

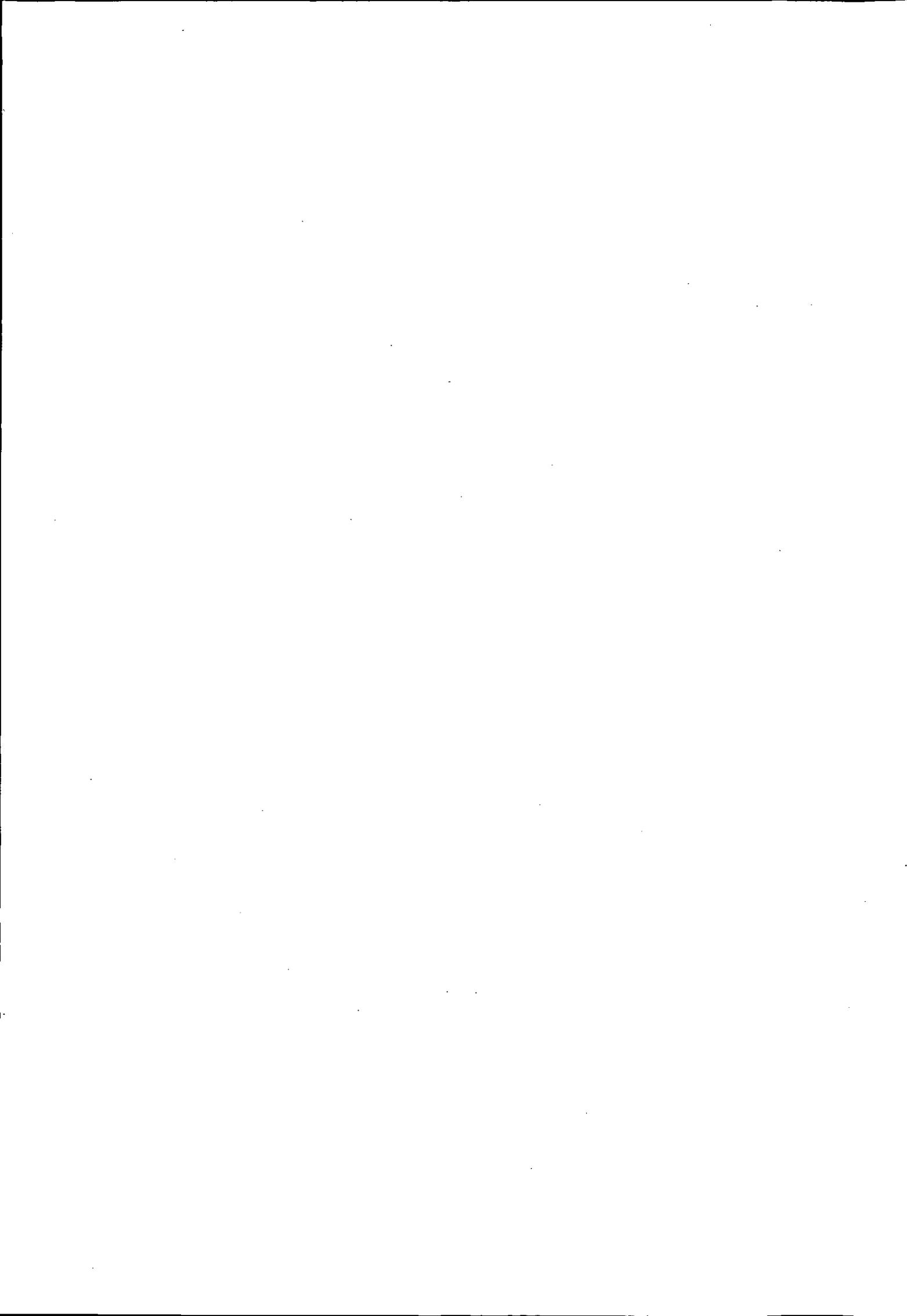
データベース構築促進及び技術開発に関する報告書

(要 旨)

平成 5 年 3 月

財団法人 データベース振興センター

本報告書要旨は、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作成したものである。





序

データベースは、わが国の情報化の進展上、重要な役割を果たすものと期待されている。今後、データベースの普及により、わが国において健全な高度情報化社会の形成が期待される。さらに、海外に対して提供可能なデータベースの整備は、国際的な情報化への貢献及び自由な情報流通の確保の観点からも必要である。しかしながら現在わが国で流通しているデータベースの中でわが国独自のものは 1/3にすぎないのが現状であり、わが国データベースサービスひいてはバランスある情報産業の健全な発展を図るためには、わが国独自のデータベースの構築及びデータベース関連技術の研究開発を強力に促進し、データベースの拡充を図る必要がある。

このような要請に応えるため、(財)データベース振興センターでは日本自転車振興会から機械工業振興資金の交付を受けて、データベースの構築及び技術開発について民間企業、団体等に対して委託事業を実施している。委託事業の内容は、社会的、経済的、国際的に重要で、また地域及び産業の発展の促進に寄与すると考えられているデータベースの構築とデータベース作成の効率化、流通の促進、利用の円滑化・容易化などに関係したソフトウェア技術・ハードウェア技術である。

本事業の推進に当たって、当財団に学識経験者の方々に構成されるデータベース構築・技術開発促進委員会（委員長 山梨学院大学教授 蓼沼良一氏）を設置している。

この「データベース構築促進及び技術開発に関する報告書（要旨）」は平成4年度のデータベースの構築促進及び技術開発促進事業として、当財団が民間企業、団体等に対して委託実施した課題についてそれぞれの成果の概要をとりまとめたものである。この成果が、データベースに興味をお持ちの方々や諸分野の皆様方のお役に立てば幸いである。

なお、平成4年度データベースの構築促進及び技術開発促進事業で実施した課題と委託先は次表のとおりである。

平成5年3月

財団法人 データベース振興センター

平成4年度 データベース構築・技術開発促進委託課題一覧

分野	課題名	委託先
社 会	1 変異タンパク質配列データベースの構築 2 新聞縮刷版見出しデータベースの構築 3 ファジィに関する文献データベースの構築 4 医療用医薬品抗生物質データベースの構築 5 交通事故調査データベースの構築 6 楽器データベースの構築 7 人体計測データベースの構築 8 大学におけるデータベース利用教育システムの プロトタイプ作成 9 先進複合材料データベースの構築 10 博物館所蔵地図資料所在情報データベースの構築調査	日本電子計算機 (株)朝日新聞社 (財)日本情報処理開発協会 (株)小田島 (財)日本自動車研究所 (株)ダイソメディアサービス (社)人間生活工学研究センター 日外アソシエーツ(株) (財)次世代金属・複合材料研究開発 協会 (財)地図情報センター
中小企業振興 地域活性化	11 地域流通最適化データベースのプロトタイプ作成 12 異分野研究のための知的オリエンテーション・ データベースシステムのプロトタイプ作成 13 在宅勤務者サポート・データベースの構築調査	(社)日本ボランティア・チェーン協会 (株)けいはんな (株)志木サテライトオフィス・ ビジネスセンター
海 外	14 銅基複合材料日本特許英文データベースの構築 15 技術協力供与機材データベースのプロトタイプ作成 16 先端産業分野における専門用語の電子辞書デー タベース化の調査研究 17 マーケティングコードの英文データベースの構築	神鋼リサーチ(株) (財)日本国際協力システム 科学技術情報研究所(株) (株)帝国データバンク
技 術	18 安全研究における多重ソース・システム構築の ための基本安全用語データベースの開発 19 3次元マッピングデータベースの技術開発 20 データベース検索サポートシステムの調査研究 21 グループウェアにおけるデータベースシステムに関 する調査研究 22 パーソナルコンピュータとLANの利用による非定形 データベースのプロトタイプ作成 23 知的資源型データベースの調査研究	(株)紀伊國屋書店 (株)日本総合技術研究所 セントラル開発(株)情報図書館 RUKIT (株)イフ・アドバタイジング (株)メイテック (株)ジャパンコミュニケーションズ インスティテュート

目 次

1.	変異タンパク質配列データベースの構築	1
1.1	概要	1
1.1.1	目的	1
1.1.2	実施内容	1
1.2	学術文献の内容調査	2
1.2.1	文献内容の概略	2
1.2.2	整備の方針	3
1.3	データベースの内容	3
1.3.1	エントリー単位の決定	4
1.3.2	項目の決定	4
1.3.3	検索キーワード	5
1.4	データの作成	5
1.4.1	データの入力作業	5
1.4.2	データ作成支援システム	5
1.5	データベースシステムの作成	6
1.5.1	検索・出力機能	6
1.5.2	検索用インデックスファイルの作成	7
1.5.3	データベースの作成	8
1.5.4	ユーザデータベースの作成	8
1.6	データベースの検索, 表示	8
1.6.1	タンパク質名による検索	8
1.6.2	アミノ酸置換パターンによる検索	8
1.6.3	著者名による検索	8
1.6.4	検索結果一覧表示	9
1.6.5	マルチ表示	9
1.6.6	ユーザファイルへの出力, ユーザファイルの表示	9
1.7	展望	9
1.7.1	用語の統一	9
1.7.2	新規データの追加	10

1.7.3	他のデータベースとの整合性	10
1.7.4	表形式以外のデータ	10
2.	新聞縮刷版見出しデータベースの構築	11
2.1	概要	11
2.2	仕様の検討	11
2.3	データ構造の決定	13
2.4	データづくり	14
2.4.1	OCRの導入	15
2.4.2	校正作業	16
2.5	システム構築	17
2.6	検証	17
2.7	今後の課題	20
3.	ファジィに関する文献データベースの構築	21
3.1	目的	21
3.2	概要	21
3.3	情報の収集・整理・選択	21
3.4	抄録作成・キーワード付与	22
3.5	書誌事項補足	22
3.6	文献情報の電子化・登録	22
3.7	運用計画	23
3.8	今後の課題	23
3.9	資料	23
4.	医療用医薬品抗生物質データベースの構築	25
4.1	目的	25
4.2	データベースの概略	25
4.2.1	データベースの構成	25
4.2.2	抗生物質データ品目構成	26
4.3	本体ファイルの内容	27

4.4	その他のファイル・辞書類	27
4.5	表示機能	27
4.5.1	本体ファイルの表示	27
4.5.2	その他の表示	29
4.6	印字機能	30
4.7	サンプルデータ	30
4.7.1	本体ファイルサンプルデータ	30
4.7.2	文献ファイルサンプルデータ	30
4.7.3	引用文献一覧	32
4.8	展望	32
5.	交通事故調査データベースの構築	33
5.1	目的と概要	33
5.2	データベースの改良と発展	33
5.2.1	データベース表の設計とデータの実装	34
5.2.2	データの補完	34
5.3	利用環境の整備（パソコンとワークステーションの連携）	35
5.3.1	ハード、ソフトの選択と試行	36
5.3.2	専用ソフトの開発	37
5.3.3	専用ソフトの効果	40
5.4	今後の課題	40
6.	楽器データベースの構築	43
6.1	目的	43
6.2	実施概要	44
6.2.1	出典別楽器データの入力	44
6.2.2	出典別楽器データの総合化	45
6.2.3	登録用テキストファイルの作成	46
6.2.4	楽器データベースへの登録	48
6.2.5	画像資料の試験登録	49
6.3	データの仕様	49

6.3.1	データの実例及び各部分の名称	49
6.3.2	データ項目	49
6.3.3	他の楽器名称との関係の表示	50
6.3.4	データの表題及び索引の登録	50
6.4	システム構成	51
6.5	データベースの機能	51
6.5.1	使用するデータベースソフト	51
6.5.2	検索機能	52
6.5.3	登録機能	53
6.6	データベースの評価と今後の課題	53
6.6.1	データベースの評価	53
6.6.2	今後の課題	54
7.	人体計測データベースの構築	55
7.1	総論	55
7.1.1	緒言	55
7.1.2	本データベースの構成	55
7.1.3	人体計測項目と体制	56
7.2	パソコンを用いたデータ前処理システム	57
7.2.1	前処理システム基本仕様	57
7.2.2	データ編集システム	58
7.2.3	画像計測システム	58
7.2.4	パソコン前処理システムのまとめ	58
7.3	採寸、統計データベースの構築	58
7.3.1	データベース基本仕様	58
7.3.2	データベースシステム概要	59
7.3.3	採寸データベース	60
7.3.4	統計データベース	61
7.3.5	画像データ仕様	62
7.3.6	データベースの運用	62
7.4	今後の課題(まとめ)	63

7.4.1	今後の課題	63
7.4.2	まとめ	64
8.	大学におけるデータベース利用教育システムのプロトタイプ作成	65
8.1	はじめに	65
8.2	データベース利用教育の事例	65
8.3	データベース利用教育における科目モデルの設定と指導マニュアル案	66
8.3.1	基本的な考え方	66
8.3.2	講義スケジュール	68
8.4	プロトタイプシステム	71
8.4.1	位置づけ	71
8.4.2	機器構成	71
8.4.3	トレーニング対象となるデータベース	71
8.4.4	システムの概要	72
8.5	今後開発する必要があるシステム	74
8.5.1	位置づけ	74
8.5.2	利用環境	74
8.5.3	機器構成	74
8.5.4	必要とされる機能	75
8.6	おわりに	76
9.	先進複合材料データベースの構築	79
9.1	概要	79
9.2	DBシステム構築の背景	80
9.2.1	先進複合材料の背景と展望	80
9.2.2	DBデータ源とデータ利用環境の位置づけ	80
9.3	方針設定	81
9.3.1	基本方針	81
9.3.2	DBの位置づけ	81
9.3.3	収録するデータ	81
9.3.4	マスタファイル	82

9.4	ハードシステムの内容	82
9.5	DB管理システム	83
9.6	RIMCOF DBの特徴	84
9.7	将来のシステム拡充	86
9.7.1	材料データベースに対する新しい要求	86
9.7.2	開発したデータベースの課題と問題解決方途	87
9.7.3	データベースの新しい潮流	87
9.7.4	オブジェクト指向データベース	88
9.7.4.1	歴史的背景	88
9.7.4.2	オブジェクト指向データモデル	89
9.7.4.3	オブジェクト指向とデータベース	89
9.7.4.4	データモデル	90
9.8	最終システム運営像	90
10.	博物館所蔵地図資料所在情報データベースの構築調査	93
10.1	調査研究の概要	93
10.1.1	調査研究の目的	93
10.1.2	実施内容	93
10.2	実施の結果	94
10.2.1	データベースの作成	94
10.2.2	データベースの利用方法	96
10.2.2.1	システムの概要	96
10.2.2.2	システムの操作	97
10.3	今後の課題	98
10.3.1	地図情報と所在機関	98
10.3.2	現在の地図利用形態	98
10.3.3	地図資料データベース化に向けて	98
11.	地域流通最適化データベースのプロトタイプ作成	101
11.1	目的と概要	101
11.1.1	はじめに	101

11.1.2	目的	101
11.1.3	概要	102
11.2	地域流通最適化データベースのプロトタイプ作成	102
11.2.1	本研究の目指すデータベースの将来像	102
11.2.2	「有望地域及び有望都市」の選定に関するシステムの構築	103
11.2.3	「適地候補ゾーン」の抽出プログラムのプロトタイプ開発	103
11.2.4	「適地候補地点」の選定に関するシステムの構築	104
11.2.5	「適地候補地点」の評価プログラムのプロトタイプ開発	105
11.2.6	「適地候補地点」の現地調査・検討	105
11.3	地域流通最適化データベースのプロトタイプに関する各種検討	106
11.3.1	テスト都市を用いたシミュレーション	106
11.3.2	データ処理作業の操作性に関する評価	106
11.4	今後の課題と結論	108
11.4.1	今後の検討課題	108
11.4.2	結論	108
12.	異分野研究のための知的オリエンテーション・データベースシステムの プロトタイプ作成	113
12.1	はじめに	113
12.2	プロトタイプシステムの開発	113
12.3	逐次メニュー方式及び一括対話方式による案内サービス例	119
12.4	今後の課題と方針	121
12.5	おわりに	122
13.	在宅勤務者サポート・データベースの構築調査	125
13.1	概要	125
13.2	在宅勤務者サポート・データベース構築の背景と意味	125
13.3	構築に当たっての問題点の整理と解決への模索	126
13.4	データベースの企画原案の構築とその検証	126
13.5	最終企画案と事業化への方策	127

14.	銅基複合材料日本特許英文データベースの構築	133
14.1	目的	133
14.2	概要	133
14.3	実施内容	133
14.3.1	情報収集	133
14.3.2	データベース構築	137
14.4	本データベース利用の効果	140
14.5	今後の課題と展望	140
14.6	英文日本特許マップの海外へのサービス資料例	140
15.	技術協力供与機材データベースのプロトタイプ作成	143
15.1	目的	143
15.2	実施内容	143
15.3	データベース概要	144
15.3.1	コンセプト	144
15.3.2	機能概要	144
15.3.3	ハードウェア	145
15.3.4	ソフトウェア	146
15.3.5	ファイル構造	146
15.3.6	画面・帳票	146
15.3.7	プロトタイプの評価	148
15.4	今後予想される効果	150
16.	先端産業分野における専門用語の電子辞書データベース化の調査研究	153
16.1	概要	153
16.1.1	目的	154
16.1.2	実施内容	155
16.2	調査研究結果の概要	157
16.2.1	日英両用電子化辞書の作成とその諸特性	157
16.2.2	和訳語・英訳語における専門分野別訳し分けの状況調査	158
16.2.3	訳語検索の連鎖による類義語クラスターの析出と辞書のシソーラス化	158

16.2.4	和文専門用語の抽出方式	159
16.2.5	和文専門用語の出現頻度分析	159
16.2.6	専門分野分類の検討	160
16.2.7	専門用語辞書の更新管理システム化	160
16.2.8	専門用語における和訳語、英訳語の特質と問題点の総合的検討	160
16.3	今後の課題	161
17.	マーケティングコードの英文データベースの構築	163
17.1	日本産業分類と帝国産業分類	163
17.2	T I Cの特色及び体系	163
17.2.1	T I Cの特色	163
17.2.2	T I Cの体系	163
17.3	帝国データバンクマーケティングコード作成の意義	164
17.4	T M Cの構成	165
17.5	英文化の背景	165
17.6	英文化の目的	166
17.7	S I C対応の意義	167
18.	安全研究における多重ソース・システム構築のための 基本安全用語データベースの開発	169
18.1	背景と目的	169
18.2	実施体制及び実施内容	170
18.2.1	実施体制	170
18.2.2	実施経過	170
18.2.3	実施内容と成果	170
18.3	安全研究における基本安全用語データベースの収集	171
18.3.1	安全研究に関する重要語の収集・評価の事例研究 (Lovastatine に関する文献事例による医薬品安全用語の収集と評価)	171
18.3.2	基本安全用語の収集(安全に関する汎用語の収集と評価)	173
18.3.3	安全関連文献の包括的調査 (安全研究に関する用語の複数データベース間での出現頻度分析)	174

18.3.4	主要データベースにおける安全用語の収集 (ENERGY SCI. & TECH., EMBASE, CASearchにおける安全用語の収集)	176
18.4	今後の展望	177
19.	3次元マッピングデータベースの技術開発	179
19.1	概要	179
19.1.1	システムの目的	179
19.1.2	実施内容	179
19.2	検討結果	180
19.2.1	システムコンセプトの検討	180
19.2.2	利用局面の確認	180
19.2.3	システム特性の設定	181
19.2.4	3次元データの所在	181
19.2.5	データソースの考え方	181
19.3	システム主要機能の設計検討	182
19.3.1	システム構成	182
19.3.2	主要機能の基本設計	183
19.4	システム構築の考え方	184
19.5	今後の課題	184
20.	データベース検索サポートシステムの調査研究	187
20.1	概要	187
20.2	実施内容	187
20.2.1	方針と内容	187
20.2.2	データベース利用者の利用目的・動向調査	187
20.2.3	調査結果	188
20.2.4	検索サポートシステムの検討	190
20.2.5	システム構成	190
20.2.6	プロトタイプ	192
20.3	今後の課題	199

21.	グループウェアにおけるデータベースシステムに関する調査研究	201
21.1	概要	201
21.2	グループウェアにおけるデータベースシステムの構築	202
21.2.1	グループウェアの技術的動向	202
21.2.2	グループウェアにおけるデータベースシステム構築技術	203
21.3	グループウェア発展に伴う情報システムの高度化並び にデータベース構築技術の進展とその課題	205
22.	パーソナルコンピュータとLANの利用に よる非定形データベースのプロトタイプ作成	207
22.1	概要	207
22.1.1	目的	207
22.1.2	基本方針	207
22.2	システム構成	208
22.2.1	ハードウェアシステム	208
22.2.2	ソフトウェアシステム	209
22.3	運用概念	209
22.3.1	情報の登録	210
22.3.2	情報の検索	211
22.4	機能概要	212
22.4.1	データベース管理サブシステム	213
22.4.2	データベース入力サブシステム	213
22.4.3	データベース検索サブシステム	214
22.5	プロトタイプの評価	214
22.5.1	操作性に関して	214
22.5.2	応答性に関して	215
22.5.3	拡張性に関して	215
22.6	今後の課題	215
23.	知的資源型データベースの調査研究	217
23.1	調査研究の目的	217

23.2	調査研究体制	217
23.3	検討項目	218
23.4	検討内容	218
23.5	主要キーワード	218
23.6	成果	225
23.7	今後の課題	226

1 変異タンパク質配列データベースの構築

日本電子計算株式会社



1. 変異タンパク質配列データベースの構築

1.1 概要

1.1.1 目的

タンパク質は、生命活動の基盤となる作用物質である。この生体高分子はわずか20種類のアミノ酸を素材とし、ペプチド結合という単一の共有結合によるアミノ酸相互の結合によりできている。タンパク質のアミノ酸配列の決定は遺伝子工学的手法で容易になり、天然のタンパク質のアミノ酸配列データは、約1年半で2倍のペースで増加しており、既に約40,000件がデータベース化されている。近年、遺伝子工学の進歩により天然のタンパク質のアミノ酸配列の一部を変え、産業に有用な人工タンパク質の研究開発が進んでおり、21世紀の主要技術の一つと目されるバイオテクノロジーの着実な果実となることが予測される。

研究者がアミノ酸配列を改変して物性や機能の改変を行おうとするとき、現在、試行錯誤的に実験を行うか、過去の文献を調査し推定する等の方法で行われている。これには研究者の膨大な時間と手間が必要とされ、有用なデータベースの出現が求められている。

そこで今回、タンパク質改変実験の過去の報告を集積し知識体系化することを目的に、変異タンパク質配列データベースの構築を実施した。本データベースは、バイオテクノロジー分野の発展に大きく貢献できるとともに、世界的学術的貢献の一助として位置づけられると期待される。

1.1.2 実施内容

(1) 学術文献の内容調査

変異タンパク質に関する学術文献について記述されている内容を調査した。

(2) データベースの内容の検討

学術文献の内容調査をふまえ、データベースとして必要な情報（項目）の検討を行った。

(3) データ作成方法の検討

データをデータベースに効率的に取り組むために、データのチェック方法などを検討した。

(4) データの作成

データの作成方法に基づきデータを約10,000件作成した。

(5) データベースシステムの作成

検索・出力機能のほか、ユーザが作成したデータの取り込み、キーワードの作成などデ

データベースを作る一連のシステムを作成した。

(6) データベースの検索

データベースシステムが仕様どおり動くかどうかを調べるために、実際の検索で行われる例を想定し、検索・評価した。

1.2 学術文献の内容調査

1.2.1 文献内容の概略

(1) 変異タンパク質

変異タンパク質は、蛋白工学的手法で作成された人工変異タンパク質と、遺伝的疾患や自然変異細菌などから古くから見い出されている自然変異タンパク質に大別できるが、人工変異タンパク質を研究している論文数の方が大勢を占めている。

(2) 変異タンパク質名

個々の変異タンパク質を識別する名称は、置換前の天然タンパク質のアミノ酸残基と置換後のアミノ酸残基名を使用している場合が多く見られるが、必ずしも統一されていない。

(3) アミノ酸残基番号

アミノ酸残基番号は、アミノ末端から振られている場合ばかりではなく、類似の他の生物由来のタンパク質のアミノ酸残基番号を参照している場合もある。その上、アミノ酸配列の長さが変化した場合、アミノ酸残基の番号づけが統一されていない。

(4) 変異タンパク質作成方法

蛋白工学的手法の中には遺伝子工学的手法（部位特異的変異、発現ベクタなど）、物理的手法（紫外線、放射能など）、科学的手法（変異試薬など）など、さまざまな手法が用いられている。

(5) 活性、性質測定条件

生物的活性や物理化学的性質を測定する際の溶媒の温度、塩濃度、pHなど実験条件が完全に一致している場合はほとんどない。

(6) 変異タンパク質作成数

測定には天然タンパク質と変異タンパク質が用意される。変異タンパク質数は通常数種類用意され、文献によっては10種類以上の場合もある。

(7) 測定結果の表示方法

天然タンパク質や他の変異タンパク質との特性の違いを比較するために、また、基質、塩濃度、経過時間などに対する変化を見るために表形式で表示される場合が多い。このと

き使用される数値は測定値そのものである場合以外に、天然タンパク質の値を1とした相対値の場合も見られる。

1.2.2 整備の方針

(1) 記述方法の統一

整備の第一歩として、データベース上での記述方法の統一を図る。この際使用する基準は、文献の中で比較的多く利用されているものに準じる。また天然のアミノ酸配列データベースとして、PIR-International(Protein Information Resource-National Biomedical Research Foundation (PIR-NBRF)), Japan International Protein Information Database (JIPID). Martinsried Institute for Protein Sequence (MIPS)が作成し、世界的によく利用されているデータベースが存在するので、このデータベースでの記述方法も参考にする。

(2) 測定結果

文献では多くの場合測定結果が表形式で記載されている。そこで本データベースでも表形式で測定データを表記する。文献にはグラフや図も掲載されているが、入力の手間がかかること、データ容量が大きいこと、システム開発に時間がかかることから、今回は、グラフ、図の収録は見合わせる。

(3) 他のデータベースとのクロスリンク

変異タンパク質に関する論文の多くは、天然タンパク質と変異タンパク質との諸特性の違いについて論じている。したがって本データベースでもこの点に重きを置くことにし、天然タンパク質に関する内容については、クロスリファレンスなどを記述し、他のデータベースを参照できるようにする。またクロスリンクする機能をもたせることにより、既存の天然タンパク質に関するデータベースと一緒に利用した場合には、直接該当する内容を引用して表示できるようにする。

1.3 データベースの内容

前章の文献調査をふまえて、エントリーの単位、作成するデータベースの項目、及び検索するためのキーワードを検討する。

1.3.1 エントリー単位の決定

エントリーの単位としては、

- (1) 1天然タンパク質を1エントリー（同じ天然タンパク質から派生した変異タンパク質をすべて一つのエントリーにまとめる）
- (2) 1文献を1エントリー
- (3) 1変異タンパク質を1エントリー

という3種類が考えられる。

(1)の場合、ある天然タンパク質についてのすべての変異タンパク質の種類を知るのにはとても都合がよいが、測定された諸特性は論文ごとに異なり、同じ特性を測定している場合でもその測定条件が一致していることはほとんどない。そのため測定結果を一つにまとめるのはあまり意味をなさず、逆に内容が多くなりすぎ見づらくなると予想される。

(2)の場合は、文献に記載されている範囲で変異タンパク質の諸特性を比較するには都合がよい。また入力する場合でも、常に新規に登録することになるので、データチェックを行いやすい。しかし一つの変異タンパク質の性質のみを表示させたり、ある特定のアミノ酸の置換を起こしたタンパク質をリストアップする等、個別に変異タンパク質を検索したり表示するには適していない。

(3)は個々の変異タンパク質の諸性質を表示したり、いろいろな検索方法を行う場合には最も適している。しかし、データをこの単位で入力するのでは重複する内容がかなり多くなる。またこの単位での表示は他の変異タンパク質の諸特性との比較には不適當である。

以上の3種類のエントリー方法の長所、短所を検討した結果、検索方法の最も簡単な(3)を採用することに決定した。しかしながら、データ入力の段階では(2)を採用して入力にかかる手間を減らし、最終的なデータエントリーにする際にプログラムで(3)に変換することにした。またエントリーの内容を表示する方法として、各エントリーごとに表示する方法以外に、複数のエントリーの内容をまとめて、変異タンパク質同士の特性を比較できるような表示方法も採用した。

1.3.2 項目の決定

エントリーの内容を見ただけで変異タンパク質の諸性質の測定を追試できることが望ましい。そのために各種実験条件について最低限必要であると思われる項目を設定する。更に詳細な内容を知ることができるように、参照している文献についての情報も含める。データは国際的な学術論文から主に収集しており、各項目の内容として用いられている用語も英語のままであることが多い。そのため項目、内容とも英語を使用する。

1.3.3 検索キーワード

(1) キーワードの取り出し方

入力内容の最終チェックを行う人間を1名に設定し、その人間の基準で用語の統一を図り、取り出しは機械処理だけで行うことにする。

(2) キーワードの決定

検索に慣れていない人でも簡単に検索できるように、その項目に現われる単語を、3文字単位で1文字分ずらしながら切り出したものを検索キーワードとして登録する。

1.4 データの作成

学術文献から人力で一つ一つデータを作成する。一つの文献で複数の変異タンパク質に対して測定が行われている場合、文献等に関する情報は共通なので、入力の手間を省きタイプミスの可能性を減らすために、1文献を1エントリーとして通常のエディタを使用して入力し、データのチェック後、プログラムを用いて変異タンパク質ごとに分割する。このようにして作成されたデータの件数は、約10,000件である。

1.4.1 データの入力作業

(a) 学術文献のコピーの取り寄せ

予め各学術雑誌ごとに担当者を定めておき、変異タンパク質に関する論文を収集する。

(b) アクセッション番号の決定

エントリーごとにユニークなアクセッション番号を決定する。

(c) TITLE, REFERENCEの入力

文献の内容をあまり詳しく読まないでも入力できるTITLE及びREFERENCEについてすぐに入力を行う。

(d) TITLE, REFERENCE以外への入力

専門の知識を有する人が論文を読み、その内容に従って、TITLE, REFERENCE以外の内容を入力する。

(e) 入力内容のチェック

データを入力していない専門知識を有する人が、プリントアウトされた入力内容と文献の内容との比較、チェックを行う。もし入力ミス等がある場合には(d)に戻す。

1.4.2 データ作成支援システム

人間が通常のエディタを使用して入力する方法では、タイプミスが避けられない。そこで、タイプされたデータを文献別に入力されている段階でワークステーションでチェックする。データは通常のテキストファイルなので、エラーを発見したらその場で修正を行う。これをエラーがなくなるまで繰り返す。文献単位のエントリーのチェックが終了したら、それを分割プログラムを利用して1変異タンパク質ごとのエントリーに分割する。

1.5 データベースシステムの作成

1.5.1 検索・出力機能

TITLE による検索の場合、変異タンパク質の表し方が著者ごとに異なっている場合が多く、原論文の変異タンパク質名とTITLE 中の変異タンパク質名が異なり、いきなり特定の変異タンパク質を指定するのは困難である。そこで、その変異タンパク質のもとになったタンパク質名やタンパク質を産する生物種名での検索ができるようにする。

更に、似たような働きをするタンパク質について同時に検索する場合もあるので、タンパク質名の一部（3文字以上）の一致でも検索できるように中間一致検索を採用する。

変異タンパク質の検索としては、アミノ酸の置換パターンによる検索も考えられる。これはアミノ酸個々の性質に重きを置いた検索方法であり、タンパク質の種類によらないアミノ酸の置換とその影響の一般則を探る場合でもある。そのため、あるアミノ酸から別の特定のアミノ酸への置換パターンでの検索を可能にしている。

TITLE 部分以外の項目のうち、著者名、雑誌名、METHOD, EXPRESSION, PURIFICATION, FUNCTION, KEYWORD, FEATUREの各項目内での検索もできる機能を組み入れる。これらの項目では統制のない自由な語が使われているので、最低3文字一致した項目を含むエントリーを検索できるようにする。

検索の結果が妥当なものであるかがすぐに確認できるように、該当するエントリーのアクセス番号とともにTITLE を表示する機能を組み込む。また検索結果はカレントリストに登録させる。

出力は通常データベース内の全項目を出力する人が多いと思われる。しかし、変異タンパク質について記載されている原論文に関する情報のみが必要であるなど、検索した内容の一部にしか興味がない場合もあるはずである。そのため、出力は基本的には全項目が出力されるようにしておき、オプション機能として指定した項目のみを表示できる機能を組み入れる。

変異タンパク質の生物活性・物理化学的性質は、天然タンパク質ばかりではなく他の変異

タンパク質と比較して表示したほうが便利な場合もある。また多くの文献では、一つのタンパク質に対して複数の変異タンパク質を作成してその諸性質を比較している。そのため、個々のエントリーの内容を順番に表示する以外に、各項目ごとにまとめて表示できる機能も組み込む。

1.5.2 検索用インデックスファイルの作成

データベースの対象となるデータの主要項目について、各項目別に検索用インデックスファイルを作成する。

(1) 基本インデックス (variant.inx)の作成

各エントリーのデータはすべて結合されてvariant.ref という一つのファイルにまとめられている。

このファイルは個々のエントリーファイルを単に順番に結合した形となっており、一つのエントリーの内容がすべて連続して収納されている。そこで基本インデックスでは各エントリーのアクセッション番号とそのエントリーが始まるアドレスとを記録しておく。これにより、アクセッション番号を指定することにより、variant.ref ファイルを頭から検索せずに、その内容にダイレクトにアクセスでき、内容の表示に要する時間を短縮している。

(2) 各種検索用インデックスファイルの作成

アクセッション番号が分かれば、(1)で示したようにその内容を短時間で表示できる。このアクセッション番号を検索するために、エントリーの内容の主要な項目ごとにインデックスファイルを作成する。

各項目に統制のない自由な言葉が使われており、またタンパク質名や実験方法など特殊な言葉も使われている。そのため検索方法としては最も単純な中間一致検索方法を採用した。一方では、タンパク質名や実験方法などほとんど同じフレーズが繰り返し現われる項目もある。そこでインデックスファイルを作成するに当たり、各項目を一行単位に分解し、その文を昇順に並びかえ、その行を含むすべてのエントリーの番号を登録した。

(3) クロス検索用インデックスファイルの作成

この変異タンパク質データベースを単独で用いても変異とそれが諸性質に及ぼす影響を知ることができる。しかしながら、既存の各種タンパク質に対するデータベース、特にPIR-Internationalのアミノ酸配列データベースや生物活性データベースとともに利用すると、天然タンパク質に関するさまざまな情報を捕うことができる。この変異タンパク質データベースの各エントリーのフォーマットは、これらのデータベースで採用しているフ

フォーマットに沿っており、(2)の各種インデックス作成用プログラムで同じようにインデックスファイルを作成できる。このため、容易にクロス検索用アクセスンファイルを作成でき、同時に検索できるばかりではなく、この変異タンパク質データベースからそれらのデータベースの対応する内容を引用することも可能である。またこのクロス検索用インデックスでは、(2)で作成した項目の内容すべてに対して3文字ごとに分解してその文字列のアドレスを登録している。これにより、より高速な検索が可能となっている。

1.5.3 データベースの作成

文献単位で入力を行った後、個別エントリーファイルを作成し、各種インデックスを作成し、データベースの作成を行った。

1.5.4 ユーザデータベースの作成

変異タンパク質データベースの利用者の中には、ある特定のタンパク質にのみ興味をもったり、自分が測定した実験結果のデータベース化など独自のデータベース作成を希望している人が予想される。そこでユーザが独自でデータベースを作成する手順を示した。

1.6 データベースの検索、表示

作成したデータベースが仕様どおり機能するかを調べるために、実際に検索されると思われる例題を幾つか想定し検索した。

1.6.1 タンパク質名による検索

タンパク質名による検索及びカレントリスト内での検索を行い、漏れなく検索できることを確認した。

1.6.2 アミノ酸置換パターンによる検索

ある任意のアミノ酸が別のアミノ酸に置き替わっている変異タンパク質の検索を行い、漏れなく検索できることを確認した。

1.6.3 著者名による検索

著者名による検索を行い、漏れなく検索できることを確認した。

1.6.4 検索結果一覧表示

検索により作成されたカレントリストの表示, 及び検索直前のカレントリストの回復を確認した。

1.6.5 マルチ表示

由来した天然タンパク質が同じ変異タンパク質について, 複数のエントリー内容をまとめて表示できることを確認した。

1.6.6 ユーザファイルへの出力, ユーザファイルの表示

データベースの一つのエントリーの内容をユーザファイルとして保存し, その内容を再び検索システムで表示できることを確認した。

1.7 展 望

1.7.1 用語の統一

今回の方法では, タンパク質名, 生物種名, 変異の記載方法, 著者名, 雑誌名については基準に従って統一が取れているが, SOURCE, METHOD, EXPRESSION, PURIFICATION及びFUNCTIONに関しては, 文献の記述に従っており, 用語の統一が完全に取れているとはいえない。またKEYWORD に関しては文献を詳細に読んだ人間の主観によるところが大きく, ほとんど統一されていない。

KEYWORD についてはこれといった基準はないので, 用語を統一するためには, 以前のエントリーで用いられた用語をリスト化しておき, KEYWORD を決定する際に常にそのリストを参照する方法が最も現実的である。この方法を利用するためには, 常にKEYWORD リストを更新しておく必要があり, KEYWORD を自動的に切り出し, リスト化するシステムを作成する必要がある。

KEYWORD に関しては用語の統一以外に, その仕様の統一も図る必要がある。由来した天然タンパク質が同じ変異タンパク質群について, その変異タンパク質群の共通した性質を示すKEYWORD を一つのエントリーに使用した場合, 他のエントリーに対しても使用しなければならない。そのため, その用語をどのエントリーに対して使用したかの対応表を作成しておく必要がある。これは結局KEYWORD インデックスそのものである。

必要なKEYWORD が漏れなくエントリーに記載されているかのチェックは, そのKEYWORD を含むエントリーのリストと, そのKEYWORD を含んでいておかしくないエントリーのリストと

の比較で行える。後者のリストは、例えばタンパク質名などでの検索から作成する。もしこの二つのリストが一致していれば、必要なKEYWORD が漏れなくエントリーに記載されていることになる。もしKEYWORD インデックスのリストが少ない場合には、幾つかのエントリーでのKEYWORD の記入漏れが考えられる。逆に多い場合には、KEYWORD の記載が適切かどうかの判断をし、不適切な場合は削除する。このような修正が簡単なシステムに変更し、更に毎日自動的に、各種インデックスの更新を行うようなメカニズムを付け加えれば、用語の統一が図れると予想される。

1.7.2 新規データの追加

変異タンパク質は今後もますます作成され、それらのデータも蓄積されていく。このデータベースは学術文献を分析する必要がある、またもととなる文献（あるいはそのコピー）が簡単に入手できる環境が必要である。今後このデータベースをアップデートしていくためには、大学等の公的機関に委託するなどの方法を取る必要がある。

1.7.3 他のデータベースとの整合性

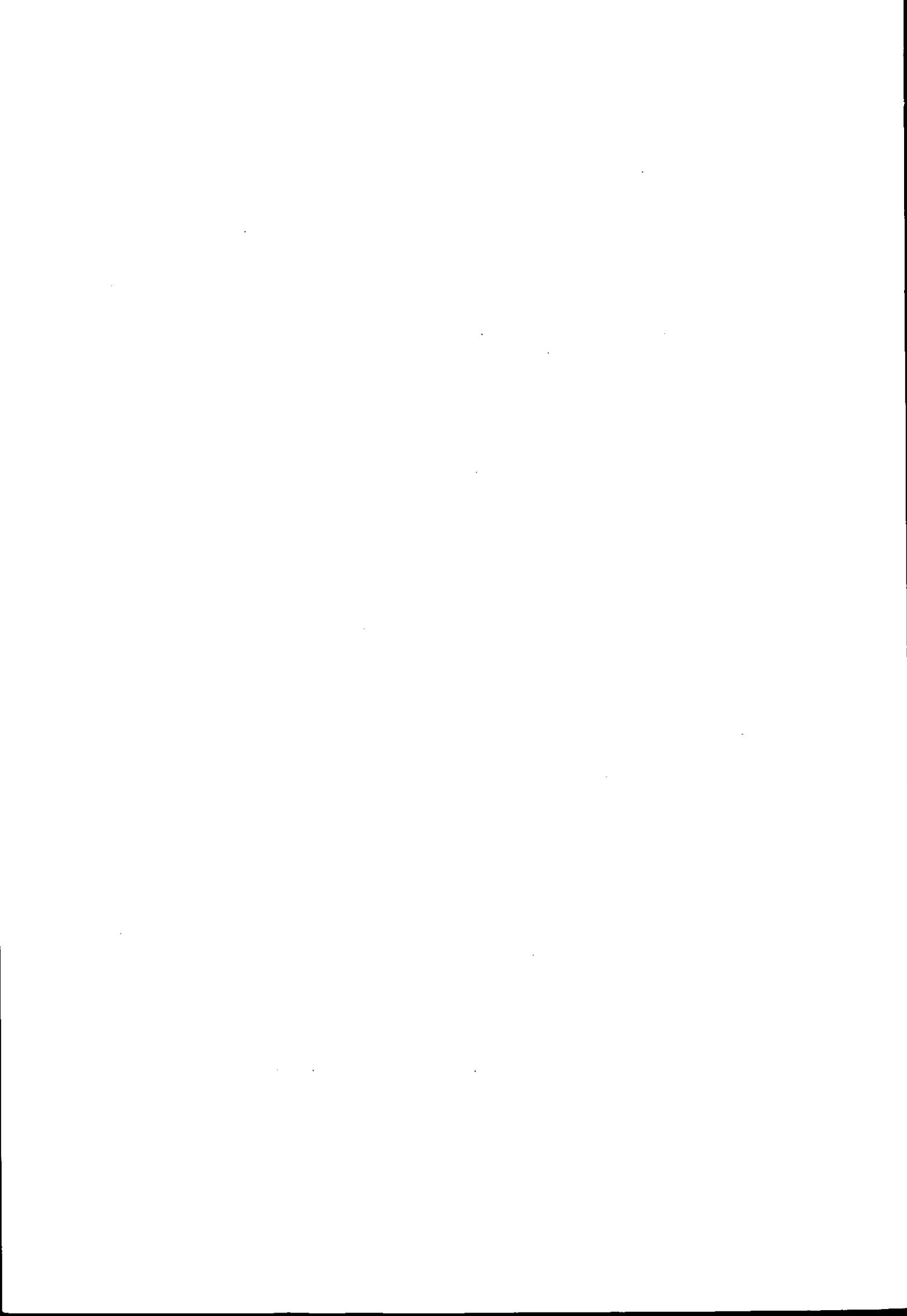
今回作成した変異タンパク質データベースでは、タンパク質名、生物種名、著者名、雑誌名の記載方法はPIR-International が採用している方法に準じた。これにより、この変異タンパク質データベースから、アミノ酸配列データベースや生物活性データベースの内容を引用することが可能となった。逆にこれら二つのデータベースが変異タンパク質の内容を参照するためには、変異タンパク質独自の項目に対応し、クロスリファレンスとして変異タンパク質のアクセッション番号を登録することが必要となる。これには、PIR-International との協力が欠かせないが、もしこれが実現すれば、変異タンパク質データベースの利用範囲が広がり、世界的学術的貢献度も増すことになる。

1.7.4 表形式以外のデータ

この変異データベースでは、生物活性や物理化学的特性の変化をすべて表形式で記載してある。しかし文献の中には特性の変化をグラフで示しているものも数多く見られる。このような場合、ある代表的な数点の情報しかデータベースに取り込んでいないことになる。ある簡単な式の組み合わせでそのグラフを表現できる場合はほとんどないので、このようなデータは、画像データとして保存するのが最善であると思われる。今後ワークステーションでの画像を取り込んだデータベースに対する技術の発展が予想され、近い将来グラフデータ等を取り込むことが容易になると期待する。

2 新聞縮刷版見出しデータベースの構築

株式会社 朝日新聞社



2 新聞縮刷版見出しデータベースの構築

2.1 概 要

朝日新聞社では1984年8月からニューメディア本部の担当で新聞記事データベースの構築を開始、1986年4月から平和情報センター（現在のジー・サーチ）と日本アイ・ビー・エムを皮切りに記事データベースの一般提供を始めた。

しかし、それ以前の記事を探すには、社内利用においても、調査部の切り抜き、あるいは各月に発行される縮刷版を操って見る等しか検索の方法はなかった。

ユーザが記事データベースの利用に慣れてくるにつれ、昔のデータを何とか同じような方法で検索できるようにならないかという要望が内外から寄せられるようになった。

そこで着目したのが当社が毎月発行している縮刷版であった。

朝日新聞記事縮刷版は、大正8年（1919年）創刊の歴史をもち、以来1年も欠かすことなく、毎月発行されてきている。しかも、巻頭には記事の見出しをそのまま項目ごとに整理した詳細な記事索引がついている。

この記事見出し索引をデータベース化し、見出し文のほか、発行日、朝夕刊の区別、面名、縮刷版における掲載ページなど、記事の所在情報を示すものをつくれば、資料探しの手間は飛躍的に改善される。長期間を通覧できるものができたら、年表、事件史といったこれまでにない新しい用途も開けるのではないか、こういった構想で構築に着手したのが縮刷版見出しデータベースである。

2.2 仕様の検討

朝日新聞縮刷版見出しは図2-1のような仕組みになっている。

まず、記事の種類によって「国土・人口」、「政治」「経済」などカットつきの大項目の見出しがついている。その下に「開発・保全」「気象」、政治関係でいえば「政党」「国会」などの中项目的見出し、更にその下に「土地対策」「環境保全」といった小项目的見出し、更にその下に「国土利用計画法」といった具体的各論の見出し、その下に各見出しが並ぶという構造になっている。各見出しの下の和数字がその号に掲載されているページ数を示し、洋数字が何段目にあるかを示す。

索引目次

週刊報告・座標……………3

国土・人口……………3

 国土・資源・人口

 〈開発・保全〉……………3

 国土開発・都市計画

 土地対策・環境保全

 〈気象〉……………3

 天気・台風・地震・

 火山・天文

政治……………4

 憲法・人権問題・政

 治思想団体・平和運

 動

 〈政党〉……………4

 政治資金・政党別動

 向

 〈国会〉……………5

 衆議院・参議院

 〈法令・法案〉……………5

 〈選挙〉……………5

 選挙制度・国会議員

 選挙・地方選挙・選

 挙違反

 〈内閣〉……………6

 首相・閣僚・閣議・

 各省庁・人事・白書

 〈財政〉……………7

 予算・税・財政投融

 資・地方財政

 〈国防〉……………7

 日米安保・防衛計画

 ・自衛隊・在日米軍

外交……………7

 領土領海・条約協定

 ・対外援助・各別別

 〈国交〉……………9

 承認・断交・大公使

 ・要人の来日・出入

 国査証・在日外国人

 ・国際犯罪

経済……………9

 経済政策・景気動向

 ・国際収支・対外経

 済協力・公取委・財

 界・経済団体・社会

 企業・決算・倒産

 〈金融〉……………10

 金融政策・金利・銀

 行・外国為替・保険

 ・証券

 〈産業〉……………11

 産業政策・各種産業

 (エネルギー・鉱工

 業・農林水産)

 〈流通・商業〉……………13

 卸小売業・百貨店ス

 ーパー・総合商社・

 倉庫業

 〈物価〉……………13

 物価対策・価格動向

 ・物価指数・料金

 〈貿易〉……………13

 輸出入動向・輸出・

 輸入・見本市・関税

運輸・通信……………14

 交通・道路・鉄道・

 航空・海運・通信・

 郵便

労働……………14

 春闘・雇川・失業・賃

 金・労働組合・争議

世界……………15

 〈国 際〉……………15

 国際会議・国交・国

 連・軍備軍縮・国際

 経済・金融・石油

 〈ア ジ ア〉……………17

 別別(各州とも同じ)

 〈中 東〉……………20

 〈ア フ リ カ〉……………20

 〈ヨ ー ロ ッ プ〉……………21

 〈ア メ リ カ〉……………22

 〈オセアニア〉……………24

文化……………24

 文化交流・学界・著

 作権・国語・文化財

 記念物

 〈科 学〉……………24

 みんなの科学・くら

 しの科学・計測・発

 明発見・観測調査・原

 子力・動植物・医学

 〈教 育〉……………25

 教科書・学校・教職

 員・学生・宗教

 〈学 芸〉……………26

 寄稿・文学・美術・

 音楽・映画・芸能

 ・演劇・舞踊・能・狂言

 〈出版・報道〉……………28

 新聞・山版・読書特

 集・放送・広告

社会……………28

 皇室関係

 〈一 般〉……………28

 風俗・習意善行・祝

 日・年中行事・戦争

 関係・宝くじ

 〈人 事〉……………29

 叙位叙勲・結婚離婚

 ・死亡

 〈公 害〉……………30

 公害対策・地盤沈下

 ・大気汚染・水質汚

 染・農業汚染・公害

 病

 〈事件・事故〉……………30

 中毒・自殺・爆発・

 遭難・交通事故・航

 空事故・火事・災害

 〈犯 罪〉……………32

 防犯・過激派事件・

 誘かい・密輸・詐欺

 横領・汚職・偽造模

 造・傷害・暴力おど

 し・殺人・強盗・窃

 盗

 〈司法・警察〉……………33

 法曹界・裁判所・行

 刑・検察・警察

 〈裁 判〉……………33

 訴訟・公判・判決

国民生活……………34

 家計・消費者・消費

 者運動

 〈衣・食・住〉……………34

 〈保健・衛生〉……………34

 病氣・医療・医薬品

 ・衛生

 〈福祉・厚生〉……………34

 社会保障(健保・年金)

 ・老人問題・身障者

 〈婦人・青少年〉……………35

 〈観光・娯楽〉……………35

 レジャー・余暇・施設

 ・趣味・基特棋・競馬

家庭面……………35

 ひととき・服飾・美

 容・料理・住まい・

 園芸

凡 例

◎各見出し語の下の和数字はページ数、洋数字は見出し語のある段数を示す。

◎見出し語の分類は一語一カ所方式をとっているもので、引いた項目にない場合は、関連のある他の項目を参照されたい。

◎欠印(⇒)は参照項目を示す。

◎2日付朝刊は休み

特集面……………36

 解説面・月間ルポ・

 にゆうさうんじ

 〈日 曜 版〉……………37

 日本史の舞台・ぶっ

 くす・紙上創作展・

 暮らしと健康・遊び

スポーツ……………38

 〈外 国〉……………38

 種目別

 〈国 内〉……………38

 種目別

地 方……………41

 地方自治・首都圏

 〈東 京 都〉……………41

 一般・都政・都議会

 ・23特別区・市部

 ・都部・島部

写 真……………42

 読者の新聞写真コン

 クール・写真特集

本 社……………42

 社説・今日の問題・

 天声人語・声・論壇

 ・世論調査・社告・事業

週刊報告

12月28日〜1月3日：一七八〇

1月4日〜10日：三六五八〇

1月11日〜17日：六二五八〇

1月18日〜24日：八九五八

座 標

生活の豊かさ 都府重人：三二九

1◇質さんの家庭論 宮下展夫：六九七

1◇過熱する生命技術 柴田鉄治：八五九

国民の意識(本社世論調査)

生活苦甚まる声 自衛隊で改選賛否ほぼ同率：六五七

本社調査に見る国民意識 不満多いが安定を評価：七四一

暮らしと政治 憲法と自衛隊 老後の暮らし・エネルギー源：七四一

仲よくする回：七四九

けなげな女性管理想像 好きな女性性 女性の進出 生まれかわるなら：七五一

どちらの名を・どこまで信用：七五五

日本人らしい平衡感覚 建前と本音をたくみに区別：七五六

調査方法：七五九

国土・人口

落ち込む出生率 昨年は史上最低 丙午の41年下回る：三一

全国で87万人 年男・年女：三九

人口動態

社会増目立つ中小都市 人口動態 移動率は7%弱る：一八七八

世帯構成(55年厚生行政基礎調査)

家庭の老齢化・ミニ化加速 高齢者世帯、増万超す この10年で倍増：五九一

〔世〕60歳以上の者のいゝ世帯の構成割合の推移：五九二

〔世〕人員別にみた世帯

数と平均世帯人員の動き

き……………五九一

4

開発・保全

〔国土開発〕

定住構想

高齢者を念頭に整備 「地域づくり」で定住研が報告書：八九二

〔都市計画〕

都市計画道路変更案を了承 都計審……………一〇一八

〔土地対策〕

首都圏の土地利用状況 10都府平方キロ航空写真に建設者が新年度着手：六七八一

宅地化へ基礎資料：六七八一

〔環境保全〕

環境アセスメント法は法令・法案

清流復活北水(紹介) 環境庁 区民の知恵集 江戸川 戸川区の親水公園：四六一

環境保全「ナショナルトラスト」導入……………七三三

土地買取りの住民に税金優遇 環境庁方針：七三三

あくまで住民主体……………七三三

湖沼保全で審判

水辺に憩いの場を 湖沼保全 中公審が答申 市街化に歯止め……………九六七

〔解説〕西濃 深刻な状況 省庁の環境法整備を九六九

湖沼保全の答申内容 要旨……………九六九

九六九

気 象

〇「災害」

大寒気団ひかえ81年始動 一三七五

北陸地方にまた寒波……………一七〇六

今冬は「第38寒害」 気象庁 なが大雪の恐れ 寒気がまた接近……………二〇九八

二〇九八

大賀 平野部は一服 内陸……………二二九

二二九

図2-1 縮刷版記事索引のサンプル

縮刷版見出しデータベース構築に当たって、まず、直面した問題は、限られたデータの中からいかに過不足なく有効な情報をひき出せるようにするかであった。

予想される問題点を洗い出してみた結果、「大」「中」「小」とついている各項目見出しが、シソーラスのように上下関係をもった統制語として使えないか検討してみた。しかし、この大項目や中小項目の見出しは時代によってかなり変動があることが分かった。時代とともに一つのテーマが項目間を移動しており、一貫したものになっていなかったのである。

結局、各論見出しや大中小各項目の見出しを上下関係無しに並立するキーワードとして加えてやれば、それなりの絞り込みの役には立つのではないかという結論に達した。

2.3 データ構造の決定

データ構造と表示の形式は図2-2のようにした。

見出し表示は2行で構成され、1行目は見出し本文、2行目が発行日、朝夕刊の区別、掲載面、そして縮刷版の掲載ページ、掲載段という所在情報を示す。

現行の朝日新聞記事データベースとの整合性を図るため、見出し表示は1行36字以内とした。そのため、長行の見出しによってはおさまりにきれいなものが出てくるので、それは本文表示(詳細表示)という形式で救済を図ることにした。

逆に訃報のように「何野何兵衛氏」としか表示がないものは、その上位項目の見出しをつけ加えて「何野何兵衛氏_訃報」という注釈めいた表現をとれるようにした。

データはMS-DOS形式のファイルとして起した。

カットつきの大項目は { } , 波形野線つきの中項目は < > , 【 】 で囲まれた小項目は [] , ゴシックの見出しは頭つき、各見出しの本文は1字下げという構造である。

これをトリガーにしてQuick Basicを使ったパソコンの変換プログラムにかけ、各上位項目を見出し文に自動的に付加させた。

次に掲載日、掲載面の設定である。

元データの見出し文の最後にある和数字はその縮刷版における掲載ページを表している。掲載年と掲載月はその縮刷版で特定されるから、これを手がかりに、掲載日、掲載面を自動発生させることにした。

まず、変換テーブルを作成する。1日の朝刊は○ページから○ページまで、1日の夕刊は×ページから×ページまで。縮刷版の各号ごとにこういったテーブルを作成し、和数字の読み替えを行い、見出し文に追加付与した。

縮刷版 D B 表示形式

【1レコードのフォーマット】

半角スペースがデータの区切り、空白データは"<>"、改行がレコードの区切り。

8412-0687	841230	M 2	1096	5	政治	財政	予算	6 0	年度	予算	政府案	反響	努力は評価
↑	↑	↑↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
1 D	年	朝面	縮	段	項	項	項		項		項		見
(年月+通しNO)	月	夕刊	刷		目	目	目		目		目		出
	日	紙	の		1	2	3		4		5		し
		面	頁										
		頁											

8412-0687 年月とその月の縮刷中の通し番号、IDとして使用、9バイト。

841230 その記事の掲載年月日、6バイト。

M 朝夕刊、M=朝刊、E=夕刊、1バイト

2 面、紙面中での頁、1～2バイト。

1096 縮刷版での頁、1～4バイト。

5 段、紙面中の位置、"1"- "15"の1～2バイト。

政治 項目1、一番大きい項目、不定長。

財政 項目2、

予算 項目3、

6 0 年度… 項目4、

反響 項目5、

努力は… 見出し、

【表示形式】

通常はA.の見出し1行・寄誌事項1行の2行表示、コマンドを打つことでB.の詳細表示モードに切り替わるようにする。

A. 見出し一覧形式（見出しは36字で切り捨て）

1	努力は評価、額には不満	米政府筋示唆	反響
	841230 朝刊	2面	1096頁 5段
2	「予算増額の支援材料に」	自民・防衛庁は期待	平和問題研最終報告書提出
	841219 朝刊	2面	704頁 5段

B. 詳細形式（末尾の「見出し」は2行になることあり）

1	努力は評価、額には不満	米政府筋示唆	反響
	841230 朝刊	2面	1096頁 5段
	ID…8412-0687		
	項目1…政治		
	項目2…財政		
	項目3…予算		
	項目4…6 0 年度予算政府案		
	項目5…反響		
	見出し…努力は評価、額には不満	米政府筋示唆	反響
2	「予算増額の支援材料に」	自民・防衛庁は期待	平和問題研最終報告書提出
	841219 朝刊	2面	704頁 5段
	ID…8412-0804		
	項目1…政治		
	項目2…国防		
	項目3…<>		
	項目4…平和問題研最終報告書提出		
	項目5…<>		
	見出し…「予算増額の支援材料に」	自民・防衛庁は期待	平和問題研最終報告書提出

図2-2 データ構造と表示方式

2.4 データづくり

縮刷版データベースは戦後編だけでも対象期間が50年近くにもなろうというものであり、その入力費用は膨大なものになることが予測された。

一番簡単な方法はワープロ外注である。しかし、単純試算してみると、1ヵ月分のデータは約16万文字であった。1字1円としても1ヵ月で16万円、年間にすれば200万円近く、10年分になればこれだけで2,000万円近いものになってしまう。

2.4.1 OCRの導入

そこで、実験的に導入したのが光学文字読取機（OCR）である。

当初は富士電機製のXP-50 を使っていたが、途中から東芝製のExpressReaderJ70という機種に変更した。

XP-50 はホストマシンにNECのパソコンPC9800型が使えるというのが魅力であった。カタログ値では文字認識速度30文字/秒、文字認識率99%ということであった。

ExpressReader(92年2月導入)は東芝のJ3100ZD がホストマシン。新しい機種だけに文字認識速度70～100字/秒、認識率99.5%以上(いずれもカタログ値)と読取速度、認識精度が大幅に向上していた。

しかし、OCRは読み取り精度が100%というわけにはいかず、読み取りミスをいかにチェックするかが大きな課題であった。

XP-50 は「21」「35」といった縦書きの連数字がまったく読めず、実際の平均的読み取り精度は92～93%くらいであった。ExpressReader は約99%。

OCRで読ませたデータはその場で修正ができる。

ExpressReader ではマシンが読み取りに自信がないものについては、通常白地に黒文字表示のところが赤字で表示される。更にその横に読み取った原稿のスキャン画像が表示され、いちいち原稿を照合することなしに、校閲作業ができることになっていた。

ところが問題なのは、赤字表示部分だけが読み損ね、あるいは読み損ね候補であるかといえばそうではなく、マシンが問題なしと判断した箇所でも結構読み損ねがあるということであった。こういった事情では、やはり全文校閲をする必要が生じる。

最終的にはOCR入力担当者は赤字表示部分だけを直す、その上で初校済みデータをMS-DOSファイルに落とし、これを原本とともに再校閲に出すという二重チェックの方法をとった。

オペレータの養成は容易である。パソコンにさわるのが初めてという学生アルバイトでも数時間の研修で一人立ちさせることができた。

1日6時間の作業時間で平均14～16万字、縮刷版にして約1ヵ月分の入力ができる。熟練したワープロ入力者が一日に打てる文字量は1万数千字というから、不完全とはいえ、素人が数人分の仕事をこなすという計算になる。

ExpressReader ではレイアウト自動解析といってスキャンした画像を自動解析し、写真やグラフ部分を除去し、文字部分を自動的に領域指定するという機能がついていた。

だが、縮刷版索引は、段ごとに仕切野が入っている比較的単純なレイアウトであるにもかかわらず、2段を1段にまとめてしまったり、逆に1段の中を更に細かく分けてしまうなど

の事態が発生し、これの再処理にかなりの手間と時間をとられた。

もう一つの問題は文字認識の精度は高いが、行頭1字下げ、2字下げという空白認識が甘いということであった。縮印版索引では、見出し本文は1行でおさまることは少なく、2行、3行にわたっている場合が多い。通常1行目は行頭1字下げ、2行目以降は2字下げという体裁になっている。

これを1本のレコードにするためには、行頭2字下げの部分はスペースと前の行末の改行記号を削除し、前の行につなぐという作業が必要になる。この行頭空白の認識がでたらめだと自動的処理が困難になる。

2.4.2 校正作業

校正には家庭在宅婦人労働力の活用を考えた。子育てなどで家庭からは離れられないものの、在宅でできる仕事ならやりたいという高学歴の女性が増えているのではないかと考えたためである。

ニューメディア本部に勤務していた退職者、あるいは知人のつてなどを頼りに10人ぐらいを募集、熟練者なら時給1,500円程度になるように作業量を設定して試みとして仕事を委託した。

一応それなりにはやれるという結果は得たが、こういった直取引の作業委託はその管理が大変であった。一人の専従要員を必要としかねぬ形勢になってきた。

次に考えたのが在宅婦人を組織しているパソコンネットとの提携であった。

私鉄系の会社の組織している半商用ネットでこういった在宅の仕事を請け負うところがあり、試験的に作業を委託した。

しかし、会員のレベルにかなりのばらつきがあり、出来、不出来にかなりの差があった。直接のやりとりでないため、書式などこちらの要望がうまく伝わらない、校正ミスが多発した場合の責任の所在が明らかでない、などの問題が生じた。

何度かネットの窓口と話し合いをもち、作業担当者を固定、担当者名も明らかにし、責任の所在を明確にすることにした。半日間の出張研修を行い、発注側の意図、作業手順を説明した。更にネット内に作業連絡ボード（電子掲示板）を設け、疑問点や要望などをパソコン通信で担当者同士がやりとりするという仕組みをつくり、ようやくほぼ満足できるレベルに達することができた。

実績として、OCRによる入力、機器本体の購入費用を無視すれば、校正費も含めて、ワープロ外注に比べ30%~40%のコストでできる感じであった。

2.5 システム構築

縮刷版データベースは商用オンラインに載せないと利用者が限定されるので、最終的には朝日新聞記事データベースと同じ方式でデータベースを構築、一般提供をめざして作業が進行中である。

しかし、この方式は時間がかかる。そこで、本格的編集を進める一方、当社が松下電器産業株式会社東京情報システム研究所と共同研究で開発を進めている新しい方式の情報検索システム「検蔵君」で、この縮刷版データベースの簡易版を構築した。

高性能のワークステーションを使い、索引語ファイルを別に作成するのではなく、テキスト全体をそのまま始めから終わりまでサーチし、該当する文字列を含む文書を拾いだし、データベースとして機能させるというものである。

「検蔵君」はUNIXマシンを検索サーバーとし、VMEバス仕様の高速検索エンジンを搭載したシステムである。検索エンジンは文字列検索専用LSI（ストリームデータプロセッサ-SDP）をVMEバス3枚の基盤に組み込んだもので、ここにメモリに読み込んだデータを一気に流し、解析を行う。

検索サーバー機としては、Solbourne（ソルボーン）、PanaStation、SunのSPARCステーションの一部などが可能だが、当社ではSolbourne5/600型（メモリ32MB、ハードディスク660MB）を使用している。

パソコンで作成したデータをイーサネット経由でワークステーションに転送し、シフトJISコードからEUCに変換、「検蔵君」データファイルに変換した。

「検蔵君」ソフトはユーザが変更可能な項目テーブルをもっており、入力するデータの書式に応じてテーブルの書換えを行えば、読み込んだデータがそのままデータベース化できる。

本業の合間を縫っての作業であったため、1ヵ月近くを要したが、データが完成してから実際にデータベース構築に要した時間は実質的には数時間であった。

2.6 検 証

「検蔵君」で構築した縮刷版見出しデータベースを使った一例を次に示す。

収容データ件数は

	朝刊	夕刊	合計
81年	52,138	20,010	72,208
82年	53,789	19,767	73,556

83年 53,296 19,969 73,265

84年 55,033 21,146 76,179

1年当たりのファイルの大きさは約8MBである。

ここでは84年におけるコンピュータの動向を探ってみた。

キーワードは「コンピュータ」と「電算機」にした。

84年における検索結果は238件であった。

ここで1番の見出しでは、字面に「コンピュータ」という言葉がないのに検索対象になっている。通常画面では見えないが上位項目の見出しが検索対象になっていることが分かる。

< 一覧表示 >

キーワード \$0 / コンピュータ | 電算機 /
検索結果 238件

1. 日立、互換路線変えず 社長表明 継続性を最重視 ソフト開発一時は負担
840120 朝刊 9 677 1
2. スパイ事件は「個人的暴走」_コンピューター (電算機)
840120 朝刊 9 677 2
3. 米製パソコン、日本で激突 IBM追いアップルも新型
840125 朝刊 8 856 1
4. ソフト保護法に疑義 米大使 通産相に公式表明_ソフト保護と著作権
840113 朝刊 9 441 5
5. 電算機ソフト 著作権法改正で保護 文化庁に中間報告 審議会小委
840120 朝刊 3 671 1
6. ソフト保護、やはり専門法通産省が猛反論 文化庁との対立鮮明に
840120 朝刊 8 676 1
7. ソフト保護万全を 電機業界首脳 通産省に要請_ソフト保護と著作権
840120 朝刊 8 676 4
8. 新電報システム電算機 IBMから調達_電電の国際資材調達問題
840121 朝刊 9 717 7
9. 自動車・電算機が増勢 米商務省の84年産業予測 対日競争強く意識_経済
840104 朝刊 9 109 1
10. 合併や買収大型化傾向 米電算機サービス業界_経済
840127 朝刊 8 932 8
11. インド向け電算機販売契約を破棄 ソ連懸念し米企業_貿易
840130 朝刊 3 1035 10
12. 「蒙古秘史」の言語電算機で解析研究_海外文化
840109 夕刊 5 315 9
13. (13) 電算機診断_情報多重列島
840122 朝刊 4 746 1
14. 三宅島の溶岩流 電算機で再現 京大防災研 避難情報システムに道_火山
840207 朝刊 3 217 1
15. 電算機データ生かせず 兼信の隊歴も入力_自衛官が同僚に発砲 死傷4人
840229 朝刊 23 1027 5
16. 第五世代コンピューター_情報化社会新語ABC
840213 朝刊 11 433 7
17. 米での生産まだ (コンピューター機器生産) _米国
840211 朝刊 8 374 5
18. 米の電算機部品メーカーを買収 日立金属の子会社_米国
840223 朝刊 8 772 8
19. コンピューター故障 東証二部で売買停止_株式市場
840222 朝刊 8 736 3
20. 画像処理に超LSI速さ中型電算機並み 日本電気 非ノイマン型実用化
840215 朝刊 8 504 4
21. 「IBMコピー」認める 関連役員、賞与カット 日立社長
840202 朝刊 8 44 5
22. 使いやすさで日本製が先行 高速電算機で米紙警告
840214 朝刊 8 468 8
23. 第5世代電算機 研究開発英と協力 5月にもスタート 当面、情報交換
840214 朝刊 9 469 1
24. IBM超大型機 機能強化し新モデルに_コンピューター (電算機)
840225 朝刊 8 860 8
25. 情報処理量12.8倍に 富士通が新ソフト IBMより低価格
840201 朝刊 8 8 1
26. 百超すユーザーが導入予定_ソフト開発・ソフト権保護
840201 朝刊 8 8 1
27. 31ビット・アドレス方式_ソフト開発・ソフト権保護
840201 朝刊 8 8 6
28. ソフト保護2省庁案まとまる 著作権法か新立法か
840215 朝刊 3 499 1
29. 文化庁 現行法内、50年間 通産省 15年、裁定制盛る
840215 朝刊 3 499 1
30. ソフト保護 通産省・文化庁の主張_ソフト開発・ソフト権保護
840215 朝刊 3 499 2
31. (解説) 成長産業 網張り争いの面も_ソフト開発・ソフト権保護
840215 朝刊 3 499 6

