原本

63－E002

高度情報処理技術者育成指針

―データベース・エンジニア編―

平成元年3月

CAIT

財団法人　日本情報処理開発協会
中央情報教育研究所

この報告書は，日本自転車振興会から競輪収益の一部である機械工業振興資金の補助を受けて，昭和63年度に実施した「高度情報処理技術者育成指針の作成」の成果をとりまとめたものです。

015032

15032

C
6
8

高度情報処理技術者育成指針

—データベース・エンジニア編—

CAIT
財団法人 日本情報処理開発協会中央情報教育研究所
蔵 書 之 印

刊 行 の 言 葉

情報化の急速な進展とコンピュータ・システムの一層の普及拡大に伴い，今日，情報処理技術者の不足は社会的問題化しつつあります。

一方，コンピュータ技術の進歩，情報処理と通信の融合，社会における情報処理システムへのニーズの多様化やその依存度の増大によって，情報処理技術者問題は量に加えて質的要求がより深刻となってきました。かつて，これほど専門技術者の量と質の相乗積を求めた工業技術は，他に存在しませんでした。

中央情報教育研究所（旧　情報処理研修センター）はこのような環境のもとで，通商産業省産業構造審議会の2000年に向けての情報化人材対策にかかわる審議結果を受けて，1987（昭和62）年6月に新組織として再出発しました。

本育成指針は新組織の重点的事業成果のひとつであり，今後要求される各種の高度情報処理技術者の役割と人材像を明確にし，それぞれの専門技術者ごとに必要な知識・技術を体系化するとともに，それらの修得方法についてのガイドラインを併せて明示したものであります。

専門技術者の種類は，産業構造審議会の提案する分類案を参考にして，アプリケーション・エンジニア，プロダクション・エンジニア，ネットワーク・エンジニア，データベースエンジニア等をはじめとして数種類にわたっており，それぞれ育成指針を作成しています。

従来，高度教育の内容は入門教育に比べ，必ずしも十分な整備が行われておらず，その体系化へのニーズは高かったわけですが，本指針はその目的を十分達成し得るものと確信いたします。

本指針が関係各方面に広く利用され，わが国の情報処理技術者の育成に寄与できることを心から念願するものであります。

終りに，本指針の作成に当って容易ならぬご尽力，ご協力を賜った各高度情報処理技術者育成指針作成委員会の委員各位に，心から感謝の意を表します。

財団法人　日本情報処理開発協会
中 央 情 報 教 育 研 究 所

高度情報処理技術者育成指針の

作成について

高度情報処理技術者育成指針の作成について

1．指針作成の背景

わが国の情報化は，産業界にとどまらず社会や家庭にまで浸透し，ソフトウェアの需要はますます増大してきている。また，開発すべきソフトウェアも巨大化・高度化し，広範囲になりつつある。企業におけるＥＤＰ部門も従来のシステム開発にとどまらず，情報のトータル管理，情報アドバイザとしての部門間調整，戦略的情報の提供などの役割を担うようになってきた。情報サービス産業も，マシンタイムや労働力の提供から脱し，高度な技術を売りものにする高付加価値産業といった真の知的集約産業として生きていくことが必要になってきている。

こうした背景の中で，ソフトウェア技術者の不足が重大な問題になっている。とくに高度な技術者，いわゆるシステム・エンジニア（ＳＥ）の不足は深刻な問題になっている。このまま放置すれば，高度情報化社会に向けて進みつつあるわが国にとって，きわめて重大な問題に発展することが予測される。通商産業省では，この問題にいかに対処すべきかを検討するために1986年３月17日，産業構造審議会情報産業部会に情報化人材対策小委員会を設置した。審議会では約１年間の審議を通して，1987年４月17日に問題点の指摘とともに，当面講ずべきいくつかの提言（以下「提言」という）を行った。

「提言」では，高度情報処理技術者（ＳＥ）の不足が今後深刻化することを指摘するとともに，その育成方向を明らかにするため，技術者の職務内容，知識・技術に着目して，限定的なものとしながらも一応の技術者分類を試みている。それによれば，まず，ソフトウェア技術者をプログラマとエンジニア（いわゆるＳＥ）に２分している。そして，エンジニアを①ユーザサイドに立った（利用部門に即した）業務を行うアプリケーション・エンジニア（ＡＥ），②利用部門とメーカの橋渡しを行うテクニカル・エンジニア（ＴＥ），③供給側のデベロップメント・エンジニア（ＤＥ），④ソフトウェア設計のトップ技術者であるプロダクション・エンジニア（ＰＥ）に分類し，次のように説明している。たとえば，ＤＥはＤＢＭＳ，開発支援ツールなどの基本ソフトウェアを作り，ＴＥはそれらの基本ソフトウェアを導入，管理

し，ＡＥはそれらを用いてアプリケーション・システムを開発する。ＰＥはそれぞれの段階において，プログラム開発のプロジェクト・リーダとしてプログラムの設計，開発に当る一方，プログラマはＰＥの指示により，プログラム設計書に基づいてプログラムを作成するといった役割分担を行う。また「提言」では，これら技術者に要求される技術・知識・需要予測等を含めて，次の表を与えている。

技術者の分類，要求される技術・知識，需要予測等

分類		職務内容等	要求される技術・知識	需給予測	主な供給源	備考
エンジニア	アプリケーション・エンジニア	業務分野に最適の情報システムを構築 要求仕様の定義，システム化の企画，実現可能性の検討，システム設計と仕様の決定，システム設計と評価，システムの運用・管理 ユーザサイドに立った業務	運用業務知識 要求仕様定義技術 システム分析・設計手法 システム評価技術 システムテスト技術 プロジェクトマネージメント能力 経営管理，経営工学の知識	新規分野の広がりにより，増加	大学の一般学科 専修学校等 エンドユーザからの転換 他の職種から就業構造転換 プログラマから	（主な所属） ユーザ ソフトウェアハウス
	テクニカル・エンジニア	ハードウェア・ソフトウェア両面での最適システムの構築運用，管理，評価 コンサルテーション 情報のＤＢ化と一元管理 ユーザとメーカの橋渡し	特定のシステム資源（ハードウェア・ソフトウェア・データベース・ネットワーク等）に関する専門知識 コンピュータのハードウェア／ソフトウェア技術，システム評価技術データベース技術，情報通信技術 複合システム化技術	急速するが，全体に占める割合は小	大学の情報関係専門学科及び関連学科 短大・高専の情報関係専門学科 専修学校等	（主な所属） システムハウス ソフトウェアハウス コンサルタント会社 ユーザ
	デベロップメント・エンジニア	特定分野の最適製品やソフトウェアを開発 汎用ソフトウェア（ＯＳ，ＤＢＭＳ，パッケージ，言語プロセッサ，開発支援ツール等），マイコン組込み製品ソフト等の開発，新技術の開発 供給サイドに立った業務	特定分野の製品開発に関する専門知識 ソフトウェア工学の専門知識 コンピュータ科学の専門知識 システム工学の専門知識 製品の開発管理技術 ハードウェアの専門知識	かなり増加	大学の情報関係専門学科及び関連学科 プログラマから	製品開発のエキスパート （主な所属） メーカ ソフトウェアハウス システムハウス
	プロダクション・エンジニア	ソフトウェア設計のコンサルテーション プログラム開発のプロジェクト・リーダ プログラムの設計・開発 工程・品質・原価管理 結合，検査 保守	ソフトウェア工学の専門知識 プログラムの設計技術 工程・品質管理技術 プログラムテスト技法 生産性向上技術，ツールの活用技術 プログラム言語の専門知識 コンピュータ科学の知識 ハードウェアの知識	増加	大学の情報関係専門学科及び関連学科 専修学校等 プログラマから	（主な所属） ソフトウェアハウス メーカ システムハウス ユーザ
プログラマ		プログラム設計書に基づいてプログラム作成 プログラムモジュール論理設計・製造 モジュール・テスト，結合テスト プログラムモジュール保守改善	ソフトウェア工学の専門知識 プログラムの設計技術 プログラミング技法 プログラムのテスト技法 プログラムの保守技術 プログラム言語の知識	技術進歩による生産性向上により，増加率は押えられ，相対に減	専修学校等 高校，短大，大学の一般学科 他の職種からの転換	キャリアパスの入口

さらに，「提言」では次のように述べている。「ソフトウェア技術者の育成については，その教育方法が確立されていない，良質な教材がないといった問題点が指摘されている。このため，産業界の真に必要とするソフトウェア技術者はいかなる技術者かといった調査を行い，それに応じたソフトウェア技術者教育の方法を研究開発し，プログラマ，アプリケーション・エンジニア，テクニカル・エンジニア，プロダクション・エンジニア育成用などの各種カリキュラムを設定するための産学共同の情報処理教育推進の中核機関を設置することが必要である。また，ＣＡＩ用コースウェア，ビデオ教材等の良質な情報処理教育用教材の研究開発を行うための機関も必要とされると同時に，教材の研究開発に当っては産業界と連携してそのニーズを的確に吸収するとともに，技術者育成のノウハウを有する各種教育機関との連携が望まれる。ソフトウェア技術者教育，なかんずくＳＥ教育の方法は未確立であり，教育方法の研究開発が必要である。また，技術者教育を行うインストラクタが不足しており，その質も低いので，情報処理教育機関や企業内のインストラクタの研修を公的機関により行う必要がある。」さらに，プログラマからＳＥへの移行教育などＳＥの教育方法が確立しておらず，ＳＥ教育の研修機関がほとんどない点も指摘している。そして，情報処理教育推進の中核機関として「情報大学校構想の推進」が提言されている。

先の「提言」に基づき，1987(昭和62) 年６月１日，この中核機関として財団法人　日本情報処理開発協会　中央情報教育研究所が発足した。当研究所の前身は情報処理研修センターであり，1970年に設立された。以来ＳＥ教育を中心に，企業人の研修機関として毎年約 2,000名，これまでに約26,000名に研修を実施してきた実績を持つ。また，71年に初級情報処理技術者育成指針，74年に中級情報処理技術者育成指針を，そして79年には上級情報処理技術者育成指針を作成した。また最近では，86年に新しい初級情報処理技術者育成指針およびそのテキストを作成している。

当研究所は，その発足と同時に，「提言」の指針ならびに社会的要請に沿って，産業界の必要とする技術者像を明らかにし，情報処理およびその関連の知識・技術の教育手法を研究開発し，カリキュラムを作成，良質な教材を作成することになった。昭和62年度以降，技術者育成の緊急性と継続性を考

えて，次の委員会を発足させ，上記の研究に着手した。

アプリケーション・エンジニア（ＡＥ）委員会

ネットワーク・エンジニア（ＮＥ）委員会

プロダクション・エンジニア（ＰＥ）委員会

システム監査エンジニア委員会

ナレッジ・エンジニア（ＫＥ）委員会

システム・エンジニア共通育成指針委員会

データベース・エンジニア（ＤＢＥ）委員会

マイクロコンピュータ応用システム開発技術者（ＭＣ）委員会

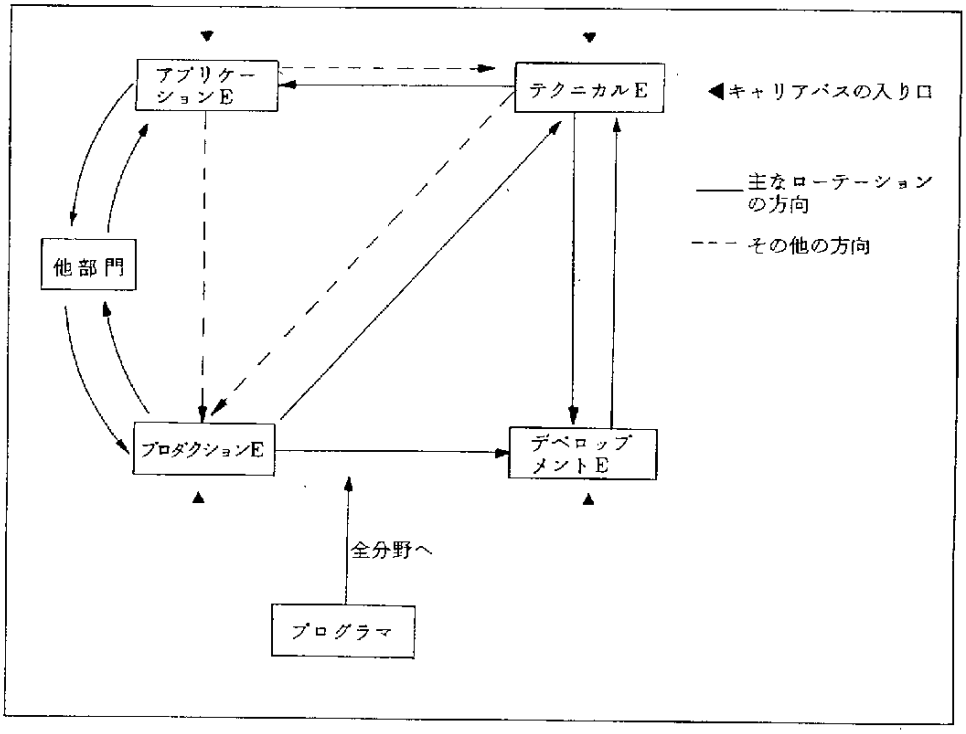
２．技術者の育成とキャリア・パス

高度情報処理技術者は短期間で育成できるものではない。そのためには企業に就職してからの不断の研修が必要である。ことに，情報処理の分野は技術進歩が急速であり，技術者は不断に技術のリフレッシュをしていかなければならない。この問題について，「提言」では以下のことを述べている。

まず，労使協調によって企業としての人材育成計画を策定し，キャリア・パスを確立し，キャリア・レベルごとの要求技術・知識等を明確にして，それに応じた教育訓練体制を確立することが必要である。企業内教育ではＯＪＴが不可欠であることはいうまでもないが，それだけでは不十分であり，集合研修，社外研修等によるＯff－ＪＴを適当に組み合わせることが重要である。さらに，実際の人材育成に当っては，個人別に育成計画を立て，ＯＪＴとＯff－ＪＴとの繰り返しにより教育訓練を実施，訓練結果の評価・フィードバックを行うことが望まれる。これからは，ＥＤＰ部門の要員にはコンピュータとシステム化に関する知識はもちろんのこと，経営全般の知識，幅広い業務知識が求められる。そのためには，ＥＤＰ部門と他部門との活発なジョブ・ローテーションが必要である。

「提言」によれば，次のようなジョブ・ローテーションの例を与えている。

ジョブ・ローテーション（例）



3．各技術者像の育成の背景

各委員会は1年近くにわたる討議を通じて，それぞれの技術者像を明らかにするとともに，育成の指針，カリキュラムを作成し，研修コースを実施する段階にまで到達した。群論は各論に譲ることとして，ここでは簡単に各技術者像とその育成の背景を要約することにする。

(1) アプリケーション・エンジニア（AE）

AE委員会では，AEについて次のように規定している。AEとは，業務部門の立場で対象業務分野における最適システムを構築するエンジニアである。そして，AEに課せられた役割と業務については，その業務に内在する問題（課題）を発掘し，把握整理するところからスタートして，最適な業務システムを開発，設計することである。AEに要求される能力をひと言で言うならば，「対象とする業務についての知識とそれを最適なものとしてシステム化する技能」である。

AEは従来から最も需要の多かったエンジニアであり，これからも新規

分野の広がりによって，ますます需要は増加するであろう。そして，情報処理の専門家からＡＥになる者もいるが，他部門から転換してＡＥになる者もいるだろう。

(2) プロダクション・エンジニア（ＰＥ）

ＰＥ委員会では，ＰＥを次のように規定している。すなわち，現存および今後実用化が予測される広い範囲でのアプリケーションに関して，共通に必要となる基礎的なソフトウェア技術を，単に経験の集積や特定のシステムやアプリケーションに依存したものとしてではなく，理論的な基礎の上に立った工学として身につけたエンジニアである。

ＰＥ養成の重要性の背景については次のような認識を持っている。

まず，社会的背景としては次のことが考えられる。今日，コンピュータの設置台数が増加し，産業，社会，個人の日常生活までもがコンピュータに依存するようになっており，高度なソフトウェア技術が必要とされている。そして，マイクロコンピュータの爆発的な普及とその影響に伴い，既製のソフトウェア・パッケージの使用が常識化し，信頼性と末端利用者にとっての使いやすさが求められるようになった。加えて，コンピュータの応用分野の拡大に対応して，ソフトウェアの生産の質量両面の改善が求められてきている。

技術的背景としては以下のことが考えられる。コンピュータのアーキテクチュの進歩により，高性能のハードウェアの機能や性能を十分に引き出すためには高い水準のソフトウェア技術が必要になった。また，データ通信，ネットワーク，分散処理等の利用環境の多様化は，これを支援するソフトウェア技術を必要とするようになった。さらに，ソフトウェア・ツールと開発環境も著しく進歩し，それらを使いこなす技術的基礎であるソフトウェア工学の修得が必要になってきた。

現在，このようなしっかりしたソフトウェア工学を身につけたエンジニアは極めて少なく，この養成は緊急の課題である。このようなエンジニアとして期待に答えられるようになるためには，相当の歳月を必要とすることや専門職種としての地位の確立が重要である。

(3) ネットワーク・エンジニア（NE）

「提言」におけるTEは，多様なシステム資源の開発・運用によっていくつかの専門分野に分かれる。これらの中で，ネットワーク関係のエンジニアの育成は緊急を要するもののひとつである。

NE委員会では，NEを次のようなエンジニアとして規定している。

すなわちNEとは，ユーザのニーズを把握し，目的とする業務について情報処理と通信を統合した最適ネットワーク・システムを企画，構築・運用のできる，またネットワーク・システムに関するコンサルテーションが行えるエンジニアである。たとえば，

① 経営戦略の一環として，あるいは社会的インフラストラクチャとして，組織間のネットワーク・システムの構築計画あるいは設計をする。

② 最適のネットーワーク・サービスの導入のための評価・選択，あるいは既存ネットワーク・システムの改善をする。

③ 情報通信システムのコンサルタントが行える。

などのエンジニアである。

NE養成の緊急性の背景としては次のような認識に立っている。

現在，情報処理と通信とは密接に融合し，通信ネットワークは情報処理システムの基本的な1要素として位置づけられるようになった。一方，情報，通信双方の新技術の開発と，1985年来の通信事業の自由化の結果，多様な通信関連システムや通信サービスが出現してきて，利用者が各々の目的に応じた最適な選択を行うには十分な知識が必要とされるようになった。このような環境のもとで，情報処理と通信の両技術をマスターし，統合してシステムを計画し，総合的な効率性，信頼性，安全性，経済性等の観点から情報通信システムを構築・運用することができる高度技術者の養成が急速に高まってきている。しかし，情報処理技術者と通信技術者とはその生い立ちが大きく異なり，両技術をカバーし得る技術者の数は極めて限られており，NEの養成は緊急の課題である。

(4) システム監査エンジニア

情報処理システム監査は，情報処理システムの企画，開発，運用などの業務とは独立した客観的な立場から総合的にシステムを点検・評価する業

務であり，その専門的エンジニアの育成は現在，社会的に強い要請がある。

システム監査エンジニア委員会では，システム監査およびそのエンジニアを次のように規定している。システム監査は，コンピュータを中核とした情報システムの企画，開発および運用に関する全業務を対象としており，システムの信頼性，安全性，効率性を高め，それによって情報化社会の健全化に資することを目的とする。

システム監査エンジニアとはシステム監査の専門家であり，監査対象から独立した客観的な立場で，コンピュータを中核とする情報システムを総合的に点検・評価し，関係者に助言・勧告するとともに，その有効利用の促進と弊害の除去とを同時に追求してシステムの健全を図る監査人である。

このエンジニア要請の背景については，次のような認識に立っている。

通商産業省は1985年１月に「システム監査基準」を作成し，1986年度から情報処理システム監査技術者試験を実施してきている。コンピュータと通信を中核としたシステムに強く依存する今日の情報化社会において，そのようなシステムのセキュリティの確保と，システムの社会的，経済的効率性および有効性の向上がますます要請されている。システム監査はこのような時代的要請に応えることを使命としている。そして，システム監査の対象となる情報システムの扱う業務は経理・財務，業務・事務管理など多岐にわたる。

だが，情報処理技術者と，経理・財務担当者，業務・事務管理担当者，そして職業会計人とはその生い立ちが大きく異なる。これらの知識・技術のカバーし得る技術者の数は現在極めて限られており，このエンジニアの要請は緊急の課題である。

(5) ナレッジ・エンジニア（ＫＥ）

1980年代に入って，わが国では急激に人工知能（ＡＩ）ブームが広がり，その過程でナレッジ・エンジニア（ＫＥ）の需要が急増した。機械翻訳，自然言語処理，パターン認識，知能ロボット等のＡＩ応用と並んで最も実用に近づいたものとして，エキスパート・システム（ＥＳ）がある。

ＫＥとはＥＳを構築する専門技術者である。ＥＳは人間のエキスパート（専門家）の持つ知識や，原理，原則，物理的法則等からなる知識ベース

と，その知識を用いて問題解決を行う推論機構からなる。ＥＳの手法は明確な処理アルゴリズムが存在しない問題に極めて有効であり，現在の知識システムの中心的役割を果たしている。

ＥＳを構築するＫＥの持つべき知識・技術は，システム構築に関する従来技術に加えて，ＡＩおよびＥＳ特有の新しい技術・知識が必要となる。

(6) データベース・エンジニア（ＤＢＥ）

今日のように，企業活動が世界的な規模に拡大し，環境がめまぐるしく変化する時代になると，必要な情報の収集を個人や１企業で行うことができにくくなっている。この事態を解消するためには，各種の最新情報を整理し，網羅的に収録したデータベースの利用が必要不可欠となってきている。

しかし，データベースの構築や利用ができる技術者が大幅に不足していることはデータベース白書等でも指摘されており，とくにデータベースの構築ができる技術者の育成が緊急課題となっている。

(7) マイクロコンピュータ応用システム開発技術者（ＭＣ）委員会

マイクロコンピュータの爆発的な普及拡大に伴い，その応用システム技術の開発と周辺技術の利用方法は目覚ましいものがあり，先端技術として重要な位置を占めるようになった。とりわけ，マイクロコンピュータを組み込むことによって機能の向上を図ったシステムはあらゆる産業分野に浸透しており，わが国の産業の土台を支えるものとなっている。このようなマイクロコンピュータの応用分野の急激な拡大は，技術者の不足を引き起こしている。当研究所では，1985年以来，㈶日本情報処理開発協会で作成してきた育成指針を引き継ぎ，今年度は上級のマイクロコンピュータ応用システム開発技術者の指針を作成した。

(8) システム・エンジニア共通

以上述べたＡＥ，ＰＥ，ＮＥ，システム監査エンジニア，ＫＥ，ＤＢＥなどは，それぞれ高度に専門化された知識・技術を必要とするが，その前提として全てのエンジニアに共通したＳＥとしての基礎的な知識・技術が存在する。その内容，および育成方法を体系的にまとめたものがシステム・エンジニア共通育成指針である。

この育成指針は２つの部分に分かれ，ひとつは知識・技術編であり，もうひとつはシステム・エンジニアに共通な資質・能力編である。

以上のような各種のエンジニアに対応する育成指針の他にも，追加すべき専門分野は多様である。また，技術や情報化の進展によって，今後，新たな専門分野出現への対応や既成の育成指針の更新も行わねばならない。

人材育成には，しっかりしたキャリア・デベロップメント・プログラム（ＣＤＰ）が重要である。現状では，各企業におけるＣＤＰは必ずしも十分確立されていない。今後，早急に解決すべき課題であろう。現在，通商産業省の主導のもとに一部のカリキュラムについてＣＡＩの開発が進められているが，教育の効果を上げるために一層の推進が期待される。また，各種の教育・研修機関が大都市，とりわけ東京に集中して著しい地域格差を生じているため，今後ＡＶ機器あるいは通信機能を用いた，地理的格差に左右されない効果的教育手段の整備が急がれる。

— データベース・エンジニア編 —

高度情報処理技術者育成指針
データベース・エンジニア編

委員構成

（敬称略・順不同）

委員長	根岸正光	学術情報センター　教授
委　員	川島　順	(財)日本特許情報機構　情報処理部　部長
	佐々木　浩	(株)日本経済新聞社　情報業務部長
	里　保徳	富士通(株)　システム本部　教育部
	鈴木茂樹	(財)日本情報処理開発協会　調査部調査課　課長
	寺村謙一	丸善(株)　取締役　情報サービス事業部長
	時実象一	(社)化学情報協会　オンライン情報部長
	野々山隆幸	横浜市立大学　商学部　助教授
	灰原一彦	(株)東京商工リサーチ　編集出版部　部長
	長谷川　昇	(特)日本科学技術情報センター　技術管理室 調査役
	三輪眞木子	(株)エポックリサーチ　代表取締役　社長
	甘利直幸	中央情報教育研究所　講師

目　　次

第0部　総　　論 …… 1
第1章　データベース・エンジニア育成の必要性 …… 1
第2章　人材像 …… 2
第3章　データベース・エンジニア育成カリキュラムの概要 …… 3

第1部　データベースの基礎 …… 5
第1章　データベースの概念 …… 6
1.1　データベースの発想 …… 6
1.2　データベースの利点 …… 7
1.3　データベースの構成要素 …… 8
1.4　データベースの種類 …… 8
第2章　データベースの構成 …… 11
1.1　データベースの3層構成法 …… 11
1.2　データモデル …… 12
第3章　データベースの管理システム …… 15
1.1　ＤＢＭＳの歴史 …… 15
1.2　ＤＢＭＳの主要機能 …… 17
1.3　データベース言語 …… 17
1.4　データベースの高信頼性 …… 18
1.5　既存ＤＢＭＳの特色 …… 18
1.6　ＩＲ（情報検索）システム …… 19

第2部　データベースシステムの構築 …… 23
第1章　要件設定 …… 25
1.1　データベース化の目的の明確化 …… 25
1.2　データベースシステムの開発要件 …… 27
第2章　システムの構築 …… 33
1.　データベース・システムの構築 …… 33

1.1　機能設計 …… 33
1.2　性能設計 …… 37
1.3　データベースの設計 …… 39
1.4　データベースの作成 …… 40
1.5　システム環境整備 …… 42
1.6　ユーザインタフェース設計 …… 42
1.7　高信頼化機能設計 …… 44
1.8　運用管理設計 …… 45
1.9　開発サポート …… 47
2.　ＤＢＭＳの選択 …… 50
2.1　業務への適合性 …… 50
2.2　機能・性能 …… 50
2.3　関連ソフトウェア・ツール …… 51
2.4　他のシステムとの結合性 …… 51
2.5　信頼性対策 …… 51
2.6　その他 …… 52
3.　ＤＢＭＳに適した効率化 …… 53
3.1　ハードウェア …… 53
3.2　Ｏ　Ｓ …… 53
3.3　基本ソフトウェア …… 54
3.4　アプリケーション・プログラム …… 54
第３章　性能評価 …… 56
1.1　性能評価の手順 …… 56
1.2　シミュレーション …… 57
1.3　ベンチマークテスト …… 57
1.4　実効時間計測ツール …… 58

第３部　データベースの構築 …… 61
第１章　ソースデータの収集 …… 63
1.1　資料の種類 …… 63

1.2　収集方法 …… 64
1.3　収集形態 …… 65
1.4　データの分析評価 …… 65
第2章　入力データの作成 …… 67
1.1　文献型データ …… 67
1.2　ファクト型データ …… 67
第3章　データの加工 …… 69
1.1　統一処理 …… 69
1.2　変換処理 …… 71
1.3　圧縮処理 …… 72
1.4　流通媒体の作成 …… 73
第4章　データ・ファイルの構造化 …… 76
1.1　構　造 …… 76
1.2　利　用 …… 77
1.3　メンテナンス …… 77
1.4　セキュリティ …… 77
第5章　データの入出力 …… 79
1.1　コード入力 …… 79
1.2　イメージ入力 …… 80
1.3　ベクトル入力 …… 81
1.4　入力データの検査 …… 81
1.5　出　力 …… 82
第6章　流通データの導入 …… 84
1.1　予備調査 …… 84
1.2　選定・契約 …… 84
1.3　ファイル入手 …… 84
1.4　内容調査 …… 84
1.5　加工処理 …… 85
1.6　システムへの組み入れ …… 85

第7章　企業内データベースの構築 …… 87
1.1　予備調査 …… 87
1.2　データ入力 …… 87

第4部　データベースの運用管理 …… 89
第1章　運用管理基準の設定 …… 90
1.1　運用体制 …… 90
1.2　データベースの運用保守 …… 91
第2章　運用体制 …… 93
1.1　責任体制 …… 93
1.2　データベース管理者 …… 93
1.3　オペレータ教育 …… 93
1.4　ユーザ・サポート …… 94
第3章　データベースの運用保守 …… 95
1.1　運用管理支援ツールの整備 …… 95
1.2　データベースの更新 …… 95
1.3　利用データの分析 …… 95
1.4　ユーザ情報の管理 …… 96
1.5　セキュリティ …… 96
1.6　データベースの再編成 …… 96

第5部　データベースの利用 …… 99
第1章　商用データベース・サービス …… 100
1.1　データベースの流通構造 …… 100
1.2　データベース・サービスの選択 …… 100
第2章　商用データベース検索技術 …… 103
1.1　文献情報検索システムとコマンド …… 103
1.2　ファクト系データベースシステムとコマンド …… 103
1.3　その他のデータベース …… 103
1.4　データのダウンロードと多角利用 …… 104

