

# データ・ベースと電子計算機

情報時代とよばれるように、多くの情報に囲まれた現代において、それらの情報を取捨選択して、効率的に活用することが必要になっている。たくさんの情報を貯えておき、必要なときに、必要な情報を容易に取出すための道具としてデータ・ベース・システムが研究され、実用化されてきている。電子計算機のハードウェア、ソフトウェア両面での発展にともなって、データ・ベースが大いに活用される気運が熟してきたように思われる。

## 1. データ・ベースの計画

データ・ベースの計画において、データ・ベース構築の要件として「データ」と「利用」の両面を検討しなければならない（図1参照）。

現在の状況として、どのようなデータをどのくらい用意するのか、またそれがどのような頻度でどのように利用されるかを把握する。それとともにデータ量の増大、利用形態の変更をふまえた将来構想をもつことが必要である。データ・ベースは一般的に大きなシステムであり、長期ビジョンをもって計画すべきものである。

データについていえば、データの質や重要度を考慮して、捨てるべきものはどんどん捨てていくという態度が必要であろう。統計的処理を行なっ

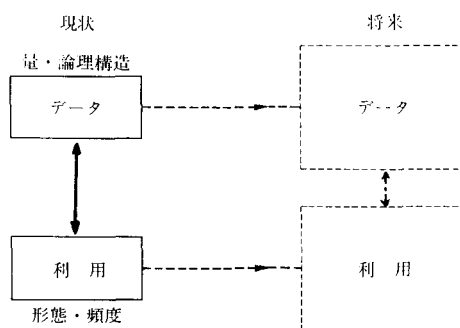


図1 データ・ベース計画の要件

て、肝要なデータだけを保存することも考えられる。データを更新するとき、更新前のデータをいつまで保存するかという原則も定めておく。時系列解析などを行なうために将来必要であるとしても、より低速でアクセスする外部記憶装置に移して保管する方法もある。

データ・ベースは多くの場合広い範囲の利用者によって活用されてこそ、その効果が発揮されるであろう。データ・ベースの利用者は図2に示すような階層によってあらわされるであろう。データ・ベース管理者はデータ・ベースの生成、利用を管理するグループである。専門利用者は、データの構造を知ったうえで、必要とあればプログラムを作成してデータ・ベースを利用できるレベルの利用者である。一般利用者は、計算機やデータ・ベースの知識を前提とすることなく、実務のためにデータ・ベースを利用する人々である。データ・ベースの計画にあたって、この一般利用者が容易に利用できるようなシステムを目標とすべきであろう。

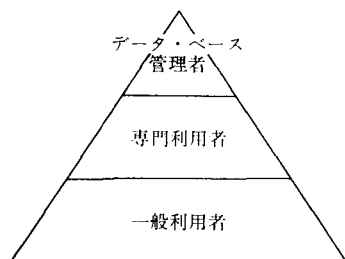


図2 データ・ベース利用者の階層

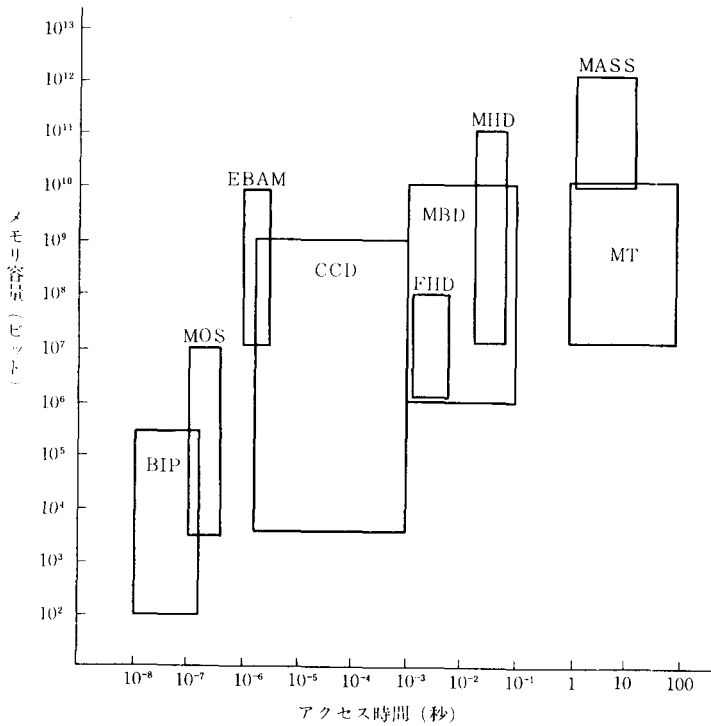


図 3 記憶装置のアクセス時間と容量

## 2. データ・ベースとハードウェア

10年ほど前にMIS(Management Information System)という言葉が流行した。経営に必要な情報が即座に得られるシステムであるが、これはまさにデータ・ベースを応用したシステムである。当時のわが国の電子計算機のハードウェアの水準

では、容易に実現できるものではなかった。しかし10年経過した現在、ハードウェアの進歩はめざましく、データ・ベース・システムが実用化される時点になったのである。

データ・ベースにおいては、記憶装置の進歩がもっとも大きなインパクトを与えているといえよう。

(1) Thesis, D. J., "An Overview of Memory Technologies", DATA-MATION, vol 24, 1 (Jan. 1978), pp 113~131.