図 2 - 3 「検蔵君」での検索例

2.7 今後の課題

縮刷版見出しデータベース第1期計画分(1981年～1990年)は、テキスト入力も終わり、全体像が姿をみせ始めている。戦後編完成に向け、引き続き年代をさかのぼっての入力も進行中である。

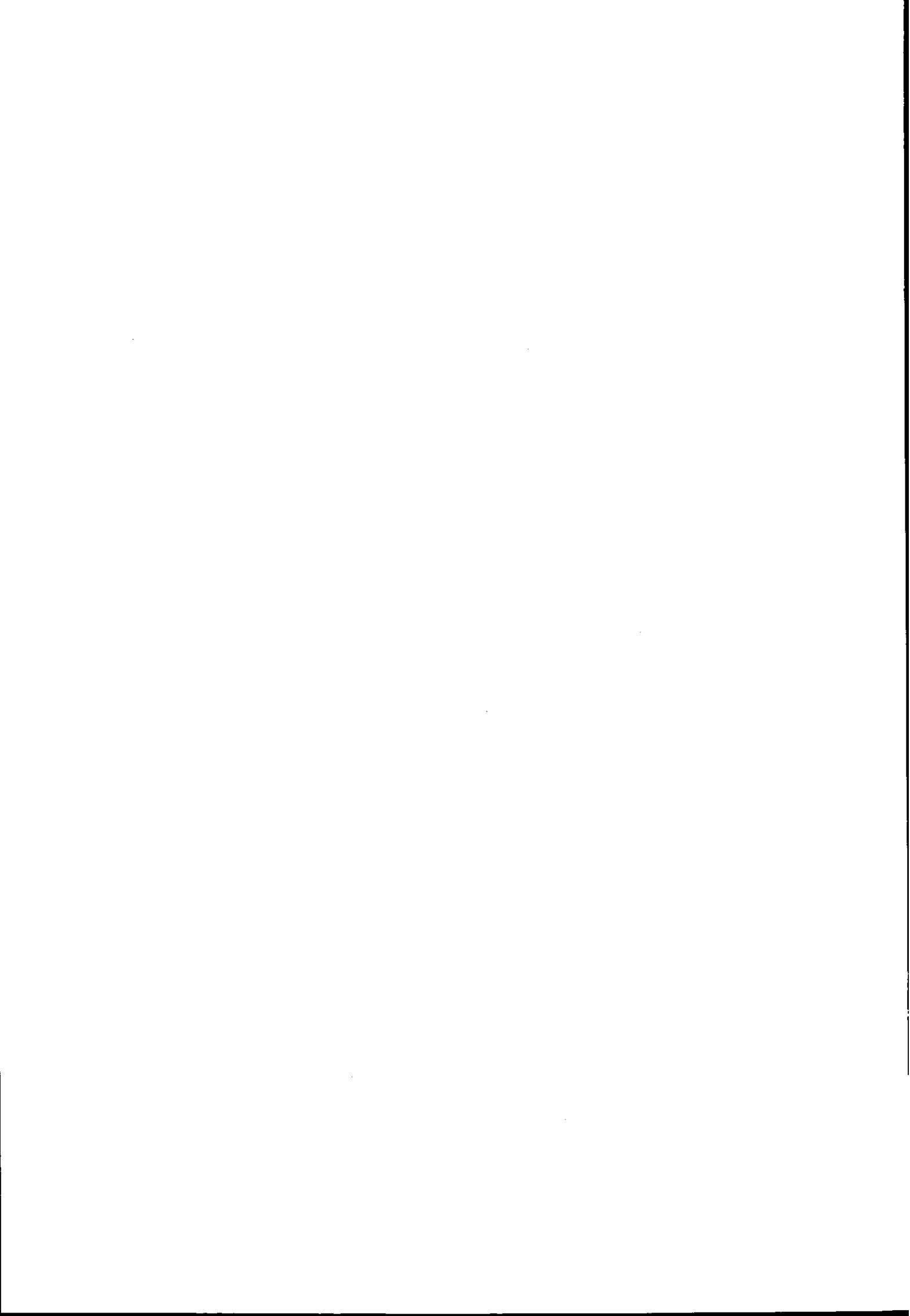
非公式ながらディストリビュータへの配信打診の結果、数社からは前向きな感触を得ている。

更に、こういったデータベースが広く世の中に広まるためには、安価で使いやすいものが必要であろうと考え、10年分をひとまとめにした電子ブック(8センチCD-ROM)化の企画も進行中である。

バブル経済崩壊後の経済不況と、当初思いもしなかった状況にまきこまれてしまったが、なんとか短期間のうちに戦後編を完成させたいと考えている。

3 ファジィに関する文献データベースの構築

財団法人 日本情報処理開発協会



3. ファジィに関する文献データベースの構築

3.1 目的

ファジィに関する文献は各分野で急速に増加しつつあり、ファジィ理論の研究を目指す研究者、ファジィ理論を製品開発に応用しようとする技術者、ファジィ理論を広く産業活動や社会活動に活用しようとする多くの人々から情報源としてのファジィに関する文献データベースの構築が望まれている。

また、ファジィ理論とその応用が健全に発展していくためにも、ファジィ理論応用の先導的立場にあるわが国の優位性を継続するためにも、現時点から、ファジィ理論とその応用に関する文献データベースを構築しておくことが必要である。

そこで、ファジィ理論とその応用に関する文献情報を網羅的に収集し、それらの情報を体系的に整備・収録して、情報の適切な利用を可能にすることにより、今後のファジィ理論の研究とその応用促進に寄与することを目的とする。

3.2 概要

ファジィに関する文献は、日本ファジィ学会を始め関連学会の学会誌・論文集や企業の技術報告書等多くの分野で増加しつつある。これらの文献の分散化を防ぎ、網羅的に収集可能な現時点からのファジィ文献データベースの構築に期待が寄せられている。そこで、「ファジィに関する文献データベースの構築」に際しては、それらの情報源の中から核となる日本ファジィ学会の発表した文献を中心に収集・整理・選択し、ファジィの専門家による抄録作成・キーワード付与を行って文献データシートを作成した。更に、作成された文献データベースに書誌事項の補足等を行ったのち、昨年度開発した「ファジィに関する文献データベースシステム」を用いて、文献情報を電子化し、データベースへ登録した。

なお、ファジィ文献の収集・選択、文献の抄録作成・キーワード付与についてはファジィの専門家集団である日本ファジィ学会の協力を得て実施した。

3.3 情報の収集・整理・選択

ファジィに関する文献は、日本ファジィ学会や関連学会の学会誌、論文集、研究会報告書、シンポジウム、大会、講演会等の論文の他に企業の技術報告書等に発表されているが、その核

となる情報源は当然のことながら日本ファジィ学会の学会誌、論文集等である。

日本ファジィ学会の発表論文数も年々増加しており、毎年1回開催されるファジィシステムシンポジウムの発表論文件数は、'90年が124件、'91年が168件、'92年が175件となっている。

そこで、今年度、構築するデータベースの情報源としては、日本ファジィ学会を始め関連学会の文献を中心として収集し、約1,700件を選択した。

3.4 抄録作成・キーワード付与

文献の全文収録を要望する声は多いが、現実的にはコストや人手の面から全文の収録は困難である。そこで、必要最小限の情報として文献に係わる書誌事項、著者名、抄録等を収録するとともに、検索効率の面からファジィ理論の体系化に基づいて作成したファジィ理論基礎用語シソーラス等を参照してキーワードを付与した。

(1) 抄録作成

昨年度開発した「ファジィに関する文献データベースシステム」ではシステム設計時に調査を行って、抄録の最大文字数を400字としている。そこで、本データベースの構築に当たってはこれに合わせて和文の抄録を作成して収録した。

(2) キーワード付与

文献データシート記入要領に従って最大12個までのキーワードを付与した。このキーワード付与については、平成2年度に実施した「ファジィ文献データベース開発計画調査」の際に作成したファジィ理論の体系化に基づくシソーラスを参照して付与した。

3.5 書誌事項補足

作成された文献データシートの書誌事項について確認し、必要があれば補足等を行って最終的な文献データシートとした。

3.6 文献情報の電子化・登録

作成された文献データシートの文献情報を昨年度開発した「ファジィ文献データベースシステム」を用いて電子化し、文献データベースへ登録した。文献データベースへ文献情報を登録する方法には一括登録と会話登録の二つの方法がある。

今年度のデータベース構築に際しては中核となるデータベースを作る立ち上がりのときであ

り、比較的短期間に電子化・登録の必要があることから一括登録で処理した。

3.7 運用計画

ファジィに関する文献データベース提供に係わる運営については、主として(財)日本情報処理開発協会が行い、データベースの構築・更新等に係わる情報収集、抄録の作成、キーワード付与等の情報提供については、本分野の専門的立場にある日本ファジィ学会及び国際ファジィ工学研究所がこれに協力する。

更新頻度は、学会誌、シンポジウムの論文集等の発行回数を考慮して、最大4回/年とする。

また、データ件数がある程度(5,000件位)の規模に達するまでは、原則として冊子体での提供を行い、対価として実費程度を徴収する。

3.8 今後の課題

ファジィに関する文献データベースの中核となるデータベースを構築したが、当面の課題と本格的な運用を行う上での課題は、以下のとおりである。

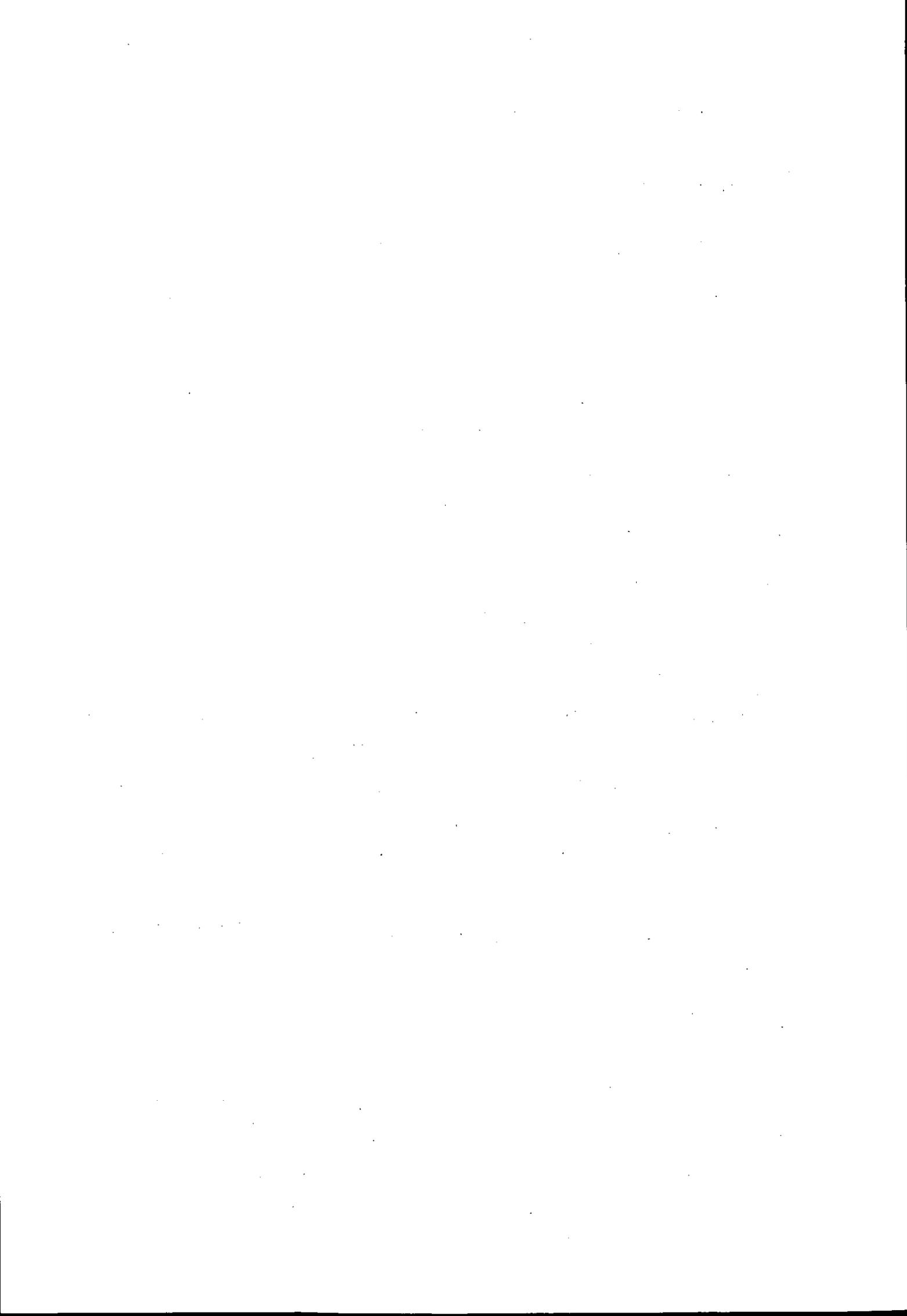
- (1) 著者名や抄録に外字(J I Sの第1水準、第2水準にない文字)が出現する。今後は所属機関名等の固有名詞にも出現の可能性がある。特に、ファジィ理論に係わる文献の抄録には数式等に演算記号、上・下つきバー、上・下つき添字がある。
- (2) 1度にまとめて発表される学会の大会論文、シンポジウムの論文集など文献は一括して電子化することになるが運用センタの負荷が大きい。
- (3) まだ、中核としてのデータベース構築であり、更に網羅的に文献を収集し、登録件数を増す必要がある。
- (4) 本格的運用に入ればファジィ理論の体系化、分野の体系化、機能の体系化の見直しが必要になるであろう。

3.9 資料

この資料は、以下の報告書より文献データベースの構築に係わる部分を抜粋して作成したものである。

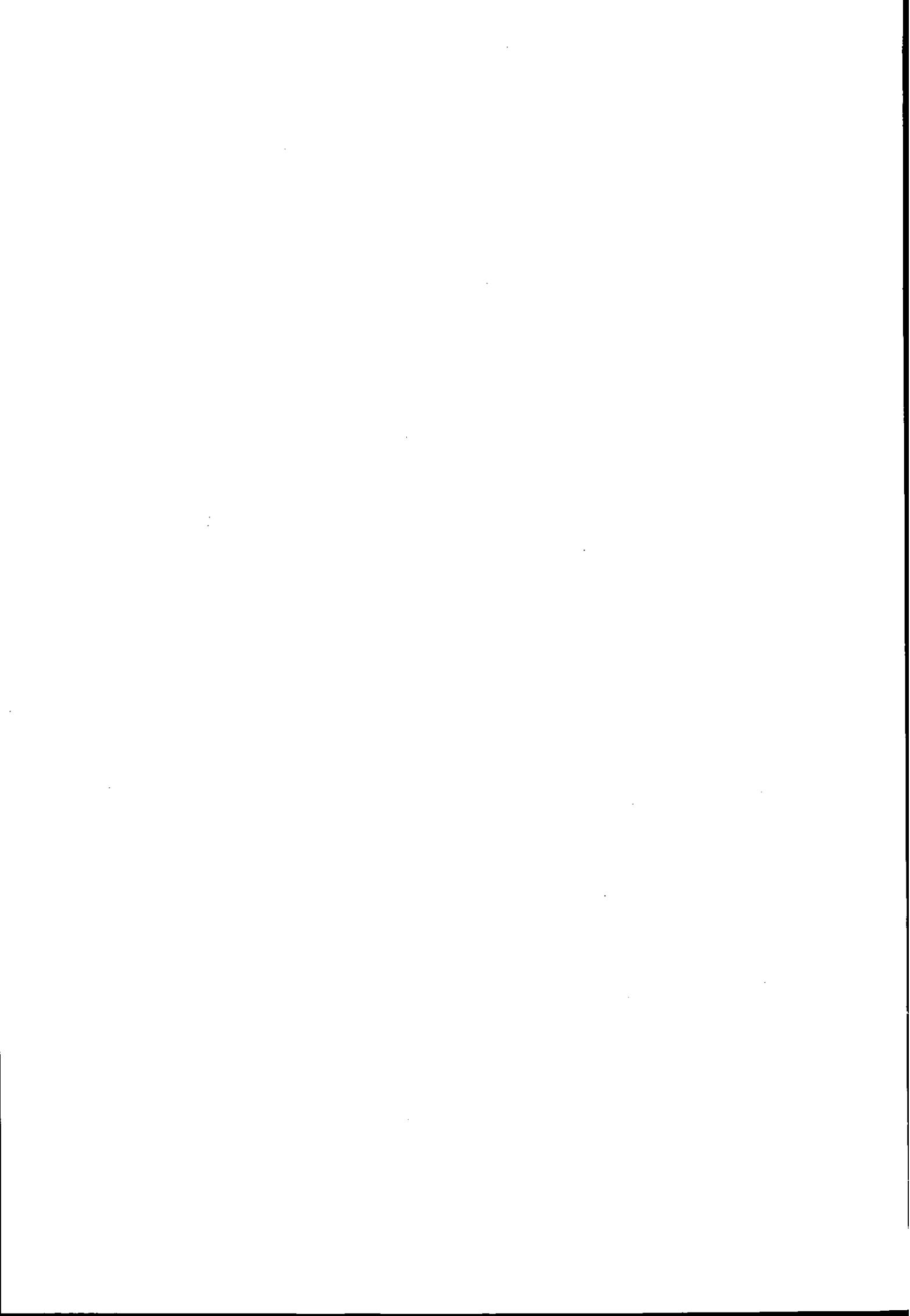
平成2年度報告書「ファジィ文献データベース開発計画調査報告書」

平成3年度報告書「ファジィに関する文献データベースシステムの開発」



4 医療用医薬品抗生物質データベースの構築

株式会社 小 田 島



4. 医療用医薬品抗生物質データベースの構築

4.1 目的

このデータベースの構築は、医療機関での抗生物質の使用をより合理的で、かつ適切なものにしていくことを大きな目的としている。

今回構築した抗生物質データベースが医療機関の診療のために使用され日常的に機能していくことが、医療費の高騰やMRSA蔓延の解決に役立つよう、ここに平成4年度に実施した委託課題「医療用医薬品抗生物質データベースの構築」の結果をとりまとめた。

4.2 データベースの概略

4.2.1 データベースの構成

本データベースは次のファイル・辞書類より構成されており、その関係を図4-1に示す。

- (1) 本体テキストファイル (以下 本体ファイル)
- (2) 医薬品辞書
- (3) 商品辞書
- (4) 薬効辞書
- (5) 会社辞書
- (6) 剤形辞書
- (7) 雑誌辞書
- (8) 文献管理データベース (以下 文献管理DB)
- (9) 文献抄録テキストファイル (以下 文献抄録ファイル)

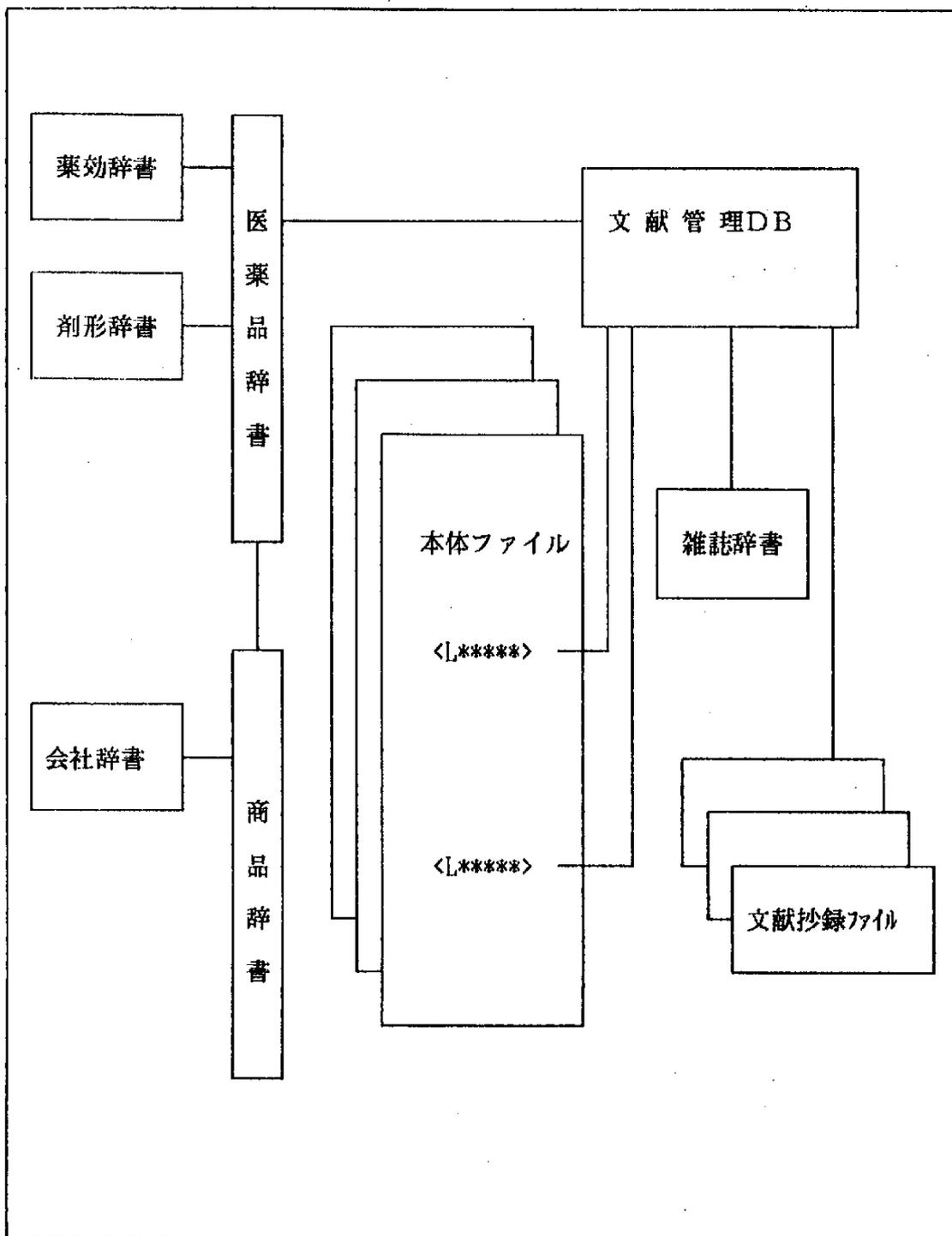


図4-1 本システムの構成

4.2.2 抗生物質データ品目構成

作成データ品目の選択の基準は、岩手医科大学病院・岩手県立中央病院・盛岡市立病院で平成4年7月に現実に採用されている抗生物質の注射剤及び内服剤とし107品目を選んだ。

4.3 本体ファイルの内容

本体ファイルのデータ項目一覧を次に示す。

- | | |
|--------------|---------------|
| 【1.情報一覧】 | 【2.医薬品基本データ】 |
| 【3.医薬品の概略】 | 【4.臨床応用】 |
| 【5.医薬品副作用】 | 【6.医薬品相互作用】 |
| 【7.薬物動態】 | 【8.病態別注意事項】 |
| 【9.投与時の注意事項】 | 【10.過量投与】 |
| 【11.他剤との比較】 | 【12.患者さんへの情報】 |
| 【13.著者】 | 【文献ファイル】 |
| 【引用文献一覧】 | 【医薬品辞書】 |

4.4 その他のファイル・辞書類

- (1) 医薬品管理、検索ツールとして以下の辞書を構築した。

商品辞書	……………	個々の製品単位で管理する辞書
医薬品辞書	……………	一般名別更に投与経路別を一単位とし医薬品を管理する辞書
薬効辞書	……………	薬効分類を管理する辞書
会社辞書	……………	会社名を管理する辞書
剤形辞書	……………	剤形を簡易に表すためのコードテーブル
雑誌辞書	……………	雑誌名を管理する辞書

- (2) 更に、引用文献を管理するための文献管理DB及び抄録を収録する文献抄録ファイルがある。

4.5 表示機能

4.5.1 本体ファイルの表示

本体ファイルはボリュームがあるので、参照したい場所に的確に移せる必要がある。その選択はカーソルを移動して該当場所を反転させ選択することにした。

- (1) 初期画面

初期画面は第1階層であるセクション番号とデータ項目一覧が表示される。

- (2) 第2階層以降の表示

薬剤によって各セクションの情報量は異なる。また各セクションについても下層構造に差があるので、初期画面でセクションを選んだあとはまず該当する本文を表示する。そして情報量の多い場合は随時、利用者の要求によりファンクションキーで現在見ている部分の見出し一覧を表示できるようにした。

で始まる行：第2層

で始まる行：第3層

で始まる行：第4層

で始まる行：第5層

① 「2.医薬品基本データ」の細見出し例

2.1.データ入力日：1993.01.31.

2.2.一般名：セフォペラゾンナトリウム，CPZ

2.7.組成：セフォペラジン，セフォビット

2.8.効能・効果：セフォペラジン，セフォビット

2.9.用法・用量：セフォペラジン，セフォビット

② 「4.臨床応用」内の細見出し例

4.1.内科

①Cefoperazoneの胆汁中移行と胆道感染症の臨床的検討<L00180>

②CPZ(Cefoperazone)の腭液及び腭組織への移行<L02052>

③内科領域各種感染症におけるCefoperazoneの臨床効果<L02053>

4.2.小児科

①小児科領域における特に腎毒性に対する検討<L02054>

③ 「5.医薬品副作用」内の細見出し例

【5.医薬品副作用】

5.1.副作用とチェック・処置

5.1.1.全身性副作用

5.1.1.1.ショック (T) (PW)

△(1)症状

△(2)チェック

△(3)処置：症状が現れた場合には投与を中止すること。(T)

5.1.2.過敏症

5.1.2.1.Skin reactins (P 1/40)

△(1)症状：発疹 (T)，じん麻疹 (T)，そう痒 (T)，紅班 (T)

△(2)チェック

△(3)処置：投与を中止し、適切な処置を行うこと。(T)

○5.1.2.2.発熱(T) (P 1/260)<L01423>

4.5.2 その他の表示

直接本体ファイルを表示する以外では初期画面の「1.情報一覧」と「14.引用文献一覧」がある。

(1) 情報一覧

これは示したすべての見出しの一覧で、見出し作成処理をした産物として生成するものをそのまま表示する。

(2) 引用文献一覧

引用文献一覧は次の例のように文献番号、出典、標題を一覧する。

L00031 日本化学療法学会雑誌 31(9):920-926, 1983

<標題>新セファロスポリン系抗生物質Cefoperazoneのヒト腸内菌叢に及ぼす影響

L00126 新薬と臨床 39(2):317-330, 1990

<標題>血液凝固系に対する影響についての検討

L00136 日本化学療法学会雑誌 29(4):375-424, 1981

<標題>呼吸器感染症に対するCefozolin との二重盲検法による薬効比較成績

(3) 文献抄録表示

本体ファイルの中で番号で表されている文献について、出典もしくは、更に詳細な抄録を見たい場合には、文献参照キーを選択し、その内容を見ることができる。

L01422 感染症学雑誌 61(11):1257-1263, 1987

<標題>小児におけるセフォペラゾンの腸内菌叢へ及ぼす影響と下痢発症頻度

更に、抄録がある場合には該当する抄録部分を次のように表示する。

L01239 腎と透析 20(1):129-133, 1986

<標題>血液透析患者におけるCefoperazoneの薬動学的研究

<抄録>

慢性腎不全のため血液透析を受けている7例(男4女3, 平均年齢41.3歳, 平均体重51.2kg, 血液透析導入前のCcr 平均2.9 ml/min)を対象に, CPZ2gを静注後経時的に得られた血中濃度をtwo-compartment open modelに従って解析した。結果は下記のとおりである。

Pharmacokinetics parameters Cefoperazone in patients with End Stage Renal Disease off and on Hemodialysis(HD)

Parameter	off HD 2g IV	on HD 2g IV
Cmax ($\mu\text{g/ml}$)	381.5	315.4
K12(/hr)	1.58	0.53
K21(/hr)	2.72	1.02
Ke1(/hr)	0.56	0.55
T1/2(/hr)	2.05	2.20
Vd(L)	8.19	9.58
AUC($\mu\text{g}\cdot\text{hr/ml}$)	687.3	575.3

以上からCPZ は腎機能障害時にもその薬物動力的にほとんど変化がみられないことにより、腎不全の患者にも安全で有効な肝排泄型抗生物質であると思われるが腎不全とともに肝機能障害を有する患者では、投与量の調節を行う必要がある。

4.6 印字機能

印字機能はファンクションキーにより随時実行することができ、また次の3通りから選択することができる。

- (1) セクションを選んでの印字
- (2) 行番号を指定しての印字
- (3) 文献の番号を選んで引用文献の印字

4.7 サンプルデータ

4.7.1 本体ファイルサンプルデータ (略)

4.7.2 文献ファイルサンプルデータ

次に7.1.セフォペラゾンで使われている文献ファイルを示す。

■ L01242

口腔領域の各種手術を行った患者45例を対象に、術前にCPZ 1gをone shot静注し血清中CPZについてはtwo compartment modelにより、また各口腔組織はthree compartment modelによる解析を行った。結果は下記のとおりである。

1) 血清中濃度におけるCPZ の薬動学的定数とSimulation式

Vd(L)	K12(/h)	K21(/h)	Kel(/h)	α (/h)	β (/h)	T1/2(min)
3.28	7.75	5.76	1.34	14.3	1.54/hr	77.3

Simulation式 $111.5e^{-0.54t} + 189.3e^{-14.3t}$

2) 各組織におけるCPZ の薬動学的定数とSimulation式

組織	K13(/h)	K31(/h)	Tmax(min)	Cmax μ g/g	T1/2(min)
歯肉	0.46	1.27	60	23.6	78.8
骨	0.12	1.44	55	5.53	78.0
嚢胞壁	0.33	0.96	75	20.3	82.6

Simulation式 $71.8e^{-0.54t} - 6.6e^{-14.3t} - 65.2e^{-1.27t}$

$14.8e^{-0.54t} - 1.7e^{-14.3t} - 13.1e^{-1.44t}$

$89.9e^{-0.54t} - 4.7e^{-14.3t} - 85.2e^{-0.96t}$

血清中濃度は30分値89 μ g/ml, 半減期77分であつた。歯肉や嚢胞壁へのCPZ 移行性は、比較的良好であり血清比30~35%であつた。顎骨への移行性は他の口腔組織より劣り血清比8%強だった。各口腔組織でのCPZ 濃度は血清中濃度推移とよく対応しており半減期は約80分だった。口腔領域主要分離菌に対する接触時間は1gワンショット静注により歯肉やのう胞壁で6時間以上, 移行率の比較的低い骨でも3時間程度認められた。

■ L01248

中, 下部胆管に完全閉塞がありPTCDを施行した7例(男性6例, 女性1例, 32~69歳)に

CPZ 2gを静注投与し、胆汁中移行を検討した。閉塞性黄だん存在下ではCPZ の排せつ遅延がみられた。CPZ は閉塞性黄だん時には腎からの排せつが促進された。肝機能障害が強く、腸肝循環が遮断されている場合でもCPZ の有効胆汁中濃度が得られる。閉塞性黄だん時の胆道感染には、PTCDを行うと抗生剤の胆汁中移行性をよくすることができる。

4.7.3 引用文献一覧

次に文献管理DBから出力したセフォペラゾンの引用文献一覧の例を示す。

- L0031 日本化学療法学会雑誌 31(9):920-926, 1983
＜標題＞新セファロsporin系抗生物質Cefoperazoneのヒト腸内菌叢に及ぼす影響
- L00126 新薬と臨床 39(2):317-330, 1990
＜標題＞血液凝固系に対する影響についての検討
- L00136 日本化学療法学会雑誌 29(4):375-424, 1981
＜標題＞呼吸器感染症に対するCefazolin との二重盲検法による薬効比較成績
- L00140 日本化学療法学会雑誌 37(9):1200-1204, 1989
＜標題＞高齢者における肝排泄型抗生剤Cefoperazoneの体内動態の検討

4.8 展 望

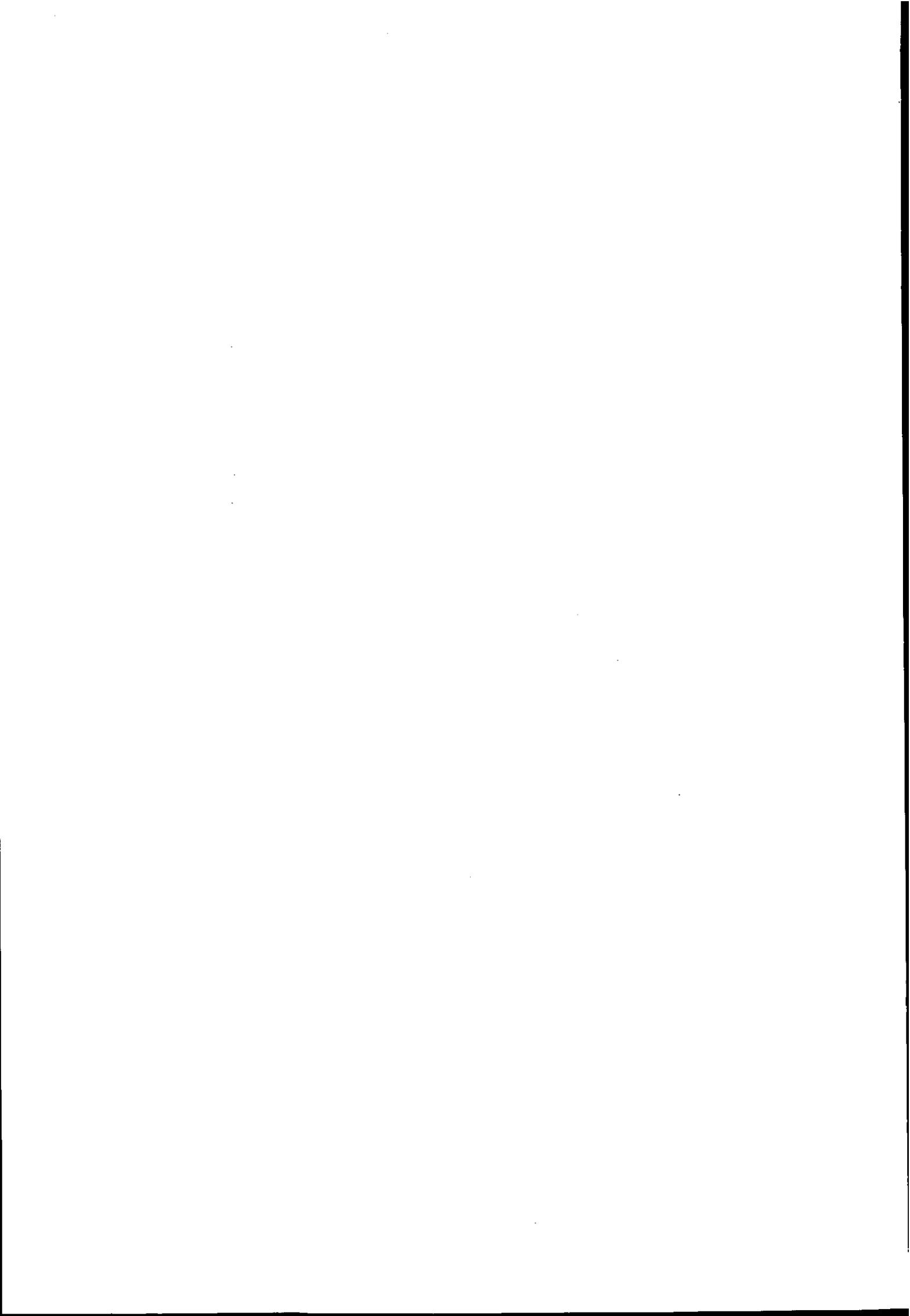
この計画は、当初汎用コンピュータによるオンラインリリースのもとに抗生物質の情報検索を可能にすることを目的にしたものであったが、オンラインデスクトップビュータのシステム構築が間に合わず、これは平成5年10月に延期し、急遽パソコンでのスタンドアロンでのシステム構築に踏み切った。パソコンレベルでこのデータベースが動くプログラムの完成は平成5年6月末を予定している。

しかし、時あたかも多剤耐性のブドウ球菌MRSAの蔓延は社会問題にさえなるにいたり、かかる時このような抗生物質データベースをリリースできることは、(財)データベース振興センターの委託事業として大きな社会的意識があるものと考えられる。

更に今回の課題の上位概念としてわれわれが意図している広範な薬効群を網羅した医療用医薬品データベースが完成したときには、医療機関における薬物療法の適切な実施に多くの貢献をなすものと確信する。

5 交通事故調査データベースの構築

財団法人 日本自動車研究所



5. 交通事故調査データベースの構築

5.1 目的と概要

交通事故の低減や交通事故時の乗員等の被害を軽減するための一方策として、交通事故調査から得られるデータをもとに事故原因や障害原因を多面的に解析し、その結果を交通政策や道路設計、車両開発等に反映させることが重要である。このような観点から、先進諸国においては古くから交通事故調査が実施されている。

交通事故調査には、大きく分けて2種類の調査があり、その一つは主として警察によって実施されている交通事故調査である。この調査は全国規模で実施され、また調査対象となる事故件数も多いことから、交通事故を統計的側面から解析する上で不可欠なものとなっている。反面、調査項目が限定されるため、道路設計や車両開発等に反映させるには必ずしも十分であるとはいえない。

そのため、多くの国では、警察による交通事故調査のほかに、特定の交通事故に対して詳細な事故調査を実施し、具体的な事故防止策や事故時の被害軽減策を検討する際の資料としている。この種の交通事故調査の代表的なものとして、わが国においては、運輸省が(財)日本自動車研究所に委託して行っている調査があるが、調査データについては、まだ本格的なデータベースの形には整備されていない。

そこで、本開発研究では平成2年度より3ヵ年の計画で、このような詳細データを多面的に解析し得るデータベースシステムを設計・開発することにし、前年度には、当研究所が受託して実施した上記の事故調査データを例にプロトタイプデータベースを作成した。最終年度に当たる本年度は、このプロトタイプデータベースの発展・改良を行い、更にデータベースに不慣れな初心者でも利用できるよう、パソコンからデータベースをアクセスするための利用環境を整備し、実用的なシステムを完成させた。

なお、本開発研究の目的を達成するため、学識経験者からなる「交通事故調査データベースの構築に関する調査研究委員会（委員長：藤原 譲 筑波大学教授）」を設置し、その下に当研究所技術者によるワーキンググループを編成し、上記委員会の指導の下に開発研究を実施した。

5.2 データベースの改良と発展

本データベースは、汎用のRDBMS（リレーショナルデータベース管理システム）である

ORACLE（商品名、米国ORACLE社製）を用いて、EWS（エンジニアリングワークステーション）Sun4（富士通社製S4/2）に構築した。

5.2.1 データベース表の設計とデータの実装

本データベースで対象としている交通事故調査は、昭和48年から現在に至るまで継続して実施されているが、調査の高度化や調査項目の修正・追加等に伴って使用される調査票が過去5回改正されている。前年度のプロトタイプデータベースでは、このうちの1種類の調査票を例にデータベース表を設計し、約300件のデータを実装した。本年度は、残りの4種類の調査票を例に別々のデータベース表を設計し、該当するデータ約1,300件を実装した。表5-1は各バージョンのデータベース表に実装したデータ数（事故件数）を示しており、本年度は総計約1,600件分のデータを実装したが、データ量が多くなっても検索や集計速度が低下するような現象は見られなかった。

このように、データベース表が複数のバージョンとなっていることから、各バージョンにまたがる検索を行う場合には、共通項目をピックアップしたビューを作成する必要がある。しかし、車両の構造が年々変化しているため昔の事故と最近の事故を合わせて解析する必要は少なく、実際には単年度ごとの検索が行われるケースが圧倒的に多いことから、データベース表が複数のバージョンに分かれていても、実用上支障となることはほとんどないものと考えられる。

表5-1 各バージョンのデータベース表に実装されているデータ数

データベース表	適用年度	事故件数	車両台数	関係人数
MOTFV01	昭和47年度～昭和60年度	703	1289	2091
MOTFV02	昭和61年度～昭和62年度	211	361	662
MOTFV03	昭和63年度～平成2年度	309	556	888
MOTFV04	平成3年度	200	341	537
MOTFV05	平成4年度	200	350	548
合計		1623	2897	4726

5.2.2 データの補完

各年度の事故調査票にはおよそ500項目以上の内容が含まれており、複数枚のシートから構成されている。そこで、これらの調査票と1対1で対応づけるため、データベース表も表

5-2に示すように複数枚に分割した。したがって、検索を行う際には各表に共通して含まれるキー項目を使って表結合を行う必要があるが、作成したデータベース表には、キー項目が欠如している表が幾つかあったため、欠如しているキー項目を補完した。

このように共通するキー項目が存在しない場合でも、外部結合の方法を用いて表結合をすることになるが、これを頻繁に使用した検索プログラムはかなり煩雑なものとなる上に、作成したデータベースの構造とSQLに相当熟知している必要がある。今回、キー項目の補完を行ったことにより、単純な表結合の技術を利用できるようになった。

表5-2 データベース表（バージョン5）の構成

表名	表の内容	項	目
MOTFV05A111	事故概要及び環境	V1~V23-5.	35項目
MOTFV05V-41	事故車両及び運転者関係	V101~V119-3.	24項目
MOTFV05V-42	事故車両及び運転者関係	V120~V133-8.	46項目
MOTFV05V-43	事故車両及び運転者関係	V134-1~V145.	50項目
MOTFV05V-44	事故車両及び運転者関係	V146-1~V162-7.	54項目
MOTFV05V-45	事故車両及び運転者関係	V163-1~V183-3.	33項目
MOTFV05V-46	事故車両及び運転者関係	V184-1~V184-19.	19項目
MOTFV05V-47	事故車両及び運転者関係	V185-1~V187-10.	21項目
MOTFV05V-48	事故車両及び運転者関係	V188-1~V188-25.	25項目
MOTFV05V-49	事故車両及び運転者関係	V189-1~V189-25.	25項目
MOTFV05V-60	二輪車事故の概要(1)	V2003~V2018.	20項目
MOTFV05V-61	二輪車事故の概要(2)	V2101-1~V2120.	33項目
MOTFV05V-62	ヘルメットの情報	V2301-1~V2315-2.	50項目
MOTFV05V-63	二輪車乗員の情報	V2201~V2212-3.	31項目
MOTFV05I---	乗員・歩行者等の情報	V501~V540.	42項目
MOTFV05O---	乗員・歩行者等の傷害接触関係	V601~V634.	39項目
MOTFV05P---	シート・年少者用補助乗車装置	V1001-1~V1017.	31項目
MOTFV05V-51	前面窓ガラス	V1101~V1104-15.	18項目
MOTFV05V-52	フード	V1201~V1209-2.	10項目
MOTFV05V-53	ステアリング	V1301~V1310.	21項目
MOTFV05V-54	ドア	V1401-1~V1406-2.	23項目
MOTFV05V-55	キャブオーバ型車	V1501-1~V1512-6.	30項目
MOTFV05V-56	タイヤ関係Ⅰ	V1601-1~V1605-10.	50項目
MOTFV05V-57	タイヤ関係Ⅱ	V1701-1~V1705-10.	50項目
MOTFV05V-58	車両火災	V1801~V1829.	38項目
MOTFV05V-59	エアバッグ	V1901~V1908-2.	21項目

5.3 利用環境の整備（パソコンとワークステーションの連携）

本データベースの利用目的は、交通事故の具体的な防止策や事故時の被害軽減策を検討する際の資料を得ることにあり、したがって、このような検討を目指す研究者や技術者が直接利用してこそ真価が発揮される。ところが、研究者・技術者が同時にデータベースの専門家である

といったケースはごく限られ、むしろ、データベースに関しては素人であると考えの方が妥当である。更に、データベースを格納してあるEWSそのものの使用にも難色を示す場合もある。

そこで、EWSに比べてより親しみやすいパソコンを利用し、更にはSQLやデータベースの構造を意識させないで検索できるようなシステムの開発を行った。また、パソコンと連携をとることにより、パソコン用の身近なアプリケーションソフトを利用して解析や図表化を比較的容易に行うことも可能になった。

5.3.1 ハード、ソフトの選択と試行

全くの初心者でも比較的短時間に操作方法を習得できるなどの理由から、パソコンとしてMacintosh（以下Macと略す）を採用した。また、EWS（Sun4）とMacの間は、Ethernet（TCP/IP）により接続した。

EWS上のデータベースをMacからアクセスするためのソフトとしては、最近オラクル社より提供されているORACLE for Macintosh (Client Version 2.0J)を用いた。このソフトは、クライアント/サーバー環境の下で、Macにクライアントとしての機能をもたせるためのソフトであり、以下のようなツールを備えている。

- (1) Mac用のアプリケーションソフト（例えばHyperCard）にSQLコマンドを組み込み、そのアプリケーションから直接ORACLEをアクセスできるようなソフトを開発するためのツール（ORACLE Access:開発者用ツール）
- (2) 容易にデータベースのアクセスやユーザの追加、テーブルの作成等が行えるHyperCardスタック（ORACLE Workshop:エンドユーザ用ツール）
- (3) ユーザ管理、テーブル管理、パフォーマンスモニター機能をもったGUI式のデータベース管理ツール（Server Manager, System Stack:データベース管理者用ツール）

そこでまず、上記のうち(2)のツールを用い、今回構築したデータベースをMacからアクセスしたところ、一応アクセスが可能であることは確かめられたが、以下のような点で実用化するには無理があつた。

- (1) かなり凝った画面設計が行われているため初心者には分かりやすいが、反面、操作時の応答が遅く、待ち時間に相当なストレスを感じる。
- (2) 表名や項目名としてデータベース表に付けられた記号がそのまま表示されるため、初心者にとっては理解しにくい。
- (3) 表結合などのSQLの知識をもっていないと、適切な条件設定が難しくなる。
- (4) カスタマイズの範囲が限定されているため、クロス集計などの機能拡張や適正化を図ることが難しい。

5.3.2 専用ソフトの開発

上記の問題点を解決するために、もう一つの開発用ツールであるORACLE Access を用いて、Mac用のアプリケーションソフトHyperCard をベースとした交通事故調査データ専用の検索ソフトを開発した。

HyperCard ではHyperTalk という専用のプログラミング言語（オブジェクト指向のインタプリタ言語）でプログラムを記述することにより、マウス操作に対応して各種の命令を実行させることができる。そこで、データベース言語SQLで記述した検索用のコマンドを、上記の開発者用ツール（ORACLE Access:インタプリタ言語へのオラクル社の標準インタフェース）を用いて、HyperTalk に埋め込むことにより、マウス操作による検索が行えるようなアプリケーションの開発が可能となる。開発に当たっては、特に以下の点に留意した。

(1) 操作時のレスポンスを極力早くすること

見栄えが多少犠牲になっても、マウス操作に対するレスポンスを優先させて、シンプルな画面設計を行った。また、HyperCard に表名や項目名等のリストをもたせ、ここで条件設定を行うようにし、EWSとの通信回数を極力少なくした。

(2) 表や項目の名称を日本語で表示させること

HyperCard にもたせた表や項目のリストに、その名称や内容を示す日本語を付記して、表や項目を選択する際に日本語が表示されるようにした。

(3) SQLを意識させないこと

表結合を自動的に行うように設計し、ユーザに表結合を意識させないようにした。

また、車両Noや事故No等で常にソートしてから出力するように設計を行うなど、ユーザ側から見えるSQLの関数や演算子は最低限のものだけに限定した。

(4) クロス集計ができること

検索結果のグループ集計のほかに、ユーザが縦列、横列を選択するだけで簡単にクロス集計が行えるようにした。

(5) 対話形式で操作が進められるようにすること

操作を促すメッセージや処理内容を示すメッセージ等を必要に応じて表示させ、間違いない操作を行えるようにした。

(6) 検索結果のファイル保存ができること

検索結果をファイル保存し、ワープロや表計算などのアプリケーションに容易に引き渡せるようにした。

(7) 画面を大きくすること

オリジナルのHyperCard は9インチの画面に合わせて設計されており、多数の表を表示

させる場合には表示が多少見づらくなるため、13インチの画面に合うように変更した。

開発したソフトのフローチャートは図5-1のようになっている。また、操作の対象になる画面は、初期画面、対象年度選択画面、検索項目選択画面、条件設定画面、検索結果表示画面、クロス集計表示画面の6枚で構成されており、更に各画面には、ボタン、サブフィールドが設定され、入力を促すようになっている。画面の例を図5-2に示す。

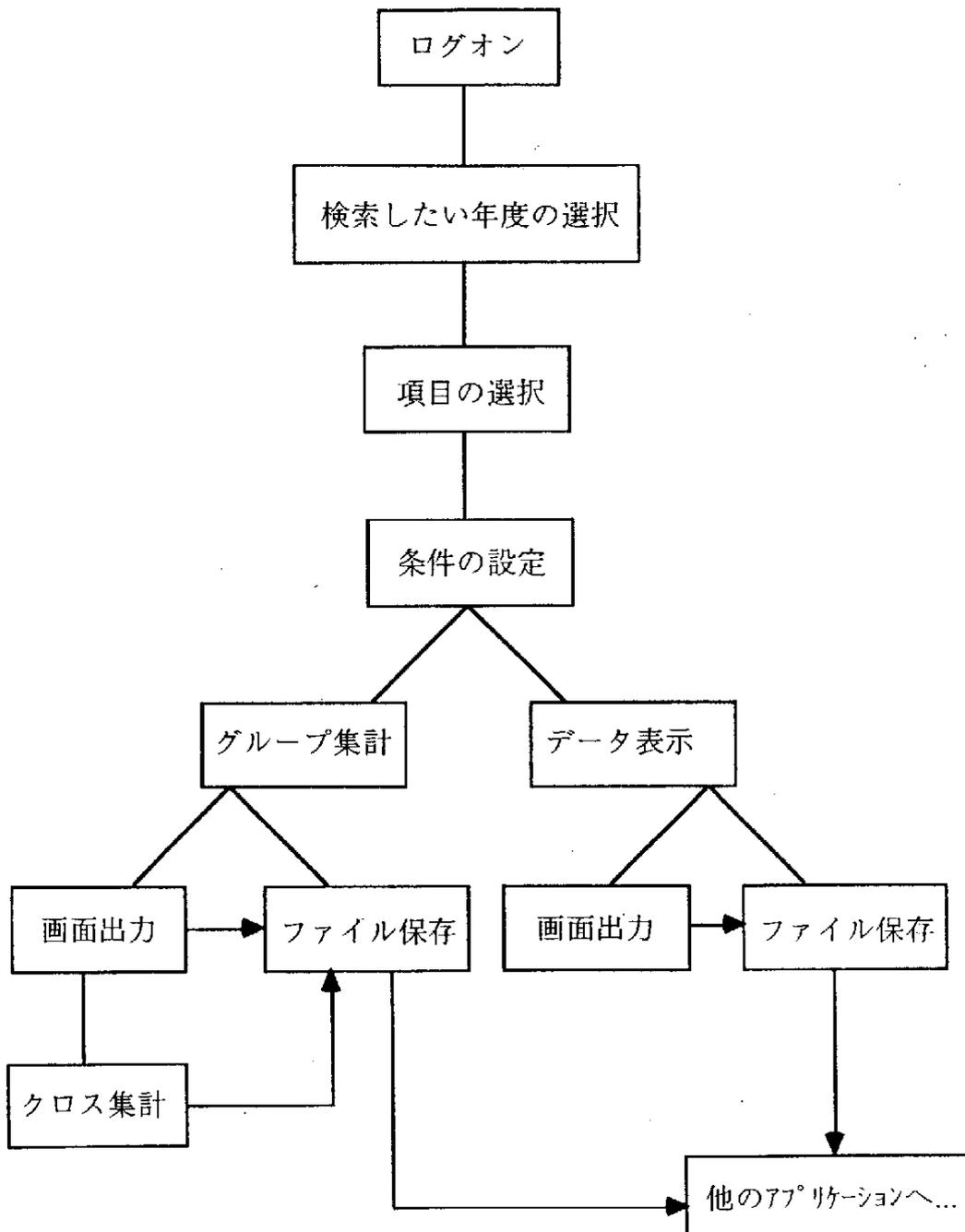


図5-1 開発した検索用ソフトのフローチャート

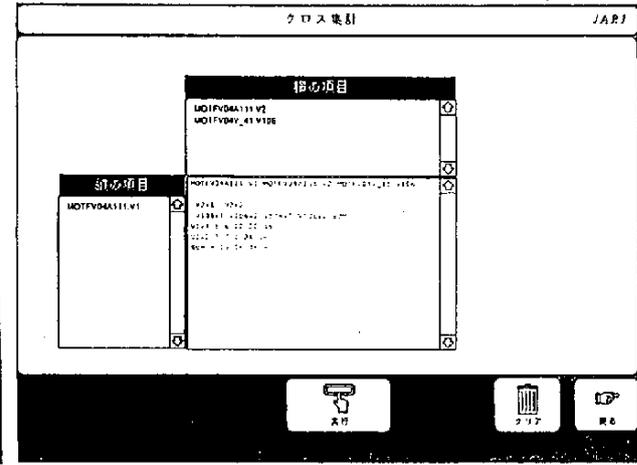
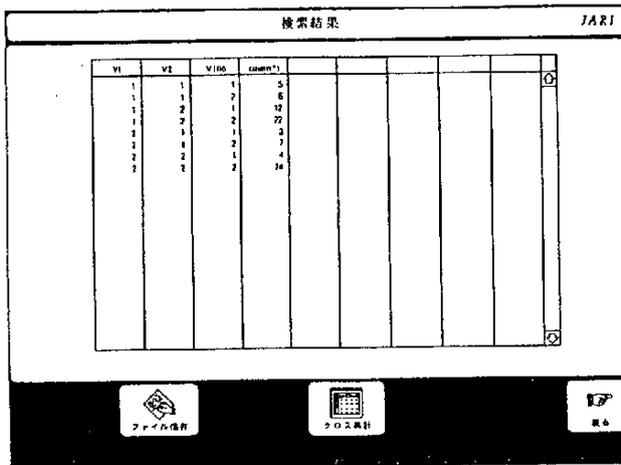
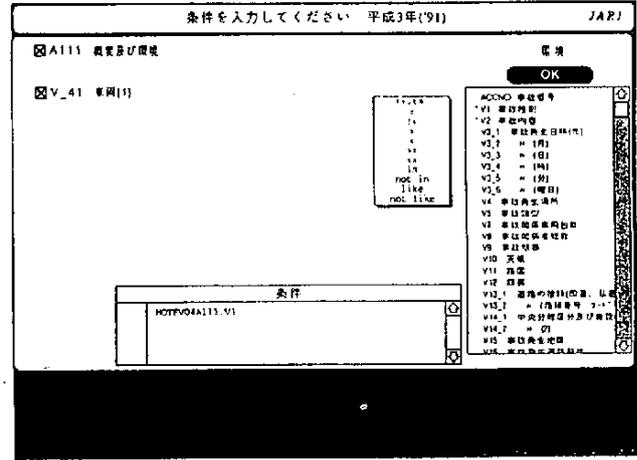
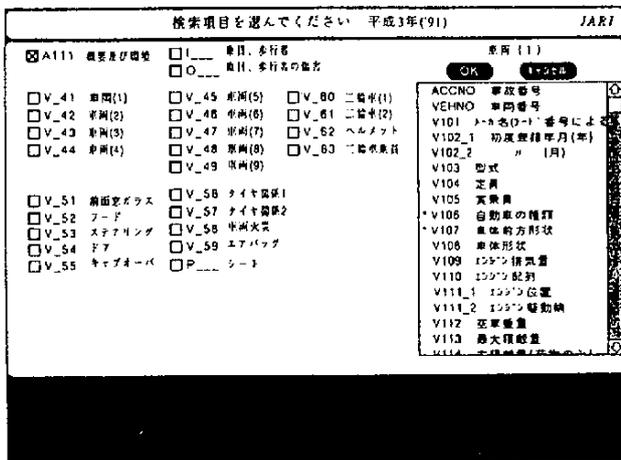
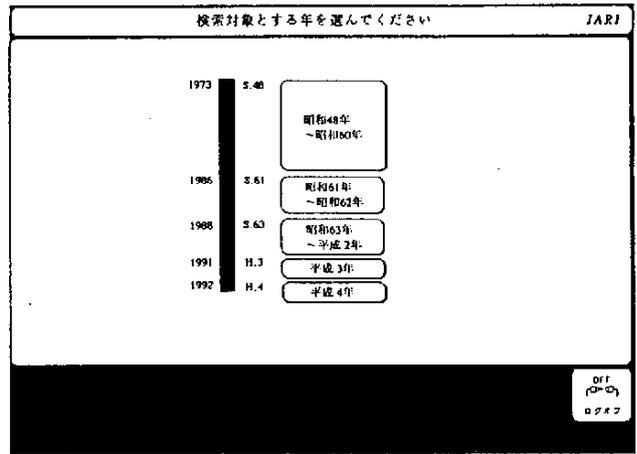
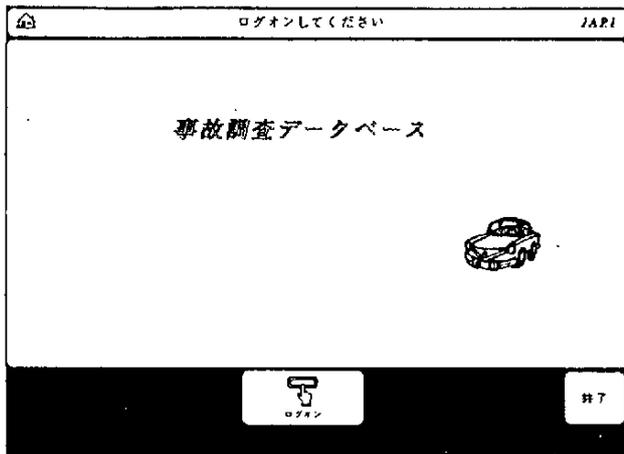


図 5 - 2 開発した検索用ソフトの表示画面の例

ただし、本データベースの理想像は画像データ等を含むマルチメディアデータベースであり、将来にわたってこれを指向して行く予定であるが、当面の課題としては以下のようなものがある。

(1) マニュアルの整備と検索ソフトの改良

本データベースを利用して目的とするデータを的確に検索するには、本データベースにどのようなデータが収録されているのかについて、あらかじめ予備知識を得ておく必要がある。本データベースの利用者は当面、当研究所の事故調査の関係者及び事故分析を担当する研究者であり、このような配慮はあまり必要ないものと考えているが、最終的には、検索したい要求に対して適切なアドバイスを与えるような検索支援スタックを、HyperCard 上に用意したいと考えている。また、今回開発したシステムに多少の改良を加えることにより、どの項目がどの位使われたか、というようなデータベースの統計情報がとれるようになるため、将来の調査項目の選定や課金情報などに利用したいと考えている。

(2) 画像のデータベース化

事故調査データのうち事故車両や事故現場の写真も重要な情報源であり、これらについてもデータベース化を図る必要がある。これら画像の処理やデータベース化については現在技術開発が急ピッチで進められている最中であり、かなり流動的な段階であるが、現時点では、図5-4に示すような、フォトCDを入力手段として、Macintosh に画像データベースを構築する方法が実現の可能性が高い妥当な方法であると思われる、できるだけ早期に試行したいと考えている。

(3) 交通事故再現シミュレーションソフトとの結合

当研究所では、事故調査や分析を支援する手段の一つとして、衝突前後の車の挙動やタイヤの軌跡を推定できるシミュレーションソフトを独自に開発している。このソフトは大型コンピュータで開発されているが、近々にワークステーションやMacintosh に移植する予定であり、その時点で本データベースとの結合を図りたいと考える。

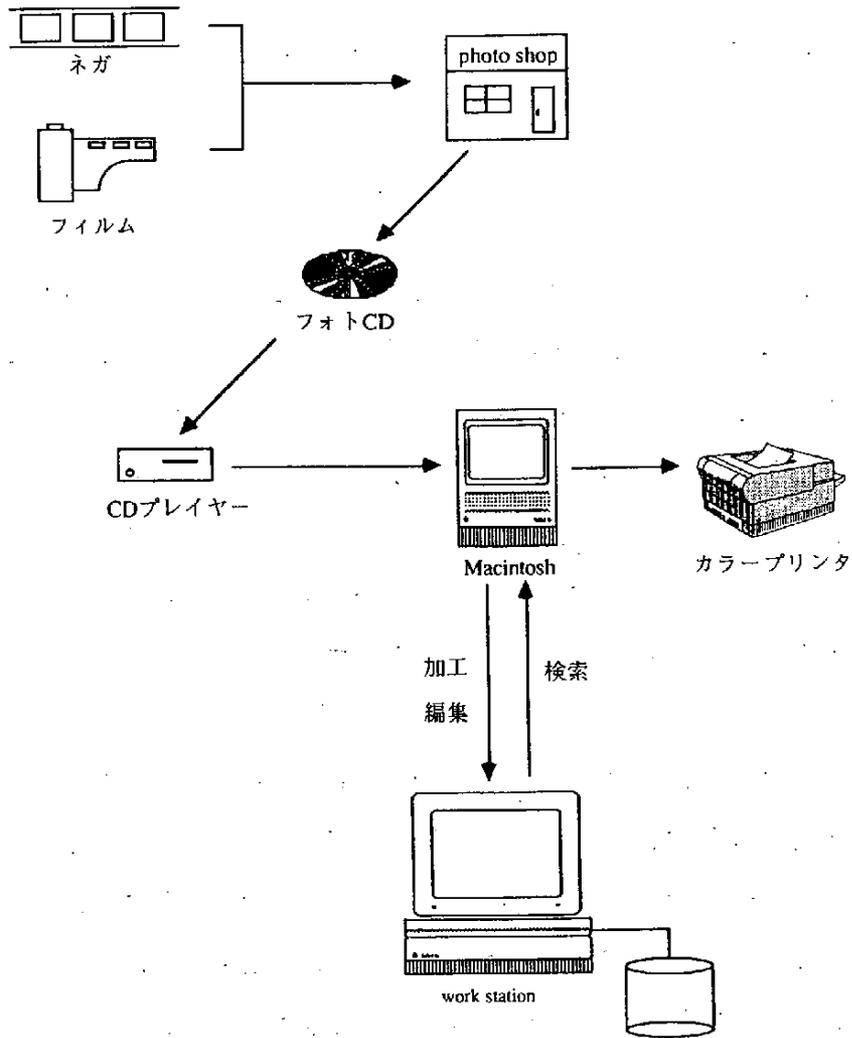


図5-4 画像データベースの基本構想

6 楽器データベースの構築

株式会社 ダイソメディアサービス



6. 楽器データベースの構築

6.1 目的

この数十年の間に、われわれが接する音楽は地域的及び時代的に拡大する傾向を見せてきた。民族音楽学の調査、研究の進展とともに世界各地の民族音楽が紹介され、一般にも民族音楽のCDやビデオテープ等が数多く流通し、それを専門家ではない人達までもが聞いて楽しむようになった。ポピュラー音楽の世界では昔から民族音楽の要素を取り入れたジャンルを次々と生み出してきたが、1980年代からは民族音楽をいっそう積極的に取り入れたポップスの総称である「ワールドミュージック」が隆盛し今日に至っている。

時代的な拡大の傾向も顕著である。西欧芸術音楽においては1960年代後半から中世、ルネサンス、バロック等の音楽を当時の楽器と演奏習慣をふまえた上で生きた音楽演奏として再現しようとする動きが起こり、専門家だけでなく一般にもこうした音楽の愛好者が非常に増加した。また、これ以降、古い時代の音楽の演奏においては、当時のオリジナル楽器を用いるのが一般的になり、古楽器の復興が盛んに行われるようになった。これと似た動きは日本の雅楽においても見られ、古代の楽器を復元することによって、長らく演奏されることのなかった古典を再現するという試みが行われている。

研究され、演奏され、耳にする音楽が地域的、時代的に拡大するにつれて、おびただしい種類の民族楽器や古楽器、及びそれを扱った資料と接する機会が増えてきた。音楽や文化人類学の研究者、学生はもちろん、音楽資料や楽器に関する資料を扱う図書館及び博物館、教師、マスコミ関係者等々、更には一般の音楽愛好者にとっても、楽器に関する情報の必要性は高くなってきた。

ところが、楽器の情報には混乱が生じやすく、この分野特有の入り組み方をしている場合が多い。同一ないしは同様の楽器が民族や国、時代によってさまざまな名称で呼ばれ、逆に、同一の名称が複数の異なる楽器に用いられたり、時代によって別の楽器を指すことも少なくないからである。別称の多いことも混乱を助長する。また、民族楽器の場合には演奏者がそのときの条件や目的に合わせて楽器を作ることが多いため個体差が大きく、楽器の類別や定義に混乱が生じることも珍しくない。

楽器データベースはそうした状況と事情をふまえ、さまざまな楽器に関する情報を整理し、それをできるだけ調べやすく、分かりやすい形で提供することを目的として計画された。そのために、次の二つの基本方針を設定しデータベースの構築作業を行うこととした。

(1) 楽器に関する総合的な情報ではなく、簡潔でその場で役立つ情報を次の2通りの手段によ

って提供する。

- ・楽器の特徴、イメージが手短かにつかめる解説
- ・専門家及び図書館用の楽器分類番号

(2) 更に詳しい情報のありかを調べるための二次資料としての機能をもたせる。そのために、まず出典別の楽器情報を収集し、それに基づいてデータを作成し、そこに出典と記載ページを表示する。

6.2 実施概要

6.2.1 出典別楽器データの入力

このデータベースの目的の一つとして楽器に関する二次情報、つまりある楽器の情報が記載されている資料についての情報を提供することがある。そのため、まず最初に楽器についての基礎資料、即ち下記に示す、現在市販されている日本語の文献と英語の代表的な文献、及び楽器分類表を対象に、出典別楽器データの収集・入力作業を行った。

◆ 文献

- ① 大阪音楽大学附属楽器博物館目録 豊中：大阪音楽大学，1984
- ② 音楽大事典 東京：平凡社，1981-1983
- ③ 楽器資料集 IV～X 東京：国立音楽大学楽器学資料館，1984-1992
- ④ 楽器の博物誌 東京：朝日新聞社，1986
- ⑤ 黒沢隆朝著 図解世界楽器大事典 東京：雄山閣，1972
- ⑥ 郡司すみ著 世界楽器入門 東京：朝日新聞社，1989
- ⑦ 佐竹悦子編 中南米の楽器 (非売品)
- ⑧ 柴田南雄著 楽器への招待 東京：新潮社，1983
- ⑨ シルク・ロードの楽器と芸能具展 楽器 東京：国立劇場，1989
- ⑩ ダイアグラムグループ編 楽器 東京：マール社，1992
- ⑪ 東京芸術大学音楽学部小泉文夫記念資料室所蔵楽器目録 東京：藝術研究振興財団，1987
- ⑫ 日本音楽大事典 東京：平凡社，1989
- ⑬ マンロウ著 柿木吾郎訳 中世：ルネサンスの楽器 東京：音楽之友社，1979
- ⑭ 民族音楽大集成 解説書 東京：キングレコード，1981
- ⑮ ソケティア著 龍村あや子訳 アフリカの音楽 東京：晶文社，1989
- ⑯ Collection of Musical Instruments. Tokyo: Kunitachi College of Music

Research Institute, 1986

⑰ Marcuse, Sibyl. Musical Instruments: A Comprehensive Dictionary. New York: W. W. Norton, 1975

⑱ New Grove Dictionary of Music and Musicians, Appendix A. London: Macmillan, 1980

⑲ New Grove Dictionary of Musical Instruments. London: Macmillan, 1984

◆ 楽器分類表

(a) ホルンボステル, ザックス著 岸辺成雄訳・増補 楽器分類表 (平凡社 音楽大事典 第2巻 「楽器学」より)

(b) AV/MARC 分類表 東京: ダイソメディアサービス, 1990 (非売品)

作業は, 各々の資料を調査しそこに掲載されている楽器名とそれに関する情報を抽出し入力するという方法により行われた。二次情報の基礎データとして用いる情報は, 原典にあるそのままの形で収集されなければならない。その点及び作業の円滑さを考慮し, 第一段階としての出典別楽器データは各々の資料の特性に合わせた形式で入力した。

6.2.2 出典別楽器データの総合化

(1) 出典別データの編集

次に, 出典別楽器データを総合化するための準備作業として, 各出典のデータを共通のデータベースソフトに読み込むために次のようにフィールド (項目) 分けを行った。

- ・ 楽器名
- ・ 異称
- ・ HS項目, AVMC項目 (いずれも楽器分類表の場合)
- ・ 楽器のタイプ・属を示す名称
- ・ 民族・地域・時代
- ・ 用途
- ・ 備考
- ・ 記載箇所 (番号/頁)
- ・ HS番号, AVMC番号 (いずれも楽器分類表の場合)
- ・ 参照楽器名

フィールド分けにおいて情報が長すぎる場合には, 略語の使用, 及び数詞のアラビア数字化, 重要性の低い部分の削除等の手段により, 各項目をフィールドの制限文字数の範囲内に縮小させた。タイプ名については, 同類の楽器をまとめて配列する目的で使用するため, 修

