

# 情報検索システム

## ＜情報の有効活用＞

### 情報システム部

今日のような情報化社会では、情報の有効活用は企業にとって重要な課題となってきた。そのための有力な手段の一つとして、最近社内外で話題になってきた情報検索システムについて、その概要・技術動向・当社への適用状況（図書システム、文書類管理・検索システム、記事・資料検索システムなど）・問題点について説明する。

#### 1 情報検索の概要

##### (1) 情報検索の定義

情報検索 (Information Retrieval) は、本来図書館学の文献検索から発生した言葉である。広義の解釈は、ファイリング、マイクロフィルムなど各種媒体を通じて目的の情報を探すことであるが、コンピュータの発達と浸透により、最近の一般的な定義は、コンピュータにより、目的の情報を端末から即時に検索することを指している。

##### (2) データベース

情報検索の前提条件として、必要な情報がコンピュータのファイルに整理・蓄積されていなければならない。即時検索可能なこのようなファイルをデータベースという。データベースの構築を支援するソフトウェアの型を第1表に示す。

第1表 データベースの型

型	創始者	システム
階層構造 (木構造)	McGee	IMS(IBM) PDM(日立)
ネットワーク構造 (CODASYL型)	Bachmann	AIM(富士通) ADBS(日電) DMS(UNIVAC)
リレーショナル型 (表形式)	Codd	SQL(IBM) QBE(IBM)

最近、情報の価値が認識され、商用データベースサービスも増加した。通産省の58年度「データベース台帳総覧」によれば、国内の商用データベースは58社、916種ある。主要な商用データベースの種類は第2表のようになる。

##### (3) 検索方法

データベースから必要な情報を抜き出す検索方法は、当社の営業配電総合オンラインシステムのように定められた属性をキーワードとして入力・

第2表 データベース (DB) の種類

DB	特色	例
文献DB	情報源を教える テキスト型	文献 JICST DIALOG
フルテキストDB	情報内容を提供 テキスト型	記事 NEEDS-IR ACE
案内DB	案内情報提供 テキスト型	農林研究課題DB
ファクトDB	情報内容を提供 数値主体	経済 NEEDS TS 株価 QUICK
画像DB	情報内容を提供 画像・図形主体	特許 PATOLIS

選択する方法とJICSTのようにキーワードの論理検索によるものがある。

属性入力・選択方式は、データベース内に存在する必要な情報は必ず得られるが、予め定められた属性からしか検索できない不自由さがある。

論理検索は、求めたい事項をキーワードとして検索できる自由度はあるが、必ずしも求める情報だけがすべて得られない。次のような検索上の問題が発生する。

- 検索漏れ (第一種誤り)
- 不必要なものまで検索 (第二種誤り)

#### 2 技術動向

情報検索が注目されてきた背景として、情報の価値が認められてきたニーズ面と、以下に述べるような最近のコンピュータ関係技術の著しい進歩によるシーズ面がある。

- コンピュータのコストパフォーマンスの向上
- 磁気ディスクの大容量、低価格化
- 光ディスクやイメージ処理端末の出現
- 端末装置の安価、高能力化
- データベース処理技術の向上

### 3 当社への適用状況

#### (1) 図書システム（60/1から実用）

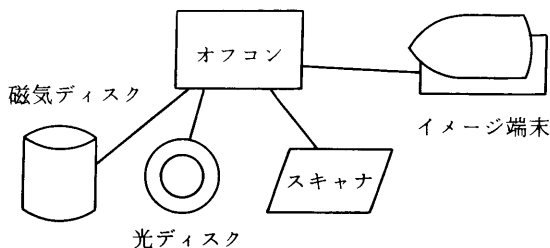
図書室には目的とする本を探すために、書名順・著者順・分類番号順・登録番号順（購入順）の4種類の図書カードが備え付けてある。しかし、内容から本を探すことはできない。

図書システムの情報検索サービスは、出版社・著者名・書名（先頭・中間の熟語指定も可能）から検索できるほか、キーワードによる論理検索も可能である。例えば、人工知能による機械翻訳の技術動向に関する本を探したい時は、下線部分の熟語をキーワードとして簡単に指定できる。

#### (2) 文書類管理・検索システム（60/2から実用）

図書システム同様に文書番号・作成年月日（以上は範囲指定）・文書種類・文書名（先頭・中間の熟語指定も可能）による検索と、キーワードによる論理検索が可能である。

図書システムと異なる点は、検索した文書などの内容（図形を含む）のイメージも出力できることである。



第1図 文書類管理・検索システムの構成

#### (3) 記事・資料検索システム（59/12から実用）

事柄・関係部門・新聞社・掲載年月日・見出し項目によって検索する。

汎用システムのため、検索項目の総文字数に制約があるので、下線部分の項目は2桁のコードとした。

検索時は、対応コードを探すことが必要となる不便さはあるが、逆に2桁のコード（数字）を入力するだけで済み、専門の人が入力を行えば簡便である。

#### (4) 社外情報検索サービスの利用

当社で現在利用している主な社外情報検索サービスは次のとおりである。

- 日経 NEEDS-IR      新聞・雑誌記事情報
- 国際金融情報センタ      カントリーリスク情報
- 日経NEEDS ECONOMY      経済情報
- JICST      科学技術文献情報
- 原子力情報センタ      原子力情報

#### 4 情報検索の問題点

日経NEEDS-IRやJICSTを利用した人は体験したことと思うが、論理検索で自分が欲する情報がすべて得られないことがある。商業用の情報検索システムでさえこの程度であり、一企業が作成したデータベースの情報検索では、以下のように種々の問題点があり、技術改善が必要となる。

- キーワードのシソーラス（同義語、上/下位語、関連語）体系の作成  
例えば、コンピュータと電算機は同義語として扱うとか、天然ガス改質器のキーワードで求まらなかったら燃料改質器で検索する。
- 簡便な入力手段  
手書き文字自動認識、音声入力装置の出現が期待される。
- 情報に全キーワードの付加  
いかに良いシソーラスを作っても、個々の情報に必要なキーワードが全部付いていなければ正しく検索されない。このためには、キーワードの自動付加機能が必要である。
- 検索時間の迅速化  
データベース、インデックスファイルなどの高度な設計が必要である。

#### 5 あとがき

我々が期待するような情報検索システムは第五世代コンピュータの出現を待たなければ実現しないが、現在のシステムでも上手に使えば十分な効果をあげることはできる。

（システム計画課）