第3章　サーチエイドの活用 …… 105
1.1　サーチエイドの種類 …… 105
1.2　サーチエイド利用の目的 …… 107
1.3　サーチエイドの選択と利用法 …… 107
1.4　サーチエイドの流通 …… 113
第4章　データベース・サービスの国際的動向 …… 115
1.1　歴　史 …… 115
1.2　アメリカのデータベース市場，業界，利用等の最新動向 …… 116
1.3　ヨーロッパ主要国の動向 …… 116
1.4　その他 …… 117
第5章　知的所有権と個人データ保護 …… 122
1.1　知的所有権 …… 122
1.2　個人データ保護 …… 123

第6部　データベースの技術動向 …… 129
第1章　マルチメディア・データベース …… 131
1.1　マルチメディア・データベースとは …… 131
1.2　マルチメディア・データベース出現の背景 …… 131
1.3　国内の研究動向 …… 132
1.4　マルチメディア文書の通信技術 …… 133
1.5　ユーザ・インタフェース …… 133
第2章　フルテキストサーチ …… 135
1.1　フルテキストサーチの概念 …… 135
1.2　マッチング・アルゴリズム …… 135
1.3　応用例 …… 136
第3章　分散型データベース …… 138
1.1　分散型データベースとは …… 138
1.2　分散型データベースシステム出現の背景 …… 138
1.3　分散型データベースシステムの実現形態 …… 139
第4章　データベースマシン …… 141

1.1　データベース基本演算の高速化 …… 141
1.2　記憶制御の高速化 …… 141
1.3　その他の高速化手法 …… 142
1.4　データベースマシンの事例紹介 …… 142
第5章　データベースへのＡＩの応用 …… 144
1.1　提供側への適用 …… 144
1.2　利用側への適用 …… 145
1.3　知的データベース管理システム …… 145
第6章　新記憶媒体 …… 147
1.1　光ディスク …… 147
1.2　光ディスクの種類 …… 147
第7章　ニューメディアの通信機能 …… 149
1.1　ビデオテックス …… 149
1.2　ＶＲＳ …… 149
1.3　ＣＡＴＶ …… 149
1.4　テレテキスト …… 149
1.5　ＡＮＳＥＲ …… 150
1.6　ＩＳＤＮ …… 150
第8章　電子出版 …… 151

第0部　総　　論

第0部　総論

第1章　データベース・エンジニア育成の必要性

今日のように，企業活動が世界的な規模に拡大し，環境がめまぐるしく変化する時代になると，必要な情報の収集を個人や1企業だけでは十分に行うことが難しくなってきている。この事態を解消するために，各種の最新情報を整理し，網羅的に収録したデータベースの利用が必要不可欠となってきている。

データベースは元来，大規模な組織体において，大型の電子計算機を用いて各種のデータを統合的に管理する方式として発達してきた。その後，コンピュータの飛躍的能力向上に伴って，このような方式をより小規模な業務にも適用することが可能になり，最近ではパーソナル・コンピュータによる個人的なデータベースの作成も行われるようになっている。一方，関連企業や一般でも有用なデータベースの利用については，これを組織内に留めずに外部にもサービスして行くという形態が，通信技術の発達とも関連して盛んに行われるようになってきた。また，企業間での効率的なデータ交換にもデータベースは活用されている。今や「データベース」は，個人データベースから企業間データベースや一般商用データベースまで，社会の隅々に行きわたりつつあるといえよう。

それだけに，データベース関連の技術者の不足には深刻なものがある。たとえば，データベースの構築状況について見ると，国産のデータベースの数は約20％（データベース白書1987による。以降の数字は同白書による）にしか過ぎず，データベースの数では大半を海外に依存しているのが現状であり，事実，ベンダの約60％からデータベースの設計・開発・構築を行う技術者の不足を訴えられている。

他方，データベースのユーザについて見ると，データベースの検索ができる研究者や調査員を約60％の企業が不足しているとしており，1社当たりで不足している要員数も約24名となっている。

このような状況を踏まえて，ネットワークを考慮したデータベースの構築や検索ができる新しいタイプの情報処理技術者（データベース・エンジニア）の育成が望まれている。

第2章　人材像

高度情報システムにおいて，その代表的資源のひとつであるデータベースに関与するデータベース・エンジニアには，データベースの作成から運用，利用に至る各段階において種々の役割が期待され，その所属も，電算機メーカ，ソフトウェア・ハウス，電算機ユーザ，データベース・プロデューサ，データベース・ベンダ，調査会社，代行検索業者など，さまざまである。

データベース・エンジニアは，テータベースの作成・供給者の立場からデータベース・システムの設計・構築・運用に携わる「データベース構築技術者」と，データベースの利用における専門技術者である「データベース利用技術者」に区分することができる。これまでの経緯では，前者に該当する人材の多くはいわゆる情報処理技術者の中から育っており，また後者については，ドキュメンタリスト系列の人材が当ってきているのが実情である。しかし，今後におけるデータベースの普及を考えれば，これを支える技術について必要な知識・技術をデータベースの視点から改めて整理し，この体系に基づいて効果的な人材の育成を図る必要がある。本育成指針では，構築技術者，利用技術者の両者共通に要求される基礎的な知識を整理・配列している。

第３章　データベース育成カリキュラムの概要

データベースの基礎
データベースの概念 データベースの構成 データベース管理システム
データベースシステムの構築
要件設定 システムの構築 データベースシステムの構築 ＤＢＭＳの選択 ＤＢＭＳに適した効率化 性能評価
データベースの構築
ソースデータの収集 入力データの作成 データの加工 データファイルの構造化 データの入出力 流通データの導入 企業内データベースの構築

データベースの運用管理
運用管理基準の設定 運用体制 データベースの運用保守
データベースの利用
商用データベース・サービス 商用データベース検索技術 サーチエイドの活用 データベース・サービスの国際的動向 知的所有権と個人データ保護
データベースの技術動向
マルチメディア・データベース フルテキストサーチ 分散型データベース データベースマシン データベースへのＡＩの応用 新記憶媒体 ニューメディア通信機能 電子出版

第1部　データベースの基礎

第１部　データベースの基礎

目　　標

データベースの基本的知識を整理する。とくにデータベースの概念，データモデルが重要であるので，これらについての知識を修得させる。

内容のあらまし

内　　容	時間数
データベースの概念 　データベースの発想 　データベースの利点 　データベースの構成要素 　データベースの種類	6 時間
データベースの構成 　マキーマ 　データモデル	3 時間
データベース管理システム 　ＤＢＭＳの歴史 　ＤＢＭＳの主要機能 　データベース言語 　データベースの高信頼性 　既存ＤＢＭＳの特色 　ＩＲシステム	3 時間
合　　計	12時間

第1章　データベースの概念

目　　標

データベースの基本的知識を整理する。とくにデータベースの概念，データモデルが重要であるので，これらについて深く勉強することが望まれる。

内　　容

1.1　データベースの発想

データベースとは，ＪＩＳの定義では１つ以上のファイルの集まりであって，ソフトウェアの内容を高度に構造化することによって検索や更新の効率化を図ったものである。従来のデータがアプリケーションプログラムごとにファイルとして作成され，管理されていたのに対して，これを全体として（あるいは管理対象別に）データベースとして管理するというのがデータベースの発想である。図1.1はファイルとデータベースのイメージを示している。

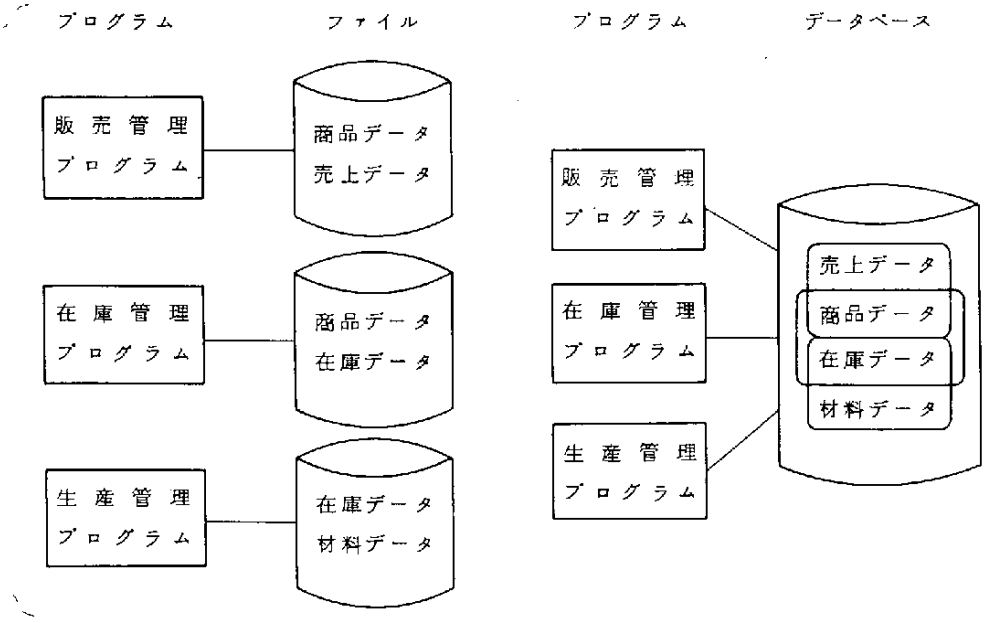


図1.1　ファイル方式とデータベース方式

1.2　データベースの利点

従来のデータのファイル別管理に代って，データベースとしてデータを管理することの利点は，

(1)　データの共有ができる。データの共有により，組織全体で所有するデータを有効に利用することができるようになる。

(2)　2次記憶装置の節約になる。データを共有することでデータの冗長度を減少させ，2次記憶装置（特に磁気ディスク）の容量を節約することができる。

(3)　システムの変更に柔軟に対応できる。従来の応用ソフトウェアに従属した形のファイルによるデータ管理では，応用ソフトウェアの変更によってファイル自身の変更が頻繁に必要になる。そこで，応用プログラムとデータベースによるデータ管理とを分離することで，応用プログラムが修正されてもデータファイルの変換（修正，再構築）をしなくてもすむようにできる。この結果，システム変更への柔軟性が高まり，システム自体の寿命が延びる。このような応用ソフトウェアとデータ管理との分離をデータ独立という。

(4)　共通に使える検索手段の提供ができる。データをデータベースとして統一的に管理することで，データベースに対する検索手段を共有することができる。データをファイルごとに管理していた状況では，検索手段もファイルごとに用意しなければならなかったが，データベースにすることでこれが1つでよいことになる。これにより応用プログラムの負担が軽減される。後に説明するように，検索手段としてはデータベース専用の検索言語による方法（検索言語方式）と，応用プログラムで利用している親言語（第3世代の手続き型高水準言語）に検索機能を組み込む方法（親言語方式）とがある。

(5)　データ管理のセキュリティが高まる。ファイルごとにデータを管理すると，組織全体での統一的なデータセキュリティ実現が難しくなるが，データベースによりデータを統一的に管理すれば，データセキュリティの統一的集中的管理が容易になる。

(6)　データの標準化が促進される。データの利用者間でデータコードや単位

の標準化が促進される。

1.3　データベースの構成要素

データベースシステムは図1.2のように構成されている。まず，全体はデータベースとデータベース管理システムにより構成されている。後者のデータベース管理システム利用者からの入力（検索要求など）を質問処理部で解析し，これをデータベース操作部に渡す。データベースの実際の操作は，データ辞書部を参照しながらデータベース操作部が実行する。データベースから抽出されたデータは，利用者の要求にしたがって出力変換部で出力様式をそろえて出力される。データベース操作の一貫性管理は一貫性制御部が行い，並行処理部がデータ操作の並行処理をすることで処理速度を向上させる。システムにトラブルがあった場合には，回復処理部がデータベースの回復を行う。その他，データベース設計を支援するデータベース設計援助システム，データベースの再構成を支援するデータベース再構成援助システムなどもデータベース管理システムに含まれている。

1.4　データベースの種類

データベースに含まれているデータの種類や検索方法でデータベースを分類することができる。この分類基準には次のようなものがある。

(1)　文献データベースとファクトデータベース

特定のキーワードを含む文献の題名，著者名，出版年，内容などを検索できるデータベースを文献データベースという。これに対して，統計データや科学的，社会的事実のデータをまとめたものをファクトデータベースという。

(2)　文字，数値，画像，映像，音声データベース

データとして文字，数値，画像，映像，音声を含むものを，それぞれ文字データベース，数値データベース，画像データベース，映像データベース，音声データベースという。

(3)　1次情報データベースの2次情報データベース

情報のもとになっているもの（原典，原著，原データなど）を1次情報

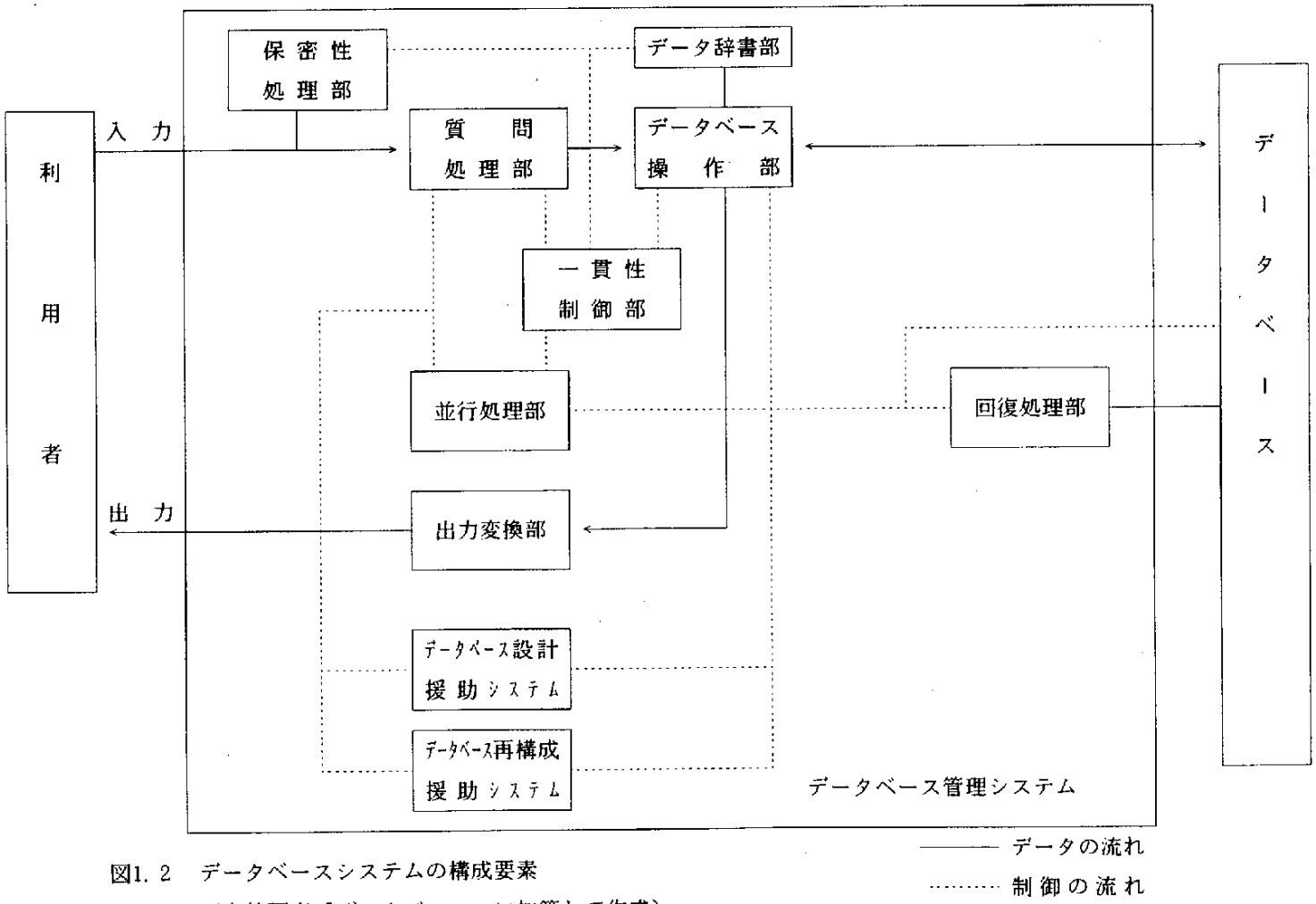


図1.2　データベースシステムの構成要素
（上林彌彦「データベース」に加筆して作成）

と呼び，これを加工，集計したものを２次情報と呼ぶ。データベースに含まれている情報が１次情報か２次情報かによって，それぞれ１次情報データベース，２次情報データベースと分類する。

(4) オンラインデータベースとバッチデータベース

オンラインで検索できるデータベースがオンラインデータベースで，バッチ形式で検索できるデータベースがバッチデータベースである。

指導上の留意点

ファイルとデータベースの発想の相違を理解させる。

用　　語

ファイル，データベース，検索手段，検索言語方式，
第３世代の手続き型高水準言語，親言語方式，データ管理，
データベース管理システム，質問処理部，データベース操作部，
データ辞書部，出力変換部，一貫性制御部，並行処理部，回復処理部，
データベース設計援助システム，データベース再構成援助システム，
文献データベース，ファクトデータベース，１次情報データベース，
２次情報データベース，オンラインデータベース，バッチデータベース

第2章　データベースの構成

目　　標

この章では，データベースの3層構成法および代表的なデータモデルについて，その特徴を説明し，理解させる。

内　　容

1.1　データベースの3層構成法

データベースは，概念的には次の3つのスキーマによって構成する。

(1)　外部スキーマ

アプリケーションごとに異なる利用者のビュー（データに対する見方）の集合で，それぞれの利用者あるいは応用プログラムによって違っている。利用者のデータの見方を反映している。

(2)　概念スキーマ

概念スキーマとは外部スキーマを統合したシステム内部でのスキーマで，データベース全体を表している。データベース管理者のデータの見方を反映している。

(3)　内部スキーマ

物理的なデータの記憶構造であり，コンピュータのデータの見方を反映している。まず，スキーマとは「データベースの記述」の意味で，ＣＯＤＡＳＹＬによって使われた用語である。このようにスキーマを3層にすることで，アプリケーションやハードウェアに対するデータ独立性を実現することができる。すなわち，アプリケーションごとにデータの見方は違うが，外部スキーマを概念スキーマから独立させることで，外部スキーマがアプリケーションごとのデータの見方の違いを許容しながら，概念スキーマがデータベース全体の統合を実現している。また，ハードウェアや物理的なデータベース構造の違いを内部スキーマで吸収してしまうことで，概念スキーマにこれらの違いが影響しなくなる。この結果，データベースの移植が容易になり，ハードウェアやデータ配置，アクセス方法の変更

にデータベースが影響されなくなる。図1．3はこの3層構造を示している。

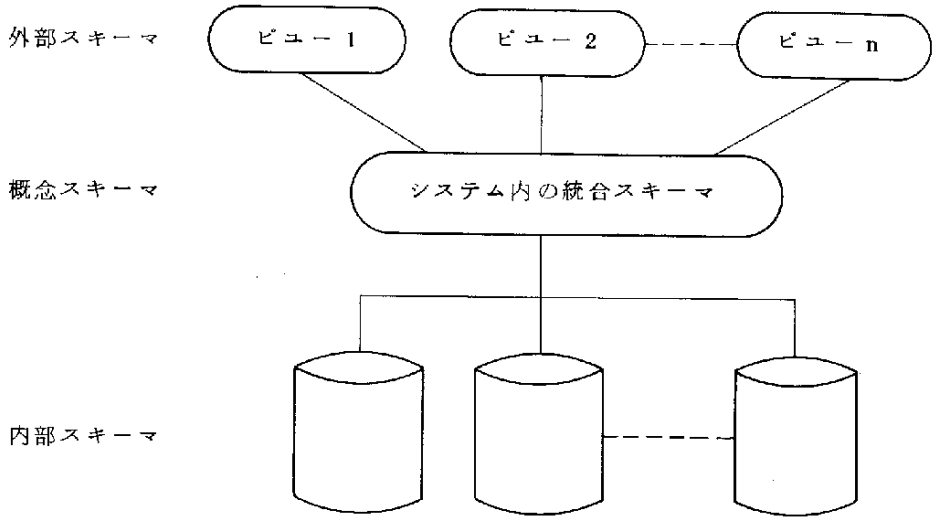


図1．3　データベースの3層スキーマ

1．2　データモデル

データモデルに含まれるデータの概念的なつながり（結びつき）をデータモデルと呼ぶ。代表的なデータモデルには次の3つがある。

(1)　階層（木）型モデル

レコード間の関係が上下，あるいは親子のツリー（木）状の階層関係を持つ。1つの親レコードは複数の子レコードを持ち得るが，1つの子レコードは1つの親レコードしか持たない。記憶効率（記憶容量のうち実際にデータが蓄えられている部分の割合）が高く，もともとの設計に適合した検索は高速だが，融通性に欠けるという欠点がある。定型化された検索中心の大容量のデータベースに利用される。

(2)　ネットワーク（網）型モデル

階層型モデルで，1つの子レコードが複数の親レコードを持てるようにしたものをネットワーク型モデルという。階層型モデルよりは融通性は少し高いが，記憶効率は少し落ちる。ネットワーク型モデルはＣＯＤＡＳＹＬ方式データベースとも呼ばれる。階層型モデルはネットワーク型モデルをより制限したものと考えられる。

(3) リレーショナル（関係）型モデル

データやデータ間の関係を表形式で表現する。この表は関係と呼ばれ，レコードに相当するものは組（タプル），フィールドに相当するものは定義域（ドメイン）または属性と呼ばれる。リレーショナル型モデルは，ＩＢＭサンホセ研究所Ｅ．Ｆ．Ｃｏｄｄによって1970年に初めて提唱されたデータモデルで，情報代数による理論的なデータ（関係）操作に適合したモデルである。リレーショナル型モデルはデータ（関係）操作における融通性が高く，処理内容による処理効率の違いは小さいが，処理効率は悪い。多様な関係操作を行う小規模のデータベースに向く。

図1．4は，以上のデータモデルの型を図式化したものである。

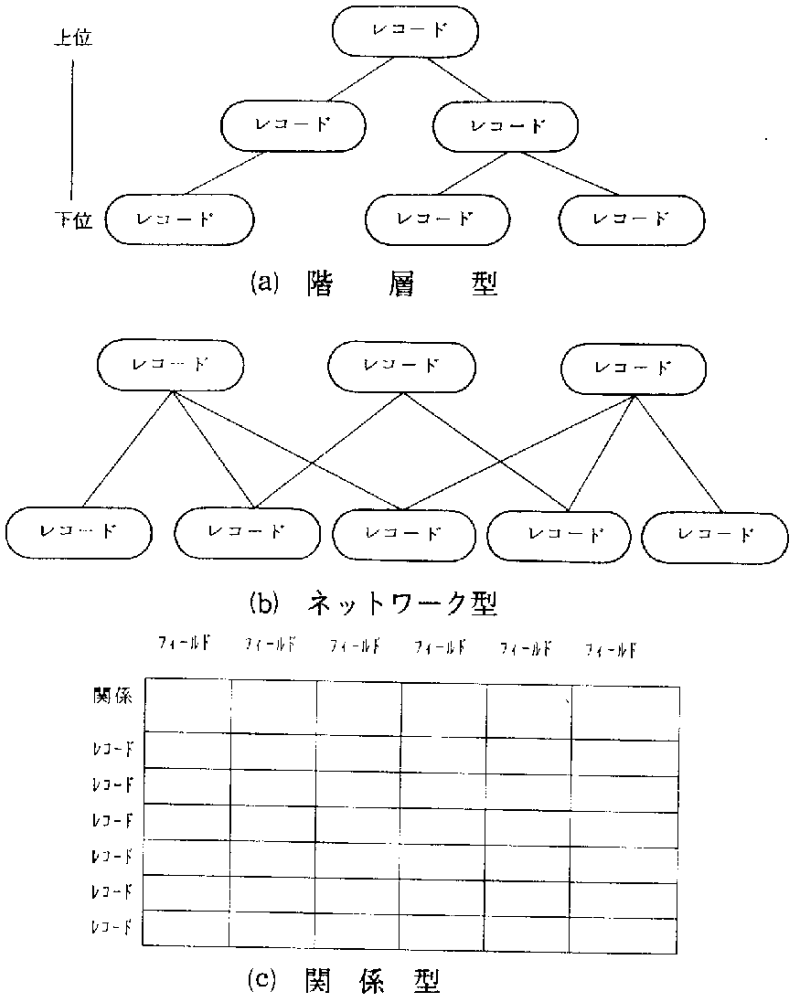


図1．4　データモデルの型

指導上の留意点

データベースやスキーマの概念，データモデルがとくに重要である。それぞれのデータモデルの特徴を理解させる必要がある。また，データモデルを記述する一般的方式であるバックマン線図の書き方も基礎的な知識である。関係型モデルでは関係操作（関係代数，関係倫理），関係正規化（第１－第４正規形），関係言語族などの知識も望まれる。

用　　語

外部スキーマ，ビュー，概念スキーマ，内部スキーマ，データベースの記述，ＣＯＤＡＳＹＬ，データ独立性，アプリケーション，データベースの移植，データ配置，アクセス方法，データモデル，階層（木）型モデル，親子のツリー（木），階層関係，記憶効率，ネットワーク（網）型モデル，ＣＯＤＡＳＹＬ方式データベース，リレーショナル（関係）型モデル，表形式，組（タプル），フィールド，定義域（ドメイン），属性，Ｅ．Ｆ．Ｃｏｄｄ，情報代数，理論的なデータ（関係）操作

第3章　データベース管理システム

目　　標

この章では，データベース管理システムの機能についての知識を修得させる。

内　　容

1.1　DBMSの歴史

データベースを管理するソフトウェアをデータベース管理システム（DBMS，Data Base Management System)と呼ぶ。

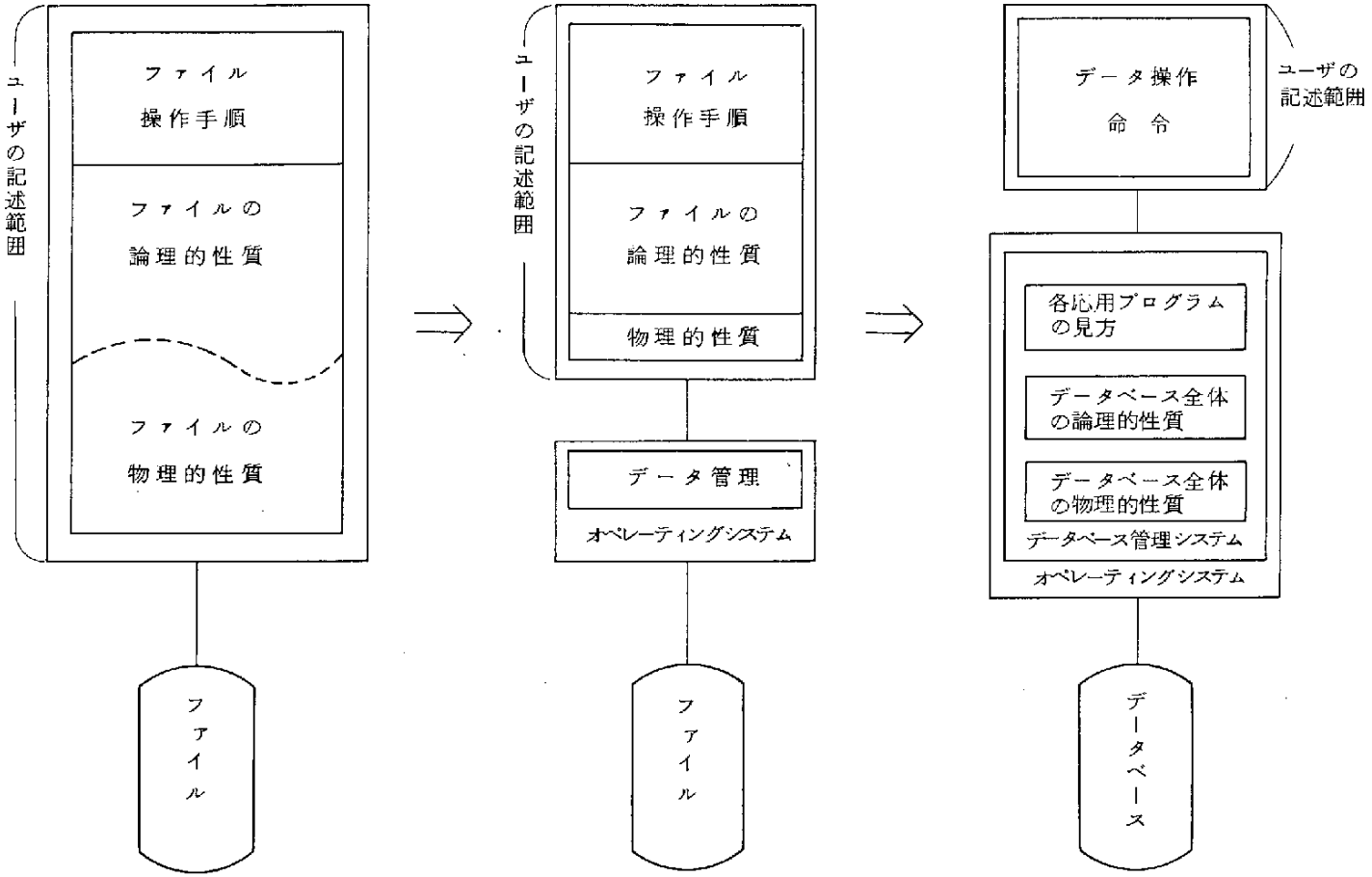
データ管理の歴史を振り返ると，図1.5のようになる。

最初は，ファイルを操作するためにはファイルの操作手順，論理的性質，物理的性質の全てについて熟知したプログラマーが，全部をプログラムで記述しなければならなかった。ここでは，ファイルの論理的性質と物理的性質との区別も不明確であった。

次に，データが資源として認識されるようになると，データ操作の共通的な部分がデータ管理プログラムとして提供されるようになった。これにより，一般の利用者はファイルの論理的性質までを記述すればよく，データの物理的性質についての認識はわずかですむようになった。データ管理プログラムが，オペレーティングシステムのもとでファイルに対する物理的操作を分担するからである。

そして，DBMSが実現すると，利用者はデータの操作についての命令を書くだけでよくなった。利用者は，従来のようにファイル操作の手順を記述する必要はなく，データベースに対する知識は最小限でよいようになった。オペレーティングシステムのもとで，DBMSがファイルの論理的性質と物理的性質も管理してくれるためである。前記のように，DBMSは外部スキーマ，概念スキーマ・内部スキーマに分かれており，内部スキーマがデータベースの物理的性質を，概念スキーマがデータベース全体の論理的性質を，そして外部スキーマが各応用プログラムの論理的な見方を記述している。

図1．5　データ管理の変遷



ＤＢＭＳは，以上のようなデータ管理の変換のなかで発展してきた。ＤＢＭＳはマギー（W.C.McGee)の源ファイル(source file）の概念がそもそもの発端であるが，これ以降，

(1) 事務用共通言語であるＣＯＢＯＬを中心としたＣＯＤＡＳＹＬ方式のデータベースの発展

(2) 汎用報告書作成プログラムから始まる，さまざまな商用データベースシステムパッケージの発展

(3) Ｅ．Ｆ．Ｃｏｄｄの関係モデルの基礎理論に基づくリレーショナルデータベースの発展

の３つの流れの中で発展してきたという歴史を持っている。

1.2 ＤＢＭＳの主要機能

(1) データベースを種々のアプリケーションが利用する場合のコマンド体系，または高水準の言語体系の付与

(2) 検索条件の解読を含むデータの検索処理

(3) ２層または３層スキーマに基づく論理構造，物理構造の分離によるデータ独立性の実現

(4) データベースの創成，再配置，更新

(5) 障害時の処理，データの完全性の確保，アクセス制御などのインテグリティ，セキュリティ制御

1.3 データベース言語

データベース言語には次のような機能がある。

(1) データ操作機能：データの検索，変更，集約（平均，最大，最小の計算），出力など。リレーショナル型の場合は集合演算や関係演算（射映，制約，選択，結合，除算）機能がある。

(2) スキーマ定義機能：スキーマ生成，変更消去，ビューの定義

(3) インターフェース機能：言語の中に組み込んで使える機能

(4) 運用機能：パスワードや利用権の管理，障害回復処理やバックアップ方式の指示

データベースに蓄積されるデータがどのような性質を持っているかを定義する言語を，データベース記述言語(ＤＤＬ，Data Description Language)あるいはデータベース定義言語という。他方，データベースのデータに対する操作をするための命令を，データベース操作用コマンドあるいはデータ操作言語（ＤＭＬ，Data Manipulation Language）という。

データベース言語の形態は，ＣＯＢＯＬなどの親言語の中にその拡張機能としてデータベース操作用コマンドを組み込み，プリプロセッサやサブルーチン・コール方式で動作させる親言語方式と，独立したデータベース言語の形を取る自立言語方式とがある。

1.4　データベースの高信頼性

データベースの信頼性を高めるために，今日のデータベース管理システムには次のようなデータ保護の機能がある。

(1)　データの呼び出し制御（アクセス・コントロール）

資格のない利用者がデータをアクセスしないように保護すること。

(2)　データの一貫制御（インテグリティ・コントロール）

データベースに対するデータ変更によって，データの一貫性が維持されるように制御すること。これには次のようなものが含まれる。

①　データの正当性確認：データベースに入力されるデータが正当であることを確認すること.