飾語が前にある場合にはその多くを倒置形に変更した。(例: small kettledrum→
kettledrum, small)

また、この段階で次の総合化用フィールドを追加した。

- ◇ レコードの配列のために用いるフィールド
(カナsort, アルファベットsort, 統一名sort)
- ◇ レコードの特定化のためのコード (アイテムcode)
- ◇ 情報の出典を示すフィールド (コレクション名)
- ◇ 楽器分類番号 (HS番号, AVMC番号)

※基礎カード出力までの過程においては、整理用データベースのファイル区分、及び同類の楽器のレコードをまとめるために使用された(ただし、この段階では大まかな観点で付与された)

(2) 整理用データベース(総合化ファイル)の作成

29フィールドのデータベースを設定し、前項の編集が完了した出典別楽器データを順次読み込み、最終的に、33,865レコードからなる整理用データベースを作成した。

6.2.3 登録用テキストファイルの作成

(1) 整理用データベースのデータの整理

楽器データベースのデータ作成のよりどころとする基礎カードを出力するために、出典別データの総合化によって構築した整理用データベースのデータの整理・編集を以下の手順で行った。

- (a) 既に大まかに付与されている楽器分類番号(「HS番号」と「AVMC番号」)を、同類の楽器を集めることが可能な程度のけた数まで精度を高めた。
- (b) これから後の作業には不要なフィールドを削除し、フィールド数を20に絞り込んだ。
- (c) 基礎カードによるデータ作成作業を容易にするために、(a)で付与した楽器分類番号を用いておおむね近い分類番号をもつ楽器をまとめ、同類の楽器のレコードをグルーピングして抽出した。この段階で、データベースのデータ作成には不要なレコードは抽出から除外した。
- (d) レコードのデータをCSV形式から改行区切り形式に変換し、各フィールドの先頭にそのフィールド名を付加した。

(2) 基礎データのカード出力

上記の整理・編集がなされたデータから基礎カード約33,500枚をプリントアウトし、配列した。

(3) データの編集・作成

基礎カードをもとにして、直接パソコンに入力する方法でデータベースのデータを作成した。基礎カードの出典名は「記載資料」として記録され、事典以外の場合にはその記載箇所（記載ページ）も付記された。複数の基礎カードが1データに集約されることが通例であるが、その場合には1データに出典名が複数列記された。出典別楽器データの対象としなかった資料を参考として用いた場合には、それも記載資料として記録した。

1 データを作成する単位については次の基準に従った。

◆ タイプを表す名称／総称的な名称

◇ 通常は楽器群の単位の名称に対して作成する。

◆ 個別楽器名

◇ 通常、同一種類の楽器の単位で作成される。

・同一種類の楽器に対する別称等は共通のデータに表示し、それに対して独立したデータは作成しない。

◇ 同一種類の楽器であるが別の側面や性格、背景をもつものが独特の名称で呼ばれる場合には、それに対して独立したデータを作成する。

「解説」の作成に当たっては、このデータベースの意図する目的から、できる限り1～2行でその楽器のイメージがわくような説明を心がけた。類似異種の楽器がある場合には相違が分かるように、より詳しい解説を行った。「解説」の作成においては、基礎カードとともに下記の資料を、この優先順序で参考として使用した。

① New Grove Dictionary of Musical Instruments. London: Macmillan, 1984

② Marcuse, Sibyl. Musical Instruments: A Comprehensive Dictionary. New York: W. W. Norton, 1975

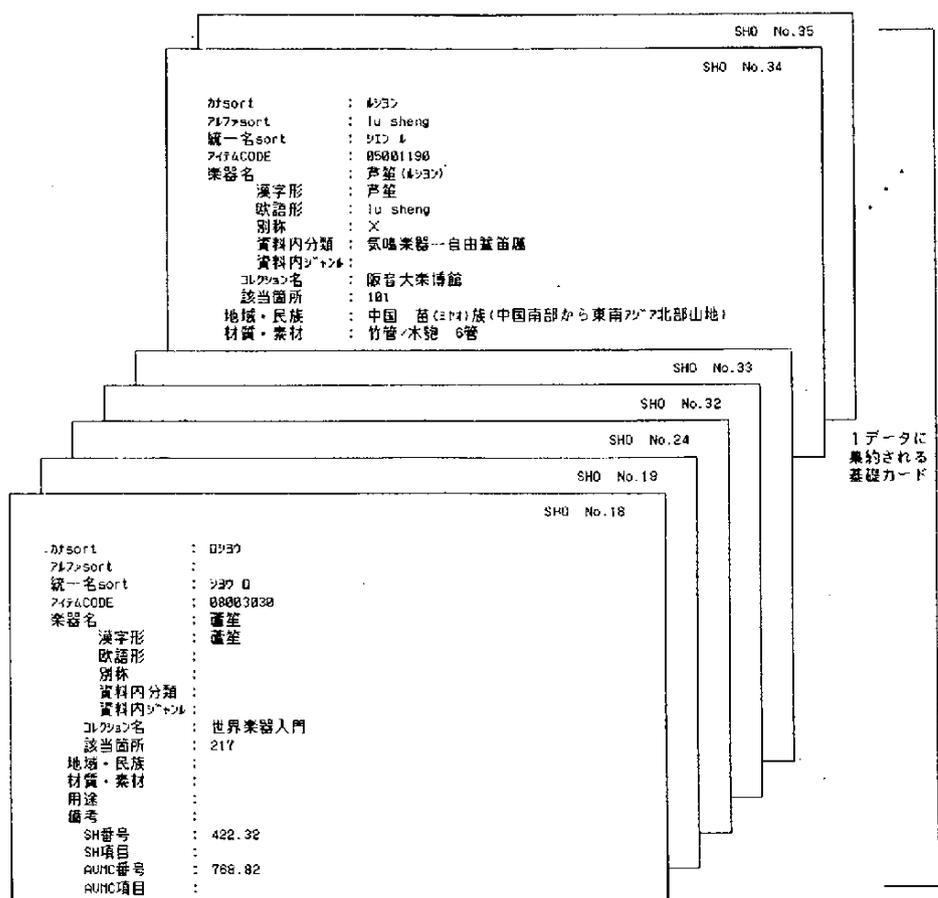
③ その他の参考資料

短い「解説」によってその楽器を表現できる場合には、「地域・民族・時代」や「材質・素材」、「用途」の項目に関する内容も「解説」の中を含めた。逆に、「解説」が長めになる場合には、その他の項目名も用いて記録した。

多数の基礎カードを集約して作成されたデータの実例を図6-1に示す。

今回の委託事業の期間内には、約3,500枚の基礎カードを処理し、約1,000件のデータを作成した。当初は、非統制楽器名データを独立したデータとして登録する予定だったが、検索機能への考慮からそれを本データに別称として表示し、そのデータの索引として登録する方法に変更したので、データの登録件数は予定よりも少なくなった。しかし、処理した基礎データの件数は計画どおりである。

作成されたデータは文書型データベースソフトに読み込むために、テキストファイルの形式で登録された。



上の基礎カードを集約して作成された楽器データベースのデータ

楽器名	: 蘆笙/芦笙 (Dyak(日本語)/Lu-sheng lu(02)sha(05)ing/lu sheng/lu-sheng(中国語・漢族))
別称	: 蘆笙/芦笙 アシカウ(日本語)/Lu-sheng(中国語); kaeng ケン(タイ語); 樓 gaena-gena(ミナ族/トン族/ス(族)); 哩斗 gadou(ミナ族); 哩杰 gajie(ミナ族)
解説	: 東南アジア、中国南西部(ミナ族/ハ族/ハ族/ツ族/族)で用いられるマツ・アカウ。細長い木製の気室に竹管が差し込まれている。マツ(1) → 葫蘆笙 → 葫蘆笙を参照
材質・素材	: 竹管(通常6) / 木製の気室 / 金属製のリット
用途	: 独奏 / 合奏・舞踏の伴奏。野外での芦笙舞(種々の祭で芦笙を吹きながら踊る)、若い男女の交流、葬儀などで重要な役割を果たす
HS番号	: 412.132
AUNC番号	: 768.82
LC番号	:
【記載資料】	
大阪音楽大学附属楽器博物館目録	
音楽大辞典 平凡社	
楽器資料集 VI 国立音楽大学楽器学資料館 998	
楽器の博物誌 朝日新聞社 49-51, 88, 127	
都司すみ 世界楽器入門 217	
シカゴの楽器と芸能員展 楽器 94	
中国少数民族楽器志 中央民族学院少数民族文学芸術研究所 130	
民族音楽大集成 解説書 “シカゴ” 133	
Collection of Musical Instruments. Kunitachi College of Music Research Institute 93, 94	
New Grove Dictionary of Music and Musicians. Appendix A	
New Grove Dictionary of Musical Instruments	

図6-1 基礎カードをもとにした楽器データベースのデータの作成

6.2.4 楽器データベースへの登録

作成されたテキストファイルを文書データベースソフトに読み込み、更に表題及び索引(キーワード)の登録を行った上で、最終的に楽器データベースのデータとして登録した。

6.2.5 画像資料の試験登録

今回の構築作業においては、試験的に下記の楽器に関する画像資料の登録を各1点ずつ行った。

- ◇ 楽器の写真
- ◇ 楽器を演奏している写真
- ◇ 歴史的な文献に見られる楽器の図

6.3 データの仕様

6.3.1 データの実例及び各部分の名称

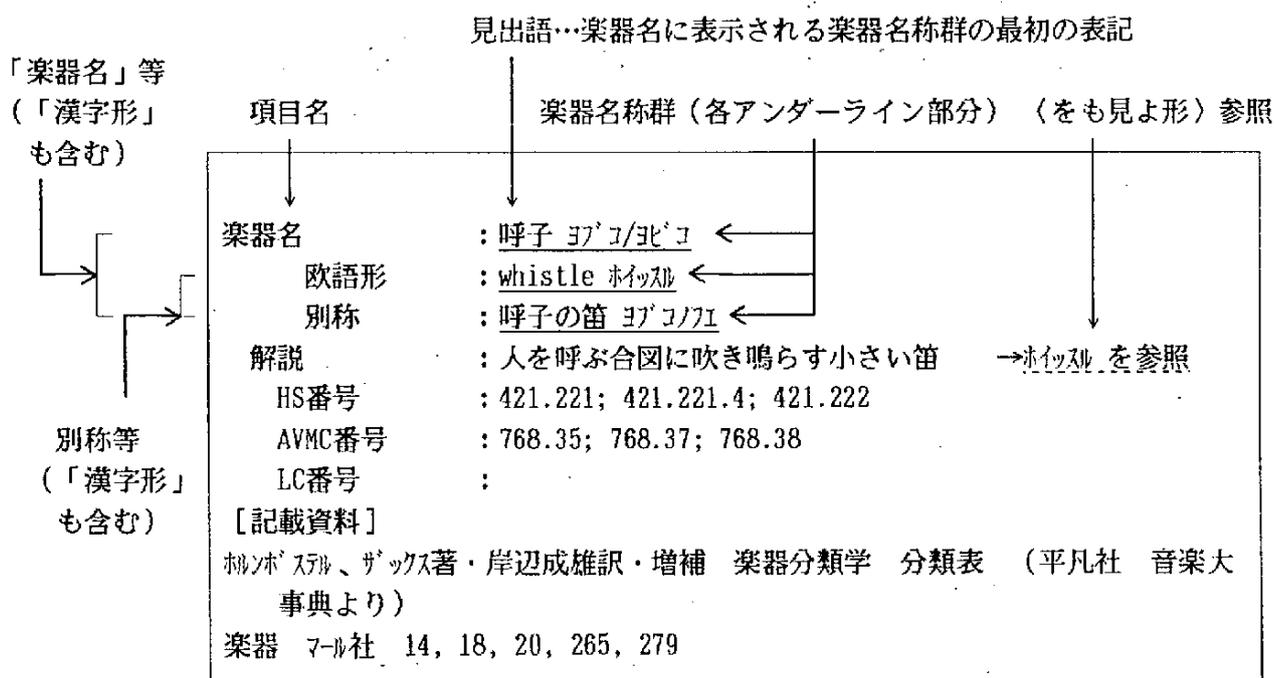


図6-2 データの部分の名称

6.3.2 データ項目

楽器データベースのデータは次の項目から構成されている。ただし、各データにおいて必ずしもすべての項目が入力されるわけではない。

- ◇ 「楽器名」等
 楽器名, 漢字形, 欧語方, 別称
- ◇ 「解説」等

解説, 地域・民族・時代, 材質・素材, 用途, 備考

◇ 楽器分類番号

HS 番号, AVMC 番号

◇ 記載資料

記載資料

6.3.3 他の楽器名称との関係の表示

下記①②の楽器データベース構築の目的及び条件による制約から, このデータベースにおいては, 楽器名称の等価関係及び連想関係, 階層関係の厳密な整理は行われていない。

① このデータベースが目的としているのは, 概念ではなく, 実際の物(楽器)と呼び名の整理である。

② 多数の言語にわたる名称を一つのデータベースにおいて取り扱う必要がある。

(概念の整理をするシソーラスは, 基本的に一つの言語の範囲で作成される。)

したがって, 各データにおいて他の楽器名称との関係を可能な範囲で表示されているが, それはシソーラスのような厳密性をもたない。

6.3.4 データの表題及び索引の登録

データの表題と索引(キーワード)については以下の基準に従って登録された。

表題	「楽器名」に表示されている楽器名称群全体
索引	①「楽器名」等に表示されているすべての楽器名称の表記
	②〈をも見よ〉参照に表示されている見出語
	③「HS 番号」, 「AVMC 番号」に表示されている個々の分類番号

(例) 6.3.1のデータの表題及び索引登録の実例

表題	呼子 ヨコ/ヨビコ
索引	① 呼子/ヨコ*/ヨビコ/ホイッスル/whistle /呼子の笛/ヨコノフエ
	② ホイッスル [ただし①に同一のものがあるので登録を省略する]
	③ 421.221*/421.221.4/421.222/768.35/768.37/768.38

(注) * の索引は他の索引の前方一致で検索されるので索引入力したとしても登録の際に削除される。

6.4 システム構成

(1) 機器構成

- ◇ パーソナルコンピュータ NEC PC-9801 RA
- ◇ ハードディスク
- ◇ ディスプレイ
- ◇ プリンタ

(2) ソフトウェア構成

- ◇ 基本ソフトウェア……………MS-DOS¹⁾ Ver. 3.30
- ◇ 日本語入力システム……………ATOK7²⁾
- ◇ データベースソフトウェア……………知子の情報³⁾ Ver. 4

6.5 データベースの機能

6.5.1 使用するデータベースソフト

楽器データベースの構築を計画するに当たって、当面は市販のデータベースソフトを使用し、パソコン単体で運用できるデータベースを構築するという基本方針を設定し、ソフトの選定を行った。その結果、次の理由からこのデータベースには文書データベースソフト「知子の情報」³⁾を用いることとした。

- ◇ 検索機能が充実している。
- ◇ データの情報量の差に柔軟に対応できる。
- ◇ MS-DOSの「ユーザ定義文字」による全角外字を使用できる。
- ◇ 扱えるデータ数と容量が楽器データベースの構築に適している。
- ◇ MS-DOSのテキストファイルの読み込みが容易である。
- ◇ 値段が安価であり、データベースを普及させやすい。

楽器データベースにおいては、文書データベースソフト「知子の情報」をそのままの形で用いる。したがって、以下に説明するデータベースの機能は、すべてこのソフトが本来有している機能の一部を楽器データベースに適用したものである。

¹⁾ MS-DOSは米国マイクロソフト社の登録商標である。

²⁾ ATOK7は株式会社ジャストシステムの登録商標である。

³⁾ 知子の情報は有限会社テグレット技術開発の登録商標である。

6.5.2 検索機能

(1) 検索の種類

検索の種類と機能を表6-1に示す。

表6-1 検索の種類と機能

検索の種類	検索対象			検索方法	半角大/小文字の区別を するかどうかの切り換え	拡張検索
	索引	表題	本文			
高速検索	○			前方一致		
普通検索	○	○		任意一致	可能	可能
全文検索	○	○	○	任意一致	可能	可能だが時間がかかる

(2) 検索対象データの指定方法

検索対象データの指定方法を図6-3に示す。

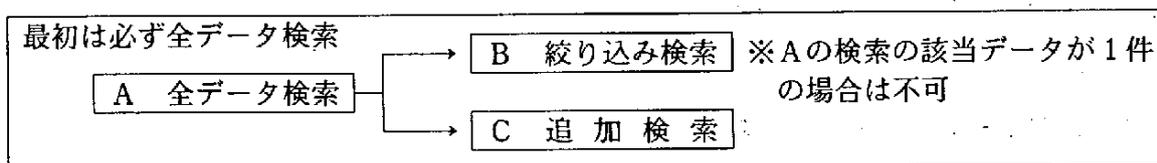


図6-3 検索対象データの指定方法

(3) 論理演算

論理演算機能を表6-2に示す。優先索引の指定をすると○が、否定索引を指定すると×が検索索引の行の左端に表示される。

表6-2 優先索引、非優先索引、否定索引

指定の種類	記号	指定方法	説明
①優先索引	○	[f・2] / [TAB]	◇必ず含まれていなければならない索引。 ★これを指定すると検索が早くなる。
②非優先索引	なし		◇その中の一部が含まれていればよいという索引。 ・一つの場合→優先索引と同等になる ・二つ以上の場合→検索開始直前に「非優先索引との一致数」(そのうち幾つかが含まれていればよい)を指定
③否定索引	×	[f・3] / [SHIFT] + [TAB]	◇その索引が含まれていないデータを検索するという指定。 ★これを指定すると検索が遅くなる。①と一緒に使うと検索時間が短くなる。(ただし、①と③では③の方が検索時に優先される。)

6.5.3 登録機能

「知子の情報」は直接データを入力して登録することもできるが、楽器データベースの構築においては通常、別に入力したデータをテキストファイル形式で保存し、それを「知子の情報」に読み込んで登録するという方法を用いた。

6.6 データベースの評価と今後の課題

6.6.1 データベースの評価

今回の作業においては、ほぼ計画に沿ったデータベースを構築することができたと思われる。ただし、普及を図れる程度のデータ件数にできるだけ早期に到達するために、できるだけ計画よりも多くのデータ作成・登録を行うことをめざしていたが、準備作業としての出典別楽器データの収集・入力の数、予定をはるかに上回ったため、それに予定以上の労力を注ぐことができなかった。

当初は、非統制楽器名データを独立したデータとして登録する予定だったが、検索機能への考慮からそれを本データに別称として表示し、そのデータの索引として登録する方法に変更したので、データの登録件数は予定よりも少なくなった。しかし、処理した基礎データの件数は計画どおりである。

内容及び特性、機能に関しては、楽器データベースは次を目標として構築された。

- ① 現在一般に使用されている楽器だけでなく、民族楽器、古楽器、電子楽器を含むあらゆる楽器を対象とする。
- ② 非統制楽器名データにより、日本で知られている自然語の楽器名のあらゆる形からの検索を可能とする。
- ③ 基本的にシソーラスの形を念頭に作成し、楽器名の階層的な関係を知ることができるようにする。
- ④ 研究者用の分類番号と図書館用の分類番号の両方を付与し、種々の分野で活用できるものとする。

構築したデータベースは、上記の目標をおおむね実現していると思われる。ただし、③の楽器名の階層関係については十分な形で表現することができなかった。即ち、英語の文献にのみある楽器名称が予想以上に多かったという事情から、すべて日本語表記にする予定だった見出語を多言語にせざるを得ず、技術的にシソーラスの形のまとめ方が難しくなったからである。そこで、結果的にシソーラスを念頭においた辞書形式に落ち着いた。

6.6.2 今後の課題

(1) 早急に取り組むべき課題

a) 基本データ二次分のテキストの追加

b) オフライン形態のデータの普及

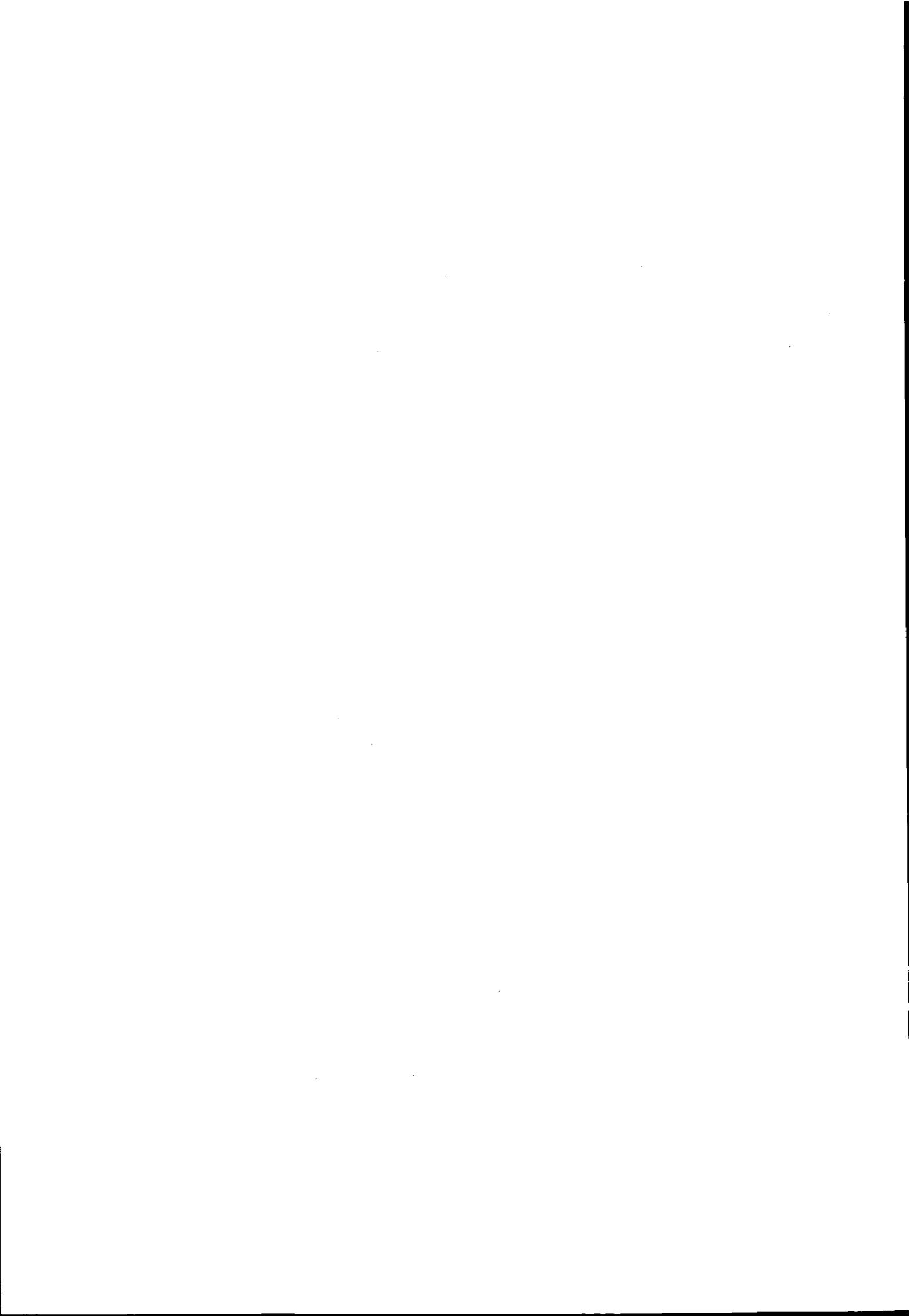
(2) 長期的な展望による課題

a) ビジュアルな部分の強化

b) オンライン化, CD-ROM化, マルチメディア化

7 人体計測データベースの構築

社団法人 人間生活工学研究センター



7. 人体計測データベースの構築

7.1 総論

7.1.1 緒言

従来の技術開発は、生産体制の合理化・高効率化を目的としたため、人間工学的視点が欠如した立場で生産が進められてきた。しかし現在、経済を始めとする情勢が変化し「量より質」「ゆとりと豊かさ」が求められる時代に入り、製品設計には使用する人間の特性に基づいた製品設計が必要となりつつある。しかし、これらの人間特性データの蓄積は極めて少なく、蓄積形態も利用を検討する際古いデータであったり、汎用性がなかったりすることが多い。

当センターは、人間特性の中でまず形態特性に関するデータベース作成を現在行っている。これらを各企業に提供することで、安全性・快適性・効率性・健康性を指向する製品を実現できると考える。

7.1.2 本データベースの構成

(1) 採寸データベース

採寸データベースは本データベース群の中核をなし、採寸データベースには1人につき178項目の人体の特徴的な部位間の距離情報であるサイズデータ及び計測地方、計測時の年齢、性別等の各個人の属性が数万人の規模で収められる。データベースの構造、機能は7.3.3で述べる。

(2) 統計データベース

統計データベースは、採寸データベースをもとに、各種の属性（年齢、性別等）で分けられたデータの統計内容を収めるデータベースである。178項目のデータすべてに対して年齢、性別等の属性による分類を行い、平均、標準偏差等の統計処理を施し、結果を統計データベースに収納する。データベース構造、機能を7.3.4で述べる。

(3) データベースシステム

図7-1にデータベースシステムの概略を示す。複数台のパソコンとホストコンピュータであるワークステーションを、イーサネット通信回路で連結した構成である。パソコンはワークステーションをアクセスする端末であると同時に、データを分散処理する使命を帯びている。

本データベース群が構成されるまでの過程を図7-2によって説明を加える。キャラバ

ンバスから当センターに送付されるデータは手計測データ、スライディングゲージデータ、三次元計測画像データの3種類である。手計測データはキーパンチャーを介してフロッピーに収録される。画像データとフロッピーに収められたスライディングゲージ及び、手計測データは一旦パソコンにて個人単位で編集される。その後画像データから特徴点間距離を算出するポインティング作業へと移行する。ポインティング作業により得られたサイズデータ及び、特徴点の座標は再度編集作業を通り光ディスクに収納される。

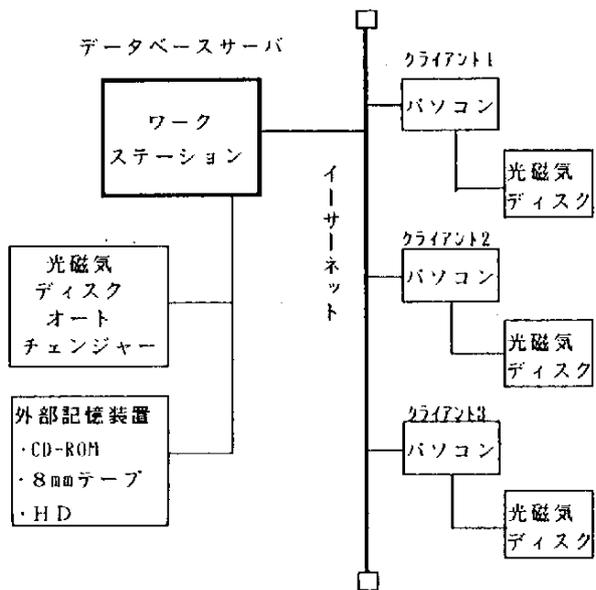


図7-1 データベースシステム

以上の作業を「パソコン前処理システム」と名付け7.2.1で簡単に説明する。前処理で得られた178項目の数値データはイーサネット回線によりパソコンからワークステーションに転送されデータベース化される。

7.1.3 人体計測項目と体制

本データベース構築のための人体計測は計測機器を登載した2台のキャラバンバスで行われている。バスは2年の歳月をかけて全国を巡回する。各地の通商産業検査所の協力を得て被計測者を集め、被服、人間工学等に関係する分野を専門とする先生方が計測を行う。被計測者は小学生以上のすべての年齢層の男女を対象にしている。以下にそれぞれの計測方法を述べる。

(1) 手計測

三次元計測器による計測が不向きな項目

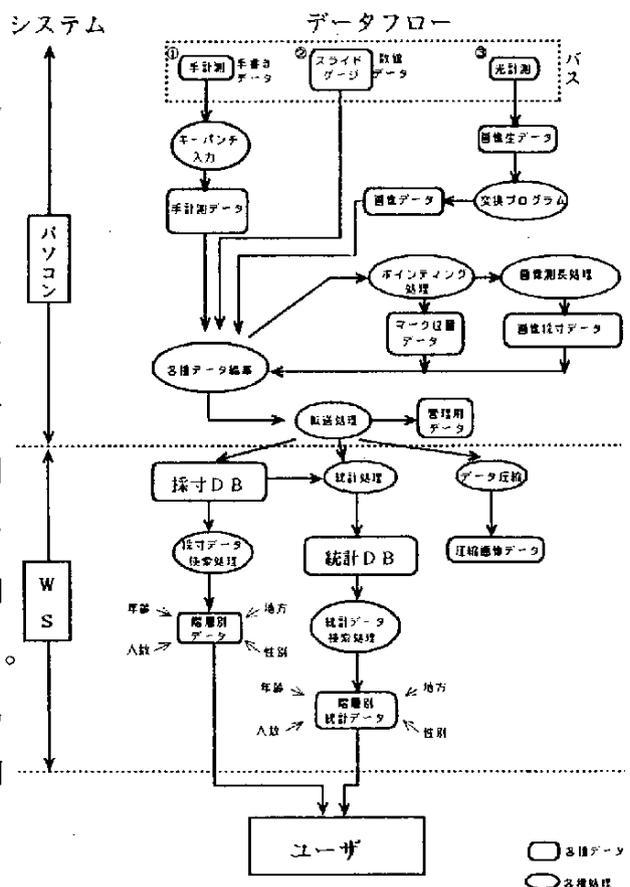


図7-2 データベース構築課程

の計測を7種類のマルチン計測器等を用いて直接計測を行っておりこれらを手計測と呼ぶ。
計測項目は65項目である。

(2) 三次元計測器

測定原理を図7-3に示す。照明・撮像系の構成は、よく知られている光切断法の構成と同じであり、測定対象の斜め上方からスリット光を測定対象の表面全体にわたって回転走査し、これを直上からテレビカメラで撮像する。実際の装置では測定精度を保証するためのパースペクティブ歪の補正を行ったり、測定の死角を少なくするために、測定対象を左右両側から走査し効領域同士を貼り合わせて最終的な形状データとしている。これらの装置を用いて、顔、手、体幹立位、体幹座位の形状計測を行っている。

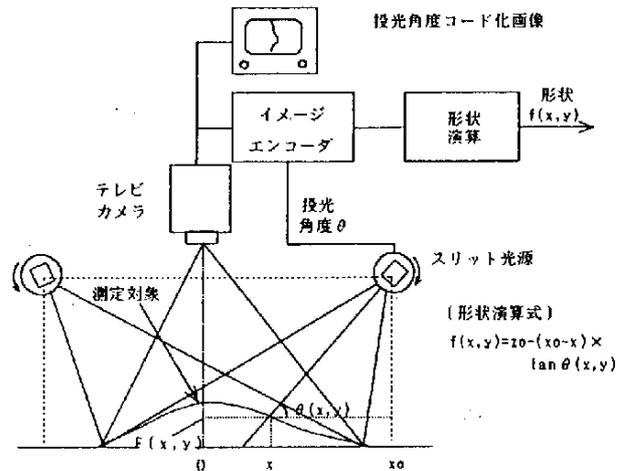


図7-3 三次元計測器測定原理

(3) スライディングスケール

スライディングスケールは、体幹を触診しなければ正確な計測ができない部位や、適当な圧力をかけて計る必要がある部位を、パソコンを介したリンクアームとプレートで接触させて計測する。計測項目は12項目である。

以上の3種類に加え、上記の計測項目より計算で求める15項目がある。

7.2 パソコンを用いたデータ前処理システム

7.2.1 前処理システム基本仕様

キャラバンパスから送られて来るデータは、ワークステーション上のデータベースに登録されるまでに、さまざまな加工処理を施す必要があるが、これらすべての処理をワークステーションで行うには負荷が高い。そこでパソコンでデータの前処理作業を行い、負荷の分散を図った。ここでは小さな単位を積み上げることで、パソコンで大きな単位を管理することを実現している。

本システムは大きく分けて、データ編集システムと画像計測システムの二つの機能をもっている。それぞれについて以下に述べる。

7.2.2 データ編集システム

計測データは1人につき5種類のデータがあるが、これらのデータはすべてバラバラのディスク（メディア）に記録されている。これらを個人ごとにまとめて、1枚のディスクに整理しデータの欠落を早期に発見する作業を、編集システムで行う。

図7-4に編集システムのデータフローを示す。編集システムではバーコード（IDナンバー）をキーにして、40人単位で、バラバラのデータを1枚の光磁気ディスクに整理している。光磁気ディスク1枚で160人分のデータが管理できる。

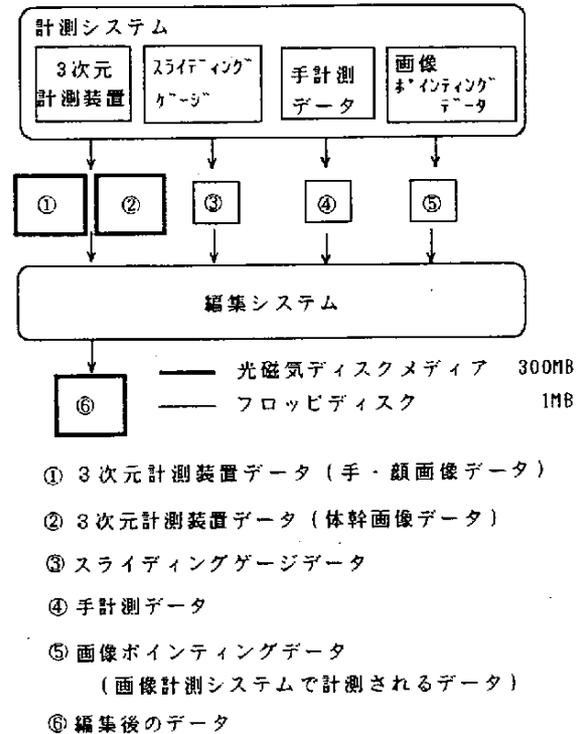


図7-4 編集作業データフロー

7.2.3 画像計測システム

画像データからパソコンで特徴点間距離を計測する作業を、画像計測システムで行う。

画像計測システムでは、ディスプレイに表示される画像に対して人体特徴点をメッセージに従ってポイントし、指定したポイントの高さやポイント間の長さを計測する。画像ポインティングデータは指定されたポイントのx y z座標の組み合わせから自動的に計算される。

108ポイントを指定して、84個の画像ポインティングデータを計測する。

7.2.4 パソコン前処理システムのまとめ

今回の人体計測で1人当たりの計測データ量は、約3MBとなる。したがって5万人のデータ総量は、約150GBとなる。WSでも150GBを管理するのは難しく、パソコンでは不可能であるが、一度に5万人のデータを管理せず、5万人のデータを分割して管理する方法を用いた。

7.3 採寸、統計データベースの構築

7.3.1 データベース基本仕様

データベース基本運用を図7-5に示す。

利用者のニーズに応えるためには、柔軟なデータベースの構築が必要とされるため、拡張

性のあるUNIXワークステーションをサーバーとし、操作性のよい小回りのきくパソコンをクライアントとして構成した。

ワークステーションが担う役割はパソコンから、ネットワークを通じ転送された数値データをもとに採寸データベースを構築し、採寸データに統計処理を施した結果を統計データベースに収納することである。画像データは元データのまま展開できるようなデータ量の圧縮を施し光磁気ディスクオートチェンジャーに保存する。パソコンとワークステーションとの役割分担を明確にし機能、コスト面等において完成度の高いシステムを構築したと考える。

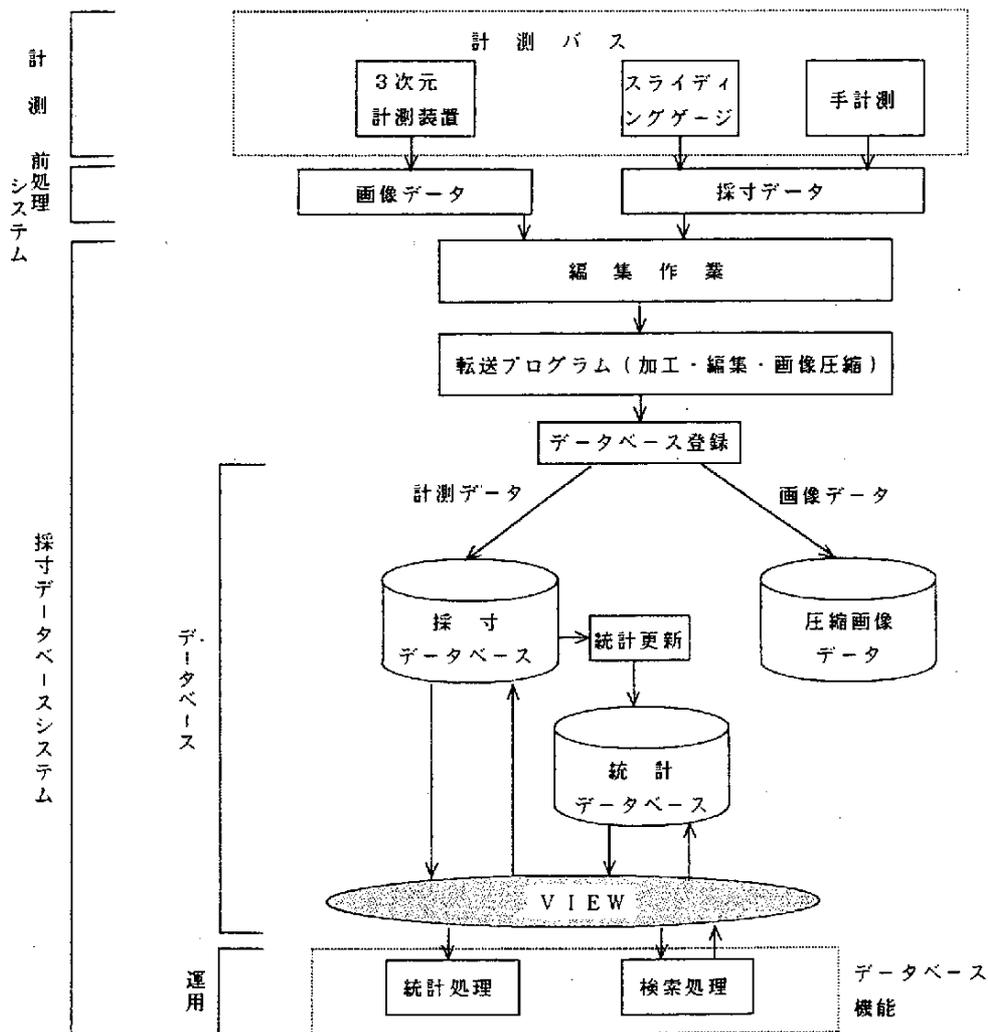


図7-5 データベース基本運用

7.3.2 データベースシステム概要

データベースソフトウェア概要を図7-6に示す。

本データベースシステムは、データベースエンジンをワークステーションに置き、ネット

ワーク上のパソコンからアプリケーションを走らせるといった本格的な分散処理環境でのリレーショナルデータベースである。

システム設計における各種基本デバイスとして以下を選定した。

- ① データベースサーバー機：AS4080-41
- ② ネットワーク：イーサーネット
- ③ データベースソフト：INFORMIX, WINGZ, ロータス123
- ④ クライアント：PC-98FA

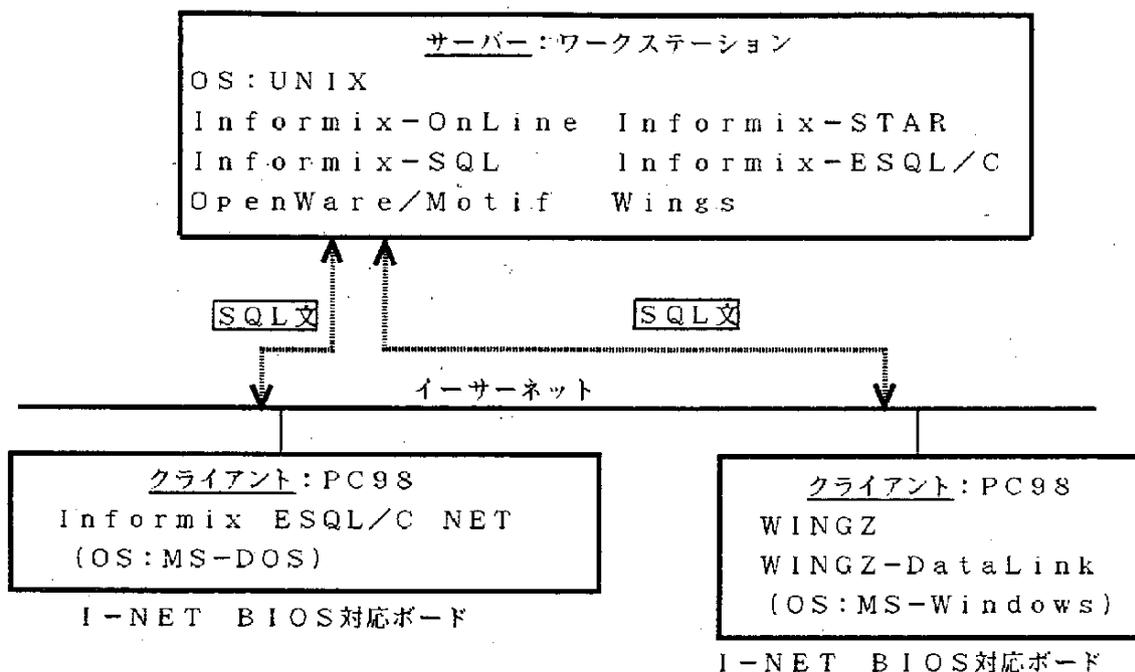


図7-6 データベースソフトウェア概要

7.3.3 採寸データベース

採寸データベースの機能を図7-7に示す。

採寸データベースは、被験者の属性情報、採寸データ、計測集団マスタの3種類のテーブルから構成されており、検索及び統計処理の基本テーブルとなる。それぞれのテーブルはID、日付、車番号等をキーにしてリレーショナルが構成される。

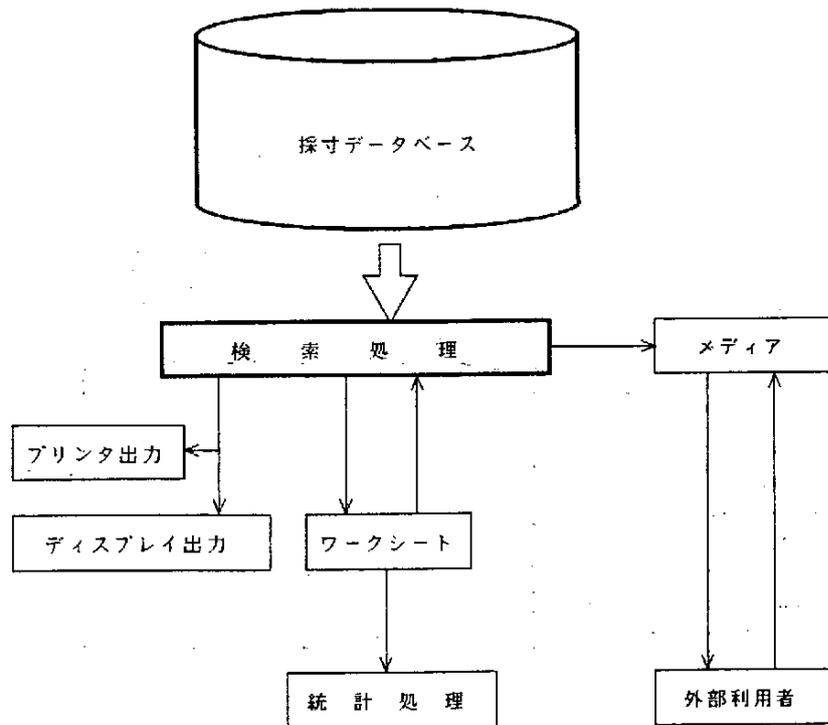


図 7-7 採寸データベースの機能

7.3.4 統計データベース

統計データベースの機能を図 7-8 に示す。

統計データベースは分類条件データと統計データの 2 種類のテーブルで構成されている。各テーブルは利用者がデータを検索、分類をするために行う条件設定に関する履歴、及び検索結果を保持し、同一条件での検索をスムーズに行うためのテーブルの集合である。

統計データは採寸データの ID をキーとして連結し、平均値、中央値、最大値、最小値、標準偏差、20 分位数、ヒストグラムのそれぞれの統計データを保持する。検索処理されたデータは必要に応じて、分類条件データと統計データとリンクした形で保存、更新される。

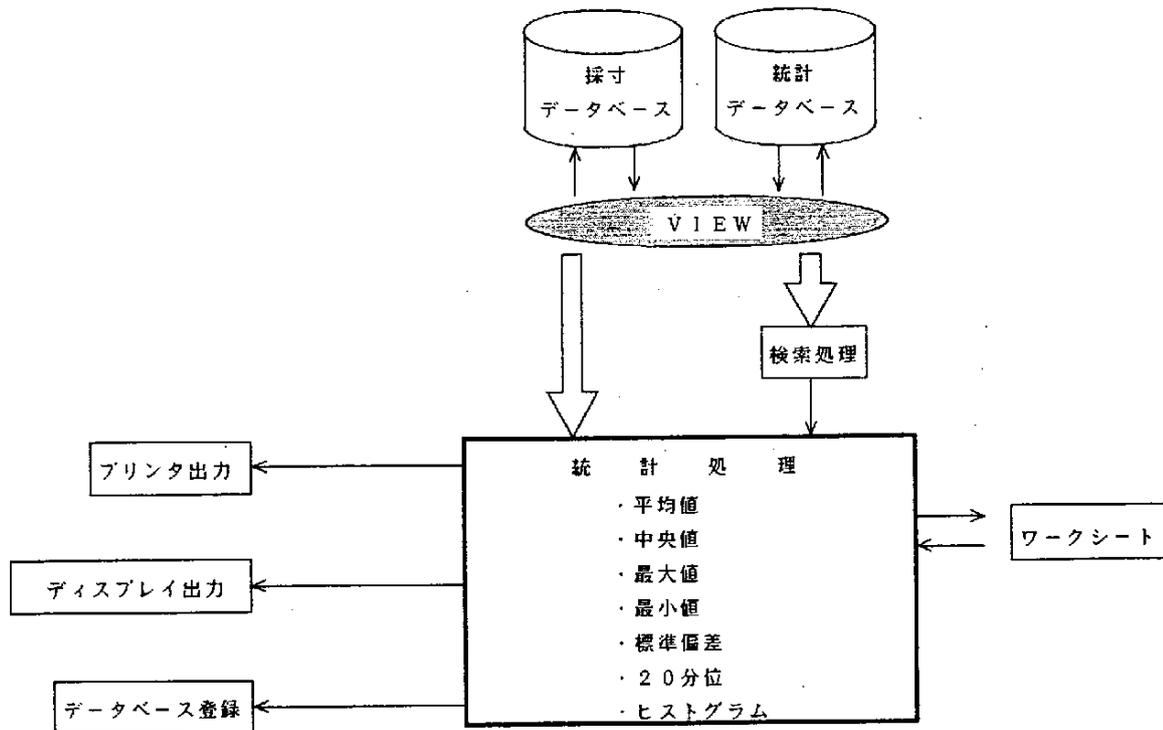


図7-8 統計データベースの機能

7.3.5 画像データ仕様

ユーザが178項目以外の計測値を要求した場合、再度画像データ呼び出しポインティング再計測を行う必要が生じる。画像データはユーザニーズに柔軟に対応できるような形で圧縮保存されている。

圧縮方法は、現在の未熟な非可逆圧縮では、画像データを欠落のない情報として残すことができないため可逆圧縮を採用した。画像圧縮率は約30%である。

7.3.6 データベースの運用

データベースの運用を図7-9に示す。

(1) データベースサービス

当センターは、各利用者の要求を満たすデータの検索を採寸データベースで行い、得られた情報をメディアにて利用者に引き渡す。利用者はデータに対して再度詳細な検索処理やグラフの作成を行うことができる。持ち帰ったデータをすぐに活用する場合は、アプリケーションソフトが当センターより提供される。

(2) 採寸データベースシステム運用手順

データベースの利用には、データベースを運用する側での利用と、利用者としての利用に大きく分かれる。主に運用する側ではデータベースを直接アクセスする方法がとられ、利用者側では、ワークシートでの提供により、情報の扱いやすさとデータベースの保護を重視した。

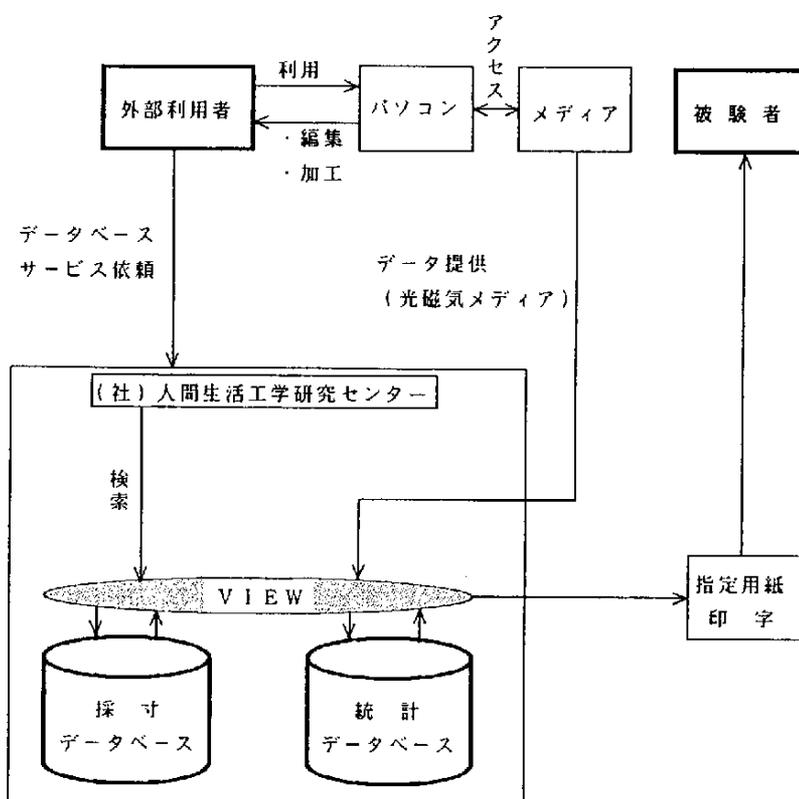


図 7-9 データベースの運用

7.4 今後の課題 (まとめ)

7.4.1 今後の課題

今後当センターにて進めていくべき課題について示す。

(1) 三次元画像のデータベース化

今回作成した採寸、統計データベースは“人体計測データベース”事業の第1ステップでありサイズデータのみをデータベース化している。

当センターが保有する人体形状に関するデータは、そのほとんどが画像で表現される形のデータである。当センターではサイズデータだけでなく、コンピュータグラフィック等の技術を用い人体モデルを動かしながら人体にフィットする最適な製品設計を行うなどの

画像データの有効利用を検討している。そのための第2ステップとして、さまざまな局面にてユーザニーズを満たす画像データベースの構築を今後1年間の予定で進める。

(2) データ提供のオンライン化

ユーザに対するサービス向上のための電話回線による外部ユーザとのオンライン化及び、オンライン化に伴うセキュリティ、プライバシー保護の検討を行う必要がある。

(3) 三次元画像データの高付加価値化

① 平均形状の検討

形状の平均、分布等の技術的可能性を検討している。形状平均手法が確立されれば、三次元画像データより日本人の平均形状等が算出可能である。

② 可動域の推定

静止三次元人体情報から人間の骨格情報等を鑑み、動的な人体の可動領域を推定する手法の技術的可能性を検討している。前述した平均形状の算出と併せて用いることにより、コンピュータ上で人体モデルを同時に見ながら製品の設計、評価が可能となる。

7.4.2 まとめ

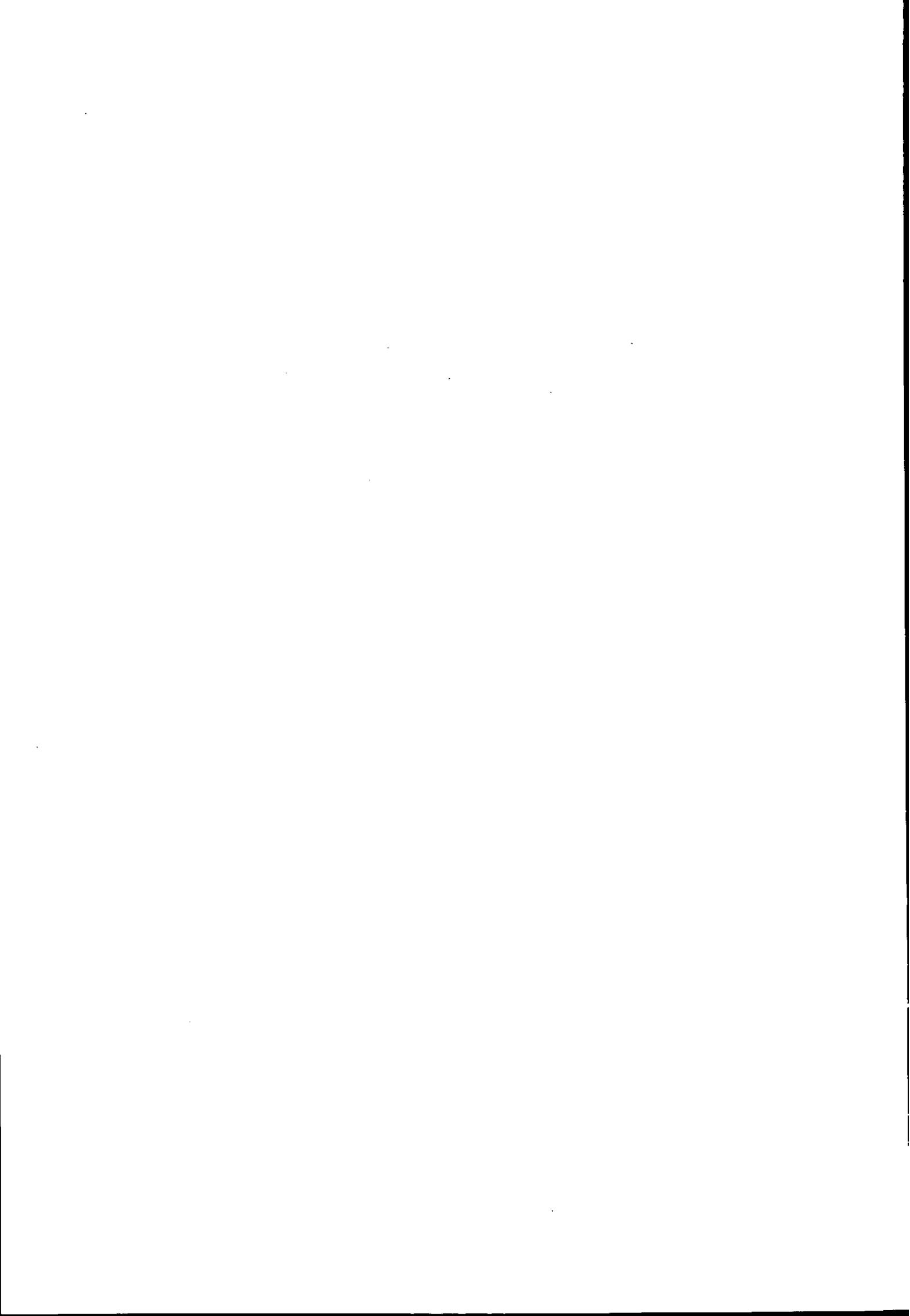
「ゆとりある豊かな生活」が求められる時代に入り、効率・コスト重視であった製品開発・技術開発も転換期を迎えている。今後は人間生活に着目した技術が不可欠となり、開発はそれを受けて展開されていくことであろう。

今回の人体形状計測は過去に類を見ない規模で行われ、国内外から注目されている。ここで得られたデータは、JIS規格をつくる時に利用されたり、人にやさしく、かつ快適なものを生むための基礎資料として使われることになるであろう。

人体計測データベースが「人に優しい、暮らしをもっと快適に」するための製品設計に用いられ、衣類を始め家具、住宅などのさまざまな製品が人間本位につくられるための一助となるように、また、生活者の満足度の向上、ひいては生活環境の質的向上のための一助となるように祈念する。

8 大学におけるデータベース利用教育
システムのプロトタイプ作成

日外アソシエーツ株式会社



- | | |
|--------------------|--|
| ③ 甲南大学 | 経済データベース |
| ④ 産能短期大学 | データベース概説
情報検索システム
データベース演習(1)
データベース演習(2) |
| ⑤ 東洋大学 | 情報検索演習 |
| ⑥ Univ. of Indiana | Information storage and retrieval |
| ⑦ 富士通(株) | サーチャー入門講座
データベース2級対策講座 |

一般に大学におけるデータベース利用教育科目は、次の三つのタイプに分けて考えることができる。

- (a) 図書館学、情報学などを専攻する学生を対象にするもの
- (b) 社会に出てすぐに役立つよう、データベースのビジネスでの利用を中心に教えるもの
- (c) データベースの利用を学習・研究にいかすことを目的とするもの

(a)におけるデータベース利用教育科目は、図書館・情報学のカリキュラムの中の科目として欠かせない科目となっているが、図書館・情報学を専攻する学生は日本の大学全体から考えれば極めて少数である。したがって、この科目を受講できる学生の数は限られているといえよう。

また(b)ではデータベース検索それ自体が学習の対象となっており、データベース検索によって得た情報で何をするのかが不鮮明である一方で、目的意識を最もはっきりもてるケースでもある。(c)は図書館・情報学を専攻するのではない学生を対象にし、学習・研究の一つの、そして効果的な手段としてデータベースを利用することを教えるものである。残念ながら、(c)に当たる授業はまだ実施例が少なく、必要性が十分に認識されているとはいいがたい。

これからの高等教育の中での情報リテラシー教育として必要なのは、情報について専門に学ぶ学生以外にも、データベース利用を含む情報処理技術一般を、各人の学習・研究と関連づけて教えることであると考えられる。それは上の例でいえば(c)のケースであり、本研究では、この(c)を更に一般化した形で、データベース利用教育における科目のモデルを示す。

8.3 データベース利用教育における科目モデルの設定と指導マニュアル案

8.3.1 基本的な考え方

ここではデータベース利用教育の授業科目として、以下のような一般学生を対象とした講義と演習、並びに各専門に応じた演習の計3科目を半期2単位(15週)で設定することを提

案する。特に前二つの科目は必修とすることが望ましい。

- ① データベース検索論（講義）
- ② データベース検索演習Ⅰ（基礎コース）
- ③ データベース検索演習Ⅱ（応用コース）

データベース検索論（講義）では、どうしても教員が主導権を握ることになるが、できるだけ学生が“情報”に対し興味をもてるように工夫しており、そのために黒板だけを使った一元的な授業は避け、ビデオテープやCD-ROM、オンラインなどさまざまな媒体を使った授業を行う。またデータベース検索演習Ⅰ（基礎コース）とデータベース検索演習Ⅱ（応用コース）では、体験学習と学生自らの興味に主眼を置き、各学生に合った教育指導（個別化教育）を行う。

指導マニュアルで共通していえることは、各回でねらい、内容を明確にし、最低限必要な知識を与えるような工夫をしていることである。教材では具体的な資料名やビデオテープ名、配布すべき資料、OHPを使用する際の図・表の名称などを挙げている。また課題では、学生が興味を覚えながら、自然と資料へ到達したり、データベースを駆使できるよう工夫している。そして検索演習における進め方では、操作法を中心に学生が慣れ親しめるような展開を図っている。

8.3.2 講義スケジュール

(1) データベース検索論（講義）

回	講義項目	講義内容
①	生活と情報と図書館	身近な情報から図書館の使い方まで
②	情報とは何か	情報一般についての理解を深める
③	学術情報（1）	レポート，卒業論文の書き方と二次資料の 使い方を把握する
④	学術情報（2）	
⑤	一般情報	ビジネスに係わる周辺情報を紹介する
⑥	ビジネス情報	実務で役立つビジネス情報を身につける
⑦	データベースとは何か	データベースの基本的な知識をもたせる
⑧	データベースの仕組み	データベース使用の際のポイントを学ぶ
⑨	データベースのサービス	主要なデータベースを理解する
⑩	検索のプロセス（1）	検索を始めるに当たって必要な知識を身に つけ，検索のプロセスを理解し，基本的検 索操作をマスターする
⑪	検索のプロセス（2）	
⑫	検索のプロセス（3）	
⑬	検索の技法（1）	目的に応じたデータベースの選定と効率の 良い検索手法を習得する
⑭	検索の技法（2）	
⑮	データベースサービスの展開	データベースサービスの将来像を学ぶ

(2) データベース検索演習 I (基礎コース)

回	講 義 項 目	講 義 内 容
①	パソコンの基本機能と操作	コンピュータの仕組みと操作法を学ぶ
②	オンディスク検索	CD-ROMデータベースを理解する
③	新聞記事情報 (1)	情報検索用フロントエンドプロセッサを使って新聞記事データベースを紹介し検索手法を身につけ現物との違いを知る
④	新聞記事情報 (2)	
⑤	新聞記事情報 (3)	
⑥	雑誌記事情報 (1)	上記検索ソフトを使って雑誌記事データベースを紹介し検索手法を身につけ、一次資料、二次資料との違いを認識する
⑦	雑誌記事情報 (2)	
⑧	雑誌記事情報 (3)	
⑨	図書情報 (1)	上記検索ソフトを使って図書データベースを紹介し検索手法を身につけ、現物と対比し違いを把握する
⑩	図書情報 (2)	
⑪	図書情報 (3)	
⑫	DIALOG	DIALOGの検索手法を理解する
⑬	会社情報 (1)	上記検索ソフトを使ってDIALOGの会社情報データベースの検索手法を学ぶ
⑭	会社情報 (2)	
⑮	まとめ	検索結果の活用法と原報入手法を学ぶ

(3) データベース検索演習Ⅱ（応用コース）

回	講義項目	講義内容
①	データベース検索演習Ⅱを始めるに当たって	コースを始めるに当たって修得しておくべきことを確認する
②	利用するデータベースの紹介	主題ごとのデータベースを理解する
③ ④ ⑤ ⑥	オンライン検索手法（１） オンライン検索手法（２） オンライン検索手法（３） オンライン検索手法（４）	オンラインデータベースの検索手法や検索戦略の構築法を学び、実例を示しながら理解を深める
⑦	検索例の紹介	オンラインでの検索例を読む
⑧	検索結果の評価	検索結果の評価法を学ぶ
⑨ ⑩ ⑪ ⑫	検索演習（１） 検索演習（２） 検索演習（３） 検索演習（４）	各自それぞれのテーマについて検索を行う、グループ演習も行う
⑬	ダウンロードデータの利用	ダウンロードデータの利用法を学ぶ
⑭	オンライン検索サービス利用の応用	これまでとりあげなかったオンライン検索サービスを紹介する
⑮	まとめ	演習内容の総括と他検索システムの紹介

8.4 プロトタイプシステム

8.4.1 位置づけ

第3章で述べたデータベース検索演習Ⅰ（基礎コース）及びデータベース検索演習Ⅱ（応用コース）ではデータベース検索演習が含まれており、演習授業で利用できるデータベース検索のためのソフトウェアが必要となる。このソフトウェアとして本報告書では次章に示す「情報検索用フロントエンドプロセッサ」を提案するが、それを開発する前に、プロトタイプシステムを実際の授業等で用いて演習教育の効果測定を行い、その結果を情報検索用フロントエンドプロセッサの設計にフィードバックさせることがここでのねらいである。

8.4.2 機器構成

ハードウェア

3.5インチ又は5インチフロッピーディスクドライブ付きパソコン。なお、演習において検索結果を印刷する必要もあるので、プリンタが装備されていることが望ましい。

ソフトウェア

オペレーティングシステムはMS-DOSとする。媒体は、システム及びトレーニング用データベースを1枚ないし2枚のフロッピーディスクに収めるものとする。

8.4.3 トレーニング対象となるデータベース

トレーニング対象とするデータベースは以下のものとする。

- ① J I C S T 科学技術文献ファイル
- ② J B C A T
- ③ J A P A N - M A R C
- ④ W H O
- ⑤ M A G A Z I N E
- ⑥ 朝日新聞記事データベース

ただし、前述のようにこのプロトタイプシステムはフロッピーベースで稼働するシステムとするために、トレーニング対象となる各データベースファイルの一部の情報（レコード）をフロッピーに収めることとする。収めるレコードの選択基準は、データベース演習授業でのプロトタイプシステムの利用の仕方に関連してくるので、主としてオフラインの授業であるデータベース検索演習Ⅰ（基礎コース）の講義内容を参考にし、そこでの検索例として適切と思われる情報を収めることとする。

8.4.4 システムの概要

このプロトタイプシステムは基本的には第5章の情報検索用フロントエンドプロセッサのオフラインの機能をサポートしている。

オフライン機能に限定した理由は、教育現場での情報機器の状況を考慮した上で、なるべく多くの利用者の試用を実現するためからである。

したがって、トレーニング対象となる幾つかのモデルデータベースをシステムに内蔵させ、それぞれのデータベース検索コマンド体系を模擬する機能を組み込むこととなる。

このシステムは情報機器の利用経験の浅い人、データベース検索を初めて行う人を対象としている。したがってデータベースによる情報検索の機能を単に提供するのではなく、入門者が誤りがちな操作の説明、キーボードの説明、更に情報検索における補助機能などが必要となる。

本プロトタイプシステムの開始から終了までの操作の流れを次に示す。

操作手順：

① 開始メニュー

② FDの使い方

システムの目的、使用するキーの説明や、ヘルプ機能などの説明

なお、ここでこのプロトタイプシステムは練習用のソフトであり実際の情報検索と操作方法や検索できる情報の種類などが異なる旨のメッセージを表示する

③ データベースファイルの選択

画面に表示されている六つのデータベースファイルの中から利用するデータベースファイルを選択する

④ データベースファイルの概要

選択したデータベースファイルの概要について説明する

⑤ 検索内容

選択したデータベースファイルの検索項目、検索方法などについて説明する

⑥ こんな時に使います

選択したデータベースを利用してどのようなことができるのかを簡単に説明する

⑦ 事例検索

あらかじめ用意されている事例検索の開始

以下のステップに従い、キーワードや演算式を入力する

- ・キーワードの入力
- ・検索結果の検討

- ・論理演算（演算式）の入力
- ・検索結果の再検討
- ・結果の表示
- ・検索の終了
- ・新規検索の開始

（以下、最初のキーワードの入力から繰り返し）

⑧ データベースの終了

データベース選択へ戻る

⑨ プロトタイプシステムの終了

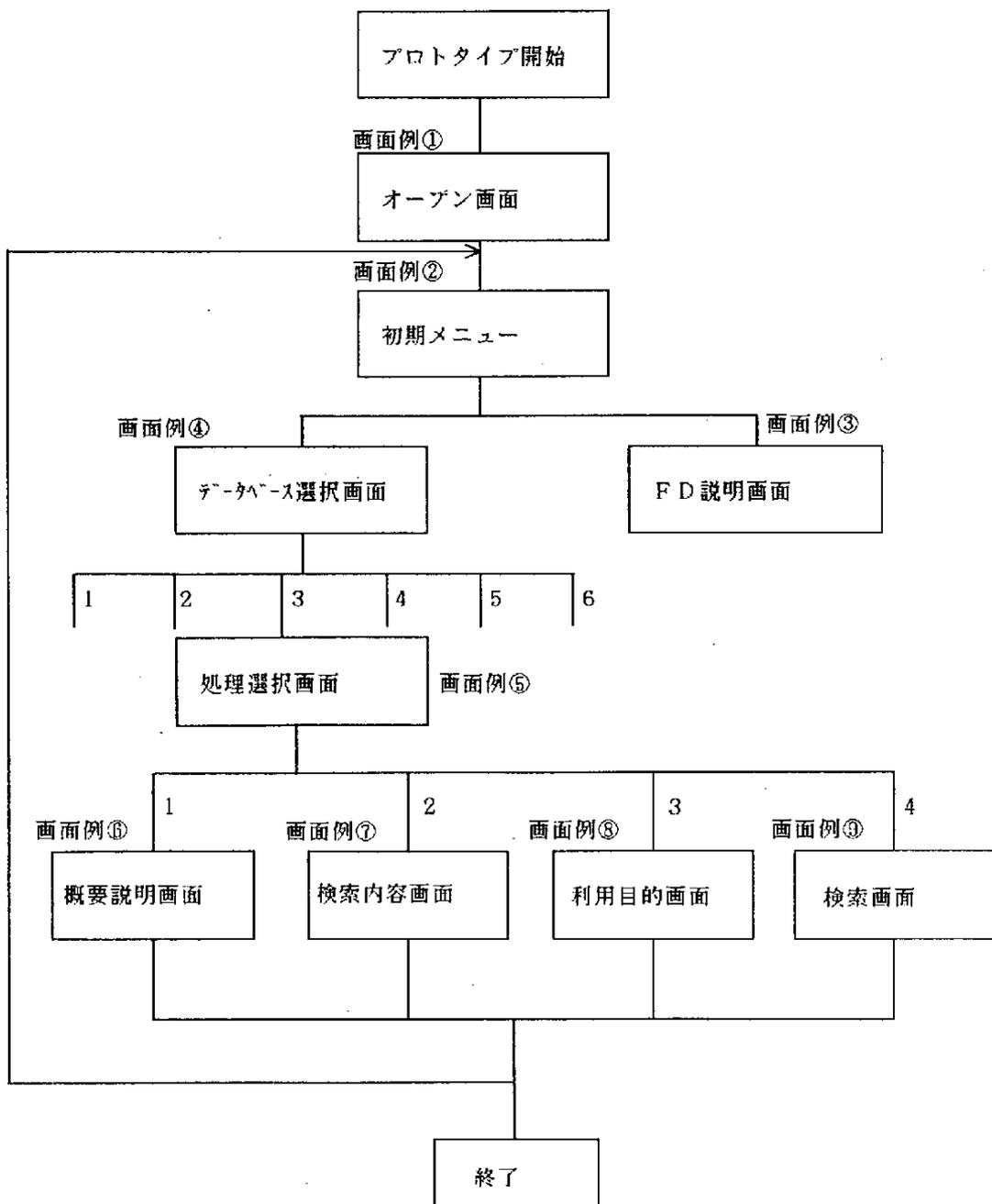


図 8-1 プロトタイプシステムのフローチャート

8.5 今後開発する必要があるシステム

ここでは、第4章のプロトタイプシステムの試用をふまえ、データベース検索演習で実際に利用するシステムとして今後検討する必要があると思われる「情報検索フロントエンドプロセッサ」について述べる。

