

解説



学術情報データベースの構成と利用

データベースサーチャの視点†

三輪 真木子††

1. はじめに

データベース産業の発展にともない、世界中の公開された情報のかなりの部分が商用データベースとして端末からオンラインで利用できるようになってきている。企業などの組織体では外部の情報を商用データベースを通じて効率的に入手するために、各種の商用データベースサービス機関と契約を結び、パスワードを入手している。また、エンドユーザに代わって検索業務を担当するサーチャと呼ばれる専門職を配置する機関も増加している。一方、企業などの組織体や個人顧客の依頼に応じて検索業務を実施するデータベース代行検索機関が100社以上あり¹⁾、ここでもサーチャが検索業務を担当している。代行検索機関は、大企業の関連会社、政府機関や自治体の外郭団体、データベースサービス機関やその代理店、調査会社などで、情報収集や調査業務の一環として検索業務を行っているものが大部分を占める。

サーチャの担当する検索業務の内容はエンドユーザや顧客の企業秘密やプライバシーに触れるものも多く、その詳細については公表されていないため、全体の動向を把握することは困難である。ここでは、すでに文献などで公表されている他の事例を参照しつつ、代行検索機関のひとつであるエポックリサーチ社の事例を中心に、最近の動向を論じる。

2. サーチャの仕事

商用データベースの普及と情報技術の発展にともない、データベースの提供形態やユーザインタフェースが多様化している。たとえば、従来オンラインで提供されてきた商用データベースと同内

容のものがCD-ROMなどのメディアで提供されたり、新聞や雑誌記事などの一般的なオンライン・データベースで従来のコマンド方式だけでなくメニュー方式によるユーザインタフェースが提供されるようになってきている。これらのCD-ROMやメニュー方式による商用データベースは、エンドユーザが自ら検索する形で普及しており、サーチャは主にコマンド方式によるオンラインデータベースの検索を担当している。最近ではエンドユーザ検索を指向する組織体が増え、コマンド方式による商用オンラインデータベースの検索をエンドユーザに開放している組織体もみられる。しかしながら、現時点ではエンドユーザ検索は増えておらず、自ら検索を行って見て検索の難しさを理解したエンドユーザからのサーチャへの検索依頼は、むしろ増加する傾向が生じている²⁾。

ここでは、サーチャが担当する検索業務の全体像と、重点事項を紹介する。

(1) 検索業務の流れ

検索業務は、エンドユーザの情報要求を把握することに始まり、検索結果をエンドユーザに提供することで完了するが、その間に行われるさまざまな手続きや意思決定を整理すると、おおむね図-1に示すような何段階かのステップとして捉えられる。

エンドユーザから検索依頼を受託する際には、質問記入用紙にエンドユーザが必要事項を記入する質問紙法と、対面または電話などによりサーチャがエンドユーザの情報要求を引き出すインタビュー法がある。いずれの方法をとるにせよ、サーチャはエンドユーザの情報要求を十分に把握し、それがデータベースから得られるような情報であるかどうかを、データベースに関する既得知識や自らの検索経験に照らして判断する。ちなみに、商用データベースには企業秘密や個人のプライバシーに関わる機密情報は含まれていない。また、

† Perspectives from a Database Searcher's Stand Point by Makiko MIWA (Epoch Research Corporation).

