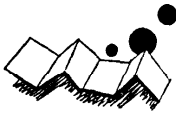


解説

データベース技術応用の過去・現在・未来†



酒井 博 敬††

1. ま え が き

計算機技術の発展のようすを示すのに、Nolan のステージ論がよく用いられる。これは発展過程に揺籃、成長、統制、および熟成があるとし、その間の投資額の推移を図示したものである¹⁾。熟成期に次世代の揺籃期が始まる。図-1 はこれにならってデータベース技術の発展を図式化してみたものである。

第1世代は階層、ネットワーク、関係の各モデルを中心とする DBMS (database management system) とその応用が発展した時期であり、第2世代はより広範な種類のデータを扱う知的情報システムが成長する時期である。1981年の VLDB 会議での話題や、設計・診断・研究開発・意思決定への応用を志向したアドバンスト・データベース・システム・シンポジウム (情報処理学会) での発表は、第2世代の萌芽を伺わせる。80年代前半はデータベース技術とその応用の質的転換期と考えられる。

2. データベース移行問題とソフトウェア工学

1977年、データベース移行問題を主題にした Data

Base Direction 会議の基調講演で、Ford 社の Roark は次のように述べている²⁾。

データベース化によって得られた利点として、同社の管理者は、要求変化への即応、新プログラム開発における10ないし20%の生産性向上、および問題分析能力の取得をあげている。とくに意思決定のための高度な問題分析においては、更新周期と制御手続きの異なる多ファイルからのデータ抽出を必要とする。ここでデータベースの存在は決定的な役割を果たす。また彼の試算では、アメリカの全企業が保有するプログラム中データベース化されたものは高々20%である。データベース技術の導入が予想以上に遅れている最大の障壁は、過去に投資したプログラムとデータの累積である。この資産を反古にすることなくデータベース化することは重大な問題である。

同会議の管理問題を取りあげたパネル・セッションでは、データベース化にともなう諸問題は技術的な面よりも人的、管理的な面にあることが強く指摘されている。データベース化は、データベース設計 (データ定義の一元化)、データの入力・保全 (データ供給者は必ずしも使用者とはかぎらない)、プログラム開発

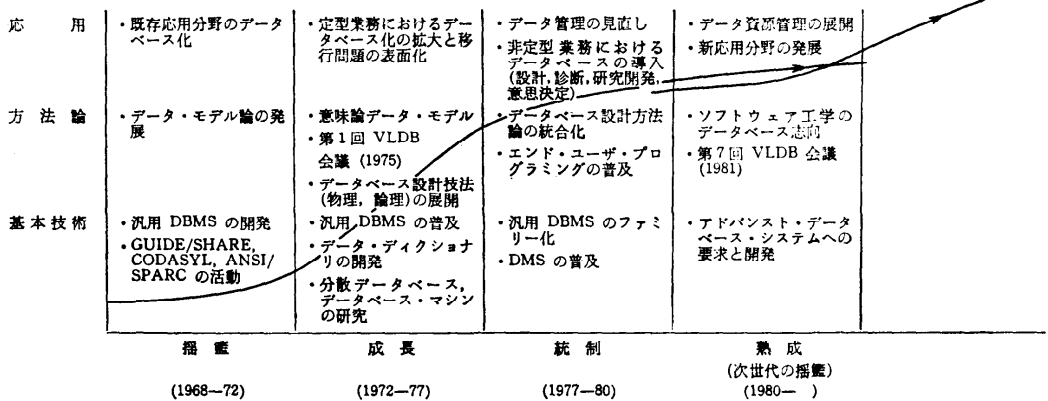


図-1 データベース技術の発展

† Applications of Database Technology: Past, Present, and Future by Hirotaka SAKAI (Hitachi Institute of Technology, Hitachi, Ltd.).
†† (株)日立製作所日立技術研修所

法(要求変化へ即応する繰返し型へ)、操作方式(オンライン化とこれにともなう組織の再編成)などにおいて変化をもたらす。これは開発者、運用者、データ供給者、およびユーザ間の責任転換を要請する。したがって上級管理者の問題認識と参画がなければデータベース化は実現しない。この意味で第1世代は管理者に多くのインパクトを与えた時期であった。