②　同時実行制御：データベースを同時にアクセスしている複数の応用プログラムが互いに干渉しないように制御すること.

③　障害回復：データベースに障害が発生したときに，データベースを最新の状態に回復させること.

1.5　既存ＤＢＭＳの特色

代表的なＤＢＭＳには次のようなものがある。

(1)　ＩＤＭＳ／Ｒ：リレーショナル型とネットワーク型の複合したＤＢＭＳである。

(2)　Datatrieve：ＤＥＣの代表的ＤＢＭＳで，簡単な英文命令で利用でき，

強力なオンラインヘルプ機能を持つが，インタープリターであるために処理効率は高くない。

(3) ＴＯＴＡＬ：シンコムシステムズ社が1968年に発表したＤＢＭＳで，ネットワーク型に近い構造をしており，簡単明快な呼び出し方式を提供している。

(4) ＩＭＳ：ＩＢＭの主力ＤＢＭＳで，典型的な階層型ＤＢＭＳである。

(5) ＡＤＡＢＡＳ：1972年にＰ．シュネルが発表した部分転置方式の索引を用いたシステムで，検索能力に優れている。

(6) ＤＢ２：ＩＢＭのリレーショナル型ＤＢＭＳである。

(7) ＳＹＳＴＥＭ2000：1969年にＭＲＩシステムズ社が発表した階層型ＤＢＭＳで，部分転置方式の索引を活用したシステムである。

(8) ＩＮＧＲＥＳ：カリフォルニア大学バークレー分校でＵＮＩＸ上に開発された関係型ＤＢＭＳで，ＱＵＥＬ，ＣＵＰＩＤなどの検索言語を持つ。

(9) ＳＹＳＴＥＭ　Ｒ：ＩＢＭサンホセ研究所で開発された関係型ＤＢＭＳで，ＳＱＬ，ＱＢＥなどの検索言語が利用できる。

(10) ＡＤＢＳ：1974年日本電気がＣＯＤＡＳＹＬのデータベース・タスク・グループ（ＤＢＴＧ）の提案を実現した汎用のＤＢＭＳで，データ構造としては階層木構造のほか，一部，網構造も扱える。

(11) ＡＤＭ／ＰＤＭ：日立製作所が開発したＤＢＭＳで，ＡＤＭはユーザのアプリケーション・システムの運用環境に応じたデータベース機能（単純構造，階層構造，階層網構造），データ・コミュニケーション機能，システム管理機能を有する。ＰＤＭはユーザのアプリケーション・システムにおいて発生するデータを蓄積し，操作する手段を提供するデータ管理システムで，データベース機能は網構造である。

(12) ＡＩＭ：富士通が開発したオンライン・データベースシステムで，ＣＯＤＡＳＹＬのＤＢＴＧ仕様に準拠した網構造型のＤＢＭＳである。このほかにもＭＵＭＰＳ，ＭＯＤＥＬ２０４などのＤＢＭＳがある。

1.6　ＩＲ（情報検索）システム

文献データベースの効率的な検索のためのデータベースシステムをＩＲ

システムという。この種のものは，機能的には一般的なＤＢＭＳを応用しても実現可能ではあるが，長大な文章データの大量な収容力，文章データからの索引語の切り出し機能，大量な索引語の組み合わせ検索など，文献検索特有の要件から汎用ＤＢＭＳでは実用的なシステムを構成できないことが多い。このため，オンライン的なデータ更新が不要である点を活かし，また，文献データに即したデータ・モデルを前提とした，効率的な検索が可能なソフトウェアとしてＩＲシステムが開発されており，商用データベースの運用などに利用されている。

指導上の留意点

ＤＢＭＳの歴史を語ることにより，ＤＢＭＳの持つべき機能や考え方を理解させる。

用　　語

データベース管理システム（ＤＢＭＳ），データ管理，ファイル操作手順，論理的性質，物理的性質，物理的操作，Ｗ．Ｃ．マギー，源ファイル，コマンド体系，言語体系，データ独立性，アクセス制御，インテグリティ，セキュリティ制御，集合演算，関係演算，スキーマ定義機能，インタフェース機能，運用機能，障害回復処理，バックアップ方式，データベース記述言語（ＤＤＬ），データ操作言語（ＤＭＬ），親言語方式，自立言語方式，データベースの高信頼性，アクセス・コントロール，データの正当性確認，同時実行制御

参考文献

データ資源管理－企業内データの有効活用を目ざして，William R. Durell著
味村重臣監修，日経マグロウヒル，1987年

データベース：理論と手法，James Martin著，成田光彰訳，
日経マグロウヒル，1986年

データベース環境の実現と管理，James Martin著，坂本・塗師・木下訳，
日経マグロウヒル，1987年
データベースシステムの基礎，植村俊亮著，オーム社，1979年
データベース，上林彌彦著，昭晃堂，1986年
データベース入門，小碇　雄著，啓学出版，1983年
情報資源管理の技法－ＥＲモデルによるデータベース設計，酒井博敬著，
オーム社，1987年
オンラインデータベース活用法，白石謙一著，日経マグロウヒル，1986年
情報の構造とデータベース，長尾・片山・植村共著，岩波書店，1983年
データベース：電子図書館の検索・活用法，滑川・下中・市川共著
東洋経済新報社，1984年
最新データベース事情，名和小太郎著，日本能率協会，1984年
データベース要論，穂鷹良介著，共立出版，1978年
データベース入門，穂鷹良介著，オーム社，1978年

第2部　データベースシステムの構築

第2部　データベースシステムの構築

目　　標

データベース管理システムの機能を十分に発揮させることができる能力，および性能の評価ができる能力を習得させる。

内容のあらまし

内　　容	時間数
要件設定 　データベース化の目的の明確化 　データベースシステムの開発要件	6時間
システムの構築 　データベース・システムの構築 　　機能設計 　　性能設計 　　データベースの設計 　　データベースの作成 　システム環境整備 　ユーザインタフェース設計 　高信頼化機能設計 　運用管理設計 　開発サポート 　ＤＢＭＳの選定 　　業務への適合性 　　機能・性能 　　関連ソフトウェア・ツール 　　他のシステムとの結合性 　　信頼性対策	18時間

内　　容	時間数
その他 ＤＢＭＳに適した効率化 ハードウェア ＯＳ 基本ソフトウェア アプリケーション・プログラム	
性能評価 性能評価の手順 シミュレーション ベンチマークテスト 実行時間計測ツール	6 時間
合　　計	30時間

第１章　要件設計

目　　標

データベースシステムの構築に当って次の要件があること，およびそれらについての理解がなされていることが必要である。

(1)　データベース化の目的が明確であること。

(2)　データベース化を実現するための要件として次の項目が明確化されていること。

①　開発手法要件

②　機能要件

③　性能要件

④　信頼性・安全性要件

⑤　運用性要件

とくに，データベース化の目的を明確化することは重要であり，よりよいデータベースを開発するための必須条件でもあるので，これら各要件を十分に理解させ，要件設定ができる能力を修得させる。

内　　容

1.1　データベース化の目的の明確化

データベースシステムを構築するに当って，データベースエンジニアはシステムのニーズが何であるのか，また，そのニーズに対してデータベースがどのように応えるのか，ということを明確にしなければならない。つまり，データベースは従来のファイルと機能面で異なるのみでなく，それを利用したシステム全体の構成概念にまで影響するものであることを理解しなければならない。

すなわち，システムをデータベース化することは，従来システムでは採用できなかったデータを中心にした考え方を取り入れることでもあり，それを中心にして応用システムを構築することになる。たとえば，従来でのＥＤＰが単に数値の計算や定まった形でのデータ編集に終始していたのに対し，デ

ータベース化を採用したＥＤＰでは，すでにデータベースに蓄えられたデータをもとにして組織内の情報処理が存在する形になる。したがって，データベースシステムの構築は，その組織全体の情報処理ニーズから総合的に考えられるべきであり，それだけに純粋技術面だけでは解決できない調整能力も必要とされる。

(1) データベースをまず知ること

データベース管理システムは何の目的で作られたかということをまず知らなければならない。つまり，従来のファイル管理システムと機能的に比較して，単なる高度なアクセス手段としてデータベースを誤解してはならない。データベースの第１の目的は，データを統合化し，共通的にデータを利用することである。データベースにおけるデータの関連を利用したデータの利用は，本当の意味で組織の情報処理を連動したシステムとすることを可能にする。現状では，従来のファイルシステムを利用した製造管理システム，営業管理システム，経理システム等はあっても，それらが連動することを可能としている組織は極めて少ないと言える。個々のシステムのつながりはデータによって可能とされるため，データが個々のシステムより共通的に利用される仕組みがなければ，その実現性はない。データベース管理システムはそれを実現するものであり，真のデータベース管理システムの目的もそこにあるといえる。

(2) データベース化の目的

データベース管理システムの目的が理解できたならば，データベースシステムを構築する目的を明確にしなければならない。考えられるデータベース化の目的には次のものがある。

① データ中心の考えかたを導入すること

② 資産としてのデータを十分に活かすこと

③ 個々に存在するシステムを統合すること

④ 欲しいときに欲しいデータが即座に得られること

⑤ データの重複をなくし，データ保守の軽減を図ること

⑥ データとプログラムの独立性を高めること

⑦ エンドユーザ言語の利用により，システム開発の効率を上げ，バック

ログをなくすこと

⑧ 多様な条件による検索を可能とすること

などが挙げられる。

(3) データの信頼性

データベースは，データの信頼性があって初めて成り立つものである。そのため，データベース化するデータの信頼性，制約事項などを十分調査する必要がある。

1.2 データベースシステムの開発要件

データベース化の目的が明確化されたならば，次にデータベースの開発要件に沿って開発が進められなくてはならない。ここでは，データベースの開発に必要となる知識について述べる。

(1) データベースシステム開発手法要件

データベースシステムの開発に限らないが，システムを開発するためにはその手順と手法が作業に先立って明確にされなければならない。とくに，データベース設計手法に何を採用するかは重要な問題である。コンピュータシステム上でシステム開発を行う場合，ほとんどはシステム開発手法やツールをメーカが準備しているので，それを利用することができる。また，外部のコンサルタントより開発手法を導入することも可能である。

いずれにしても，前項で示したデータベース化の目的を実現するための最も効率のよい開発手法を，開発に先立って見つけ出すことが重要である。

(2) 機能要件

システムに対するニーズを満たすために，データベースシステムがどのような機能を備えなければならないかということを明確にし，それについての理解を深める。データベースシステムに対する機能要件には次のものが考えられる。

① 検索機能

ア．どのような条件で目的とするデータを選択するのか？

→条件の種類

イ．どの程度動的に条件が設定されるのか？

→動的な処理の種類

ウ．どの範囲での検索がなされるのか？

→検索の対象データ

② データの加工編集機能

ア．データベースに蓄積されたデータをどの程度加工，編集してプログラムまたは端末に渡すのか？

イ．加工に対する関数として何が必要か？

ｅｘ．ＳＵＭ，ＡＶＧ，ＭＡＸ，ＭＩＮ，・・・・・

③ データ更新機能

ア．データ更新で必要とされる機能の種類は何なのか？

ｅｘ．ＳＴＯＲＥ，ＤＥＬＥＴＥ，ＭＯＤＩＦＹ，・・・・

イ．どの程度自動的なデータ更新が必要なのか？

ウ．データの論理矛盾に対してどのようなチェックを必要とするのか？

④ インタフェース機能

ア．どのホスト言語からデータベース機能を利用するのか？

ｅｘ．ＣＯＢＯＬ，ＦＯＲＴＲＡＮ，ＰＬ／Ｉ，・・・

イ．どのアプリケーションプログラムとのデータ授受があるのか？

ウ．どの端末ソフトウェアとのデータ授受があるのか？

→ＭＭＬ（マイクロメインシームリンク）

エ．どのようなエンドユーザ言語とのインタフェースが必要なのか？または，どのようなエンドユーザ言語が用意されているのか？

⑤ ユーティリティ機能

ア．どのようなデータベースに対するユーティリティ機能が必要なのか？

ｅｘ．セーブ，ロード，プリント

イ．データベース格納レコード件数，その処理時間はどの程度まで要求されるのか？

(3) 性能要件

前項で，データベースシステムに必要とされる機能要件が明確化されたが，次にそれぞれに対する性能要件には何があるのかということを理解す

る。性能要件として次のものが考えられる。

① 応答時間

与えられた条件に対して，どの程度の応答時間でなければならないのか？　応答時間を左右する要素として次のものがある。

ア．検索母集団のレコード件数

イ．検索の条件（ＡＮＤ，ＯＲ，・・・）

ウ．ファイル間の演算（ＪＯＩＮ，・・・）

エ．キーの有無

オ．加工の有無　　等

② プログラム処理時間

バッチモードで大量にデータ処理（検索，更新，加工，・・・）する場合の処理時間がどの程度でなければならないのか？　処理時間を左右する要素として次のものがある。

ア　処理トランザクションの件数

イ．１トランザクションの処理時間

ウ．ＳＯＲＴの有無とその処理件数

エ．プリントその他編集内容　　等

③ ユーティリティ処理時間

大規模データベースを効率よく運用するためには，データベースに対するユーティリティ利用のための処理時間がどの程度でなければならないのか？　その内容には次のものがある。

ア．論理データロード時間

イ．論理データセーブ時間

ウ．物理データロード時間

エ．物理データセーブ時間

オ．論理データ追加時間

カ．論理データ削除時間

キ．アライン（ディスク領域の再利用化）時間

ク．各種データベースリポート取得時間　　等

(4) 信頼性・安全性要件

データベースシステムが運用されるに当って必要とされるその信頼性・安全性について理解する。その内容には次のものが考えられる。

① 信頼性

信頼性とは，システムに障害が発生したとしてもシステムを障害発生前の状態に復帰させ，かつ処理を続行し，障害によってシステムが停止したり誤動作しないための障害に対する免疫の度合いをいう。信頼性を高めるためにはどのような機能が必要とされているのか，その内容について理解する。信頼性を高めるための仕組みには次のものがある。

ア．ジャーナル取得機能（論理，物理２種あり）

→データベース更新前，更新後

イ．データベース復元機能

→ＴＢＯ（トランザクションバックアウト）

→ファイルリストアとフォワードリカバリ

ウ．チェックポイントの設定とその取得

エ．２重書き

オ．自動エラー修正機能　　等

② 安全性

安全性とは，許されていない人やプログラムからのデータベースへのアクセスを食い止めることであり，別名セキュリティとも呼ばれている。安全性には何があるのか，その内容について理解をする。安全性の内容には次のものがある。

ア．セキュリティの対象

a．データベース ── ＤＤＬ（ディレクトリ）
　　└── データ

b．プログラム

c．ユーティリティ

d．オンライン画面

e．エンドユーザ言語

f．端　末

イ．セキュリティの方法

a．パスワード

b．ＩＤカード

c．端末の鍵

d．部門による方法

e．時間による方法

f．指紋等の自動認識　　等

(5) 運用性要件

データベースシステムを効率よく運用するための要件について理解する。運用性要件としては次のものがある。

① データベース利用状況管理

どのデータベース（表）が誰からどの程度利用されているかなど，データベースの利用に関するリポートの出力を行う。

② スキーマ／サブスキーマ／プログラムの関連の把握

どのプログラムからどのサブスキーマ（または仮想表）を利用しているのか，また，どのサブスキーマからどのスキーマを利用しているのかなどのリポートの出力を行う。

③ 課金管理

データベースの内容をサービスするに当って，それ相応の課金管理がなされなくてはならない。課金の対象となる要件は，データベース資源に限れば次のものが考えられる。

ア．データベースの種類

イ．検索回数

ウ．検索レコード件数

エ．表示レコード件数

オ．ＭＭＬの有無

カ．グラフ／その他機能の利用

キ．中間データの保持容量と期間　　等

指導上の留意点

データベースシステムの構築に当り，要件設定は最も上流工程で重要であるので，実習を伴う指導が望ましい。

用　　語

開発手法要件，機能要件，性能要件，信頼性・安全性要件，運用要件，データベース管理システム，エンドユーザ言語，データベースシステム，データベース設計手法，検索機能，データの加工編集機能，データ更新機能，インターフェース機能，ユーティリティ機能，応答時間，検索母集団，検索の条件，ファイル間演算，プログラム処理時間，ユーティリティ処理時間，大規模データベース，論理データロード時間，論理データセーブ時間，物理データロード時間，物理データセーブ時間，論理データ追加時間，論理データ削除時間，アライン時間，リポート取得時間，ジャーナル取得機能，データベース復元機能，トランザクションバックアウト，ファイルリストア，フォワードリカバリ，チェックポイント，二重書き，自動エラー修正機能，セキュリティ，ＤＤＬ，パスワード，ＩＤカード，データベース利用状況管理，スキーマ，サブスキーマ，課金管理，検索回数，検索レコード件数，表示レコード件数，中間データ

第２章　システムの構築

1.　データベース・システムの構築

目　　標

データベースの構築に当っては，要求仕様書の内容に照らして既存のデータベース管理システム（ＤＢＭＳ）を選定し，構築するのが一般的であるが，既存のＤＢＭＳでは要求仕様書の内容を満たすことができない場合には，ＤＢＭＳの改善も実施しなければならない。ここでは，データベース・システムの構築ができる能力の修得のほか，ＤＢＭＳの評価やＤＢＭＳの改善に必要な知識も合わせて修得することを目標とする。

内　　容

1.1　機能設計

従来のファイルによる情報処理システムは，プログラムとファイルが対応したシステムであったため，①データ項目の重複が多く，ファイルが冗長であったこと　②プログラムやファイルに変更があったとき，プログラムとファイルの双方に負担がかかったことなどで，システムの管理，保守，変更，統合に膨大な費用がかかっていた。

この問題を解決するには，処理中心のアプローチからデータ中心のアプローチへという発想の転換が必要であり，この結果，データベースという概念が生まれたのである。

データベースという発想の誕生により，冗長性のなくなったシステムから最新のデータが即時に抽出できるようになった。

(1)　データベース自体の機能

①　データ定義

データ定義言語（ＤＤＬ，Data Definition Language）によるデータ要素の特性の定義，再定義，表示，削除等の機能の説明，およびスキーマ，サブスキーマ，格納媒体の初期化等についても合わせて説明する。

② データインテグリティ

オンラインシステムの特徴のひとつにデータインテグリティがあるが，この適時性および正確性の保障機能について説明する。

③ 排他機能

オンラインシステムにおいては，データの更新中に同一データに対するアクセスを排除するようになっているが，この機能について説明する。たとえば，山田太郎が預金の引き出しをしている時に給与の振込があった場合，排他制御が行われていないと預金額が正しく更新されない場合が出てくる。

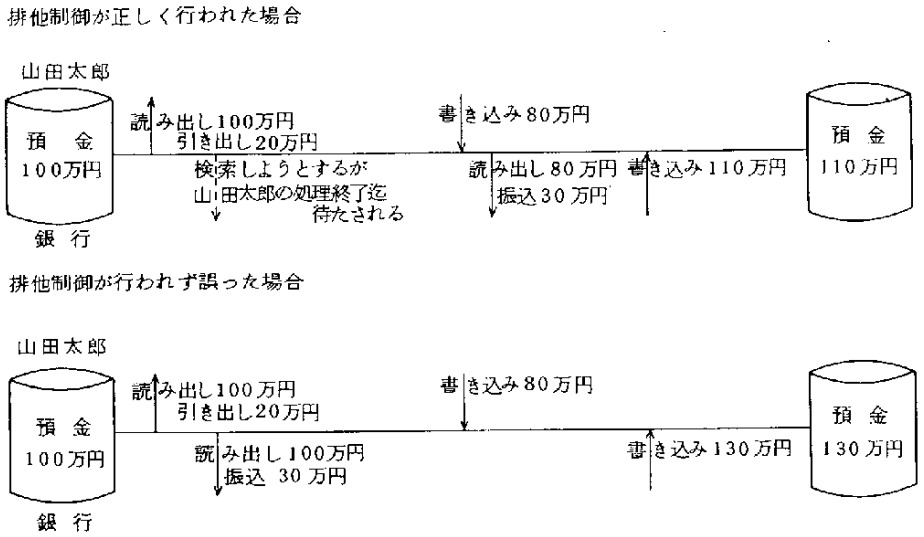


図2.1 排他処理の説明

また，排他制御の方法については，データベースのアクセス時にＤＢＭＳが自動的に領域を占有する方式と，最初のデータベースのアクセス時に応用プログラムがオープンしている全エクステントを占有する方式があることも説明する。

・自動的に領域を占有する方式では，データベースのアクセス時にＤＢＭＳが自動的に領域を占有するが，トランザクション終了時は，ＤＢＭＳが自動的に領域の占有を解除する。

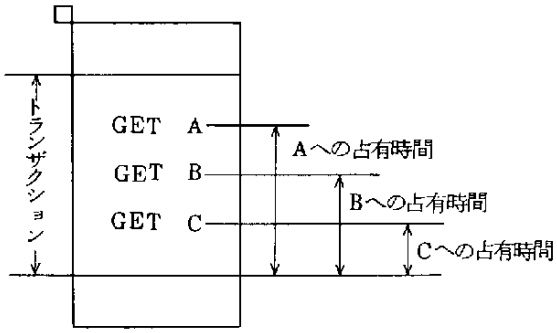


図2.2　自動的に領域を占有する方式

・全エクステントを一括して占有する方式では，該当トランザクションで最初のデータベースアクセス時に応用プログラムがオープンしている全エクステントを占有し，トランザクション終了時に全エクステントを解除する。

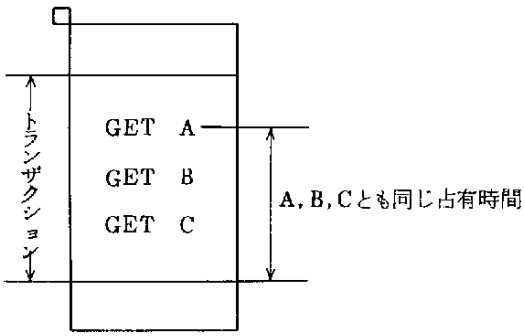


図2.3　全ｴｸｽﾃﾝﾄを一括して占有する方式

④　デッドロックの自動解消機能

複数のユーザにデータベースへの同時アクセスを許可しているとデッドロック状況が発生する場合があるが，オンラインのＤＢＭＳにはデッドロックを自動的に検出し，解消する機能が必要であることを説明する。

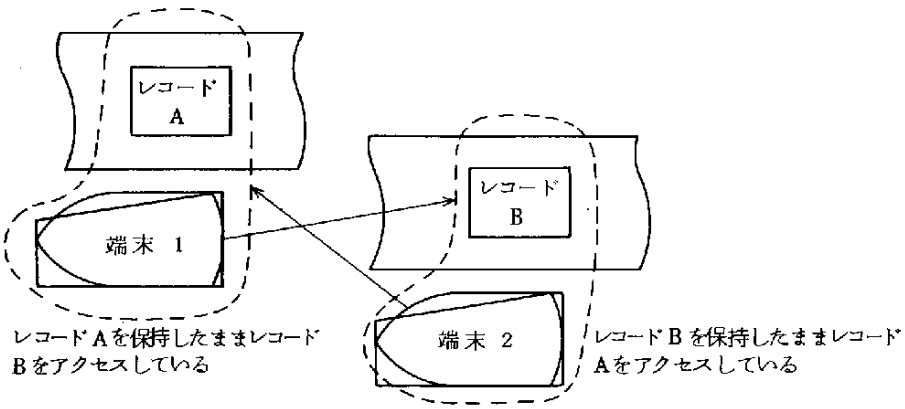


図2.4　デッドロック状況

⑤　分散データベース

分散データ処理には処理の分散とデータの分散とがあるが，分散データベースの理論と技術は，他のデータベースの概念と比べると初期段階のものである点を考慮して説明する必要がある。

分散データベースを定義すると，パーティションに区分されたデータベースで，しかもデータベースの制御に複数のＤＢＭ（データベースマネージャ）が関与していなければならない。なお，パーティション化方式には，垂直パーティション化方式（全レコード階層が各パーティションに入っている）と水平パーティション化方式（レコード階層の一部だけが各パーティションに入っている）とがある。

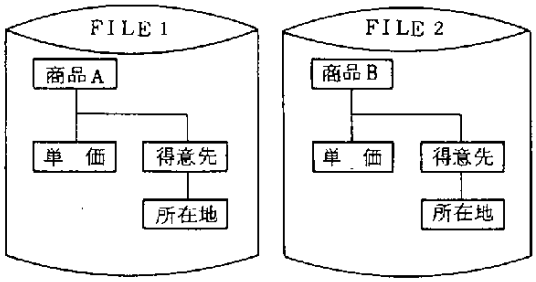


図2.5　階層データベースの垂直パーティション化方式

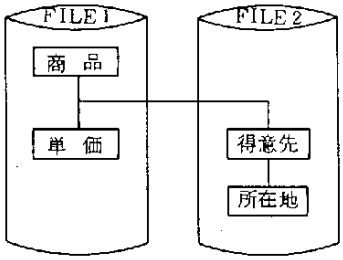


図2.6　階層データベースの水平パーティション化方式

また，データを分散することによるメリットについても以下の観点から説明する。

ア．アプリケーション処理効率（データを使用される場所に配置できる）

イ．データアベラビリティ向上（ノードの一部が故障しても，全部または一部を除き，データベースを使用できる）

ウ．組織上の柔軟性の確保（データ処理上の理由で，データ資源の集中化を図る必要がなくなる）

エ．データ精度の向上（パーティションを入力／利用部門の近くに置けるので精度が向上する）

オ．システム拡張の簡易化（ノードをネットワークに追加しやすい）

カ．パフォーマンスの改善（入力／利用部門に近いため応答時間の短縮が図れる）

キ．処理能力の増大（単一データベースと比べて大量の仕事をこなせる）

ク．ネットワーク全域へのアクセス（いかなるデータにもネットワーク内の所在に関係なくアクセスできる）

(2) データベース利用の機能

① データアクセス

利用者とＤＢＭＳとのインタフェースをとるのが操作言語（ＤＭＬ，Data Base Manipulation Language)であり，ＤＭＬは一般的に１問１答方式でデータベース処理を行っている。