8.5.1 位置づけ

この情報検索用フロントエンドプロセッサは、データベース検索演習Ⅰ、Ⅱでの演習用ソフトウェアとして利用するものである。データベース演習Ⅰでは前述のとおりオフラインで演習を行い、データベース演習Ⅱでは実際のデータベースにオンラインで接続して演習を行うので、この情報検索用フロントエンドプロセッサは、オフライン、オンラインの双方の機能が必要となる。

8.5.2 利用環境

本ソフトウェアの利用環境としては、近い将来の技術進歩を考慮して、以下の三つの環境を想定する。

① パソコン

パソコンと公衆回線をモデムを介して直接接続する利用環境

② パソコン+LAN

複数のパソコンをLANで接続し、ゲートウェイ（例えば、パソコン又はワークステーション）を介して公衆回線に接続する利用環境

③ ワークステーション+LAN

ワークステーション上で実行されている本ソフトウェアから、直接公衆回線に接続するか又はゲートウェイを介して公衆回線等に接続する利用環境

8.5.3 機器構成

ハードウェア

本ソフトウェアの実行機器は、パソコン並びにワークステーション、ハードディスク、CD-ROM、LAN、プリンタで構成される。ただし、ソフトウェアの種類としては、パソコン版とワークステーション版を分ける方が現実的である。

ソフトウェア

オペレーティングシステムは、パソコン版はMS-DOS、ワークステーション版

はUNIXとする。ワークステーション版は、ウィンドウシステムとして、X、OPEN-WINDOW、MOTIFをサポートする。

8.5.4 必要とされる機能

(1) オフラインサポート（基本的にはプロトタイプシステムと同じ）

① 各種トレーニング用データベース

トレーニング用データベースの情報量が少なすぎると、データベース検索演習Ⅰの実習での情報検索のリアリティに欠けると思われるので、プロトタイプシステムの試用結果も参考にして、ある程度の情報量を確保する必要がある。そのために、フロッピーでは容量が少ない場合には、CD-ROMなどの媒体でトレーニング用データベースを供給するか、ハードディスクにトレーニング用データベースをインストールして使うことも検討する必要がある。

トレーニング用データベースの種類は、プロトタイプシステムの6種類の他に、追加する必要があると判断されたデータベースがあれば追加するものとする。

② 検索コマンドのシミュレーション

各トレーニング用データベースの検索言語体系をシミュレートするのに加え、検索手順等の説明を必要な箇所に表示する。

(2) オンラインサポート

① 通信機能

公衆回線を使ってオンラインデータベースと通信を行うために、通信機能を備える必要がある。通信機能に加えて、オートログイン機能、画面表示スクロールアップ、スクロールダウン機能、画面表示のカット&ペースト機能、端末属性の設定、キャラクターコードの選択、通信速度の設定などのホストの通信属性との整合を図るための機能も必要であろう。

② データベースディレクトリ機能

利用者が幾つものデータベースに試行錯誤的に接続して検索する煩わしさを省くために、実際の検索に先駆けて、入力した検索語を含むレコードが各データベースに何件あるかを表示する機能である。本システムでは、サポートするデータベースの索引情報を統合したファイルをCD-ROMなどで供給することによりこの機能を実現する。

③ インテリジェントヘルプ機能

データベース検索コマンド体系はデータベースごとに異なるので、利用者はデータベースごとに検索コマンドを覚えなければならない。この負担を軽減するために、ヘルプ

機能を提供しているデータベース検索システムもある。しかし、これらのヘルプメッセージが英語であったり、必ずしも分かりやすいメッセージばかりとはいえないのが現状である。そこで、現在接続しているデータベースの検索コマンドのヘルプメッセージを、利用者の検索状況に対応させ、利用者の要求に応じて表示するのが、インテリジェントヘルプ機能である。この情報検索フロントエンドプロセッサは、接続しているデータベース及び利用者の検索状況を逐次把握していて、利用者の要求（ホットキーなどを決めておく）があれば、その状況に即したヘルプメッセージをウィンドウで表示する。

④ ダウンロードデータの加工機能

検索した情報を加工するための機能である。

・ダウンロード機能

この機能は、検索結果を端末であるパソコンなどのディスクに保存するための機能である。検索結果に対して後で加工する必要がある場合に利用する機能である。

・エディタ機能

表示結果の加工などの際には、通信ソフトを終了させることなくエディタが使える必要がある。

・データコンバート機能

検索結果を日本語ワープロ、表計算ソフト、データベースソフトなどで使うためには、データ形式やファイル形式などの変換を必要とする場合がある。詳しい知識がなくても簡単に目的のアプリケーションソフトのデータ形式、ファイル形式にコンバートする機能である。

(例、LOTUS 1-2-3、一太郎、DBASEへ)

8.6 おわりに

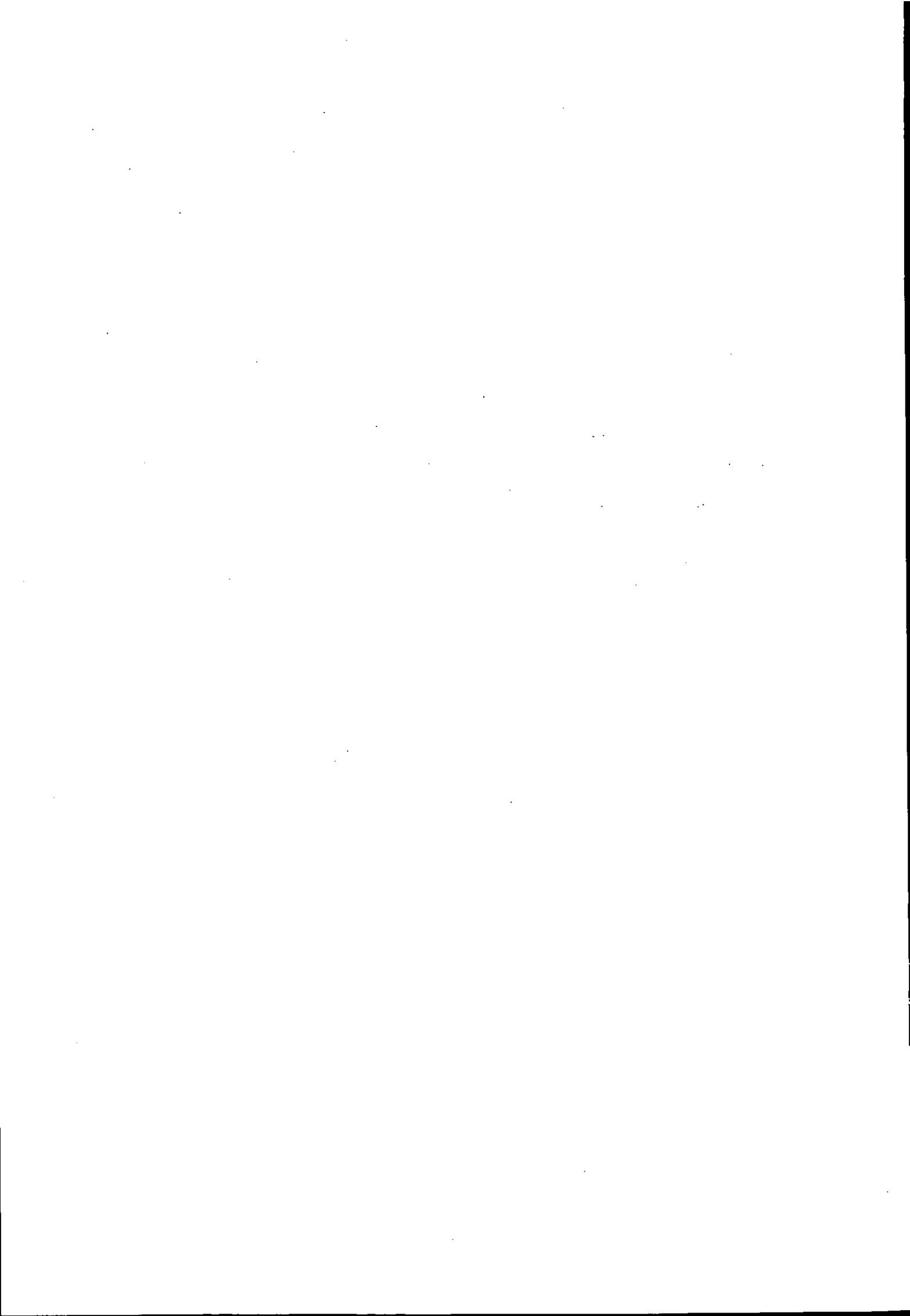
わが国の大学におけるデータベース利用教育の実施には、現在のところ、人や設備の面でも、また経費的な側面においても、さまざまな制約を受ける状況にある。しかし一方において、情報を効果的に活用する能力である情報リテラシーの修得の必要性はますます高まってきており、一部の大学では、担当者がそうした制約に苦慮しつつ授業を実践しているのが現状である。このような状況に対して、委員会及びワーキンググループは昨年夏から、どのようなカリキュラム内容（シラバス）と教材（特にシステム）とが用意されれば、データベース利用教育担当者にとって有益なツールとなり、なおかつ、データベース利用教育システムの確立に貢献できるであろうか、といった観点から検討を進め、その結果より具体的な指針の提示をめざし、本研

究報告書をまとめた。なお、今後の課題としては、以下の2点が挙げられる。

第1番目の課題は、大学におけるデータベース利用教育を、その大学における講義要項に明確に位置づけることの必要性である。各大学においては、できる限り早急に、また可能な範囲内で、明確な基本方針をもって、すべての学部学生にデータベース利用教育を教授する環境を設定されることが望まれる。

第2番目の課題は、本研究報告書に記載されているカリキュラム案に従って講義、演習を行った際の、教育効果の測定方法を確立する必要がある。それは自己評価という観点からも、受講した学生自身が、客観的に自分の到達達成度が判明できるような、判定機能を伴ったシステムの確立が、次のステップとして望まれる。

以上のように、残された問題もあるが、本研究報告書が、わが国の大学におけるデータベース利用教育の実施に当たって、参考となりうるカリキュラム及び教材システムとして利用され、更に、今後のデータベース利用教育に関する活発な議論の契機となることを期待する。



9 先進複合材料データベースの構築

財団法人 次世代金属・複合材料研究開発協会



9. 先進複合材料データベースの構築

9.1 概 要

先端産業（航空、宇宙、エネルギー、医療、レジャー等）に必要な有機、金属及びセラミックス複合材料の特性に関するDB（データベース）を構築するため、WS（ワークステーション）の導入とシステム基本部を構築した。また、PC（パソコン）のデータを移植し、PCによるプロトタイプレベルのDBからハード及びソフトを高性能化して、プロトタイプで実施した成果を更に発展させた。これにより材料の開発及び開発材料の利用を支援し、国際レベルの先進複合材料DBを産業界に提供する基礎固めをすることができた。

その実施内容と成果概要は以下のとおりである。

(1) DB構築用ハードシステムの導入

導入したシステムはUNIX WSであり、データ検索の高速化、マルチユーザ化、データ通信等の機能によりサーバー/クライアント方式によるグループウェア（個々の研究者がデータを相互に利用できるシステム）を備えている。

(2) DBMS（DB管理システム）の基本部構築

本DB管理システムの主要部には市販のリレーショナル型を選定した。

また、電話回線によるBBS（電子掲示板）機能を付加した。

(3) データ収録、編集用周辺システムの整備

DB操作編集用PCを新設しWSへイーサネット接続したためにデータの入力編集が容易になった。

ファックスモデムを用意し、編集したデータは本システムのファイルから直接ユーザにG3仕様ファックスで配布できるようにした。

(4) PC用プロトタイプDBデータの移植による材料DBの試行

① 複合材料用素材、プリプレグ、製品等の文献、カタログ、ハンドブックによる特性値（力学、物理学、熱物性等）のデータを収録した。

② 収録する試験計測データには次世代産業基礎技術研究開発制度で試作された複合材料のデータを一部含めた。

③ 温度条件、主応力条件等の検索条件により所要目的の材料や材料の組み合わせを選択できる。

これらの実施体制として、主任研究員の下に専任研究員1名及び嘱託研究員1名及び複合材料関係官学産の委員・専門委員10人からなる委員会形式の審議機関を設けた。

9.2 DBシステム構築の背景

9.2.1 先進複合材料の背景と展望

本DBの構築の対象としている材料は先進複合材料と呼ばれているものである。近代産業で最初に華々しく実用化された複合材料はガラス繊維強化プラスチック（GFRP）であるが、それ以降40余年の間にプラスチック基材料（FRP）、金属基材料（FRM）、セラミック基材料（FRC）等の開発と進歩がなされている。

特に近10数年に開発された高性能複合材料を先進複合材料と呼んでいる。

その特長は在来の材料に比し、高い比強度や比弾性率等の優れた力学的特性値をもっている。それ故、宇宙や航空機等の先端産業にとって不可欠の材料になってきた。先進材料を含む複合材料の将来的市場は1985年のDu Pont社のHoney 副社長の予測では西暦2030年には金額面で現在の金属構造材に替わり50%を複合材料が占める趨勢であると出ている。

9.2.2 DBデータ源とデータ利用環境の位置づけ

開発された新材料使用製品の製作過程（工程）とそのデータの発生関連状況を図9-1に示す。工程作業結果で発生するデータは記録の対象及びDBの入力対象となるがDB利用目的に応じてDB収録の対象が選択されることになる。

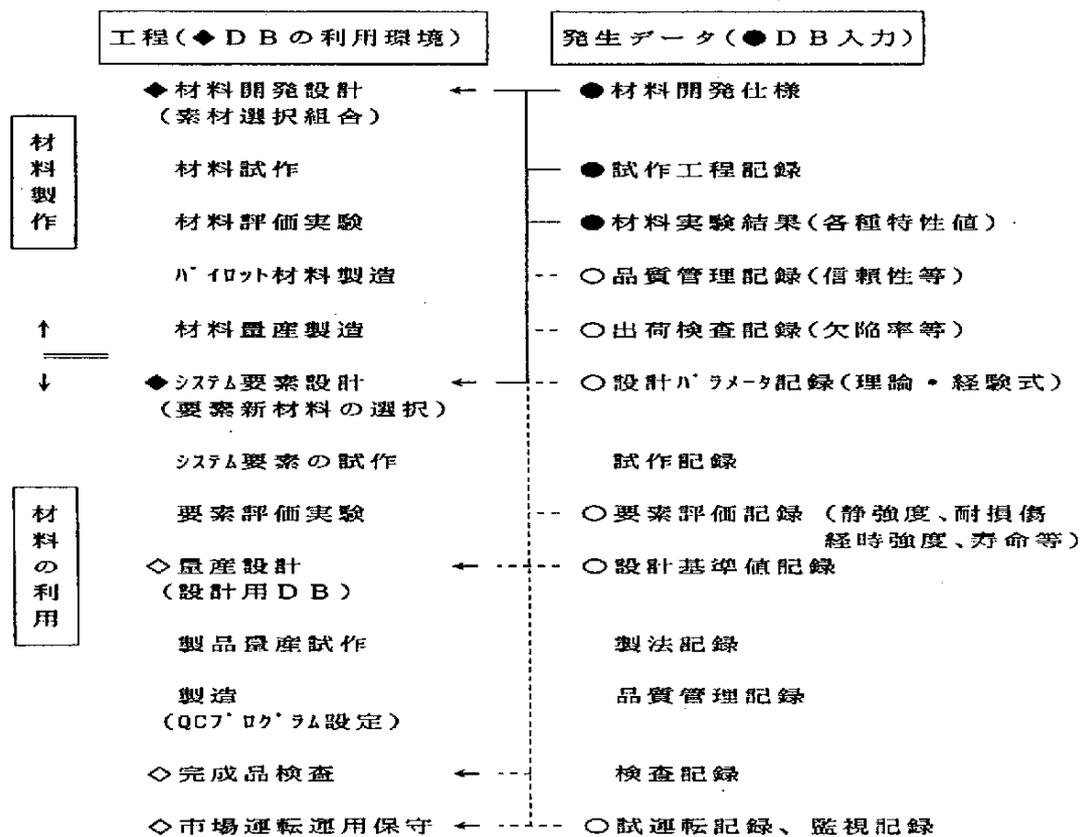


図9-1 製品製作過程とデータの関連図

今回のDBへの入力対象は表中に●印で示したファクトデータと文献等による二次データである。また、利用環境としては◆印を付した工程段階である。実線は両者の関連性を示す。○、◇印は設計用のデータ源とその主な利用環境を示し、点線は今回の対象DBとの関連性を表示している。

9.3 方針設定

9.3.1 基本方針

先に開発した先進複合材料プロトタイプのDBシステムを踏襲し、材料開発や利用方法に関する現在までの産業界の成果や経験則、知識等を整理し活用するDBの構築を基本方針とした。更に、下記のようなDB構築の技術環境の進展をふまえたシステム構成の方向性とした。

- ① ISDN通信システム環境への移行
- ② 画像処理技術の進展
- ③ オブジェクト指向のDB構造化技術への潮流
- ④ AI機能等知的DBへの発展

特に③は現在のDB構造化技術の一つの頂点であり、④の本格的なAI機能DB達成のための重要なステップと予想されているので、当該DBでは、複合材料に関するこの構造化DBモデルの構築を最終像とした。

9.3.2 DBの位置づけ

産業界における当該DBの位置づけは、図9-4に示すDB管理者及び材料開発者の相当位置にある文献データファイル、ソースデータファイル及びマスタファイルとする。これらのファイルはファイル単位での利用が可能となるように収録されるので一つのDBとみなされる。

9.3.3 収録するデータ

収録するデータファイルは大別して、文献値データと実験結果のファクトデータである。収録するに当たり、各地に分散している研究拠点からのデータ登録が可能とする。

データ内容は、複合材料用素材、プリプレグ、製品等の文献、カタログ、ハンドブックによる特性値（力学、物理学、熱物性等）や試験計測値データとするが、特に先進複合材料の利用が更に加速するための試験研究課題である疲労やクリープ、衝撃などに対する耐性、熱

や酸化・腐食に対する耐性など産業界のニーズへの対応を指向する。

9.3.4 マスタファイル

上記データファイルを整理編集して再収録するファイルをマスタファイルと呼ぶが、マスタファイルではDB構造、変数名や単位を統一したものとなり、製品の設計段階で本DBのユーザが設定する設計用DBに対して、データ内容の整合性を維持できるように配慮し、常に進展する知見やデータにより、スパイラルアップするシステムを目指す。

即ち、主として材料開発段階で利用するDBのデータは製品設計段階で利用する設計用DBの利用目的と連携し、製品設計フェーズの進展（概念設計、形状決定、詳細設計）の各段階に対応した特性値の取得、その項目範囲やデータ量の拡大等への柔軟な対応ができる構造を指向する。

9.4 ハードシステムの内容

今回導入したシステムは既設のパソコンによるプロトタイプDBの運用環境との整合性を保持し、他の研究機関とのオンラインによるグループ作業の利便性等を考慮し、SUN SPARC STATION IPX 及びPCによるサーバー/クライアント方式とした。

表9-1にRIMCOF（当協会英文略称）DBセンターのハード構成を示す。

表9-1 RIMCOFセンターのハード構成表

データベース構築用システム

ワークステーション本体

主記憶32MB	1台
内蔵ディスク1.4G	1台
内蔵ディスク424M	1台
5G磁気テープ装置	1台
644M CD-ROM装置	1台
16インチカラーCRT	1台
キーボード	1台

DB操作用端末パソコン	1台
FAXアダプタ	1台
レーザープリンタ	1台
8ポートハブ	1台
イーサネットボード	2台
モデム	2台

PC通信システム

専用パソコン	1台	既設
HDDディスク	1台	既設
モデム	1台	既設

検索用端末システム

多用途向けパソコン	1台	既設
HDDディスク	1台	既設
CD-ROM	1台	既設
光磁気ディスク	1台	既設
画像スキャナ	1台	既設
FAXアダプタ	1台	既設
レーザープリンタ	1台	既設

9.5 DB管理システム

構築したDBの管理システム(DBMS)としてリレーショナル型(RDBMS)を採用し

た。これはプロトタイプのDB構造を踏襲するためでもあるが、RDBMSとしての次のような特長や効果がある。

- ① 表形式のデータ構造であり、データはアプリケーションのプログラムに依存せず独立しているためデータの取扱いに柔軟性がある。
- ② 国際標準規格となっているSQL言語（DB問い合わせ）を使用でき、データ処理の標準化によりDB開発効率の向上が期待できる。
- ③ サーバー/クライアント方式によるネットワークによる分散化ができるので、システムの稼働性向上、リソースの経済的利用やリスクの分散が期待できる。
- ④ データの整合性の点検・管理がしやすい。

本DBとしては各種の市販RDBMSのうちから、日本国内での利用実績の多いDBMSであるInfomix系を採用し、DBMSのFront EndにInfomix SQL, 4GL, ESQL/Cを選択した。また、Back Endエンジンには画像処理可能なInfomix Onlineを選択している。

9.6 RIMCOF DBの特徴

(1) RIMCOF DBの階層構造

当協会のDBは協会の英文名称の頭文字を採りRIMCOF DBと名づけている。このDBの階層構造を今回の構築対象である先進複合材料（ACM）DBについて展開すると図9-2のようになる。

RIMCOF DB

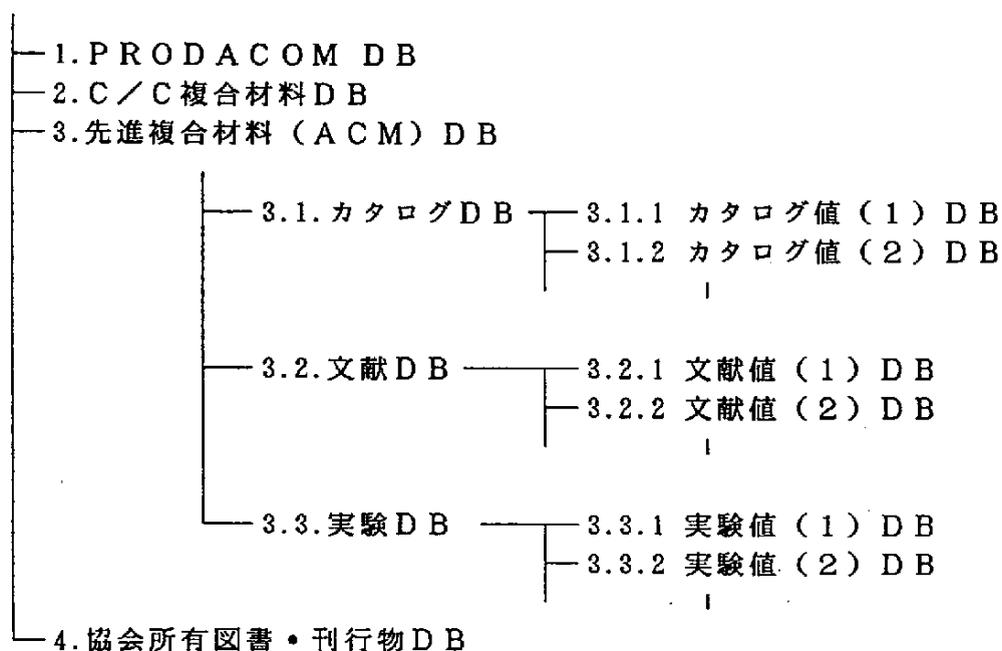


図9-2 先進複合材料（ACM）DBの階層構造展開

(2) データ構造

各種のACMDBを構築するに当たり、DB内各ファイルのデータ構造を設定する必要がある。そのためには材料の特性値項目を抽出・分類・整理して検討した。また、特性値項目間に存在する理論や知識・経験の関連性を把握して、データ収録用の各ファイル(表)にある項目間の関係(Relation)づけを行った。複合材料の一般的なリレーションは図9-3の例のような実体関連図で表現される。この基本的な機能と収録するデータの内容に応じて、DB内の各ファイルのデータ構造を設定している。

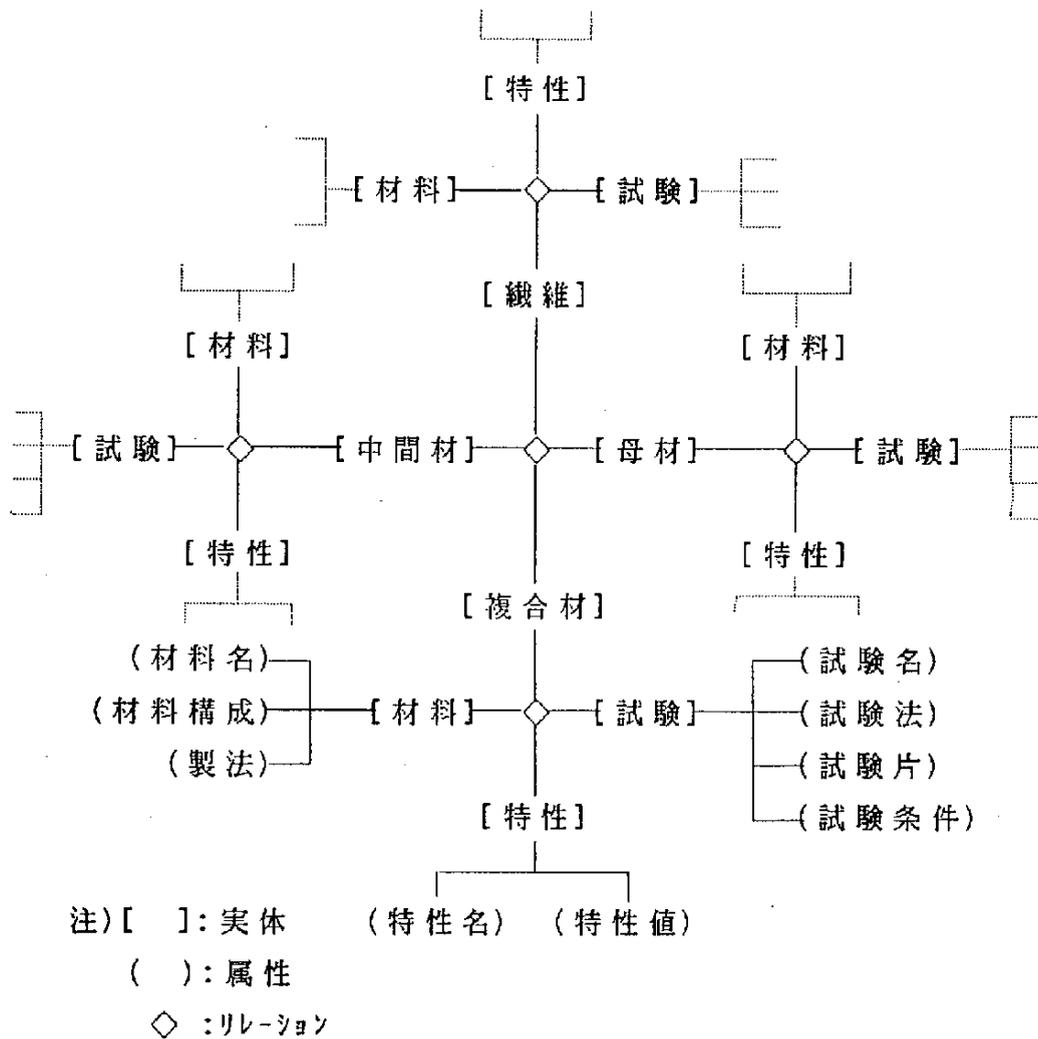


図9-3 複合材料実体関連図例

(3) 入力作業と検索利用

RIMCOF DBへのデータ入力はグループ作業で行えるよう配慮し、オフライン又はオンラインでのシステム化を行った。

文献値DBにオフラインで入力する場合は協会が発行している「ACMデータシート登録

マニュアル」に従って実施する。

実験値DBへの入力には端末研究拠点からBBS通信により「RIMCOF-BBS バイナリファイル アップロード手順」に従ってファイルデータを電話回線で登録する。この場合には入力データファイルを圧縮送信し、協会ホスト局で圧縮を解凍復元し、PC通信専用ホストが管理するDBに編集、入力を行う。ホストはWS（ワークステーション）と接続されているのでWSのDBにデータを変換入力する。

RIMCOF DBを検索利用の場合、現状では協会の検索用端末PCにより目的DBにアクセスするか、各研究拠点よりBBSによりアクセスすることになる。

BBSによる場合には現状のPC通信網で行う場合とWSに直接ログオンする場合がある。前者の場合、協会BBSホスト上にWS上のDBデータを再展開したものにアクセスすることになる。後者の場合は利用者許可者に別途発行する操作マニュアル（「複合材料研究開発DB/rimcof」の起動及び基本操作）を参考にして利用することになる。

9.7 将来のシステム拡充

9.7.1 材料データベースに対する新しい要求

新しい機械構造物には、超高温、超高压、超軽量、超高強度、超低温など、従来にない環境条件及び特性が要求され、これに応えるために多くの新しい材料が創生されてきた。

このような多くの新材料は、従来の慣用的な材料、即ち鉄鋼やアルミニウムなどと異なり、材料の特性のバランスが余り良くなく、その材料の著しい特長を十分に発揮させるには、その材料の使用条件、使用方法にかなりの配慮が必要となる。即ち機械構造の設計においては、材料の適材適所を十分に検討する必要がある。材料のデータベースはこうした複雑な設計作業の中で、設計者を強力に支援するシステムとして極めて有用である。設計者は多くの候補材料の比較検討を即座に行い、設計を進めることができる。

材料データに関する設計側からの要求事項を表9-2に示す。材料データベースはこうした要求に応えるものでなければならない。この中には標準的な材料データベースが備えるべきもの、及び利用者が独自に構築する必要があるものが存在する。このため、材料データベースはデータの検索及び出力のみならず、追加や修正の機能も必要になる。この場合、既存データは修正できず、また入力されたデータが蓄積されているデータと同一の水準で取り扱われる必要がある。

表 9 - 2 材料データに関する設計要求

要求内容	: 要求レベル
データの種類	: 多数データ, 単一データ, 内挿データ, ノウハウデータ
材料の認知度	: 確立された材料, 改良された材料, 社内材料, 革新的材料
信頼性	: 材料標準, 信頼限界, 正確さ, 感度解析, 追跡性
標準化	: 国際的標準, 国内標準, 顧客標準, 社内標準, 未標準, 要検討

9.7.2 開発したデータベースの課題と問題解決方途

開発したデータベースの今後検討しなければならない課題は次のとおりである。

- 1) 加工, 処理, あるいは経済性など, 材料特性の中でも従来余り考慮されなかったデータに対しても対応が可能なように, 特性データの構造を柔軟にする必要がある。
- 2) 設計の段階に応じて粒度(粗さ及び細かさ)の異なるデータを提供できるような階層的データ構造を実現する必要がある。
- 3) 材料データベースと設計/解析システムとの有機的な統合を図る必要がある。
- 4) 新規データの入力, 既存データの検索において知的な支援ができるようにする。

これらの課題はデータベースの新しい潮流のもとで克服できると思われる。

その潮流とは, リレーショナルデータベースからオブジェクト指向データベースへの進展である。

9.7.3 データベースの新しい潮流

現在のデータベースとして最もよく用いられている方式は, リレーショナルデータベースである。リレーショナルとは関係という意味であり, 現実の実体を計算機の中のデータとして表現する際のモデル化の一つの方法である。主要なモデル化は, 関係モデル, 階層モデル, そして網モデルである。関係モデルは実体の属性集合を二次元の表形式で表現するものであり, こうした表を複数枚, 階層的に結合すれば現実のデータをうまく表現することができる。また, 表形式は人間にとっても理解が容易で, 更に表同士を演算して新しい関係を作成するなどの操作が関係演算という数学に基づいて行われるため, 論理的に堅牢である, などの特長がある。また, 現在広く用いられていること, ファイルへの記憶が容易に行えるなどの周辺的長所もある。

しかしながら、最近では幾つかの欠点も指摘されてきた。それらは、第1に、二次元の表形式では、二次元の関係しか表現できず、多くの次元データは別の関係表として結合され、互いの関係が見えにくいことである。材料データベースでは、基本的な実体として、材料、試験、結果（特性）が存在し、材料には複数の試験が対応し、試験には複数の結果が対応する。この対応関係は1枚の表にはならず、複数の表が複雑に対応し、全体としての関係が特定の関係表に束縛されてしまう。

第2の欠点は、クラス階層構造と、クラス-インスタンス関係がなく、すべてのデータが個別的なデータとなり、データ自体の階層性がない。もし、データも階層性を有するならば、例えば、アルミニウムという上位クラスを設け、その下位クラスとして種々の合金を考えたとき、上位クラスに与えられた一般的なデータを下位のクラスはデフォルトとして参照することができる。そして、下位のクラスでそのデータが与えられた場合にはそのデータが参照される。こうしたメカニズムは継承（インヘリタンス）と呼ばれる。継承機能は通常のリレーショナルデータベースには備わっていない。

第3に、データの検索・出力以外の解析など複雑な処理をデータベースと有機的に結合することは困難である。現在では幾つかの計算機言語で組まれたプログラムの中でデータベースを呼び出して、そこからデータを得て解析や設計を進める環境は実現されている。しかしながら、データベースの中でデータ駆動型の種々のプログラムを起動させることは、一般には困難である。データの一貫性制御などの動作はデータベースの機能として備わっているが、更に複雑な処理を組み込むことはできない場合が多い。

この他にも幾つかの欠点が挙げられるが、こうした欠点を克服するデータモデルがオブジェクト指向である。

9.7.4 オブジェクト指向データベース

9.7.4.1 歴史的背景

第1世代のデータベースはファイルシステムであった。第2世代は階層データベースシステムである。第3世代はCODASYLデータベースシステムである。第2、第3世代のデータベースでは一つの環境の中で多くのユーザが統合化されたデータベースを利用することが可能になった。しかしながら、データの独立性と操作の面倒さから第4世代のデータベースであるリレーショナルデータベースが開発された。この特徴は宣言的な問い合わせという概念である。第5世代のデータベースの特徴は、第1にデータモデルが豊富であること、第2にビジネスのデータ処理を越えて、さまざまな応用に柔軟に対応できる多

くの機能を有していること、などである。

1980年代に新しいデータベースが幾つか開発された。その方向は二つあって、一つはリレーショナルデータベースを拡張するもの、他の一つはオブジェクト指向データベースである。両者の相違はデータモデルとデータベース言語にある。リレーショナルデータベースを拡張する方法では、まず、リレーショナルモデルとリレーショナル問い合わせ言語から出発し、モデルの拡張と意味的操作を行う。POSTGRESは最も有名な第5世代の拡張リレーショナルデータベースである。一方、オブジェクト指向を用いた方法では、ORION, Ontos, IRISなどがある。重要なことは、拡張されたリレーショナルデータモデルよりもオブジェクト指向データモデルのほうがはるかに自然であるということである。

9.7.4.2 オブジェクト指向データモデル

オブジェクト指向データモデルとリレーショナルデータモデルには大きな相違がある。オブジェクト指向データモデルではオブジェクト指向の概念、即ちデータのカプセル化、インヘリタンス（継承）、ポリモアフィズム（多態性）という概念が含まれている。これらの概念は従来のデータモデルにはない。オブジェクト指向データモデルを用いることにより、対象実体が即座に、素直に表現でき、意味的なデータ操作が行え、そしてオブジェクト指向計算機言語との親和性が極めて高い。こうして、設計がオブジェクト指向によって行える。

オブジェクト指向言語を用いる場合には、オブジェクト指向データベースと完全に統合化することも可能である。このときデータベースを操作する上でのプログラム上でのインタフェースの問題はなくなる。

オブジェクト指向の考え方は複雑なソフトウェアシステムを設計、開発、保守する上で、将来最も有望な方法である。この考え方が将来の知識ベースシステムにおける知識の表現において基本的な方法として定着するとき、オブジェクト指向データベースはこうした環境において最も重要なデータベースとなる。

9.7.4.3 オブジェクト指向とデータベース

オブジェクト指向は、ソフトウェアの設計と製作における革新技術であり、オブジェクト指向データベースはこのオブジェクト指向というアプローチに基づいて構築される。即ち、応用における実体とその関係をオブジェクト指向によって表現する。オブジェクト指向の考え方は高性能知的システムの開発には欠かせないものとなっている。

データモデルとは実世界のオブジェクト（実体），それらについての制約，及びオブジェクト間の関係を論理的に組織化したものである。オブジェクト指向の概念によって構築されるデータモデルをオブジェクト指向データモデルと呼ぶ。オブジェクト指向データベースはオブジェクトの集合であり，その挙動，状態，そして関係がオブジェクト指向データモデルによって定義されている。オブジェクト指向データモデルは，CAD/CAE/CASE/CAM，知識ベースシステム，マルチメディア情報システム，先進的ユーザインタフェースなどの応用において将来の最も重要なデータモデルとなる。

9.7.4.4 データモデル

オブジェクト指向データモデルはオブジェクト指向によって行うデータのモデル化である。残念ながら，オブジェクト指向言語はそれぞれ多少異なった方法をもっているためオブジェクト指向について普遍的な方法は存在しない。しかしながら，共通の重要なデータのモデル化についての概念は存在し，次のようになる。

- ① オブジェクト及びオブジェクトのアイデンティファイア
- ② アトリビュート（属性）とメソッド
- ③ カプセル化及びメッセージパッシング
- ④ クラス
- ⑤ クラス階層構造及びインヘリタンス
- ⑥ インスタンス

このようなデータモデルに基づいて構築されるデータベースがオブジェクト指向データベースであるが，こうした期待に反してオブジェクト指向で組み立てられる言語，知識ベース，そしてデータベースに標準化されたものはない。このため，例えば商業的データベースをオブジェクト指向で構築しても無意味であると考える人もいる。確かにリレーショナルデータベースは現在確固たる地位を築いている。しかしながら，オブジェクト指向データベースはいずれ本格的な普及が始まると思われる。それまでに基本的な検討を行い，リレーショナルデータベースとオブジェクト指向データベースの相互変換の道を探ることは極めて重要であろう。

9.8 最終システム運営像

最終的に構築した先進複合材料を一般利用する場合，フロッピーディスク（FD）又はCD-ROMによるオフライン利用と通信回線によるオンライン利用が考えられる。この両者の長

短を比較すると、オンラインでは一般にどんな機種でもアクセスできるがFD/CD-ROMを活用するにはPCの機種対応をどうするかが問題となる。しかし、オンラインの場合もホストでの経費負担をどのようにするか、問題がないとはいえない。

しかし、将来の技術環境をふまえて考えると検索は知的機能のあるオンラインで行い、そのデータ利用はFD/CD-ROMによるオフラインメディアによるデータ利用が最適組み合わせと予想されている。いずれにしてもデータプロデューサとして両者に対応可能な体制の構築を目指している。

先進複合材料DBの利用を実現した例として、近く発売される先進複合材料データを紹介する。

これは先に、次世代産業基盤技術研究開発制度で先進複合材料を開発し、このデータを汎用機システム「PRODACOM」DBとして収録したが、これを平成3年度の重要データベースの開発として「プロトタイプの構築」を受託実施した際、PCシステムで活用できるようにデータを変換し、複合材料DBプロトタイプとした。今回、一般にオフラインで活用できるように市販のPC用DBソフト（dBASE相当類）を用い、更にFDに付加した専用プログラムによりこのデータを用いて複合材料の設計に利用できるようにしてディストリビュータから発売することになった。

これにより先端産業界のボトムアップが期待でき、国家的プロジェクトの有効化メリットは計り知れないものとなる。また、産業界にとっても初めての出来事であり、材料ファクトデータの取り込みとDB化に弾みがつくことは間違いない。この意味において、時代を画するものとなる。

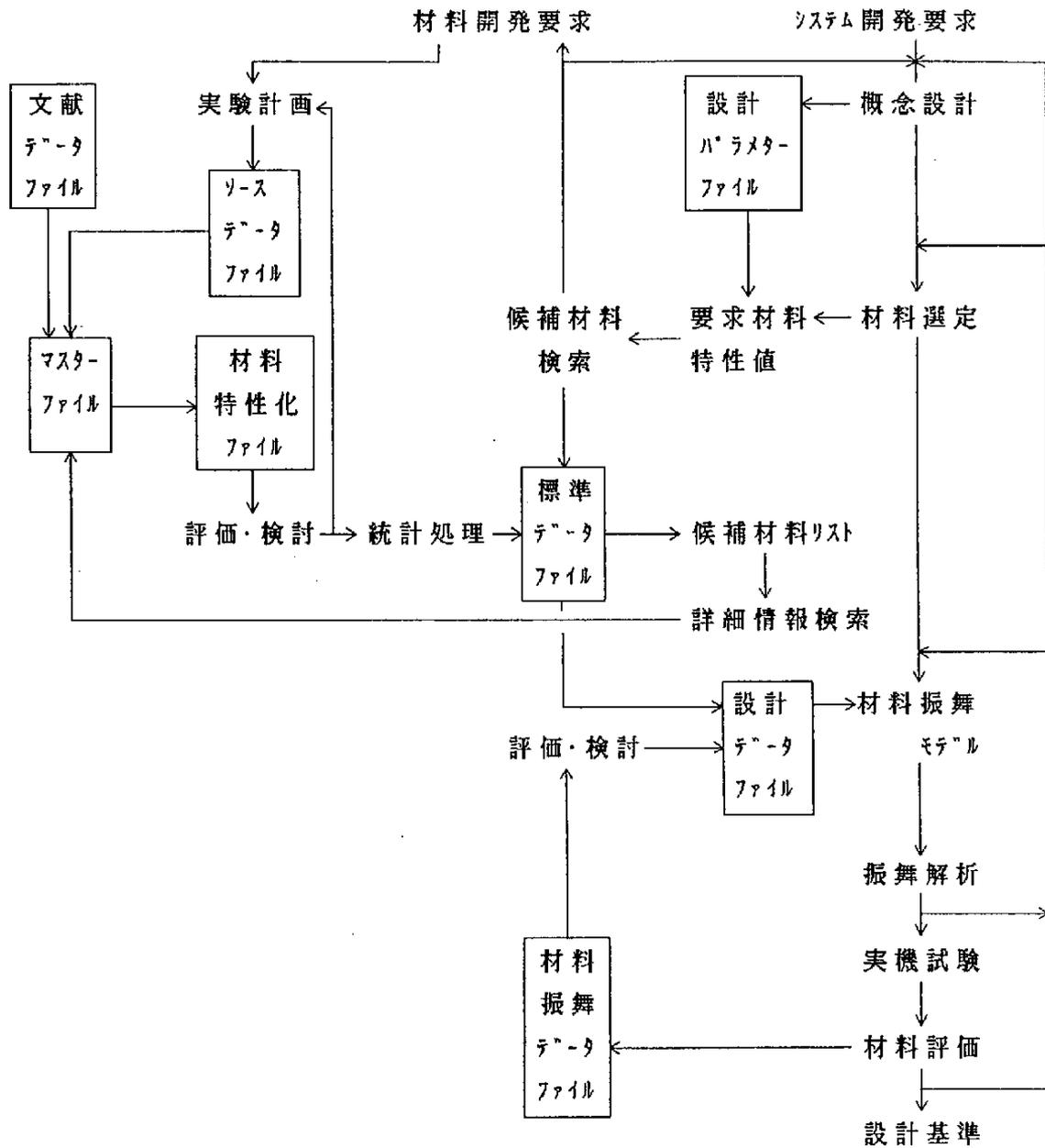


図9-4 材料DBと材料開発・システム開発の関係

10 博物館所蔵地図資料所在情報
データベースの構築調査

財団法人 地図情報センター



10. 博物館所蔵地図資料所在情報データベースの構築調査

10.1 調査研究の概要

10.1.1 調査研究の目的

地図は地上の諸事象を図上に表現し、多方面にわたるデータを包含した極めて実用性に富む情報源の一つである。現在、これらの地図の多くは、作成機関は勿論のこととして、研究機関、更に、図書館、博物館などの機関において所蔵・保存されている。しかしながら、それらの各機関において、所蔵・保存されている地図及び地図関係資料を、ただちに網羅的に把握することは難しい状況にある。

一方において、地理・地図関係の学者、研究者又は学生、出版・マスコミ関係者を始め一般市民等からの地図及び地図関係資料に関する問い合わせが急増している現状である。必要な地図及び地図関係資料に関してパーソナルコンピュータなどを使ってすぐさま検索し、情報を得ることができるというのが究極の目的ではあるが、その前段階として所蔵状況を調査し、所在情報の提供を図るのが急務である。

このような状況をふまえ、今回、博物館所蔵地図資料所在情報データベースの構築調査を行った。

10.1.2 実施内容

平成4年度においては、約2,000館の博物館を対象に、地図及び地図関係資料の所蔵状況調査を実施し、各博物館所蔵の地図関係資料の数量、種類、形態、作製年代、発行所（国）についてのデータにより、わが国博物館における地図及び地図関係資料の所在情報データベースの構築を行った。

(1) 調査票の作成

地図にはさまざまな種類のものがあり、その分類には基本的に形態、大きさ、素材、発行年（時代）、描かれている範囲（大陸別、国別、地域別など）、縮尺、発行所などが考えられなければならない。これらのさまざまな地図を構成している要素について、今回のデータベースの構築は所在情報に関するものであることから、地図の所在を知る上で最低限必要なものに絞った。

博物館所蔵地図資料所在情報データベースの構築調査のための調査票に必要な要素としては〔形態〕、〔発行年〕、〔発行所〕（発行国）を選定し、調査票の設問事項の設定を行った。

〔形態〕については、一枚もの地図、地図帳、及び地球儀を選定し、その他、立体地図模型、電子地図、壁面や屏風など描かれている素材の違いによるものなどは特殊形態のものとして一括することとした。

〔発行年〕については、近代測量による地図であるかどうかで大きく分け、外国の場合も勘案して、19世紀までのものと、20世紀のものに区分することとした。

更に、19世紀までの一枚ものの地図を古地図、地図帳を古地図帳とすることとし、この古地図（19世紀までのもの）については、手描きと版ものの区別を行った回答を求めた。

〔発行所〕（発行国）については、日本発行と外国発行とに2大別した。

以上の他、〔地図関係の書籍・雑誌〕の有無及び〔地図関係目録〕の有無についても設問項目を設けた。

(2) 調査票の配布と回収

〔調査票〕の配布対象となる博物館名等に関する資料としては、(財)日本博物館協会（1991年11月）発行の「全国博物館園職員録」を使用した。

実際に発送配布した博物館数は 1,922、有効回答数 1,186であった。

なお、有効回答数 1,186は、調査票配布数 1,922の62%に当たる。

10.2 実施の結果

10.2.1 データベース作成

(1) データベース項目

本データベースでは、入力ツールとして調査票「地図資料ご所蔵状況及び地図展に関する調査」に基づき、データ入力用データベースシートを設計してある。これはデータベースのマネジメントシステムではなく、あくまでも最終的にMS-DOSのプレーンテキストへコンバートするためのものである。表10-1にフィールド項目とデータ形式の一覧表を示す。

なお、データ長は、特に記載がある場合を除き、1バイトコード（キャラクタ）によって計算している。また、可変長データについては特に制限を設けなかったが、概ね40バイト長を限度として入力している。

(2) データの入力

データは、調査票回収段階でフェイスシートと調査対象リストとのチェックの上、入力用に設計されたデータベース画面からキーボードにより入力を行った。入力用プログラムとしては、市販のパーソナルコンピュータ用データベースソフトウェアであるThe card 3+を採用した。今回の調査票においては入力の際にリレーショナルにリンクさせるべきデータがな

いこと、入力ミスの少ない画面設計ができること、入力の訂正が容易かつ迅速にできること、入力値の範囲演算、論理演算ができることなどが採用の理由である。

なお、調査票の関係でコーディングシートを使用していないので、それに伴う入力・検定ミス等に対処するため、作業用フラグを別途用意した。また、あり得ない数値に対するエラーメッセージやトラップ、調査項目の区切りに対するメッセージビープなどを設定してある。

表10-1 フィールド名・データ形式・データ内容

フェイスシート

FN	フィールド名	調査表項目名	データ形式・長	データ内容	備考
1	番号	番号	##### 8バイト長	自治体コード+シリアル番号	シリアル番号は任意
2	博物館名	博物館名	可変長 2バイトコード 容認	返送された調査表の名称に準拠する	財団などの法人名称は割愛
3	読み	読み	可変長 2バイトコード 容認	返送された調査表の名称に準拠する	財団などの法人名称は割愛
4	郵便番号	郵便番号	##### 6バイト長	平成4年度郵便番号表に準拠	
5	住所	住所	可変長 2バイトコード 容認	返送された調査表の住所に準拠する	
6	電話番号	電話番号	市街局番 局番 番号 /12バイト長	返送された調査表の電話番号に準拠する	FAX番号、内線は割愛

1-1 一枚もの地図

FN	フィールド名	調査表項目名	データ形式・長	データ内容	備考
7	1 日本	1 枚もの地図 (日本発行)	4バイト長	実数/0/-/+	NUL値不可
8	1 外国	1 枚もの地図 (外国発行)	4バイト長	実数/0/-/+	NUL値不可
9	1 地形日	地形図 (日本発行・国家機関作成)	4バイト長	実数/0/-/+	NUL値不可
10	1 地形外	地形図 (外国発行・国家機関作成)	4バイト長	実数/0/-/+	NUL値不可
11	1 地質日	地質図 (日本発行)	4バイト長	実数/0/-/+	NUL値不可
12	1 地質外	地質図 (外国発行)	4バイト長	実数/0/-/+	NUL値不可
13	1 海図日	海図 (日本発行)	4バイト長	実数/0/-/+	NUL値不可
14	1 海図外	海図 (外国発行)	4バイト長	実数/0/-/+	NUL値不可
15	1 地籍	地籍図	4バイト長	実数/0/-/+	NUL値不可
16	1 その他	その他 1 枚もの	4バイト長	実数/0/-/+	NUL値不可

1-2 一枚もの古地図

FN	フィールド名	調査表項目名	データ形式・長	データ内容	備考
17	古1日手	古地図 (日本発行・手描き)	4バイト長	実数/0/-/+	NUL値不可
18	古1日版	古地図 (日本発行・版もの)	4バイト長	実数/0/-/+	NUL値不可
19	古1外手	古地図 (外国発行・手描き)	4バイト長	実数/0/-/+	NUL値不可
20	古1外版	古地図 (外国発行・版もの)	4バイト長	実数/0/-/+	NUL値不可
21	他1日手	その他特殊 (日本発行・手描き)	4バイト長	実数/0/-/+	NUL値不可
22	他1日版	その他特殊 (日本発行・版もの)	4バイト長	実数/0/-/+	NUL値不可
23	他1外手	その他特殊 (外国発行・手描き)	4バイト長	実数/0/-/+	NUL値不可
24	他1外版	その他特殊 (外国発行・版もの)	4バイト長	実数/0/-/+	NUL値不可

10.2.2 データベースの利用方法

10.2.2.1 システムの概要

本データベースは図10-1に示すような、今日最も普及しているパソコンシステムにより使用することができる。このシステムには、MS-DOSによって作動するパソコン、ディスプレイ、プリンタが必要である。更に、本システムを作動させるためのカード型データベースのソフトウェアをパソコン本体のハードディスク上にインストールしておくことが最低限必要である。

このシステムでは、次のような項目からの検索が可能である。

- (1) 都道府県・市町村コード
- (2) 博物館名
- (3) 郵便番号
- (4) 住所
- (5) 電話番号

各項目により検索されたデータは、各カード型データベースソフトウェアの有する集計機能や作図機能を用いることで、二次的な加工を行うことも可能である。

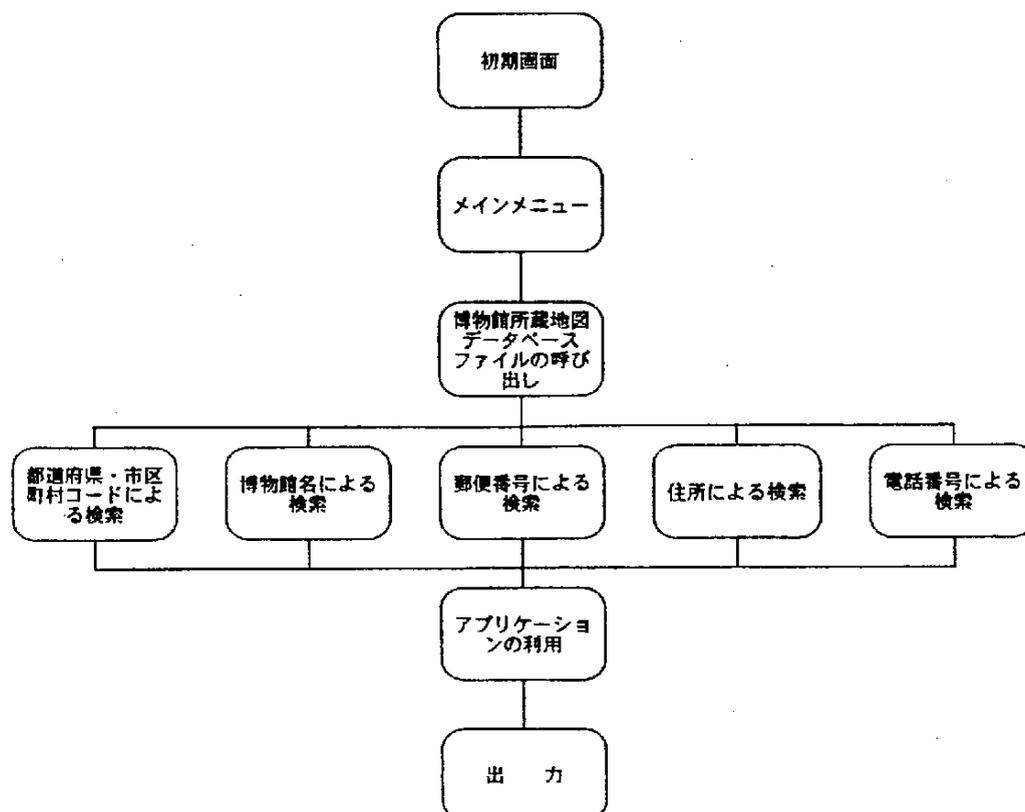


図10-1 システムの構成

10.2.2.2 システムの操作

前項システムに基づく検索例の手順例は以下のとおりである。

検索条件設定 (図-1~2) (都道府県コードによる検索)

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	<input type="checkbox"/>
2: 博物館名	<input type="checkbox"/>
3: 番号がわかる	<input type="checkbox"/>
4: 郵便番号	<input type="checkbox"/>
5: 住所	<input type="checkbox"/>
6: 電話番号	<input type="checkbox"/>
7: 一都道府県(日本)	<input checked="" type="checkbox"/>
8: 一都道府県(外国)	<input checked="" type="checkbox"/>
9: 都道府県(日本)	<input checked="" type="checkbox"/>
10: 都道府県(外国)	<input checked="" type="checkbox"/>
11: 都道府県(日本)	<input checked="" type="checkbox"/>
12: 都道府県(外国)	<input checked="" type="checkbox"/>
13: 都道府県(日本)	<input checked="" type="checkbox"/>
14: 都道府県(外国)	<input checked="" type="checkbox"/>

1 頁 1 行 7 5 6

図-1

11) 博物館地図 DB (1196件) 検索済 184件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	13101-02
2: 博物館名	海防記念館
3: 番号がわかる	伊人館(伊賀町人かん)
4: 郵便番号	180
5: 住所	東京都千代田区水田町1-1-1
6: 電話番号	03-3581-1951
7: 一都道府県(日本)	17
8: 一都道府県(外国)	0
9: 都道府県(日本)	--
10: 都道府県(外国)	--
11: 都道府県(日本)	--
12: 都道府県(外国)	--
13: 都道府県(日本)	--
14: 都道府県(外国)	--

カード表示 100 順昇 設定 0 (M7:ESCキー)

図-2

検索条件設定 (図-3~4) (博物館による検索)

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	<input checked="" type="checkbox"/>
2: 博物館名	<input checked="" type="checkbox"/>
3: 番号がわかる	<input type="checkbox"/>
4: 郵便番号	<input type="checkbox"/>
5: 住所	<input type="checkbox"/>
6: 電話番号	<input type="checkbox"/>
7: 一都道府県(日本)	<input type="checkbox"/>
8: 一都道府県(外国)	<input type="checkbox"/>
9: 都道府県(日本)	<input type="checkbox"/>
10: 都道府県(外国)	<input type="checkbox"/>
11: 都道府県(日本)	<input type="checkbox"/>
12: 都道府県(外国)	<input type="checkbox"/>
13: 都道府県(日本)	<input type="checkbox"/>
14: 都道府県(外国)	<input type="checkbox"/>

検索済 179 表示 100 順昇 設定 0

図-3

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	2 01232-07
2: 博物館名	北沢町歴史民俗資料館
3: 番号がわかる	4 00229-01
4: 郵便番号	5 00324-01
5: 住所	6 02467-01
6: 電話番号	7 02491-01
7: 一都道府県(日本)	8 02441-01
8: 一都道府県(外国)	9 02445-01
9: 都道府県(日本)	10 02370-02
10: 都道府県(外国)	11 03211-01
11: 都道府県(日本)	12 03213-01
12: 都道府県(外国)	13 03205-01
13: 都道府県(日本)	14 03307-02
14: 都道府県(外国)	15 03242-01
15: 都道府県(日本)	16 03265-01
16: 都道府県(外国)	17 03265-01
17: 都道府県(日本)	18 03202-01
18: 都道府県(外国)	19 03207-01

表示 100 順昇 設定 0

図-4

検索条件設定 (図-5~6) (複数条件による検索)

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

25: 博物館名	<input checked="" type="checkbox"/>
26: 博物館名(外国)	<input checked="" type="checkbox"/>
27: 博物館名(手)	<input checked="" type="checkbox"/>
28: 博物館名(外国)	<input checked="" type="checkbox"/>
29: 博物館名(外国)	<input checked="" type="checkbox"/>
30: 博物館名(手)	<input checked="" type="checkbox"/>
31: 博物館名(手)	<input checked="" type="checkbox"/>
32: 博物館名(手)	<input checked="" type="checkbox"/>
33: 博物館名(手)	<input checked="" type="checkbox"/>
34: 博物館名(手)	<input checked="" type="checkbox"/>
35: 博物館名(手)	<input checked="" type="checkbox"/>
36: 博物館名(手)	<input checked="" type="checkbox"/>
37: 博物館名(手)	<input checked="" type="checkbox"/>
38: 博物館名(手)	<input checked="" type="checkbox"/>
39: 博物館名(手)	<input checked="" type="checkbox"/>
40: 博物館名(手)	<input checked="" type="checkbox"/>

検索済 24 表示 100 順昇 設定 0

図-5

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	2 01556-01
2: 博物館名	3 04101-03
3: 番号がわかる	4 00229-01
4: 郵便番号	5 00324-01
5: 住所	6 02467-01
6: 電話番号	7 02491-01
7: 一都道府県(日本)	8 02441-01
8: 一都道府県(外国)	9 02445-01
9: 都道府県(日本)	10 02370-02
10: 都道府県(外国)	11 03211-01
11: 都道府県(日本)	12 03213-01
12: 都道府県(外国)	13 03205-01
13: 都道府県(日本)	14 03307-02
14: 都道府県(外国)	15 03242-01
15: 都道府県(日本)	16 03265-01
16: 都道府県(外国)	17 03265-01
17: 都道府県(日本)	18 03202-01
18: 都道府県(外国)	19 03207-01

表示 100 順昇 設定 0

図-6

検索データの集計 (図-7~8) (一枚もの地図の集計例)

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	
2: 博物館名	
3: 番号がわかる	
4: 郵便番号	
5: 住所	
6: 電話番号	
7: 一都道府県(日本)	最大-1000 最小-平均
8: 一都道府県(外国)	
9: 都道府県(日本)	
10: 都道府県(外国)	
11: 都道府県(日本)	
12: 都道府県(外国)	
13: 都道府県(日本)	
14: 都道府県(外国)	

1 頁 1 行 10 5 6

図-7

11) 博物館地図 DB (1196件) 検索済 179件 (検索済分 1880件) Nmap3Pro

(1) 項目名

1: 番号	
2: 博物館名	
3: 番号がわかる	
4: 郵便番号	
5: 住所	
6: 電話番号	
7: 一都道府県(日本)	最大-1000 最小-平均-101.16120
8: 一都道府県(外国)	
9: 都道府県(日本)	
10: 都道府県(外国)	
11: 都道府県(日本)	
12: 都道府県(外国)	
13: 都道府県(日本)	
14: 都道府県(外国)	

1 頁 1 行 10 5 6

図-8

10.3 今後の課題

10.3.1 地図情報と所在機関

現在、わが国において地図情報に対する機関としては、次のようなものがある。

(1) 地図作成機関

建設省国土地理院，建設省，通産省工業技術院地質調査所，通産省，農林水産省，海上保安庁水路部，気象庁，運輸省，国土庁，科学技術庁，総理府統計局，郵政省，文化庁，環境庁，北海道開発庁，防衛庁，法務省，自治省，地方公共団体，大学。

(2) 地図関係機関

日本国際地図学会，(財)日本地図センター，(財)地図情報センター，(社)地図協会，(財)日本水路協会，(財)国土地理協会，(財)日本リモートセンシング技術センター，東京地学協会，日本地図資料協会。

(3) 地図所蔵機関

図書館（国立国会図書館など），博物館（東京国立博物館，神戸市立博物館など），文書館「国立公文書館，埼玉県立文書館など），大学（地理学教室など），研究所（アジア経済研究所など），報道機関（(社)共同通信社など），(財)地図情報センター。

10.3.2 現在の地図利用形態

一般の利用者や専門の研究者が利用する形態で一番多いと考えられるのは、地図そのものの閲覧である。博物館・図書館・文書館などの施設で地図を利用する場合、一般の利用者はまず図書・地図カードなどで必要な地図の検索から始まり、地図の取り出し（職員による）、閲覧、必要に応じてのコピーや写真撮影、あるいは複写依頼などが考えられる。

次にどのような地図が存在しているかなど所在情報を得るためのカード等による検索であろう。また、地図展などの開催に伴う地図の貸し出しも考えられる。

10.3.3 地図資料データベース化に向けて

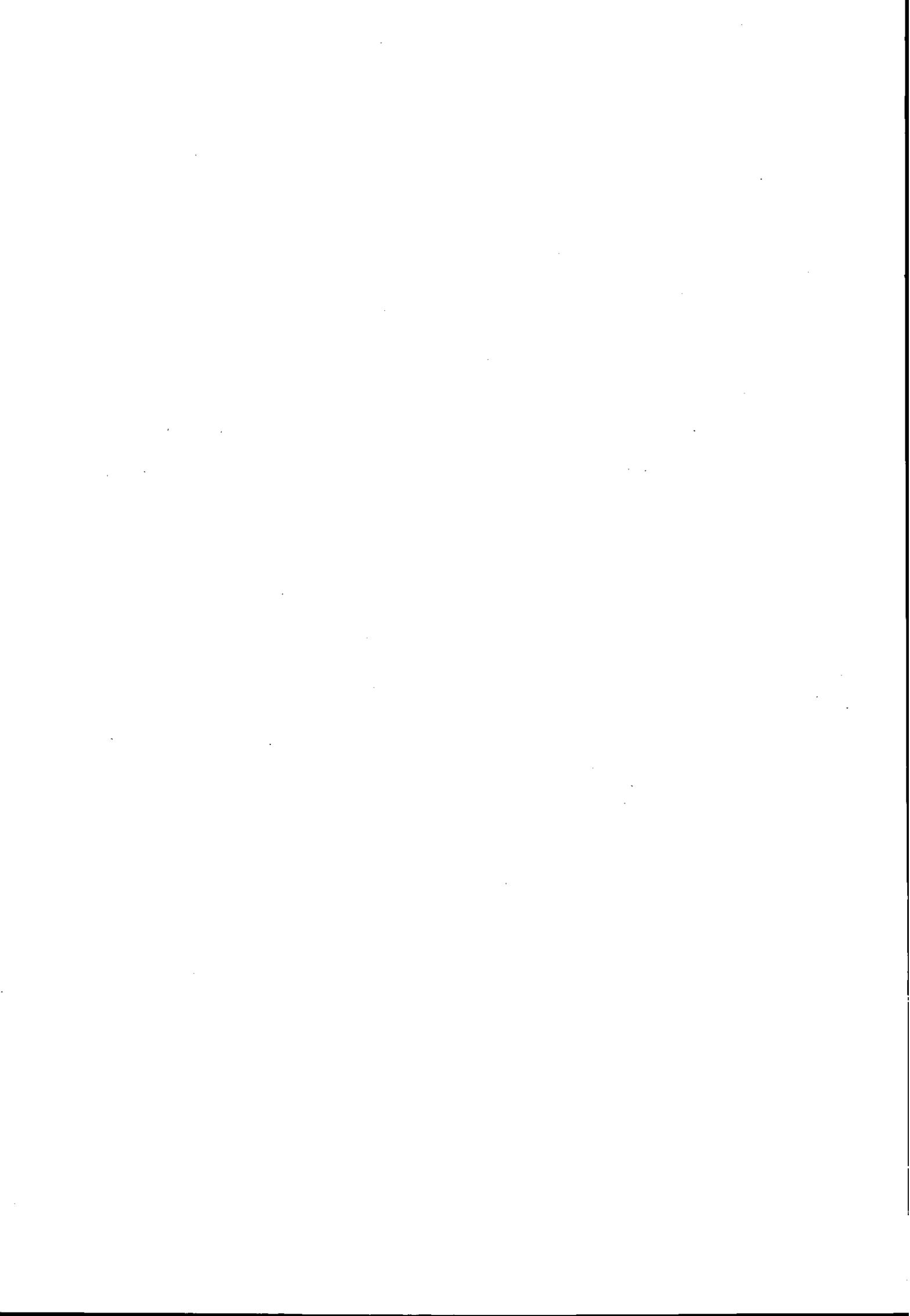
人類のこれまでの長い歴史の蓄積の中で、さまざまな地図が作成されてきた。それらは、その地域のその時代の情報を反映した画像情報資料であるにもかかわらず、この地図に関するデータベース化は遅れた状況におかれている。

国土基本図及び国土数値情報のデータベース化のほか、住宅地図や海図などの一部には、現在普及しつつあるCD-ROMなどに地図情報を入力することにより、パソコンのディスプレイ上での閲覧が可能となっている。しかし、地図研究者などは本物の地図に触れること

を望むことになろうし、また、古地図と称されるもののなかには、細部を見ないと（本物と贋作の識別など）役に立たない場合もある。特に紙質・色彩などを研究する際には、本物の地図使用が必須であることも考慮に入れなければならない。このようなところに、地図の保存と閲覧、そしてデータベース化のジレンマともいえる問題が発生する。