技術的な面では実務、理論両面で多くの技法とツールが開発された。少なくとも何らかの設計技法が現実に応用され、またデータベース設計もビジネスとして成り立っている。文書化を含む設計方法論の導入は、設計手順を標準化し、また設計、モデリング、モニタリングなどの諸ツールの利用指針になるという意味で重視されるようになった。

データベース工学は管理的な面でも技術的な面でもソフトウェア工学としてとらえられるべきである。たとえば多くの実務家と理論家が参加した1979年のER(Entity-Relationship)会議は、ERモデルを利用した情報のモデリングと分析を主題としている⁸⁾。ここでは要求分析と定義、情報モデリングとメタ・データ設計、データ分析とデータベース設計技法、多モデルDBMS、設計ツール、応用事例など、ソフトウェア工学におけるER概念の導入が議論された。

データベース・システムにおいて展開された方法論や理論は、単にDBMSによって管理されるデータだけではなく、より広い範囲のデータへ向けられる必要がある。Gotliebは1980年のVLDB会議で、Census, UTLAS, PRESTEL, PLATO, MEDLARS, ORBIT, SITAなどの大データベースが、データベース技術の対象としてとりあげられていないという問題提起をしている⁹⁾。これらのデータベースはなお成長を続けており、その運用の中でデータベースがいかに作られ、また使われるかという現実的な問題が提起される。Gotliebは、このような実在システムにおけるデータ・コミュニケーション、概念モデル、性能評価、データ保全などの問題に、データベース技術者および研究者はもっと目を向けるべきであると強調している。

3. DBMS および DMS 製品の普及とその影響

データベースの歴史はDBMSの普及とデータベース概念形成の歴史であったといえよう。1960年代後半から出現した汎用DBMSは、GUIDE/SHARE要望、CODASYL DBTG報告、関係モデル論、ANSI/

SPARC報告などによる概念形成に支えられて普及した。70年代前半はDBMSの優劣論、データ・モデル論、DBMS導入指針などの話題で賑わった。この時期はユーザにとってどのDBMSを使用すべきかということより、ベンダにとってどのようなDBMSを提供すべきかの議論が先行した¹¹⁾。

データの独立性、データの共用、データの完全性といったデータベース本来の目標を応用分野において実現することは容易でない。DBMSを導入しても複数の応用プログラムが一つのデータベースを共用することは稀であり、また新しい応用の追加に際してデータベースが別個に作られることが多い。しかし一方ではDBMSを単なるアクセス法として受けとめ、問合わせ言語、報告書作成、レコードの分類・選択のような周辺機能を利用してプログラム開発の生産性向上に役立てているユーザも多い。1981年にData Decisions社が行った調査では、ユーザ満足度7点以上(10点満点)のDBMSあるいはファイル管理システムとしてあげられているのは次の製品である⁸⁾。

QUICKJOB (System Support Software), SAS (SAS Institute), DMSII (Burroughs), IDMS (Cullinane), IMAGE (Hewlett-Packard), DYI 250 (DY-LAKOR), EASYTRIEVE (Pansophic), DM-IV (Honeywell), ADABAS (Software AG), DMS (Univac), INQUIRE (Infodata), DATACOM/DB (Applied Data Research)。

現在DBMSベンダは50社に達し、DBMS市場はブームを迎えている。Frost & Sullivan社は、DBMSソフトウェア製品の売上は年率25ないし35%で伸びており、80年代末までに4,000億円市場になると予測している。

DBMSあるいはファイル管理システムに、非手順言語、対話型問合わせ言語、入力フォーマット設計、報告書作成、CRTスクリーン・フォーマット定義、図形処理、コンピュータ・レター生成、統計・財務分析プログラム、データ・ディクショナリなどの機能を組みこんだものは一般にDMS(data management system)とよばれ、エンド・ユーザを含む幅広い層に普及し始めている。DMS製品には、RAMIS II (Mathematic Products Group), FOCUS (Information Builders), NOMAD (National CSS), INFO (Henco), USER-11 (North County Computer Services), CREATE (Complete Computer Systems)などのほか、Four-Phase, HP, Microdata, Prime, Wangな