ここでは，その代表的な機能であるデータの検索，挿入，削除，追加などの更新機能のほか，ＤＭＬを用いた各種テーブルやインデックス等の定義，変更，削除などの機能についても合わせて説明する。

② データディクショナリ／ディレクトリ（ＤＤ／Ｄ）

メタ・データ（データ項目の名称，長さ，属性等データ自身に関するデータ記述）の貯蔵所であるＤＤ／Ｄの目的，およびＤＢＭＳとの関連について説明する。

1.2 性能設計

① 応答時間

応答時間とは，ＤＢＭＳ環境下でシステムがユーザ要求に応えるのに要する時間である。この時間の相当部分は，アプリケーション・プログラムの出すデータベース呼び出しのサービスに費やされているので，ここを短くすることが応答時間の短縮につながることを説明する。

ホスト常駐ＤＢＭＳの応答時間＝ａ＋ｂ＋ｃ＋ｄ（ef＋gh＋i)である。

ａ：ネットワーク通過時間

ｂ：通信管理・メッセージ制御のための時間

ｃ：アプリケーション・プログラム実行時間

ｄ：ＤＢＭＳに対するデータベース要求回数

d：DBMSに対するデータベース要求回数

e：DBMS要求当りの平均I／O回数

f：平均I／O実行時間

g：DBMS要求当りのレコード待ち回数

h：平均レコード待ち時間

i：DBMS要求当りのDBMS・データ管理のための時間

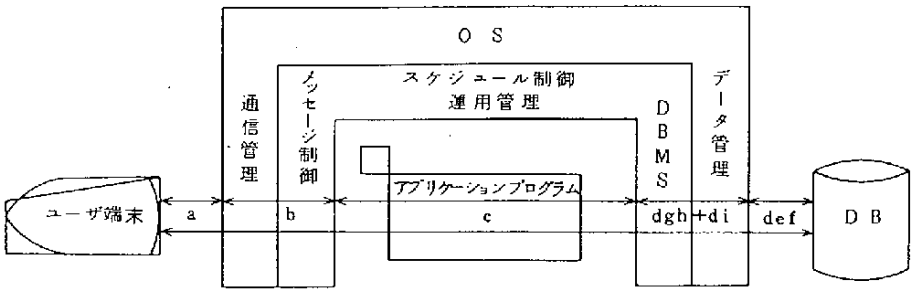


図2.7　応　答　時　間

② 容　量

データベース環境の成否はデータの編成方法とアクセス方法に大きく左右されるが，ここではデータを直接アクセス記憶装置に格納する方法について説明する。

格納方法には，大きく分けるとランダム化方式とインデックス方式とがある。そこで，この2つの方法について説明するとともに，オーバー・フロー領域や再編成（再配置，再構成）についても合わせて説明する。

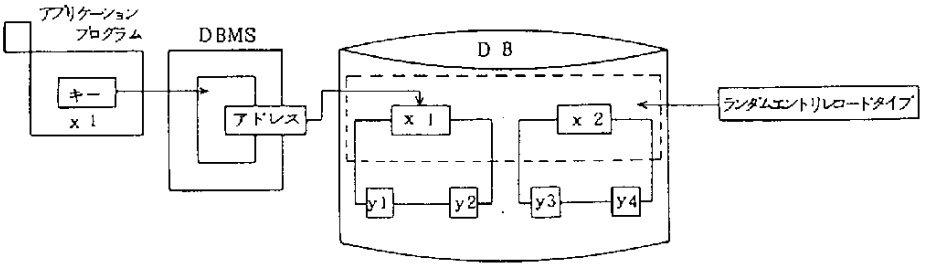


図2.8　ランダムエントリ

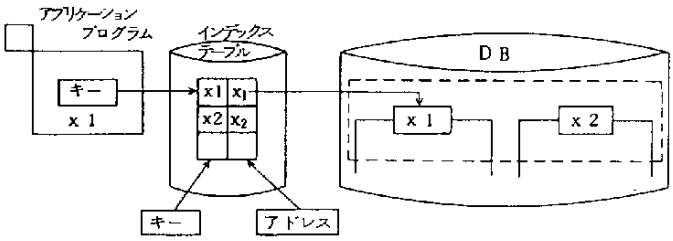


図2.9　インデックスエントリ

③　取扱量

単位時間当りの取扱量を把握することは，データベースの格納方法決定，応答時間算出の重要な要素であることを説明する。

1.3　データベースの設計

データベースの設計に当っては，適用業務を分析し，業務からシステムの機能設計を行う。ここでは，これらの分析が完了した後，実際にデータベース化するのに必要な論理構造設計，格納構造設計，物理構造設計について説明する。

①　論理設計

レコードの種類と内容，レコードタイプ間の関係（主従，対等），データベースへのエントリ方法（インデックス，拡張インデックス，ランダム，ダイレクト，システムエントリ）および効率などについて説明する。

②　格納構造設計

データの配置の決定に当っては，直接ファイルに割り当てず，１つまたは複数の格納空間を設定し，これにレコードタイプを配置することを説明する。このとき，レコードを格納できない場合に備えたオーバーフロー領域や，１つの格納空間内に配置されたレコードタイプ群を運用上のまとまりで分割する考え方についても合わせて説明する。

③　物理構造の設計

格納空間へのデータの配置が決まると，次に格納空間を物理的な直接アクセス外部記憶装置に割り当てることになるが，これに先立って，ページ

サイズの決定や記憶媒体の選定（空きがあるかどうかを事前に調べる）が必要であることを説明する。

1.4　データベースの作成

データベースの作成に当っては，①～③を経て，以下に示すデータベース作成へとつながることを説明する。

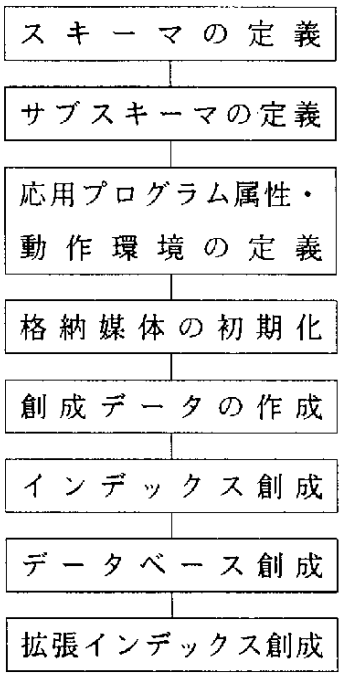


図2.10　ベータベース作成手順

① データモデル

ア．階層型モデル

階層型モデルの特徴は，親(parent)のレコード(segment) と子(child)のレコードの関係がＩ：Ｎに限っていること，すなわち，１つの子レコードは複数の親レコードと関係を持てないことである。しかし，現実には１つの子レコードが複数の親レコードと関係を持たねばならない場合もありうるので，この回避方法（複数の木（tree))に分割したデータ構造を設定することで回避できる）についても合わせて理解させる。

イ．ネットワーク型モデル

ネットワーク型モデルの特徴は，レコード間の関係がM：Nの場合でもそのまま表現できるデータモデルであるため，現実の事象を実態に近い形で表現できることを理解させる。

なお，ネットワーク型モデルの説明では，ＣＯＤＡＳＹＬ　ＤＢＴＧを例にその構造を理解させるとよい。

ウ．関係型モデル（リレーショナルモデル）

階層型モデルやネットワーク型モデルのいずれを使用するにしても，目的とするデータを得るためにはレコード間の親子関係を知っておく必要があった。Ｅ．Ｆ．Ｃｏｄｄによって提案された関係型モデルは，レコード間の関係を意識する必要がないことを以下の観点から説明し，理解させる。

a．データはリレーション（relation）と呼ばれる表形式で表現できることを理解させる。また，ユーザはあらかじめ設定されているデータベースの構造に拘束されずに，ユーザ自身で複数のリレーショナルにまたがるビュウ（view）と呼ばれる外部スキーマの設定と，それに基づくデータベースの操作が可能であることを理解させる。

b．関係型モデルをＥ－Ｒ（事象－関係性）モデルで表現する方法を理解させるとともに，さらにこれを正規化（normalization)するプロセスも理解させる。

c．正規化したリレーションに対しては，選択（selection)，制約(restriction)，射影（projection)，結合(join)，削除（division）といった関係演算の意味を理解させる。なお，正規化の概念には，すべての属性の値が原子的である第１正規形（First Normal Form)の他に，第２正規形（２ＮＦ）から第５正規形（５ＮＦ）まであることに言及できることが好ましい。

② アドレス指定

データを探し出すうえで最も基本的なことであるアドレス指定には，直接アドレス指定と記号アドレス指定があることを説明し，その方法についても説明する。

1.5　システム環境整備

(1)　ハードウェア

本体（ＣＰＵ，メモリ），直接アクセスファイル（ドラム，ディスク，光ディスク），チャンネルについては，ＤＢＭＳを有効に運用するのに必要な速度や容量などについて説明する。

(2)　ソフトウェア

ＯＳ（ジョブ管理，タスク管理，仮想記憶制御），通信管理，データ管理，メッセージ制御，環境，運用管理，スケジュール管理などとＤＢＭＳの関連について説明する。

1.6　ユーザインタフェース設計

(1)　検　索

検索についてはコマンド方式，メニュー方式などがあるので，それぞれの特徴を説明する。また，検索において，中間結果をどのレベルまで残しておくかが検索速度や渡り検索に重大な影響があることも，合わせて説明する。

(2)　加工・編集

数値データの加工や予測，グラフ化，表出力，計算機能のほか，日本語での情報処理も端末がパソコン化しているため，編集が容易であるということを説明する（本体でこれらの処理は可能であるので，効率を考慮してホストでするか端末でするかを決定する必要がある）。また，データのダウンロードやアップロードによる機能の負荷分散についても合わせて説明する。

(3)　ホスト言語

アプリケーション・プログラムの作成に当っては，ＣＯＢＯＬ，ＦＯＲＴＲＡＮ，ＰＬ／Ｉ，アセンブラなどで作成できるが，データベースのアクセスは，各言語ごとに異なったデータベース操作言語で記述されることを説明する。また，アプリケーション・プログラムの登録に当っては，このプログラムの属性や動作環境を定義する必要があることも説明する。

(4) ユーティリティ

ユーティリティには物理レベル，システムレベル，アプリケーションレベルの3つのレベルのものがあることを説明する。物理レベルのユーティリティについては，ビットやバイト単位のレベルの処理をするので，ＤＢＭＳの変化に対応しにくいことを説明する。また，システムレベルのユーティリティは，一般に言われている“ユーザ出口”を利用したものであればＤＢＭＳの変化にある程度対応できること，さらに，アプリケーションレベルのユーティリティならばＤＢＭＳの変化でも影響がないことなども合わせて説明する。

(5) マルチメディア

イメージ，映像，音声，ＦＡＸなどの情報を組み合わせた対応も必要となるので，これについても説明する。

(6) 出力端末制御

出力端末装置は物理的に大きさに制限があるので，分割して出力しなければならない場合があることを説明し，これらに関するアプリケーション・プログラムの負担の軽減を図る必要があることも合わせて説明する。

(7) ユーザサポート

コマンドの入力方法を教えてくれるＨＥＬＰ機能，データベースが多数ある場合の内容を紹介してくれるクリアリング機能などについても説明する。

1.7 高信頼化機能設計

(1) ＤＢセキュリティ

ＤＢ利用に関する障害には，アプリケーション・プログラム障害，システム障害，データ障害の3種類があることを説明する。

(2) リカバリ

アプリケーション・プログラム障害，システム障害（本体，回線，端末），データ障害について，それぞれのリカバリ方法を説明する。

① アプリケーション・プログラム障害

アプリケーション・プログラム実行中の異常に対しては，トランザクション単位に一括して行っているので，トランザクション開始前の状態に復旧することを説明する。

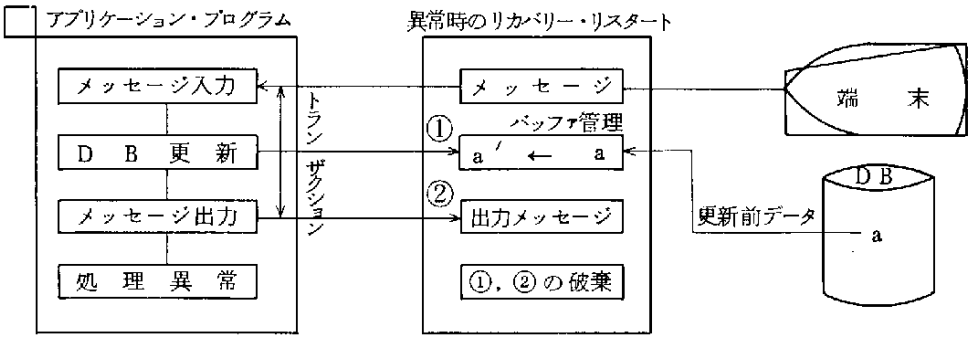


図2.11 アプリケーション・プログラム処理異常時のリカバリー・リスタート

② システム障害

ダウンリカバリは，システムが再開された後にアプリケーション・プログラムの起動と並行して行うが，トランザクションについては終了処理中のものは終了時の状態に，実行途中のものは開始前の状態にすることを説明する。また，システムダウン時に処理中の入出力メッセージは無効となることも説明する。

なお，ログ・ファイルはデータベースの更新部分のデータを時系列的に記録したもので，ビフォア・イメージとアフタ・イメージとがあることも合わせて説明する。

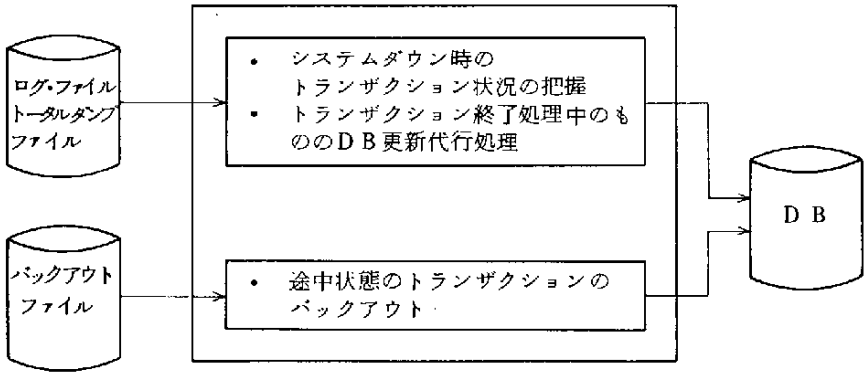


図2.12　システムダウン時のリカバリー・リスタート

③　データ障害

データセットリカバリは，データの内容を退避した全ダンプデータと，それ以降の更新データを使用しての再現処理が原則であることを説明する。

(3)　ユーザ・チェック

データベース使用に当っては，端末使用，プログラム使用，ＤＢ参照，ＤＢ更新およびＤＢの構造の更新等について，これを実施しようとした者の確認・許可の方法について説明する。

(4)　２重化

システムの障害処理と合わせて，障害が起こりにくい構造，たとえば本体の複数化，回線の２重化，ファイルの２重化などについても，コストパフォーマンスを考慮した説明を行う。

(5)　データの独立性

種々の蓄積データの変更に対して，それを利用しているプログラムが影響を受けないよう設計することが保守時の費用低減に重要であることを説明する。

1.8　運用管理設計

(1)　モニタリング

システムの運用状況を記憶するために収集している情報には，大きく分けてアプリケーション・プログラムごとの情報とトランザクションごとの情報とがあり，別個に管理する必要があることを説明する。

なお，トランザクションごとの情報は，いわゆるデータベース利用に関する課金情報なので，ここではアプリケーション・プログラムごとの情報を，さらに資源の負荷状況分析（チューニングのための性能分析用情報）とアプリケーション・プログラムによる各種資源の使用量の把握，および異常事態発生状況の把握とに分けて説明する。

①　資源の負荷状況

資源の負荷状況には，利用待ち行列の発生状況，資源の負荷状況，エクステントの競合状況，各種バッファの使用状況等を把握することがＤＢＭＳのチューニングに必要であることを説明する。

②　アプリケーション・プログラムによる各種資源の使用量の把握，および異常事態発生状況把握

アプリケーション・プログラムを特定ユーザが利用する場合（たとえば，総務部が人事情報ＤＢのユーザである場合）には，トランザクション単位の情報の把握は不要であるが，課金のためには，当該アプリケーション・プログラムに関するトランザクション件数やデータベースなどの利用に関する情報が必要であることを説明する。また，アプリケーション・プログラム単位に起こる特異事象発生情報（アラーム情報，デッドロック事象等の情報）も，ＤＢＭＳのチューニング情報として必要であることも説明する。

(2)　課　金

オンラインシステムにおいてはトランザクションごとにシステムのユーザが異なる場合が多いので，このようなことが考えられるオンライン・システムでの課金情報はトランザクションごとに必要であることを説明する。トランザクションごとの情報収集では，トランザクション処理時間，データベースの使用量などが課金情報として必要なので，これらの情報収集が

必要なことを説明すると同時に，トランザクション処理状況の把握がＤＢＭＳのチューニング情報として必要であることも合わせて説明する。

(3) システム運用の省力化

人手による作業の軽減や操作ミスの減少を図ることを目的として，システムの自動化運用の必要があることを説明し，これらの機能についても説明する。

① システムの利用状況・ダウン状況の把握機能

② アプリケーション・プログラムの利用状況の把握機能

③ ＤＢ利用状況・障害状況の把握機能

④ 端末の利用状況の把握機能

⑤ システムの立ち上げ，停止機能

⑥ 障害復旧機能

⑦ 既存データを利用したデータベース創成機能

(4) データベースの生成・再編成

運用段階で最も厳しい作業はデータベースの生成あるいは再編成である。これは全データが対象となるからである。しかし，処理効率，スペース効率を考えると，再編成は必ず行わねばならない作業である。したがって，ここではデータベースの生成，再編成の必要性および考慮すべき事項について説明する。

1.9 開発サポート

(1) テスト支援機能

アプリケーション・プログラムの開発は，モジュールテスト，総合テスト，システムテストといった3段階を経た開発が一般的であることを説明する。また，オンライン・システム開発に当っては十分な環境や大量のデータが用意されてないことが多く，これらの各段階でテストの支援をするためには，オンライン・システムをバッチで動かしたり，テスト・データを用意するシステムが必要であること，アプリケーション・プログラム開発に関する情報の一元管理が必要であることを説明する。

(2) 画面フォマット作成支援機能

アプリケーション・プログラムの開発に当っては，画面フォーマットの作成が必要であり，かなりの労力が必要であるので，画面フォーマットが作成できる支援システムが必要であることを説明する。

指導上の留意点

データベースシステムの構築に当っては，ＯＳ機能，ネットワーク構成，ファイルの概念が関連知識として必要なので，これについても説明する。

用　　語

ＤＤＬ，ＤＭＬ，スキーマ，サブスキーマ，データインテグリティ，
排他制御，トランザクション，エクステント，デッドロック，
分散データベース，ＤＢＭ，処理の分散，データの分散，
垂直パーティション化方式，水平パーティション化方式，アベラビリティ，
パフォーマンス，ネットワーク，テーブル，インデックス，
オペレーティングシステム，通信制御装置，回線，伝送手順，端末，
会話処理，集信処理，配信処理，交換処理，通信管理，メッセージ制御，
データ管理，ＤＢＭＳ，データベース，端末ソフト，同期通信，非同期通信，
ＤＤ／Ｄ，応答時間，直接アクセス記憶装置，格納方法，ランダム化方式，
インデックス方式，オーバーフロー領域，再編成，再配置，再構築，端末数，
待ち時間，回線数，回線速度，ＣＰＵタイム，データ編成方法，アクセス方法，
論理構造，格納構造，処理構造，エントリ方式，拡張インデックス，ダイレクト，
システムエントリ，格納空間，レコードタイプ，ページサイズ，階層モデル，
Ｅ－Ｒモデル，ネットワークモデル，リレーショナルモデル，親，セグメント，
子，木，事象，ＣＯＤＡＳＹＬ　ＤＢＴＧ，Ｅ．Ｆ．Ｃｏｄｄ，リレーション，
ビュウ，正規化，選択，制約，射影，結合，削除，関係演算，第１正規形，
第５正規形，アドレス指定，直接アドレス指定，記号アドレス指定，光ディスク，
チャネル，ファクシミリ，イメージスキャナ，ジョブ管理，タスク管理，
仮想記憶制御，環境管理，運用管理，スケジュール管理，

オンラインデータベース，ＤＤＸ，電話網，特定回線，伝送制御手順，
端末制御手順，タッチパネル，渡り検索，ダウンロード，アップロード，
マルチメディア，イメージ，映像，画像，音声，出力端末制御，ユーザサポート，
ＨＥＬＰ機能，クリアリング機能，データベースセキュリティ，リカバリ，
リカバリ・リスタート，トランザクション単位，
アプリケーション・プログラム障害，システム障害，データ障害，
システムダウン，バックアウト，ログファイル，全ダンプデータ，
ユーザチェック，２重化，コストパフォーマンス，モニタリング，課金情報，
チューニング，資源の集荷状況，使用量把握，異常事態発生状況把握，
モジュールテスト，結合テスト，システムテスト

2. ＤＢＭＳの選択

目　　標

データベースシステムの構築に当っては，要求定義を明確化したうえで，既存のＤＢＭＳを評価選定してデータベースシステムの構築を行う。ここでは，既存のＤＢＭＳを選択するために必要な評価知識や技術を修得させることを目標とする。

内　　容

2.1　業務への適合性

(1)　ＤＢＭＳの制限

ＤＢＭＳの制限事項として，データ量，アイテム数，アイテム長，キーの数，文字の種類（含図形）などを主とした量に関する制限があった場合，システムを構築するうえで十分であるか。

(2)　データモデル

現在，構築しようとしているデータベースにふさわしい構造であるか。

2.2　機能・性能

(1)　ＤＢＭＳ利用に関する制限

予定しているコンピュータ，ＯＳ，メモリーサイズ，直接アクセスファイル数量で使用可能であるか。

(2)　応答時間

予定しているコンピュータ，ＯＳ，メモリーサイズ，直接アクセスファイル，回線（速度，数）端末数で予定通りの応答時間に納まるか。

(3)　ユーザ・インタフェース

エンドユーザが満足できるシステム（マルチメディア対応も含め）が構築できるか。

2.3　関連ソフトウェア・ツール

(1)　アプリケーション開発機能

言語，エディタなどのアプリケーション開発のサポート機能は十分か。

(2)　利用可能ソフトウェア

ユーティリティ，アプリケーション，パッケージなど利用したい品揃えが十分か。

2.4　他のシステムとの結合性

(1)　データ処理変更への柔軟性

既存システムからの移行性や開発後の変更に対する柔軟性は十分か。

(2)　データディクショナリ

データディクショナリの有無，および市販されているデータディクショナリの流用が可能か。

(3)　他のＤＢＭＳの通信機能

システム間通信機能が備わっているか。

(4)　異種ＤＢＭＳとのインタフェース

異種ＤＢＭＳとのインタフェースが容易にとれるか。

2.5　信頼性対策

(1)　セキュリティ

アプリケーション・プログラム障害，システム障害，データ障害に対するリカバリは機能のほか，時間的にも満足できるものか。

(2)　機密保護

端末使用，ＤＢ参照，ＤＢ更新，アプリケーション・プログラム使用などに対する確認・許可の方法は適切か。

(3)　２重化

本体，回線，ファイル等障害が起こりにくいような構成にできるか。

2.6　その他

(1)　ドキュメンテーション

導入しようとしているＤＢＭＳのドキュメンテーションは十分か。

(2)　ＤＢＭＳのサポート要員

ＤＢＭＳの導入に十分なサポート要員がいるのか。

(3)　使用料

ＤＢＭＳの使用料は妥当なものか。

(4)　教育体制

ユーザへの教育体制は整っているのか。

指導上の留意点

ＤＢＭＳの選定に当っては，データベース・システム構築の要求定義能力のほか，オンラインデータベース・システムの構築に必要なＯＳやネットワークの知識が伴わないと，評価項目の抽出や評価自体ができないので，これらについても指導することが望ましい。

用　　語

データ量，アイテム数，アイテム長，キーの数，文字種，データ構造，ＯＳ，メモリサイズ，直径アクセスファイル，応答時間，回線数，回線速度，端末数，エンドユーザ，ヒューマン・インタフェース，マルチメディア，アプリケーション，開発サポート機能，ユーティリティ，アプリケーション・パッケージ，結合性，柔軟性，移行性，データディクショナリ，ＤＢＭＳ間通信機能，セキュリティ，アプリケーション・プログラム障害，システム障害，データ障害，リカバリ，ユーザ・チェック，端末使用，ＤＢ参照，ＤＢ更新，アプリケーション・プログラム使用，ファイル２重化，回線２重化，本体２重化，ドキュメンテーション，サポート要員，使用料，教育体制

3. ＤＢＭＳに適した効率化

目　　標

データベースシステムの構築に当っては，ＤＢＭＳなどの改善が必要な場合もある。ここでは，ユーザができる範囲のＤＢＭＳの改善方法，とくにチューニングについて知識，技術の修得を目標とする。

内　　容

ＤＢＭＳの性能は，具体的アプリケーションとＤＢＭＳの適合度合によって大きく変動する。総合的な調整の結果として，ＤＢＭＳ自体が備えている各種のパラメータ，ＯＳ側のパラメータやファイルの配置など，最高の性能が発揮される。こうしたユーザ側での改良は，エンドユーザからの要望のほか，各種モニタリング・システムからの情報に基づいて行うことができることを説明する。また，データベース・システムの改善は「調査」の範囲に留めた方がよく，さらに大きな効果が得られる場合が多いことも合わせて説明する。

3.1　ハードウェア

エンドユーザからの要望やモニタリング・システムの負荷情報により，不足している端末，ファイル，メモリなどを知ることができるので，メーカのサポートの範囲内でコストパフォーマンスを考慮した増設，再配置などで効率化を図る。

3.2　ＯＳ

ＤＢＭＳでは，メモリやディスクの割当，使用を全面的にＯＳにまかせているものや，ＯＳの基本機能だけを用いてかなりの部分をＤＢＭＳ独自に管理するものなどさまざまである。いずれにしても，ＤＢＭＳの性能向上はタスクの優先度，その他ＯＳ側のパラメータ設定に依存する部分も多く，システム・ジェネレーションから再検討する必要もある。

3.3　基本ソフトウェア

ＤＢＭＳには“ユーザ出口”（ベンダがユーザに基本ソフトの一部を開放している窓口）があるので，ここにアプリケーションに適合したルーチンを作成して組み込み，効率化を図る。

3.4　アプリケーション・プログラム

エンドユーザからの要望やモニタリングによって得られたＩ／Ｏ回数，ＣＰＵタイム，ＤＢ呼び出し回数などから，効率の悪いアプリケーション・プログラムは改良を図る必要がある。一般的には，次の５点に注意することでデータベースの効率化，すなわち，最も処理時間のかかるＩ／Ｏ回数の減少とスペース効率を上げることが可能である。

(1)　データベースを構成する各ファイルの使用状況に応じて，ディスク，ＩＣディスク，メモリ・ディスクなど，記憶装置の階層への適切な配置を行う。

(2)　アプリケーション・プログラムによるファイルのブロック化で，ファイルのアクセス回数を減らす。

①　データの発生規則（データ発生件数が規則的である場合は，アプリケーションプログラムでデータのブロック化を図った方がよいが，変動がある場合はＤＢＭＳにブロック化させた方が効率がよい．）

②　データの更新パターン（データの更新パターンが規則的である場合は，アプリケーション・プログラムでデータのブロック化を図った方がよいが，ランダムの場合はＤＢＭＳにブロック化させた方が効率がよい．）

③　データのアクセスパターン（データのアクセスが規則的である場合は，アプリケーション・プログラムでデータのブロック化を図った方が効率がよいが，ランダムにアクセスされるならばＤＢＭＳにブロック化させた方が効率がよい．）

④　データのサイズ（データのサイズが小さければ，アプリケーション・プログラムでデータのブロック化を図った方が効率がよいが，大きければＤＢＭＳにデータをブロック化させた方が効率がよい．）

(3)　頻繁にアクセスされる可能性が高いデータは，オーバーフローエリアに

入る可能性が最も少ない場所で構造化を図る。

(4) ブロック長，レコード長，バッファ数，オーバーフローエリアなどのＤＢの物理的なサイズの適正化を図る。

(5) ＤＢの再編成，ガーベージコレクションを適時に行うことで，処理効率とスペース効率が向上することが多い。

指導上の留意点

ここでは，アプリケーション・プログラムの作成の仕方で効果の上がる方法を具体的に説明する。

用　　語

モニタリング・システム，負荷情報，端末，ファイル，メモリ，コストパフォーマンス，ＯＳ，パラメータ，システム・ジェネレーション，ユーザ出口，基本ソフト，アプリケーション・プログラム，ＩＣディスク，メモリ・ディスク，I/O 回数，ＣＰＵタイム，ＤＢ呼び出し回数，スペース効率，ファイルのブロック化，データの発生規則，データの更新パターン，ランダム，アクセスパターン，データのサイズ，オーバーフローエリア，ブロック長，レコード長，バッファ数，物理的サイズ，再編成，処理効率，ガベージコレクション