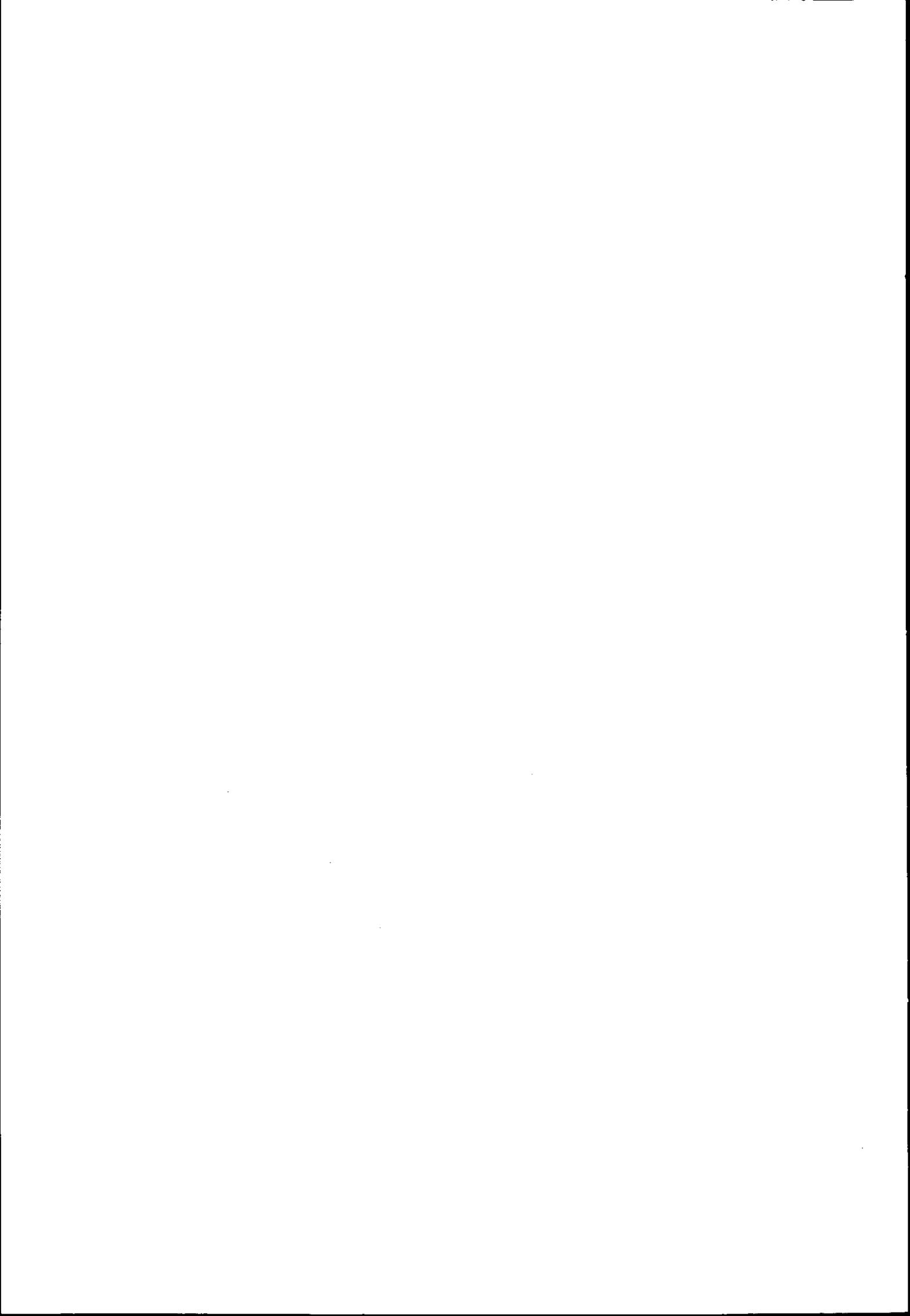
1989年以来、われわれは地図の所在情報データベース化をすすめ、図書館、博物館の一部について調査を行ってきた。今回は未調査博物館を対象として調査を行い、博物館所蔵地図資料所在情報データベースの構築を行ったものである。

今後の課題としては、調査未着手の文書館等の機関を対象とした地図の所在情報データベース化を完結することであると考え。次の段階として、所蔵されている地図についての図名等を始めとする、各種の属性を付加した地図データベースの構築がなされなければならない。しかしながら、この分野は書誌データベースとしても膨大な数量であり、今後の一層の調査研究が必要である。



11 地域流通最適化データベースの プロトタイプ作成

社団法人 日本ボランティア・チェーン協会



11. 地域流通最適化データベースのプロトタイプ作成

11.1 目的と概要

11.1.1 はじめに

わが国流通業界は未曾有の転換期に直面している。対外的に3年前の日米構造協議を契機とする、大店法の画期的な改正、及びこれに伴う内外にわたる流通規制の緩和と開放政策は、流通秩序の一大地殻変動を惹起しており、これにバブル不況が、個人消費の減退とともに加わり、一層大きな激動となっている。

特に、商業集積の立地について、消費者のカーライフは消費者の来店手段の革命をもたらし、徒歩・自転車を前提とした、商店街から自動車を基調とした郊外店へと大きく移行している。

これらの三大与件 -①政策の自由化 ②バブル不況 ③カーショッピングは相俟って、中小商店のあり方にも重大な革新を迫っており、その生き残りのために中小店も郊外立地への移行が必須の命題となっている。

本研究は、こうした大きな社会的要請に応えようとするものである。

(株)日本ボランティア・チェーン協会は、中小専門店約6万軒を擁する全国組織の団体であるので、協会としても重要な課題である。

11.1.2 目的

消費者のカーライフに基づく、カーショッピングに適合する店舗の最適立地を科学的にかつ適確に探查評価するシステムとプログラムを研究するためのプロトタイプを作成することが目的である。

小売業は、日々の生活を支える地域密着産業であり、かつ、消費者に、便利でかつ支持されるものでなければならない。

消費者は、個人単位であるので、中小商品に対しても、大企業と同様の物差しで取捨選択を自由に行う。これに評価されて始めて、中小店も生き残りうるのである。

したがって、地域ごとに、消費者に最も支持される商業立地を選び出す必要があり、それは、大手企業との複合立地であることも容認する必要がある。特に、中小店にとって、新規出店は、失敗を許されない。それにも拘わらず、他方において「出店」という業務については、中小店側にほとんど経験も、ノウハウもないのが通例である。

かかる矛盾を解決するため、当協会付属の地域流通研究所においては、第1年度において、

①問題の所在 ②多様なアプローチ ③適地判断上、視覚に訴える地図が重要な要素となるため、重ね合わせ方式によるラスタ情報为基础とする立地選定策を中心にした調査研究を終えた。

平成4年の第2年度の研究においては、更にこれを、視覚に訴える背景地図情報をカラー画像でもちながら、道路を考慮した自動車による人口移動のシミュレーションを可能にしたプログラムのプロトタイプを開発し、具体的都市についてそのシミュレーションを可能にするデータベースのプロトタイプを作成しようとするものである。

11.1.3 概 要

「地域流通最適化データベース」は、最適立地探索を行うための支援データベースである。最適立地探索は、次の5段階のステップで行われる。

第1段階：有望地域、有望都市の選定

第2段階：適地候補ゾーンの抽出

第3段階：適地候補地点の選定

第4段階：適地候補地点の評価

第5段階：適地候補地点の実地調査・検討

本年度はプロトタイプ作成であり、最適立地探索システム全体を動かすに際して核となる部分のプロトタイプを作成した。

具体的には、まず、各ステップについて再検討を行い、人間による判断が適切であるステップ（第1段階、第3段階、第5段階）と、コンピュータによる処理が適切であるステップ（第2段階、第4段階）について分け、前者については、資料整備、既存データベースの改良、判断手順の整備を行い、後者については、プロトタイプのソフトを開発した。これらを用いて、第1段階～第5段階の各ステップを実在の2都市を用いて行い、データベースのプロトタイプを作成した。このシミュレーションにより、今年度の研究の評価と、今後の進め方の方向性を得ることができた。

11.2 地域流通最適化データベースのプロトタイプ作成

11.2.1 本研究の目指すデータベースの将来像

中小店が集団で、単独、又は大手流通企業と複合立地で郊外型SC（ショッピングセンター）を建設しようとする際、商圈の実情や競合関係等から判断して、幾つかある適地候補の中から最適立地をコンピュータで選べるような、システム、プログラム、及

びこれらをふまえたデータベースを構築しようとするものである。

人口メッシュデータの1 km単位という制約がある程度克服されれば、郊外型SCのみならず、SM（スーパーマーケット）、専門大型店、外食等の立地についても応用が可能になると思われる。

また、将来的には、本年度のプロトタイプでは実現できなかった、自動車による右折左折の負荷等、人間の心理を反映した諸要素を解析の際に入れることや、解析の評価基準について種々工夫することにより、より消費者の行動に近いシミュレーションができるデータベース作成を考えている。

11.2.2 「有望地域及び有望都市」の選定に関するシステムの構築

このシステムは、最適立地探索のための第1段階目のステップである。

ここでは、①「有望地域の選定」及び、その地域内における②郊外型SCの「有望都市」の選定を行う。

①の地域選定は、府県別に大型店の普及度のデータを検討することにより可能である。これは、商業統計表、並びに国勢調査データより必要データを府県別にコンピュータに入力し、大型店売場面積当たりの支持人口を出すことにより一次的な判断を行うものである。②の都市選定は、①で出した地域に関して、市町村別に、郊外型大型店の普及度のデータを収集整備することにより判断できる。このため、スーパーマーケット名鑑のデータのうち、店舗規模、店舗立地、駐車台数により、郊外型大型店の要素を満たす店舗が未整備である都市を検索することによって一次的な判断を行う。

今回、この①と②の検索が可能な形に、既存のデータベースを整備した。条件検索画面は対話式であり、コンピュータに馴染みの薄いユーザでも、容易に求める検索が行え、自動的に印刷できるようになった。

11.2.3 「適地候補ゾーン」の抽出プログラムのプロトタイプ開発

このプログラムは、最適立地探索のための第2段階目のステップを可能にするものであり、これを使用して、「適地候補ゾーン」の抽出を行うのに必要なデータ（地図、人口、道路データ）を入力し、実行（フェーズ1と名づける）により「適地候補ゾーン」を抽出した。

このアルゴリズムの概要は次のとおりである。

- ① 各種データを入力する。
- ② 階層グラフを構成し、その枝の長さを計算する。
- ③ 評価基準の値 x (km) 圏内 n (人以上)

- ④ 道路グラフの各頂点 p に対して以下の操作を行う。
- ・階層グラフ上で、 p から各メッシュ頂点までの最短路を求める。
 - ・ p の集積人数を計算する。
 - ・ p の集積人数が n 人以上なら候補頂点として選ぶ。
- ⑤ 候補頂点と両頂点が候補となる枝を画面上に出力する。
- 必要であれば、PIXEL DIOで出力する。
- データベース入力操作は、次のように行う。
- ① 特定地域における地図をカラーキャナー (PIXEL DIO) により入力する。この地図データは、道路や障害物の入力画面の背景や候補ゾーンの紙への出力に用いられる。
- ② 人口データは、統計に用いられる標準地域基準メッシュ (1 kmメッシュ) の矩形区画で区切られた地図画面を見ながら、当該メッシュの中に、人口をマウスを用いて入力していく。
- ③ 道路データは、同じく背景地図を利用して計算機の画面からマウスによって入力する。
- 「適地候補ゾーン」抽出評価基準は、考えられる幾つかの評価基準のうち、プロトタイプでは、「道路沿い距離 x km 圏内人口 n 万人以上」を用いた。パラメータ x 、 n を変えて抽出アルゴリズムを再び起動することにより、適地候補ゾーンを絞り込むことや広げることも可能である。
- 「適地候補ゾーン」の出力は、評価基準を満たす道路グラフの頂点と、両頂点が基準を満たす枝を画面の背景地図に上書きすることにより行う。また、画面から、基準を満たす (候補) 頂点の集積人数を見ることがもできる。最終的な出力はPIXEL DIOを用いて、フルカラーで地図データ・道路データ・候補ゾーンを同一紙に出力することもできる。

11.2.4 「適地候補地点」の選定に関するシステムの構築

このシステムは、最適立地探索のための第3段階目のステップである。

ここでは、「適地候補地点」の選定を行う。

第2段階で抽出した「適地候補ゾーン」に沿って、まず、出店可能な地点を、都市計画図による土地用途区分、及び住宅地図による土地使用状況等から判断し、次にその上で、出店者側の要求する条件を満足する地点を捜し出すことである。

費用対効果を考えると、「適地候補ゾーン」さえ与えられれば、人間による判断の方が適確かつ時間もかからないため、このシステムは判断項目 (探索項目) を整備することにより、適宜行うこととした。その適地の探索項目については、リストの形式で記入票を作成した。

(何故なら、適地の要素は、項目は同じでも、出店希望者により内容が異なるため)

11.2.5 「適地候補地点」の評価プログラムのプロトタイプ開発

このプログラムは、最適立地探索のための第4段階目のステップを可能にするものであり、これを使用して、第3段階で探索した地点について、既存大型店をも含めて比較評価を行う。必要なデータのうち地図データ、人口データ、道路データについては、「適地候補ゾーン」抽出の際に使用したデータを用い、店舗データについて入力を行う。この実行（フェーズⅡと名づける）により、地図上の全人数のうち、既存店を含め、候補地点を利用すると予想される期待流入人数を計算することができ、この値により、各地点の評価を行うことができる。

このアルゴリズムの概要は次のとおりである。

- ① 各種データを入力する。
- ② 階層グラフを構成し、その枝の長さを計算する。
- ③ 階層グラフ上で候補店及び各競合店から客メッシュへの最短距離を計算する。
- ④ 修正ハフモデルを用いて、候補店及び各競合店への期待流入人数を算出する。
- ⑤ 結果を画面もしくはファイルに出力する。

データ入力操作は次のように行う。

- ① 地図、人口、道路データについては、「適地候補ゾーン」抽出の際に入力したデータを使用する。
- ② 競合店データ及び候補地点データの輸入は、道路節点と同様に計算機の画面に映された地図データ上でマウスによって行う。競合店に関するデータとしては、売場面積、駐車場台数、駐車場階数等を入力することができる。

候補地点のデータには、同じ項目について、想定されるデータを入力することにより、店舗比較のシミュレーション対象にすることが可能になる。

「適地候補地点」評価基準は、考えられる幾つかの評価基準のうち、プロトタイプでは、「地図上の全人数のうち候補店を利用すると予想される期待流入人数」を用いる。この期待流入人数の具体的な計算は修正ハフモデルを用いて行う。

「適地候補地点」の出力は、候補店及び各競合店の期待流入人数を画面上の店舗地点に行う。この値は、フェーズⅡを実行した後、売場面積などの店舗データと同様に画面上で参照することができる。

11.2.6 「適地候補地点」の現地調査・検討

これは、最適立地探索のための最終ステップである。

第4段階で評価した「適地候補地点」を、現地で評価の確認を行う。

今回、プロトタイプ作成に当たり、テスト都市について、現地調査を行ったが、基本的に

は通常経営者が業界常識と経験によって判断するのと同様な形のパターンになったと思われる。

11.3 地域流通最適化データベースのプロトタイプに関する各種検討

11.3.1 テスト都市を用いたシミュレーション

今回作成したプロトタイプの評価を行うため、最適立地探索システムの第1段階において有望都市と選定されたN市をテスト都市として取り上げ、その周辺20kmの範囲について、地図データ、人口データ、道路データ、店舗データからなるデータベースのプロトタイプを作成し、「適地候補ゾーン」の抽出と、「適地候補地点」の評価について行った。

11.3.2 データ処理作業の操作性に関する評価

(1) ハードディスク構成

プロトタイプでもあり、かつ、中小店支援ツールとしての役割を念頭においているため、基本的に、できるだけコンパクトで、投資負担のかからないハードディスク構成を心掛けた。

EWS	SONY NWS-3460 (CPU R3000)
	主記憶 8MB (最大16MB) ;
	ハードディスク 415MB ;
	MIPS値 17MIPS
光磁気ディスクユニット	SONY RMD-S550
光磁気ディスク	SONY EDM-1DA1S 記憶容量 594MBYTES
スキャナー/プリンタ	CANON CLC-500 PIXELDIO
	CANON IPU-10

(2) データ入力作業に関する評価

データ入力作業は、マンマシンインタフェースの工夫により、Xウィンドウ上でのマウス操作により容易に行うことができる。入力画面については、図11-1参照。

入力作業時間は、テスト都市について次のようであった。

縮尺5万分の1の地形図20km四方のエリアを入力するのに3.5時間、基準メッシュ(1km)400個に男女別人口を入力するのに1.5時間、道路(687本)入力に5時間、店舗(62店)入力に4.5時間を要した。

mapウィンドウのレスポンス時間は、背景地図データの容量に大きく左右されるが、

地図データ容量が 4,158,537byteのN都市については、背景地図データがある場合で、map起動に45秒、再描画に12秒、背景地図なしの場合、map起動に2秒、再描画に2秒である。

入力作業には、単純な数字データ入力ではなく、地図上の位置確認という判断が伴うため、このような時間がかかったが、2日間あれば1市町村のビジュアルなデータが入力できることになるので、プロトタイプとしては効率的であると思われる。

(3) データ解析操作に関する評価

フェーズ1、フェーズ2を解析するための実行条件を入力する操作は、データ入力と同じように、Xウィンドウ上でのマウス操作により容易に行うことができる。

フェーズ1の応答時間は、節点数が622、道路数698、店舗数62のN都市の場合、解析実行時間と再描画時間を合わせて、12分10秒であり、フェーズ2の応答時間は、1分35秒であった。

フェーズ1の応答時間については少し長く感じるかもしれないが、解析結果がビジュアルにゾーンの色変換により表現されるのでその有用性は大きく、EWSという限られたメモリ対効果から考えれば、実用に耐えられると思われる。

(4) データ出力に関する評価

解析結果の出力表現は、誰が見ても一目で分かるようにビジュアルなことに重点を置き開発を行った。その結果、フェーズ1の実行条件(x km以内にn万人以上の商圈をもつ点)を満たす道路上の節点を黄色で、両端の節点が条件を満たす場合は道路を赤色、片端の節点が条件を満たす場合は道路を橙色に変わるよう、視覚的に表現される。また、各節点の解析結果の数字を読むこともできる。

フェーズ2に関しては、解析結果は各店舗を表すノードをマウスでクリックすることにより、解析結果を数字で読むことができる。

これらの画面出力は、色と数字で端的に表現されるため、分かりやすく、かつ説得力をもつので、コンピュータに馴染みの薄い小売業者にとっても、魅力的な支援ツールとなる。

(5) データ解析結果に関する評価

解析結果については、系数的な問題を更に解決すべきであるが、基本的には通常経営者が業界常識と経験によって判断するのと同様な形のパターンになったと思われる。今後の課題は後述するが、プロトタイプのレベルとしては、概ね評価できるとと思われる。

11.4 今後の課題と結論

11.4.1 今後の検討課題

今回作成したプロトタイプ「適地候補ゾーン」抽出プログラム、並びに「適地候補地点」の評価プログラムに関して、今後更に、次のような課題に取り組む必要がある。

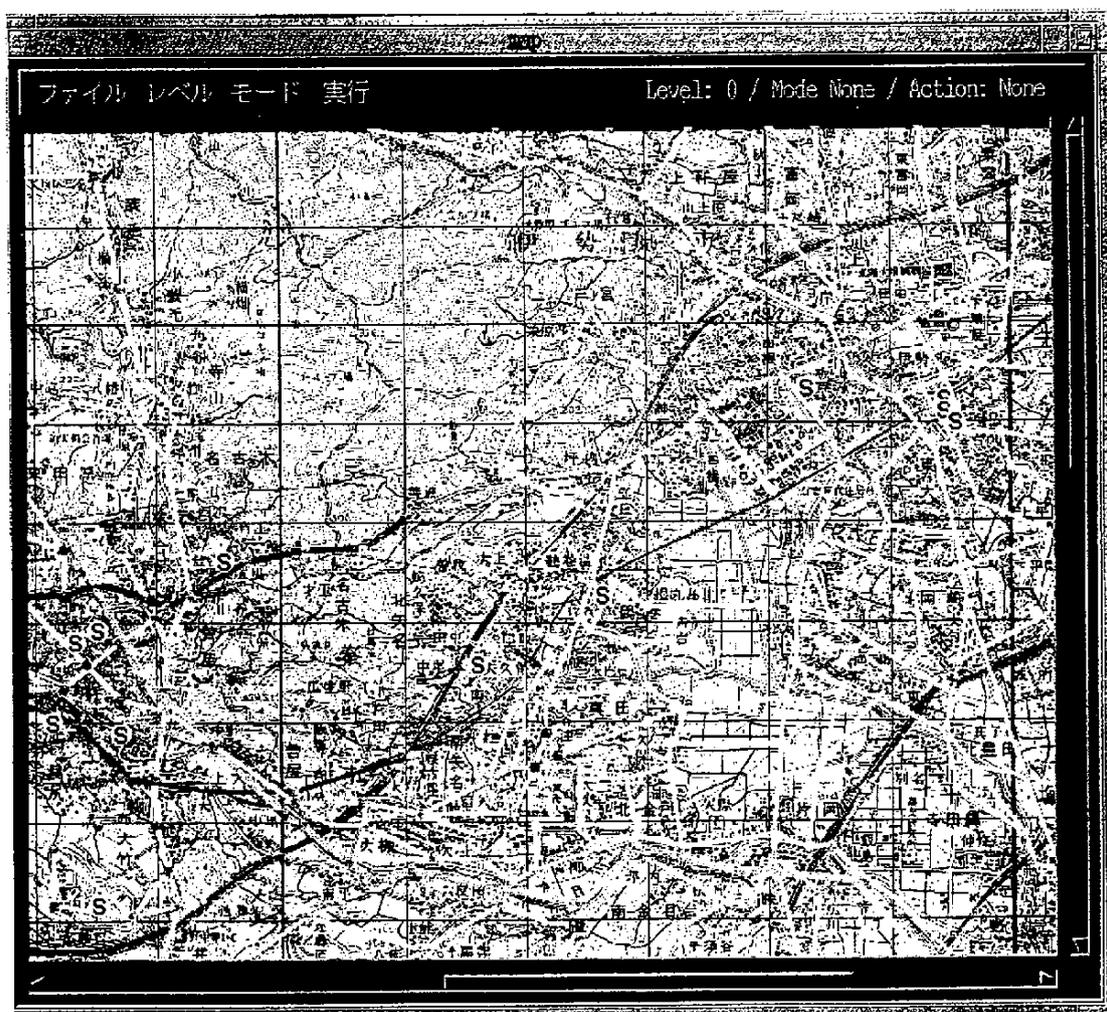
- ① 他の評価基準の採用検討
- ② メッシュの細分化による近似精度の向上とその効果の検証
- ③ メッシュ間グラフと道路グラフの結合に関する工夫
- ④ 交差点における右折、左折、直進の区別
- ⑤ 距離による集積人数の補正
- ⑥ 付加情報レベルの利用
- ⑦ 計算時間の短縮
- ⑧ メッシュの付加情報区分の利用
- ⑨ 自動車でのアクセスの他、自転車や徒歩でのアクセスの導入
- ⑩ 各種換算パラメータの妥当性の検証

11.4.2 結論

本データベースを将来的に活用するユーザ層として、コンピュータにあまり馴染みのない中小店のオーナー等も念頭においているため、分かりやすくビジュアルに表現することが外せない要件である。そこで、文字情報、画像情報等を融合したマルチメディアを用いて開発する必要があり、その結果、地図データという画像情報を取り扱うだけでも処理データ量が大きく、EWSレベルが最低必要になる。しかし、最近のダウンサイジングの流れの中でEWSが普及してきているため、価格的にも入手可能な範囲になっていることは、今後の事業化を目指す際に明るい材料である。

本年度は時間的制約上、11.4.1に指摘したような課題については実現していないが、次期システムにおいて改善する目途をつけている。

本年度のプロトタイプを基礎として、更に地域流通最適化データベースを充実させ、これを活用することにより、各々の地域流通において、小売店集団が最適立地に商業集積の形成を行うことができれば、小売店主の立地選定に関するメリットがあるだけでなく、地域商業は近代化し、その結果、地域経済振興に貢献し、地域消費者への福祉を増進することができる。これは大変意義の深いものであり、引き続き完成を目指して研究を進めたいシステムである。



Exec Menu

確認 取消

ファイル名 _____

半径 (km) _____

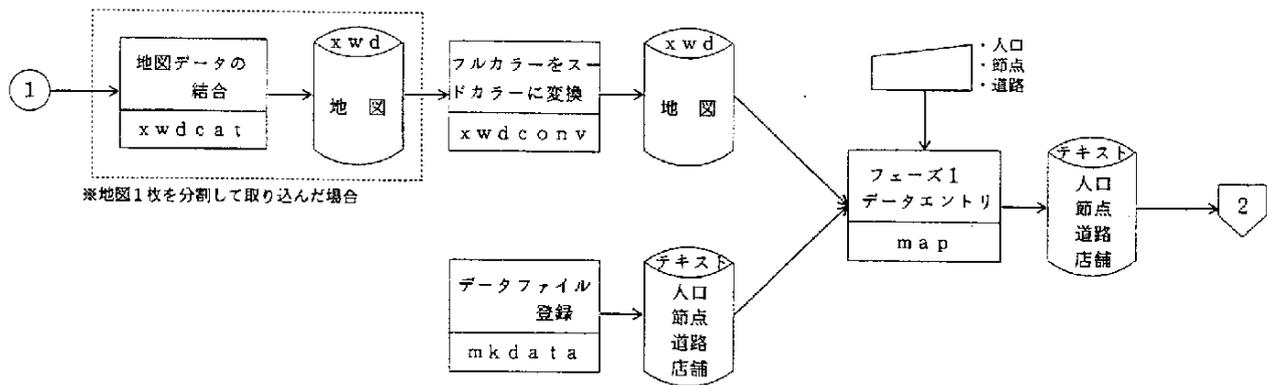
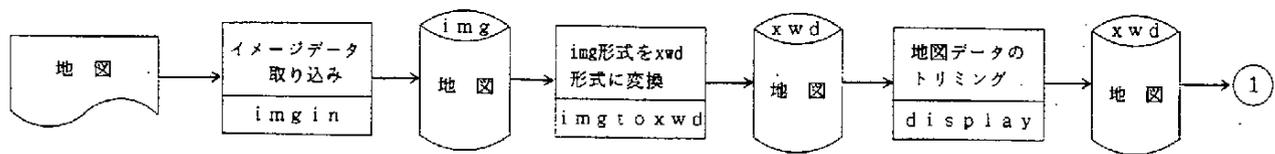
人口 (人) _____

Exec Menu (Phase 2)

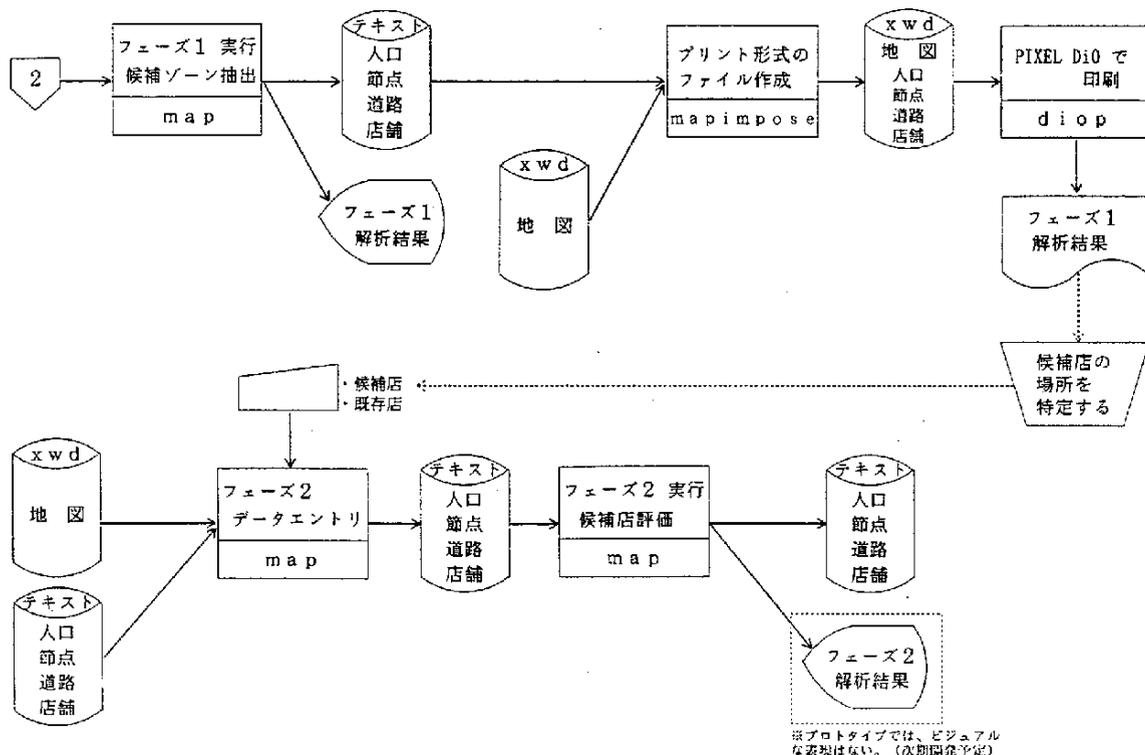
確認 取消

ファイル名 _____

図11-1 入力画面



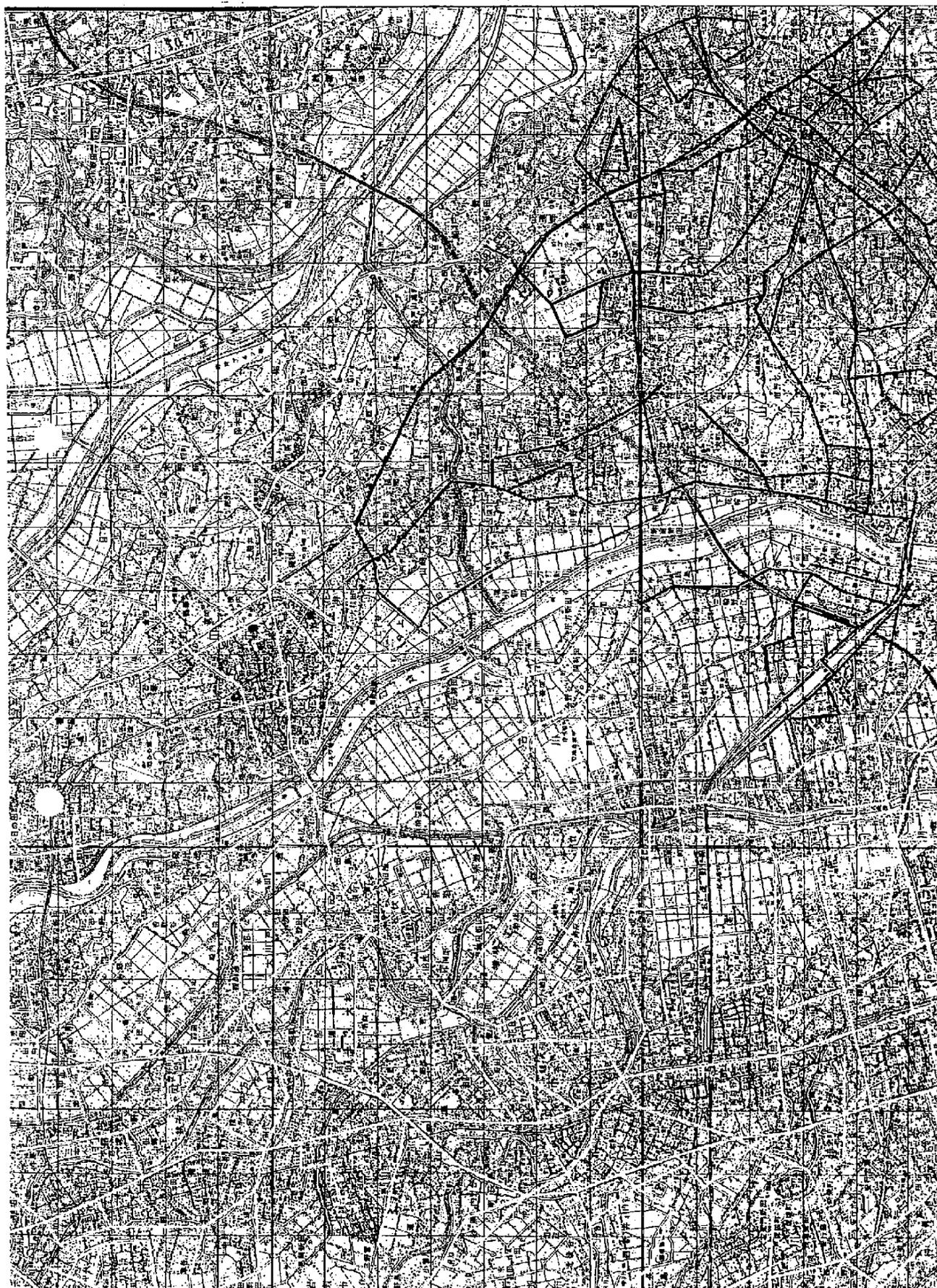
処理フローチャート 1/2



処理フローチャート 2/2

図11-2 処理フローチャート

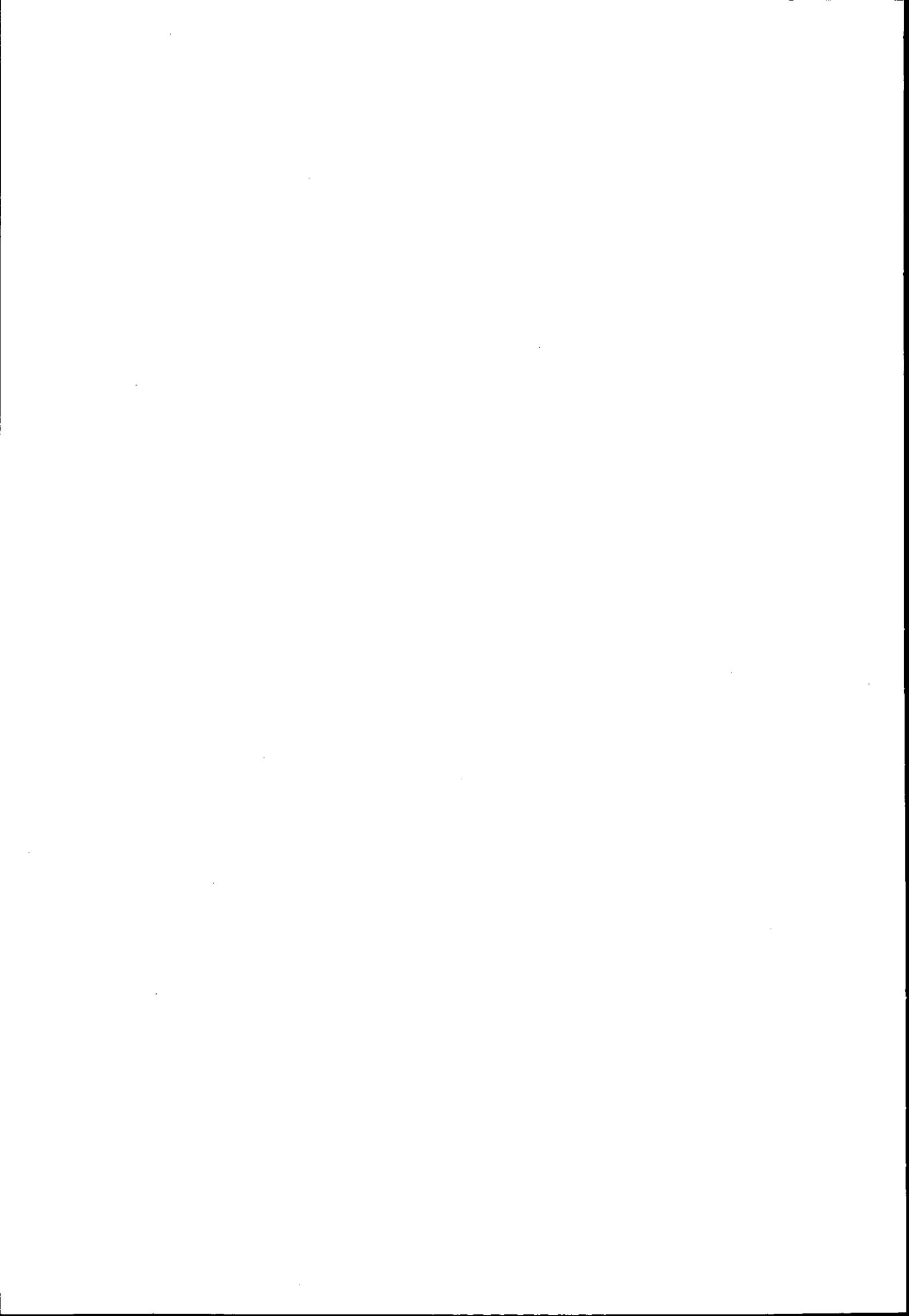
(参考) フェーズI 出力





12 異分野研究のための知的オリエンテーション・
データベースシステムのプロトタイプ作成

株式会社 けいはんな



12. 異分野研究のための知的オリエンテーション・データベースシステムのプロトタイプ作成

12.1 はじめに

産業技術の急速な発展などにより、企業の異分野進出のニーズは益々高まっている。このため、これまでの自分の専門分野以外の分野について学習、勉学あるいは研究する（以下、単に学習あるいは研究するという）必要性の生じた人々が急速に増加している。そこで、まず異分野学習のために必要となるさまざまな知識・情報をデータベース化したデータベースを構築する。そして、このデータベースをもとに、異分野の学習を効率的に行うことができるように、知識工学的手法を活用して、適切なオリエンテーションやガイダンスのできる知的オリエンテーション・データベースシステムを構築すると、異分野の学習効率が向上し、人材の活用や人材の配置・転換の効率化が促進されるとともに、異分野の研究者・学習者相互の交流も促進され、企業活動や産業の振興に役立つことになる。しかし、このようなシステムはこれまでに類例のない新しいデータベースシステムとなる。

本研究の目的は、このような異分野研究のための知的オリエンテーション・データベースシステムのプロトタイプシステムを構築し、その稼働性を実証するとともに、本格システムの実現に向け、利用上の問題点、操作性、サービス内容等について検討・修正を行うことである。

12.2 プロトタイプシステムの開発

(1) 対象データ・情報

異分野学習あるいは研究を行うに当たって重要となる知識・情報・ノウハウの提供サービスを行う知的オリエンテーション・データベースシステムの実現には種々のデータ・情報が必要となる。システムの有用性の向上の観点からはデータの収集の範囲は広ければ広いほど望ましい。しかし、システムの構築・運営の経済性の観点からは収集範囲はできるだけ少ない方が望ましい。そこで、このトレードオフを考慮して、本システムが提供すべきデータ・情報の当面の目標を次のようにすることにした。

- (a) 大学・学部・学科及び大学院大学・研究科・専攻に関する情報
- (b) 各大学・学部の履修要覧に記載の情報
- (c) 各大学院・研究科の履修案内に記載の講義・講座及び研究者情報
- (d) 各大学の学位授与情報と博士論文

(e) 文献情報紹介者、研究者へのアンケートあるいはインタビューなどによる論文発表情報

(f) キーワードによる検索における同義語の検索もれを防ぐため、専門分野にも対処可能な同義語辞書（キーワードは英語化されているものもあるので、科学技術用英和辞書も必要）。

プロトタイプシステムの構築に当たって収集したサンプルデータは17種、約82キロバイトである。

(2) データベース管理システムの開発言語とスキーマ

異分野学習あるいは研究のための知的オリエンテーション・データベースシステムは12.1節で述べたようにこれまでに前例のないデータベースシステムとなる。そこで、昨年 of 調査研究で明らかにした知的オリエンテーション・データベースシステムの各種案内サービスのイメージをもとに、ワークステーション (SUN-SparkStation) 上にプロトタイプシステムを構築することにした。

本データベースシステムの構築に当たっては知識工学的手法の適用の容易さを考慮して、システムの構築言語としてPrologを用いることにした。対象データの構造は多種多様な構造をもつが、対象データをPrologにおける事実として記述できることを明らかにした。また、Prologの不定長のリスト機能を利用することにより、複雑な構造をもつデータでもそのままに近い形で容易にデータベース化できる。

対象データは一般に

$$\text{datal} (A_1, A_2, \dots, A_n).$$

と表すことができる。ここで、 A_1, A_2, \dots, A_n は対象データの各項目を表す。Prologでは、ユーザにより A_1, A_2, \dots, A_n のうち、幾つかの項が指定されたとき、他の項を出力するのはこの言語の原理上の基本操作となっているので、このような検索については何の問題もない。

以下では、単独のデータでは検索できないデータの検索法について述べる。次のようにデータモデルにより定義された二つのスキーマがあるとする。

$$\text{datal} (A_1, A_2, \dots, A_n).$$
$$\text{data2} (B_1, B_2, \dots, B_m).$$

ここで、 m は一般には n と異なる正整数である。 datal の k 番目の項目 A_k と data2 の j 番目の項目 B_j が関連をもっている場合、 datal と data2 の両方の情報をもつ新たなデータ $\text{data1} \&2$ をPrologの定義を以下に示すように行うことにより、新たなサブスキーマとして生成することができる。

$data1\&2 (A1, A2, \dots, An, B1, B2, \dots, Bm) :-$
 $data1 (A1, A2, \dots, Ak, \dots, An),$
 $data2 (B1, B2, \dots, Bj, \dots, Bm), def(Ak, Bj).$

ここで、 $def(X, Y)$ は X と Y の関連を示す Prolog の定義又はデータである。 X と Y は直接的な関連が無くとも、Prolog の特徴である推論機能を用いて関連づけることもできる。例えば、 A_k が学部や学科などの組織・機関の情報を表し、 B_j が研究者名を表す情報とすると、ある学科のある研究者という関連を示す情報が $def(A_k, B_j)$ である（極端な関連の例として A_k と B_j が同一のときは、 A_k と B_j をキーとして $data1$ と $data2$ の両方の情報をもつ新たなデータ $data1\&2$ を得ることができ、この場合は $def(A_k, B_j)$ の項は上式の定義式においてなくてもよく、 $data1\&2$ の中の B_j も省略できる）。この定義式により、例えば $data1$ のある項目の入力情報に対し（部分情報も可）、 $data1$ の情報だけでなく $data2$ の情報をも得ることができるようになる。

以上の定義を繰り返し行い、以下のようなスキーマを生成できる。

$all_data (A1, \dots, An, B1, \dots, Bm, C1, \dots, D1, \dots).$

上のスキーマにおいて、ユーザにより幾つかの項が指定されたとき、ユーザの必要とする他の項を出力できることは明らかである。

上に説明したように、本システムでは、収集されたデータはすべてそのまま知識ベースとして利用可能な形態となるので、単独でも多様な利用が可能となる。更に、ユーザの要求する情報を種々のデータを組み合わせて有機的に新たに生成し、提供することも可能となる。必要となる組み合わせデータは要求されたときに生成されるので、データとして記憶しておかねばならない情報は顕著に節約され、記憶容量がわずかで済む。このほか、単独のデータでは提供不可能な高度な内容の情報をユーザが必要としている場合でも、システムが種々のデータを組み合わせて加工することにより作り出して、提供できるという特徴をもたすことが可能となることが分かる。

(3) 逐次メニュー方式による案内サービスシステムの開発

従来のデータベースのコマンドやキーワードによるデータの検索は検索の効率は高いものの、初心者には必ずしも便利とはいえない側面をもっている。本システムでは、データベースの性格上ユーザは利用が初めてでデータベースシステム利用法を知らない場合が多いと予想される。このため、本システムの初めての利用者でも本システムの機能を十分に使いこなせるように、知識工学的手法を適用して、システム利用に関するノウハウ的知識のデータベース化を図ったユーザフレンドリーなシステムを実現することにした。まず、初心者向けの検索方式として、逐次対話型のメニュー方式によるデータベース利用により、初めての人で

も必要とするデータを容易に検索できるシステムとするための種々の創意工夫をこらした検索方式を開発した。本システムでは階層的に順次提供される各メニュー画面において、検索支援関連情報を知識ベース化した知識ベースに基づいて、必要と考えられる情報が一覧表示されるようになっているので、キーワードによる検索を除いて、基本的にはシステムのメッセージに従って該当の数字を入力していけば希望の情報を知ることができる。

構築した逐次メニュー方式のデータ案内サービスシステムは以下のとおりである。

- (a) 大学構成学部名や学科名、学科定員数などを紹介する大学・学部・学科案内サービス
 - (b) 各大学の講座の研究スタッフ、研究分野及び現在進行中の研究テーマについて紹介するだけでなく、研究テーマについてのキーワードによる講座の逆検索も行うことができるような大学の講座案内サービスシステム
 - (c) 第一線で研究している大学の助手を含む研究者、学者の所属、地位、研究テーマなどの紹介を行うことのできるような研究者案内サービスシステム
 - (d) 大学、学科などを指定すると、その学科における授業科目名、必要単位数、授業内容の概要などを紹介する科目履修案内サービスシステム
 - (e) 著者名、学位論文名及びその目次に使われている術語をすべてキーワードの対象とすることができるような学位論文案内サービスシステム
 - (f) 各研究者による各自の専門分野における発表論文群などの文献紹介を行うことのできる専門分野（文献）案内サービスシステム
- (4) 一括対話方式による案内サービスシステムの開発

本システムで採用したメニュー方式によるデータベースの利用は初心者でもデータベースを容易に利用できるという利点がある反面、2回目以降の検索のときはメニューの階層をたどるのが面倒であったり、人によっては思考に合わない部分があったりする。また、あらかじめ想定したデータの検索しかできないという短所もある。このため、本システムでは、メニューを横断的・垂直的に一括してたどって検索を行うことのできる一括対話方式による対話型案内方式も開発し、提供可能な情報の中の各項目を任意に組み合わせた情報を簡単に検索できることを可能とした。

構築した一括対話方式による案内サービスシステムは、機関・組織の属性案内、人（研究者）の属性案内及び学習・研究分野案内の3種類のシステムであり、学習・研究分野案内は更に科目履修案内、学位論文案内及び専門分野（文献）案内の3種類のシステムに分かれている。

図12-1に本システムの各種案内サービスの概略図を示す。

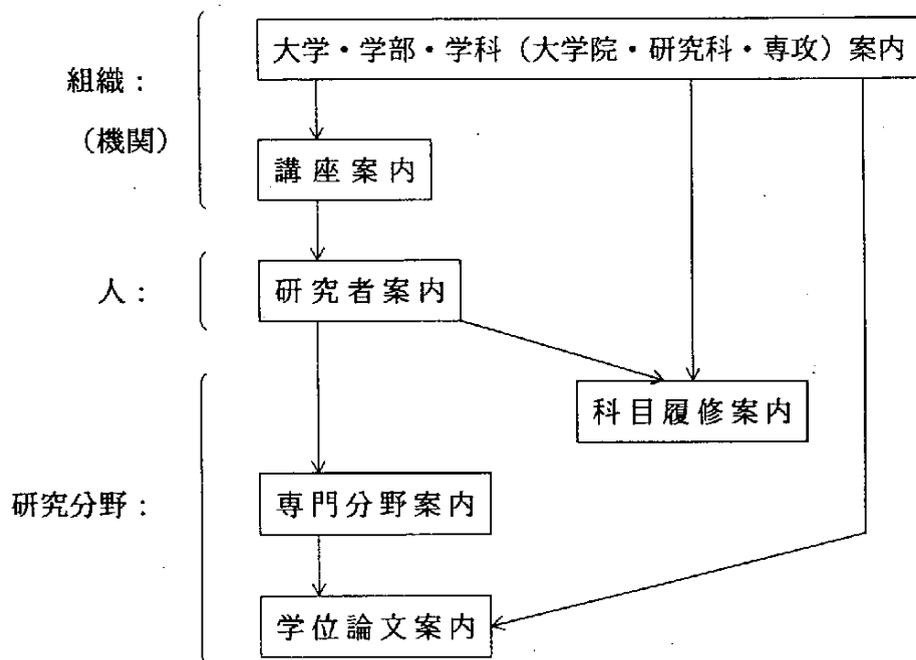


図12-1 本システムの各種案内サービスの概略図

図12-2に開発した本プロトタイプシステムの構成を示す。図に示すように、本システムでは、イニシャルメニュー（初期メニュー）から逐次メニュー方式あるいは一括対話方式のメインメニューにジャンプし、ここから各種の案内サービスを受けるように構成されている。

各種の案内サービスを受けるに当たっては、種々の支援をシステムより受けれるように、システム内には種々の知識ベースが構築されており、必要に応じて随時参照されるようになっている。構築された知識ベースは、府県名に関する知識ベース、市町村名に関する知識ベース、地位に関する知識ベース、著者の英文表示の知識ベース、専門用語同義語及び和英同義語の知識ベースの5種類である。

(5) インテリジェント機能

本システムに組み込まれたインテリジェントな機能は以下のとおりとなる。

- (a) 収集されたデータはすべてそのまま知識ベースとして利用可能な形態となるので、単独でも多様な利用が可能となる。更に、ユーザの要求する情報を種々のデータを組み合わせで有機的に新たに生成し、提供することも可能となる。必要となる組み合わせデータは要求されたときに生成されるので、データとして記憶しておかねばならない情報は顕著に節約され、記憶容量がわずかで済む。このほか、単独のデータでは提供不可能な高度な内容の情報をユーザが必要としている場合でも、システムが種々のデータを組み合わせで加工することにより作り出して、提供できるという特徴をもたすことが可能となった。短所としては、生成に時間を要するので、情報検索の応答性が悪くなる欠点がある。ただし、プログラムの構築が簡単でデータの保守・更新が容易となる利点がある。

- (b) 不確定文字?や不確定文字列*を使用することができるので、人名や国際学会名など曖昧な情報、曖昧な記憶しかない場合にも容易に情報検索することができる。また、正確な情報を有している場合には、入力の省力化が図れる。
- (c) 複雑な論理式をそのまま解釈できるので、高度な論理式を用いて、効率よく必要情報を検索できる。
- (d) 知識ベースとして専門用語の同義語辞書及び和英同義語辞書を持ち、同義語をユーザに提示し、使用するかどうか確認するようになっているので、キーワード検索時に検索もれが生じにくい。

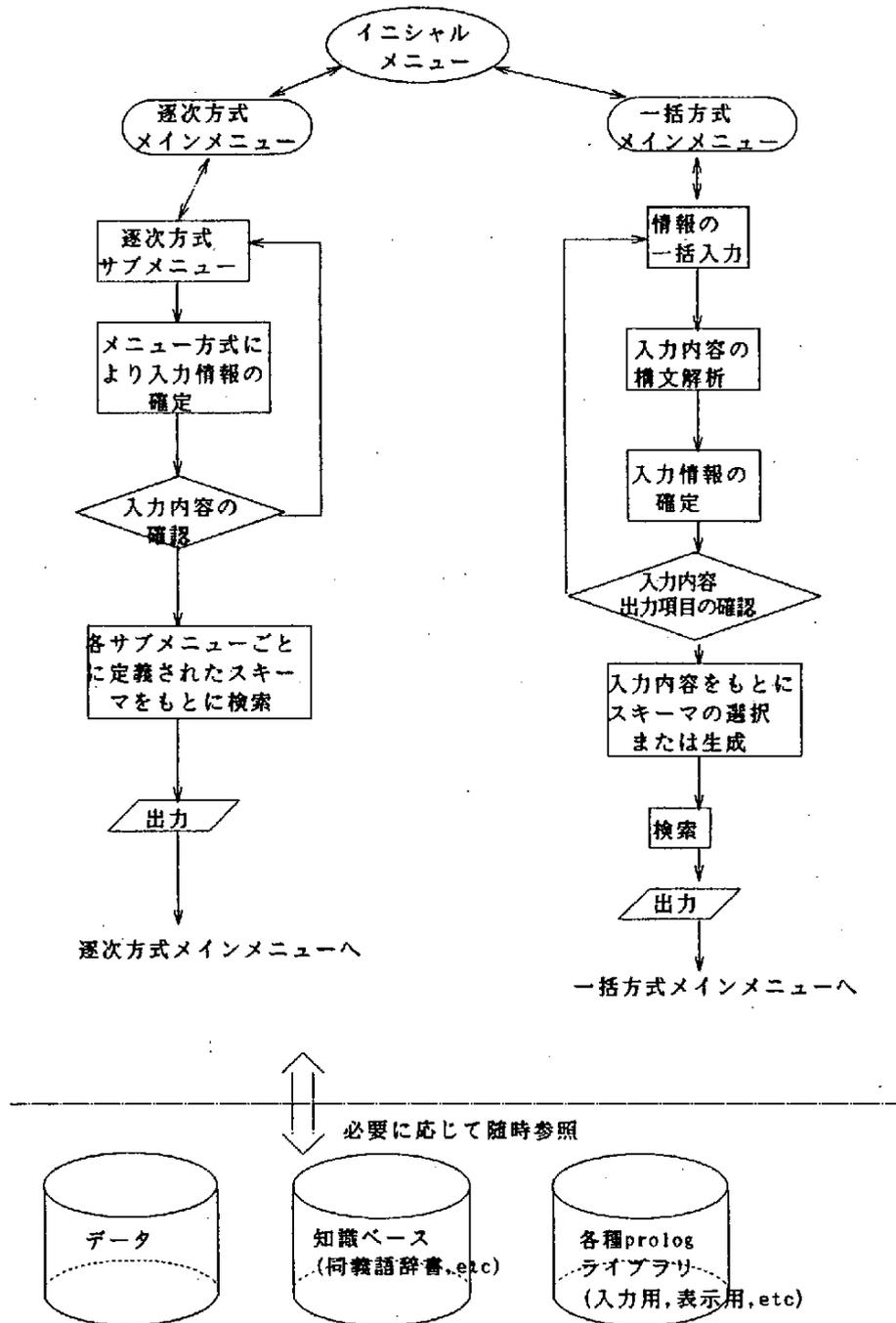


図12-2 開発した本プロトタイプシステムの構成

- (e) 知識ベースとして著者名の英文表示辞書をもっているのに、日本語の著者名から英文の文献も検索できる。
- (f) 知識ベースとして、府県名や市町村名に関する辞書や人の地位に関する辞書を有しているので、利用が便利となる。
- (g) 一括対話方式では、“大学”、“学部”、“学科”など、システムが判定できる語句は、ユーザが入力しなくても、システムが自動的に判定するようになっている。また、不必要な項目については、入力を打ち切ってもシステムが不要であると判定できるようになっている。これらの機能はユーザに余計な負担をかけないため、ユーザがスマートに本システムを使いこなすことを可能とする。

12.3 逐次メニュー方式及び一括対話方式による案内サービス例

構築した逐次メニュー方式及び一括対話方式による案内サービスシステムを利用して、実際にどのようにして、どのようなサービスを受けられるかについて、具体的な案内サービス例を取り上げ説明した。

図12-3に本案内サービスシステムの逐次メニュー方式のメインメニューを示す。

異分野研究のための
インテリジェント・オリエンテーション・データベースシステム

メインメニュー

現在のインテリジェント・オリエンテーション・データベースシステムのサービス内容は次のとおりです。番号で選んでリターンキーを押して下さい！

機関・組織案内

1 大学案内 (U)	1 1 大学院案内 (GU)
2 学部案内 (F)	1 2 大学院研究科案内 (GF)
3 学科案内 (D)	1 3 大学院専攻案内 (GD)
4 講座案内 (L)	

人の属性案内

5 研究者案内 (R)

学習・研究分野案内

- 6 科目履修案内 (S) 16 大学院科目履修案内 (GS)
- 7 学位論文案内 (DT)
- 8 専門分野 (文献) 案内 (SF)

- i (I) イニシャルメニューへ
- 0 終了

KIDS>

図12-3 本案内サービスシステムにおける逐次メニュー方式のメインメニュー

図12-4 は本案内サービスシステムの一括対話方式のメインメニューを示す。

異分野研究のための

インテリジェント・オリエンテーション・データベースシステム

一括対話方式

知りたいことはなんですか？ アルファベットを入力してリターンキーを押して下さい！

知りたい対象

1 人の属性

n : 勤務先・所属情報, 研究者姓名, 地位 (教授, 助教授, 講師, 助手)

研究テーマなど

2 機関・組織の属性

o : 大学, 学部, 学科名, 学科定員数, 講座名, 講座の構成, 講座の活動 (研究
テーマ), 法人形態 (国公立の別), 住所・連絡先など

3 学習・研究分野の情報

k : 履修科目情報

d : 学位論文

p : 文献 (論文)

i : イニシャルメニューへ戻る

0 : 終了

KIDS>

図12-4 本案内サービスシステムにおける一括対話方式のメインメニュー

図12-3及び図12-4のメインメニューより、種々のデータベース検索サービスを受けることができる。本システムではシステム自体がインテリジェントな機能を有しているので、初心者にもデータベースの熟練者にも高度な検索方式をスマートに利用できることを種々の具体例で示し、詳細に説明したが、ここではその内容については紙面の制約上省略する。

12.4 今後の課題と方針

本研究の今後に残された課題と方針を述べると次のようになる。

(1) データの入手体制の確立と著作権の問題

必要情報を収集してデータベース化し、データベース検索サービスを提供するに当たっては、各情報の入手体制の確立のほかに、データの著作権の問題についても十分検討する必要がある。

(2) 自然言語（日本語）による対話方式の開発

本システムでは、メニューを横断的・垂直的に一括してたどって検索を行うことのできる一括対話方式による対話型案内方式も検索方式として構築し、提供可能な情報の中の各項目を任意に組み合わせた情報をハイスピードで検索できることを可能とした。しかし、この方法では検索を行う際に情報の各項目の入力の順序が規定されるという難点がある。実際には、情報検索を受けたい人が自由な発想によって、まるで人に尋ねているような感じで検索項目の入力の順序には制約されずに検索することを可能とすべきであろう。人間の思考に近い自然言語的な対話による検索を可能とする日本語検索システムを構築することが必要になると考えられる。

(3) ユーザサービス機能の強化

(a) ユーザによる直接書き込み機能の強化 本システムの文献案内においては、最新データへのアクセスが可能となるようにするため、文献情報の紹介者がデータを直接書き込むことを可能とした。本システムにおいては、同義語辞書ファイル及び科学技術用英和辞書ファイルが備えられており、その内容が情報検索に活用される構成となっているが、同義語辞書及び科学技術用英和辞書ファイルの一層の充実のためには、ユーザによる直接書き込み機能をもたせることも同様に必要になると考えられる。

(b) ヘルプ機能の充実 構築した本プロトタイプシステムでは、最小限のヘルプ機能は有しているが、オンラインマニュアルやオンラインの検索例表示の提供サービスなど、ヘルプ機能を更に充実させる必要がある。成果報告書で示した対話例を本システムの詳しい対話の仕方の説明としてオンラインマニュアルの一つとして組み込む方針である。

(4) 今後の方針

異分野研究のための知的オリエンテーション・データベースシステムの本格的サービスの実施としては、当社(株)けいはんなの提供している情報ネットワークサービス「けいはんなネット」を通じて行うことを予定している。サービスするデータ・情報としては12.2節で挙げた(a)から(f)までが考えられるが、最初はこの中の(a)大学・学部・学科及び大学院大学・研究科・専攻に関する情報、(b)各大学・学部の履修要覧に記載の情報、(c)各大学院・研究科の履修案内に記載の講義・講座及び研究者情報の中で、大学と学部を限って行うことが経済性の点からは得策と思われる。

サービス実施のためには、本格的システムを構築することが必要となるが、そのためには商業化に向けてシステムの使い易さを一層洗練する必要がある。このため、既に構築されているプロトタイプシステムを利用して試験的なオンラインサービスを行い、本委員会の委員による存在し得るかも知れないバグ取りを行うほか、システムの操作性をより向上させるための意見を収集する必要があると考えられる。

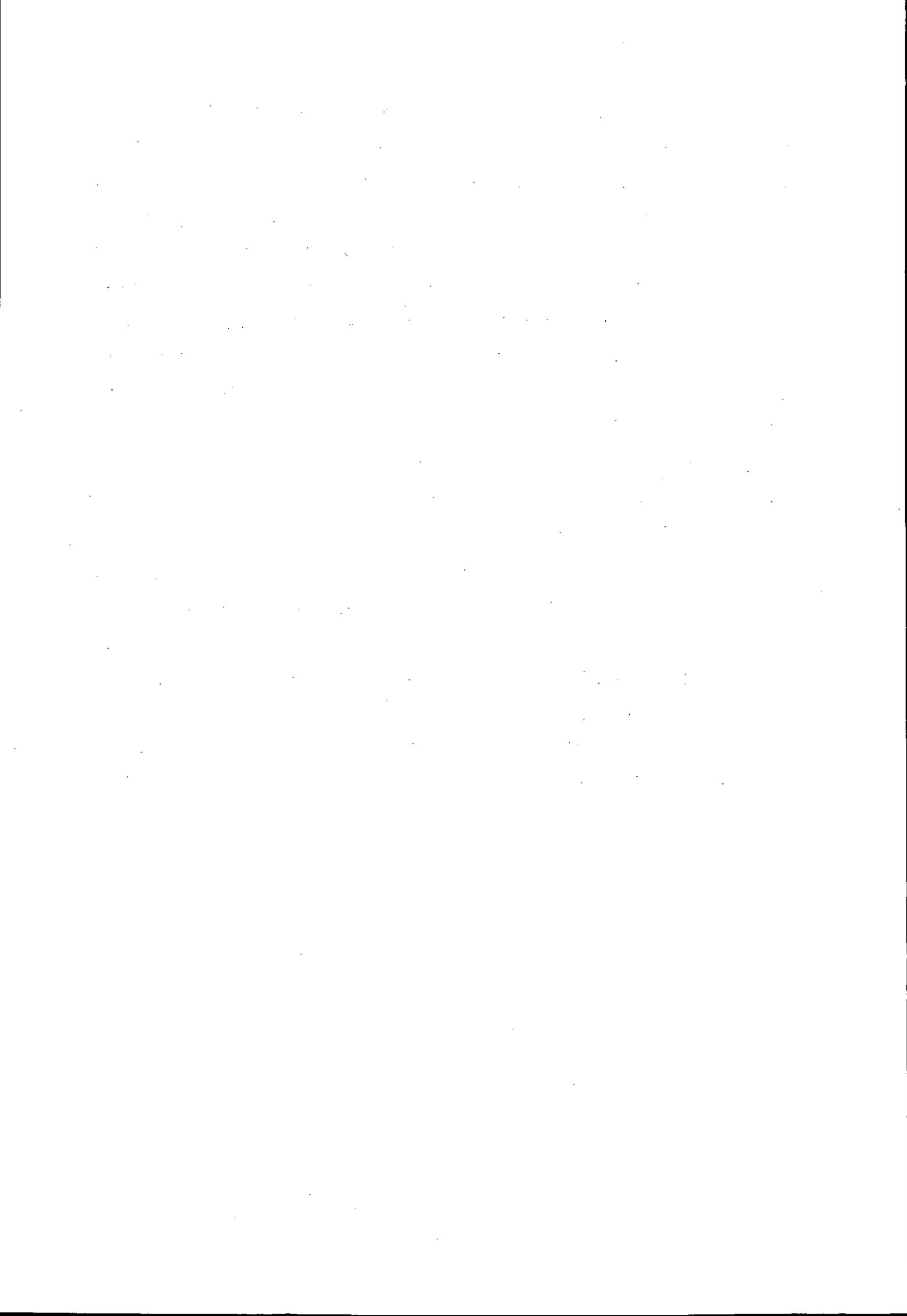
更に、本システムはサンワークステーション上にスタンドアロンで構築されているが、これを「けいはんなネット」上で稼働できるように移植する必要もある。また、データベースシステムの維持・管理・更新を行うことは、コスト的にもマンパワー的にも多大の負担がかかる。このため、これをサポートする機能を強化することも重要な課題となる。これらも、今後に残された課題である。

12.5 おわりに

本研究において行った調査研究内容及び得られた成果より、(1)異分野学習あるいは研究を行

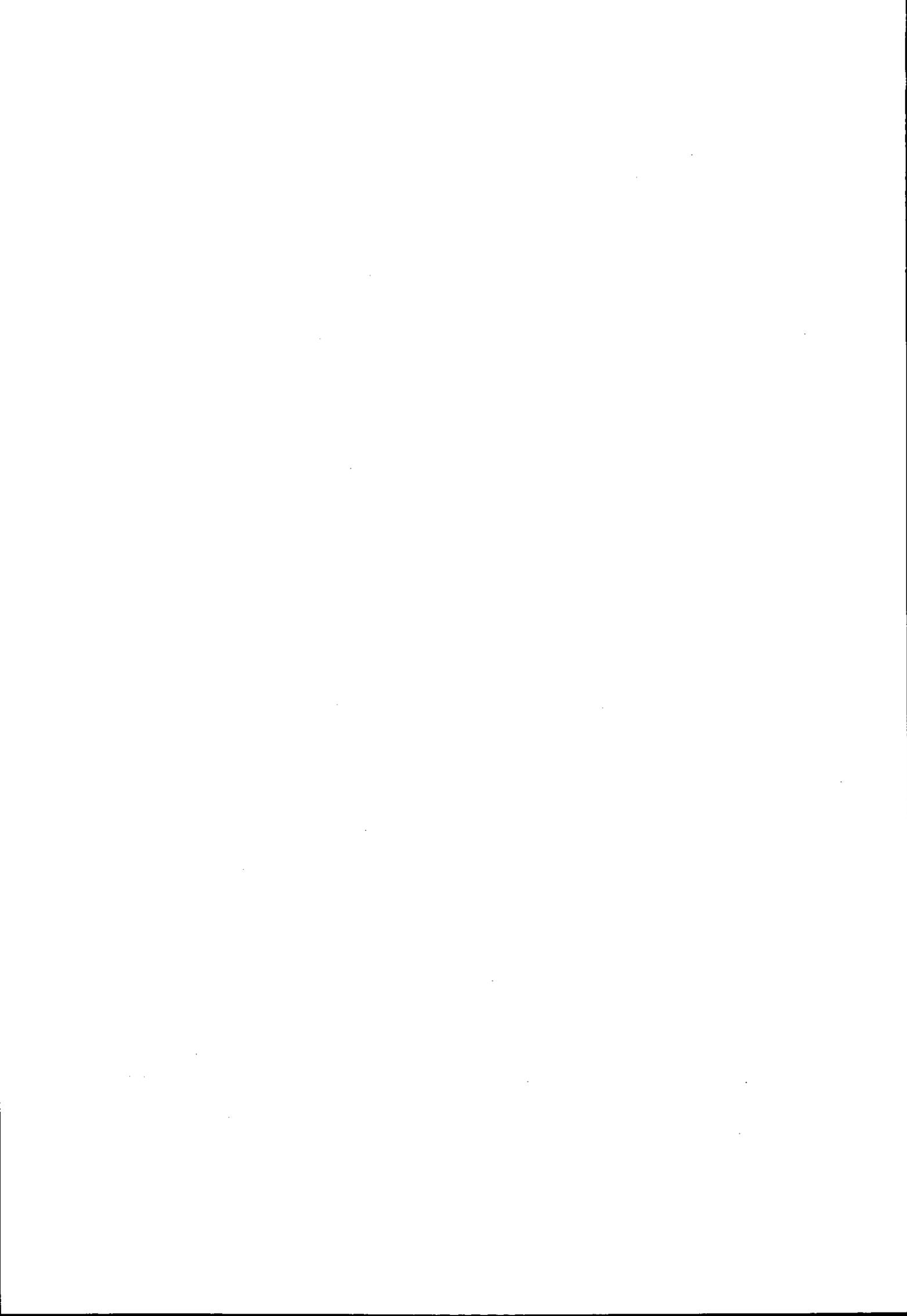
うに当たって重要となる知識・情報・ノウハウの提供サービスを行う知的オリエンテーション・データベースシステムのサービスの内容や有用性を明らかにすることができた。(2)複雑な構造をもち、属性も多種にわたり、形式が一定でないデータを、Prologを用いることにより、データをそのままに近い形で容易にデータベース化することが可能となった。(3)本提案データベースシステムの本格的システムの実現のもととなるプロトタイプシステムが構築された。プロトタイプシステムの構築に当たっては、(a)各種知識ベースをもたず、(b)データ自身も知識ベースとして利用する、(c)Prologの推論機能などの言語自身の特徴を生かすことにより、インテリジェントで、ユーザの使いやすいデータベースシステムとすることができた。(4)プロトタイプシステム構築の過程で、本システムの操作性、サービス内容等に多くの修正を加えるとともに、改善を図ることができた。

今後予想される効果を列挙すると以下のとおりである。データベースサービス実施に向けた本格的システムを構築すると、委託者の提供している情報ネットワークサービス「けいはんなネット」で、実際のサービス化を行うことができるようになる。また、本提案システムが本格的に構築され、データ案内サービスが実施された場合(1)大学と企業の研究者相互の交流が促進される。(2)大学や公的な各種研究機関で研究している研究者にも有用なデータベースシステムとなり、わが国の先端技術の一層の開発に貢献することができるようになる。(3)生涯学習を行う上で有用な指針を提供できるシステムとして利用できる。(4)大学へ進学する高校生の合理的な大学選択のための重要な情報を与えることもでき、国際ネットワークで利用できるようにすると、海外で、わが国への留学生が留学先の大学や研究室を選定するのに利用できることで、国際的な情報化へ多大な貢献ができることになる。