どのミニ・コンピュータ DMS がある。

DMS 製品はいわゆるエンド・ユーザ・プログラミングの普及を促している³⁾。この状況が出現した背景には、データ処理部門とユーザ部門間のプログラム需給均衡の破綻、ソフトウェア人口不足の深刻化、およびデータ処理の 50% はデータベース応用向きであるという事実がある。エンド・ユーザは人事、営業、経理、資材などの部門管理者、アナリスト、エンジニアなどデータ処理部門以外の人々を指す。DMS によるエンド・ユーザ・プログラミングは、簡単な報告書作成から、複雑な応用システム開発にわたっている。エンド・ユーザがその業務環境をより豊かで、より人間的なものに再設計するために、計算機の活用を探究し始めたということができる。DMS はワード・プロセッサ、電子メール、電子カレンダーなどとともに、オフィス・システムの主要成分となっている。

DMS の普及を促す要因の一つにミニ・コンピュータ向き DMS の出現があげられる。データベースが実用的意味をもつのは、ファイル容量 50 M バイト、主記憶容量 128 K バイト以上とされているが、ミニ・コンピュータはこの条件に合う。強力な DMS はサービス性がよく、安価でもある。エンド・ユーザ・プログラミング志向とあいまって、部門コンピュータの設置とミニ DMS の利用は急増する気配をみせている。その機器構成には、①独立型（部門専用コンピュータ）、②ネットワーク型、および③後置型がある。現在①が多いがその延長上に②がある。②の環境でローカル・ユーザの DMS 利用が増えると分散データベース・システムに発展する。技術的にもデータの所在ノードをブラック・ボックス化する分散問合わせ処理、ネットワーク内データの複数コピーの内容一致を制御する同期化処理が可能な時期になっている。③は同じデータベースを異種メインフレーム・コンピュータで共用できること、また容量の限界もミニ・コンピュータの追加によって安価に解消できることから普及し始めている。

DMS の利用はプログラム開発方法論にも影響を与えつつある。ソフトウェア・プロトタイプング (SP: software prototyping) とよばれるものである⁴⁾。これはプログラム開発においてはじめから最終製品を作るのではなく、プロトタイプを開発し、これに改良を加えながら最終目標に接近する方法である。これは次の性質をもつ。すなわち、① SP は短期間に安価に開発されなければならない。このため種々の型のツール

が使われる。DMS はもっとも強力なツールである。

② SP は単なる机上のアイデアではなく、ワーキング・システムとしてそれ自体実用に供される。③ SP はユーザ要求、システム設計構造、データおよび制御構造などに関する仮定を検証する目的をもった繰返し過程である。その使用を通してエンド・ユーザとシステム設計者は新しい要求と改良点を発見し、これを次の SP 版に反映させる。

エンド・ユーザ・プログラミング、ミニ・コンピュータ向き DMS、ソフトウェア・プロトタイプングなどの普及によって広範なユーザ層がデータ資源の価値を認識し、データベース化の必要性を理解する環境が形成されつつある。

4. データ管理者とデータ分析

データベース化の障壁は技術的問題ではなく管理的問題にある。DBA (database administrator) 100 人中 70 人が、上級管理者の支援不足、ユーザの非協力、組織間調整の不備といった政策的、管理的な問題をあげている。エンド・ユーザ・プログラミングやソフトウェア・プロトタイプングは、一方でプログラム生産性の向上、データベース技術の局所的修得といった利点をもたらす反面、他方でより高い視点からの政策を誤るとデータ共用に関する管理不在を助長する危険をはらんでいる。