第 3 章　性能評価

目　　標

構築する，または構築されたデータベースシステムが，性能，要件を満足するか否かの評価を行うための知識の修得を行う。

内　　容

1..1　性能評価の手順

データベースに対する性能評価の手順は次のようになると考えられる。

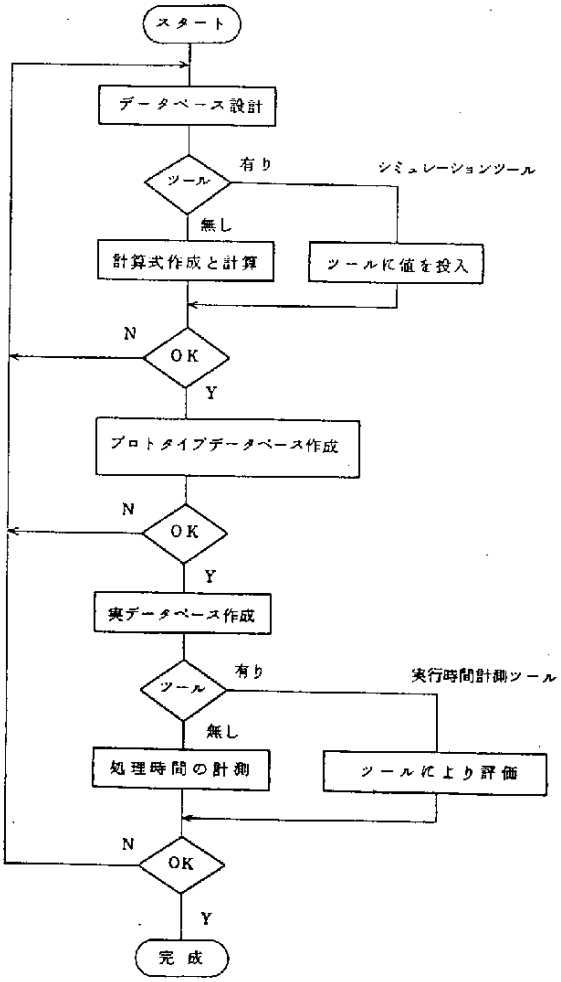


図2.13　性能評価手順

1.2　シミュレーション

実際にデータベースを構築しないで，目的とするデータベースの性能の予測を行う。そのためのツールがあれば，データベースに対してパラメータを与えるだけでよいが，もしツールがなければ計算式を作成しなければならない。

・計算式の作成

性能を評価するために必要な値は処理経過時間である。処理経過時間はＩ／Ｏ時間とＣＰＵ時間の和であり，データベース管理システムのそれぞれの機能を利用した場合の数式をたてる。数式はＩ／Ｏ回数から求められる。ＣＰＵ時間は実験的に求められた値より数式を立てる。たとえば，インデックスレベルが１であるデータの検索時間は次のような数式になる。

$T = l \times (i + c)$

l：インデックスレベル

i：Ｉ／Ｏ時間

c：ＣＰＵ時間

このようにして各機能単位に数式を作成し，処理時間の算定を行う。この値はシングルジョブでの環境の値であるため，さらにシステムの混雑度や資源の共有度によってその値の補正を行う。

1.3　ベンチマークテスト

計算式やシミュレーションツールによって性能を求めるほかに，実際のデータの一部を利用して性能テストを行う，いわゆるベンチマークテストによる性能評価方法がある。

より高い精度を求めるために，実際の目的システムと同様のプログラムを作成してベンチマークテストを行った方がよいが，一般的にはデータベースユーティリティプログラムやエンドユーザ言語を利用する。たとえば，応答時間に関しては，検索条件を端末から投入し，答えが返ってくるまでの時間をストップウオッチにて計測する方法がある。

1.4　実行時間計測ツール

実際のデータベースが完成した時点で，実行時間よりデータベースの性能評価を行うツールについて理解する。実行時間計測ツールはほとんどデータベース管理システムと一体となっているか，またはオペレーティングシステムと一体となっている。その計測内容の主なるものには次のものがある。

(1)　データベース毎のＩ／Ｏ回数

(2)　各機能単位のＩ／Ｏ回数（時間）とＣＰＵ時間

(3)　Ｉ／Ｏ要求に対する実Ｉ／Ｏ回数（バッファ効果の評価）

(4)　キー部分に対するＩ／Ｏ回数

(5)　ファイル演算に対するＩ／Ｏ回数とワークＩ／Ｏ回数

(6)　データベースコマンドの利用回数

(7)　データベースディスク領域の利用状況　　等

指導上の留意点

性能評価のポイント，ツールの使い方，ツールの作り方など，データベースの性能評価技術を修得させる。

用　　語

性能要件，ツール，シミュレーション・ツール，プロトタイプ，データベース，実データベース，実行時間計測ツール，処理経過時間，Ｉ／Ｏ時間，ＣＰＵ時間，Ｉ／Ｏ回数，インデックスレベル，シングルジョブ，混雑度，ベンチマークテスト，オペレーティングシステム，バッファ効果の評価，ファイル演算，ワークＩ／Ｏ回数，データベースディスク領域

参考文献

最新データベース管理システム（日経マグロウヒル社）

ＡＩＭ概説書（富士通）

データベース言語ＮＤＬ（ＪＩＳＸ3004）

データベース言語ＳＱＬ（ＪＩＳＸ3005）

最近のデータベース理論（日本ソフトウェア科学会）

第3部　データベースの構築

第3部　データベースの構築

目　　標

データベースの構築は，対象とするデータベースの種類により，その入力手段，ファイル構成，利用技術などが著しく異なっている。ここでは，その代表的なものとして，文献型データベース，数値型データベースおよび画像型データベースの3つを取り上げ，ソースデータの収集，入力データの作成，データの加工，データファイルの構造化，データの入力，流通データの導入について説明し，理解させる。なお，数値型データベースについては，商用データベースで利用の多い経済データを例にとって説明する。

内容のあらまし

内　　容	時間数
ソースデータの収集 　資料の種類 　収集方法 　収集形態 　データの分析評価	3時間
入力データの作成 　文献型データ 　ファクト型データ	3時間
データの加工 　統一処理 　変換処理 　圧縮処理 　流通媒体の作成	3時間

内　　容	時間数
データ・ファイルの構造化 　構　造 　利　用 　メンテナンス性 　セキュリティ	6時間
データの入出力 　コード入力 　イメージ入力 　ベクトル入力 　入力データの検査 　出　力	6時間
流通データの導入 　予備調査 　選定・契約 　内容調査 　加工処理 　システムへの組み入れ	6時間
企業内データベースの構築 　予備調査 　データ入力	3時間
合　　計	30時間

第1章　ソースデータの収集

目　　標

データベースのためのソースデータは，販売データ，在庫データなど企業内・組織内で発生するものと，文献データ，統計データなど外部で発生するものとに分かれる。前者のための収集方式は，一般のアプリケーション・システム構築において扱われるが，後者はデータベースシステム固有の問題となる。この章では，データベースを構築する際，その基礎となるデータを収集する実務について説明するが，文献型データベースと数値型データベースとではソースデータの種類や収集方法がかなり異なっているので，両者別々に説明する。

内　　容

1.1　資料の種類

(1)　文献型データベース

文献型データベースで対象とする資料は，その使用目的や内容によって種々の形態のものが存在している。

主要な資料を形態別に列記すると次のようになる。

①図書　②雑誌　③テクニカルレポート　④会議資料　⑤学位論文　⑥特許資料　⑦新聞　⑧政府刊行物　⑨規格　⑩カタログ（製品取扱説明書）　⑪視聴覚資料

また，これらの資料の中で連続的に発行されるものは逐次刊行物と呼ばれているが，逐次刊行物のカテゴリーに属する資料には次のようなものがある。

① 定期刊行物（新聞，雑誌）

② 不定期刊行物（各種研究機関の報告書，紀要，特許公報など）

③ 年報・年鑑類（各年刊なども含む）

④ 定期的に行われる会議の議事録，予稿集

⑤ モノグラフシリーズ（同一の総合タイトルのもとに刊行される一連の単行本）

(2) ファクト型データベース

ファクト型データベースのうち，経済データに関する主なものには次のようなものがある。

① リアルタイムな情報

株式，債券など，時々刻々取引価格が変動する情報

② 統計情報

統計を作成する諸官庁，地方公共団体，日銀，業界団体が提供する情報

③ 企業情報

財務データ，決算書，決算短信，営業報告書，有価証券報告書など

1.2 収集方法

文献型データベースの資料の入手は購入によるものが大部分であるから，場合によってはかなりの金額を要する。また，入手した資料は利用しやすいように整理・保管しなければならないので，そのためのスペースも必要である。このように，資料は予算と情報部門の人員および施設に制約されることが多い。なお，資料の収集，整理・保管には細々とした事務が付随するので，これにコンピュータを利用して自動化している例も多い。

数値データベースでは，データの発生の現場，またはデータを作成する機関から直接入手するのが通例であり，そのデータの収集能力が数値型データベース業者の決め手ともなっている。

(1) リアルタイムな情報

株式，債券など，時々刻々取引価格が変動する情報に関しては，それぞれの取引所から直接オンラインでリアルタイムに入手する。

(2) 統計情報

統計データなどは，統計を作成する諸官庁，地方公共団体，日銀，業界団体から，直接，発表時点で定期的に軽印刷物で入手する。最近では，諸官庁が指定統計作成に当って電算処理を取り入れているため，磁気テープおよびフロッピーディスクによる提供も増えている。

(3) 企業財務情報

財務データに関しては各企業から直接入手するが，通常，非上場企業では，商法の定める営業報告書ベースの決算書，または税務申告用の決算書の写しを入手する。

上場企業では，取引所で発表する決算短信（速報ベース）や株主総会に発表する営業報告書，企業会計原則に基づいた有価証券報告書を企業から直接入手するか，証券取引所または大蔵省で閲覧して入手する方法がある。

1.3 収集形態

(1) 紙

(2) マイクロフィルム，マイクロフィッシュ

(3) 磁気記録媒体

① 磁気テープ

② フロッピーディスク

(4) 光記録媒体

① 光ディスク

② CD－ROM

③ CD－I

(5) 通信の利用

① オンライン

② ファクシミリ

③ テレックス

1.4 データの分析評価

(1) 雑 誌 類

雑誌類には原著論文，レビュー，抄録，紹介記事，広告記事等，性格の異なる論文や記事が掲載されているので，収集目的に応じて文献を取捨選択しなければならない。

(2) 特許文献

特許権の調査を目的としたデータベースの構築には，対象範囲の特許文献（特許公報）が網羅的に収集されていることが必要である。たとえ技術的に劣るものでも，権利文献としては重要な意味を持つので割愛してはならない。

指導上の留意点

ソースデータの収集では，文献型データとファクト型データとではかなり収集方法が異なり，とくに株価情報のようにリアルタイムでないと意味がないものがあるなど，データの特性による相違を説明する。

用　　語

文献型データベース，数値型データベース，ソースデータ，収集方法，図書，雑誌，テクニカルレポート，会議資料，学位論文，特許資料，新聞，政府刊行物，規格，カタログ，視聴覚資料，逐次刊行物，定期刊行物，不定期刊行物，年報・年鑑類，議事録，予稿集，モノグラフシリーズ，ファクト型データベース，リアルタイム，統計情報，企業情報，磁気テープ，フロッピーディスク，企業財務情報，収集形態，マイクロフィルム，マイクロフィッシュ，磁気記録媒体，光記録媒体，光ディスク，ＣＤ－ＲＯＭ，ＣＤ－Ｉ，ファクシミリ，テレックス，原著論文，レビュー，抄録，紹介記事，広告記事，特許文献，権利文献

第2章　入力データの作成

目　　標

入力データの作成に当って，文献型データ，ファクト型データのそれぞれについて，作成要領，インデキシング，辞書，データディクショナルなどの整備や作成方法について理解させ，実際にこれらの作業が行える能力を修得させる。

内　　容

1.1　文献型データ

収集された文献データを入力データの型にする過程では，記入用紙のフォーマット作成からアブストラクト化，インデキシング，辞書やデータディクショナリなどの選択や作成といった工程がある。

また，最近は大容量ファイルの出現やユーザニーズから，文献データもアブストラクトで提供するばかりでなく，全文で提供する業者が増加している。

1.2　ファクト型データ

収集されたファクトデータには，大別すると数値データと図型データとがあり，数値データの中には，さらに経済データ，企業データ，実験データなどがある。また図型データも，図面データの他に化学式のデータもあり，ファクト型データは扱う内容でさまざまのものがある。

このため，そのフォーマット化やファイル化に当って，媒体の選定にはユーザ端末の普及状況などを考慮して決める必要がある。

指導上の留意点

文献型データとファクト型データの特徴，商用データベース，企業内データベースの相違を簡潔に説明する。

用　　語

文献データ，ファクト型データ，インデキシング，辞書，
データディクショナリ，アブストラクト化，全文データベース，数値データ，
図型データ，経済データ，企業データ，実験データ，化学式，ユーザ端末，
商用データベース，企業内データベース

第3章　データの加工

目　　標

既に蓄積されたデータが，各々の使用目的によって再加工を要求されることがある。その目的の主なものとして，以下のものがあげられる。

① データの精度または内容をさらに高めるための統一処理
たとえば，通貨のドル換算，人名，会社名の統一，用語の表記法の統一，同義語または類似語の統一.

② 使用するシステムに適合させるための変換処理
たとえば，文字コードの変換，フォーマットの変換，企業コード，地域コードの追加や置換.

③ とくに，イメージデータの場合における圧縮，復元処理.

④ データを第3者に提供するための流通媒体の作製
たとえば，MT，フロッピー，光ディスクによるデータの提供または販売，CD－ROMによるデータベースの作製.

ここでは，データの加工の目的とその種類，およびその加工に用いられる技術または特殊な機器等について説明するとともに，その問題点を指摘して理解させる。

内　　容

1.1　統一処理

データを蓄積する場合には，当然，作成基準に基づいてデータシートを作成し，一定の入力方式で入力されるため，誤りのない完全なデータが蓄積されているのが普通である。だが，実際にはデータ作成基準の改変，とりわけ人間系に関係するデータ処理工程における個人差や精度の違いによって，エラーデータや不統一のデータが混入している場合が多い。また，自然語を扱うシステムにおいては用語の表記法上の不統一など，検索を行ううえで障害になる点も少なくない。

通常，データの統一を行うためにはテーブルを用いるが，その機能として

次の３つがある。

①ＩＤ番号の割当て

②同義語，類似語の集合を作成

③統一のための用語処理の履歴を記録

◎　統一処理の注意事項

(1)　人　名

漢字名，カナ名，英文名（ローマ字）の３通りを登録することが多い。データベースの種類によっては厳密な漢字名を要求されることもある。カナ名は人によってまちまちであるので，漢字の読み音から自動的に変換することは危険である。英文名はカナ名が正確に登録されていれば自動的に変換することは可能である。だが，人によってはローマ字の表記法にも執着する場合があるので注意を要する。

(2)　会 社 名

漢字名，カナ名，英文名（ローマ字）の３通りを登録することが多い。正式な会社名のほかに略称や慣用名を用いる場合があるので，統一テーブルに登録する必要がある。英文名は，その会社自体においても必ずしも統一されていない場合が多いので注意を要する。

とくに，会社名は変更，合併等により名前が変わる場合が多いので，常にメンテナンスを行い，その履歴をデータとしてテーブルの上に記録しておくことが大事である。また，同名の会社がかなりあるので，それを区別するためのコード，府県コード，業種コードのような形でテーブルに登録しておくと便利である。

(3)　用　語

漢字名，カナ名，英文名の３種類を登録しておくことが望ましい。

自然語の文章を対象にするときは，表記法の統一を行うことが大切である。そのためには，文章から抽出された用語のインバーテットファイルを作成し，そのファイル上で用語を統一するのが能率的である。

1.2　変換処理

物理的な変換処理には，文字コードの変換とテープフォーマットの変換の

2つがその主なものである。いずれも既製のデータベースを自社システムに取り込む場合や，データベースを利用するシステムの変更等でその変換作業が発生する。とくに，マスタテープフォーマットを変更した場合，既に自社内で運用されているほとんどのシステムを改造する必要が起こってくるので，テープフォーマットの設計に当っては拡張性の高いものを作るように心がけるべきである。

また，数値型データベースの場合，各種のコードを用意して検索のキーとして用いており，経済データにおける代表的なコード化としては次のものがある。

(1) 企業（会社）コード

上場企業の場合，証券コード（4桁）を利用するのが一般的であるが，上場以外の企業を扱う場合も多いので，データベース機関によってユニークな企業（会社）コード（6桁または9桁）を作成して企業のキーとしている。企業コードは，企業の合併，称号変更，本店移転等によって変更しないことを原則としている。

(2) 地域（住所）コード

府県コード，市町村コードなど，ＪＩＳとしての地域コード（5桁）のほかに，府県，市区郡，町村名，大字小字等に対処するための住所コード（6桁ないし11桁）があり，市町村の合併，町村名の変更，住所表示の変更に伴うメンテナンスを国土地理院の情報に基づいて行っている。

(3) 業種コード

日本標準産業分類による業種コード（4桁）があり，大分類，中分類，小分類，細分類が定められているが，この他に各々の目的によりユニークな業種分類を行っている機関もある。また，さらに細分類したものとして商品コード，製品コードといったものもある。

(4) その他のコード類

金融機関統一コード（金融機関名4桁，支店名3桁），事業所コード，統計データの各項目ごとのコードなど各種のコードがあり，必要に応じて利用されている。このほか，数値型データについては月間データを4半期ごとのデータに変換したり，千円単位を百万円単位に変換したり，円をド

ルに変換したりする単位の変換が必要である。

1.3 圧縮処理

画像型データベースの１種であるイメージ情報は，コード情報のデータに比べて情報量が極めて多いので，種々の方法でデータを圧縮している。

ちなみに，Ａ４版(210mm× 297mm）の書類を線密度10本×10本／mmで入力した場合のドット情報としての情報量は，210mm×297mm×10本×10本÷８ビット＝780KB となる。このドット情報を２値化し，デジタル情報として入力したデータの中では，連続した同色の画素の列を他の符号に置き換えることでデータ量を圧縮する。再生の際には逆の手順でそれを復元する。この圧縮，復元技術のうち，主なものを次に説明する。

(1) ランレングス方式

走査により，横１列に入力されたデータの中で連続している同色の画素をひとつのデータとして取り扱い，同色の画素の種類とその連続数を符号化して記録する方法で，１次元圧縮方式の１種である。圧縮率は，標準テキストチャートの場合で１／４といわれている。

(2) MH (Modified Huffman) 方式

MH方式は，ＣＣＩＴＴ勧告のグループ３のファクシミリ装置に用いられている１次元符号化方式である（ＣＣＩＴＴ　Ｔ・４参照）。MH方式はランレングス方式と同様に，連続している同色の画素の連続数（ラン長）をカウントしてそれを符号化したものであるが，符号化の際に発生頻度の高いラン長のものには短いビット数の符号を付すなどの改良を加えている。圧縮率は，標準テキストチャートの場合で１／８といわれている。

(3) MR (Modified Read) 方式

MR方式は，グループ３のファクシミリ装置に用いられている１次元符号化方式のオプションとして拡張したものである（ＣＣＩＴＴ　Ｔ・４参照）。この圧縮方式は，前ラインの変化点（白→黒または黒→白）と現ラインの変化点の位置の差を符号化するもので，２次元符号化方式ともいわている。圧縮率は，標準テキストチャートの場合で１／15といわれている。

(4) MMR (Modified Modified Read) 方式

MMR方式は，グループ4のファクシミリ装置に用いられている2次元符号化方式である（CCITT　T・6参照）。この圧縮方式は，MR方式と同様なアルゴリズムを使用しているが，最初のラインを符号化する際に，参照ラインとして全て白ランのラインを仮定して行う，伝送上のエラーチェックのためのEOL (End of Line)の冗長データを削除するなどの改良を加えたものである。圧縮率は，標準テキストチャートの場合で1／20といわれている。

(5) DYUV (Delta YUV) 方式

DYUV方式はカラー画像等を扱うCD－Iの圧縮等に用いられるもので，自然画データのRGB (Red, Green, Blue) の各データを輝度Yと色差U，Vに分離し，さらに各Y，U，Vごとにそれぞれの前値との差分を符号化するものである。CD－Iでは，自然画データをRGB各々8ビットの計24ビットで1画素を構成する。このDYUV圧縮により約1／3に圧縮される。

1.4 流通媒体の作成

(1) MT等によるデータの提供

データベースの中身を第3者に提供する場合，通常MT，フロッピーディスク，光ディスク等の媒体を使用する。データベースの中身を部分的に提供する場合には，マスタテープより必要とする期間のデータまたは項目を抽出して，送付テープを作成する抽出処理が必要になる。また，この際，提供先のシステムに合せるために，コード変換やテープフォーマットの変換を行う場合が多い。

さらに，データベースの内容を常に最新の状態に変更するシステムにおいては，すでに送付したデータの中身を更新するためのトランザクションテープを作って送付する必要がある。

(2) CD－ROMによるデータベースの製作

CD－ROMは，レーザ光による情報を記録したガラス基盤より作成したプレス用原盤を用いて，ちょうどレコード盤を作るのと同じようにプレ

スまたはモールデングによってプラスチック製のディスクを製造する。このため，１度に多量の複製をすることができる。したがって，１タイトル当りの複製枚数が多ければ多いほど，ＣＤ－ＲＯＭ１枚当りの製造コストが低減する。これが電子出版といわれる所以であって，再生用の機器が普及し，そのコストが低減し，一般化してくれば，今まで紙によって発行されていた出版物のかなりの量がＣＤ－ＲＯＭに移行する可能性は高いといえよう。

① 編　集

ディジタル化された情報が記録されているユーザデータを，ＣＤ－ＲＯＭ作成に必要なデータフォーマットに編集し，１セクター2048バイト単位に区切ってＣＤ－ＲＯＭマスタテープを作成する。この際，効率のよい検索を行うためには，検索システムとの関連において，ファイル構造，ディレクトリ構造，インデキシング構造を決定する必要がある。さらに，データを特定のシステムのみによって読み取れるようにする必要があれば，この工程でデータの暗号化を行う。

② データプレマスタリング

ＣＤ－ＲＯＭマスタテープのデータは，プレマスタリングシステムによりチェックおよびセクタアドレス情報（ヘッダー）の付加が行われ，モード１の場合は，誤り訂正符号のＥＤＣ，ＥＣＣが演算処理され，それが記録された後に同期信号データが付加される。このようにしてプレマスタリングされたデータは，３／４インチのＶＣＲ（ビデオカセットレコーダ）テープに出力される。

③ ディスクマスタリング

ＶＣＲテープに記録されたデータは，ＣＤマスタコードカッターにおいて，レーザ光線によりフォトレジストが塗付されたガラス基盤にピット情報として記録される。次いで，ピット情報が記録されたガラス基盤上にＮｉメッキを施し，それを剝離してＮｉマスタを作成する。

④ 成型工程

このようにして作成されたＮｉマスタの原盤を用いて，通常のプレスまたはインジェクションモールデングによって透明プラスチック製のデ

またはインジェクションモールデングによって透明プラスチック製のディスクにピット情報を転写する。そして，最後にこのディスクの表面に反射膜，保護膜を被覆してＣＤ－ＲＯＭを完成する。

指導上の留意点

データの加工には，統一処理，変換処理，復元処理，流通媒体の作成などがあるが，このうちの統一処理，変換処理，媒体作成については，一応の実務能力がつくように指導する。また，復元処理については特別の場合を除き必要ないので，説明する程度でもよい。

用　　語

統一処理，同義語，類似語，変換処理，イメージデータ，圧縮処理，復元処理，流通媒体，作成基準，データシート，入力方式，自然語，テーブル，用語処理，表記法の統一，インバーテットファイル，文字コード，テープフォーマット，マスタテープ，企業コード，地域コード，業種コード，画像型データベース，イメージ情報，情報量，線密度，ドット情報，２値化，ディジタル情報，画素の列，再生，復元技術，ランレングス方式，１次元圧縮方式，圧縮率，ＭＨ方式，ＣＣＩＴＴ，ファクシミリ，１次元符号化方式，連続数，ＭＲ方式，変化点，ＭＭＲ方式，ＥＯＬ，ＤＹＵＶ方式，カラー画像，ＲＧＢ，輝度，色素，電子出版，検索システム，ファイル構造，ディレクトリ構造，インデキシング構造，データプレマスタリング，プレマスタリングシステム，セクタアドレス情報，モード１，ＥＤＣ，ＥＣＣ，同期信号データ，ＶＣＲ，ＣＤマスタコードカッタ，レーザ光線，フォトレジスト，インジェクションモールデング

第4章　データ・ファイルの構造化

目　　標

ＤＢＭＳでは，階層型，網型，関係型の３種のデータ・モデルが扱われる。データベースの適用業務の分析により，これに適したデータ・モデルの選定とＤＢＭＳの選定で行われる。この章では，データ・ファイルの構造化，すなわちデータベースそのものの設計をＤＢＭＳの機能と関連させながら説明する。

内　　容

1.1　構　　造

データベースは，多くのデータ源から収集されるデータを統合的に収容する複合ファイルである。たとえば，人事データベースでは，従業員の生年月日，給与などは従業員に付随するデータであるが，その所属部門はそれ自体独立したデータとして管理されるべきもので，個々の従業員と部門とはＤＢＭＳの機能によって相互に結びつけられなければならない。データベース化の目的に応じて，何をデータベースにおける独立した「主体」とみなすかが決定され，これに即して，主体の属性として扱うデータと主体間の関係として保持されるべきデータが区分されることになる。「主体」は物理的にはレコードに対応するので，前者のデータを「レコード内データ」，後者のデータを「レコード間データ」ということがある。データ・モデルはレコード間データの論理的な構造を表現するもので，データベース化の適用対象に応じて３種のモデル間の得失がある。

(1)　構　造

具体的事例に基づき，レコード内データとレコード間データの区分，適切なデータ・モデル（階層型，網型，関係型）の選択，そこへのデータの割付について説明する。

(2)　レコード項目

データ項目には，検索対象項目と検索の結果に付随して出力されるだけ

の項目がある。また，更新頻度，更新方式（オンライン／バッチ），必ず存在する必修項目と該当するときだけ現れる選択項目，同種項目が繰り返して収容される場合の固定回繰り返しと可変回繰り返しのような区分がある。項目の長さも固定長と可変長とがある。

文献型データでは可変長，可変繰り返し，選択項目が多い。その場合，長さや繰り返し回数の最大，平均，最低を予め調査してデータベースの設計をする必要がある。数値データは一般に固定長（最大長）として扱うが，時系列データなど可変回繰り返し項目になるものも多い。

1.2 索　引

ＤＢＭＳではデータの呼出し速度を向上させるため，必要に応じて索引を生成しておくことができる。データベースの利用目的に応じて，どのような検索要求が生じ、それを満足するために索引を付すべき項目の選定について説明する。また，ＤＢＭＳでは各種の索引構成手法が採用されており，それぞれに得失があるので，これらを具体的データの事例と関連させて説明する。

1.3 メンテナンス性

① ファイルの構造化に当っては，項目の追加や削除といった拡張性について説明し，これらの性能はＤＢＭＳの構造によるものが多いことを説明して理解させる。

② データベースはデータ量が膨大なため，更新の必要がある場合は更新のしやすい構造にしておかねばならないことを説明し，理解させる。ただし，文献型データベースなどでは，データの追加は行うが内容の更新は行わないものも多い。また，更新の際には索引も更新する必要があり，この処理時間も問題になる。

1.4 セキュリティ

データベースの信頼性・安全性については，アクセス権限の設定を行い，被害を最小限にする。また，データファイルの維持・更新の方法やタイミングについて説明し，理解させる。

グについて説明し，理解させる。

指導上の留意点

構造化に当っては，利用者の立場に立ち，しかもメンテナンス性を考慮したものでなくてはならないことを指導する。

用　　語

ＤＢＭＳ，階層型，網型，関係型，データ・モデル，適用業務，構造化，データ源，複合ファイル，主体，属性，レコード内データ，レコード間データ，検索要求，索引構成手法，メンテナンス性，セキュリティ，追加，前除，拡張性，安全性

第５章　データの入出力

目　　標

この章では，実際にデータを入力する際の入力手段の種類，特徴，使用する機器などについて説明するとともに，入力についての問題点，注意事項などを指摘する。

入力は入力データの検査が重要であるので，検査方法およびその際の注意事項について説明する。また，文献型データベースが主としてラインモードで出力されるのに対して，数値型データベースの場合は画面対応型が多いことについても触れる。

内　　容

1.1　コード入力

パンチマシン，漢字入力機，ワープロ端末機等を用いて，コード化された英数字や文書などを入力する。入力に使用される機器は，入力コードに応じて各々開発されている。最近では，パターン認識技術を用いて光学的，音声的に文書を入力する技術が開発されている。

(1)　入力コード

①　ＥＢＣＤＩＣ（Extended Binary-Coded Decimal Interchange Code）

ＢＣＤコード（Benary-Coded Decimal code)を拡張して制定されたものである。８単位で構成される符号を使用し，英字の大文字，小文字，特殊記号，制御文字 256種類を表すことができる。

②　ＥＢＣＤＩＫ（Extended Binary-Coded Decimal Interchange Kana code）

８単位で構成される符号を使用し，基本的にはＥＢＣＤＩＣと同じコードであるが，このコードではアルファベットの小文字の代わりにカナ文字を使用している。

③　漢字コード　　ＪＩＳ第１水準，ＪＩＳ第２水準，その他

④　ＡＳＣＩＩ（米国標準コード：American Standard Code for Infor-

mation Interchange)

米国の規格協会であるＡＮＳＩ (American National Standard Institute)が制定し，1963年以来米国のデータ伝送で使われている。７ビットからなり，英字，数字，特殊記号，一部の機能符号を表すことができる。なお，８ビットのＡＣＩＩも開発されており，これはＡＳＣＩＩ－８と呼ばれている。

