来データベース・クリアリング情報のオンライン提供と共に, その情報を収録するデータベースの提供が可能となろう。

3. 著作権にかかわる法体制の整備

わが国の著作権法では, データベースについては何の言及もなされていないため, 契約当事者以外の第三者による海賊行為からデータベースを保護する手立てはないといえる。一方, 米国では1980年12月に著作権法が改正され, データベースは編集著作物であり著作権保護の対象となる, とされている。

このような動向を受けて, 文化庁は昭和60年9月, データベースと著作権に関する最終報告書を取りまとめた。これは, データベースを著作権法の枠組みの中で保護するという前提に立ち, データベースはそれ自体で著作権を持つことを主張している。しかし, この報告書は, データベース作成者の権利保護については十分だが, デイストリビューターの権利については不十分であるという声も聞かれる¹⁰⁾。今後, この点について検討を重ね,

しっかりとした法体制が築かれることが期待される。

4. ネットワークの整備

わが国には中小規模のシステムが多数存在し, しかも標準化が遅れているため, 利用者にとって不便な点が非常に多い。しかし, 今後も米国のような巨大システムが出現するとは思われないので, 近い将来, それぞれの分野でデータベース・サービス業者によるサブ・ネットワークを築き, さらにそれぞれの中心となるデータベース・サービス業者を結ぶネットワークを構築していくことが望まれる。そのためには関連技術の開発, 通信回線や標準化に関する問題の解決が急がれる。

また, 国産データベースの海外流通を促進するためには, 国際ネットワークの整備, 海外サービス体制の強化などが必要である。そのためには, 多大な費用と長い時間を要することはいうまでもない。しかし, 日本が世界の情報の10%を占めるようになった現在, その情報を海外に提供することは, 国際社会の一員である日本の責務である。

VI. おわりに

データベース・サービスが一つの産業として独立するためには, 多年にわたる莫大な資金と人材の投入が必要である。米国では過去20余年にわたり多様な分野でデータベースの構築に多額の助成を継続的に行い, 政府保有のデータも公開されてきた。日本では最近やっと政府の助成が行われるようになったばかりである。今後わが国のデータベース業発展の鍵を握るのは, このような政府の動きであろう。

通商産業省はデータベース・サービスの重要性を認識し, データベース振興センターを設立すると共に日本データベース協会(DINA)の設立を支援してきた。さらに, 産業構造審議会情報産業部会に設置されたニューメディアに関する小委員会, 情報提供サービス振興小委員会を通じてさまざまな取り組みを行っている。

例えば, 昭和58年12月に発表された, ニューメディアに関する小委員会での検討結果では, データベース・サービスを今後の情報化の進展を促進する上での重要な産業として位置づけ, 早急に対応すべき重要課題としてこれら産業の基盤整備の必要性を指摘している。また, 海外への提供が可能なデータベースを整備するための提言も行っている。さらに, 情報提供サービス振興小委員会では, 昭和60年11月にまとめた中間報告の中で, 多種多様なデータベースの流通を担当する総合型デイストリビュー

利用の側面からみた国産データベースの特徴

ーターや特定分野の専門型ディストリビューターの育成が、データベースの利用価値を増大し、円滑な流通を図る上で重要な課題であることを指摘している。

一方、データベース・サービス産業を確立し、データベースの構築と流通を促進し、知的資産の共同利用による高度情報化社会の実現と発展に寄与することを目的として発足した日本データベース協会は、データベースの構築の助成と利用、流通の促進、政府保有データの公開や統計データセンターの構想に関して提言を行ってきた。

データベースの海外依存度を縮小し、国産データベースをさらに発展させるためには、上述したように政府と民間が緊密な連携を保ち、データベース産業の育成をはかることが極めて重要なのである。

謝 辞

本稿の基になりましたアンケート調査に御協力いただいた各位に対して、心から感謝いたします。

- 1) パネルディスカッション：国産データベースの構築と提供はいかにあるべきか。情報管理。27(11), p. 937-47 (1985)
- 2) 通商産業省。データベース台帳総覧 昭和59年度版。日本情報処理開発協会, 1983, 1242p.
- 3) データベース・サービスに関するユーザー意識調査。データベース振興センター, 1985, 160p.
- 4) JOIS 利用アンケート調査結果。JOIS ニュース, no. 10, p. 5-7 (1981)
- 5) 細野公男他。オンライン情報検索が冊子体購入に与える影響。情報管理, 27(11), p. 971-80 (1985)
- 6) データベース白書86年版。p. 306-7.
- 7) 郵政省電気通信局。データ通信業実態調査報告書。東京, データ通信協会, 1985, 95p.
- 8) 名和小太郎。ファクト・データベースの現状と課題, 特集：データベースの現状と課題, 情報通信学会誌, 2(3), p. 41-5 (1984)
- 9) 行政管理庁行政管理局。行政情報システムの社会的課題に関する調査研究報告書。東京, 旭リサーチ・センター, 1983, 445p.
- 10) 科学技術庁編。総合レビュー 科学技術情報活動の現状と展望 7巻。データベースの高度利用。大蔵省印刷局, 1985, p. 145-55.