企業が投資してきたデータ処理の目的はよいデータからよい情報を得ることにあつたはずである。悪いデータからは悪い情報しか得られず、それは誤った決定をもたらす、企業の収益性と生産性を低下させる。データは資金、人、資材と同様、特定部門に私有されるべきものではなく、共有資源として管理され、企業目標に沿って活用されるべきである。この認識は 70 年代に DBMS の普及とともに顕在化してきた。全社視点から文書データを含めた企業データを管理する機能をもつ DA (data administrator) を DBA から分離して設置することの必要性が叫ばれ、データ処理部門にも情報管理部あるいは情報システム部などの名が付けられ始めた。DBA が計算機データ・システムの実現に関与するのに対し、DA は企業の全データの性質と用途を把握し、データの収集と利用の仕方を設定し、これをモニタし、変更する責任をもつ。このような DA 機能の必要性は多くの場所で論じられてきたが、データ価値の評価基準がなく、DA 機能を実現する方法論も確立されていないのが現状である。

DA が企業データを分析するために、概念モデル設計は一つの効果的な手段を提供する^{9),7)}。これは企業が必要とするデータ対象を識別し、その基本性質と用途を明らかにし、概念モデルとして表現することである。概念モデルは、実体と実体間の関係のようなデータの静的性質の記述と、データに対して施される操作の特性のような動的性質の記述からなる。さらに企業活動や企業環境の成長、変化に対応して、データの静的および動的性質そのものを変更するための理念、すなわち企業方針、規準、制約などに関する記述も概念モデルの表現に含まれる。最近 ER モデル、TAXIS などの意味表現能力にすぐれた概念モデル向きデータ・モデルが開発され、実用されるようになってきた。これらのデータ・モデルは、単にデータベース設計の手段としての局所的な適用にとどまらず、広義のデータ分析の方法論として発展させるべきであろう。また DA 機能を遂行する上級管理者にとっても、このようなモデリング技法に関する訓練は有効かつ必要である。

5. アドバンスド・データベース・システム

データベースへの移行が問題となっている一方で、従来データベース技術の直接の対象の外にあった応用分野が多く出現している。CAD/CAM (computer aided design and manufacturing)、診断、研究、開発、意思決定などの分野である。従来のデータベース・システムでは、情報構造表現とデータ利用形態はきびしく限定されていた。新しい応用分野はこの制限を越える多くの要素を含んでいる。データ対象として図形、画像、テキストなどを含み、また実体間の関係を一般的な形で表現する知識を扱う。利用形態は非定型的かつ創造的である。この分野を仮にアドバンスド・データベース・システム (ADBS: advanced database system) とよぶ¹⁰⁾。

意思決定への応用の一環として、最近 DMS とともに計画、予測、統計解析などのプログラムを利用する管理者が増えている。これは実用的 DSS (decision support system) とでもいうべき分野で、FCS-EPS (財務計画)、RAMIS II (DMS)、SAS (統計解析)、TELL-A-GRAF (グラフィックス) といった組み合わせが用いられている⁹⁾。DSS は既存ファイルから情報要求に合致するデータを検索、集計、リストする報告型と、比較的簡単なプログラムを用いる解析型に分れる。いずれも現実問題への適用において効果的であることが経験されており、その普及とともにより高度な

DSS へ発展することが予想される。

CAD/CAM 向きデータベース、いわゆるエンジニアリング・データベースの分野では問題がより顕在化している。CAD/CAM パッケージあるいは対話型グラフィックス・システムの普及によって、図形データ、NC データ、仕様書など多くのデータが累積した。これらの設計、製造データの利用は局所的機能の自動化に止まっており、全社的な設計、製造工程の流れの中では共用されていない。生産性改善のためには CAD/CAM 機能の統合を必要とするが、ここでもデータベース移行問題が表面化している。この状況の中で、デジタル表現の図面および図形データの応用プログラム間流通手順として、これらのデータの共通形式を規定する IGES (initial graphics exchange standard) が注目されている。

エンジニアリング・データベースはさらに次のようなより本質的な問題をもっている。

(1) 技術革新、法的制約などの要因によって、設計、製造方法が絶えず変化する。これは動的なデータ構造の変化、応用プログラムの変更を要請する。

(2) 図形情報の扱いが重要な部分を占める。図形情報の操作は複雑であり、単純な形状変化が多く情報の変化をひき起す。このため意味的矛盾が生じないように図形モデルに対する完全性制御が必要である。