(2) パターン入力

① 光学的文字読み取り（ＯＣＲ）

光学的文字読み取り機を用いてペーパー上に書かれた文字情報を読み取り，パターン認識によりコード情報に変換する。

② 音声入力

人間の音声をパターン認識によりコード情報に変換する。

③ 手書き入力

手書き入力されたイメージ情報をパターン認識によりコード情報に変換する。

1.2 イメージ入力

イメージ入力の方式には，スキャナを用いるディジタル方式と高解像度ＴＶカラーテレビを用いるアナログ方式とがある。

(1) ディジタル方式

① モノクロ

モノクロのスキャナを用いて入力する。解像度は通常８本／㎜と16本／㎜の２種類がある。

② ハーフトーン

モノクロのスキャナを用いて入力するが，階調モードを選択することによってハーフトーンの画像や写真を入力することができる。

③ カラー

カラースキャナを用いて入力する。

(2) 高解像度ＴＶカラーテレビ

高画質の静止画の画像が入力できる。

1.3　ベクトル入力

入力された画像に位置情報を与える入力方式で，地図や有機化合物の構造式の入力に用いられる。これには，デジタイザーを用いて入力するメッシュ型とグラフィックス入力機を用いるグラフィックス型とがある。

1.4　入力データの検査

入力されたデータをチェックするには，人手によるもの（マニュアルチェック）と機械によるもの（マシンチェック）とがある。そのチェック方式は以下のとおりである。

(1)　プルーフトータルチェック

プルーフリストとインプット伝票を照合し，インプット内容をチェックして確認する。

(2)　モニタチェック

インプット内容をディスプレイに表示するか，モニタリストにプリントアウトしてインプット内容をチェック，確認する。

(3)　リミットチェック

数値データの上限値，下限値をチェックする。

(4)　対応チェック

データの取り得る値をテーブル化し（コンスタントテーブルという），入力データがコンスタントテーブルにあるかないかをチェックする。

(5)　組み合わせチェック

1レコード中に相互関連のある数種のデータが含まれているとき，それらのデータの組み合わせが理論的に許されるものかどうかをチェックする。

(6)　カラムずれチェック

データ項目の区切り記号やレコード区切り記号を予め規定しておき，入力データの該当箇所が同一の記号になっているかをチェックする。

(7)　バッチトータル（ハッシュトータル）チェック

バッチ単位，またはハッシュ単位にある特定の項目の合計値を求め，これをデータとともに出力する。入力時，電子計算機でも別に合計値を計算し，入力された合計値と比較して一致するかをチェックする。トータルチ

ェックと呼ぶこともある。

(8) シーケンスチェック

入力データがある一定の順序に並んでいなければならない場合，その通りに並んでいるかをチェックする。

(9) ニューメリックチェック

数値項目に数字以外の文字が混入していないかをチェックする。

(10) チェックデジットチェック

チェックデジットを使用してチェックする。

(11) バランスチェック

バッチ単位，2種類以上の計算結果が等しいことが予めわかっている場合，計算機でそれを計算し，計算結果が等しいかどうかをチェックする。

(12) ダブルレコードチェック

マスタファイルを更新時に，トランザクションファイルに同一レコードが複数個ある場合は，このレコードをエラーとして検出する。

(13) マンマッチレコードのチェック

マスタファイルのキー項目とトランザクションファイルのキー項目を照合し，不照合のレコードをエラーとして検出する。

(14) オーバーフローのチェック

計算機で四則演算を行った結果，桁数が予め用意した桁よりも大きい時の桁あふれ（オーバーフロー）をチェックする。

(15) サインチェック

計算結果の符号がマイナス（プラス）になっては不都合な場合に行うチェックである。

1.5 出　力

文献型データベースが主としてラインモードで出力するのに対し，数値型データベースの場合，画面対応型で1画面に一定の情報をまとめて表示する場合が多い。また，1画面に表，グラフ等で表示する場合があり，その点でのソフト上の対応も必要である。

数値型データベースの利用に当って，データそのものの利用に加えて計量

モデルや予測モデルの必要な場合がある。さらに，分析用のソフト，モデル・シミュレーション用のソフト，統計解析用のソフトなどを用いてデータを加工，分析することも多く，ソフトを独自に開発し，データとともに提供することが必要である。

（例）経済分析，時系列分析，モデル・シュミレーション，産業連関，グロスセリション分析，ポートフォリオ分析，財務分析，企業評価システム，マクロモデル，金融モデル，経済長期・短期予測モデル　　等

指導上の留意点

コード入力，イメージ入力，ベクトル入力および出力については説明する程度で構わないが，入力データの検査については実際の作業ができる能力まで修得させる。

用　　語

検査方法，ラインモード，画面対応，コード入力，入力コード，ＥＢＣＤＩＣ，ＢＣＤコード，ＥＢＣＤＩＫ，漢字コード，ＡＳＣＩＩ，パターン入力，ＯＣＲ，音声入力，手書き入力，イメージ入力，スキャナ，ディジタル方式，アナログ方式，モノクロ，ハーフトーン，カラー，階調解像度，階調モード，ベクトル入力，デジタイザ，メッシュ型，グラフィックス入力機，グラフィックス型，マニュアルチェック，マシンチェック，プルーフトータルチェック，プルーフリスト，モニタチェック，リミットチェック，対応チェック，コンスタントテーブル，組み合せチェック，カラムずれチェック，ハッシュトータル，シーケンスチェック，ニューメリックチェック，チェックデジット，バランスチェック，ダブルレコードチェック，アンマッチレコード，オーバーフロー，サインチェック

第6章　流通データの導入

目　　標

この章では，流通データをデータベース化するに当って必要な導入上の留意点，およびシステムへの組み入れに関する知識を修得させる。

内　　容

1.1　予備調査

予備調査では，提供されるデータの提供形態，提供の継続性の見通し，内容の信頼性，出典，データ量，提供を受ける際に支払う代価等を調査する必要があることを説明し，理解させる。

1.2　選定・契約

選定・契約では，提供されるデータが利用目的に合致したものであるのか，利用上の問題点はないのか，第3者に提供してよいのか，さらに提供する場合の条件は何か，著作権上のトラブルはないのかなどを調査する必要があることを説明し，理解させる。

1.3　ファイル入手

ファイル入手に当っては，提供先のフォーマットのままなのか，こちらの指定したフォーマットに変換してくれるのか（ただし，こちらが読めるだけのフォーマット変換なのか），また，それら提供フォーマットは継続性があるのか，こちらの都合で変更が可能なのかなどを提供先と取り決める必要があることを説明する。

1.4　内容調査

内容調査では，予備調査，選定・契約時，ファイル入手時に得た仕様通りになっているか，データの単位等は明確であるか，使用されているコードなどは既存の体系であるのか，端末ソフトウェア等はそのまま使えるのかにつ

いて具体的に調査する必要があることを説明し，理解させる。

1.5　加工処理

(1)　フォーマット変換

流通データの導入に当っては，既存のシステムの入力，更新処理に合わせるための不足項目の追加，不要項目の削除，既存システムで用いている単位にデータをそろえるなどのデータの加工処理が必要であることを説明し，習得させる。

(2)　インデックス整備

流通データの導入に当っては，既存のデータベースのインデックスとの整合性のチェック，インデックスの読み替え，追加等が必要なことを説明し，習得させる。

1.6　システムへの組み入れ

流通データの導入に当っては，既存のシステムのどの部分に組み込むのが適切であるかを評価させるとともに，既存データとの更新のタイミングのずれの調整について説明し，習得させる。

指導上の留意点

流通データの導入に当っては，ドキュメントとデータとの整合性，提供形態，信頼性，継続性，著作権上のクリアなどのチェックや了承が必要であることを十分に説明する。

用　　語

提供形態，提供の継続性，信頼性，出典，契約，著作権，フォーマット，仕様通り，端末ソフトウェア，流通データ，既存システム，更新処理，不足項目，不要項目，加工処理，インデックス整備，システムへの組み込み，更新のタイミング

参考文献

「情報管理入門」　㈳日本ドクメンテーション協会発行（1980年10月）

第７章　企業内データベースの構築

目　　標

この章では，企業内データをデータベース化するに当って必要な導入上の留意点，およびシステムへの組み込みに関する知識を修得させる。

内　　容

1.1　予備調査

予備調査では，既存のシステムから移行すべきデータや新規に作成するデータを調査する必要があること，また，データの入手先，入手したデータの変換，投入データの正当性を調査する必要があることを理解させる。

とくに，データベースの移行については，同一データについて複数個の入手先が存在することがあり，しかも，それらが互いに矛盾しあうこともあるので，そのことに注意することも説明する。

1.2　データ入力

入力データがすでにコンピュータに乗っている場合と，新規に投入する場合に分けて説明する。

(1)　入力データがすでにコンピュータに乗っている場合

データの記憶媒体，記憶形式を考慮し，変換方法について検討を行うことを説明する。

(2)　新規に投入する場合

データの入力方法，編集の方法と時期，データベースへのデータ投入，ユーティリティの用意および編集後のチェック方法，担当部署について検討を行うことを説明する。

指導上の留意点

企業内データの導入に当っては，正当性のあるデータの入手についての検討，および編集後のデータチェック担当部署を検討することが重要であることを理解させる。

用　　語

企業内データ，予備調査，データ入手先，データの変換，データの正当性，記憶媒体，記憶形式，変換方法，入力方法，編集方法，編集時期，チェック方法

第4部　データベースの運用管理

第4部　データベースの運用管理

目　　標

データベースの運用管理に必要な知識・技術を習得させる。

内容のあらまし

内　　容	時間数
運用管理基準の設定 　運用体制 　データベースの運用保守	3時間
運用体制 　責任体制 　データベース管理者 　オペレータ教育 　ユーザ・サポート	3時間
データベースの運用保守 　運用管理支援ツールの整備 　データベースの更新 　利用データの分析 　ユーザ情報の管理 　セキュリティ 　データベースの再編成	6時間
合　　計	12時間

第１章　運用管理基準の設定

目　　標

この章では，データベースの運用管理に関する運用体制，および運用保守に関するマニュアル化のポイントを学習する。

内　　容

1.1　運用体制

(1)　責任体制

①　平常時対応

②　緊急時対応

(2)　データベース管理者

①　データ管理

②　ＤＢＭＳを含むソフトウェア管理

③　コンピュータおよびネットワーク管理

④　顧客管理

(3)　オペレータ教育

データベース・システムの運用

(4)　ユーザサポート

①　マニュアル作成

ユーザーズ・マニュアルの作成更新

②　ユーザ教育

導入時教育

③　質問対応

ア．加入申込

イ．検索時の質問

ウ．代行検索時

1.2　データベースの運用保守

(1)　運用管理支援ツールの整備

①　モニタリング

②　バックアップ

(2)　データベースの更新

①　データの入手

②　整理・加工

③　追加・修正・削除

(3)　利用データの分析

①　検索の効率化

②　データベースの拡充

(4)　ユーザ情報の管理

①　ユーザ情報ファイルの管理

ア．パスワード等の管理

イ．キープされたワークファイルの管理

②　課　　金

ア．利用時間・

イ．利用内容

ウ．ワークファイルのキープ

エ．ネットワーク利用

オ．出力内容

(5)　セキュリティ

①　障害対策

ア．原因追及

イ．障害復旧

ウ．代替案

エ．課金の割引

オ．障害予防管理

カ．変更管理

② 機密保護

ア．データベース・システム

イ．ソフトウェア

ウ．データ

エ．利用可能レベル

オ．プライバシ保護

指導上の留意点

データベースの運用管理では，社内データベースと商用データベースとでは運用基準が異ってくることを説明し，理解させる。

用　　語

運用体制，責任体制，平常時対応，緊密時対応，データベース管理者，ＤＢＭＳ，ソフトウェア管理者，コンピュータ管理者，ネットワーク管理者，顧客管理者，オペレータ教育，データベース・システム，データ更新，ユーザサポート，ユーザーズ・マニュアル，ユーザ教育，導入時教育，ニューズレター，質問対応，加入申込，代行検索，運用管理支援ツール，モニタリング，バックアップ，パスワード，課金，利用時間，利用内容，ネットワーク利用，セキュリティ，障害対策，原因追及，障害復旧，代替案，障害予防管理，変更管理，機密保護，プライバシ保護

第２章　運用体制

目　　標

この章では，データベース検索システムの運用に携わる要員およびその役割について説明し，理解させる。

内容のあらまし

1.1　責任体制

データベース・システムの運用に当っては，システムが常に正常に作動するように障害予防までカバーできる体制が望ましい。また，一担障害が発生した場合は，被害を最少限度に食い止め，素早い復旧が可能な体制作りが必要である。

1.2　データベース管理者

データベースの管理者は，データベースの維持・更新のために以下のことを実施する。

(1)　データベースの定義と構築

(2)　データベースの維持・管理（資源管理，再編成等）

(3)　データベース利用者への説明資料の作成

(4)　顧客情報を把握し，運用の円滑化を図る。

1.3　オペレータ教育

オペレータは，平常時のデータベース検索システムの運転はもちろん，緊急時にはシステムの復旧を行なうわけであるから，日頃よりこれらのことが滞りなくできるような教育を行う必要がある。

1.4　ユーザ・サポート

(1)　マニュアル作成

データベースのマニュアルには，データベースを操作するためのマニュアルと，顧客が必要とする情報がどこにあるかを示すマニュアルとが必要である。

(2)　ユーザ教育

ユーザ教育は，普通はデータベースの導入時に実施される。この時点では，端末の操作の仕方と所在情報から顧客の求めている情報がどのデータベースに収められているかを見極める教育がなされる。また，検索機能やデータベースの品揃え等の増加についてはニューズレター（広報誌）で知らされるのが一般的である。

(3)　質問対応

データベースのベンダには，加入申し込みや検索方法，所在に関する問い合わせなどがあるので，これらに対応する要員が必要である。また，代行検索やＳＤＩサービス（検索式を登録しておいて定期的に検索を受けるサービス）を実施しているベンダも，ユーザが真に求めているものを確定するための要員が必要である。

指導上の留意点

データベース・サービスを実施するに当っては，常に障害の予防に努めるべきであるが，障害発生時は被害を最小限にした迅速な復旧体制が必要であることを説明し，理解させる。

用　　語

データベース検索システム，障害予防，復旧体制，所在情報，ニューズレター，ＳＩＤ，検索式

第3章　データベースの運用保守

目　　標

この章では，データベースの運用保守に必要なツールや仕組みにはどのようなものがあるかを説明し，理解させる。

内　　容

1.1　運用管理支援ツールの整備

運用管理支援ツールには，データベース，コンピュータ，ネットワークなどの利用状況を記録して，効率の悪い箇所を改善するための情報を提供するモニタリング・システム，システムダウン時の状況や復旧のための情報を提供をするバックアップ・システムなどがある。

1.2　データベースの更新

データベースの更新は，以下の手順にしたがって実施される。

(1)　ソースデータの収集

(2)　入力データの作成（アブストラクト作成，インデキシング，ディクショナリ作成等）

(3)　データ加工（統一，変換，圧縮等）

(4)　データファイルの構造化

(5)　データ入力

(6)　データベースの更新

1.3　利用データの分析

モニタリング・システムで収集されたデータから利用頻度の高いデータベース，利用の多い検索方法，使用時間，ファイル使用量等が分析され，データベースの拡充や検索の効率化が図られる。

1.4　ユーザ情報の管理

(1)　ユーザ情報ファイルの管理

ユーザ情報ファイルの管理では，1ユーザが1パスワードとは限らないので，これらパスワードの管理やベンダが用意したワークファイルの利用状況の管理等を行う。

(2)　課　　金

課金はパスワードごとに実施するのが原則なので，パスワード，使用データベースごとの接続時間，出力形態，ネットワーク利用時間等の情報を収集する必要がある。

1.5　セキュリティ

(1)　障害対策

データベースサービスの実施に当って，障害予防策を施すのは当然ではあるが，障害が発生した際は障害対策マニュアルにしたがって，原因追及，障害復旧を試みる。また，復旧に時間がかかるようであるならば代替案の実施，課金の割引措置等が必要になってくる。

なお，ハードウェア，ソフトウェア，ネットワーク，データなどの変更履歴や障害履歴も障害対策時の参考となるので管理しておく必要がある。

(2)　機密保護

機密保護については各レベルごとに異なるが，いずれの場合もパスワードによる管理がなされているのが一般的であり，次のようなレベルがある。

①　システムレベル（ハードウェア，ソフトウェア，ネットワーク等の利用や更新に関するもの）

②　データレベル（参照できるデータベース，項目に制限を加えているもの）

1.6　データベースの再編成

データベースのオーバーフロー域にレコードがたまったり，使用できない空スペースが記憶装置上に出現すると効率が悪くなるため，再編成が必要となる。

再編成に当っては，全てのデータが対象となるので非常に時間がかかる。なお，障害等の回復時間も考慮し，十分な作業時間をとっておく必要がある。

指導上の留意点

運用保守については，知識レベルではなく実施できるレベルまで育成する。

用　　語

運用管理支援ツール，モニタリング・システム，システム・ダウン，バックアップ・システム，パスワード，ワークファイル，接続時間，出力形態，ネットワーク利用時間，障害対策マニュアル，変更履歴，障害履歴

第５部　データベースの利用

第5部　データベースの利用

目　　標

商用データベースサービスの効果的な検索技術を習得するとともに，国際的動向，知的所有権，個人データ保護に関する知識も習得する。

内容のあらまし

内　　容	時間数
商用データベース・サービス 　データベースの流通構造 　データベース・サービスの選択	6時間
商用データベース検索技術 　文献情報検索システムとコマンド 　ファクト系データベース検索システムとコマンド 　その他のデータベース 　データのダウンロードと多角利用	9時間
サーチエイドの活用 　サーチエイドの種類 　サーチエイド利用の目的 　サーチエイドの選択と利用法 　サーチエイドの流通	3時間
データベース・サービスの国際的動向 　歴　史 　アメリカのデータベース市場，業界，利用等の最新動向 　ヨーロッパ主要国の動向	3時間
知的所有権と個人データ保護 　知的所有権 　個人データ保護	3時間
合　　計	24時間

第1章　商用データベース・サービス

目　　標

オンラインで提供されている商用データベース・サービスを利用するには，端末装置を設置し，データベース・サービス業者と契約を結んでパスワードを入手するとともに，データ通信回線の利用手続をすることが必要となる。また，これらの商用データベース・サービスを利用するには，数多くのデータベースの中から欲しい情報の入っているものを的確に選ぶことが必要である。ここでは，商用データベースの流通構造と関連業者各々の役割り，およびデータベース選択のツールとノウハウについて説明し，理解させる。

内　　容

1.1　データベースの流通構造

オンライン・データベースの流通には以下の業者がかかわっている。各々の役割りと具体例を説明する。

(1)　プロデューサ

(2)　ベンダ

(3)　付加価値通信網

(4)　情報検索業者

(5)　ゲートウェイ・サービス

1.2　データベース・サービスの選択

(1)　データベースの種類

オンラインで提供されているデータベースにはどのようなものがあるかを以下の視点から分類し，代表例を紹介する。

①　タイプ別（ファクト・データベースとリファレンス・データベース）

②　主題別（ビジネス，科学技術，一般，その他）

(2)　データベースの選択基準

データベースを選択するに当って着目すべきポイントを挙げ，具体例を

紹介する。

(3) データベース選択のツール

データベースを選択する際に使用するツールを紹介する。

① データベース・ディレクトリ

② データベースのデータベース

③ データベース・クリアリング・システム

④ クロスファイル・インデックス

(4) サービス・システムの選択

オンライン・データベースにアクセスするための端末装置，通信ソフト，検索システムの内容，役割り，選択基準，選択方法を説明する。

① 端末装置（パソコン，モデム）

② 通信ソフト

③ 検索システム（基本機能，付加機能，マンマシン・インタフェース）

④ 通信ネットワーク

(5) ユーザ支援機能

オンライン・データベース・サービスのユーザ支援機能の種類，内容を実態に即して説明し，理解させる。

① ユーザ教育

② 利用マニュアルとニューズレター

③ ヘルプ・デスク

(6) 契約形態と利用料金

オンライン・データベースの契約形態，利用料金の設定は業者やデータベースの種類によりまちまちであるが，これをいくつかのパターンに類型化するとともに，料金の比較方法を紹介する。

指導上の留意点

実例を数多く利用し，実態に即した指導を行う。

用　　語

プロデューサ，ヘルプデスク，ベンダ，パスワード，
付加価値通信網，ユーザーＩＤ，情報検索業者，サーチャ，
ゲートウェイ・サービス，データベース・ディレクトリ，
ファクト・データベース，パソコン通信，リファレンス・データベース，
データベース・クリアリング・システム，クロスファイル・インデックス，
モデム，通信ソフト，検索システム

第2章　商用データベース検索技術

目　　標

商用データベースを検索するための機器，通信手段，ソフトウェアについて基本的な知識を得る。典型的な情報検索サービスの利用方法，基本的なコマンドを知り，検索の手順を理解する。

内　　容

1.1　文献情報検索システムとコマンド

典型的な文献情報検索システムの構成，コマンド，検索方法について説明する。

(1) 文献情報検索システムの基本（システム構成，ファイル，レコード，データ項目，オフラインプリントなど）について説明する。

(2) 転置索引の構成，検索のアルゴリズム，とくにブール演算子，近接演算子，全文検索について述べる。

(3) 文献データベースの特徴を理解して検索する方法（データベースの選択，検索フィールドの選択，検索語の選択，シソーラスの利用など）について述べる。

(4) 複数データベースのクロスファイル検索の有効性と方法について述べる。

(5) 原文献の入手，翻訳について述べる。

1.2　ファクト系データベース検索システムとコマンド

典型的なファクト系データベース（たとえば会社信用情報）検索システムの基本的構成について説明する。

1.3　その他のデータベース

図形情報など，その他のデータベース・サービスについて説明する。

1.4　データのダウンロードと多角利用

検索データのダウンロード，データの編集，蓄積について述べる。その際に発生する著作権問題にも触れる。社内システムとの統合利用についても論ずる。

指導上の留意点

データベースの検索技術習得に当っては，実際に端末を操作することが望ましい。

用　　語

文献情報検索システム，コマンド，検索方法，システム構成，ファイル，レコード，データ項目，オフラインプリント，転置索引，検索アルゴリズム，ブール演算子，近接演算子，全文検索，検索フィールド，検索語，シソーラス，クロスファイル検索，原文献，ファクト系データベース，図形情報，ダウンロード

第3章　サーチエイドの活用

目　　標

データベースのサーチエイド（検索補助資料）類を活用し，データベースのレコード構成，収録データ，収録期間，索引方針，検索手法などについて，検索する立場として効率的検索を実現するための知識を修得する。

内　　容

1.1　サーチエイドの種類

基本的には，すべてのデータベース作成機関やベンダ（データベース作成機関が作成したデータベースを集め，自社のソフトでユーザーにデータを提供するオンライン・システム業者）により，個別のサーチエイドが発行されている。サーチエイドの種類は大別すると次の通りである。

(1)　ユーザーズ・ガイド

データベース作成機関の紹介に始まり，サンプルレコード，レコードを構成する各データ項目の詳細，収録期間，統制索引語を付与する際の方針が中心に解説されている。とくに，初めて検索するデータベースの場合には，当ガイドの事前のチェックは不可欠である。

例：METADEX User Manual　(METADEX データベース)

JICST 科学技術文献ファイルガイド（JICST 科学技術文献ファイル）

(2)　シソーラス

データベース作成機関の索引者が各レコードを索引する際，事前に設定されている統制索引語を参照するためのリストである。検索者にとっては，データベース作成機関によって統制されている索引語を調べ，それを検索語とすることにより効率的検索が可能となる。一般に，シソーラスでは統制索引語はアルファベット順か，あいうえお順に配列されており，ある索引語に対する索引語間の階層関係も表示されている。

また，シソーラスは，基本的には印刷物としてデータベース作成機関よ

り出版されているが，そのなかにはシソーラス自身を機械可読形式のオンライン・ファイルとして提供し，データベース検索のための補助ファイルとしているデータベース作成機関やベンダもある。同様に，最近注目されているＣＤ－ＲＯＭに収録されたデータベースにもシソーラスを搭載しているものがある。

例：Thesaurus of Metallurgical Terms （METADEXデータベース）

JICST 科学技術用語シソーラス（JICST 科学技術文献ファイル）

(3) 分類コード表

分類は，元来データベースに対する印刷体抄録誌や，索引誌中に各レコードを収録する際の配列順序を決定するために導入されたものであるが，現在ではオンライン検索時に有効な検索キーとなっている。ほとんどの分類は，データベースが収録対象としている分野をさらに細分化したものである。分類は，一般的に英数文字でコード化されており，コード表にその階層関係や分類タイトルが記載されている。

例：Alloys Index Names and Codes List (METADEXデータベース)

JICST 科学技術分類表（JICST 科学技術文献ファイル）

(4) 収録対象誌リスト

データベースによってチェックの対象とされている雑誌，会議録，本などのリストである。各誌の正式名，省略名，コーデン・コード，国際標準逐次刊行物番号（ＩＳＳＮ），国際標準図書番号（ＩＳＢＮ），出版社などが記載されている。研究者が普段読んでいる興味ある雑誌が，どのデータベース作成機関によってチェックされているかを調べるのに便利である。

例：Source Journals in Metals and Materials (METADEXデータベース)

JICST 資料所蔵目録（JICST 科学技術文献ファイル）

(5) ニュースレター

データベース中のデータ項目は固定的ではなく，データベースによっては追加データを入力したり，利用者からの要望を取り入れて分類体系の変更などを実施することがある。検索者はこれらの変更を常に知る必要があるが，そのための最適な資料がデータベース作成機関より発行される月刊，季刊などのニュースレターである。

例：Materials Information (METADEXデータベース)

JOISニュース (JICST 科学技術文献ファイル)

(6) その他

データベース作成機関より発行されるサーチエイド類のほかに，ベンダより発行されているユーザーズ・マニュアル（ベンダ・システムのコマンド体系を中心に説明した検索手法の解説書），契約ユーザーへ送付されるニュースレター，各データベースを解説する資料などがある。また，興味ある分野の情報がどのデータベースに収録されているかなどを調べるための資料として，データベース・ディレクトリが他機関より発行されており，なかには印刷体に対応してオンライン・ファイル化されているディレクトリもある。

例：Searching DIALOG (DIALOG INFORMATION SERVICES, INC.)

Directory of Online Databases(CUADRA/ELSEVIER)

1.2 サーチエイド利用の目的

データベースの収録対象分野，レコード構成，各データ項目などを理解し，費用，時間，検索精度などの点で効果的検索を実現することが最大の目的である。とくに重要なのがシソーラスであり，シソーラスを無視した検索を行ってしまうと適合レコードを漏らしてしまう恐れがある。

データベース作成機関の索引者は，レコード中に記述されている概念（内容）を適切な統制索引語に置き換えて索引付けを行う。したがって，検索者は検索しようとする概念をできる限り索引者と同じ立場で，適切な検索語に置き換えて検索するのが望ましい。そのために，シソーラスが各データベース作成機関より発行されており，検索者はリストから検索語とするための統制索引語を選択するのである。

1.3 サーチエイドの選択と利用法

代表的データベース作成機関より発行されている前記の具体例を挙げ，具体的利用法を理解する。ここでは，英国のＩＥＥ(Institute of Electrical Engineers)情報部が作成機関であるＩＮＳＰＥＣデータベースを例とする。

当データベースは，物理学，電気工学，電子工学，コンピュータ，制御工学，情報工学の分野において世界最大の英語のデータベースである。

(1) ユーザーズ・ガイド

図5. 1のように，データベースの紹介や収録対象分野が解説されている。

THE INSPEC DATABASE

The INSPEC database is conceptually a single machine-readable file of classified and indexed bibliographic records, from which the whole range of INSPEC publications and services is derived.

The database as held on INSPEC's own computer system necessarily contains many data elements which are of local or 'housekeeping' significance, and are used only for in-house production purposes. This *'Users' Guide'* is concerned with the database as seen by the user, whether by direct interaction with the machine-readable file in a batch or on-line retrieval system, or through the content of the abstracts journals and other publications.

In Part 2, the overall coverage of the INSPEC database is defined in terms of subject scope and literature types. Part 3 and Part 4 deal respectively with the bibliographic content of an INSPEC record, and the classification and indexing practices used in the database. Part 5 consists of an annotated set of specimen records, one of each literature type.

2.1 *Subject coverage*

The INSPEC database covers the following main subject areas:

Physics
Electrical engineering and electronics
Computers and computing, control theory and technology

These make up the three sections of the database. The detailed coverage of each section can be seen by examining the complete classification scheme (included as Appendix D). It must be understood, however, that the appearance of a subject in the INSPEC classification scheme does not guarantee that all material on all aspects of that particular subject is covered. The extent to which a given subject is covered can only be assessed by consideration of:

(a) the degree to which that subject and its various aspects are central to or on the fringe of the main subjects of the database.

(b) the extent to which material on the various aspects of the subject is scattered in publications devoted primarily to subjects outside the INSPEC subject scope.