†† (株)エポックリサーチ

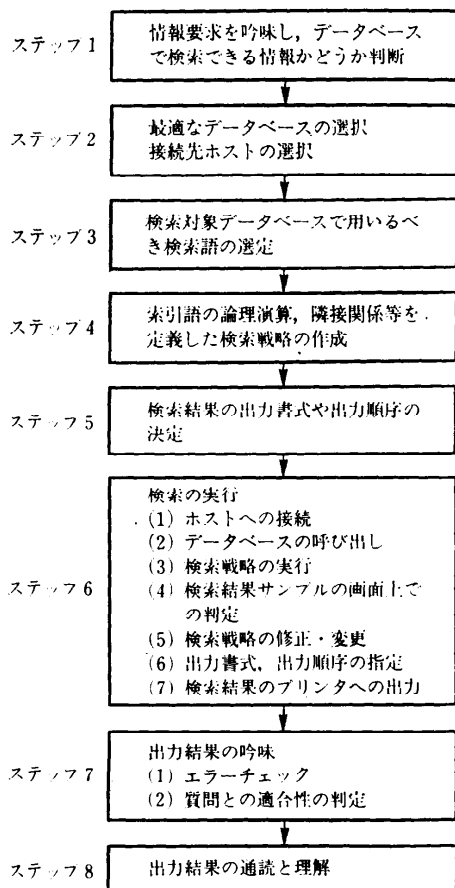


図-1 情報検索の手順

発展途上国や独立国家共同体（旧ソ連）ならびに東欧諸国の情報はデータベースにはほとんど含まれていない。さらに、商用データベースが構築される以前の昔にさかのぼる情報は、遡及入力が行われている図書目録などの一部のデータベースを例外として、ほとんどデータベース化されていない。

情報要求の中には、新たな研究開発に着手するに先だって行われる重複研究調査や特許調査のように、関連のありそうな文献や情報をもれなく検索することを要求される「高再現率検索」と、もれがあってもよいが関連する主な文献や情報のみを検索するよう要求される「高精度検索」があり、サーチャは質問記入用紙の記入事項やインタビューから、エンドユーザの情報要求がどちらに該当するかを判断しなくてはならない。また、検索に先だって費用を見積るよう要求される場合には、情報要求を満たせそうなデータベースで主な

検索語のポスティング数（出現頻度）を確認する「概略検索」を行うこともある。

(2) データベース選択の視点

ひとつのデータベースの中には数十万から数百万件のレコードが収録されているので、多くのデータベースを対象として検索業務を担当するデータベースサーチャは、理論上は数十億件のレコードにアクセスできることになる。データベースサーチャの仕事の真髄は、この数十億件のレコードの中から求める情報を含むレコードをいかに効率的に探し出せるかという点に集約されるといえよう。ただし、最短時間で求めるレコードにたどり着けるか否かは、最初にアクセスするデータベースの選択にかかっている。もし、求める情報が含まれていないデータベースを最初に選択すると、その検索に要する接続時間料金、通信料金、およびサーチャの人件費は不用の出費となり、情報入手までの時間の遅れも生じる。

情報要求に最も適したデータベースを選択するノウハウの大部分は、サーチャの検索経験とデータベースに関する知識に依存するが、補助資料としてデータベースディレクトリを参照することもできる。データベースを選択する際には、個々のデータベースについて表-1に示す点を把握していることが必要である。

データベースディレクトリの中には、商用データベースとしてオンラインで提供されているものもあるので、最初にこのデータベースにアクセスするというのもひとつの方法である。さらに、ひとつのサービスシステムの中で利用できる複数のデータベースのおのおのにおける検索語のポスティング数を確認できる「クロスファイルインデックス」と呼ばれるファイルが提供されている場合には、このファイルにアクセスして主な検索語の出現頻度を確認するという方法も利用できる。また、複数のデータベースを一括検索できるサービスシステムでは、関連のありそうなデータベースをすべて選択して横断的に検索することも可能である。

(3) 検索戦略

現在提供されている商用データベースでは、索引ファイルに登録されている文字列や記号を端末側から入力した検索語と照合し、それらが一致した場合のみ検索結果として出力できる完全照合方

表-1 データベース選択の着眼点

重点項目	概要
収録主題分野	科学技術分野の専門的なデータベースにおける境界領域の収録範囲
情報源の形態	文献で「単行書」、「雑誌記事」、「特許」、「学位論文」などの構成比
収録開始年	遡及検索で、いつごろの情報までさかのぼって検索できるか
タイムラグと更新頻度	最新情報を求める際に、どの程度新しい情報が得られるか
収録情報の網羅性	文献で、特定の雑誌の記事がもれなく収録されているか
収録情報の対象地域	地域性のあるデータベースではその地域、国際的なデータベースでは収録情報の国別地域別構成比
データベースの記述言語	検索結果を読んで内容の適否を判断できるか
収録データ項目	必要なデータ（抄録、全文、図表、財務データなど）が含まれているか
検索補助資料	そのデータベースの検索に必要なシソーラスや分類表はあるか