(3) データ要素間の関連性が強い。たとえば各種規格、標準データ、設計技術情報、製造情報、検査情報、図形情報は互いに関係をもつ。このため高度な関連性維持機能が必要である。

(4) 設計過程は情報の質的および量的変化をとともなう。たとえば企画設計、基本設計、詳細設計、実施設計などの過程において、情報量が増加するだけでなく、数値情報の精度が上がり、構造の複雑さが増す。とくに図形情報の変化は著しい。したがってデータ構造の動的変化を扱うモデリングと制御が要請される。

(5) 設計は試行錯誤による発見的過程である。計算結果は決定値ではなく、設計の進行過程で変更される。すべての変更は検討のための試行、すなわち仮定に過ぎない場合が多い。このため高度な対話的操作が求められる。

(6) 応用プログラムは設計、製造技術のノウハウであり、データと同様の価値をもつ資源として管理されるべきものである。それは目的とする実体を発見あるいは構成する手続きとして共用される必要がある。

ここにあげた特性は ADBS に共通したものであ

る。ADBS に対するこのような要請はそのまま次世代の情報システムへの要請として受けとめられている。また研究面でも、データベース工学と知識工学の融合、実体の動的モデリングやアクティブ・データベースの開発などが注目されている。ADBS を含む現実の応用分野からの多様な要請は80年代のデータベース・システムの方向を支配するものである。

6 む す び

データベース技術応用は、データ資源の管理統合領域の拡大と高度利用志向、システム開発へのエンド・ユーザの参画、および ADBS の発展にみられるように大きく変貌しつつある。この動向は企業におけるデータ資源管理に対する考え方と体制の見直しを要請し、また制御構造を中心に展開されてきたソフトウェア工学のデータベース概念による見直しを促している。さらにデータベース工学と知識工学の融合、図形、画像、テキストなどのより広い範囲の対象を含み、動的に変化する実世界を忠実に表現し、制御するモデリング技法の開発を要請している。80年代はデータベース技術とその応用の相互作用が活性化し、これによって情報システムが影響を受け、発展する時期である。

参 考 文 献

- 1) Berg, J. L. ed.: Data Base Directions II: The Conversion Problem, Data Base & SIGMOD Record, Vol. 12, No. 3 (1982).
- 2) Bracchi, G., Furtado, A. and Pelagatti, G.: Constraint Specification in Evolutionary Date Base Design, Proc. IFIP TC-8 Working Conf. on Formal Models and Practical Tools for Information Systems Design, pp. 149-165 (1979).
- 3) Canning, R. G.: 'Programming' by End Users, EDP ANALYZER Vol. 19, No. 5 (1981).
- 4) Canning, R. G.: Developing Systems by Prototyping, EDP ANALYZER Vol. 19, No. 9 (1981).
- 5) Canning, R. G.: Interesting Decision Support Systems, EDP ANALYZER Vol. 20, No. 3 (1982).
- 6) Chen, P. P. ed.: Proc. Int. Conf. on Entity-Relationship Approach to Systems Analysis and Design (1979).
- 7) Davenport, R. A.: Data Administration-The Need for a New Function, Proc. Information Processing 80, pp. 505-510 (1980).
- 8) Data Decisions: Systems Software Survey, Datamation Vol. 27, No. 13, pp. 108-144 (1981).
- 9) Gottlieb, C. G.: Some Large Questions about Very Large Databases, Proc. 6th Int. Conf. on VLDB, pp. 3-7 (1980).
- 10) 情報処理学会編: アドバンスト・データベース・システム—設計・診断・研究開発・意思決定のツールとして—シンポジウム論文集 (1981).
- 11) 椿 正明: データベース設計と運用の方向, Computer Report 1981/6-1982/1.
- 12) Nolan, R. L.: Restructuring the Data Processing Organization for Data Resource Management, Proc. Information Processing 77, pp. 261-265 (1977).

(昭和 57 年 6 月 1 日受付)