13 在宅勤務者サポート・データベースの構築調整

株式会社 志木サテライトオフィス・ビジネスセンター



13. 在宅勤務者サポート・データベースの構築調査

13.1 概 要

「在宅勤務」は、情報社会における特徴的な働き方になると予想されているが、日本の将来における、その普及の道筋が明確に示されているわけではない。「在宅勤務者サポート・データベース」は、就業者の意識やニーズをふまえた上で、就業者側及び在宅に適した仕事を有している企業や個人にとって、真に役立つデータベースとして構想されたものであり、それは情報社会の健全な発達をより促進させるデータベースとなると考えられる。

本構築調査では、上記のようなデータベースを成立させるための諸条件を子細に検討し、そこでの問題点を整理した上で、その解決策となる、データベース構築に当たっての企画案を創り出した。

13.2 在宅勤務者サポート・データベース構築の背景と意味

アメリカにおける在宅勤務者の実数は、600万人以上といわれるが、この数年間は年々増加傾向にあり、今後とも増える勢いは衰えない。日本では、アメリカとマネージメントの仕方や企業風土、国土の状況等が異なり、さしあたっては正規社員でありながら在宅勤務を行うというテレコミューター（電子通勤をする人）は、そう急激に増加するとは思えないが、しかしながら、主婦層、高齢者層、身障者層、大学院生層、フリーランス層といった各層において、在宅勤務に対してのニーズが高い。

特に、日本独特のM字型就業曲線を描く主婦層における「子育て中の就業意欲の高さ」や、また高齢者層の「経験を活かしたホワイトカラー的職業への就業志向」は根強いものがある。

その一方で、仕事を発注する企業側も、大手中堅企業では新規事業立ち上げ時の短期契約社員に代表されるような頭脳労働系業務を行う在宅勤務者（ネットワークカー）、また中小零細企業においては、名刺入力等に代表されるようなビジネスサポート・代行系業務を行う在宅勤務者に対してのニーズが高い。

ワーカー側のニーズと、クライアント企業側のニーズを、効率よく結びつけ、双方にとってメリットが大きい仕組みが創造されれば、埋もれている潜在労働力の発掘とそれに伴った企業活動の発展へとつながるといえる。在宅勤務者サポート・データベースは、そうした役割を担うものとして位置づけることができる。

また、「在宅勤務者サポート・データベース」が広く普及し、多くの人々が実際に在宅勤務

を行うようになれば、結果としてそれは一極集中の是正にも寄与することとなる。通信等を用いた遠隔地での勤務形態（フレックスワーク）や在宅勤務は、分散型就業形態の具現化ととらえられるわけだが、これは集中を抑える効果的な方策なのである。

上記のような点に、「在宅勤務者サポート・データベース」構築の積極的な意味を見いだすことが可能である。

13.3 構築に当たっての問題点の整理と解決への模索

心理的側面検討部会、組織的側面検討部会、経済的側面検討部会、技術的側面検討部会という4検討部会を組織し、四つの側面より在宅勤務者サポート・データベースの問題点について子細に検討し、また学識経験者からの意見聴取を行った。

各検討部会で語られた内容は多岐にわたったが、最も代表的な課題という形で、要約すれば、次のようになる。

- ・ ワーカー一人一人が意欲をもって係わり、成長し続けられる仕組みの創出
（ワーカーの側面にたって本データベースのあり方を検討した心理学的側面検討部会では、ワーカーの成長をサポートする仕組み等を考察）
- ・ 在宅勤務者への継続的な仕事の流れの創出
（仕事を出す側の企業に焦点を当てた組織的側面検討部会では、いかにしたら在宅勤務者へ仕事が継続して流れるかを中心的に考察）
- ・ データベース提供センターの役割の明確化と経営的に成立する仕組みの創出
（本データベースの提供がビジネスとして成立するか否かを論じた経済的側面検討部会では、企業とワーカーを結ぶ効果的な仕組みとデータベース提供センターの役割について考察）
- ・ データベースのコンセプトを最もよく具現化する技術・システムの選択
（技術的側面検討部会では、このデータベースに係わる各人にとって使いやすく、かつ本データベースのコンセプトを最もよく体現する最適の技術やシステムについて考察）

学識経験者としては、テレワーキングの専門家でもある都市計画学者、身障者に対してのコンピュータ教育を研究する教育学者から話を伺ったが、ともに新しい勤務形態に対しての期待が語られ、在宅勤務者サポート・データベース構築の意義が強調された形となった。

13.4 データベースの企画原案の構築とその検証

アンケート調査を実施するに当たり、在宅勤務者サポート・データベースの企画原案を作成

した。

「パソコン通信を用いた人材登録データベース」という概念を基本にすえたものであり、ワーカーのランクの設定や教育システム、データベースを提供するセンターの役割やペイメント等に関して記述したものである。これを在宅勤務に対しての関心が高いと思われる主婦層（ヤングミセス～シニアミセスの年齢階層別3層）、高齢者層、身障者層、フリーランス層（男性、女性）、学生（大学院生）層、各5名ずつ計40名に対してアンケート調査を行った。

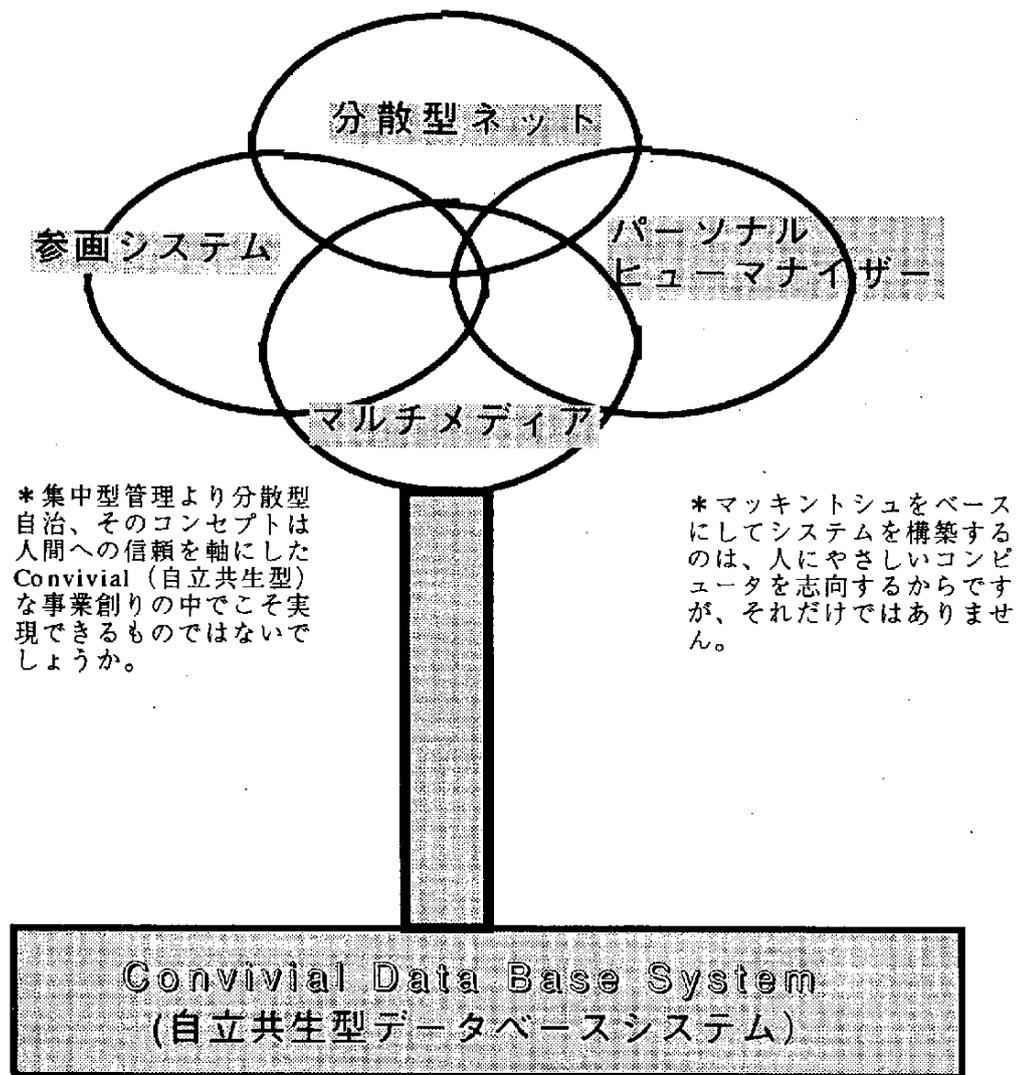
結果は、おおむね在宅勤務者サポート・データベースに対して好意的な回答が多かったが、子細にみると、主婦層で特に在宅勤務に対しての関心が高く、特に30才代半ばまでの若年（ヤングミセス）層で、自らデータベースに登録したいという積極性をもった者が多かった。しかしその一方で、継続して仕事がくるのかといった不安やまたセンターの役割に関してのシビアな意見を記した人も多かった。

男性シニア層は、定年退職後の生活の一選択肢として在宅勤務を考えてはいるが、この在宅勤務者サポート・データベースがさしあたっての重要関心事とはなりにくいようである。フリーランス層にとっては、営業窓口の拡大ととらえる人が多いが、このシステムに横のネットワーク作りを期待している様子も伺える。

ワーカー側としては、登録料に見合った収入が確実に得られるのか、といった点がやはり大きな問題としてとらえられている。

13.5 最終企画案と事業化への方策

最終企画案では、本データベースが、あくまでも在宅勤務者側にたち、個々人のワーカーの成長と利益を優先したデータベースシステムであるということを確認した上で、まずは『自立共生型データベースシステム』という概念を提示し、①分散型ネット、②パーソナルヒューマンライザー、③マルチメディア、④参画システムという特徴を備えたりレーショナル型のオープンシステムを提案した。



*ダウンサイジングやオープンシステム化の流れをワーカーの立場から継承し、情報の双方向通信により相互に成長しあえるビジネス環境創りをめざします。

図13-1 システムの基本コンセプト(1)

事業組織の概略とネットワークのシステム構造は、図13-2、図13-3に示すとおりである。

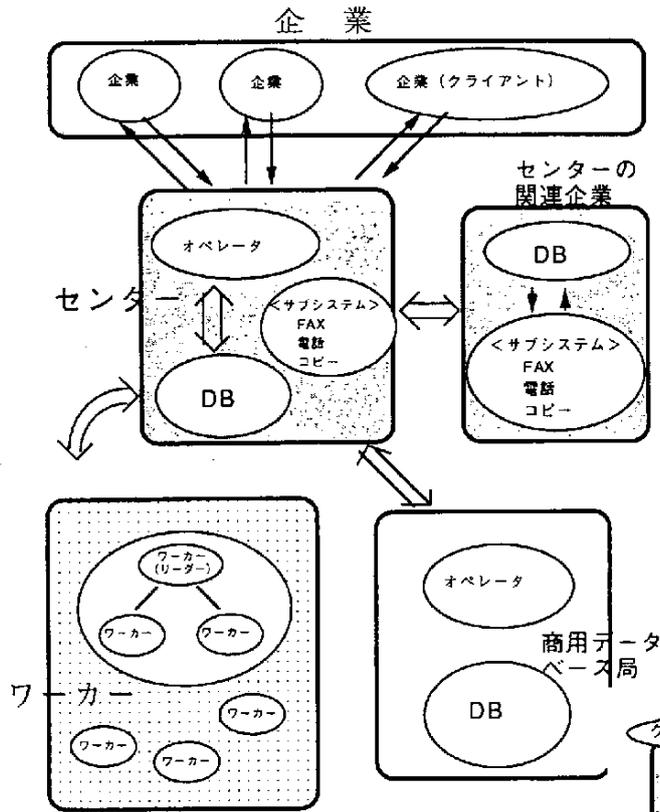


図13-2 事業組織の構造

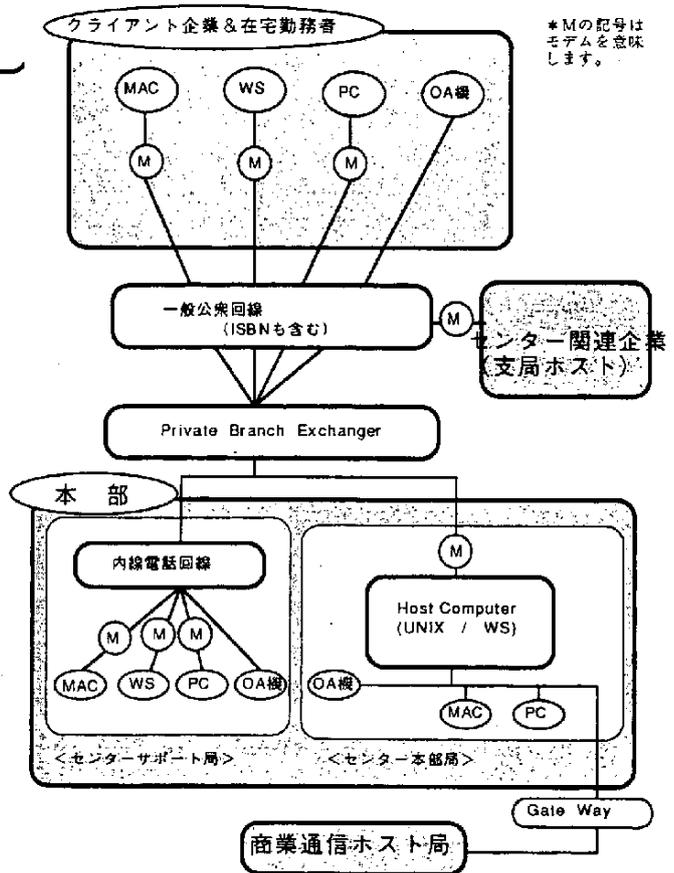


図13-3 DBシステムとネットワーキング構造

データベース自体の中身は、〈雇用情報データ〉〈教育・研修データ〉〈ライフスタイル支援データ〉〈業務・管理データ〉〈研究・開発データ〉の五つの基本的なデータ部分からなり、TEXT系のデータとともに画像系データ、グラフィック系データも蓄積できるように設計される。またそれに併設して会員のSIGコーナー等からなる〈交流コーナー〉が開設される。

- 〈交流コーナー〉
- ① センター事務局からお知らせコーナー
 - ② 会員SIGコーナー
 - ③ 企業情報コーナー（電子メールで各メンバーが自由にアクセスして雇用情報のごく一部を知ることができる）

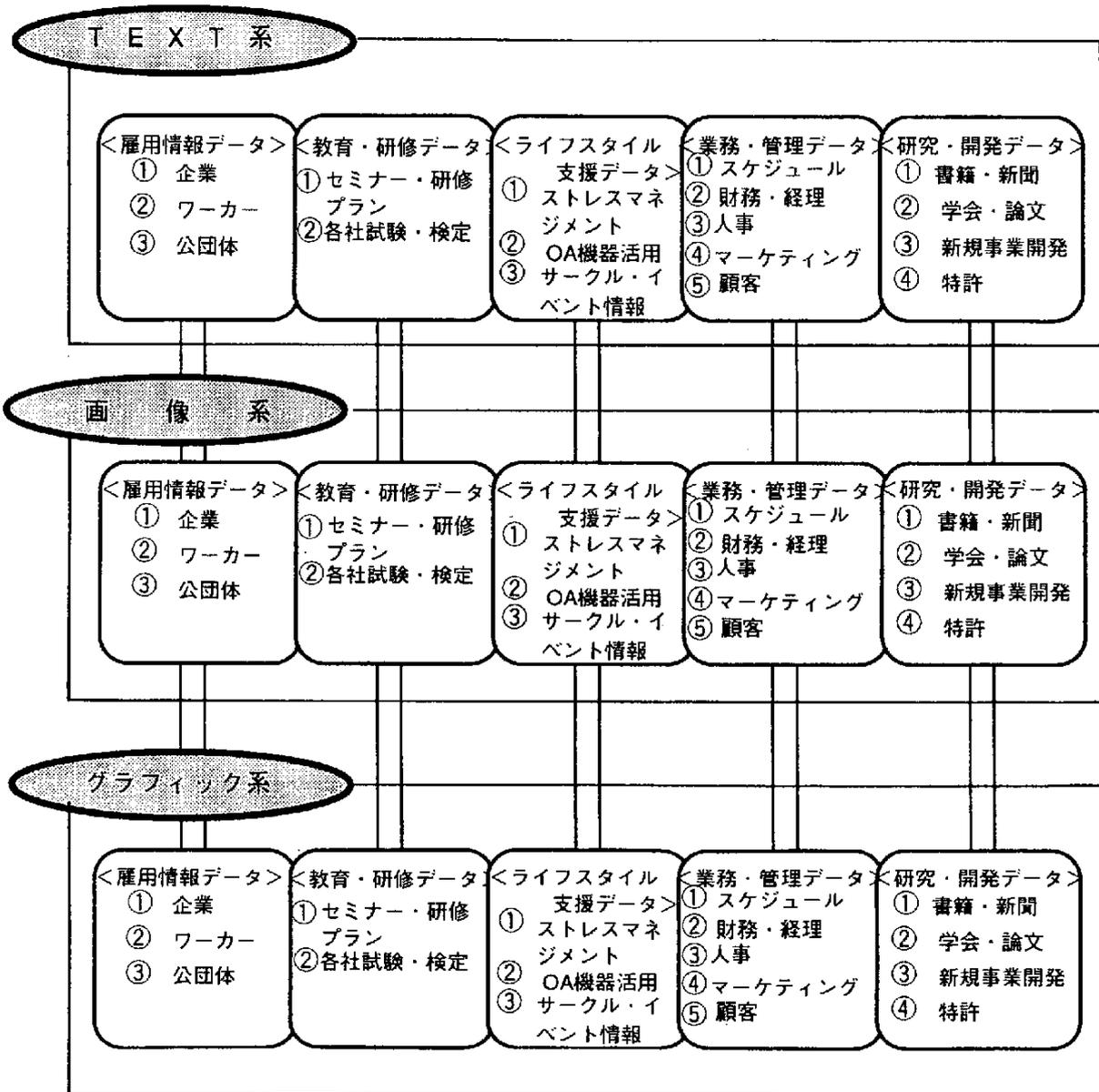


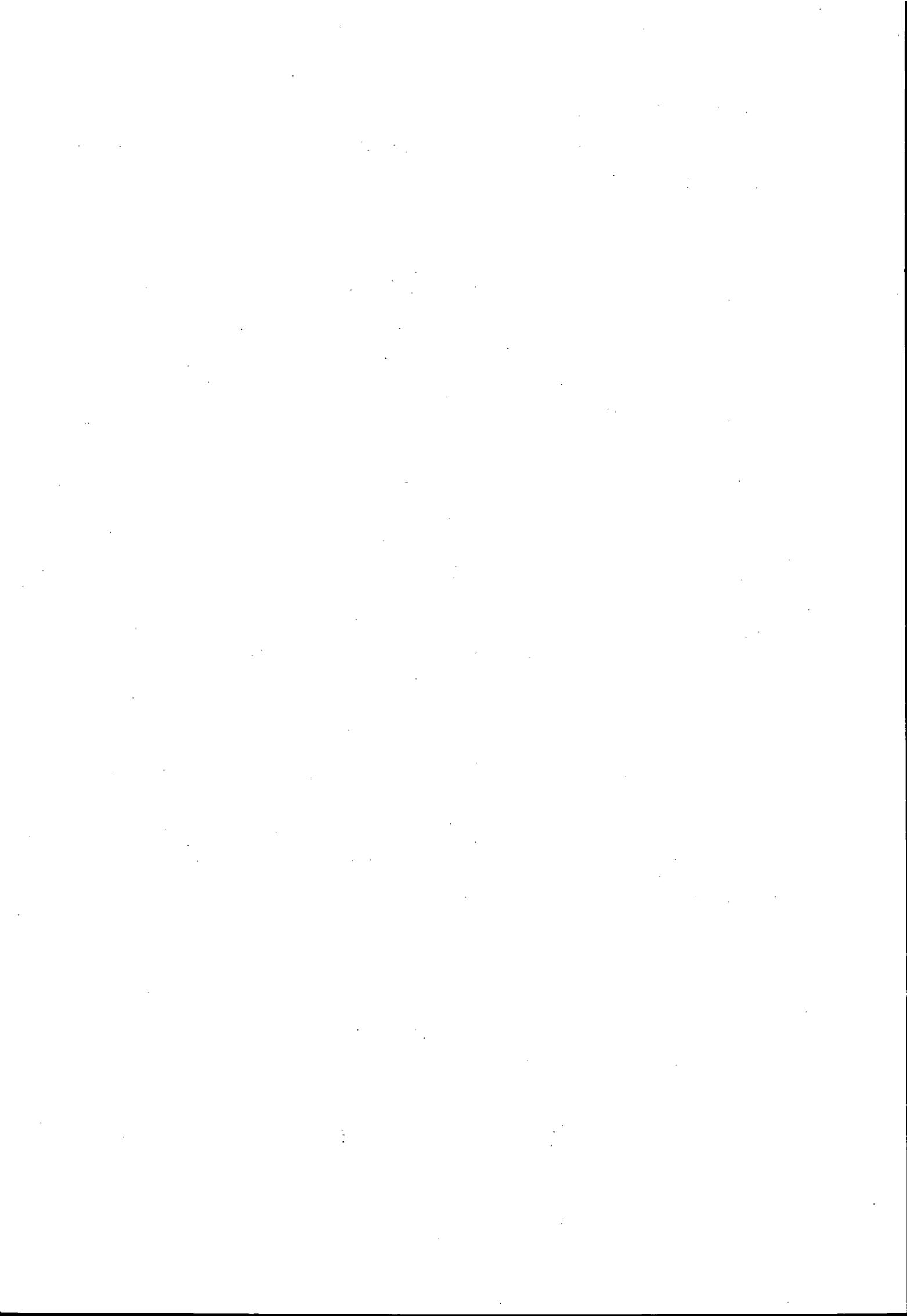
図13-4 データ内容の分類体系

事業化への方策としては、第1フェーズ（システム開発期）、第2フェーズ（地域ネット拡大期）、第3フェーズ（利潤還元期）での事業展開を考えており、約3年のうちに会員数を1万名にするという企画案とした。

	第1フェーズ (システム開発期)	第2フェーズ (地域ネット拡大期)	第3フェーズ (利潤還元期)
	'93.4	'94.4	'95.4 '96.4
事業計画内容	8月 12月 4th DIM 設計期 プロトタイプ 作成期 実験 導入期	8月 12月 協賛企業の 公募 UNIX/WS 移行 センター局員 の拡充	6月 新規組織発足 商用ネットとの共 同事業&世界的な ネットの拡大
必要経費	DBソフト 4th DIM 他 60万 DBハード MAC FX/ci } ハードディスク } 200万 モデム } FAX配送機 } オペレータ人件費 (1人) 1万/1日×200日=200万 システム開発費・メンテ 250万 通信費等 5万/1ヶ月×12=60万 計 770万	DBソフト (ATOSON-1) 350万 DBハード (UNIX/WS) 250万 FAX配送機 50万 オペレータ人件費 (3人) 200万×3人=600万 システム開発費・メンテ 250万 通信費等 20万/1ヶ月×12=240万 計 1,740万	ソフト系 1,000万 ハード系 2,500万 人件費/メンテ 1,500万 通信費等 1,000万 計 6,000万
組織	志木サテライト+公共事業団	志木サテライト+公共事業団 協賛企業	社団or新規事業団企業
会員	800名	3,000名	10,000名

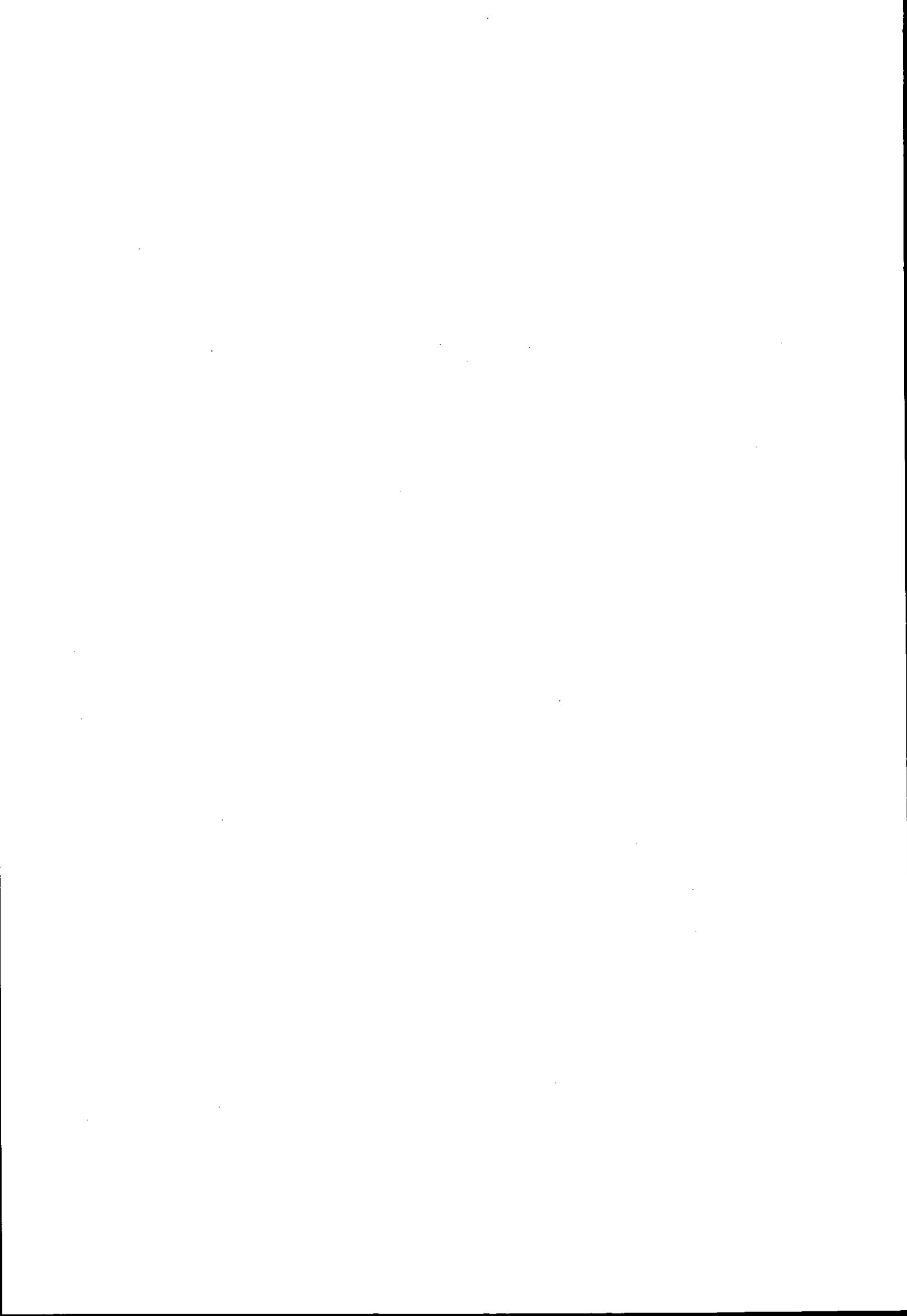
*事業化を大きく二つに分割すれば、2年後までを実験的な事業期、それ以後を本格的なビジネス拡大期と分けられる。上記の例は更に最初の時期をシステム開発期として三つのフェーズに分割している。

図13-5 今後の事業展開計画案



14 銅基複合材料日本特許英文データベースの構築

神鋼リサーチ株式会社



14. 銅基複合材料日本特許英文データベースの構築

14.1 目 的

昭和62年及び昭和63年には、(財)データベース振興センターからの委託によりSiC系ファインセラミックス及び銅基複合材料の特許データベースを構築し、その後平成元年及び平成2年は、メニュー追加及びデータ更新を実施し、5種類のデータベースを構築した。平成3年及び昨年は、そのデータベースを使用し、グラフ解析結果も含めて日本特許発明者にPRを実施した結果、多数の日本特許発明者に利用されるようになった。

平成3年度のデータベース白書にも記載されている如く、日本情報に対する諸外国からの関心は非常に高いので、今回はビジネスとしての採算性に幾分不安はあるが、銅基複合材料日本特許英文データベース構築の委託を受け、その情報を加工し、海外へのサービスを検討するために、英文日本特許マップ調査集を作成した。

14.2 概 要

日本特許の英文データベースについては、既に国内外でサービスされているが、それは特許書誌事項と抄録文であり、更に効率的に日本特許情報が活用されるためには、銅基複合材料日本特許英文データベースを構築し、その情報を加工して海外にサービスすることが必要である。そのためそのサービスの検討をした。既に(財)データベース振興センターの委託を受けてデータベースを構築しデータ更新を完了している銅基複合材料特許データベースの情報(和文1,162件)について、現在オンラインでサービスされている日本特許英文データベースにより情報(和51年10月以降のデータ)を収録し、dBASE IIIを使用し、公開特許番号、発明の名称、出願日、出願人、発明者、IPC及びキーワードを収集し、データベースの構築を行い、検索及びグラフ解析が容易に実施できるようにし、海外へのサービスを検討するため、英文日本特許マップ調査集を作成した。

14.3 実 施 内 容

14.3.1 情報収集

(1) 予備検索

下記の式で予備検索を行い、銅基複合材料に関するフリーキーワード及びIPCの選択を

行った。昭和55年1月から昭和62年12月までを検索期間とし、下記S4式の資料によりフリーキーワードの選択を行った。昭和48年、昭和52年、昭和55年、昭和59年及び昭和62年を検索期間とし、下記S5式の資料によりIPCの選択を行った。

- F FK=(F355572 + F355699 + F355755 + F356038) S 1
 (銅) (銅基) (銅合金) (銅粉)
- F FK=(F213283 + F213297 + F283557) S 2
 (黄銅) (黄銅粉) (青銅)
- F XK=(R031 + R032) S 3
 (粉末冶金) (繊維強化複合金属)
- F S 3 * (S 1 + S 2) → オフラインにより要約文手配 S 4
- F S 4 # C22C9/00 → オフラインにより抄録文手配 S 5
 (銅基合金)

(2) 銅基複合材料PATOLIS 検索式

予備検索により入手した特許公報抄録文、特許公報要約文により、銅基複合材料に関する特許公報を選択し、IPCとフリーキーワードのものを追加し、検索式を作成した。

銅基複合材料PATOLIS 検索式

- F IC=(C22C1/00 + C22C26/00 + C22C29/00 + C22C30/00 + C22C32/00
 [合金の製造] [ダイヤモンドを含む金] [炭化物、酸化物、ほう化物、窒化物又は珪化物を基とする合金] [各成分を50%未満含有する合金] [ほう化物、炭化物、窒化物、酸化物、珪化物又はその他の金属化合物]
 + C10M7/02) S 1
 (無機化合物)

- F IC=(B21D37/20 + B22D23/04 + B22F3/00 + B22F5/00
 [単一の他のサブクラスでカバーされない作業による工具の製造] [浸漬による鑄造] [成形又は焼結方法に特徴のある金属質粉からの工作物/物品製造] [製品の特殊な形状に特徴のある金属質粉からの工作物/物品製造]
 + B22F7/00 + B23K9/26 + B23K20/00 + B23K35/02
 [粉末焼結による金属粉からなる複合層、複合工作物/物品製造] [電極の付属具] [衝撃又は他の圧力を加えることによる非電氣的接合] [溶接棒、溶接電極等の機械的形狀]

+ B23K35/30 + B32B5/12 + B23P1/12 + B29D3/02) S 2

主成分が 1550℃以下 の融点をも つもの	隣接する層の織 維又は繊維条の 相対的配列を特 徴とする積層体	工作物に対 する電極の 運動又は位 置ざめ	挿入物をも った物品で 補強材をも つもの
---------------------------------	--	--------------------------------	--------------------------------

F IC=(F16C27/02 + F16C33/14 + F16C33/24 + F16J15/34

滑り 軸受	軸受又は軸受 部品の特別な 製造方法：な じみ運転	種々の材料か らなる種々の 範囲の滑り面 をもつ滑り軸 受部品	一つの部材に多少 半径方向の面に対 して押しつけられ たスリップリング を持つ密封装置
----------	------------------------------------	---	---

+ F16K5/22 + F28D15/00 + H01H1/02 + H01H33/66 + H01L21/58) S 3

栓又はコッ クで潤滑に 関し特徴の あるもの	閉鎖管中の中間 熱伝達媒体が流 路壁を通り抜け る熱交換装置	材質を 特徴と する接 点	真空ス イッチ	支持体上 への半導 体装置の マウント
---------------------------------	---	------------------------	------------	------------------------------

F FK=(F213283 + F213297 + F267894 + F283557 + F355572 + F355699
(黄銅) (黄銅粉) (純銅) (青銅) (銅) (銅基)

+ F355755 + F356004 + F356011 + F356038 + F356041) S 4
(銅合金) (銅被覆) (銅被膜) (銅粉) (銅粉末)

F IC=(C22C9/00) S 5
(銅基合金)

F (S 1 + S 2 + S 3) * (S 4 + S 5) S 6

F XK=(R031 + R032) S 7
(粉末冶金) (纖維強化複合金属)

F FK=(F249540 + F249574 + F249560 + F249699 + F399723 + F399630) S 8
(焼結) (焼結体) (焼結材) (焼成) (粉末) (粉体)

F FK=(F163797 + F255212 + F356004 + F396511 + F398009 + F428941) S 9
(含浸) (浸透) (銅被覆) (複合) (分散) (溶浸)

F FK=(F173564 * F291885 + F317157 * F173564 + F176680 * F398616
(強化) (纖維) (短纖維) (強化) (均一) (分布)
+ F206568 * F435386) S10
(硬質) (粒子)

F S 5 * (S 7 + S 8 + S 9 + S10) S11

F (S 6 + S11) S12

(3) 情報収集及び情報選択

前項のS12式により、昭和46年7月から平成4年4月30日までに公開された特許情報5,270件を収集し、発泡金属、超電導材料、クラッド材、積層材及びロウ材以外の銅基複合材料に関する特許情報1,162件を選択した(適合率22%)。

(4) 日本特許英文データ収集

前項で収集・選択した銅基複合材料に関する特許情報1,162件のうち、昭和51年10月から平成4年4月30日までに公開された情報1,039件について、ORBITの日本特許の英文ファイルを使用し、次の要領で情報を収集した。

1. 55-122804/PN;55-122841/PN;55-128554/PN;55-134102/PN;55-134103/PN;55-134143/PN;55-134147/PN;55-136557/PN;55-138006/PN;55-138036/PN;55-138037/PN;55-138038/PN;55-145102/PN;55-145107/PN;55-145135/PN
2. 55-160437/PN;55-161004/PN;55-161005/PN;55-164049/PN;55-164050/PN;56-5901/PN;56-9344/PN;56-9345/PN;56-13449/PN;56-13451/PN;56-19832/PN;56-20137/PN;56-22018/PN;56-25944/PN;56-25945/PN
3. 56-41348/PN;56-44732/PN;56-45034/PN;56-45035/PN;56-47533/PN;56-51543/PN;56-55501/PN;56-55502/PN;56-62909/PN;56-69335/PN;56-72144/PN;56-77353/PN;56-84431/PN;56-84433/PN;56-133442/PN
4. 56-136942/PN;56-136943/PN;56-139641/PN;56-152905/PN;56-153629/PN;56-15670/PN;56-161647/PN;56-169739/PN;57-2847/PN;57-5832/PN;57-7872/PN;57-9850/PN;57-22417/PN;57-32304/PN;57-46413/PN
5. 57-54236/PN;57-57419/PN;57-57420/PN;57-57801/PN;57-57803/PN;57-57804/PN;57-67141/PN;57-70245/PN;57-76143/PN;57-79137/PN;57-87004/PN;57-87005/PN;57-101603/PN;57-101627/PN;57-108231/PN
6. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15
7. 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30
8. 31 OR 32 OR 33 OR 34 OR 35 OR 36 OR 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45
9. 46 OR 47 OR 48 OR 49 OR 50 OR 51 OR 52 OR 53 OR 54 OR 55 OR 56 OR 57 OR 58 OR 59 OR 60
10. 61 OR 62 OR 63 OR 64 OR 65 OR 66 OR 67 OR 68 OR 69 OR 70 OR 71 OR 72 OR 73 OR 74 OR 75
11. 76 OR 77 OR 78 OR 79 OR 80
12. PRT SS 81 PULL 1-75
13. AUTOCOST ON
14. STOP
15. Y

14.3.2 データベース構築

(1) 収録データ項目と収録法

公開特許番号，発明の名称，出願日，出願人，発明者，国際特許分類（IPC）及びキーワードを収録した。

1) 公開特許番号

例えば，S60-003821などとけた数を10けたに統一して入力した。

2) 出願日

例えば，83.06.22などと左づめで入力した。

3) 発明の名称，出願人及び発明者

ORBIT のファイルに準じて入力した。

4) 国際特許分類（IPC）

例えば，C22C-001/09 などと11けたに統一し，左づめで入力した。

5) キーワード

和文日本特許に収録されているフリーキーワードを集計し，表14-1の分類基準を作成し，それに基づいてキーワードを付与するとともに，発明の名称に使用されている単語をキーワードとして収録した。

表14-1 分類基準

分 類 基 準		収録件数
添 加 成 分	ALUMINUM (A1)	134
	CHROME (A2)	235
	TIN (A3)	208
	NICKEL (A4)	180
	IRON (A5)	167
	TITANIUM (A6)	115
	COBALT (A7)	125
	MOLYBDENUM (A8)	128
	GRAPHITE OR CARBON (A9)	228
製 法	DISPERSION REINFORCE (B1)	163
	FIBER REINFORCEMENT (B2)	109
	PRECIPITATION REINFO(B3)	19
	INFILTRATION METHOD (B4)	86
	INTERNAL OXIDIZATION (B5)	23
用 途	ELECTRODE (C1)	134
	ELECTRICAL CONTACT (C2)	216
	BEARING (C3)	79
	SLIDING MATERIAL (C4)	114
	FRICTIONAL MATERIAL (C5)	52
特 性	ABRASION RESISTANCE (D1)	226
	STRENGTH (D2)	245
	ELECTRICAL CONDUCTIV (D3)	103
	WELD RESISTANCE (D4)	145
	WEAR RESISTANCE (D5)	69
	HEAT RESISTANCE (D6)	65
	CORROSION RESISTANCE (D7)	36
	WORKABILITY (D8)	25
	TOUGHNESS (D9)	31
	FITNESS (D10)	21

(2) ファイル構造

レコード実例を調査し、ファイル構造を次のとおりとした。

USE CUENBASE

・ LIST STRUCTURE

データベースの構造 : B: CUENBASE.dbf

データ・レコードの数 : 1039

最終更新日付 : 01/07/93

表14-2 ファイル構造

番号	フィールド	タイプ	幅	備考
1	NO	数値型	4	レコード番号 公開特許番号 出願日 発明の名称
2	PN_A	文字型	10	
3	AP_DATE	文字型	10	
4	TI	文字型	128	
5	PA_1	文字型	40	} 出願人
6	PA_2	文字型	40	
7	PA_3	文字型	40	
8	PA_4	文字型	40	
9	IN_1	文字型	35	} 発明者
10	IN_2	文字型	35	
11	IN_3	文字型	35	
12	IN_4	文字型	35	
13	IN_5	文字型	35	
14	IN_6	文字型	35	
15	IC_1	文字型	15	} 国際特許分類 (IPC)
16	IC_2	文字型	15	
17	IC_3	文字型	15	
18	IC_4	文字型	15	
19	IC_5	文字型	15	
20	IC_6	文字型	15	
21	TEC 1	文字型	24	} キーワード
22	TEC 2	文字型	24	
23	TEC 3	文字型	24	
24	TEC 4	文字型	24	
25	TEC 5	文字型	24	
26	TEC 6	文字型	24	
27	TEC 7	文字型	24	
28	TEC 8	文字型	24	
29	TEC 9	文字型	24	
30	TEC10	文字型	24	
31	TEC11	文字型	24	} キーワード
32	TEC12	文字型	24	
33	TEC13	文字型	24	
34	TEC14	文字型	24	
35	TEC15	文字型	24	
36	TEC16	文字型	24	
37	TEC17	文字型	24	
38	TEC18	文字型	24	
39	TEC19	文字型	24	
40	TEC20	文字型	24	
41	TEC21	文字型	24	
42	TEC22	文字型	24	
43	TEC23	文字型	24	
44	TEC24	文字型	24	
45	TEC25	文字型	24	
合計	—	—	1213	レコード長

14.4 本データベース利用の効果

- (1) 和文の銅基複合材料日本特許データベースについては、(財)データベース振興センターの委託を得て、PATOLIS の詳細な検索式を作成して情報収集を実施し、更に情報を選択し、その情報について改めてキーワードを付与しており、研究開発の狙いの検討や特許侵害に関する調査のために有効に活用されている。同一のテーマについて英文データベースを構築したので海外に対しても同様に精度及び加工度の高い情報を提供でき、各種テーマに応用することにより日本情報に対する海外からのニーズに応え、日本情報の国際的流通の促進に寄与することができる。
- (2) 日本語に慣れていない英語を母国語とする国内外の人々、特に研究開発担当者が研究開発の狙いの検討や特許侵害に関する調査のために有効に活用でき、研究開発の重複が避けられる。
- (3) 本データベースは、各種の英文特許マップ作成のために活用できるので技術動向の調査が容易であり、適切な開発方針の決定のために使用でき、研究開発の効率化に寄与する。

14.5 今後の課題と展望

- (1) 日本特許英文情報の海外へのサービスについては、ビジネスとしての採算性に幾分不安はあるが、日本情報に対する諸外国からの関心は非常に高いので、精度及び加工度の高い情報を提供し採算性の検討を行う。
- (2) 海外へのPR方法としてはダイレクトメールによる方法や代理店を活用する方法等があるが、いずれにしても、類似テーマを同時にPRしてPR効率を良くするために、類似のテーマのメニューを追加する必要がある。

14.6 英文日本特許マップの海外へのサービス資料例

dBASE IIIデータのグラフ作成ソフトは各種のものが発売されているが、今回はテクノリサーチ(株)の特許マップ作成ソフトを使用し、特許マップデータ解析を実施した。解析項目は下記のとおりであり、詳細は報告書を参照下さい。

14.6.1. 情報解析結果（総括編）

14.6.1.1. 出願人ランキング及び発明者ランキング

14.6.1.2. 主要出願人毎の特許出願件数の推移

14.6.1.3. 主要出願人毎の発明者数（重複を除く）推移

14.6.1.4. 主要分類（IPC）毎の特許出願件数の推移

14.6.1.5. 技術動向（キーワード毎の特許出願件数の推移）

14.6.1.6. 技術関連

14.6.1.7. 新規参入出願人リスト

14.6.2. 情報解析結果（個別企業編）

（1）明電舎

（2）トヨタ自動車

（3）三菱電気

（4）田中貴金属工業

（5）東 芝

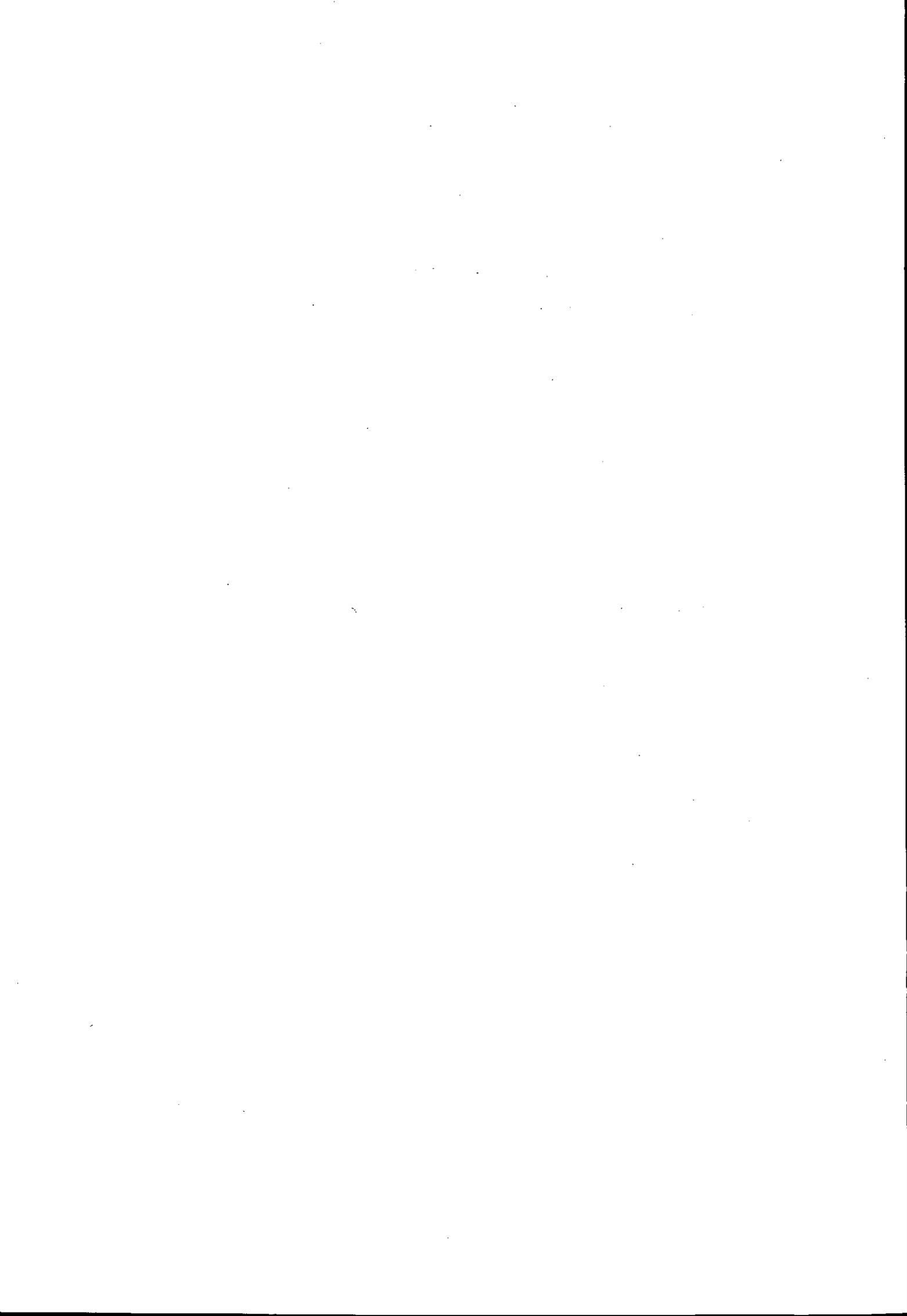
（6）三菱マテリアル

（7）住友電工

（8）日 立

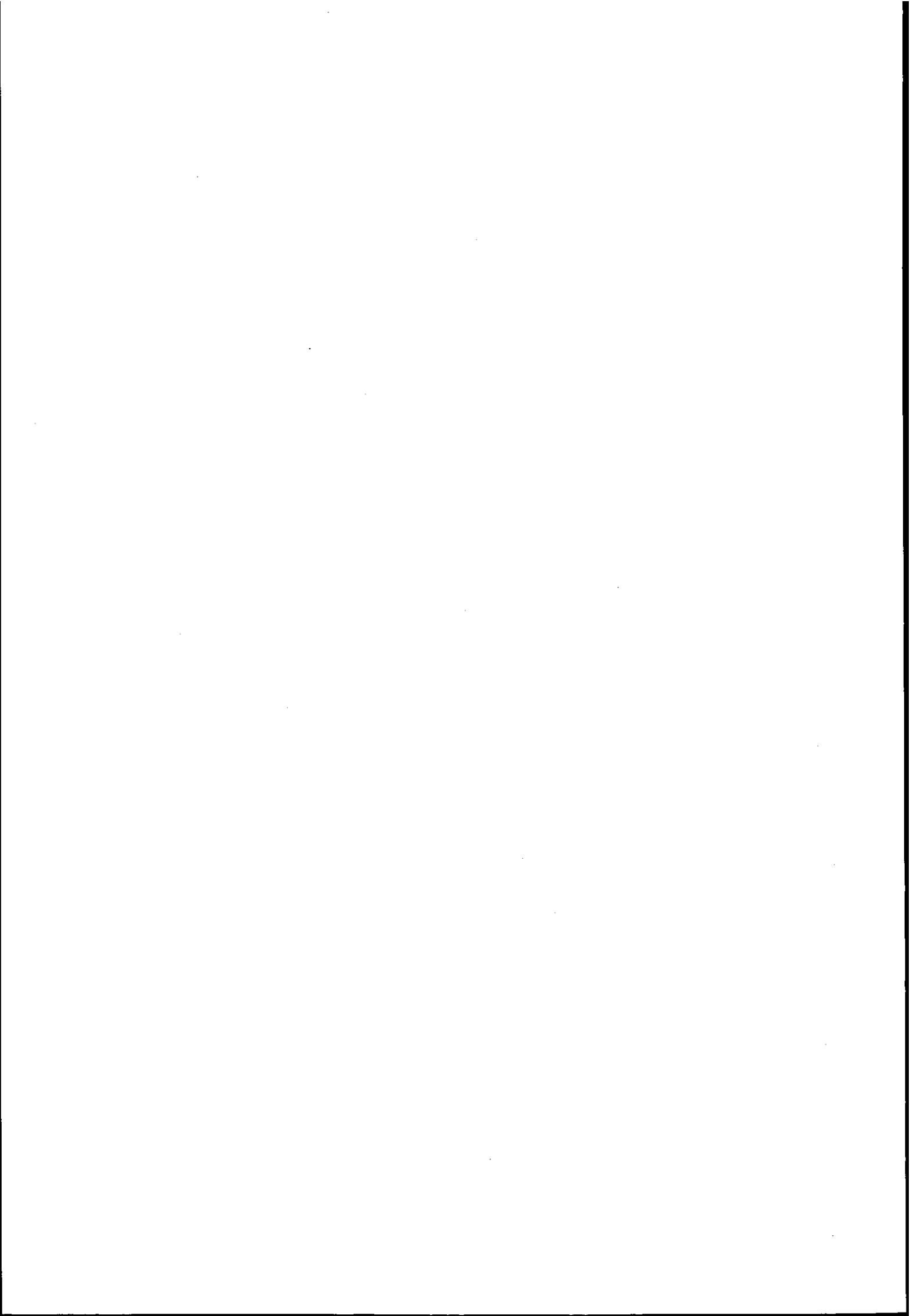
（9）東芝タンガロイ

（10）日立粉末冶金



15 技術協力供与機材データベースの
プロトタイプ作成

財団法人 日本国際協力システム



15. 技術協力供与機材データベースのプロトタイプ作成

15.1 目的

わが国の実施している政府開発援助の中で技術協力は、研修員の受け入れ、機材の供与、専門家の派遣等の協力により、開発途上国の人造り、国造りに大きく貢献をしてきた。特に、わが国の供与した機材はプロジェクトの実施に不可欠であり、これによって技術移転がスムーズに行われ大きな成果を挙げてきた。しかしながら一方で、援助する側のきめ細かなデータの収集、整備が必ずしも十分に行われてこなかった側面も一部見られる。

したがって、供与される機材の情報を体系的にデータベース化し、技術協力関連の機関がそれを共有し利用することは、質の高い効果的かつ効率的な技術協力を実施する上で必須のものであり、急務となっている。

当財団では、平成2年度に「政府開発援助（ODA）に関するデータベースの構築調査」を実施した。ここで提起したシステム化構想のうち、本年度は技術協力関連機材のデータの整備をするための準備段階として、わが国の技術協力の主体的な実施機関となっている国際協力事業団が実施した種々のプロジェクト方式技術協力の中から代表的なものを取り上げ、将来における体系的なデータベースの構築に資するためにプロトタイプを作成した。

15.2 実施内容

本事業を推進するに当たり、「技術協力供与機材データベース検討委員会」を組織し、全体の方針・計画の調整及び仕様の検討の指針を討議し、作業部会で具体的作業を実施した。

対象とする業務に関係する事業部に対するアンケート調査を実施し、エンドユーザのニーズの把握を行い、仕様の概要を作成し、プログラムの第一バージョンを作成し、マスタファイルの入力、機材データの入力を行った。

これを検討委員会でデモンストレーションを実施し、更に詳細な機能の検討を行った。この結果を受け、プログラムに対しては微調整を行い、データベース構築のための問題点を明確化し、今後の方針等を作成した。

15.3 データベース概要

15.3.1 コンセプト

このデータベースは、前項のアンケートの結果を参考に、図15-1のような用途と対象部門について、プロジェクトの機材選定等に直接従事している各事業部の担当者を具体的なユーザーと想定して構築した。

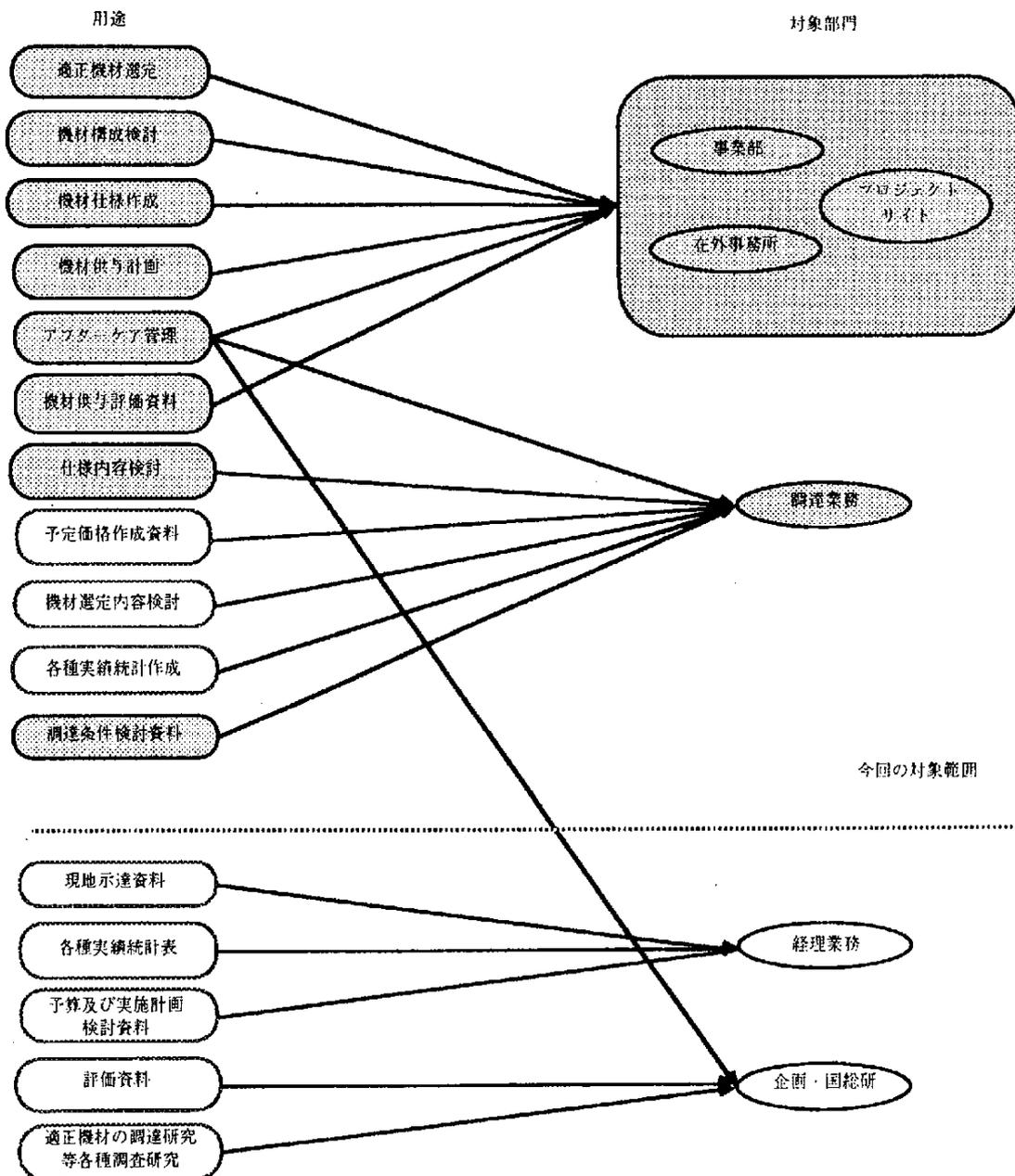


図15-1 用途と対象部門

15.3.2 機能概要

プロジェクトに供与した機材について以下のデータ項目をキーワード(複合条件設定可能)

にしてそれぞれ検索し、一覧表示、帳票の出力をする。また、必要に応じ、集計並びに統計処理等計算機能を有する。

(キーワードとなる主な項目)

- ・国別
- ・年度別
- ・事業費別
- ・プロジェクト種別別
- ・プロジェクト別

(1) 次の項目が画面上で表示、閲覧できる。

- ・ 供与機材関連データ

(機材名、型式、単価、数量、仕様、電源、付属品、スペアパーツ、その他)

- ・ 納入商社関連情報

(商社名、連絡先、担当者、現地代理店、その他)

- ・ メーカー情報

(メーカー名、連絡先、担当者、現地代理店、その他)

- ・ 契約関係情報

(契約番号、契約相手方、契約年月日、契約金額、その他)

- ・ プロジェクト関係情報

(国名、プロジェクト名、プロジェクト種別)

(2) 機材の使用現況が分かる。

- ・ 保守管理状況が分かる。
- ・ 維持・修理状況が分かる。

(3) 特定の機材種別(将来的には機材別)に、メーカー名、納入実績、納入可能商社が分かる。

(4) 検索機能

自由キー検索を基本とする。

(データ入力後、今後のシステム化に向けて、キーワード、シソーラス作成の要不要について検討することとする)

(5) 国別、プロジェクト別、機材種別によって分類した供与実績を帳票形式で出力する。

15.3.3 ハードウェア

プロトタイプを構築するハードウェアは、Apple社のMacintosh IIciを採用した。理由は以下のとおりである。

- ・マンマシンインタフェースが優れていること
- ・コンピュータの知識を必要とせず、エンドユーザレベルで容易に本データベースを活用できるようにできること
- ・GUIが優れていること
- ・将来ネットワーク上で運用する際に想定しているUNIXサーバーマシンとの接続性が優れていること
- ・システムの拡張性に優れていること

Macintosh II ciのシステム構成は、

- ・CPU 25MHz 68030
- ・メモリ 5MB
- ・ハードディスク 230MB
- ・OS 漢字トーク 6.0.7

15.3.4 ソフトウェア

リレーショナルデータベースの4th DIMENSION v2.1.1-Jを採用した。

15.3.5 ファイル構造

マスタファイル

- ・業者マスタファイル (+代理店情報サブファイル)
- ・業種マスタファイル
- ・プロジェクトマスタファイル
- ・プロジェクト種別マスタファイル
- ・機材種別マスタファイル
- ・事業費マスタファイル
- ・国マスタファイル

データファイル

- ・購入機材情報ファイル (+購入物品サブファイル)
- ・契約情報ファイル (+契約情報サブファイル)
- ・使用状況情報ファイル

15.3.6 画面・帳票

入力画面等の画面の構成は次のとおり。

メニューバー : 各種マスタ, 供与機材DB, 編集, ファイル

各種マスタ : 業者マスタ登録, 業者マスター一覧/更新
 業種マスタ登録, 業種マスター一覧/更新
 プロジェクトマスタ登録, プロジェクトマスター一覧/更新
 プロジェクト種別登録, プロジェクト種別一覧/更新
 機材種別マスタ登録, 機材種別マスター一覧/更新
 事業費マスタ登録, 事業費マスター一覧/更新
 国マスタ登録, 国マスター一覧/更新

供与機材DB : 購入契約登録, 購入契約一覧/更新
 機材情報一覧/更新
 使用状況情報一覧/更新

(出力リスト)

国別, プロジェクト別機材供与実績表
 購入契約別機材供与実績表
 プロジェクト別機材供与実績表
 機材別機材供与実績表
 機材製造会社別機材供与実績表
 納入商社別 (価格有り) 機材供与実績表
 納入商社別 (担当部課) 機材供与実績表

また, 主な画面のサンプルを次に示す。

購入機材一覧表示/更新(9/25)			
【契約情報参照】			削除
プロジェクトコード	0483101	水道技術訓練センター	キャンセル
契約番号	63120902	予算年度 1987	登録
業者コード	006860209	藤田商事株式会社	
契約方法			
契約年月日	88.12.21	納期年月日 89.04.28	
契約金額	4,500,000	見積金額	
			← →
【機材情報】			
↓ダブルクリックすると入力できます		合計金額	¥4,226,600
商 品 名	製 造 元	価 格	
コンピュータソフトウェア	(株) プラビオスインターナショナル	510,000	
デジタルマンメーター	富士電機	1,141,000	
溶存酸素計	薬研製作所	606,000	
PH計	電気科学		
ジャーテスター	セントラル科学	204,000	
ピーター	大科電機	102,000	
PH測定器	アドバンテック	301,000	
ホテイングスターラー	アドバンテック	120,000	
投水器	融会社	224,000	
下ロビン	サンブラテック	3,900	
下ロビン	東洋計量器	13,800	
金属アスピレータ	東洋計量器	2,900	
洗浄用大かご	サンワカケン	28,800	
洗浄用小かご	サンワカケン	10,000	

図15-2 購入機材一覧表示/更新画面の例

カスタム				
プロジェクト名	納期年月日	No.	機材名	製造
作物保護強化	89.05.31	1	独-日-英「科学用語語源辞典」	同学社
作物保護強化	89.05.31	2	科学用語語源辞典	同学社
作物保護強化	89.05.31	3	有用マメ科植物ハンドブック	財団法人大豆輸入協
作物保護強化	89.05.31	4	図説熱帯植物集	広川書店
作物保護強化	89.05.31	5	Modelling in Behavioural Ecology : An introductory	Croom Helm London
作物保護強化	89.05.31	6	The estimation of animal abundance and related	Griffin London
作物保護強化	89.05.31	7	Ecology (paperback) by M.Begonetal	Blackwell Scientific
作物保護強化	89.05.31	8	Insect on grain legumes in Northern Avstralia	University of Gbeens
作物保護強化	89.05.31	9	Insect outbreaks by eds P.Barbosa &J.C.Shults	Academic PRESS
作物保護強化	89.05.31	10	Agricultural insect pests of the tropics and their	Cambridge Universit
作物保護強化	89.05.31	11	Ecological theory and in tegrated pest management	Wiley Interscience
作物保護強化	89.05.31	12	Biological control on the phyll oplane (1985)	ASP Press
作物保護強化	89.05.31	13	Genetic basis of bio-chemical mechanisms of plant	APS Press

図15-3 出力帳票内容表示 (プロジェクト別の例)

15.3.7 プロトタイプの評価

(1) データ量

目標は各事業部（5事業部）から2プロジェクトを選び、合計10プロジェクトの供与実績を入力して結果を検討することであったが、さまざまな理由で、機材が入力できたのは、6プロジェクトとなった。

1機材に関する情報を1レコードとした場合、6プロジェクトで末端のレコードは約400レコードとなり、データ量は2MBとなった。ただし、使用状況情報については、入手できず未入力であること、機材種別等のコード類が未定のため、本来見込まれるデータ量の半分程度と判断される。

また、システムもコンパイル後、2MBの容量となった。

(2) 問題点

① オリジナルデータのフォーマットが不統一

オリジナルデータのフォーマットを決める必要がある。最も基本的な契約書付属の納入機材内訳書の品名のくくりかたが統一されておらず、入力時の作業に支障がある。困るパターンの一つは、

品名	仕様	価格
「コンピュータシステム一式」	P C 9801 2台	トータル
	外付けハードディスク 120MB	金額
	熱転写プリンタ	

というもので、本データベース入力の最小単位は上記の例では仕様部分の個々の機器を個々のレコードとして管理したいのだが、「コンピュータシステム一式」の価格しか情報がないため、個々の機器の納入価格、仕様等の基礎データが把握できない。

② 機材種別又は機材コードが未確定

現在検索方式が自由キー検索としているが、登録した膨大な供与機材品目名のうち、求める機材の品名をいちいち記憶しておかないと、検索時に抽出できなくなる。また、契約書の内訳書の記述にも統一性がないので、同じ機材が別の名前で登録されていた場合、もしくはタイプミス等によるわずかな違いで検索できない例がある。このため、ワイルドカードを用いて、ある程度あいまいな検索キーでも検索できるように設計したが、限界があるようである。

このため、機材種別コード又は機材コードを定めて、用途や分類によって絞り込み検索をする方式も取り入れているが、機材種別又は機材コードを作成するに至らなかった。コードによる検索を実用化する方がインデックスづけすることにより、検索及びソートのスピードの向上が図れるので真剣に検討する必要がある。また、機材の品名を記憶するユーザ側の負担がなくなるので効果は大きいと思われる。

③ 現地調達機材、第三国調達機材情報の入手方法

国内調達機材については、フォーマットの若干の相違はあるものの、契約書でオリジナルデータは捕捉できるが、現地調達機材、第三国調達機材の情報については、現在のところ、恒常的に一定のフォーマットで入手する方法がない。したがって、今回のデータ入力では、現地調達機材、第三国調達機材情報は割愛した。

また、現地調達機材、第三国調達機材の過去の実績についてのフォローの可能性については、プロジェクトにより情報の保管状況が異なるため困難と思われる。

今後、プロジェクトの機材の現地調達、第三国調達のポーションは一層増えていくと思われるので、データベースのデータとして取り込む必要性が大である。

④ 図形データ、写真データの取り扱い

検討委員会の中で、将来的にプロジェクトサイトの機材の配置図、機材の図面データ、供与機材のマニュアルの原本管理について取り扱う必要があるという意見があった。本システム及びソフトウェアでは、P I C T形式でのデータの取り込みが可能ではあるが、

対象が膨大であるため、パソコンで取り扱えるレベルではない。また、一部扱うにしても、パフォーマンスを犠牲にしなければならない。したがって、光ファイリングシステムとの連携を含めたハードウェア構成等について検討する必要がある。

(3) 改善のための懸案事項

① オリジナルデータのフォーマットの統一

前項で指摘したように現在のオリジナルデータは、かなりフォーマットが不統一で、手作業で入力するにも効率が悪い。国際協力事業団で交わす契約書の内訳書の指定フォーマットを作成し、納入業者に義務づけることは最低限検討すべきである。

また、Excel等の表計算ソフトが既に民間企業では浸透していることを考えると、契約書とともに内訳書の表をフロッピーにに入れて提出してもらう等の改善が考えられる。この場合の入力作業の軽減・効率化は絶大と考えられる。

② 機材種別又は機材コードの確定

検討委員会では、1) JIS の分類を参考にする、2) JICST の分類を参考にする、3) SITCコード(国連で使用)を参考にする等の意見が出されたが、現在のところ決め手はないのが実状である。機材種別コード並びに機材コードの検討については、次年度の課題とした。

この機材種別コード並びに機材コードの策定は、本データベースの運用上最もキーとなる部分でもあり、また、将来的に機材購送の手続きにPOSシステムを導入することも考えられるので慎重かつ早急に制定したい。