式が採用されている³⁾。そのため、特定の主題に関わる情報を求めて検索を行う場合には、データベース中でその主題が表現されている形式に合わせて検索語（言葉、分類コード、記号など）を指定することが必要である。たとえば、化合物の検索では、探したい化合物の名称だけでなく、構造式や構造図を検索キーとして検索を行う場合が多い。検索主題を言葉で表現する場合には、あらかじめ決められている統制言語をシソーラスなどから選定して検索語とする統制語方式と、本文中で用いられている言葉がそのまま検索語として利用できる自由キーワード方式があり、両者を併用しているデータベースも多い。どちらの方式を採用しているかによって、検索語の選定方法が異なる。

情報要求の中に複数の主題が含まれているような場合には、検索目的に合わせて主題間の関係を定義することが必要である。

検索戦略とは、情報要求の中から検索主題を抽出し、それらを検索語に置き換えるとともに、検索主題間の関係を論理演算子や近接演算子などの演算子を用いて表現したもので、実際の検索に先だって準備しておく。最初に準備した検索戦略で検索を行った結果が情報要求と適合しなければ、検索中に準備した検索戦略の一部を修正するか、いったん検索を中断して再度検索戦略を作成し直す。実際の検索では適合情報を最初の検索で得られることはほとんどないので、図-1のステップ3からステップ7で何度か試行錯誤を繰り返す場合が多い。

(4) 検索結果の評価と加工

サーチャの受託する検索業務は、検索結果として得られたデータベースからの出力リストをエンドユーザに引き渡せば完了するものとみなされているが、実際には、検索結果に対する評価や加工を依頼されることも多い。特に、全文データベースが増加して、一度の検索で大量の検索結果が得られるようになると、得られた検索結果を吟味して必要な情報のみを抽出したり、検索結果を分析して要約するような作業を、エンドユーザから依頼されることが増えている。また、外国語の出力結果の日本語への翻訳を依頼されることもある。

情報処理技術や通信技術の発達によって、短時間で大量の情報が入手できるようになったが、人間がそれらの情報を吸収し理解する速度にはほとんど変化がみられないため、データベースサーチャには、エンドユーザに代わってデータベースを検索するだけでなく、得られた検索結果をエンドユーザに代わって読み、理解し、分析することまでもが期待されはじめているのである。

3. 二次情報データベースの利用

商用データベースは、収録情報の特性から、二次情報データベースと一次情報データベースに分類される⁴⁾(表-2 参照)。二次情報データベースは、より多くの情報や原文を求める利用者にその情報源を提示する機能をもち、文献の書誌情報(索引、抄録を含む)を収録する書誌データベースと、名鑑(名簿、企業ディレクトリを含む)などの案内データベースに分けられる。このようなデータベースの特性は、利用者や利用方法にも反

表-2 Directory of Online Databases のデータベース分類

データベースの種類	概 要
Reference Databases (二次情報データベース)	より多くの情報や全文を求める利用者に、その情報源を示し紹介する
Bibliographic (書誌データベース)	雑誌記事、報告書、特許文献、学位論文、会議録、単行書、新聞記事などの印刷された文献の書誌情報で、抄録をとまうものも含む
Referral (案内データベース)	案内情報で、出版されていない情報に関する要約や概要をとまうものもある
Source Databases (一次情報データベース)	手を加えていない生データ、文献の全文、もしくは電子媒体による流通のためにのみ作成された素材
Numeric (数値データベース)	生の実態調査データや、統計的に処理された要約データで、通常は特定の変数(特定の製品や産業の生産・出荷など)の長期間にわたる測定値を表す時系列データ
Textual-Numeric (テキスト数値データベース)	多くの項目をとまうレコードからなるデータベースで、一部の項目に数値データを含むもの、物性データなどの辞書やハンドブック形式のものを含む
Full Text (全文データベース)	新聞記事、仕様書、判例、ニューズレターなどの全文を含むデータベース
Software (ソフトウェア・データベース)	端末側のコンピュータでダウンロードして利用できるコンピュータ・プログラムからなるデータベース
Images (画像データベース)	化合物の構造図、写真、地図、商標、ロゴマークや、イラスト、設計図素などの画像を含むデータベース