For example, take the subject of *materials science and metallurgy*. Although this subject is included in [illegible] hysics section of the database, only certain aspects fall wi[illegible] the coverage of that [illegible] materials. Ot[illegible] and trea[illegible]

図5. 1　INSPEC USERS' GUIDE THIRD EDITION より

(2) シソーラス

図5.2のように，シソーラスには非統制語から統制語への参照，各統制語間の階層関係，分類コードへの参照，索引開始年月などが記載されている。

たとえば，図5.2のように索引者がサーボ機構（SERVO SYSTEM）について検索する場合，シソーラスを無視してSERVO SYSTEM（S）のみを検索語としてしまうと，タイトル，抄録などを検索対象とした自由語検索になってしまう。そのような場合，SERVO SYSTEMSに対する統制索引語であるSERVO MECHANISMSを使った検索ではなくなり，多くの検索漏れが生じる恐れがある。各用語はシソーラス中でUSE参照により統制されており，さらにSERVO MECHANISMSに対する上位，下位，関連語などの階層関係も定義されている。上位，下位，関連語などを検索語に加えるか否かは検索主題によっても左右されるし，最終的には検索者あるいは依頼者の判断による。

包括的にできるだけ漏れなく検索する場合には，統制索引語のSERVO MECHANISMS，下位語のSERVO MOTORSや自由語のSERVO SYSTEM（S），SERVO SYSTEM（S），SERVO MECHANISM（S），SERVO MOTOR（S）などを検索語として選択するのが望ましい。

servo systems
USE servomechanisms

servomechanisms
UF servo systems
NT servomotors
BT controllers
TT control equipment
RT actuators
position control
position measurement
robots
CC C3200
DI January 1973

servomotors
UF servos
NT synchros
BT servomechanisms
TT control equipment
RT actuators
amplidynes
position control
position measurement
tachometers
CC B8340 B8380 C3260
DI January 1973

servos
USE servomotors

USE : USE Reference
統制語
（非統制語から統制語への参照）

UF : Used For Reference
非統制語
（統制語から非統制語への参照）

NT : Narrow(er) Term
下位語

BT : Broad(er) Term
上位語

TT : Top Term
最上位語
(File 12, 13, 14, のみ)

RT : Related Term
関連語

CC※ : Class (ification)Code
オンライン用分類コード

DI※ : Date of Input
索引開始年月

PT※ : Prior Term
旧索引用語

※ INSPEC Thesaurus 特有のものです。

図5.2 INSPEC Thesaurus 1989 より

(3) 分類コード表

図5.3のように，各主題内容に対する英数文字のコードが設定されている。当コード自身も階層関係を有しており，たとえばAは物理学セクション，7400番台は超伝導(Superconductivity)に関する分類で，数字の桁数が多くなるにしたがって細分類となる。検索者は上位，または下位分類を検索語として自由に選択することができる。分類コードを検索語として選択する目的には，一般に次の3つが考えられる。

① 検索主題中に用語では表現しにくい，広い概念が含まれる場合。

例：食品中の界面活性剤

② 種々の分野にまたがる用語による検索結果から，非適合レコードを除くスクリーンとしての役割の場合。

③ 包括的に漏れなく検索するために，用語と分類コードを併用する場合。

A7400	**Superconductivity** *(see also B3200 Superconducting materials and devices)* *1977-. 1973-76, use A8400; before, use A1724*
A7410	**Occurrence, critical temperature** *1977-. 1973-76, use A8420; before, use A1724*
A7420	**Theory** *1977-. 1973-76, use A8410; before, use A1724*
A7420D	Phenomenological and two-fluid theories *1977-. 1973-76, use A8410; before, use A1724*
A7420F	BCS theory and its applications *1977-. 1973-76, use A8410; before, use A1724*
A7430	**General properties** *1977-. 1973-76, use A8420, A8430; before, use A1724*
A7430C	Magnetization curves, Meissner effect, penetration depth *1977-. 1973-76, use A8420; before, use A1724*
A7430E	Thermodynamic properties; thermal conductivity *1977-. 1973-76, use A8420; before, use A1724*
A7430G	Response to electromagnetic fields, nuclear magnetic resonance, ultrasonic attenuation *1977-. 1973-76, use A8420, A8430; before, use A1724*
A7440	**Fluctuations and critical effects** *(inc. noise, chaos, nonequilibrium superconductivity, lo[illegible]ation, etc.)* *1977-. 1973-76, use A84[illegible] use A1724*

[illegible]ity effects, [illegible]
[illegible]-76

図5.3　INSPEC Classification 1988より

(4) 収録対象誌リスト

図5.4のように，収録対象誌の情報が記載されている。誌名，コーデン・コード，国際標準逐次刊行物番号（ＩＳＳＮ）などを検索キーとして，特定雑誌などに収録されている情報に絞って検索することができる。

INSPEC List of Journals and other Serial Sources

Journal of Chemical Education
Coden: JCEDA8 ISSN: 0021-9584
J. Chem. Educ. (USA) 20th and Northampton Streets, Easton, PA 18042, USA. Covered from 1968—
Frequency: 12/yr. BLLD: 4956.00000

Journal of Chemical Information and Computer Sciences
Coden: JCISD8 ISSN: 0095-2338
J. Chem. Inf. & Comput. Sci. (USA) [Formerly: *Journal of Chemical Documentation*] American Chemical Society, 1115 Sixteenth Street, N.W., Washington, DC 20036, USA. Covered from vol.15, no.1 Feb. 1975
Frequency: 4/yr. BLLD: 4956.80000

Journal of Chemical Physics
Coden: JCPSA6 ISSN: 0021-9606 *Abstracted Completely*
J. Chem. Phys. (USA) American Institute of Physics, 335 East 45th Street, New York, NY 10017, USA. Covered from 1968—
Frequency: 24/yr. BLLD: 4957.00000

図5.4 INSPEC List of Journals and other Serial Sources. Jan 1986より

(5) ニュースレター

図5.5のように，新しいデータ項目の追加などに関する詳細が基本的に無料で送付されてくる。また，当データベースを搭載している各ベンダ・システムを使って検索する際の検索式や注意点も解説されている。

inspec matters **USER UPDATE**

No. 30 SEPTEMBER 1987

Chemical & Numerical Data Indexing Revisited: A round-up of online system implementations

by Jeff Pache, Manager, INSPEC Tape Services

The new indexing (for material added to the Database from January 1987 onwards) has now been implemented on the following systems:

BRS	Dialog
CAN/OLE	ESA-IRS (Chemical Only as yet)
Cedocar	Orbit
Data-Star	

It is expected that ESA-IRS will soon implement the Numerical Indexing and that STN will implement both later this year.

[illegible] new fields and the principles be[illegible]
'INSPEC Matters No. 4[illegible]
[illegible] the following [illegible]

Orbit Field: MF
Searches:
Silicon in any role: /MF SI
Cadmium Mercury Telluride:
/MF 'CD/SS' (L) 'HG/SS' (L) 'TE/SS'
(Quotes and spaces are required because /SS is a reserved word.)

Numerical Data Indexing
BRS Field: NI
Searches:
Wavelength of 0.0000106m:
WAVELENGTH WITH '1.06' ADJ E-05.NI.
Power of 1 to 10 GW: POWER WITH E+09 WITH W.NI.
(Retrieves 1.0 to 9.9 . . . GW only)

CAN/OLE Field: ND
[illegible]hes:

図5.5 INSPEC Matters No.30 Sep 1987 より

(6) その他

図5.6は，データベースを検索する際のデータ項目ごとの検索タグなどを確認するための資料である。

SAMPLE RECORD

DIALOG Accession Number

722477 A87025836, B87012482
AN = AlGaAs/GaAs distributed feedback laser diodes grown by MOCVD /TI
AU = Ohata, T.; Honda, K.; Hirata, S.; Tamamura, K.; Ishikawa, H.; Miyahara, K.; Mori, Y.; Kojima, C.
CS = Sony Corp. Res. Center, Yokohama, Japan
Sponsor: Allied/Bendix Aerosp. Technol.; AT&T Bell Labs.; American Cyanamid; et al
PY =
JN = J. Cryst. Growth (Netherlands) vol.77, no.1-3 637-42 Sept. 1986
CO = CODEN: JCRGAE ISSN: 0022-0248
SN = Metalorganic Vapor Phase Epitaxy 1986. Proceedings of the Third International Conference 13-17 April 1986 Universal City, CA, USA CT=
CY =
U. S. Copyright Clearance Center Code: 0022-0248/86/$03.50 CL=
TC = Treatment: EXPERIMENTAL;
DT = Document Type: CONFERENCE PAPER
LA = Languages: ENGLISH

FILES 12,13

INSPEC
DIALOG FILES 12,13

SEARCH OPTIONS

BASIC INDEX

PAGE	SEARCH SUFFIX*	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
12-8	/AB	AB	Abstract	Word	S CLADDING(W)LAYER/AB
12-11	/DE	DE	Descriptor[1]	Word & Phrase	S FEEDBACK(W)LASER?/DE S III-V SEMICONDUCTORS/DE
12-15	/ID	ID	Identifier[2]	Word & Phrase	S LASER(W)DIODE?/ID S MESA-STRIPE GEOMETRY/ID
12-17	/TI	TI	Title and Conference Title[3]	Word	S DISTRIBUTED(W)FEEDBACK/TI

*If no suffix is specified, all Basic Index fields are searched.
[1]Also /DF.
[2]Also /IF.
[3]Searchable using CT= or /TI.

ADDITIONAL INDEXES

BIBLIOGRAPHIC FIELDS

PAGE	SEARCH	DISPLAY	FIELD NAME

図5.6 DIALOG INSPEC データベース解説資料より

1.4 サーチエイドの流通

サーチエイドの種類は，各データベースまたはベンダよりリストが発行されている。日本では，データベース作成機関，またはそのデータベ

ースを保有しているベンダの日本代理店からの購入も可能である。ベンダ代理店を通して利用契約を結べば，ベンダから送付されるニュースレターでサーチエイド新版発行などのニュースが送られる一方，無料で入手できるものもある。

指導上の留意点

サーチエイドの活用に関してもっとも注意が必要なのは，まずサーチエイド類（特にシソーラスや分類コード表）は，データベースによって毎年あるいは5年ごとなど，頻度はさまざまであるが常に改訂されるということである。シソーラスや分類コード表の改訂新版には，新しく統制索引語として導入された用語や新分類コードなどが含まれている。したがって，検索者はよく利用するデータベースの最新版サーチエイド類を必ず用意しなければならない。

一方，統制索引語や分類コードの年代による変更が問題である。ある主題内容に関する分類コードがある年代以降変更されている場合，最新分類体系によって過去のレコードをすべて再分類付与しているデータベースは問題ないが，再分類付与を行っていないデータベースでは年代ごとに有効な分類を使い分ける必要がある。このような場合，旧版も最新版と同様に検索式作成のために利用する必要があるので廃棄するわけにはいかない。

用　　語

サーチエイド，ベンダ，シソーラス，分類コード

第4章　データベース・サービスの国際的動向

目　　標

データベース・エンジニアは，データベースにかかわる知識を総合的に身につけておく必要があるとの観点から，国際的な市場，業界，技術，施策等の動向からとくに重要なものについて，できる限り日本との比較の上で説明し，理解させる。

内　　容

1.1　歴　　史

海外のデータベース・サービスがどういう経緯で発展してきているかについて理解させる。これによって，現在のサービスがどのような段階に位置しているのか，また，日本は欧米に比較してどういう状況にあるのかが分かるからである。

発展経緯はアメリカをベースにすると次のような3ステップに整理できる。

(1)　第1期：1960年代

政府助成，国家プロジェクト，技術移転等によるデータベースの整備育成。商用データベース萌芽期。

(2)　第2期：1970年代

科学技術情報を主体に商用データベースが発展。

国際ネットに乗ってアメリカのデータベースが海外へ流通。

(3)　第3期：1980年代

ビジネス情報が主流になり，データベースが多様化。

ユーザもビジネスマンなど裾野が拡大。

こうした過程で，アメリカは日本に対して，均衡のとれた情報の流通（シンメトリカル・アクセス）を要求している。一方，日本やヨーロッパは，アメリカの巨大データベースによる情報の大量流入への対応を迫られている。

1.2　アメリカのデータベース市場，業界，利用等の最新動向

国際的に最も進んでいるアメリカのデータベースの実情を以下の項目にそって説明し，理解させる。

(1)　市場動向

①　分野別市場　……　金融，ビジネス，経済，信用，法律，判例，科学技術などの主分野の市場推移について

②　メディア別市場　…　オンライン，MT，FD，CD－ROMなど主なメディアの市場規模推移について

(2)　業界動向

①　流通機構　……　プロデューサ，ベンダ，VAN，インフォメーション・ブローカ，政府機関などの分業体制，役割分担について

②　ベンダの特色　……　スーパー型，分野特化型，すき間型などの分類，特色と代表的ベンダについて

(3)　新しい話題

最近のアメリカのデータベースをめぐるホット・トピックスのうち，とりわけ日本とかかわりのあるものについて解説する。

①　CD－ROM

②　ゲートウェイ

③　AI

④　日本技術文献法

⑤　企業吸収・合併

⑥　トランザクション・サービス　　等

1.3　ヨーロッパ主要国の動向

ヨーロッパのデータベース産業は，従来，政府主導型で日本と似ているといわれているが，実情は必ずしもそうでないこと，各国がそれぞれ独自の方針を探していること，各国の施策が新局面を迎えていることなどを理解させる。

(1)　イギリス　…　民間活力を活かすのが基本方針で，ロイターの金融情報

など世界的ベンダも出現。

(2) フランス … QUESTEL／DARC（化学および特許データベースに特長）を政府助成で構築する一方，ビデオテックスを重視。

(3) 西ドイツ … 従来の政府主導を見直し，新情報政策で民間の競争力強化を重視。

1.4 その他

情報をめぐる国際的話題に関し，組織動向，背景等について理解させる。

① GATTにおけるサービス貿易

② TDF（国際情報流通）

③ 日本情報の海外提供

指導上の留意点

この章では，業界，技術等の細部には立ち入らずに，国際的データベース動向の中で日本がどのような位置づけにあるのかを理解させることにある。したがって，海外の動向を日米，日欧比較の視点で説明し，理解させることが重要である。

情報流通には国境がないから，国内の事象だけではなく，日頃から新聞，雑誌，TVなどで広く海外の出来事に関心を持つよう指導する。

表5.1　データベース・サービス発展の3ステップ

段　階	局　　面	時　　期
1	商用以前	1960年代
	政府の強力な支援 助成 ナショナル・プロジェクト	
2	商用化（技術主導）	1970年代
	科学技術分野主流 商用サービスの出現，本格化 金融等一部領域で利益計上 ユーザは技術者や学者など専門家	
3	市場拡大（ニーズ主導）	1980年代
	ビジネス分野主流 競争の激化 利益計上サービスの増大 ユーザ層はビジネスマン一般に拡大	

表5.2　アメリカにおけるデータベースの新しい潮流

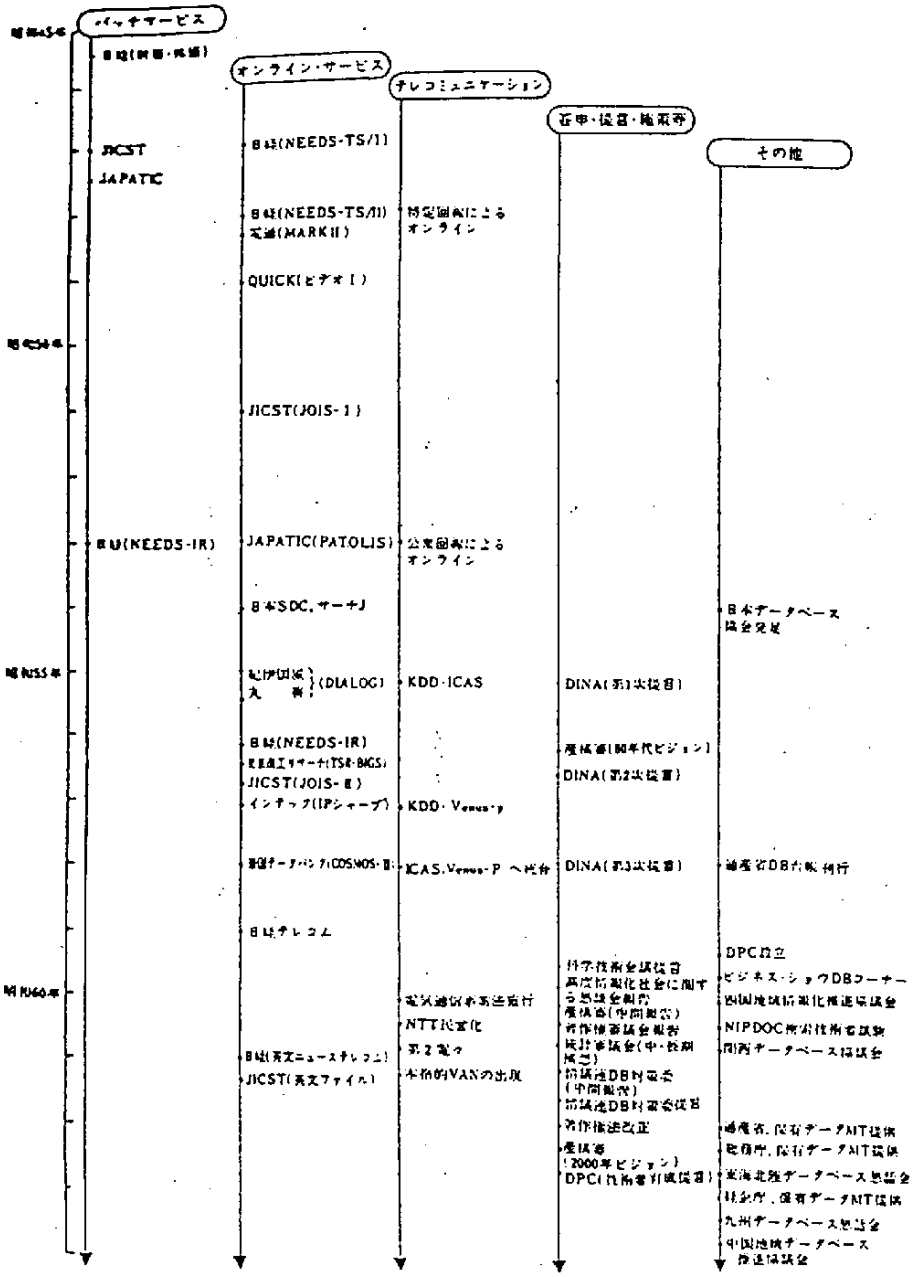
インテグレーション	ゲートウェイ 付加価値（サービス統合） ユーザ・フレンドリィ （AI，ｴｷｽﾊﾟｰﾄ•ｼｽﾃﾑ等） 企業吸収／合併
トランザクション・サービス	金融・ディーリング ショッピング 予約 バンキング
グローバライゼーション	金融 企業吸収／合併

表5.3　海外のデータベース年表

西暦	関　連　施　策	主なできごと
1951	米，国立科学財団(NSF) 設立	
1953		IBM,科学技術用IBM701発表
1955		IBM,事務計算用IBM702発表
1957	ソ連，スプートニク打上げ成功	
1958	米，エクスプローラ打上げ成功 米,NASA 設立	
1959		CAS,情報システム構築開始
1960		MIT,TSS 開発
1961	ソ連，有人衛星飛行成功 米，アポロ計画発表	CAS,Chemical Title発行
1963	米，ワインバーグ報告	NASA，STARの開発 NLM,Index Medicus 開発 SDC,米空軍委託でCIRCOL開発 NLM,MEDLARS 開発（バッチ）
1964	米,CFSTI（後，NTIS）設立	GE,GE/MARK開始
1965		CAS，Post 開発 IBM，IMS開発 IBM，IBM360 （汎用）発表
1966	米，情報公開法	ロッキード，DIALOG開発 ロッキード，NASA/RECON開始
1968		IBM，BROWSER開発
1969		ARPANET 稼働 BIOSIS，Previews開発 CODASYL，DBTG レポート発表
1970	米，NTIS設立 米，ＥＦコッド，リレーショナルデータベース概念発表	
1971		IBM，IBM370 発表 SDC,ORBIT Ⅱ開発 NLM，MEDLINE開始
1972		Software AG,リレーショナルDBMS(ADABAS)開発 SDC,Search開始 ロッキード，DIALOG開始 ニューヨークタイムズ，インフォメーションバンク開始
1974	米，プライバシー法案可決	
1975		Telenet，VANサービス開始
1976	米，情報公開法改定	BRS．サービス開始
1977	仏，プライバシー法案下院通過 加，プライバシー法成立	Tymnet Inc.，VANサービス開始
1978	仏，ノラ，マンク報告	
1979	加，クライン報告	Dow Jones，News Retrieval サービス開始 Compu Serve，CIS開始
1980	米，文書業務合理化法	Source,The Source 開始
1984		EURONET DIANE 完成
1985	EC，EUTELSAT利用の高速 デジタル通信網オープン（７月）	
1986	米，日本技術文献法の制定（８月） 英，Pergamon InfoLine，ORBETを買収（９月）	DIALOG，CD-ROMの第１弾発表（12月）
1987	英，ウォーリック会議（９月）	
1988	米，Knight，Ridder,ORBETを買収（７月） 米，マクミラン社，BRS を買収（11月）	韓国初のデータベース展示会 ＜Into-Seoul開催＞

（出所：データベース白書　1989年版）

表5.4　わが国データベース・サービスをめぐる経緯



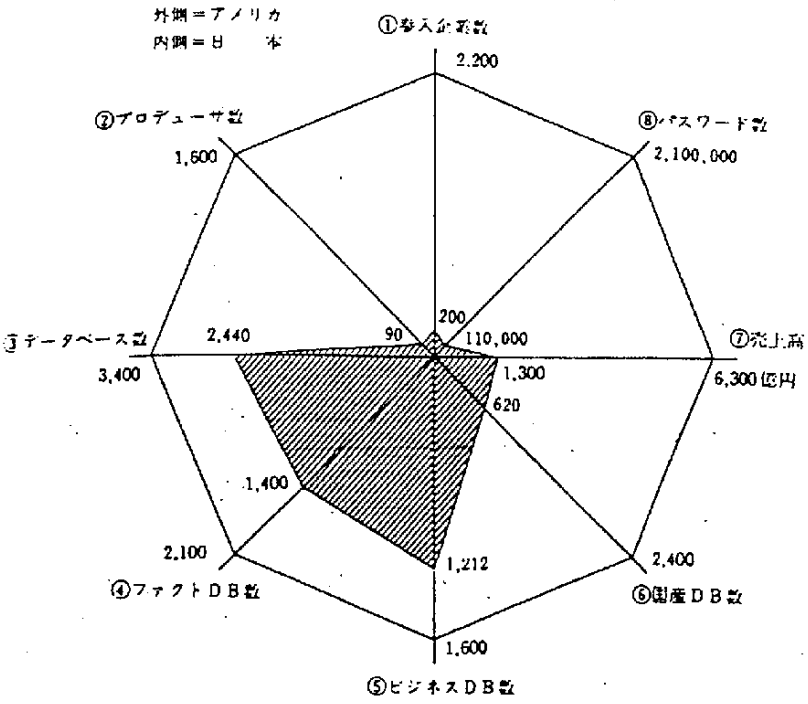


図5.7　日米データベース事情比較イメージ（1987年）

用　　語

スプートニク・ショック，ワインバーク報告，ノラ・マンク・レポート，流通機構，ＣＤ－ＲＯＭ，ゲートウェイ，シンメトリカル・アクセス，ＴＤＦ（11Ｆ），トランザクション・サービス，スーパー型ディストリビュータ，プロデューサ，インフォメーション・ブローカ，政府の振興施策，非課税障壁，ＡＩ

第5章　知的所有権と個人データ保護

目　　標

この章では，データベース・サービスにかかわる法制的課題のうち，知的所有権と個人データ保護について述べる。法制度はそれ自体厳密に構成された体系であり，したがって固有の論理，用語などを持っている。それは技術的体系と必ずしも一致しないので，その内容を正確に理解することが必要である。また，この章で述べる課題は法制的に見て新しいものであり，判断が不明確な領域が残っている。これについても十分に配慮する必要がある。

内　　容

1.1　知的所有権

(1)　著作権全般

データベースにかかわる知的所有権は，著作権法で保護されることになっている。したがって，まず著作権法の理解が前提となるので，内容について簡単な説明をする。

①　知的所有権とは何か．著作権は工業所有権（特許を代表例として）とどこが違うか．

②　著作権で保護される対象は「表現」である．

③　著作権には情報のコピーに対して，「複製」「翻案」の2種類がある．

(2)　データベースとのかかわり

データベースについて，とくに著作権法を意識しなければならない点を説明する。

①　原データとの関係

原データに著作権のある場合（文章情報，画像情報など）とない場合（数値情報，法律・判例など）とがある．

②　原データの抄録について

原データ作成者へ許諾を求めることが必要か否かについては，業界によって慣行が異なる．

③ ディストリビュータ権利

理論上は2次著作権は有線送信権があるが，権利の内容は未確定である.

④ ダウン・ローディング

複製に相当する.

⑤ 法人著作

法15条に注意する必要がある.

⑥ データベース開発，運用の外注

データ作成者が権利を主張するためには，契約に明示する必要がある.

⑦ シソーラス

シソーラスは著作物である.

(3) データベース取引の現状

現実の取引関係を著作権法のもとで検討，整理してみる。

① データ作成者〜プロデューサ間

② プロデューサ〜ディストリビュータ間

③ ディストリビュータ〜ユーザ間

1.2 個人データ保護

「プライバシー」という言葉は，コンピュータ化された情報に対しては，「自分に関する情報をコントロールする権利」という意味にまで拡張されている。ただし，最近は「プライバシー保護」(マニュアル・データまで含む）というよりも，むしろ「個人データ保護」（コンピュータ処理データのみを扱う）という概念で把握することが通例になっている。

(1) 個人データの定義

1988年に総務庁が公表した個人情報保護法案によれば，個人データを「生存する個人に関する情報であって，氏名，番号，記号その他の個人別に付された識別項目または個人に関する記述により個人を識別できるもの」と定義している。

(2) 制度の枠組み

個人データ保護に関する法律は，国によって次のように異なる。

① 政府所有のデータのみを対象とするのか，民間の持つデータまで含むのか.

② 個人データのみを対象とするのか，法人データまで含むのか.

③ コンピュータ・データのみを対象とするのか，一般文書のデータまで含むのか.