③ 現地調達機材、第三国調達機材情報の入手方法の検討

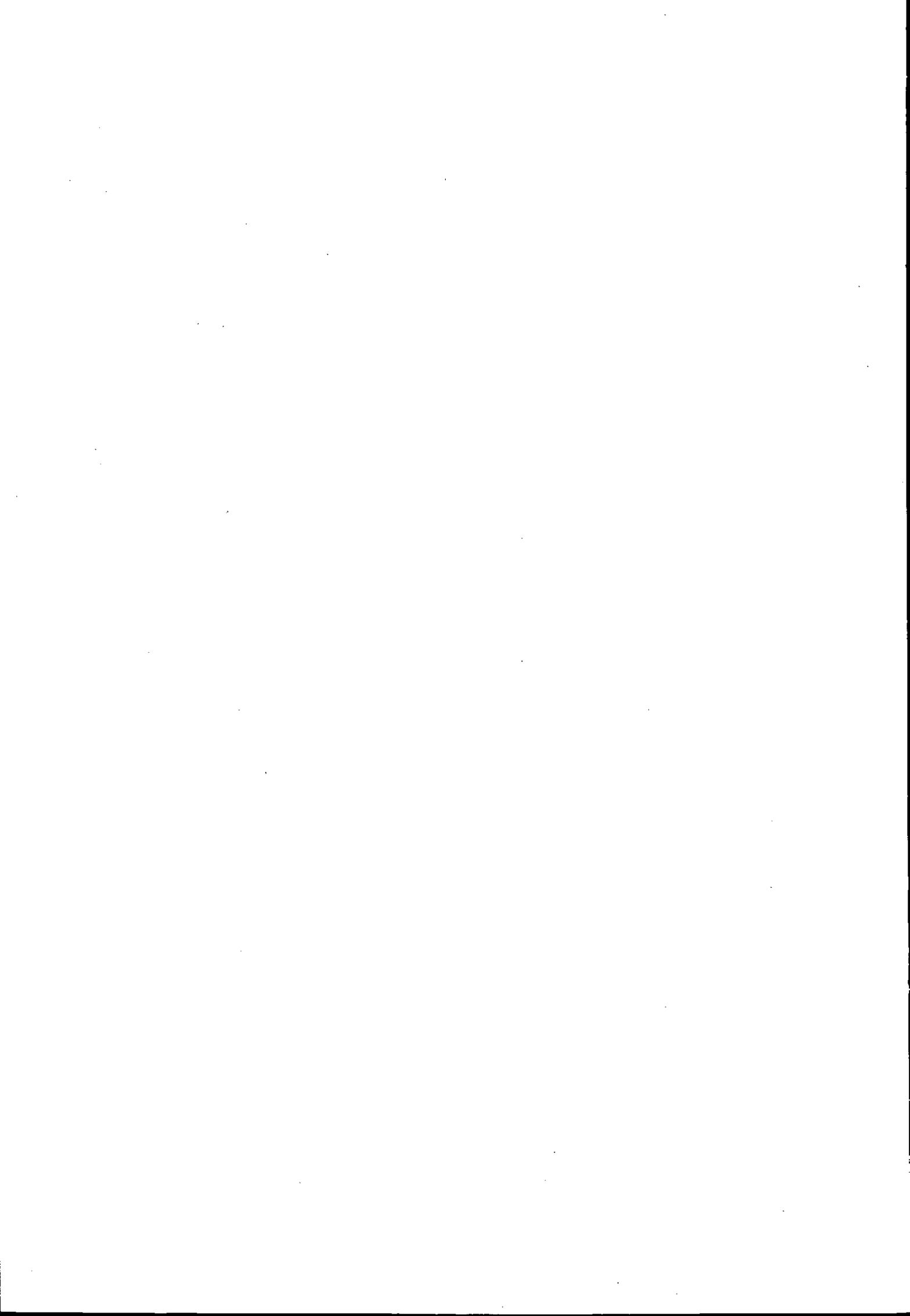
アンケートでも多数の意見があったが、機材の現地取り扱い状況の情報、現地の代理店情報、メンテナンス可能性等の情報に対する要望が大変強い。今後ますます、機材の現地調達、第三国調達が増えていく傾向を考えると、今から情報整備に着手する必要がある。そのためには、国際協力事業団でのプロジェクトサイトからの情報の入手方法、提出フォーマット等をシステムティックに構成してもらう必要がある。データ入力に適した形のものになるように、今後国際協力事業団と十分すり合わせをしながら検討していく必要がある。

15.4 今後予想される効果

(1) 技術協力の効果的・効率的実施に資することが可能となる。

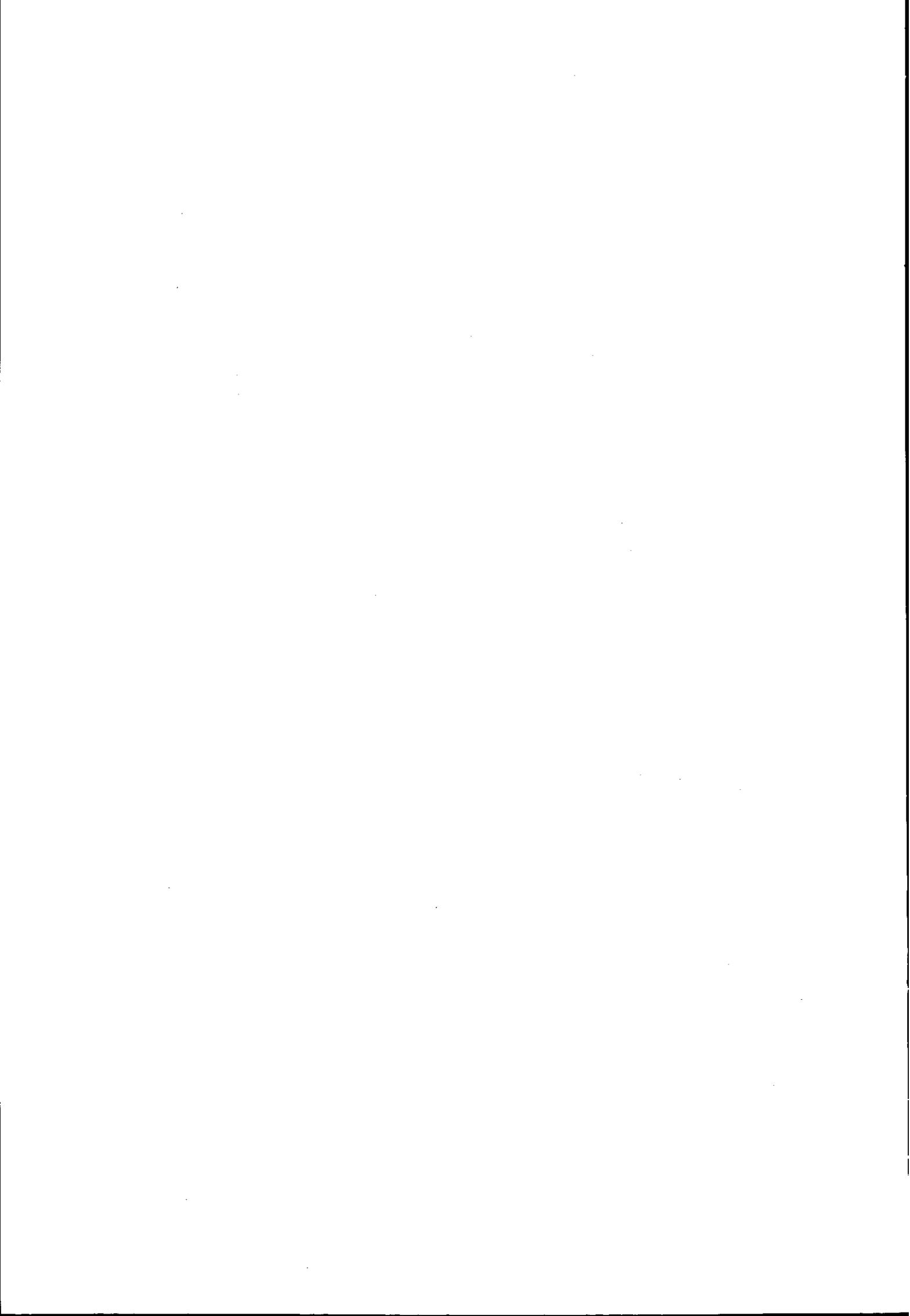
- ・迅速な機材仕様の作成

- ・効率的な購送業務の実施
- (2) 個々のプロジェクトに供与した資機材を把握することにより、整合性のとれた機材の選定、購送が可能となり、機材供与事業の精度・効果が向上する。
- ・適正機材構成策定
 - ・適正機材選定
- (3) 進行中のプロジェクト及び終了プロジェクトに対してのフォローアップ、アフターケアを迅速かつ正確に実施することが可能となるとともに機材の保守管理が容易となる。
- ・保守管理の改善
 - ・トラブル時の迅速な対応
 - ・アフターケア管理の充実と強化
- (4) 現地及び第三国調達物品の購入実績情報整備により、現地調達可能性、メンテナンス可能性が的確に把握できる。
- ・現地調達の適正な実施と管理



16 先端産業分野における専門用語の
電子辞書データベース化の調査研究

科学技術情報研究所株式会社



16. 先端産業分野における専門用語の電子辞書データベース化の調査研究

16.1 概要

本調査研究「先端産業分野における専門用語の電子辞書データベース化の調査研究」は、辞書の主体的な維持管理者として、ユーザではなく、むしろ翻訳システムの開発・供給者を想定し、それらにおける効率的な先端専門用語の検出と、翻訳用辞書データベースの経常的更新を支援するシステムのありようを中心課題として検討を行ってきた。つまり、ユーザはこうして常時更新されてゆく辞書の、経常的な配布を受けることによって、常に最新の環境において、翻訳システムを使ってゆくことができるという図式を追求するものとした平成3年度及び4年度の2年間にわたって調査研究を進めてきた。平成3年度においては、英→和の翻訳を主題として、先端専門用語のための英日翻訳辞書の更新・維持管理の方法について検討した。昨年度の調査の結果をふまえて、平成4年度においては、和→英の機械翻訳用辞書のありようを主題の一つとして検討した。ここで、昨年度の調査研究の結果得られた重要な知見として、翻訳用専門用語辞書における「英→和」と「和→英」の統合的な更新・維持管理という構想がある。機械翻訳システムは、英日、日英ともに既に出回りつつある。つまり、辞書の需要としては、英和、和英の双方ともに既に高まっている。このことは情報流通の国際化がますます進展してゆくことの反映でもあろう。したがって、海外の英語文献に発見される新出英語をいち早く英和辞書に登録すると同時に、わが国の和文文献から発見される日本語新出用語に、適切な英訳を与えて、日本情報の海外発信をより迅速、円滑にしてゆくことも、機械翻訳システムに期待される重大な機能になっている。情報交流の濃密化に伴って、このような新訳語は急速に普及し、英→和として得られた新語は、早晩、和→英の方向に還流して出現するし、また反対の方向、つまり、日本語起源の英訳語のわが国への還流という事例も、当然ながら多発する。

つまり、ここでは、「英→和」あるいは「和→英」という一方向の翻訳に着目した辞書の更新管理は、非常な不効率を招来するであろう。したがって、先端用語辞書に関しては、英和、和英の双方向からの利用を当初から折り込んだ統合的な辞書管理の方式が必要である。いわば、英和、和英の両引き辞書、兼用辞書ということを念頭においた検討が必要であり、また可能でありかつまた有効であるというのが、昨年度の「英→和」主体の調査の中で判明してきた。

この知見に即して、平成4年度の調査研究においては、和英両引きの機械翻訳用専門用語辞書の更新システムを中心課題に設定して、調査研究を進めた。即ち、両引き辞書を設計し、既存の電子化辞書数種を併合しつつ、こうした両引き辞書を実際に生成し、これに英和、和英双

方向からの高速検索機能を付加した。次に、この両引き辞書の諸特性について、統計的な分析を行った。即ち、英語については、句を構成する単語数と適用専門分野の関係、日本語については、漢字だけからなる語、カタカナ語等々、字種の構成を、その適用分野別の変位をみた。また、いわゆる訳し分けの状況に関して、昨年度の英語の和訳の際の分野別の訳し分けと対照させて、新たに、日本語の英訳の際の訳し分けの状況を統計した。更に、日本語の用語からスタートして、その英訳語を得、次にその和訳語を得るとような、連鎖的な翻訳をしたときに、どのような結果がえられるかを検討して、潜在的な類義語の群がどのように形成されているかを調査した。

次に、日本語の電子化文献を収集し、これをテストデータとして、上記の両引き辞書と突き合わせながら、新出の日本語専門用語の検出のシステム化に関して検討し、同時に、日本語の頻度統計を行った。この際、日本語処理に固有の問題である分かち書きについても検討し、辞書登録語との完全一致と部分一致の双方を試みた。なお、本調査研究では、専門分野分類の設定が重要な点であることに鑑み、学協会へのアンケート調査等を実施して、適正な分野分類を設定し、これを両引き辞書の編成や分野別分析などに用いた。更に、辞書更新のシステム化について、英文と和文の双方から、出現頻度統計と連鎖的検索の手法を用いて新出語を析出する方式と、機械翻訳システム自体を応用してゆく方式について検討した。また、上記の各分析を総合的に見渡したところで、専門用語における和英翻訳、英和翻訳の特質と問題点について検討した。

機械翻訳の実用化、普及に当たっては、新語、特に新出専門用語の発見とその辞書登録が、結局において決定的役割を担うものと思われる。本調査研究が、このような専門用語辞書のデータベース化、その更新整備のシステム化に対して、基礎的データと設計上の指針を与えるものとなれば幸いである。

16.1.1 目的

機械翻訳システムの実用的な翻訳精度・性能は、翻訳手法の高度化という要因もさることながら、むしろ専門用語辞書の充実の度合に依存するところが非常に大きいということが、機械翻訳システムの実用事例の積み重ねの中で、改めて確認されつつあるようである。したがって、電子化された専門用語辞書をデータベース的観点に基づいて、効果的、効率的に作成し、またこれを常時更新してゆく必要がある。この場合、こうした専門辞書の拡充・維持管理を、これまでのシステムのように、個別ユーザの問題であるとして放置してしまうのではなく、翻訳システムの供給者における経常的ユーザサービスの一環としてとらえた上で、そのための辞書データベースの整備手法を検討するのが有効である。

本調査研究では、先端産業各分野における専門用語の発生・使用状況を実態的に調査・分析することにより、専門用語辞書のためのデータベース構築・運用に関する基礎データを提供し、同時に、このような辞書管理のシステム化における有効な方法、体制について提言することを目的とする。

平成3年度には、英日機械翻訳用の辞書データベースの整備を研究主題として、エレクトロニクス、情報、材料の3専門分野を中心とする、250学協会の5,700の英文論文から400万語を抽出して、用語の出現頻度、共出現頻度、頻度の時系列変化等について調査分析を行い、電子化専門辞書整備の方法を検討した。

平成4年度においては、昨年度の英文専門用語に関する分析を更に進めるほか、新たに日英機械翻訳のための辞書の整備に関して調査研究を行い、この両者を総合して、英日・日英双方向の機械翻訳に適用できる「和英両引きの専門用語辞書データベース」の構築・更新のシステム化について、調査研究を実施した。

機械翻訳システムのための専門用語辞書の更新・維持管理においては、日々新たに大量に発生する先端的専門用語を検出し、それらを翻訳用辞書へ反映させてゆくことが、翻訳システムの実用性能の決め手ともいえることは、上述のとおりである。平成3年度の調査研究結果として、こうした先端的専門用語に関する辞書の整備に当たっては、英語文献をソースとする「英→和」翻訳用の辞書の整備と、日本語文献をもとにする「和→英」翻訳の辞書整備とは、これらを別個のものとして扱うのではなく、英和と和英を統合化して考えるのが有効であるとの結論を得た。確かに、国際的情報流通の進展が著しい現今の環境下において、先端的専門用語については、ことさら翻訳需要が高く、しかもこれは「英語→日本語」翻訳と「日本語→英語」翻訳の要求が同時に発生するはずである。

本調査研究では、このような問題認識に基づいて、英文と和文の専門用語に関する実態調査を行い、それらを比較検討しながら分析して、英日と日英の専門用語辞書データベースの効率的な整備・更新の方法を明らかにする。即ち、原文献から専門用語を能率的に検出・検定するための手法と、英日・日英辞書の相互変換の可能性に関する研究を実施する。これに基づいて、この種の辞書システムに係わる基礎的データと有効な管理手法や重要な知見を提供することにより、機械翻訳システムの精度の向上や自然語処理に係わる各種システムの開発等に幅広く貢献するものとする。

16.1.2 実施内容

本調査研究は、先端的な専門用語を迅速・適確に見つけだして、機械翻訳用辞書にそれを経常的に反映させてゆく方法について検討するものである。ここでは、和文と英文の電子化

文献と、電子化用語辞書を応用して、専門用語の使用実態を統計的に明らかにしながら、検討を進めた。

本調査研究の実施主体である科学技術情報研究所（J-TIES）は、これまで欧米の機関・企業と提携した調査業務を多数実施しており、先端的産業・技術の動向全般にわたって基礎的な資料、知識を集積している。本調査研究においては、こうした蓄積をもとにし、更に関係諸方面から必要な資料を収集し、これに対して分析を加えていった。一方、こうした過程全体にわたって指導を仰ぎ、また分析結果の評価を担当頂くべく、専門家により構成された「評価委員会」を組織した。即ち、委員会における調査研究全体の方針策定、評価分析と、J-TIESにおける調査作業の実施という分担関係により、調査研究を推進した。

本調査の具体的手順は概略次のとおりである。

(1) 昨年度調査研究結果のレビューと本年度調査計画の策定

先端的専門用語辞書の更新管理においては、英和・和英の統合的処理が有効であることを確認し、この認識に基づいて本年度の調査研究計画を策定した。

(2) 専門分野分類の検討と設定

両引き辞書における用語の適用分野分類の設定に当たって、学協会宛のアンケート調査を実施し、これらをふまえて、自然科学分野を対象に22の専門分野分類を設定した。

(3) 電子化両引き辞書の設計と作成及び検索システムの開発

既存の英日対訳電子化専門用語辞書4種類について、訳語を整理統合して、和英・英和双方から検索できる辞書を設計、作成した。また、これを高速に検索できるシステムを開発付加した。

(4) 両引き辞書の統計的特性の分析

作成された両引き辞書について、英語用語の単語数、日本語用語の字種構成と適用分野の関係を統計し、分野別の専門用語の全般的傾向の把握を行った。

(5) 両引き辞書における訳し分け状況の日英比較分析

英和、和英の双方について、主として専門分野別の訳し分けが行われるが、こうした訳し分けの現状について統計的に分析し、また特異事例を検出、検討した。

(6) 訳語のつながりからみた類義語群の形成状況の分析

英和検索によって得られる和訳語（群）を、和英検索に投入して英訳語（群）を得るという過程を連鎖させることにより、類義語クラスターの析出を試みた。ここでの知見は、辞書更新システムや機械翻訳システムを設計する上で重要である。

(7) 調査対象文献の収集と電子化

わが国の代表的な学協会誌等の論文要旨（和文・英文）と通信社による配信記事（和文

・英文)の相当量を収集、電子化ファイルとし、両引き辞書との突き合わせ、出現頻度統計等の実験対象データに用いた。

(8) 日本語分かち書き手法の調査と分かち書きシステムの開発

和文文献から専門用語を抽出するためには、まず和文の分かち書き処理が必要である。このとき、用語辞書を用いずに、字種構成から判別して高速でなお一定の精度が得られる手法について調査の上、分かち書きのシステムを開発して、本調査研究に用いた。

(9) 日本語文献における専門用語検出システムの検討

和文の文献から専門用語を抜き出し、これを辞書と突き合わせながら、新出・未登録の専門語であると確認の上、辞書登録を行うという一連の過程について検討した。

(10) 日本語専門用語の出現頻度統計とその分析

和文の上記調査対象文献について、日本語の専門用語の出現頻度の統計調査を行った。この際、分かち書きの問題があるので、辞書登録用語との全体・完全一致のほか、部分的な一致についても検索して統計し、その結果を検討した。

(11) 出現頻度と連鎖検索による日英両用辞書の更新方式の検討

上記のような本年度の調査分析を、昨年度の調査研究結果と総合して、日英両用の専門用語辞書の更新・維持管理の方式について検討した。

(12) 専門語の検出と辞書更新に対する機械翻訳システムの応用可能性の検討

本調査研究では、機械翻訳システムの運用とは一応独立の工程として、専門用語辞書整備の方法を検討してきた。しかし、今後は機械翻訳システム自体を応用した辞書更新方式も有効と目されるので、この点についても検討を加えた。

(13) 専門用語における和訳語、英訳語の特質と問題点の総合的検討

上記のような調査分析をふまえて、わが国と海外における先端的な専門用語発生メカニズム、またその国際間での流過程など、現状における基本的な枠組みにさかのぼって、翻訳システムのあり方との関係を検討した。

16.2 調査研究結果の概要

16.2.1 日英両用電子化辞書の作成とその諸特性

既存の電子化英和对訳専門用語辞書3種と、昨年度の調査研究の過程で得られた新出専門語を収録した独自の辞書の4者について、訳語や適用分野区分等を整理しつつ統合して、和→英、英→和の検索が可能な両引き辞書ファイルを開発した。この辞書は、訳し分けの関係を分解して、「1英文用語対1和文用語」という対応関係をもつものとし、全体で約27万エ

ントリーからなる。異なり語数では、英語21万5千語、日本語21万2千語になる。

全体的にみると、日本語の方が英語よりも1.4%ほど語彙が小さいということであるが、ここでは、日英ほぼ同数であることが注目されよう。英→和における訳し分けと、和→英における訳し分けが、全体としてはほぼ同数という結果であるから、この点に関するかぎり、日英翻訳と英日翻訳の難易度は同等ということでもある。

次に、英文用語については、用語を構成する単語数、和文用語については、その字種構成（漢字・ひらがな・カタカナなど）と、適用分野の関係について統計して分析した。これは、例えば漢語の専門語の多い分野と、カタカナ語の多用される分野といった比較検討であり、分野別専門用語辞書の整備方針に関して、有用な知見が含まれる。

16.2.2 和訳語・英訳語における専門分野別訳し分けの状況調査

昨年度の調査研究では、一つの英文専門用語に対して、複数の和訳語が当てられるという、訳し分けの状況を2種の既存電子化辞書について調査した。この場合、訳し分けが専門分野別になっているのであれば、翻訳処理上の問題はかなり軽減される。即ち、翻訳対象文献の分野区分を予め指定してやればよいからである。しかし、調査の結果をみると、訳し分けが分野ごとに特定されるとはいえない例も相当あった。

そこで、本年度では、4種の辞書を統合した和英両引き辞書を用いることにより、対象範囲を拡大すると同時に、日本語からみた英訳語の訳し分けという方向についても調査を行い、英→和と和→英の比較を行った。その結果、例えば土木や電気電子分野では、英→和に比べて和→英の際の訳し分けが多い（日本語語彙が小さい）のに対して、応用物理では逆に英語語彙の方が小さいといった、分野別の状況が読みとれる。専門分野別に機械翻訳辞書を編集・整備するに当たっては、このような分野別の特性を見極めた上で、必要な調整を行う必要がある。

16.2.3 訳語検索の連鎖による類義語クラスターの析出と辞書のシソーラス化

今回の調査研究において開発した英和・和英両引き専門用語辞書の機能を応用して、英→和→英→和→・・・という訳語の連鎖的な検索を行って、翻訳語を媒介とした類義語クラスターの析出を試みた。この際、例えば「亜鉛」を和英で引いて「ZINC」を得、次に「ZINC」を英和で引いて「亜鉛」だけが得られるという「単純往復」が発見されるのであれば、翻訳にとっては、これが最も好ましい状況であるといえる。しかし、現実には、このような1対1の往復関係はそれほど多くないと思われる。むしろ、同義語や類義語そして訳し分けなど結合関係を介して、多数の用語がからみ合った状況も多いのではなかろうか。こうした仮説

に基づいて、幾つかの和英の用語を起点として、その関連用語をいわば手繰りよせてゆくという操作により、日英の専門用語がどのような結合関係を形成しているのかについて、その概略的感触を得るものとした。翻訳辞書の整備に当たっては、単純な対訳辞書的構成では、翻訳精度にもおよそ限界があると考えられ、ここに述べたような分析を通じて、用語間・訳語間の関連性を内蔵する、まさしくデータベース的なアプローチが必要であろう。これは、従来からシソーラスと称されてきたものに近いので、翻訳辞書のシソーラス化ということもできるが、利用目的を異にしており、したがって、構成や編成の手法に関して新たな研究を要するものであろう。

16.2.4 和文専門用語の抽出方式

和文の文献における専門的用語の出現状況を調査するについては、まず、単位語の抽出が必要である。即ち、分かち書きの工程である。現今、分かち書きの処理は、予め用意された辞書との突き合わせ処理を含む方式が有力である。しかし、辞書の構築方法自体を検討対象とする本調査研究のような場合には、辞書検索を内蔵する分かち書き処理は有効ではない。したがって、ここでは、辞書を用いない分かち書きの手法について調査・検討し、国立国語研究所で開発された方法を参考にして、字種の変わり目に注目した文章分割の方法を設計し、その処理プログラムを開発した。端的には、例えば漢字の連鎖、カタカナの連鎖などを語として設定してゆく方法であり、これにより抽出された文字列を辞書と突き合わせながら、専門用語を検出してゆく手順を検討した。

即ち、分かち書きにより取り出された「語」を、一般語からなる辞書と上述の専門用語辞書とに突き合わせて辞書末登録の語を発見し、これらについて、更に詳細に検定するという手順、方法について、電子化された和文文献を用いて実験しつつ検討した。なお、辞書との突き合わせ処理では、文字列の部分的一致についても検出するようにして、検索精度を高めた。また、未登録語の検定では、前項に述べた連鎖検索の手法の適用可能性を検討した。

16.2.5 和文専門用語の出現頻度分析

前項における調査検討結果をふまえて、和文の電子化文献における専門用語の出現頻度の統計を行った。この際、用語例が辞書登録用語と部分的に一致する場合についても統計し、各用語の独立／連接使用の実態についても明らかにするものとした。統計対象の文献データとしては、通信社の一般記事と情報関連分野の科学技術論文要旨を用い、それぞれにおける高頻度用語を調査し、また上記の部分一致と完全一致の用例数の比率による統計も行った。例えば、他の用語との连接的使用の多い用語については、翻訳の際に注意が必要であり、む

しろ連接例の方を既に熟した専門用語と認定し、これを新たに辞書登録するといった対応が有効である。本統計調査によって、このような専門用語辞書維持管理に関する有用な知見が、その具体例とともに得られている。

16.2.6 専門分野分類の検討

上述における両引き辞書の開発や、専門用語の抽出、出現頻度統計等全般にわたって、専門分野別の検討が重要である。そして、その分野分類自体が、本調査研究にとっては前提的に必要とされる重要な項目であった。これは、機械翻訳のための専門用語辞書の分野別編成とも関連する問題である。そこで、本調査研究の一環として、自然科学系の学協会に対してアンケート調査を実施して、その見解を徴した。これを大学の学科・講座編成、学会会議の研究連絡委員会構成、既存の専門用語辞書にみられる分野分類等と照合して、独自の分類体系を設定し、本調査全体についてこれを適用した。

16.2.7 専門用語辞書の更新管理システム化

情報化と技術革新の進展が急速な今日的状況において、日ごとに発生する専門的用語群に対して、翻訳用の用語辞書は、これに遅滞なく追従できるものでなければならない。この観点からして、翻訳用専門用語辞書のデータベース化と、その更新管理のシステム化は不可欠であろう。本調査研究では、この点について次の2方面から検討を加えた。

その第一は、出現頻度統計と連鎖的検索による検定を活用した、日英両用専門用語辞書の更新方式の検討である。ここでは、例えば、翻訳システムの開発・供給者において、和文と英文の文献における用語の出現状況を、適切な統計を採取しながら常時監視し、十分な検定作業の上で、新出用語を認定・登録したのち、これをユーザに経常的に配布してゆくといった体制が想定されることになる。

第二は、専門語の検出と辞書更新に対する、機械翻訳システムそのものの応用という視点である。本調査研究では、専門用語辞書の整備を、機械翻訳システムの運用とは一応独立の工程として運用することを念頭に検討してきた。しかし、今後は機械翻訳システム自体を応用した辞書更新方式も有効と目される。そこで、「辞書更新管理システムとしての機械翻訳システム」という観点に立って、その機能要件、可能性等について検討を加え、今後への示唆を与えるものとした。

16.2.8 専門用語における和訳語、英訳語の特質と問題点の総合的検討

上述のような調査分析の全体をふまえつつ、わが国及び外国における先端的な専門用語の

発生メカニズムや、その国際間での流通・浸透の過程などについて、先端科学技術分野を中心として、現状の基本的な枠組みを明らかにし、こうした現状に有効に対応できるような機械翻訳システムと、そのための専門用語辞書更新システムのあり方を検討した。翻訳とは基本的に文化的な所為であって、各国の社会的環境が、その濃淡はともあれ、翻訳対象の文献に反映されている。したがって、科学技術文献の翻訳においても、その研究の社会的・文化的背景に対する一定の理解が必要である。このような意味で、翻訳用辞書というものは、単なる言葉の置き換え表ではない。本項では、こうした視点もふまえて、専門用語の翻訳に係わる基本的な問題を検討した。

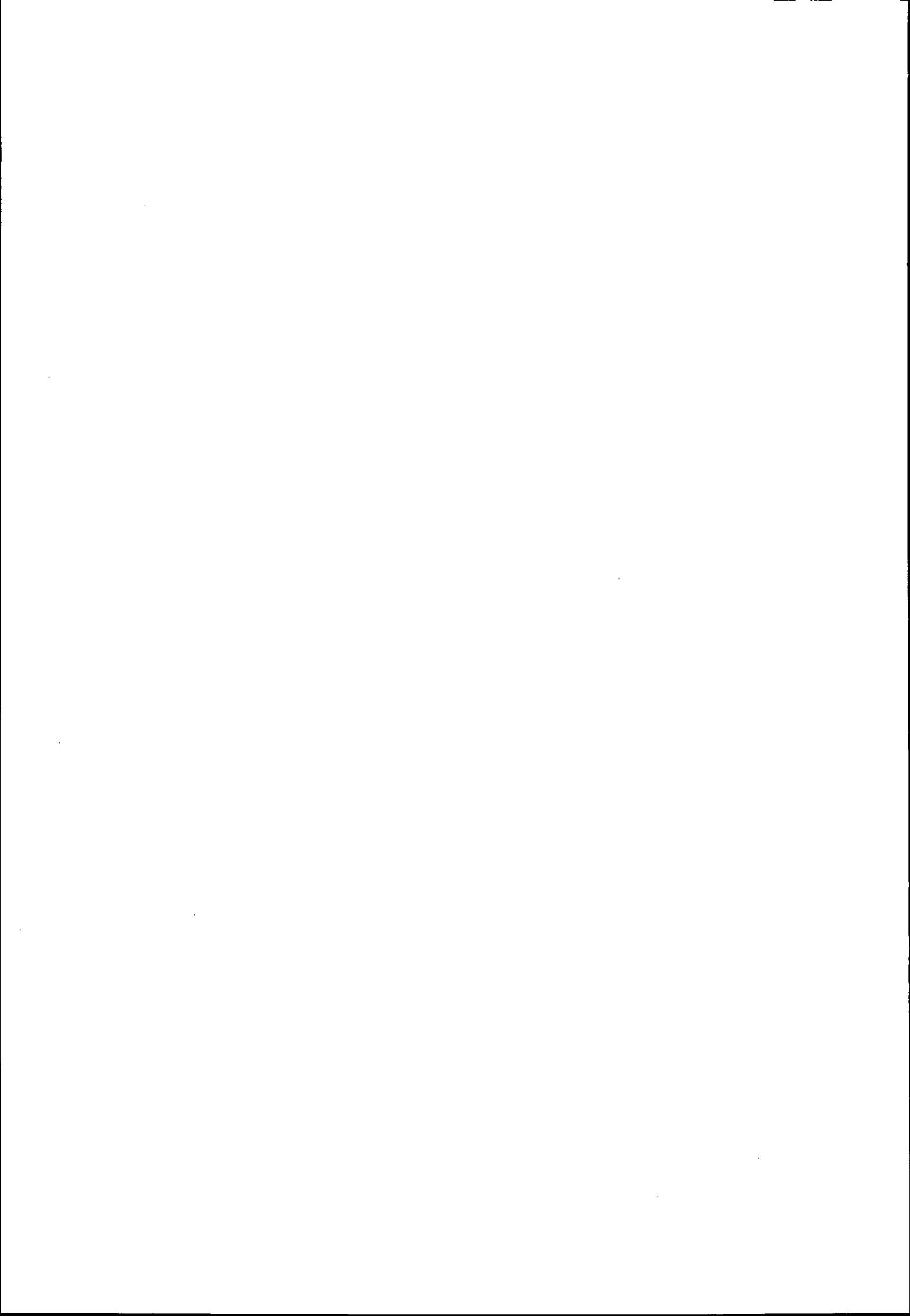
16.3 今後の課題

これまで開発を手掛けてきた専門用語辞書システムと機械翻訳辞書システムにつき、これらの包括できる分野の一層の拡充をめざすとともに、システムの一層の効率化を図り、予想される多くのユーザの要請に広く応えられるようにすることがのぞまれる。



17 マーケティングコードの英文データベースの構築

株式会社 帝国データバンク



17. マーケティングコードの英文データベースの構築

17.1 日本産業分類と帝国産業分類

産業分類は、比較・分析の目的に見合って適度に細分化されたものであること、時系列での分析を行うために安定したものであること、普遍性をもったものであることなどの条件を満たす必要がある。統計資料を作成する場合、各人が全く異なる定義の下に設定された産業分類を使用すると、それぞれの資料を比較・分析することが非常に困難になる。

これを考慮して、帝国データバンクでは、総務庁（旧行政管理庁）の日本標準産業分類をスタンダードコードとし、これに商用データベースとして活用するためのノウハウを付加した帝国産業分類（TEIKOKU DATABANK INDUSTRIAL CODE: 以下T I C）を設定した。

このT I Cは、日本標準産業分類を母体に、永年にわたり多業種に及ぶ企業の経営状態について調査活動を行ってきた帝国データバンクの業種分類ノウハウを加味したもので、1956年5月に完成、運用が開始された。

その後、企業情報のデータベース化による業種分類ニーズの多様化に伴い、1975年10月までに四次にわたる部分改訂を行い、現在に至っている。

17.2 T I C の特色及び体系

17.2.1 T I Cの特色

T I Cは、企業の経済活動に関するフルレポートである「信用調査報告書」、企業財務に関するデータベース [COSMOS 1]、企業の概要に関するデータベース [COSMOS 2]、企業の倒産件数の集計や産業ごとの景気動向指標などの統計資料など、帝国データバンクの保有する企業情報のあらゆる分野で使用されている。

例えば、[COSMOS 2]では、主業、従業として1企業に対して最大二つのT I Cが付加されている。その結果、企業のセールスランキング（同じ分類の中で売上げが何位に位置しているか）の提供や『東京都港区にある情報サービス業を営む企業』あてのダイレクトメール用のアドレスタグの出力等、さまざまな活用が可能となっている。

17.2.2 T I Cの体系

T I Cは、大分類／中分類／小分類の3段階分類から成り立っている。

大分類は 14項目 アルファベット [A] ～ [N]

[例] A 農業

中分類は 90項目 アラビア数字 2けた [01] ~ [99]

[例] 12 原油, 天然ガス鉱業

小分類は 673項目 アラビア数字 3けたから5けた

[例] 121 原油鉱業

321-3 ガラス容器製造業

409-21 石油卸売業

17.3 帝国データバンクマーケティングコード作成の意義

産業の進展で、ニューサービスや新業態の増加に伴う分類体系自体に対する見直しの声や、企業のきめ細かいマーケティング戦略導入により、T I Cのマイナーチェンジでは対応しきれない、更なる業種の細分化が望まれ始めた。

帝国データバンクでは、T I Cの細分化要望に応え、新しい産業分類体系 [マーケティングコード T M C] を設定した。T M Cは、ニューサービスや新業態の積極的な収録により、さまざまな変化を遂げる産業形態をできる限り包含した体系となっている。T M Cの特色は次のとおりである。

① 現状の産業情勢に則している。

コードが膨大で多岐にわたることにより、運用に支障をきたしてはいけぬ。コード化を判断する際には、現状の市場規模及び将来性を材料としてやみくもな設定は避けている。

② 実用的でマーケティングニーズに合致している。

マーケティングコードは、標準産業分類、T I Cの小分類はもとより、営業部門で収集する業種によるデータベース抽出条件ニーズなどを、さまざまな角度から検証し、活発に活動している業種、サービスを中心に収録している。

③ 実際の運用のしやすさを重視したコードである。

コードの運用に際しての最大のポイントは、コード探しの容易さである。データベースへのコーディングやそのメンテナンス作業が効率よくできるかどうかが生産性に大きく係わっている。そのためにはコードが適当な数であること、コードの名称が具体的で分かりやすいものであることが必要である。コード名称は、実際の当社の調査資料や一般的表現を参考にしつつ設定している。また、言いまわしや表現を換えたものも、類語として収録し、なるべく検索に個人差がでないように配慮した。

④ T I C小分類に則している。

T I Cが、帝国データバンク社員のみならず多くのユーザに企業情報データベースを通じて親しまれていることに着目し、その小分類 673項目を利用してマーケティングコードを設定した。5けたで構成されているT I C小分類に枝番2けたを付加し、7けたとしている。

⑤ 業態での抽出が可能である。

T I Cでは表現できなかった量販店、訪問販売、通信販売、賃加工等の一部「業態」にも小分類に沿ってマーケティングコードを設定し、抽出可能となっている。

17.4 T M C の 構 成

マーケティングコードは、

- 1) マーケティングコード及び名称
- 2) マーケティングコードに対応するT I C
- 3) マーケティングコード名称を他の表現で言い替えた類語より構成されている。

① マーケティングコード

マーケティングコードはT I C《5けた》に枝番《2けた》を付加した《7けた》で構成されている。総設定数は3,162である。

② 類語について

マーケティングコード名称を言い替えたものである。例えば、電化製品卸、電気製品卸、家庭電化製品卸は類語であり、どの言葉からでもマーケティングコードに行き当たるように設定されている。

17.5 英文化の背景

ジェットロではアンケート方式で輸入取引関心企業・輸入取引関心商品のデータベース化に着手していた。ジェットロの構想は海外のバイヤーが輸出したい商品で検索し、日本側で同商品を取り扱っている業者を抽出できるデータベースの構築であった。

これに対して当社では、企業情報データベース(COSMOS2)の和文・英文データベースを活用するようジェットロに提案した。

当社の提案に対しては関心を示しながらも次の点が問題となった。(当社では帝国産業分類「日本標準産業分類に準拠」を使用)

- ① ジェットロの構想では単品で検索できることが前提である。
- ② 産業分類は統計向きで分類が粗く、構想外である。

③ 商品分類はH Sコード（88年から国際的に採用された新しい関税分類コード）が望ましい。などが焦点になった。

当社では、H Sコード（国際関税コード）、S I Cコード（Standard Industrial Classification Manual）に関して研究を続けていたが、ともに日本の商品コードとしては日本特有の商品に対応できず、独自のコードづけを余儀なくされるなど、いずれのコードを採用しても体系の維持は難しく実情に合わず、現データベースで使用している産業分類との整合性からも両コードの採用を見送った。

17.6 英文化の目的

海外から日本企業の情報に関心が高いとのことで、ダイアログ（DIALOG Informetion Services, Inc）から熱心な勧めがあったほか、海外及び国内のディストリビュータからも英文C O S M O S 2の海外提供について多くのアプローチがあった。

当社ではデータベースの国際流通の側面から、1991年4月からダイアログをディストリビュータとして日本企業C O S M O S 2の英文版5万社（現在は18万社）のサービスを開始した。次いで国内においても、日本経済新聞社（日経テレコン）、ジー・サーチをディストリビュータとしてサービスを開始した。

前述のとおり、初期の利用は単に日本の企業情報を知りたいとするニーズからエンドユーザが輸出したい商品（希望する商品名で検索）を扱うパートナーを探したいとするニーズは拡大、多様化されつつある。

これらのニーズに応えるには言葉の問題を解決することが先決になるが、現状ではニーズに応えられる既存の体系も手段もなく、独自に英文化することは避けられない。

しかし、ニーズのサイズも把握できない状況では費用対効果が最大のネックであることが議論の対象になったが、わが国から海外向けに日本企業の情報を英文で提供しているのが当社のみであることから、これらの国際的ニーズに応えることは避けて通れないとの結論を得た。

極言すれば、現在のダイアログ及び国内のディストリビュータを通じてのサービスが初期の日本企業の情報を知る手段としての利用に加えて、マーケティングコードをキーにした検索は単にサービス内容の拡大に留まらず、日本貿易摩擦解消の一助にもなるものと考えたからである。

このような背景をふまえてマーケティングコードの英文化とデータベースの構築を通じてコードマスタの整備に着手した。

17.7 S I C 対応の意義

課題応募時においては単にマーケティングコードを英文化するとの発想であった。

企画段階における検討ではHSコード（関税分類コード）、S I Cコード（Standard Industrial Classification Manual）の対応も考慮した。しかし、日本標準産業分類がベースになっている関係からわが国特有の商品などの対応に問題が多く、本課題が完成した後に再度、検討することで保留にした。

本課題が採択され、委託契約締結後の最初の取り組みとしてマーケティングコード英文化に向けての基礎調査を行うこととした。

委託テーマはマーケティングコードを海外のエンドユーザが使用可能な英文翻訳と利用しやすい体系の整備を主目的とした。

主旨は次のとおりである。

・ダイアログの企業データベースにおける標準産業分類の採択状況

- ① ダイアログで提供されている企業データベースのうち、米国企業データベースの大部分（92%）では、米国S I Cコードもしくはこれを細分化したものを採用している。
- ② ダイアログで提供されている企業データベースのうち、米国以外の企業データベースでは、米国S I Cコードを採用しているものが38%、自国S I Cコードのみが16%（英国、日本）、K O M P A S Sコード（U C S準拠）を採用しているものが22%である。
- ③ 以上の結果から、米国のユーザを念頭において産業分類コードをつけるのならば米国S I Cコードを付与することが望ましい。
- ④ マーケティングコード英訳作業と並行してS I Cコードへの変換作業を行う可能性を検討すべきである。

・マーケティングコードの英文化及び米国S I Cコードの比較

- ① 対応する分野の米国S I Cコードを日本語に訳す。
- ② マーケティングコードを英文化するに当たり、該当する訳語があればそれを採用し、末尾に対応する米国S I Cコードを付加する。
- ③ マーケティングコードに該当する米国S I Cコードがなければ、独自の訳語を作成する。

なお、ヒアリング時にダイアログの海外での利用が日欧各50%であるとの報告もあり、同報告書を受けて社内プロジェクトで検討した。

検討の結果

- ① 英文化の主旨が国際的ニーズに応え、データベースの国際流通を促進することが基本であること。

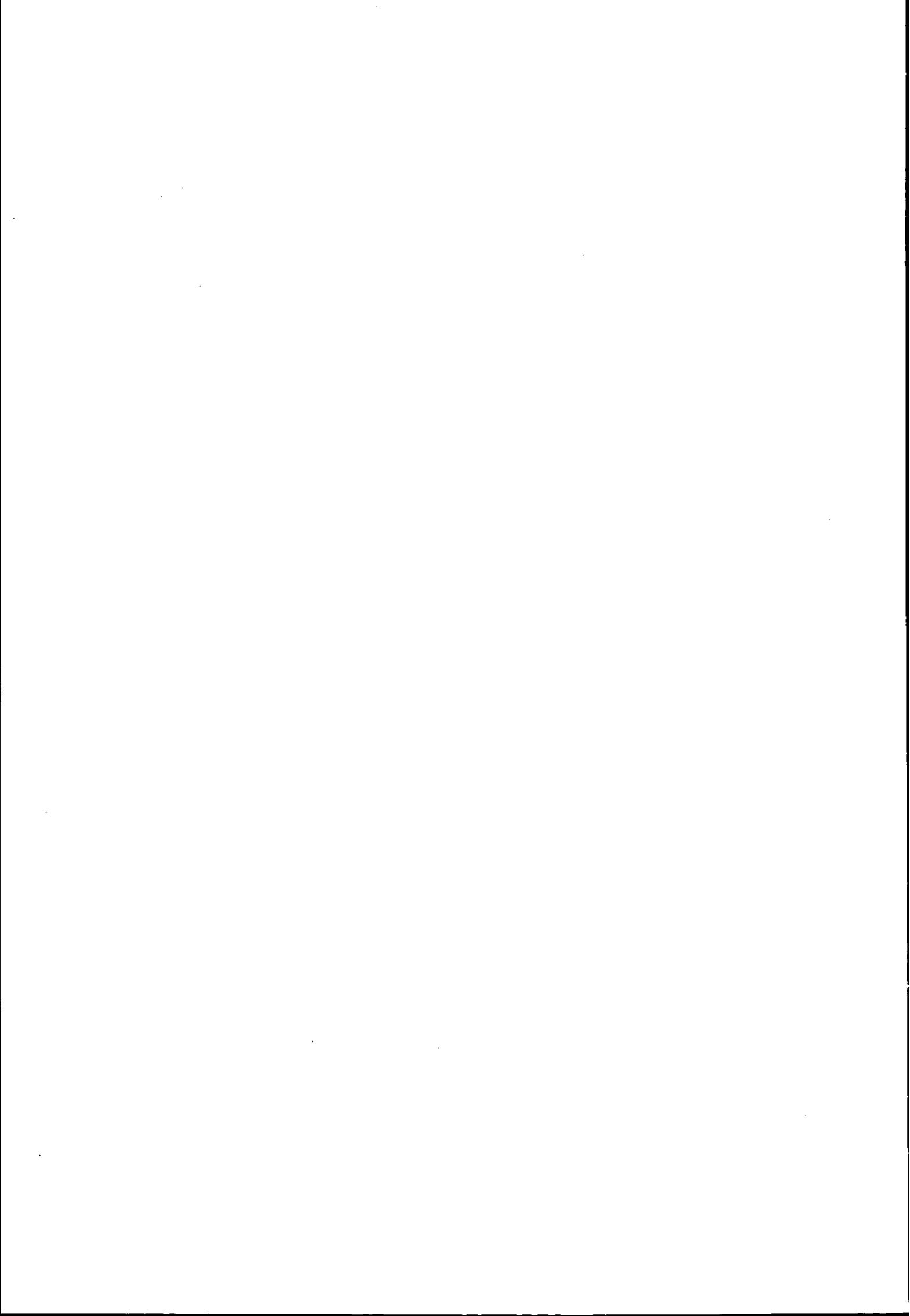
② 企画段階での検討のとおり、国際的に流通する米国S I Cコードとの対応が可能であれば対応可能分のみでも促進した方がよいこと。

③ マーケティングコードの英文化が主体であるが、米国S I Cコードと対応できた分は米国S I Cコードをキーにしたマーケティングコードとの対応表も作成すること。

以上の結論をもって、当初企画した中で最も使用頻度の高い米国S I Cコードとの対応をつけることにした。

18 安全研究における多重シソーラス・システム
構築のための基本安全用語データベースの開発

株式会社 紀伊國屋書店



18. 安全研究における多重シソーラス・システム構築 のための基本安全用語データベースの開発

18.1 背景と目的

「安全」問題は、労働、輸送、土木、建築、原子力、環境、化学、食品、医薬品に至る科学技術全般の広い分野で扱われている。したがって、安全問題に関する情報を収集するためには、これらの分野それぞれをカバーしている多数の文献データベースを探索しなければならない。現在、文献データベースは、それぞれ個別の索引語体系によって管理されているため、同一の概念に対する索引語は個々のデータベースによって異なっている。このため、安全問題においては、一つのテーマについての情報を得るために、データベースごとに索引語を調べなければならないという、たいへんな労力・時間を要してしまう。また、全文検索機能を用いて標題・抄録中の用語を検索しても、同一の概念に対し分野固有の異なった表現があり、複数のデータベースに対し精度の高い検索を行うことは困難を極める。

「安全」情報について、このような問題をも解決するためには、「安全」の中心的な概念について、分野ごとに又はその分野を代表するデータベースで使用されている用語・索引語間で相互変換を可能にする必要がある。このためには、上記の分野における「安全」概念に密接に関連する用語をデータベースごとに再整理・体系化し、更に個々のデータベースで使用・管理されている「安全」関連の索引語、分野固有の表現とのリンケージの情報を有する、言い替えれば複数のシソーラス、複数データベース間での用語の相互アクセスを可能とする「安全」分野の基本用語知識ベースを開発することが求められる。

このような知識ベースの実現には、下記に挙げる調査・用語収集が必要とされる。

- 1) 安全用語の事前調査
- 2) 安全情報のデータベースの調査及び基本安全用語の収集、分類
- 3) 上記の用語による複数データベースに対するクロスファイル探索、安全の観点からのデータベース及び用語の評価
- 4) 安全情報の収集に有効な複数のデータベースを対象とした安全関連用語の収集

本研究は以上の点を主眼におき綿密な調査・分析を実施したものである。

18.2 実施体制及び実施内容

18.2.1 実施体制

この調査・開発を推進するに当たっては学識経験者と専門家で構成される「安全用語検討委員会」を設置し、この委員会の協力を得て、具体的な調査活動を行った。なお、委員会の構成メンバーは次のとおりである。

委員長	藤原 鎮男	神奈川大学	知識情報研究所	所長
委員	山本 晴彦	神奈川大学	理学部 応用生物科学科	教授
委員	河村 正一	神奈川大学	理学部 化学科	教授
委員	高木 伸司	神奈川大学	理学部 化学科	教授
委員	小幡 行雄	神奈川大学	理学部 情報科学科	教授
委員	後藤 智範	神奈川大学	理学部 情報科学科	助教授

18.2.2 実施経過

平成4年7月15日	委託契約の締結
平成4年8月11日	第1回 安全用語検討委員会（研究調査の概要説明）
平成4年9月22日	第2回 安全用語検討委員会（医薬品安全用語の事例研究を説明）
平成4年11月5日	第3回 安全用語検討委員会（FID' 92における研究発表の報告）
平成4年11月24日	第4回 安全用語検討委員会（ENERGY SCI. & TECH. からの用語抽出の結果について報告）
平成4年12月15日	第5回 安全用語検討委員会（EMBASEからの用語抽出の結果報告）
平成5年1月26日	第6回 安全用語検討委員会（CAS からの用語抽出の結果報告）
平成5年2月3日	第7回 安全用語検討委員会（本年度問題点について検討）
平成5年2月17日	第8回 安全用語検討委員会（来年度の課題について検討）

18.2.3 実施内容と成果

(1) 実施内容

本研究は、安全情報に関する複数のシソーラス、複数データベース間での用語の相互アクセスを可能とする「安全」分野の基本用語知識ベースの開発を意図し、下記の四つの基礎的な調査研究を実施した。

- 1) 事前調査としての医薬品の安全用語の評価
- 2) 安全情報のデータベースからの基本安全用語の収集、分類

3) 上記の用語による複数データベースに対するクロスファイル探索, 安全の観点からのデータベースの大規模な調査, 及びこれに基づく用語の客観的評価

4) 安全情報の収集に有効なデータベースを対象とした安全関連用語の収集

18.3.4の用語収集においては, EMBASE (医学・生物学), CASearch (化学), ENERGY SCI. & TECH. (エネルギー分野) の3種類のデータベースについて, 安全関連用語を収集することができた。

(2) 成果

実施内容の3) について下記の成果が得られた。

1) 「安全」に関する情報が, 80以上のデータベースに分散していることを実証的に示した。

2) 「安全」概念を構成する九つの基本テーマについて, 80の個々のデータベースのレコード数を計量的分析することにより, 下記の事実が明らかになった。

a) MEDLINE, EMBASEなどの包括的な医学分野のデータベース, ENERGY SCI. & TECH. などの包括的な非医学分野のデータベースに, 「安全」情報が量的に多く含まれていることが判明した。

b) 一方, TOXLINE, CHEMICAL, SAFETY NEWSBASEなどの毒性・化学物資に関する専門的なデータベース, 及びENVIROLINE, POLLUTION ABSTRACT などの環境分野の専門的なデータベースには, 「安全」情報が高い比率で含まれていることが判明した。

また, 実施内容の3), 4) から下記の事実が判明した。

本研究で使用した, NIOSHTICから抽出・類型化した9種類のカテゴリーの「安全」用語を組み合わせ検索することにより, EMBASE (医学・生物学), CASearch (化学), ENERGY SCI. & TECH. (エネルギー分野) の3種類のデータベースでは, 「安全」情報だけを高い精度で得ることができた。

更に, 実施内容の4) から下記の成果を得た。

a) EMBASE (医学・生物学), CASearch (化学), ENERGY SCI. & TECH. (エネルギー分野) の3種類のデータベースから, 「安全」に関連する約200の用語を抽出できた。

b) これらの用語には, 同一の概念でも分野又はデータベース固有に使用される表現があることが明らかになった。

18.3 安全研究における基本安全用語データベースの収集

18.3.1 安全研究に関する重要語の収集・評価の事例研究 (Lovastatine に関する文献事例によ

る医薬品安全用語の収集と評価)

多くの分野で重要な課題であり、かつ学際的なテーマである「安全問題」について既存の多種多様なデータベースのより一層効果的な利用を図るシステム構築の第一段階として、検索用語を選定するための事例研究を行った。

本研究では、事例研究の調査対象分野として医薬品分野を採り挙げ、医薬品の安全性に関連する汎用語を、専門家の経験により選定・分類した。これらの用語の有効性を客観的に、評価するために、特定テーマについて文献検索を行い、得られた文献(抄録)を「安全」を主題として扱っている文献と、そうではない文献とに分類、それぞれの文献群で、専門家により選定された用語の出現頻度分析を行った。

専門家の経験によって選定された医薬品の安全性に関する用語は約 100語であった。調査対象テーマとしては、医学分野のデータベースには多く文献が含まれていると考えられる *lovastatine* (高コレステロール低減の治療薬) を採り挙げ、DIALOGシステムにより検索を行った結果、485文献を得た。これらの文献を上述のように、「安全」に言及している程度により分類しその結果、安全を主要な内容としている記述程度 2 に分類されたものは46件で全体の約10%に相当し、そのうち26件、56%は副作用についての症例報告であった。その他、薬品の市販後における副作用経験をまとめたものが8件、毒性試験で副作用となる可能性のある毒性を指摘したものが7件などが主なものであった。これに対して、安全性とは関連の低い文献が、245件、52%であった。

安全を主たる記載内容とする文献、即ち、安全についての記述程度 2 とした文献群と、安全について記述程度 0 の文献群との間で、各レコードの標題、抄録、索引語(Mesh ディスクリプタ)の項目において、医薬品の安全用語の出現頻度を比較することにより、個々の用語の安全用語としての妥当性評価を試みた。これらの用語には複合語も多く、単数形と複数形の綴の相違など語尾変化をさせて使用する語も少なくないため、単語に分割し、標題、抄録、索引語での出現頻度、及び出現頻度の全単語総数中の比率、即ち相対出現頻度を算出した。

標題項目では上述の調査対象の単語の出現頻度はすべて1~2と低いため、出現頻度分析による評価から除外することとした。

抄録項目では、*acute, adverse, events, safety, symptoms, toxicity(toxicities, toxicic), trial(trials)* の7単語は、記述程度 2 の文献群での相対出現頻度が、記述程度 0 の文献群のその10倍以上であることが判明した。一方、出現頻度が記述程度 0 群の方が記述程度 2 群よりも10倍以上となる単語はなかった。

索引語項目についても、標題、抄録項目で採用したのと同様な分析を行った。drug, effec

ts, trials の3語を除くすべての語は、記述程度2群での出現頻度が記述程度0群での出現頻度が10倍以上であることが分かる。このことはacute, adverse, failure, toxicityの4語は、索引語項目では安全性について言及した文献群において、安全性について言及していない文献群よりも高い頻度で出現することが明らかにされた。

18.3.2 基本安全用語の収集（安全に関する汎用語の収集と評価）

一般に安全問題のような既存のさまざまな学問分野からの研究を要するテーマは、情報アクセスという観点から、複数データベースにわたる情報の分散、データベース間での用語の不統一、という問題が提起されている。

「安全」情報について、このような問題を少しでも解決するためには、「安全」の中心的な概念について、分野ごとに、又はその分野を代表するデータベースで使用されている用語（索引語を含む）を調査する必要がある。本研究では、「安全」についての基本用語を収集・選定することを意図し、下記に挙げる調査及び用語収集・評価を行った。

1) 「安全」情報を含むデータベースの現状

2) 「安全」情報を収録しているデータベースからの基本用語の収集

「安全問題」は化学物質と密接な関係があり、主として化学物質の関係する「安全問題」については、既にデータベースは列挙されている。そこで、本研究では現在利用可能なオンライン検索システムのデータベースマニュアル、データベースのディレクトリなどを用いて、「安全問題」を集中的に収録しているデータベースを調査した。その結果、NIOSH, HSELINE, CSNB, CIS Abstracts の四つのデータベースはいずれも労働安全衛生分野の情報を主要な収録対象としている書誌データベースであることが分かった。

本研究では、労働安全衛生分野を主要な収録対象としているデータベースの中で最も蓄積量が多いNIOSH（以下NIOSHと略す）を用語収集の対象データベースとして選定した。書誌データベースには、用語データを含む項目として標題、抄録、索引語などがあるが、用語収集の対象としてはより標準化された用語が含まれる索引語項目が妥当と考え、NIOSH データベースの1990年度の全レコードを対象に索引語項目に含まれる用語を収集の対象とした。

検索の結果、NIOSH の1990年の総レコード数は 4,294であった。4,294すべてのレコードを対象に、索引語項目からディスクリプタ（索引語）を抽出した。抽出された索引語を、大文字-小文字変換、記号削除などのさまざまな処理をし、アルファベット順にソートし、すべての索引語の出現頻度を算出した。索引語の総数（種類）は 2,435であった。

これらの索引語を2人の専門家によりそれぞれの観点から用語を整理・分析し、可能な限り包括的な概念を表す単語を抽出し、労働安全衛生の観点から基本となる約50の用語を選定

した。表18-1に選定された用語及びそのカテゴリーの一例を挙げる。これらの用語には労働安全衛生固有の用語も含まれているが、多くの用語は18.3.1で収集した医薬品分野の用語と重複していることが分かった。

表18-1 選定された用語及びそのカテゴリーの一例

安全性一般	risk, safety
危険因子との接触経路	exposure, inhalation, radiation
危険因子	contamination, pollution, dusts, fatigue, etc.
一般的被害	accident, poisoning, explosion, hazard, biohazard
健康被害	disorder, damage, disease, injury, symptoms, etc.

18.3.3 安全関連文献の包括的調査（安全研究に関する用語の複数データベース間での出現頻度分析）

安全概念を構成する幾つかのテーマについて、データベースごとのそれぞれの収録件数が、事前に分かっていたら非常に効率的に情報の収集が実現できる。

本研究は、安全概念を表現する幾つかの用語について、複数のデータベースにおいて、これらの用語のレコード件数をクロス分析をすることにより、そのテーマに対する個々のデータベースの関連性を明らかにすることを意図したものである。

安全概念を構成する基本的な用語としては、18.3.2の調査研究で得られた基本安全用語を9種類のカテゴリーに分類した。調査対象データベースの主題分野としては科学技術全般を採り挙げた。検索システムとしてDIALOGを使用し、クロスファイル検索用のデータベースであるDIALINDEXで科学技術全般に対応するテーマであるALLSCIENCEに含まれるデータベースを調査対象とした。このテーマには164種類ものデータベースが含まれ、また書誌データベース、物性データベース、全文データベースなどさまざまな種類がある。データベースの種類によりレコードを構成する項目は異なっているため、異種のデータベースを比較対象とすることは妥当ではない。そこで、本研究では、調査対象とするデータベースの種類を書誌データベースに限定し、更に特許及びpascalなど非英語圏のデータベースを除外し、合計85種類のデータベースを調査対象とした。

9種類の各カテゴリーに対応する検索式として、各カテゴリーに含まれる用語の前方一致及び論理和を用いて検索を実施した。検索結果を、9（カテゴリー数）×85（データベース

数) のレコード数の表にまとめた。

本研究では、各カテゴリに対するデータベースの特性を評価するために、個々のデータベースの保有レコード量、及び各カテゴリの全体でのレコード量を考慮し、下記の2種類の評価尺度を考案した。

$$\text{評価尺度 1} = \frac{\text{各データベースの各カテゴリに対するレコード数}}{\text{各カテゴリの総レコード数}} \times 100$$

$$\text{評価尺度 2} = \frac{\text{各データベースの各カテゴリに対するレコード数}}{\text{各データベースの保有レコード総数}} \times 100$$

上述の9×85のレコード数の表をもとに上記の両尺度の値を計算し、カテゴリごとにデータベースを分析し、カテゴリ間の比較を行った。

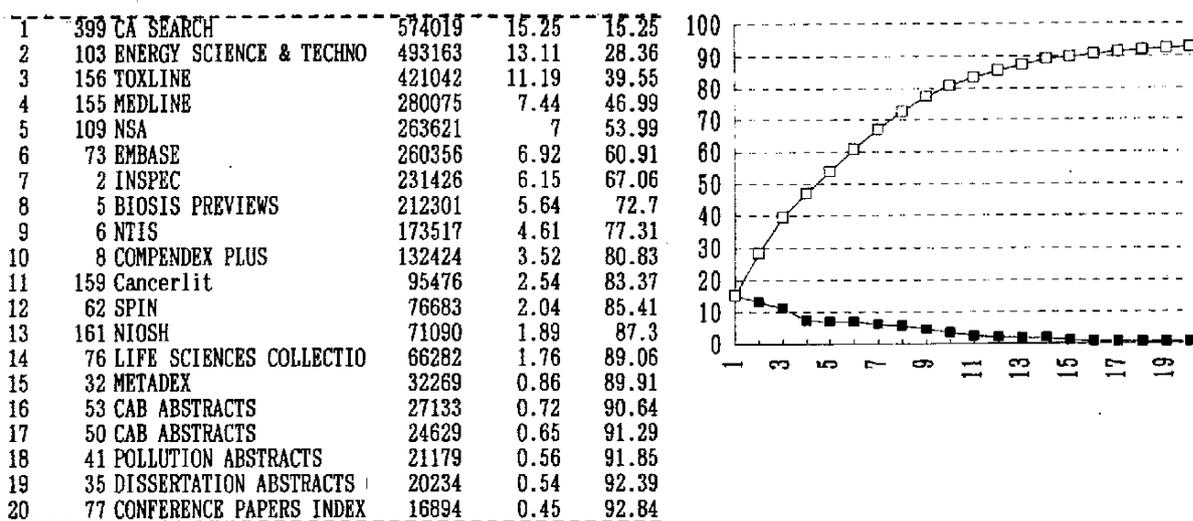


図18-1 カテゴリ1におけるレコード量, 尺度1, 累積比率 (上位20)

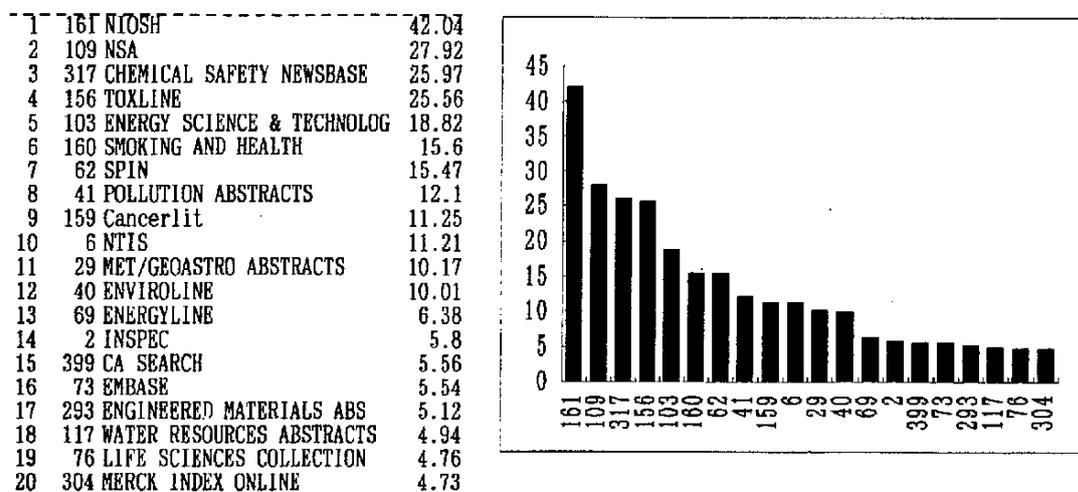


図18-2 カテゴリ1における尺度2 (上位20位)

具体例として、図18-1、図18-2にカテゴリー1即ち、危険因子との接触経路に関する汎用語（exposure, inhalation, radiation）において、尺度1及び尺度2の上位20位までのデータベース名とその値について、表・グラフを示す。

尺度1では、どのカテゴリーにおいても概ねMEDLINE, EMBASEなどの包括的かつ総レコード量の多いデータベースが上位にくることが明らかになった。一方、尺度2では、TOXLINE, SMOKING AND HEALTH, CHEMICAL, SAFETY NEWSBASE など特定主題領域だけを収録対象とし、総レコード量が比較的小量のデータベースが上位にくる傾向があることが判明した。また、いずれのカテゴリーでも医学・生物学分野のデータベースが上位にランクすることが明らかになった。下記の六つのデータベースは、二つの異なった尺度で共通して上位にランクされる回数の多いことが明らかになった。

OCCUPATIONAL SAFETY & HEALTH (NIOSH)	EMBASE (EXCERPTA MEDICA)
TOXLINE	MEDLINE
Cancerlit	NTIS

これらのデータベースは安全文献を収集するために非常に有効なデータベースと考えることができる。

18.3.4 主要データベースにおける安全用語の収集（ENERGY SCI. & TECH., EMBASE, CASearchにおける安全用語の収集）

これまでの研究において、18.3.2では安全問題を形成する主要な概念を表現している用語を、労働安全衛生分野のデータベースの索引語群から抽出し、カテゴリー化した。18.3.3では、これらのカテゴリーに対する科学技術分野の書誌データベースにおけるレコード件数を相互比較することによって、安全の観点から個々のデータベースの特性、更にカテゴリーに含まれる用語の特性を明らかにした。本研究は、次の段階である個々の既存分野のデータベースにおいて、安全概念を表現する用語を抽出・収集し、18.3.2の調査によって得られた安全に関する基本カテゴリーをもとに、整理・体系化を意図したものである。

18.3.3に述べられた分析結果をもとに、更に用語収集者の専門領域、データベースの収録対象分野の網羅性・利用可能性などを検討した結果、本年度の用語収集対象として、EMBASE（医学・生物学）、CASearch（化学）、ENERGY SCI. & TECH.（エネルギー分野）の三つのデータベースを対象とした。

用語収集の方法として、情報アクセスに有効な用語を収集するという本プロジェクトの主旨を満たすために、上記のデータベースに含まれる書誌データ（標題、抄録、キーワード）から抽出するという方法を採用した。このためには、精度の高い、言い替えれば安全をテ-

マとしている文献を高い比率で得ることが必要である。これを実現するために、各カテゴリーについては、論理和を用い、カテゴリー間では論理積を用いるという検索式を採用した。

上述の探索戦略によりEMBASE, CASearch, ENERGY SCI. & TECH. のデータベースの検索結果は、それぞれ、946, 1,595, 624件であった。いずれも500件を収集対象とし、専門家により500件すべてを読み、分脈も考慮し安全に関連する用語を抽出した。抽出された用語数は、各々のデータベースで、それぞれ約200語であった。表18-2, 表18-3にENERGY SCI. & TECH. 及びEMBASEの各データベースより抽出された安全関連用語の一例を挙げる。

表18-2 ENERGY SCI. & TECH. から抽出された安全関連用語の一例

arsenic exposure	chemical waste
asbestos exposure	chernobyl accident
asbestos fibers	childhood cancer
asbestos induced scarring	chromium exposure

表18-3 EMBASEから抽出された安全関連用語の一例

fatal falls	neurotoxic potential
hazard management	noise injury
hazard model	occupational exposure to organic solvent
hazardous noise	on-duty injury

18.4 今後の展望

本研究は、安全情報に関する複数のシソーラス、複数データベース間での用語の相互アクセスを可能とする「安全」分野の基本用語知識ベースの開発を意図し、下記の四つの基礎的な調査研究を実施した。

- 1) 事前調査としての医薬品の安全用語の評価
- 2) 安全情報のデータベース (NIOSH) からの基本安全用語の収集、分類
- 3) 上記の用語によるDIALOGの85の書誌データベースに対するクロスファイル探索、安全の観

点からのデータベース及び用語の評価

4) 生物・医学, 化学, エネルギーの各分野を網羅的にカバーするEMBASE, CASearch, ENERGY SCI. & TECH. データベースを対象とした安全関連用語の収集