映されている。

(1) 書誌データベースの検索

書誌データベースの検索は、既往文献調査に利用される。具体的には、以下のような目的で利用されている。

- 特定著者の業績調査：たとえば、招待講演者を選定したり講演者の紹介を作成することを目的としている。個人の経歴調査と並行して行う場合もある。

- 特定主題の既往文献調査：研究着手時に参考事例として利用できる既往文献を探したり、重複研究がないかを確認するために、特許や学術論文などの文献調査を実施する。

- 図書の選定および発注：特定分野の単行書や雑誌(記事)を収集する目的で、文献調査を実施し、得られた文献リストの中から必要なものを選定する。原報発注機能などを用いて図書を発注する場合もある。

- 目録作成の補助：図書館などで所蔵図書の目録データを作成する際に、すでにデータベースに収録されている当該図書の目録データを流用する目的で利用する。この場合、検索結果は端末側でダウンロードされ、必要に応じて修正や情報の付加を行った後に、図書館内の目録データベースに取り込まれる。複数の図書館の目録データを一体化した書誌ユーティリティも、この方式を採用し

ている。

(2) 案内データベースの検索

案内データベースの典型的な用途は、機関調査および人物調査で、以下のような目的で利用される。

- 企業概要調査：企業名は分かっているが業種や規模や所在地が不明であったり、取引先候補や競争相手などの特定企業に関してより詳細な情報を入手したいという目的で、企業の財務データ、株主や経営者の名前、関連会社、メインバンク、業種、取扱製品などを調べる際に、この種の調査を行う。取引先や販売先を選定するために、売上高、所在地、業種、といった特定の条件を満たす企業について一括して検索を行う場合もある。

- 人物調査：名前は分かっているが経歴が不明であったり、特定の人物に関して詳しい経歴を入手したいという目的で、個人の学歴や職歴の調査をする場合に行う。委員会の委員や講演会の講師を選定する際の基礎データとして、特定分野の専門家などについて一括して検索を行う場合もある。

- ダイレクトメール宛名ラベルの作成：特定条件を満たす企業や個人を捜し出し、データベース中に収録されている住所や所在地のデータをダイレクトメール宛名ラベルの書式で出力して利用する場合もある。

4. 一次情報データベースの利用

手を加えていない生データや文献の全文などを含むデータベースを、一次情報データベースと呼ぶ。コンピュータの記憶容量の増大と出版・印刷過程へのコンピュータの導入にともなう電子化テキストの増大により、一次情報データベースの数は急激に増大している⁵⁾(図-2参照)。一次情報データベースは二次情報データベースとは異なり、作成時における索引付けや抄録作成といった人手による作業が省力化できるため、作成コストが低く時間のずれも少ないので、今後ますます増加することが予測される。

一次情報データベースは、実態調査データや統計データなどの数値データベース、表側・表頭などに文字データをともなうテキスト数値データベース、新聞・雑誌記事、判例、議事録、ニューズレターなどの全文データベース、パブリック・ドメインのプログラムなどからなるソフトウェア・データベース、および地図、写真、設計図などの画像データベースに分類され、以下のような目的で利用されている。

(1) 数値データベースの検索

数値のみからなるデータベースの検索は、対象分野の特性や数値の意味を十分に理解していないと困難であるため、この種の検索は通常エンドユーザによって行われており、サーチに検索を依頼されることはほとんどない。

(2) テキスト数値データベース

物性データや、意識調査、市場調査、国勢調査などの調査結果の表などが、表頭・表側や概要を

説明するテキストとともに収録されており、以下のような目的で検索されている。

- 特定物質の同定：化学および医薬分野の製造業における研究開発部門で、実験などの過程で生成された化合物を同定したり、その特性（毒性、副作用など）を把握したりする目的で利用されている。この種のデータベースの検索は、専門分野（化学、薬学など）の基礎知識を有するエンドユーザやサーチャが担当している。