図 3 および 図 4 に現在の記憶装置のアクセス時間、メモリ容量、ビット当りのコストを示す<sup>(1)</sup>。図中の略号はつぎの装置を意味する。

- BIP : bipolar random access memories
- MOS : MOS (metal-oxide-semiconductor) random access memories
- EBAM : electron beam accessed memories
- CCD : charge coupled devices
- MBD : magnetic bubble devices
- FHD : fixed head discs
- MHD : moving head discs
- MT : magnetic tape units
- MASS : mass strage systems

図に見られるように、記憶容量の飛躍的な増加と、それともなうビットあたりのコストの低下はいちじ

るしい。また主記憶装置と外部記憶装置(ディスクや磁気テープなど)の中間を埋める記憶装置が開発されつつある。このような記憶装置の発展は、今後も続くであろうと期待されている。

さらに、計算機本体の処理速度と主記憶装置の増大および端末機器の安価な提供により、時分割による会話型の検索が可能になっている。タイプ

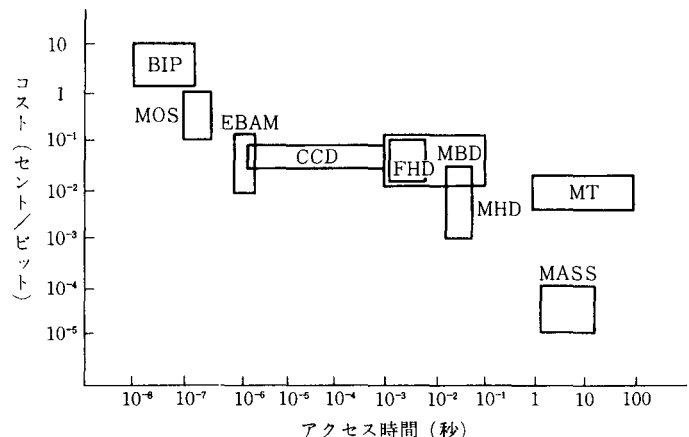


図 4 記憶装置のアクセス時間とコスト

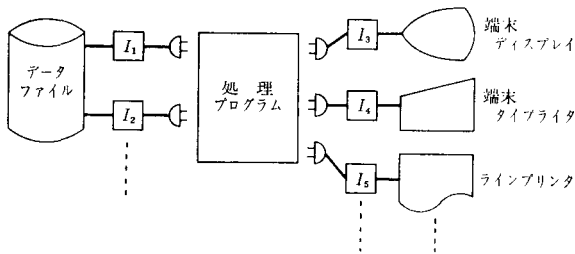


図 5 機種独立なプログラムの構造

ライタ端末のみならずディスプレイ端末の普及によって、データ・ベースの手軽な利用が可能になったことに注目したい。グラフィックディスプレイ装置を使用すれば検索されたデータをグラフで表現することができて、非常に使いやすくなる。

### 3. データ・ベースとソフトウェア

データ・ベースを設計するにあたって、データの論理的な構造をいかにして記憶装置上の物理的構造に置き換え、いかに検索するかが重点になる。米国において事務処理応用のための共通プログラム言語を開発する目的で設立された CODASYL (The Conference on Data Systems Languages) は COBOL の開発を行なった他、データ・ベース用共通言語の検討を精力的に行なっている。データ・モデルすなわちデータの論理的構造を統一的にあらわそうという試みが多く発表されている。CODASYL で研究されたモデルの他、最近大いに注目されているのが E. F. Codd の提案した関係モデルである。そのようなデータ・モデルにしたがった構造を記憶装置上に実現し、検索するためのソフトウェアシステムは DBMS (Data Base Management System) とよばれる。多くの汎用 DBMS が開発されており、電子計算機の各機種に既製のものが用意されている。DBMS はかなり大きなシステムとなるので、とくに専用的に効率のよいシステムをつくるのでなければ、なるべく既存のものを利用したほうが得策であろう。