日本では個人データ保護法はない。ただし，地方公共団体では個人情報保護条例のあるところが多い。また，民間セクターでも個人情報を扱う業界にはさまざまの公的，私的規制がある。

(3) 制度の理念

行政管理庁は，1982 年に『個人データの処理に伴うプライバシー保護対策』という報告書を公表している。これによれば，プライバシー保護として次のような「5原則」を掲げている。

① 収集制限の原則

② 利用制限の原則

③ 個人参加の原則

④ 適正管理の原則

⑤ 責任明確化の原則

なお，規制対象としてはコンピュータ・データとマニュアル・データとを，また官庁データと民間データとをそれぞれ含むとしている（法人データについては別途検討の要ありとしている）。

この5原則は，ＯＥＣＤが80年に発表した『プライバシー保護と個人データの国際流通についてのガイドラインに関する理事会勧告』を整理したものである。

(4) 具体的な手続き

公共領域では，統計法で個人データの保護を規定している。また，住民基本台帳法が1985年に改正され，その閲覧に制限が加えられるようになった。地方公共団体の条例には，個人データの扱いについてさまざまの規定を設けている。

民間領域では，さまざまの分野で自主規制のルールが設けられている。まず，国民生活審議会は81年に『消費者信用情報機関の適正な運営につい

て』という報告を提出している。ここには，個人信用情報システムについて次のような配慮事項が示されている。

① 運営方法の明確化，機関の存在と趣旨の周知

② 同意文言による事前承認

③ 本人への開示，事実に反する場合の訂正

④ 登録する情報の範囲の適正化

⑤ 情報登録基準の明確化

⑥ データの漏洩防止，目的外使用の禁止

⑦ 会員の資質保持

このほかにも，貸金業法，割賦販売法では情報機関の指定が義務づけられている。当然ながら，この種の機関に対しては個人情報保護についてのルールが制定されている。

1987年，金融情報システム・センターが金融分野における個人情報保護の取り扱いについてガイドラインを発表した。翌88年には，(財)日本情報処理開発協会が民間の各事業分野における個人情報保護についてのガイドラインを発表している。

(5) 地方公共団体の条例に対する配慮

これは公的領域のものではあるが，さまざまな点で企業にも影響する。たとえば，次のような点である。

① 住民基本台帳の閲覧

② 業務の受託

③ データベースの納入

(6) 個人データの国際流通に関する配慮

データベース利用を含むデータ通信は国際化をしており，海外センターにおける情報処理，海外センターと日本間の国際通信において，各国の制度に基づいた規制を受ける。

まず，海外において情報処理を行う事業者は，その国のプライバシー保護法を遵守する義務があることになる。また，プライバシー情報については国際通信を禁止される場合がある。たとえば，ＥＣ条約ではプライバシー保護制度のない国への個人情報の流出を禁止している。この条項に対応

して国内法を制定している国に，フランス，イギリスなどがある。

指導上の留意点

データベースを著作権で保護しようという判断は，1986年の著作権法改正によって明確となった。これは新しい秩序であり，まだこれに基づいた判例はない。一般に，システム技術者はものごとの白黒をはっきりとつけたがる性格を持っているが，こと法律では，判断に裁量の余地が残される場合が多い。この不明瞭な点を明確にした場合，将来の技術発展の可能性が法的に拘束されることも考えられる。

他方，個人データについてはプライバシー保護の問題などがあるが，法体系の整備は遅れているのが現状である。また，この問題では，日本国内でもまだ統一した見解には至っていない。したがって，この扱いに関しては慎重な配慮が必要である。

用　　語

知的所有権，著作権，著作物性，原著作物，２次著作物，編集著作物，データベースの著作物，複製，翻案，有線送信権，ダウン・ローディング，プライバシー，個人データ，センシティブ・データ，ＯＥＣＤ勧告（８原則），ＥＣプライバシー保護条約，プライバシー保護５原則（行政管理庁），指定統計のＸ表示（秘匿データ），住民基本台帳，信用情報センター，メイル・プレイファリング・サービス（ＭＰＳ）

参考文献

加戸守行：著作権法逐条講義（３訂），著作権資料協会，79年

吉田大輔：データベース・サービスと著作権問題，著作権資料協会，81年

著作権審議会第７小委員会（データベースおよびニューメディア関係）報告書，85年，文化庁

名和小太郎：最新データベース事情，11章，日本能率協会，84年

データベース・サービス業連絡懇談会：活用データベース，９章，東洋

経済新報社，85年
データベース振興センター：データベース白書1986，Ⅱ－三章，86年
名和小太郎：データベースの著作権，『法とコンピュータ』４号，86年
高石義一：コンピュータ法務，４章，にじゅういち出版，87年
米国議会技術評価局（日本電子工業振興協会訳）：電子・情報時代の知的所有権，日経マグロウヒル，87年
松田政行：コンピュータ・ビジネス・ロー，５章，商事法務研究会，87年
名和小太郎：情報を商品にする方法，14～15章，東洋経済新報社，88年
行政管理庁行政管理局：プライバシー保護の現状と将来，ぎょうせい，82年
行政管理研究センター：世界のプライバシー法（改訂），ぎょうせい，82年
経済規格庁国民生活局：情報化時代の消費者政策，大蔵省印刷局，85年
名和小太郎：電子仕掛けの神，４章，７章，勁草書房，86年
経済規格庁国民生活局：民間部門における個人情報の保護（調査編，資料編）大蔵省印刷局，87年
総務庁行政管理局：行政機関における個人情報保護対策，ぎょうせい，87年
金融情報システムセンター：金融情報システム白書（昭和62年版），財経詳報社，88年
名和小太郎：ＩＣカードとプライバシー保護，『電気通信時報』５巻２号，88年
堀部政男：高度情報社会とプライバシー保護，岩波書店，88年

第6部　データベースの技術動向

第6部　データベースの技術動向

目　　標

データベースに関連した新しい技術動向に関する知識を習得する。

内容のあらまし

内　　容	時間数
マルチメディアデータベース 　マルチメディアデータベースとは 　マルチメディアデータベース出現の背景 　国内の研究動向 　マルチメディア文書の通信技術 　ユーザインタフェース	2時間
フルテキストサーチ 　フルテキストサーチの概念 　マッチングアルゴリズム 　応用例	1時間
分散型データベース 　分散型データベースとは 　分散型データベース出現の背景 　分散型データベースシステムの実現形態	2時間
データベースマシン 　データベース基本演算の高速化 　記憶制御の高速化 　その他の高速化手法 　データベースマシンの事例紹介	2時間

内　　容	時間数
データベースへのＡＩの応用 　提供側への適用 　利用側への適用 　知識データベース管理システム	2時間
新記憶媒体 　光ディスク 　光ディスクの種々	1時間
ニューメディア通信機能 　ビデオテックス 　ＶＲＳ 　ＣＡＴＶ 　テレテキスト 　ＡＮＳＥＲ 　ＩＳＤＮ	1時間
電子出版	1時間
合　　計	12時間

第1章　マルチメディア・データベース

目　　標

マルチメディア・データベースの概念，出現の背景，国内の研究動向，通信技術，ユーザ・インタフェースなどについて，システムを構築する立場として必要な知識の修得を目標とする。

内　　容

1.1　マルチメディア・データベースとは

現在のところ，マルチメディア・データベース・システム（Multimedia Database System）は実用化されていない。しかし，これに対するニーズは高まってきており，実用化へ向けて国内外で研究開発が進んでいる。

マルチメディア・データベースにおけるメディアとは何か。従来のデータベースが文字・数字を主体に扱っていたのに対して，マルチメディア・データベースでは，さらに文章（テキスト），図形，イメージ，音声といった新しいタイプのデータも取り扱うことになる。しかも，文字・数字データベース，文章データベース，図形データベース，イメージ・データベース，音声データベースというように個々に構築，管理，運用するのではなく，それが統一的に行われるのがマルチメディア・データベースであるといえる。

1.2　マルチメディア・データベース出現の背景

マルチメディア・データベースのニーズと実用化に対する期待は高い。マルチメディア・データベースが可能となってきた理由として，以下の3点が挙げられる。

(1)　データベース管理技術の発展

データベースの研究が進み，階層型およびネットワーク型のＤＢＭＳ（データベース管理システム）に加えて，今日では，データ構造の柔軟性と高度のユーザインタフェースの提供が容易なリレーショナル型のＤＢＭＳが主流になるまで発展している。

(2) 通信技術の発展

ディジタル通信をはじめ，光通信，衛星通信等，通信技術に関する技術革新は激しい。その結果，文字，数値，文章，図形，イメージ，音声等の異なるメディアを同時に送ることが可能となっている。

(3) 大容量蓄積技術の発展

光ディスクなどの大容量の蓄積技術が進歩している。

1.3 国内の研究動向

通商産業省の工業技術院の大型プロジェクトとして，1985年末から91年末までの計画で「電子計算機相互運用データベース・システム」がスタートしている。この計画は，工業技術院電子技術総合研究所とＩＮＴＡＰ（情報処理相互運用技術協会）が中心に進めており，研究開発の主要テーマのひとつとしてマルチメディア・データベース・システムの開発がある。このプロジェクトの全体イメージを図6.1に示しておく。

東京大学では，文字（テキスト），図形（グラフ），画像（イメージ），音，映像（動画）が取り扱えるワークステーションＰＩＥ（Professional's Integrated Environment）を開発中である。構成は図6.2に示すとおりであるが，入出力装置として光ディスク装置やファクシミリ装置を用いており，ＬＡＮを経由して他のワークステーションとの交信も可能なように設計されている。

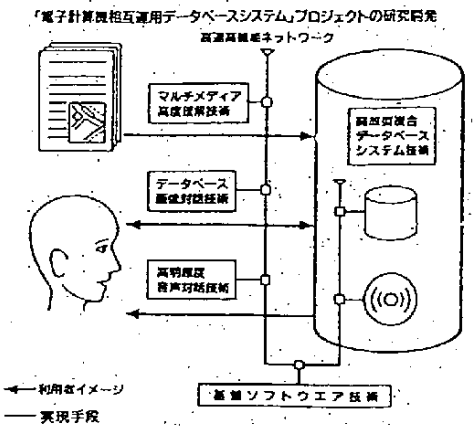


図6.1

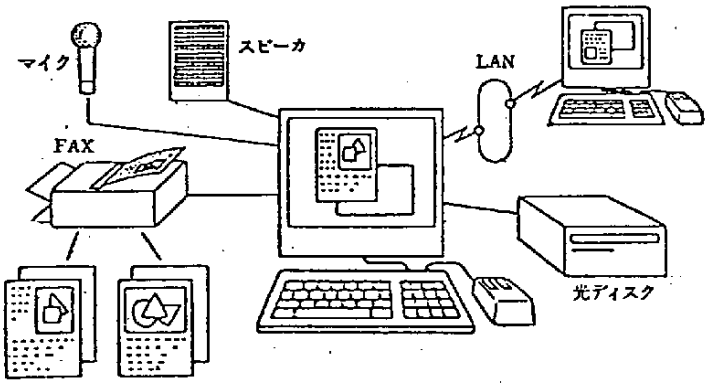


図6. 2

1. 4　マルチメディア文書の通信技術

マルチメディア・データベースの普及には，そこに蓄積されたデータで構成するオフィス文書（マルチメディア文書）を通信する技術が不可欠である。このような性質を持ったオフィス文書を相互に通信するには，そのためのプロトコルが必要である。

文書通信プロトコルの国際標準化について，まずＩＳＯ（国際標準化機構）がＴＣ９７／ＳＣ１８で，交換規約，構造規約，表現規約のすべてについて検討を進めている。また，ＣＣＩＴＴにおける文書通信に関連するプロトコルとしては，まずＳＧⅦのＭＨＳがある。これは国際標準電子メールプロトコルであり，ほぼ交換規約に対応する。次にＳＧⅧのテレマティークサービスがあるが，これは非音声系の通信サービスの総称であり，1984年には文章とイメージから構成されるマルチメディア文書を対象としたミクストモード通信の標準規約を勧告した。さらに，1986年にはミクストモード通信の改版が勧告された。

マルチメディア文書の通信はまだ研究段階にあるが，今後はワークステーションが持つ機能との結合，国際標準プロトコルとメーカ個別のプロトコルとの接続が課題となろう。

1. 5　ユーザ・インタフェース

マルチメディア・データベースのユーザは，従来のデータベースのユーザ

に比べて，コンピュータに対する親しみがより少ないと考えられる。したがって，ユーザ・インタフェースとしては，視覚的インタフェースや自然言語（日本語，英語）等の知的インタフェースの応用も必要となる。

指導上の留意点

マルチメディア・データベースの実用化に向けて解決すべき課題は多い。とくに重要な技術は，ＤＢＭＳ，データモデル，通信規約，ユーザ・インタフェース，大容量蓄積技術である。

用　　語

マルチメディア・データベース，マルチメディア文書，ユーザ・インタフェース

第2章　フルテキストサーチ

目　　標

フルテキストサーチの概念，代表的な手法（アルゴリズム）とその応用例などについて説明し，システムを構築する立場として必要な知識を修得することを目標とする。

内　　容

1.1　フルテキストサーチの概念

データベース・システムや情報検索システムでは，利用者の質問に対する応答性を向上させるためにインバーテッド（転置）ファイルを作成する場合が多い。

データベースに格納されるデータの形態は文字データや数値データが多いが，最近では図形データや画像データの割合も増えてきている。文字データベースでは，文献検索に代表されるように書誌的事項等の2次情報を対象に検索し，フルテキスト（全文）については表示（出力）に限る場合が多い。

フルテキスト・データベースの提供が今後も増える傾向にあるだけでなく，フルテキストを対象に文や語の位置なども考慮してサーチするフルテキストサーチのニーズが高まってきている。したがって，フルテキストサーチはデータベース技術として欠くことのできない技術のひとつである。

1.2　マッチング・アルゴリズム

フルテキストサーチは，文字列のマッチング（照合）を基本技術としている。なかでも，マッチング・アルゴリズムの研究は古くから行われており，これまでにもいくつかの手法が示されている。

(1)　非前処理方式

最も古典的なアルゴリズムとして，テキストの内容をそのままサーチ用の辞書（索引ファイル）などに登録する方法がある。辞書の構造としては，

ハッシング，二分岐（binary tree），B-tree が採用されている。

(2) 前処理方式

フルテキストサーチを高速で処理するには，テキストに対して何らかの前処理が必要である。

① 移動ロード方式

移動ロード方式（shifted preloading）は，単語を1文字ずつ移動させて複数パターンを登録することにより，高速でフルテキストサーチを可能とするものである。

② 番号（signature)付与方式

単語，熟語，句等，任意の文字例に対して数桁の番号を付与し，サーチ時にはこの数字を用いる方式である。これは，文字列より数字を対象にサーチした方が高速処理が可能というコンピュータの特性を利用したものである。

1.3 応用例

情報検索システムにおけるフルテキストサーチの一例を紹介する。

文　字　例	項目番号	文 番 号	単語番号
unstructured	1	1	1
data	1	1	2
bases	1	1	3

図6.3　フルテキストサーチ

この表は，対象とする文字例がどの項目に該当し（項目番号)，フルテキスト中のどの文中に出現するかを示している（文番号)。また，隣接している文字列（単語）は何か（単語番号）もわかる。このような表（すなわちインバーテッドファイル）をあらかじめ作成しておけば，文中の文字位置を指定した高速サーチが可能となる。

指導上の留意点

フルテキストサーチを高速で処理するには効果的な前処理が必要である。その結果，単に応答時間が短縮されるだけでなく，必要記憶容量の削減，あいまい検索，スペルミスの検出等も可能となる。なお，ハードウェアによるフルテキストサーチの研究も進んでいる。米国ＧＥ社をはじめ，国内では日本電気，日立製作所などでフルテキストサーチ用の高速検索専用装置の製品化が進められている。

用　　語

テキストサーチ，マッチング・アルゴリズム，インバーテッド・ファイル

第3章　分散型データベース

目　　標

分散型データベースの概念，出現の背景，実現形態などについて，システムを構築する立場として必要な知識を修得することを目標とする。

内　　容

1.1　分散型データベースとは

分散型データベースとは，地理的に分散して発生したデータを1カ所に集中させず，論理的には全体をひとつのデータベースとして運用する形態をいう。

集中型データベースに対して，分散型データベースの利点を挙げると以下のようになる。

(1)　一部の障害が全体に波及しないので信頼性が高くなる。

(2)　場所間（センターとローカル）のやりとりが減少するので，通信コストが削限できる。

(3)　システムの拡張など構成上の柔軟性に優れている。

1.2　分散型データベースシステム出現の背景

SystemR*や分散型ＩＮＧＲＥＳなど，すでに実用化されている分散型ＤＢＭＳがある。この他にも，多くの分散型ＤＢＭＳの研究開発が進められている。このように，研究開発および実用化へ向けての動きが活発になった理由は次の3点に集約できる。

(1)　リレーショナル型ＤＢＭＳが実用化された。

(2)　コンピュータネットワーク技術が確立された。

(3)　分散型データベースに対するユーザニーズが高まった。

1.3　分散型データベースシステムの実現形態

ここでは，分散型データベースシステムの実現形態を考える。各サイドにあるＤＢＭＳが，同種の場合と異種の場合では分散型データベースシステムの実現形態は異なる。データモデルとＤＢＭＳの２つとも異なるＤＢＭＳで分散型データベースシステムを実現する場合が，最も複雑な形となる。

(1)　ＤＢＭＳが同種の場合

最も単純な同種のＤＢＭＳで分散型データベースシステムを実現する場合には，すでにあったＤＢＭＳに設計変更を加える場合が多い。SystemR*や分散型ＩＮＧＲＥＳなどがその例であり，データモデルとしてリレーショナルモデルを採用しているのが特徴である。

(2)　ＤＢＭＳが異種の場合

異なるＤＢＭＳを統合して分散型データベースシステムを実現する場合には，データモデルとＤＢＭＳ自身の異種性を克服する必要がある。採用しているデータモデルが同じでも，ＤＢＭＳの実現方法が異なる場合と２つとも異なる場合がある。

使用するＤＢＭＳが異種である分散型データベースでは，実用化の域に達した例はない。統合のための共通データモデルとしては，リレーショナルモデルを採用する場合が多い。だが，実際には異なるデータモデル間の変換が十分にできるのかという課題が残る。また，異種のＤＢＭＳを同時に運用する場合の障害回復機能についても調整する必要がある。

ＯＳＩ（開放型システム間相互接続：Open Systems Inter-connection）では，互いに離れたシステム内にあるデータベースにアクセスするための規格ＲＤＡ（Remote Database Access）を検討中である。ＲＤＡは，７層の最上位である応用層上に実装される。ＤＢＭＳどうしが通信する際に使用するデータモデルはリレーショナルモデルであり，アクセス言語としてはＩＳＯならびにＪＩＳ規格に適合しているＳＱＬ（Structured Query Language）が使用される。ＲＤＡだけでは，異種のＤＢＭＳからなる分散型データベースシステムを実現することはできない。だが，利用者サイドに立った場合，この概念に近い使い方ができる。

指導上の留意点

分散型データベースシステムを構築する際に用いるＤＢＭＳには，同種であるか異種であるかに関係なく，共通した課題がある。

まず，質問処理の最適化（guery optimization）がある。利用者の質問（問い合わせ）の解釈と実行を，どのサイドでどういう手順で行うかがポイントである。そして，最適化の評価はコストで行われる。このためには，データ（リレーションの一部あるいは全体）を重複させることもありうる。

ＤＢＭＳが異なると，とくに同時実行制御，機密保護機能および障害回復機能などのデータ制御機能に関する調整が重要なテーマになる。

用　　語

分散型データベース，ＤＢＭＳ，同種，異種，ＯＳＩ，ＲＤＡ

第4章　データベースマシン

目　　標

データベースの主要演算とデータベースマシンの特長機能の関連を説明し，各種データベースマシンの特長を理解させる。

内　　容

1.1　データベース基本演算の高速化

関係データベース処理を中心に，基本演算と対応するハードウェア・アルゴリズムを説明する。

(1)　ソート

各種アルゴリズムが提案されているが，並列性が高く，かつ効率のよい方式はまだ考案されていない。

①　パイプライン・マージソート

②　パイプライン・ヒープソート

③　バイトニックソート

(2)　サ ー チ

サーチ対象が整列済か否かにより2つに大別できる。さらに，サーチキーが1個か複数（一括サーチ）による分類も可能である。

(3)　ジョイン（多数テーブル間結合）

(4)　プロジェクション（重複の排除）

(5)　アグリゲーション（総和等の集計演算）

1.2　記憶制御の高速化

データベース処理において，ディスクからの読出時間は常に大きな問題である。さらに，読み出したデータのうち，必要データは10％で残りの90％は不要である（90－10ルール）とも言われている。この観点からの高速化手法について説明する。

(1) オンザフライ (On the Fly) 処理

ストリーム処理の一例であり，ディスクからの読み出し・サーチ・転送を重畳させることができる。

(2) 主記憶データベース

アクセス頻度の高いデータベースを主記憶上に常駐させる方式であり，入出力処理起動のオーバヘッドも削減できる。ワンレベルストレッジの概念もその一例と考えられる。

1.3 その他の高速化手法

(1) クラスタリングと並列処理

ディスクなど，記憶媒体におけるデータ格納技術法としての静的クラスタリングと，演算の実行時にキー値等により動的に分割を施す動的クラスタリングに大別される。いずれにも並列処理との整合性があることを説明する。

(2) インデックス

(3) ハッシング

(4) 連想メモリ

以上は，基本手法として簡単に説明する。

(5) 同期排他制御

(6) プロテクション制御

(8) 障害回復制御

(9) 可変長データ，可変構造データ

以上は，実用システムとしては必要な機能ではあるが，高速化の観点からは阻害要因となりやすいことを理解させる。

1.4 データベースマシンの事例紹介

各データベースマシンの特長技術と目的とする処理内容を明確化する。

(1) ＲＡＰ(Relational Associative Processor)

On the fly 処理の代表例として説明する。

(2) ＤＢＣ(Date Base Computer)

クラスタリングにより大規模データベース(10^{10} バイト程度）を可能にした例として説明する。

(3) Teradata社ＤＢＣ／1012

商用データベースマシンおよび並列マシンの代表例として説明する。Ｙネットについても説明が必要である。

(4) ＧＲＡＣＥ

日本の代表的データベースマシンであり，動的クラスタリング，パイプラインマージソータ，機能ディスク等について説明する。

指導上の留意点

ここで挙げた高速化のための各種手法は，データベースシステムの構築時に参考になるような形で紹介する。

用　　語

関係代数演算，マージ・ソート，ヒープソート，バイトニックソート，オンザフライ処理，ストリーム処理，主記憶データベース，クラスタリング，Ｙネット

第５章　データベースへのＡＩの応用

目　　標

この章では，データベースの構築・作成支援，運用面の支援，利用面でのユーザ支援等に影響を与えるであろう人工知能（ＡＩ）について説明する。

内　　容

1.1　提供側への適用

(1)　データベース構築

①　データベース設計に対しては，データのモデル化，論理構造，物理構造の定義，インデックス設定の知識などのＤＢ設計エキスパートシステムの適用.

②　設計に伴う性能予測をするためのアクセス頻度，応答速度，データ構造や配置調整のためのシミュレーション・エキスパートシステムの適用.

③　アプリケーション開発に対する，知的ソフトＣＡＤエキスパートシステムの適用.

(2)　データベース作成

①　ＯＣＲや音声認識に導入されている単語知識，文法，文脈知識，図形，画像で適用されている線分・形状知識など，人工知能技術の進展に伴ったデータ入力支援.

②　文献データベースのキーワード付け，抄録作成，外国語の翻訳など自然言語処理技術の応用によるデータベースの作成支援.

(3)　データベース運用

①　アクセス状況に応じたデータの再配置，データ格納域の利用状況監視と自動調整等データベースの再編成処理を行うデータベース運用エキスパートシステムの適用.

②　異常検出，アラーム出力，リカバリ等の運用状況変更に伴うシステム評価，シミュレーション等を行うシステム運用エキスパートシステムの適用.

③ ユーザ管理面でも，ユーザを識別するための音声認識，画像認識等人工知能技術の応用.

1.2 利用側への適用

(1) 利用の準備

データベースの検索訓練や自然言語処理技術を応用したクリアリング・システム。

(2) データベースの利用

マン・マシン・インタフェースの改善，共通の検索言語の実現のほか，自然語による検索やリアルタイム翻訳などへの応用。

(3) 内容の高度利用

データベース中の内容を高度に加工するための各種エキスパートシステムへのデータの取り込み。

1.3 知識データベース管理システム

ＤＢＭＳに関しては，管理するデータの多様性と広汎な属性関係を的確で柔軟に定義するために，意味論モデルに基づいた知的ＤＢＭＳが考えられる。

指導上の留意点

これらの技術は，一部を除いて実用化されてはいないが，近い将来の技術として知っておいた方がよい知識である。

用　　語

人工知能，ＡＩ，ＤＢ設計エキスパートシステム，
シミュレーション・エキスパートシステム，
知識ソフトＣＡＤエキスパートシステム，ＯＣＲ，音声認識，単語知識，
文法，文脈知識，線分・形状知識，自然言語処理技術，
データベース運用エキスパートシステム，
システム運用エキスパートシステム，画像認識，リアルタイム翻訳，

知識データベース管理システム，意味論モデル，知的ＤＢＭＳ

第６章　新記憶媒体

目　　標

この章では，その容量から画期的な記憶媒体としてデータベースのハード，サービスに変化をもたらすことが期待されている光ディスクについて説明する。

内　　容

1.1　光ディスク

光ディスクは，レーザ光による光の針を用いてデータの書き込みと読み出しを行うもので，高密度記録による大容量と低コストを最大の特徴としている。

このため，図形や写真など，今までのメモリでは実用化できなかったイメージデータをＤＢ化することを可能にした。たとえば，光ディスク１枚にＡ４版のイメージデータを入力する際，そのコストは１円程度にしか過ぎない。また，１枚当たりでＡ４版６万枚のデータを収納することもできる。さらに，ランダムアクセス性，ファクシミリ伝送網や光ＬＡＮとの整合性からデータベースの画期的展開をもたらすことも期待できる。このように，大容量かつ低価格のパッケージメディアとして，データベース拠点を従来の集中型から分散型へ変貌させるのに十分の能力を保持している。

1.2　光ディスクの種類

光ディスクは大別すると，ＣＤ－ＲＯＭ，ＤＡＤ，家庭用ビデオディスクなどのように再生専用のものと記録可能のものとに分類され，さらに追記型と書換型とに分類される。

記録可能なものは，通常のオンラインデータベース用として優れた性能があり，再生専用のＣＤ－ＲＯＭは大容量かつ低価格のパッケージメディアとして，個人向けのデータベースが販売されている。

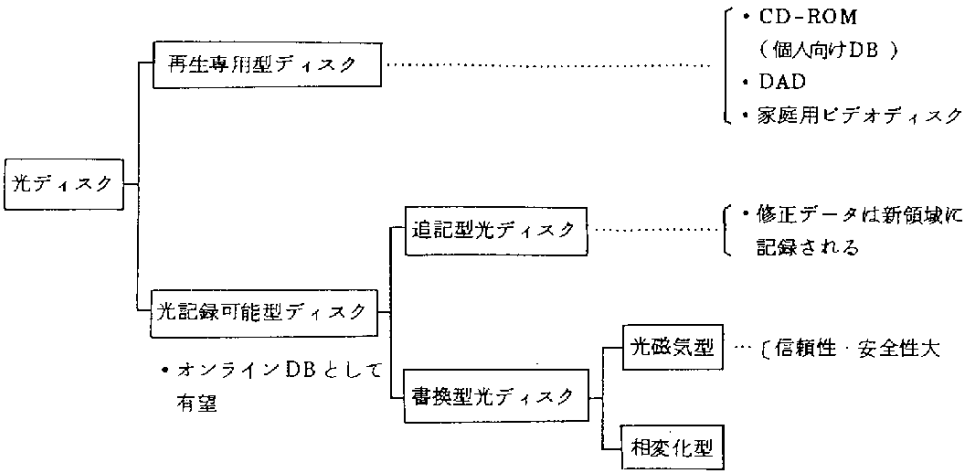


図6.4　光ディスクの種類

指導上の留意点

実用化されているメディアであることを説明するとともに，イメージ情報の記録に優れた特徴があること，文字情報と組み合せることで一層効果的であることなども合わせて説明する。

用　　語

光ディスク，レーザ光，高密度記録，大容量，低コスト，イメージデータ，ランダムアクセス性，ファクシミリ伝送網，光ＬＡＮ，パッケージメディア，集中型，分散型，ＣＤ－ＲＯＭ，ＤＡＤ，家庭用ビデオディスク，再生専用型ディスク，光記録可能型ディスク，追記型光ディスク，書換型光ディスク，光磁気型，相変化型

第７章　ニューメディア通信機能

目　　標

この章では，オンライン・データベースとして将来性のあるニューメディアについての知識を学習する。

内　　容

1.1　ビデオテックス

ビデオテックスは，コンピュータとテレビ受像機，パソコン等の端末を通信回線で結び，静止画像情報を提供する双方向図形通信システムで，通信と処理の融合によって生まれた新しいメディアである。

ビデオテックスには，一般公衆を対象としたパブリック型と企業や特定地域を対象としたプライベート型とがあり，プライベート型は広汎メディアとして有力であるため，地域密着型の情報分野での利用が期待されている。

1.2　ＶＲＳ

ＶＲＳは，コンピュータとテレビ受像機などを結んだ画像情報システムで，広帯域伝送路を使用しているため，音声付カラー画像を提供できる。

1.3　ＣＡＴＶ

ＣＡＴＶは，双方向機能と情報提供手段としての有用性が注目されており，都市型ＣＡＴＶに関心が高まっている。

1.4　テレテキスト

テレテキストとは，テレビ放送電波のすき間を利用して，放送中の番組とは別に文字や図形で構成された静止画情報を提供するものである。テレテキストは，数多くの情報の中から受け手がニーズに応じて欲しい情報を選択でき，幅広い分野の情報提供用のツールとしての可能性がある。

1.5　ANSER

ANSERは，電話器からの問い合わせに対してコンピュータからの音声応答をするメディア変換システムであり，現在，残高照会や入金通知，解約受付，オーダ・エントリなどに利用されている。

1.6　ISDN

音声，ファクシミリ，動画，データなどを多重化して電話による交信ができるので，ビジネスをはじめとして遠隔地教育等その利用範囲は広い。

指導上の留意点

ここで挙げたニューメディアは，今までの概念とは異ったデータベースが提供できることを説明する。

用　　語

ビデオテックス，通信回線，静止画像情報，双方向図形通信システム，パブリック型，プライベート型，VRS，画像情報システム，広帯域伝送路，音声付カラー画像，CATV，双方向機能，都市型CATV，テレテキスト，テレビ放送電波，ANSER，ISDN，遠隔地教育

第８章　電子出版

目　　標

データベース関連の新しい動きとしての電子出版に関する知識を習得する。

内　　容

データベース・サービスとのかかわりで見ると，電子出版は電子技術の進展により，大型コンピュータを用いる電算写植（ＣＴＳ）からパソコンを用いる電子編集出版（ＤＴＰ）まで，多様な姿でデータのコンピュータ可読ファイル化を容易にし，このことがデータベース・サービスの新たな展開を促すところとなりつつある。電子出版をサービスとしてとらえると，①オンライン・データベース型（オンライン情報検索サービス，ビデオテックス），②パッケージ型（フロッピーディスク，ＣＤ－ＲＯＭ，ＣＤ－Ｉ，ＩＣカード），③コンピュータ通信型（電子雑誌）のようになる。データベース・サービスは，現在，オンライン・サービスが中心となっており，さまざまなメディアが交錯する電子出版の世界では，ＣＤ－ＲＯＭが，①情報単価が安い，②大量生産できる，③パソコンに利用しやすいなどの理由により，大量情報の蓄積・流通においてオンライン・サービスに次ぐ位置となろう。

指導上の留意点

コンピュータ出版の副産物として，電子出版はデータベースの供給源となりうることを学習させる。

用　　語

電子出版，電子写植，電子編集出版，電子雑誌

参考文献

増永良文，“マルチメディアデータベース総論”情報処理，vol.28，No.6，June，1987

原田史郎，“オフィスシステムとマルチメディア文書通信技術”情報処理，vol.28，No.11，Nov.，1987

牧之内顕文，尹博道，“マルチメディアデータベースの利用者インタフェース”情報処理，vol. 28,No.6，June. 1987

G. H. Gonnet，“Unstructured Data Bases or Very Efficient Text Searching”，Proceedings of the Second ACM SIGACT-SIGMOD Symposium on Principles of Database Systems

増永良文，“分散型データベースシステム”情報処理，vol. 28,No.4，Apr.，1987

A. Tojo，“The Interoperable Database System Project and OSI Promotion in Japan.”

Proceedings of International Symposium on Interoperable Information Systems，Feb.

田中譲：サーチ，ソート・ハードウェアと記号処理への応用，情報処理，Vol. 23,No.8 (1982)

Ozkarahan，E. A. etal.:RAP-An Associative Processor for Data Base Management，Proc. AFIPS，Vol. 44

Banerjee，J. etal.: DBC-A Database Computer for Very Large Databases，IEEE. Trans. Comput.，Vol. C-28 (1979)

喜連川：データベースマシン，データベース・システム～最新技術と応用の展望～，ソフト・リサーチ・センタ

データベース白書 '86

データベース白書 '88

— 禁 無 断 転 載 —

平成元年3月発行

発行所　財団法人　日本情報処理開発協会
中央情報教育研究所
〒105 東京都港区浜松町2丁目4番1号
（世界貿易センタービル　7階）
TEL. 03(435)6511　（代表）