- 市場動向の把握：特定業界の動向を分析するために、既存の市場調査結果を利用するような場合に、この種のデータベースを検索する。新たに市場調査を実施する際の予備知識を得る目的で利用することもある。

- 意識調査：特定の問題に対する一般大衆もしくは特定グループに属する人々の意識の変化を追跡し、政策策定や商品企画の参考として利用するといった目的で、既存の意識調査結果を検索する。

(3) 全文データベース

新聞・雑誌記事や判例、議事録、ニューズレターでは、印刷形態と全文データベースの両方で利用できるものが増えている。印刷媒体ではなくあえてデータベースを利用するのは、以下のような場合である。

- 情報の即時的入手：日本では、全国紙の大部分と地方紙および専門紙の一部が全文データベースとして提供されており、ものによっては印刷された新聞より早く情報を入手できるものもある。欧米では各地の主要新聞や専門分野のニューズレター、ビジネス雑誌、および専門誌の一部が全文データベースとして提供されており、日本から印刷物を郵送で入手するより全文データベースを利用したほうが早く情報が入手できる。

- 切抜き作業の省力化：取引先や競合他者に関する記事、特定分野の政策や省庁の通達に関する記事などを毎日切り抜いて回覧したりスクラップする代わりに、「検索テーマ」として登録しておく、関連記事が自動的に検索される。記事を紙面イメージのデジタル画像として光ディスクに登録し、要求に応じて顧客のファックスに出力するサービスも、この目的で利用されている。

- 特定主題のニュースや記事の入手：全文データベースの中には、書誌データベースと同様の主

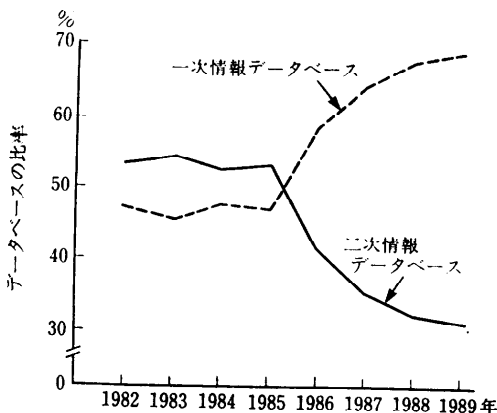


図-2 一次情報データベースと二次情報データベースの経年変化

題索引や分類が付与されているものもある。この種の全文データベースでは、書誌データベースと同様、特定主題の記事を探す目的で検索を行い、該当する記事があれば全文を出力する。索引や分類がなくとも、タイトルやテキスト中の言葉（固有名詞など）により求める情報を検索することも可能である。外国語のニュースや記事の場合には、印刷物を入手して目次や本文の全てに目を通すより全文データベースを検索するほうが探す手間やコストを節約できるために、あえて全文データベースを検索する場合も多い。

●特定分野の動向調査：特定の社会現象や政策動向を長期間にわたって調査し、変動要因や背景事情を把握する際には、関連する新聞・雑誌記事を適及的に収集し、それらを分析する手法が有効である。この目的のために記事を探して入手する手段として、データベースが利用される。この種の調査では、出力される記事が膨大な数にのぼることが多い。

(4) ソフトウェア・データベース

パブリック・ドメインのソフトウェアを収録したデータベースは、そのソフトウェアを利用するエンドユーザが自ら検索し、必要に応じてダウンロードして利用しているため、サーチャが代行することはない。

(5) 画像データベース

商用データベースとしてオンラインで一般に提供されている画像データベースは、商標・ロゴマークや化合物の構造図などに限定されている。大規模な製造業には、企業内に商標調査部門を設け、専属のサーチャを配置しているものもある。また、化合物の構造式検索は、化学・医薬分野の研究・開発機関で書誌データベースや物性データベースの検索と併せてサーチャが担当している。