データ・ベースは長期にわたって利用できるように設計しなければならないが、将来の機種変更

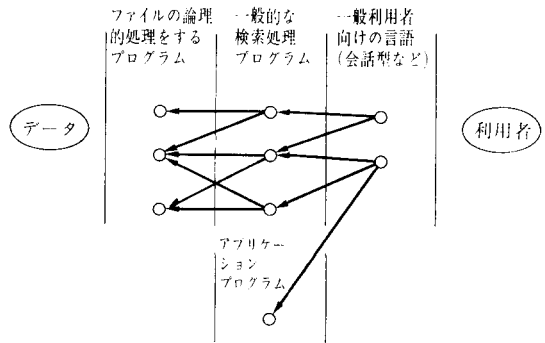


図 6 処理プログラムの構造

などにもなって、大きな変更が必要にならないよう留意しなければならない。その意味でデータもプログラムも機械に依存した部分を小さくして、機種に独立な構造にすることが望ましい。図 5 には、プログラムの構造の概念図を示す。周辺機器とは独立に処理プログラムを構成し、データ・ファイルや出力機器へはプラグで配線を接続できるような形としておく。各機器とのインターフェイスは図の中の  $I_1, I_2$  などに示した部分に、個別のプログラムを作成してゆけばよい。この  $I_1, I_2$  などのプログラムをできるだけ小さくできれば、変換が楽になる。

図 6 には、処理プログラムの構造の概念図を示す。ファイルの論理的処理、一般的な検索処理までは汎用 DBMS によって実現されるであろうが、さらに一般利用者向けに使いやすい言語を作成することが望ましいと思われる。たとえば、端末機器を用いて、データ・ベースの使用法を学びながら使えるような仕組みが考えられる。また、そのデータ・ベースの目的に応じたアプリケーションプログラム (統計処理) などを組み込むと使

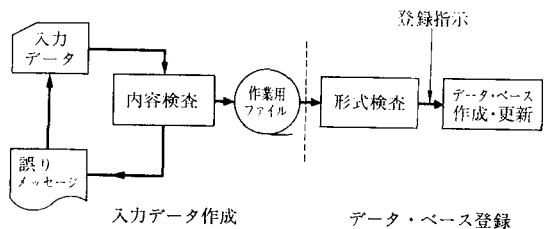


図 7 データ・ベース作成・更新の手続き

いやしくなるであろう。

#### 4. データ・ベースの管理

データ・ベースの円滑かつ安全な運用のためにデータ・ベース管理者のグループを組織する必要がある。その主な業務はつぎのとおりである。

##### (1) データ・ベースの生成と更新

データ・ベースの生成・更新の手続きを管理する。データ定義はデータ・ベース管理者の監督のもとに行ない、データの生成・更新の許可を与える。データ・ベースの生成・更新において、不完全なデータを登録することは混乱のもとになるので慎重に行なわなければならない。図7に示すような手続きが考えられる。データそのものの入力と整備は、データ収集側の人によって、その内容の検査を行なったうえで作業用ファイルへ入れておく。適切な時期に、データ・ベース管理者がそれをデータ・ベース本体へ登録する。データ更新のスケジュールをきちんと立てないと、データ間の整合性が保てなくなって、利用者に迷惑をかけることになる。

##### (2) データ・ベースの機密保護と保全

データ・ベースは「誰にでも容易に使える」と同時に、「勝手に使わせない」機能が必要である。機密データを特定の利用者のみが利用できるようにすることと、データ・ベースを破壊するような使い方をされないための用心が必要になる。データ・ベース管理者はデータに錠（ロック）をかけておき、利用者はそのレベルに応じた鍵（キー）を渡されて、許されるデータだけを利用することになる。キーは暗証の形で各利用者に与えられるであろう。また、事故によってデータ・ベースが破壊されたときの復旧対策も必要である。

##### (3) 利用状況の把握と記憶構造の変更

利用状況を把握するため、利用記録をシステムが自動的に取る機能をもつことが望ましい。利用効率の向上と記憶装置の効率的な利用を目的として、データの記憶構造の変更が可能になるよう設

計すべきであろう。

#### 5. データ・ベースのコストと効率

データ・ベースのコストはつぎのような項目からなる。

##### (1) 初期費用

データ・ベース設計費  
ソフトウェア開発費  
入力データ整備費  
その他設備費

##### (2) 経常費用

更新用入力データ整備費  
記憶装置専有費  
データ・ベース維持費  
データ・ベース利用の直接費

普通のソフトウェアの場合は、開発費を別にすれば計算機使用の直接費が大きい。しかしデータ・ベースでは、入力データの整備費、データ・ベース維持費が大きくなるであろう。また、記憶装置が安くなっても、その専有費用を無視することはできない。利用時の検索応答時間を短くするためには、高速アクセスの記憶装置のコストの増加、データ・ベース更新の際の内部構造の変更に要するコストの増加が見込まれる。コストと効率とのバランスをいかにするかが、大きな問題となるであろう。緊急度や利用頻度に応じて、検索効率に段階を設けることが考えられる。記憶装置を階層的に準備してデータを割り当てることが考えられよう。たとえば第1次外部記憶として可動ヘッドのディスク、第2次記憶装置としてマストレージを利用し、さらに使用頻度の低いデータは磁気テープに格納するといった方策が検討される。コストと効率のトレードオフは恰好なORの問題であろう。OR的な見地からデータ・ベースがもっと検討されてよいのではなかろうか。

つねかわ・じゅんきち 1932年生  
(株)日本科学技術研修所