5. 分野別にみたデータベース利用の特徴

日本で利用できる商用データベースのディレクトリとして広く利用されている『データベース台帳総覧』では、収録対象のデータベースをその内容から、「ビジネス」、「自然化学・技術」、「社会・人文科学」、「一般」の4分野に分類している。1991年12月末時点におけるデータベースの分野別分布は、図-3のようになっている⁶⁾。以下では

これらの分野ごとに、データベース利用の特徴を述べる。

(1) 自然科学・技術分野のデータベース利用

『データベース・サービスに関するユーザの意識調査』⁷⁾ (以下『意識調査』)によれば、この分野でよく利用されているデータベースは、特許関係(特許・実用新案ファイル、WPI、商標ファイル、CLAIMS など)、医学関係(MEDLINE および JICST 医中誌国内医学文献ファイル)、化学関係(CA-SEARCH および CA)のほか、科学技術分野全般を対象とする JICST 科学技術文献ファイル、物理学・通信・コンピュータ分野の INSPEC 生物学分野の BIOSIS で、いずれも抄録を含む書誌データベースである。したがって、自然科学技術分野におけるデータベースの利用は、特定分野の文献調査を主たる目的としているものと想定される。

科学技術分野の書誌データベースの多くは、主題別に分かれているが、おのおのの主題について世界中の文献を収録しているので、特定のテーマに関する世界中の重要文献を網羅的に検索することが可能である。なお、この分野のデータベースは、ディスクリプタや主題分類コードなどの統制言語により索引されているものが多いため、主題による検索を行う場合にはシソーラスや主題分類表を参照する必要がある。

特にこの分野のデータベースの利用が盛んな医

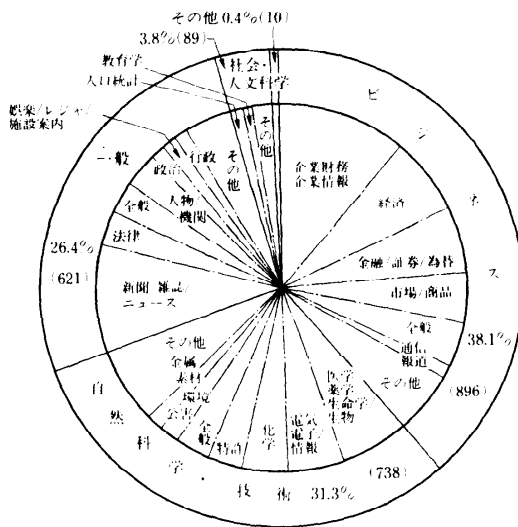


図-3 商用データベースの分野別分布

薬品製造および石油・化学工業では、ほぼ90%の企業がデータベースを利用しており、その大部分では専任のサーチャを配属している。これらのサーチャには、検索技能のみならず専門分野の基礎知識が要求されるため、薬学や化学分野の出身者が多い。この業界では、基礎研究着手時の特許調査や重複研究調査、応用研究段階における競合他社の動向調査、製品開発段階における市場動向調査といったように、研究開発の諸段階でさまざまなデータベースが利用されており、データベースは業務に不可欠なツールとして定着している。機器製造業や食品工業などでも同様の傾向がみられる。

(2) ビジネス分野におけるデータベース利用

『意識調査』によれば、この分野でよく利用されているデータベースは、COSMOS 2などの企業情報データベースである。企業情報データベースは国別に作成されているものが多いため、これらを利用する際も国別・地域別のアプローチが必要である。最近では欧米各国の企業情報データベースが利用できるようになったので、企業名と所在地が分かれば、検索が可能である。企業情報データベースでは企業を業種により分類している。国内では業種分類に独自の業種コードを採用しているものが多い。一方、米国や欧州では各国の標準産業分類コード(SICコード)を用いているものが多いため、業種からアプローチするにはコード表が必要である。企業情報データベースは、企業信用調査やダイレクトメール宛名ラベル作成などのほか、業界内の企業動向を長期的に捉えるためにも利用される。

一方、金融・証券関係では、ディーラやアナリスト向けに取引状況をリアルタイムで提供するデータベースが数多く提供されている。彼らは、ディーリング・ルームに設置された複数のモニタ端末の画面に世界中の各取引所から提供される株価や取引状況を参照しながら、取引業務を行っている。

ビジネス分野では、経済統計などの数値データベースも数多く提供されているが、この種のデータベースを利用するには検索だけでなく目的に応じた統計処理も必要となるため、エンドユーザによる利用が主体となっている。

(3) 社会・人文科学分野におけるデータベ

ス利用

図-3に示されているとおり、社会・人文科学分野のデータベースはきわめて少ない。その理由は、当該分野のデータベースの需要が少ないためと考えられる。商用データベースが事業として成り立つためには、有料で情報を入手する意思のある潜在利用者が数多く存在していなければならない。ところが、この分野のデータベースの潜在利用者は研究者に限られ、彼らの求める情報は歴史的なものから新しいものまで多岐に及んでいるので、それらをすべて満たすようなデータベースの構築には多大なコストがかかる。それにも関わらず、研究者には高額な料金を支払うことを期待できないため、費用対効果の観点からみて収益性に乏しい。そのため、日本ではこの分野での商用データベースの構築があまり進んでいない。

一方、教育分野のERIC、心理学分野のPsyc-INFO、社会学分野のSociological Abstractsなどの公的機関や学会などの非営利機関で作成されている海外の社会科学分野のデータベースは長い歴史をもち、よく利用されている。日本でも、ERICはサーチャ教育用のツールとして活用されている。

(4) 一般分野におけるデータベース利用

『意識調査』によれば、この分野でよく利用されているデータベースは、日経新聞記事ファイル(NIKKEIファイル)と朝日新聞記事データベースである。つまり、一般分野では新聞記事の全文データベースが、主として切抜き業務の代替手段として利用されている場合が多い。全文データベースでは、多くの場合ディスクリプタや分類コードなどの統制語による索引が記事単位で付けられているため、記事中のパラグラフなどの特定の情報を検索しようとする場合には自由キーワードに頼らざるをえない。そのため、長期にわたる業界動向や市場動向を把握するための補助的ツールとして新聞記事データベースを利用する場合には、検索もれやノイズを排除するために、自由キーワードとして用いる用語の選定に十分な配慮が必要である。この種のデータベースの多くはメニュー方式で提供されており、特にサーチャによる検索を必要としないため、企業ではエンドユーザが利用している場合が多い。このほか、人物・機関情報などの参考データベースも、この分

野に含まれ、やはりエンドユーザに利用されている。

法律分野のデータベースは、日本ではまだ数が少ないためあまり利用されていないが、米国ではこの分野のデータベースが最もよく利用されている。これは、米国が判例法を採用しており、裁判にあたって過去の類似判例を包括的に調査する必要があるため、判例の全文データベースが充実しているからである。米国では、LEXIS や WESTLAW といった判例および法令のデータベースが弁護士およびその支援をするサーチャによって利用されており、法曹界では業務に不可欠のツールとなっている。

6. 既存データベースの限界と将来への期待

現在利用できる商用データベースとその検索システムには多くの問題点があり、それが改善されないかぎりだれもが気軽に利用できる道具とはなりえない。ここでは、これまで多くの商用データベースの検索を実施した経験に基づいて、サーチャの立場から現在の商用データベースや検索システムの問題点と改善への要望を提示する*。

(1) コマンドとメニュー

コマンド方式の検索システムでは、サービスシステムごとにコマンドやその指定方法が異なるため、あらゆる関連情報を網羅的に検索しようとすると、検索手法の習得にかなりの時間と忍耐力を要する。一方、メニュー方式では、検索結果にたどり着くまでに5~10の画面を経なければならず、検索自体に時間がかかりすぎる。

これらの点を改善し、コマンドにせよメニューにせよ、利用者からみてより分かりやすく記憶しやすく、しかも短時間で目的の情報にたどり着ける検索システムの開発を望みたい。

(2) 検索言語

統制語方式のデータベースでは、個々のデータベースで採用している索引語彙をシソーラスや分類表で調べたうえでないと検索語を選定できないため、検索の事前準備にかなり手間取る。自由キーワード方式では、同じ概念や同じ対象物に異なる名称(同義語、類義語、略語など)があるため、もれのない検索は困難である。外国語のデー

タベースを検索する際には、検索語を外国語で指定するので、外国語の幅広い語彙知識が必要であり、出力結果を読み取って適否を判断する際には外国語のかんりの読解力が必要である。

データベースの検索におけるこの種の語彙知識と外国語読解力の必要性を軽減するために、利用者が自らもつ語彙の中から自国語で検索語を指定すれば、データベースで用いられている語彙に自動的に変換してくれるようなフロントエンド機能と、出力結果を翻訳する機能を付加してほしい。

(3) 端末側の入力装置

既存の商用情報検索システムは、メニュー方式であれコマンド方式であれ、キーボードから検索条件を入力するものが大部分である。キーボードの操作は、慣れた人にとってはなんでもないのである。特には、キーボードから日本語で検索条件を入力する際には、半角・全角の切り替え、改行キーと送信キーの使い分けなどが端末装置のメーカーによって異なるため、熟練者でも戸惑うことが多い。

第一段階として端末装置のキーボード標準化を、さらには、キーボード操作を必要としない端末装置の出現を期待する。

(4) 出力結果の確認と理解

検索結果の出力情報を画面上で読んで理解することは、先にも述べたように目に負担がかかりすぎるため、通常はプリンタに検索結果を出力し、検索終了後にそれに目を通す。全文データベースの検索や高再現率検索では大量の出力結果が得られることも多く、1時間で検索した出力結果に丸一日をかけて目を通すことも少なくない。今後、利用できるデータベースの種類や収録データ量が増加するに従い、出力される情報の量はさらに増大し、内容も玉石混交となることが予測される。

情報の電子化により情報を検索し入手するために要する時間は大幅に減少したが、人間が情報を読んで理解するために要する時間は依然として変わらない。

人間の読解力や読むスピードを強化したり、出力情報の品質評価を支援するような情報システムの出現を期待したいが、それに先だって、多様な形態ですでに蓄積されている膨大な量の電子化情報の中から必要な情報のみを厳選して検索できる、より高度な検索機能の開発が望まれる。

* 以下に述べる問題点と要望は、情報メディア研究会における報告¹⁾を要約したものである。

参考文献

- 1) データベース代行検索機関ディレクトリー, p. 105, 情報科学技術協会, 東京 (1989).
- 2) 社内サーチャーの育成, 情報の科学と技術, Vol. 41, No. 5, pp. 392-403 (1991).
- 3) Belkin, N. J. and Croft, W. B.: Retrieval Techniques, *Annual Review of Information Science and Technology*, Chap. 4, Vol. 22, pp. 109-145 (1987).
- 4) Using the Directory, *Directory of Online Databases*, Vol. 12, Nos. 3&4, Gale Research, Detroit, pp. vii-viii (1991).
- 5) データベース白書, データベース振興センター, 東京 (1986-1992).
- 6) 平成3年度データベース台帳総覧案内パンフレット, データベース振興センター, 東京.
- 7) データベース白書, データベース振興センター, 東京 (1986-1992), pp. 48-68.
- 8) 三輪真木子; 情報検索システムにおけるユーザーインターフェイスの条件, 情報処理学会研究報告 91-IM-3 情処研報, Vol. 91, No. 77 (1991).

(平成4年3月4日受付)



三輪真木子 (正会員)

1951年生. 1973年日本女子大学文学部史学科(東洋史専攻)卒業. 1979年米国ピッツバーグ大学大学院図書館情報学修士課程修了. 1983年慶應義塾大学大学院文学研究科図書館・情報学博士課程修了. 1981年筑波大学学術情報処理センター技官就任. 1983年(株)エポックリサーチ取締役就任. 現在同社代表取締役. 情報管理, 情報検索の研究開発に従事. 慶應義塾大学, 東洋英和女学院大学兼任講師. 著書「サーチャーの時代第2版」丸善, 「垂米利加解読新書」世界文化社(共著者一ノ渡勝彦)他. 訳書「図書館サービスの評価」丸善(共訳者中村倫子). アメリカ情報学会, 三田図書館・情報学会各会員.