残された課題としては,

1) 上記4)において収集した用語を2)において行ったのと同様に, 分類, 類型化して整理する。

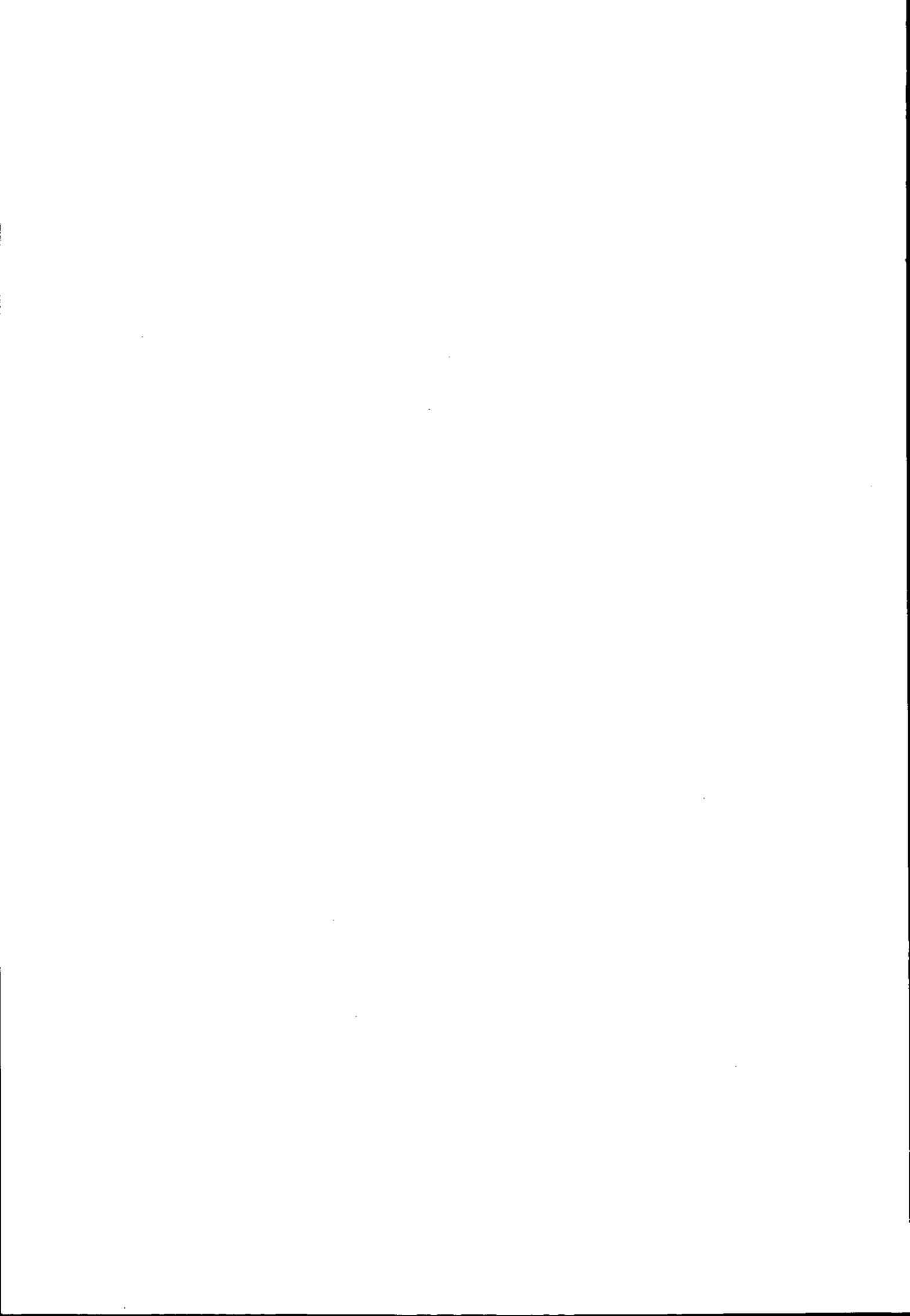
2) 収集整理した4領域の用語について, 比較し, 統合化する。領域間で同一の用語が異なった概念, あるいは同一の概念に異なる用語が用いられている場合には, これを解消する方法を検討する。

3) 2)と平行して, 交通・運輸, 土木・建築, 農業・風水害などの自然災害, スポーツ等の各分野における科学技術に関する安全用語の収集とそれらとの結合を検討する。

4) 1)から3)までを終了した段階で, DIALOGなどの商用オンライン検索システムにおいて, 安全に関する情報を効率的かつ網羅的に検索することを可能とするシステムの構築を図る。などが挙げられる。これらを通じて, 当初の目的である「安全」分野の基本用語知識ベースを完成させるべく努力したいと考えている。

19 3次元マッピングデータベースの技術開発

株式会社 日本総合技術研究所



19. 3次元マッピングデータベースの技術開発

19.1 概要

地図データベースを包括した3次元マッピングデータベースの社会的背景と関連技術をふまえて概念設計を行い、システムのフレームを明確化するプロトタイプを開発した。

19.1.1 システムの目的

地図データベースは行政機関、自治体、都市ユーティリティ事業者など各界で利用されており、高度な機能を保有したシステムが多数存在する。地図データベースは二次元の地図情報のみを管理の対象としてきたが、都市構造の複雑化に伴い従来の面的な扱いでは不十分なものとなってきている。

都市における建造物、道路に限らず都市ユーティリティ（ガス、電気など）もその形状は複雑化の一途をたどっており、3次元形状データベースは地図情報と地表面、建造物などの3次元形状データを統合することによりこうした社会的趨勢に対応する、より実体に接近した地域のデータベースを目指すシステムである。特にニーズの高い都市開発支援や都市管理用に用途を絞ったデータベースの技術開発を行った。

19.1.2 実施内容

本データベース構築に先だって行った社会的ニーズの整理をふまえて、管理対象となるデータの種類、内容、精度、表示法、情報源などの整理を行った。その上でデータの選定、整備方針を定め、システムの基礎的な機能を設計した。それをもとに評価用のプロトタイプシステムの開発を行い、3次元データ（3Dデータ）の表現方法などを検討した。

① データの選定・内容検討

上記の検討を受けて3次元マッピングデータベースの管理対象となる形状データ、属性データをニーズと照合しながら選定した。その際、想定ユーザとしては自治体、公官庁などを想定し、社会情報基盤となるものを選んだ。

② データベース基礎機能設計

ファイル形式の検討をふまえて、データベースの基礎的な機能（表示機能、検索機能、集計機能など）についてのシステム構成を検討した。システムの入出力様式についても概要を整理した。

③ プロトタイプ開発及び評価

特定の地区についての3Dデータを実際に入力し、既成の3Dアプリケーションとデータベースの組み合わせによってプロトタイプを構成する。

プロトタイプのシステム構成としては現在プロトタイピングに最適の環境であることを基準にして以下のものを採用した。

ハードウェアはMacintoshを想定している。システムの構成はアプリケーションを統合するアプリケーションと都市の3Dデータを入出力するモデラー、コンピュータグラフィックス(CG)として画像を作成するアプリケーション、そして各種シミュレーションを担当するアプリケーションからなる。

19.2 検討結果

3次元マッピングデータベースへのニーズ検討によって利用局面とシステムの特徴を定め、それをもとに機能設計を実施した。

プロトタイプはDTM (Digital Terrain Model)、コンピュータマッピングデータベースと3次元情報の結合、地表及び建造物の3次元表示を含んだものを構成した。

19.2.1 システムコンセプトの検討

都市開発及び都市管理業務においてはその処理すべき地域が広範囲であり、関係者も多数に及ぶ他、扱う情報量も膨大なものがある。このような事情により幾つかの自治体はコンピュータを用いた地理的な情報管理システムを既に運用している。しかしながら都市空間の高度利用がここ数年の間に著しく進むに従って、従来の面的な地域情報データの管理では開発業務、都市インフラの設備管理業務に支障をきたす可能性がある。

3次元マッピングデータベースシステム（以下システムと略す）は立体的な都市空間の可視化を自治体全域にわたって繰り広げることによってこうした点を補い、同時に複雑化した都市インフラをより効率的に管理することを目標とするものである。

19.2.2 利用局面の確認

3次元マッピングデータベースシステムが必要とされる利用局面は

- ① 都市開発業務
- ② 都市計画業務
- ③ 都市管理業務

に大別されることを確認し、その現状と3次元情報へのニーズを整理した。

19.2.3 システム特性の設定

(1) 既製システムとの機能分化

コンピュータマッピングは都市計画、都市管理などに深く浸透し、それぞれの業務に応じた機能、性能、形態をとっている。新システムはこれと役割分担を行う。また、コンピュータマッピングの機能補完も考慮すべきであろう。

(2) ハードウェア、ソフトウェア（OS、データ等）の資源活用

新システムではハード、周辺機器の共有を始め設備の共有を重視する。また既製システムのデータ蓄積も無視できない。このハード、データ資源を有効に利用できるシステムが望ましい。

(3) 運用方式とメンテナンスの容易性

3次元マッピングデータベースの表示対象となるデータは極めて多種多様であり、データ更新作業が膨大になる。本システムが恒久的に利用されるためにも可能な限り自動的かつ最小限にデータ更新できる仕組みが求められる。

19.2.4 3次元データの所在

3次元情報を表示するに当たって最も利用しやすいデータは国土地理院の国土数値情報である。国土地理院では標高に関する情報として等高線データと格子型のメッシュデータを用意している。

① 等高線データ

等高線のデジタルデータは縮尺1：25,000の地形図用のものを用意しており、1：25,000の地形図の作図用の用途が多い。等高線データは線ベクターの集合という形態をとっている。

② 格子型メッシュ標高データ

格子型メッシュ標高データは、縮尺1：25,000の地形図をスキャニングして、200×200のメッシュを切り等高線をもとに計算して求められている。これらの作業は国土地理院の開発したアプリケーションで自動化されている。

19.2.5 データソースの考え方

ここでは既存のコンピュータマッピング保有データのうちで本システムに取り込むべき情報を選定する。システムの特性にリストアップしたように本システムはコンピュータマッピングのデータベースとのデータ交換に重点を置く。その基本はデータソースとしての取り込みである。きめ細かな都市計画もしくは都市管理情報を設定することでシステムのセットア

ップを容易にすることができる。これはデータ変換さえ構築できれば自動的に3次元表示がある程度まで保証されることでもある。ここで都市計画関連情報を例にとればそうした情報として次の3種がある。

- (1) 規制界種データとは土地の用途規制の境界線を示したものである。このデータは原則的に曲線情報である。この曲線データを地表面に投影する手続きはサーフェスマッピングと呼ばれる手法が適しているが、表現精度の限界もある。
- (2) 土地利用現況図区分は地表面への写像（マッピング）を行うタイプのものである。このデータのうち住宅、施設データは建物現況図のうちで取り扱うため景観シミュレーション中では表示する必要はないものと考えられる。したがって、一部の項目についてのみ地表面への写像を行う。他の区分情報については適宜サーフェスマッピングを行う。
- (3) 3次元建造物での表示（テクスチャマッピング）は次のように設定する。建造物のディテールを表現する場合は建物種別ごとにテクスチャを割り当てる。このテクスチャは予め妥当なデータをライブラリに登録しておき3次元表示させる際に結合する。

19.3 システム主要機能の設計検討

19.3.1 システム構成

(1) 機能構成の考え方

本システムは最新のハード、ソフトのコンピュータテクノロジーを最大限に取り入れたものであるべきである。とりわけ次のような特色をもたせることによって差別化を図ることが望ましい。

- ・汎用性：3次元表示に重点を置くことはいうまでもないがその適用範囲が都市情報の広汎な局面をカバーできるものでなければならない。特に都市インフラの要素を表現できることが鍵となる。
- ・将来性：ハードウェアの進歩は日進月歩であり、アーキテクチャが旧弊なものとなりがちである。ハードウェアは時代の主流となるものであることが好ましい。
- ・協調性：既成システム、特にコンピュータマッピングとの情報統合を図ることが重要である。
- ・容易性：可能な限りマニュアル不要、メンテナンスフリーであること。操作の容易性も当然考慮されるべきである。

(2) 基本データベースについて

都市開発及び都市管理の業務支援に関する基本的な3次元情報として必要なデータに

ついてリストアップするとともに内容についても整理する。

① DTMデータベース

地形情報の3次元モデルをデータベースとして管理するシステム。管理対象としては自治体全域をもつことが望ましい。地形（標高）情報は原則的に暦年変化がわずかであり更新の手間はそれほどかからないと考えられる。このデータが3次元表示の基礎情報であり、ベースマップの投影などにも有効である。

② 境界情報データベース

市町村区界、都道府県境などの境界線情報を地表面に投影したもの。したがって、3次元曲線データとなる。

③ 都市情報データベース

都市情報システムのベースマップを3次元へ投影するためのものである。もしくは3次元オブジェクトとするためのベースデータとして用いることもできる。

19.3.2 主要機能の基本設計

(1) データ変換機能

3次元形状データ（3D都市モデル）を扱う環境を本システムは標準的に保持することとする。これにより従来のシステムと共存できよう。特に従来のコンピュータマッピングとの連動を目指して、外部データの取り込み（Import）、吐き出し（Export）が本システムの機能として具備されることが不可欠である。

(2) 各種表示機能の設定

プロトタイプシステムで構成した基幹機能のアウトラインをまとめる。これらの機能は3次元情報をフルに活かせるものを基準にして構成された。

① 2D-3D描画機能

・モデリング機能

3D形状データの作成を行う。主として基本モデルの修正と開発該当区域の全体フレームの中に取り込むなど付加的に用いる場合が多い。このため機能深度はそれほど求められない。

・レンダリング機能

いわゆる景観シミュレーション機能がこれに対応する。視野、視点を自由に指定し景観を表示する。

・アニメーション機能

連続的な画像を予め設定したシナリオに沿って自動生成する。動線に沿った視点移動は

この機能の一部であるとも考えることができる。

② シミュレーション機能

・形態制限表示機能

形態制限の綿密な検討は開発者に委ねるのが原則であるがラフな評価が別モジュールで対応可能である。

(i) 容積率評価

(ii) 建坪率評価

(iii) 斜線制限表示

(iv) 日影規制表示

・連続視点移動機能

ウォークスルーは建築CAD、CG系では常識的な機能となりつつある。このため導入のポイントは機能の有無ではなく、視点移動機能の画像の品質とレスポンス、使いやすさとなる。

19.4 システム構築の考え方

プロトタイプシステムは機能構成を簡素なものとするためにモジュール化し、各モジュールのデータをDBMSでハンドリングするために最終的に一括する。また、プロトタイプ構築に際しての3次元情報の作成は図19-1のような簡易的な手続きによった。

19.5 今後の課題

既製システム、特にマッピングデータベースとのデータ交換を形状のみにとどまらず属性情報をも包括させることが望ましい。これは都市管理に本システムを適用するために必要性が高い。そのためにはSQL、informixといった汎用データベースの知識が求められるだけでなく、グラフィックデータの交換フォーマットについてのノウハウも必要となる。広範囲なソフトウェア技術者の動員が望まれる。

いうまでもなく景観画面の作成はCG技術をフルに必要とする。CGソフトを構成する三つのモジュール、モデラー、レンダラー、アニメーターについても基礎からの開発に際しては各々専門技術者が必要となる。ネットワーク構築も技術的蓄積が必要であり、セキュリティ機能設計、異機種接続などが要求され、これらを統合するシステムインテグレーションが要求される。

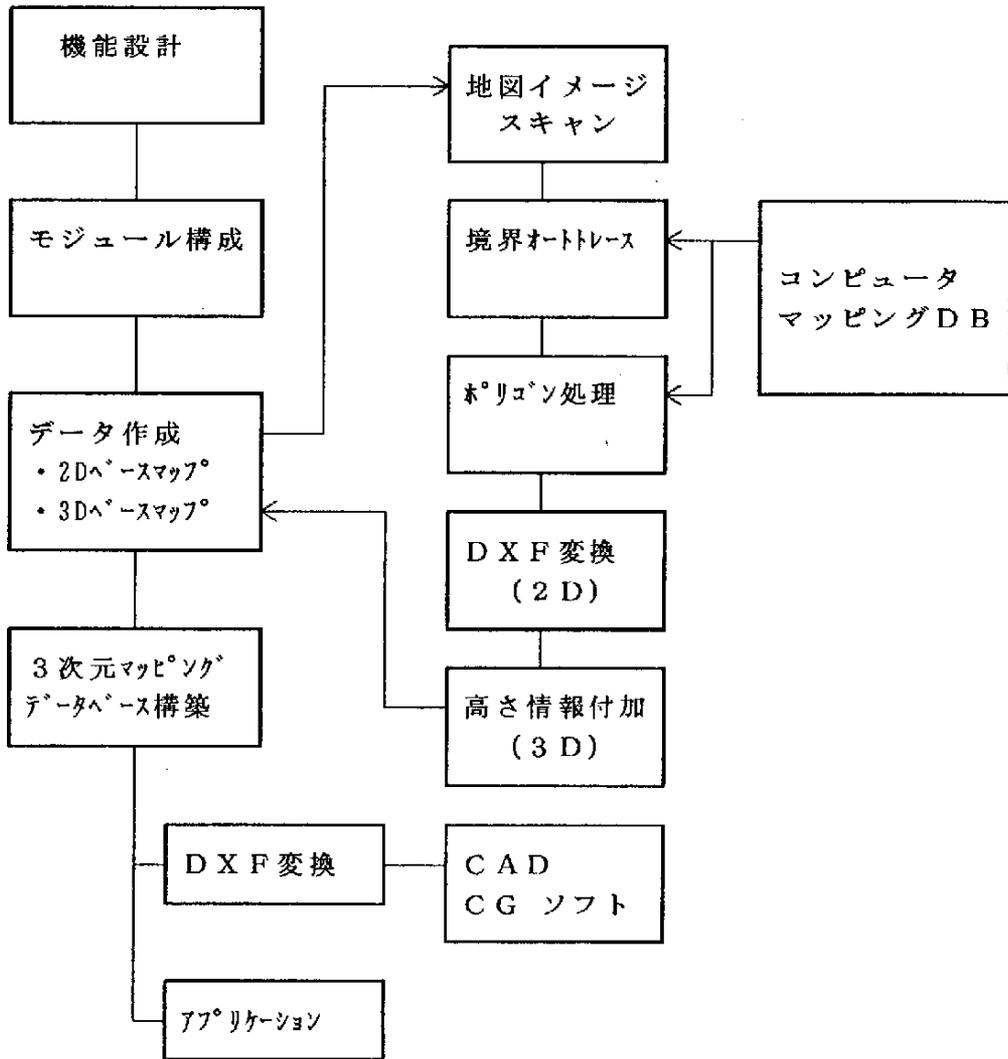
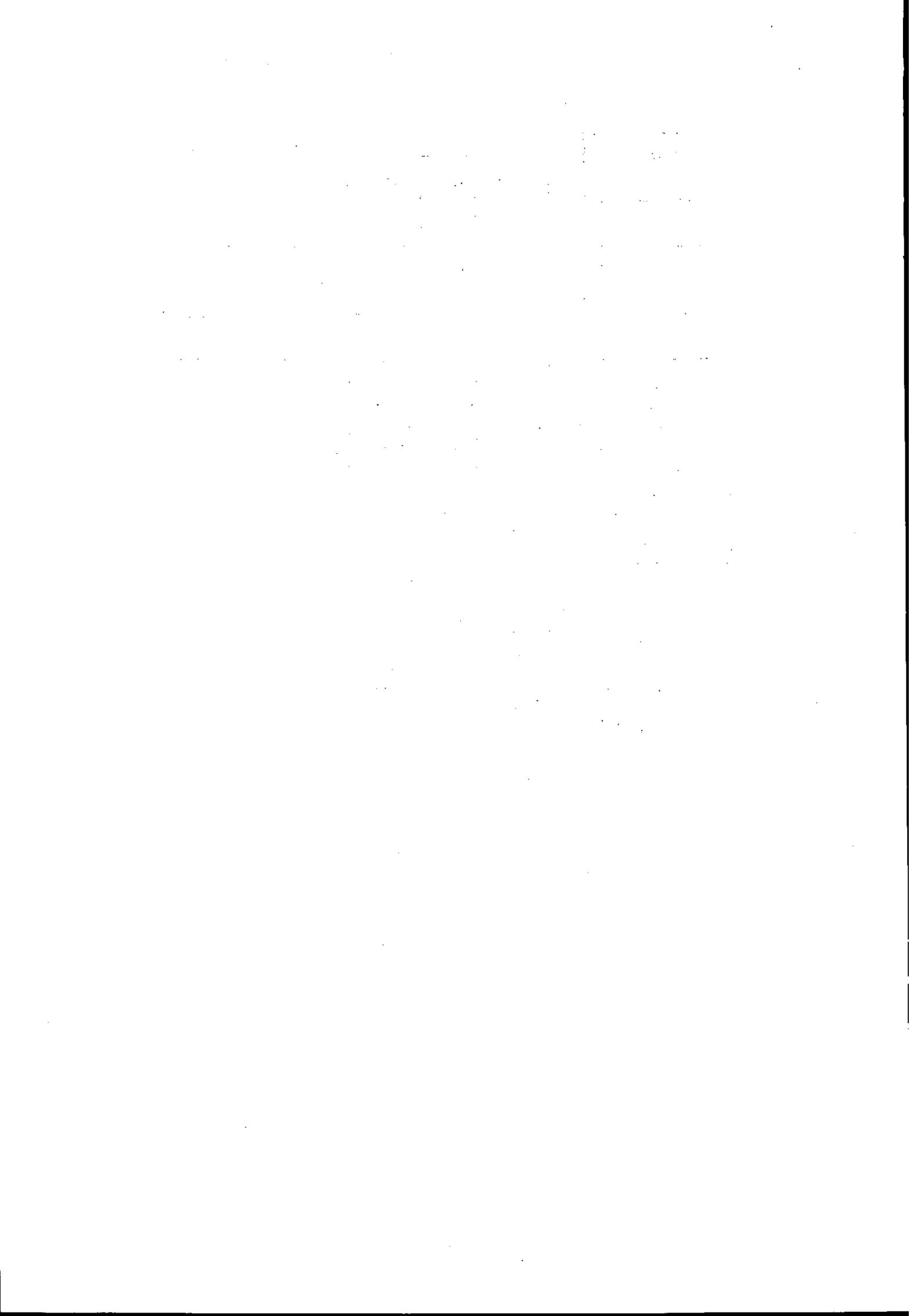
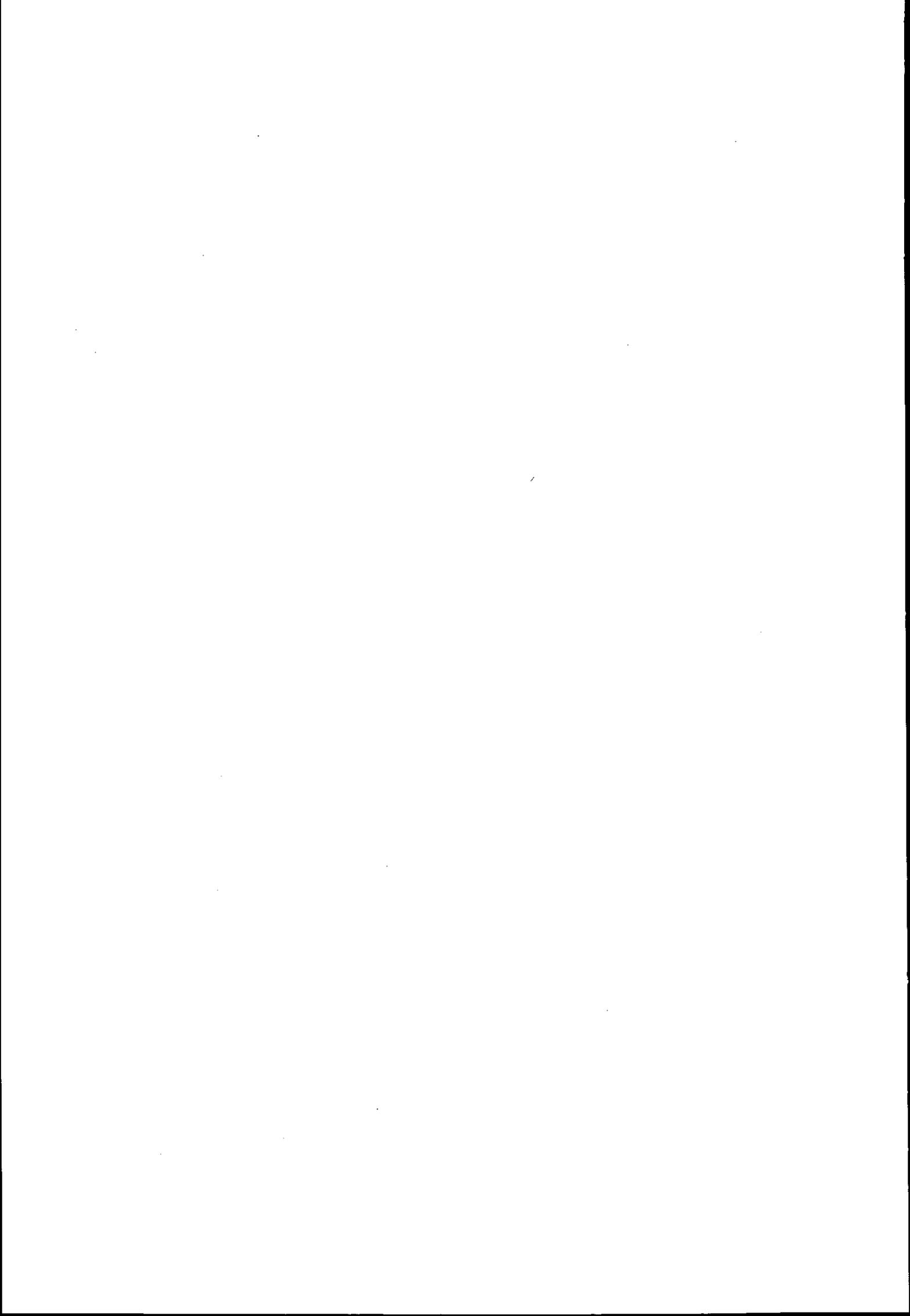


図19-1 システム構築手順



20 データベース検索サポートシステムの調査研究

セントラル開発株式会社 情報図書館RUKIT



20. データベース検索サポートシステムの調査研究

20.1 概 要

昨今データベースの数は順調に増加し、また通信ネットワーク、端末機等インフラストラクチャーが整備され始めている。一方ユーザ側の視点に立つと、数が増えたことによる弊害として必要な情報の所在が分かりにくくなってきている。更に、現在稼働しているほとんどのシステムは、分野別ファイルで構成されており、必ずしもニーズに即したメニュー構造にはなっていないのが現状である。したがって、検索を行う際には、どのデータベースあるいはデータファイルにアクセスするか利用者が予め意志決定する必要があり、熟練した利用者には差し支え無いものの、初心者や利用頻度が少ない利用者、あるいは今後裾野が広がってくるとされるデータベース利用者のことを考慮すると、これらのことが普及啓発の妨げの一因になっているといっても過言ではない。そこで、検索の目的や意向を論理的に解析してデータベースファイルに導くサポートシステムを考案することがデータベース利用の動機づけに寄与し、強いてはデータベース業界の普及啓蒙につながることを期待される。

本研究の目的は、情報図書館RUKITの活動をふまえてデータベース利用者の検索目的や主題、利用データベースなどの傾向を調査し、利用者がデータベースシステムあるいは、ファイルを意識することなく、検索目的から意向に沿ったファイルを導くシステムを研究開発するものである。

20.2 実施内容

20.2.1 方針と内容

本事業を四つの視点に分け以下の手順により調査・開発研究を行った。

- ① データベースの利用動向の調査
- ② 収集した情報の分析
- ③ サポートシステムの開発
- ④ サポートシステムの検証

20.2.2 データベース利用者の利用目的・動向調査

通常データベースはさまざまな目的で利用されているが、エンドユーザの必要とするデータ検索は、システムやファイルの性質上単純な検索パターンで目的が達成される場合と複雑

な検索手法又は予備知識が必要となるケースに2分される。

いずれの場合も、ある程度データベースあるいは端末機となるコンピュータの利用経験をもつユーザを想定した構造になっている。このため、頻繁に使用するデータベース又はファイル以外にアクセスする必要がある場合、経験者であってもマニュアルやリファレンスツールを参照するなど事前準備が必要である。また、多くのデータファイルはその名称が提供者の固有名詞や属性を表す名称であり、必ずしもファイル内容がその名称を忠実に表現しているとは限らない。

そこで、実際の検索に立会い、利用者がどのような目的で情報図書館RUKITに来館し、検索主題が何であり、それらを満たすためにどのデータベースシステム及びファイルを選択したか、その一連の検索プロセスについてヒアリング調査を実施した。

20.2.3 調査結果

平成3年8月～11月に情報図書館RUKITを利用した延べ3,642件の検索のうち、当館職員が検索に立会いヒアリングが可能であったケース2,581件について結果を集計分析した。

○ 全般的な傾向

表20-1に来館者の利用動向調査結果の分析マトリックスを示す。縦軸に利用目的をとり、横軸は検索主題とした。また今回特に重要となるファイル選択の傾向を分析し、利用目的と相関させた。頻度が高かった上位3位までファイル名及び回数を付記した。このマトリックスのフォームにデータを整理し更に、利用目的別及び検索主題別件数の昇順にソートした。

この分析表から、どのような目的でデータ検索がなされ、その際に選択されたファイルの利用傾向、検索主題の傾向が把握できる。

検索主題 利用目的	財務 企業分析 金融 為替	社会 問題 犯罪 風俗 教育 歴史	食品 医薬 医学 バイオ 農業	個人 情報 連絡 先	化学 全般 化学 工業 化学 製品	情報 通信 コンピ ュータ OA	芸術 娯楽 レジ マー スポ ーツ	出版 マス コミ 放送	住宅 建設 土木 不動産	エネ ル ギー 環境	人事 経営 M&A	エレク トロニ クス 家電	政治 外交 防衛 法律	ファッ ション 繊維 紙 日用品 雑貨	自動車 航空機 船舶 宇宙 輸送	金属 工作 機械 FA	計	利用頻度1位 のファイル	利用頻度2位 のファイル	利用頻度3位 のファイル
技術情報の調査	15	20	94	4	140	49	3	4	39	28	1	36	4	5	15	14	471	JICST109)	新聞一括(52)	日経四紙(41)
執筆-番組作成のヒント	17	89	53	76	0	6	46	35	3	9	11	1	10	5	2	2	365	新聞一括(83)	WHO(60)	MAGAZINE(33)
企業の信用調査	272	10	5	7	2	9	3	2	1	0	24	2	1	2	1	2	343	COSMOS2(101)	TSR・BIGS(49)	日経四紙(36)
会議資料作成	35	29	11	8	4	19	16	15	23	16	11	1	16	11	4	2	221	新聞一括(40)	日経四紙(25)	MAGAZINE(16)
人物人材調査	12	27	18	83	2	1	23	19	1	5	3	2	6	4	8	0	214	WHO(65)	新聞一括(30)	日経四紙(16)
食生活学習・生活・趣味	11	33	12	30	7	38	18	16	3	5	6	1	5	0	3	2	190	BOOK(22)	WHO(15)	日経四紙(14)
製品開発/動向調査	7	27	44	0	15	20	4	3	4	5	3	7	0	15	4	10	168	日経四紙(14)	JICST12)	新聞一括(7)
競合会社の動向調査	68	6	15	0	8	5	8	2	6	6	10	4	1	4	5	3	151	日経四紙(28)	新聞一括(22)	TSR・BIGS(18)
業界/市場動向調査	52	4	9	1	11	15	6	2	5	12	10	6	0	2	9	2	146	日経四紙(22)	MDB(19)	MAGAZINE(12)
顧客開拓	54	5	10	5	6	2	6	1	16	2	5	6	5	4	2	2	131	COSMOS2(24)	日経四紙(15)	新聞一括(8)
政治・文化活動	8	36	3	9	5	1	13	1	3	6	1	0	4	1	0	0	91	新聞一括(20)	日経四紙(16)	朝日新聞(7)
特許資料・裁判資料	2	3	4	0	16	5	1	2	0	3	2	9	5	2	1	1	56	特許(18)	WPI(8)	JAPIO(6)
就職・転職・求人	13	4	0	6	0	1	1	2	0	0	6	0	0	0	1	0	34	日経四紙(7)	WHO(5)	COSMOS2(4)
計	566	293	278	229	216	171	148	104	104	97	93	75	57	55	55	40	2581	新聞一括(533)	日経四紙(445)	COSMOS2(293)

表20-1 利用動向調査結果分析表

20.2.4 検索サポートシステムの検討

○ 仕様決定までのプロセス

本サポートシステムの利用対象者は、最終的に情報を必要とするエンドユーザにおいたことに特色がある。つまり、検索担当者やサーチャー業務に従事している利用者などサポートを特に必要としないユーザには、通常の検索が可能な通信環境を提供しつつ、それ以外のユーザが必要に迫られた際に、検索例や事例を参照しながら簡単にファイル選択が可能な環境も混在させることである。これらの条件を満たすために以下の点に留意しながら要求仕様を考慮した。

- ① どのレベルの利用者を想定するか。
- ② 全要求の何割をサポートシステムで解決できるか。
- ③ クロスファイル検索をどうするか。
- ④ 対象とするデータベース又はファイルをどの程度考慮するか。
- ⑤ ファイルがゲートウェイ接続と直接接続のどちらもサポートしている場合、どちらを選択するか。
- ⑥ 同じ属性をもつファイルは優先順位をもたせるか、階層を深くしてはどうか。
- ⑦ 特許、化学等専門的な知識が必要なファイルをどうするか。
- ⑧ 社内LAN等、ネットワーク上での複数利用も考慮する。
- ⑨ ログインした後、どのレベルまでサポートするか。
- ⑩ システム上でマニュアル的な仕組みが考えられるか。

委員会における討議及び技術的問題などを考慮し、⑧は仕様を含めるが、⑦⑩は今後の課題とした。

これらはいずれも委員会の過程で議論された点である。これらの点をふまえ、次節以降で述べる仕様をプロトタイプとして策定した。検索サポートシステム概念図を図20-1に示す。

20.2.5 システム構成

○ システムの概念とプロトタイプ作成の目的

プロトタイプ作成に当たり、

- ・データベース利用のみならずパソコンの初心者でも容易に操作ができること
- ・情報を取り出すまでの利用料金（回線料金、データベース利用料金）が極力安くなること
- ・データベース利用方法に関するアシスタント業務が軽減されること、もしくはアシスタント要員が少なくてすむこと

といった、基本的な条件の他にプロトタイプ固有の条件も考慮する必要がある。即ち

- ・ある程度の仕様変更が容易にできること（特にメニュー構成）
- ・技術的な問題やコスト的な問題点が把握できること
- ・実用的にもある程度耐えられること

等である。言い替えれば、「プロトタイプ」開発に当たっては、本格的な実用に耐える必要はないと考えられるが、一方では、システム検討中に起こる仕様変更に対応できるばかりでなく、操作性や機能において実用化をイメージできるものでなければならない。更には、ハードウェアにおいても機種が限定される仕様であってはならない。

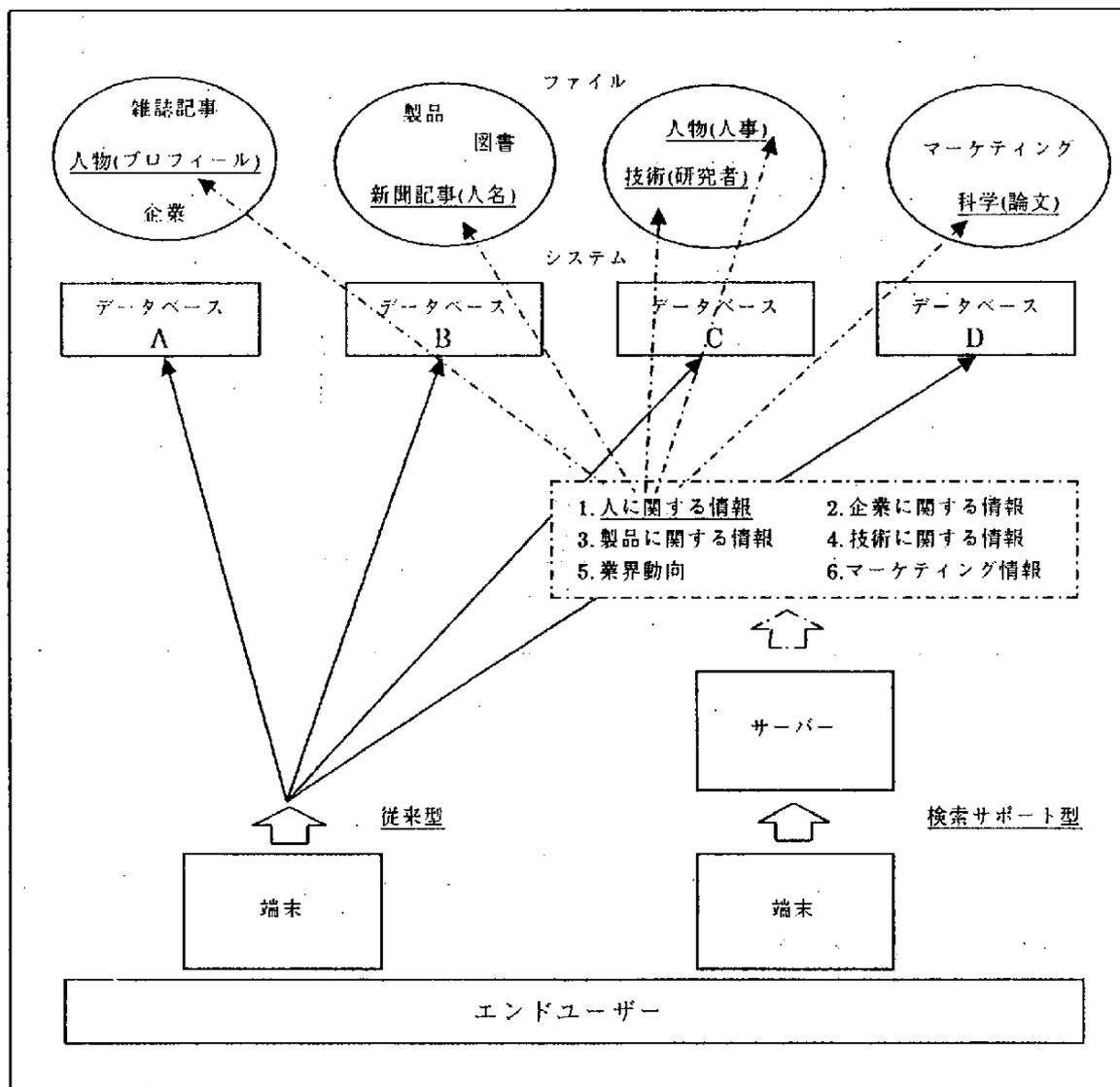


図20-1 検索サポートシステム概念図

そこでこれらのことをふまえて、以下のことに特に留意してプロトタイプを作成した。

- ・ハードウェアやOSは汎用的なものを念頭におき設計する。
- ・画面仕様及び操作性は実用化に耐えるものにする。
- ・画面の仕様変更は容易にできるようにする。
- ・通信中や誤操作によるエラー処理及び明らかに実現可能なもので本検討で予算的に困難な機能は削除する。

システムの概略は図20-2のとおりである。

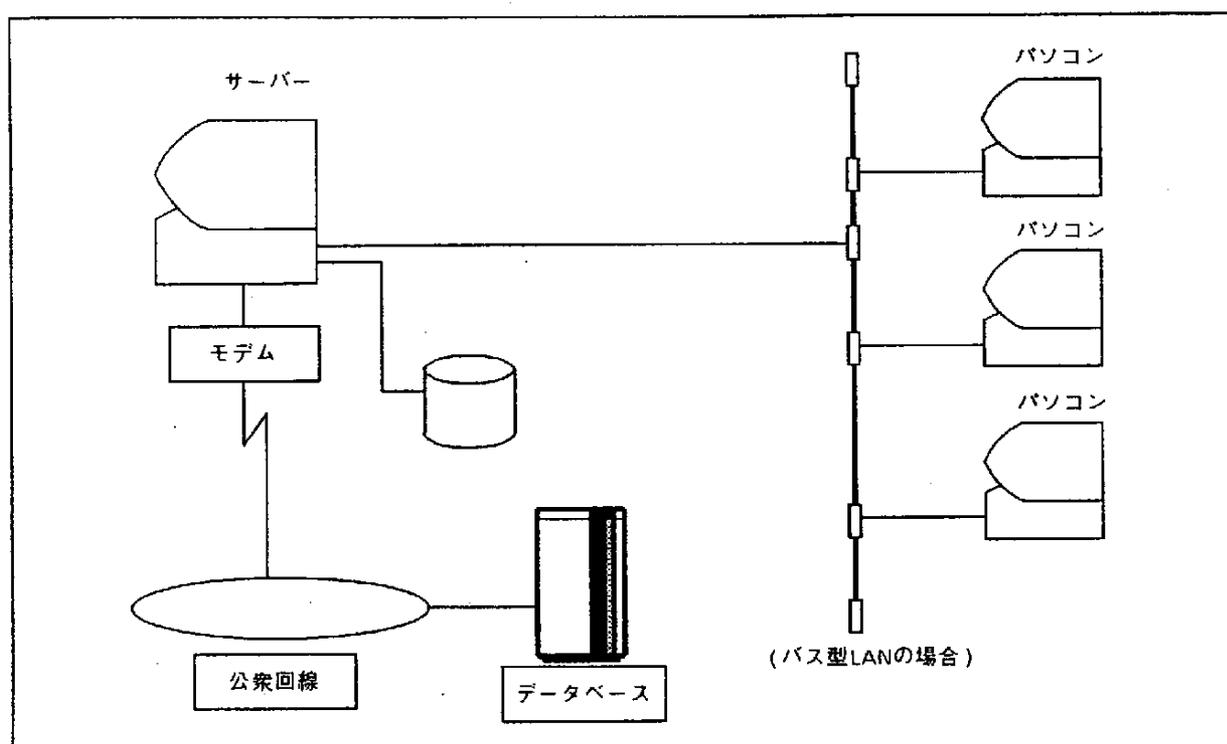
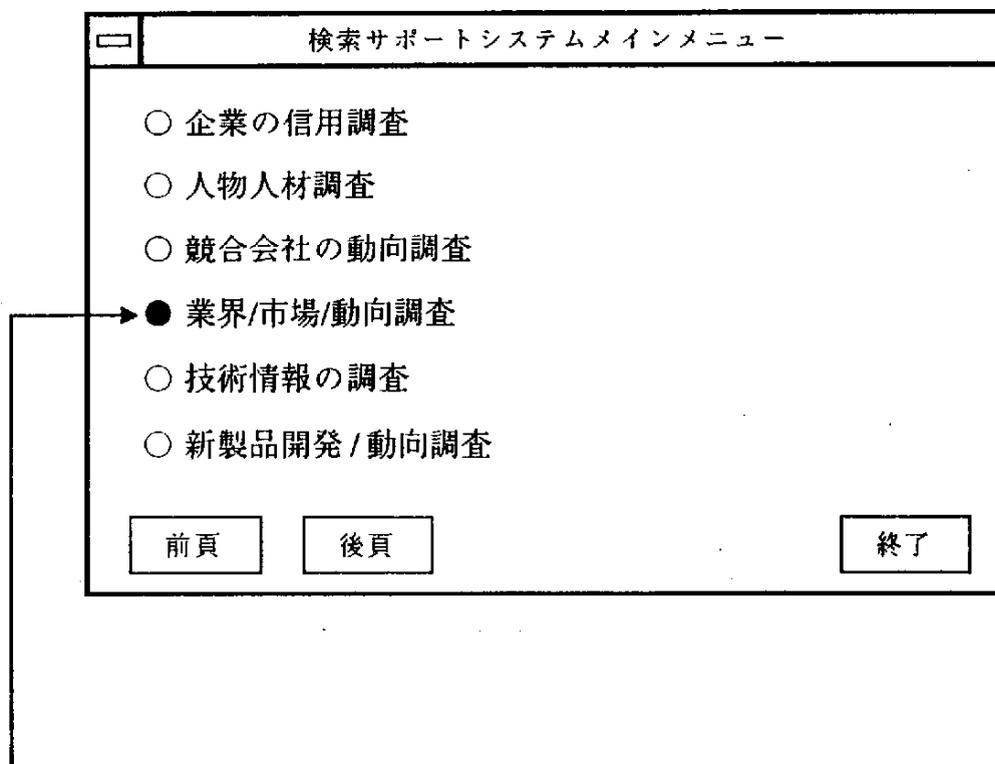


図20-2 システム概略図

20.2.6 プロトタイプ

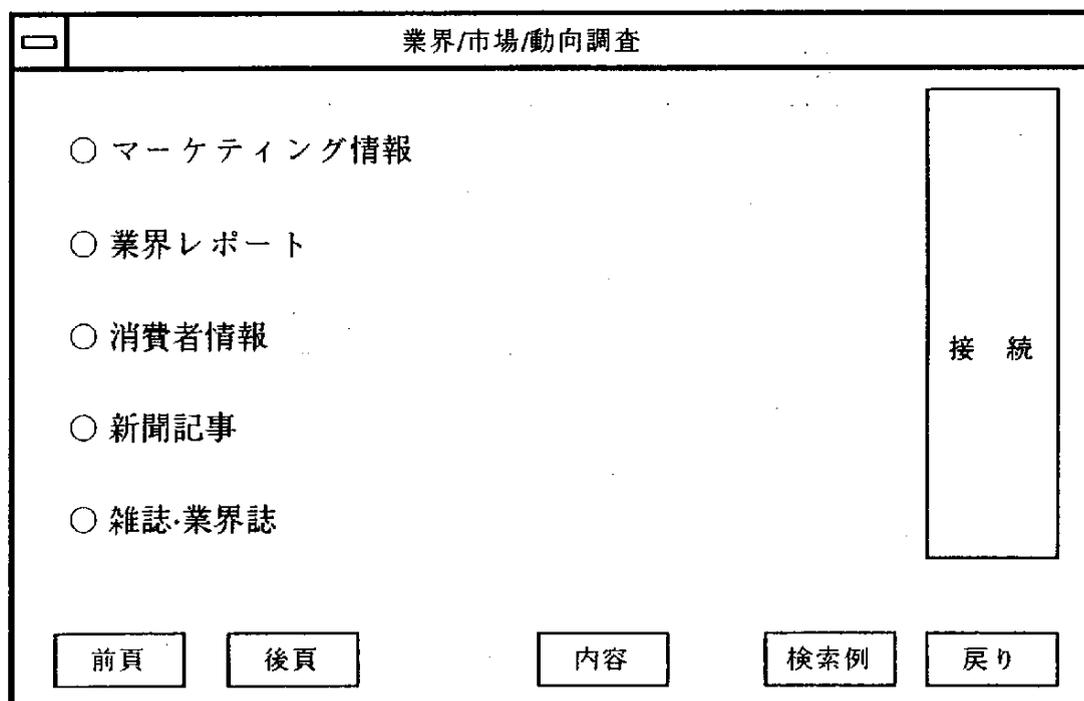
本サポートシステムの操作フローと、構築した目的別項目及び補助機能の内容を実際の端末画面のイメージで例示する。

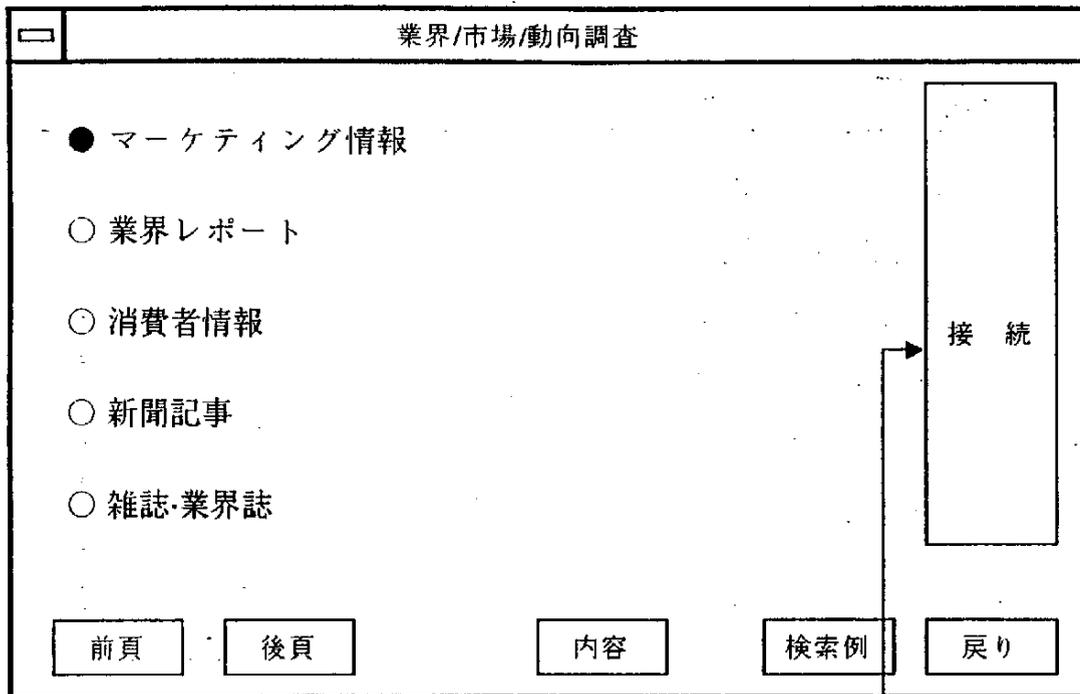
○ 操作フロー



① 各項目をマウスで選択する。選択されると表示が○→●に変わる。

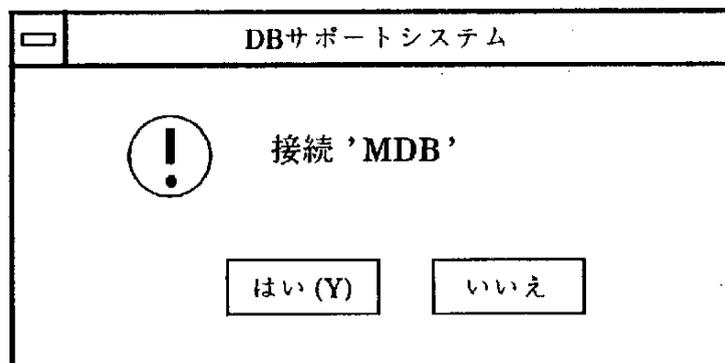
② 選択後、 ボタンをクリックすると各項目のサブメニューウィンドウがオープンする。





③ サブメニューに表示された各項目をマウスで選択する、**接続** をクリックすると対応するファイルに接続するための確認ウィンドウがオープンする。

接続する場合は **はい (Y)** キャンセルする場合は **いいえ (N)** をクリックする。



④ サブメニューに表示された各項目ごとの「内容」や「検索例」を参照したい場合は、項目を選択後、それぞれのボタンを選択することにより、サンプルウィンドウがオープンする。

以下に各サブメニュー及び、対応するサンプルウィンドウを例示する。

○目的別項目画面

企業の信用調査				
<ul style="list-style-type: none">○ 企業概要○ 企業財務情報○ 海外系企業・海外進出企業○ 海外企業情報○ 医薬品企業○ 経営者情報	接 続			
<input type="button" value="前頁"/>	<input type="button" value="後頁"/>	<input type="button" value="内容"/>	<input type="button" value="検索例"/>	<input type="button" value="戻り"/>

人物人材調査				
<ul style="list-style-type: none">○ 取材/インタビュー新聞記事○ 取材/インタビュー雑誌記事○ 人事情報○ 著名人プロフィール○ 外国人プロフィール○ 執筆文献/論文○ 執筆書籍○ 経営者情報	接 続			
<input type="button" value="前頁"/>	<input type="button" value="後頁"/>	<input type="button" value="内容"/>	<input type="button" value="検索例"/>	<input type="button" value="戻り"/>

競合会社の動向調査				
<input type="radio"/> 新聞記事	接 続			
<input type="radio"/> 雑誌・業界誌				
<input type="radio"/> 技術文献				
<input type="radio"/> 新製品情報				
<input type="radio"/> 会社業績				
前頁	後頁	内容	検索例	戻り

業界/市場/動向調査				
<input type="radio"/> マーケティング情報	接 続			
<input type="radio"/> 業界レポート				
<input type="radio"/> 消費者情報				
<input type="radio"/> 新聞記事				
<input type="radio"/> 雑誌・業界誌				
前頁	後頁	内容	検索例	戻り

技術情報の調査

- 技術情報(タイトル、出典)
- 科学技術情報
- 新聞記事
- 雑誌・業界誌

接 続

前頁 後頁 内容 検索例 戻り

新製品開発 / 動向調査

- 新聞記事
- プレスリリース
- 新製品情報

接 続

前頁 後頁 内容 検索例 戻り

○補助機能画面（企業の信用調査）

内 容 : 企業概要						
* * * 会 社 概 要 1 * * *						
XX/XX/XX XX:XX	企業番号 (984032786)					
【商号】 富士通エフ・アイ・ピー 株式会社						
【住所】 〒105 東京都港区新橋6-1-11 秀和御成門ビル	【電話番号】 03-3433-2251					
【業種】 (主) 8522 情報処理サービス業 (従)						
【創業】 年 月	【評点】 XX 点					
【設立】 昭和 52 年 11 月	【従業員】 2,068 人					
【資本金】 48,000 万円	【株式】 非上場					
【株主数】 人	【法人・個人】 法人					
【売上高順位】 全国 1,449 社 X,XXX 位	【調査レポート有無】 有					
県 584 社 XXX 位	【調査時点】 19XX 年					
【業績】						
決算期	売上高(百万)	利益金(万)	配当%	資本%	B/S	申告所得(万)
88.03	X,XXX,XXX	XXX,XXX	XX	XX	無	88.03 X,XXX,XXX
89.03	X,XXX,XXX	XXX,XXX	XX	XX	無	89.03 X,XXX,XXX
【取引銀行】						
第一勧銀(日比谷)	協和(新橋)	三菱(新橋)				
横浜(丸ノ内)	大和(丸ノ内)	太陽神戸(丸ノ内)				
接 続	検 索 例	複 写	閉 じ る			

検 索 例 : 企業概要						
<p>企業情報(特定検索)サービスに接続されました。</p> <p><<情報提供元:帝国データバンク>></p> <p>>選択番号と検索データを例のように入力して下さい。</p> <p>(例1:2,03-3123-1111 例2:3,7-財*@(0:前方一致検索))</p> <p>1 : 企業番号検索 2 : 代表電話番号検索 3 : 企業名検索 (, 地域指定)</p> <p>E : 前の操作に戻る</p> <p>-> 3,7777777777-</p> <p>1 件 見つかりました。</p> <p>>次のいずれかを選択して下さい。</p> <p>D : 企業一覧を表示する S : 再度検索を行う</p> <p>E : 前の操作に戻る T : DB選択に戻る</p> <p>-> D</p> <p style="text-align: center;">< 帝国データバンク 企業一覧情報 > XX/XX/XX XX:XX</p> <p>選択番号 <商 号> <住 所> <業 種></p> <p>1 富士通エフ・アイ・ピー 株式会社 東京都港区新橋6-1-11 情報処理サービ</p> <p>>選択番号を入力して下さい。または次のいずれかを選択して下さい。</p> <p>A : 次の10件を表示する B : 前の10件を表示する</p> <p>N : 次の100件を表示する S : 再度検索を行う E : 前の操作に戻る</p> <p>-> 1</p>						
接 続	内 容	複 写	閉 じ る			

20.3 今後の課題

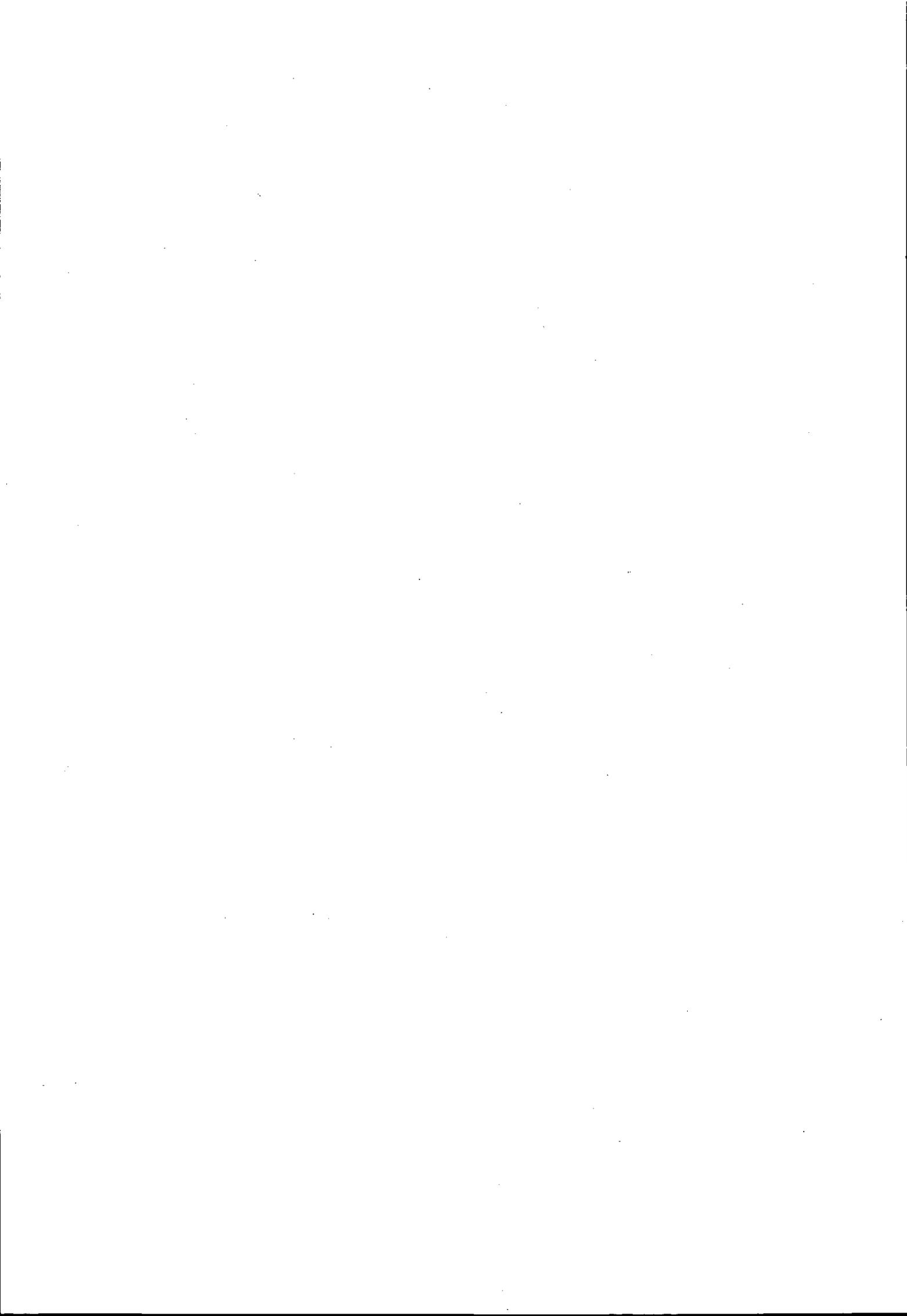
今回の調査研究の最も重要な成果は、検索サポートシステムの思想や、仕組みが情報図書館 RUKIT の活動を通じて多くの利用者からヒントを得、さまざまなニーズがサポートシステムのコンセプトに反映できたことである。

一方、本システム構築では、初心者や経験の浅い利用者を対象にしたため、パターン化が可能な単純検索には向くが、調査研究の過程では、課題も多く残された。これらを解決するには、更に数多くの調査分析と技術開発が不可欠である。

また、今回の研究成果は利用者のニーズを最優先に開発されたものであることから、データベース提供会社やプロデューサとも連携し、よりユーザフレンドリーなシステム開発を期待したいものである。本検索サポートシステムを更に発展させ、誰にでも簡単にデータベースにアクセスできる環境を整えることが、商用データベースの普及・利用促進にもつながると思われる。

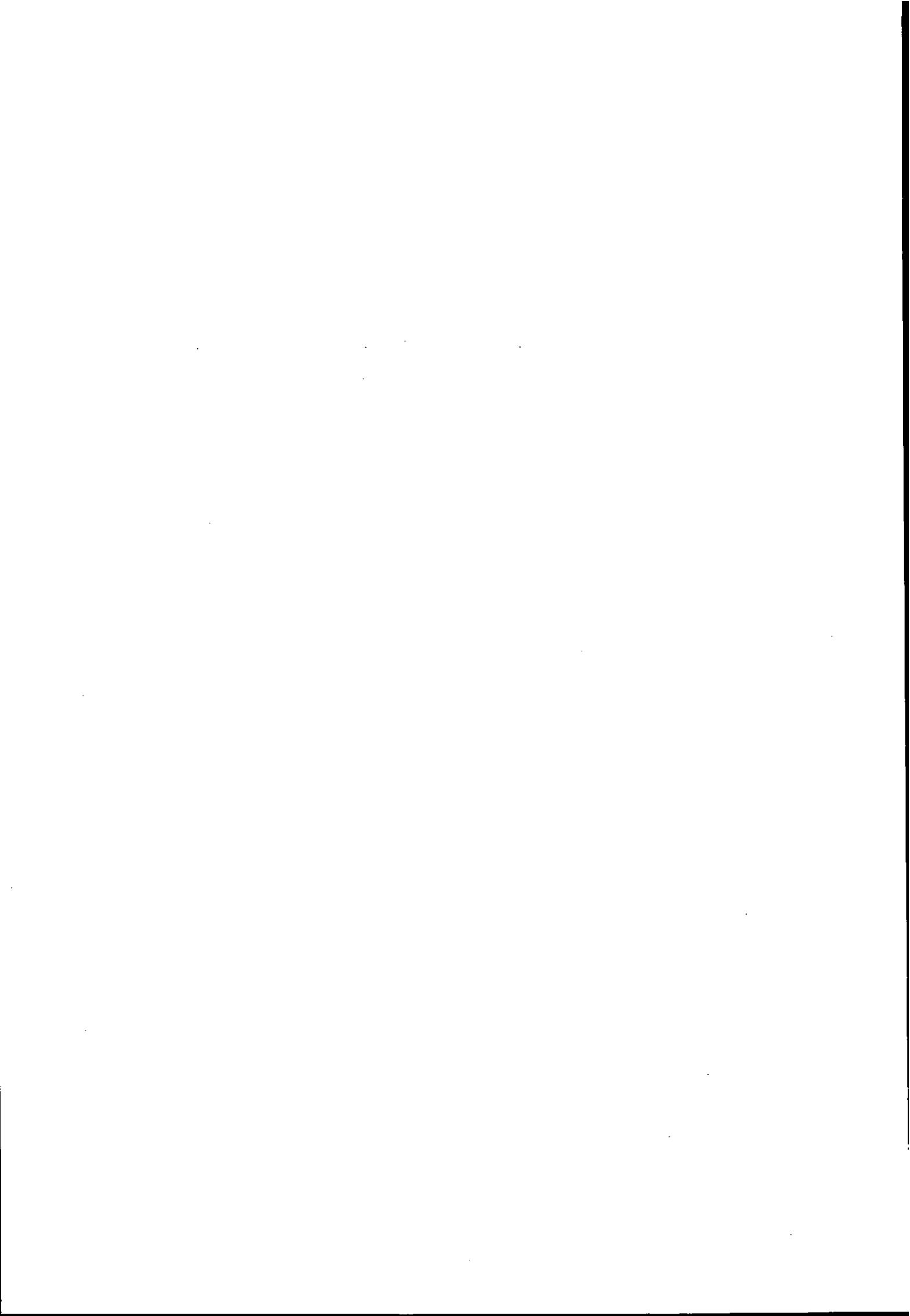
そこで、特許や化学情報のように専門家対象の高度な支援機能をもつシステムも含めより有用な検索サポートシステムについて、更に研究していくことが重要である。

更に、商用データベースのサービスを超越付加価値利用を創りだすことの端緒が得られたので、今後この方向への積極的展開を図る必要がある。



21 グループウェアにおけるデータベース
システムに関する調査研究

株式会社 イフ・アドバタイジング



21. グループウェアにおけるデータベースシステムに関する調査研究

21.1 概 要

これまでの情報システムは、個人の作業の支援が中心であったが、今後は、グループ作業の支援、いわゆる“グループウェア”が情報システムの主要な目的となる。しかし、グループウェアの遂行にはさまざまな技術的課題の解決が必要であり、中でも各個人、グループで利用されるデータベースの実現、管理についての検討が極めて重要となる。つまり、各個人のデータベースシステムとグループデータベースシステムから構成される分散型データベースシステムについての詳細な調査・研究が必要となる。このため、グループウェアにおけるデータベースシステムに関し、特にユーザにおける情報システムの高度化を視点として、その技術と展望、ユーザにおける活用、課題等さまざまな側面から調査研究を行った。

本報告書では、

I 課題性の研究

について述べた後、

本論については；

II グループウェアにおけるデータベースシステムの構築

1章： グループウェアの技術的動向

2章： グループウェアにおけるデータベースシステム構築技術

III グループウェア発展に伴う情報システムの高度化並びにデータベース構築技術の進展とその課題

から構成されている。

1章では、グループウェアについての考え方、グループウェアの分類、各種グループウェアの現状と技術的動向について解説した。

2章では、グループウェアにおけるデータベースシステムの基本技術、データベースシステムがもつ特徴、グループウェアにおけるデータベースシステムの設計等についての調査研究の成果を紹介した。

また、III グループウェア発展に伴う情報システムの高度化並びにデータベース構築技術の進展とその課題では、今後の高度情報化社会におけるオフィスワークにおけるグループウェアとそれを支えるデータベースの重要性、グループウェア発展に伴う解決すべき課題等について考察した。

以下にその概要を述べる。

今回の調査研究の成果が、今後のデータベース構築技術の向上に役立つことになれば幸いである。

21.2 グループウェアにおけるデータベースシステムの構築

21.2.1 グループウェアの技術的動向

グループウェアの言葉の由来、考え方、システムコンセプト等について簡単に紹介した後、グループウェアの分類に基づき、それぞれについて解説を行った。

(1) グループウェアの分類

グループウェアは、その使用方法によって時間的特性と空間的特性によって分類される。ここでは、これらの分類について解説を行った。

(2) 即時直接対面型のグループウェア

この分野での代表的なグループウェアとしては、電子会議システムがある。ここでは、同システムについて、システムの概要、技術動向及び具体例等を紹介した。

(3) 即時遠隔分散型グループウェア

グループのメンバーが遠隔地に分散されている環境において、メンバー間での協調作業をリアルタイムで支援するものとして、LAN、WAN等のネットワークの利用が前提となる。また、この種のグループウェアでは、メンバーが分散していることから、アクセス制御、作業の進行管理、同時実行制御等をいかに行うかが重要な技術課題となる。こうした背景から、以下のシステムについて、現在の研究開発状況等について調査研究を行った。

① グループエディタ

テキスト文章とかグラフィックデータをワークステーション上で編集するシステム

② ウィンドウ共有／画面共有システム

ウィンドウや画面がアプリケーションと独立に共有できるシステム

③ ワークステーション／ビデオ会議システム

テキスト、グラフィック、音声、動画等マルチメディア対応のシステム

(4) 非即時遠隔分散型グループウェア

これは、オフィス内あるいは遠隔オフィスに分散されたメンバーが電子メールを用いてメッセージの交換を行いながら協調的に作業を進めていくためのシステムである。しかし、対象としている作業が非同期であるため、なんらかの協調モデルを用いて支援する必要がある。

このため、以下のシステムについて、システムの概要、具体的開発例、技術的課題等について調査研究を行った。

① ビジネス・コーディネーションシステム

言語行動理論に基づいて会話をモデル化した対話支援システム

② 議論支援システム

議論を「問題」、「見解」、「意見」という枠組みで表現したモデルを基本にしたシステム

③ オフィスワークフローシステム

オフィスにおける定型的なプロセスを手続きモデルとして構造化し、電子メールを用いて支援するシステム

④ メッセージ構造化システム

作業間メッセージを集中制御メカニズムをもたずに協調作業の支援を行うシステム

⑤ 協調文書作成支援システム

電子メール、ファイルサーバー等を用いて非同期で文書の協同執筆、レビュー等を行うシステム

21.2.2 グループウェアにおけるデータベースシステム構築技術

(1) データベースシステムの技術的動向

近年のコンピュータネットワークの発展、パーソナルコンピュータ（PC）、ワークステーション等の進展に伴い、情報システムは急速に分散化の傾向にある。こうした中で、データベースシステムも、複数のコンピュータ内にデータベースを分散させた形態や、DBMSの機能をサーバーとクライアントに分散させた機能分散の形態となってきた。したがって、この分野で重要とされるデータモデルである、リレーショナルモデルについてデータ構造、操作演算等について調査研究を行った。

(2) グループウェアデータベースシステムの基本技術

グループウェアデータベースシステムを実現するための基本技術として、オブジェクト指向データベースシステムと分散型データベースシステムがある。特に、分散型データベースシステムにおいてはシステムの自律性と協調が重要となることをふまえ、これらについて調査研究を行った。

① オブジェクト指向データベースシステム

ここでは、オブジェクト、メッセージパッシング、階層化と継承等まず基本概念につ

いて考察し、実際のプログラミングの例、リレーショナルモデルでのオブジェクト指向、オブジェクト指向データベースシステム利用の効用等について解説した。

② グループウェアにおけるデータベースシステムの自律性と協調

グループウェアでは、複数のデータベースシステムがネットワークで結合された形態となるため、自律的な分散型データベースシステムを連邦的に統合する問題について調査研究を行った。

このため、分散型データベースシステムについての考え方を明らかにした後、多データベースシステムの異種性、自律性、分散性について考察を行った。グループウェアでは、こうした自律型分散データベースシステムを統合的に利用する必要があるため、管理者を設定してコントロールを行う完全統合化あるいは全体のコントロールを行わない非統合化を含め、種々のレベルでのデータベースシステムの統合化について調査研究を行った。更に、分散型データベースシステムの参照モデルをベースに独立性について考察し、システムアーキテクチャ及びそのためのトランザクション管理について調査研究を行った。

③ グループウェアにおける分散型データベースシステムの設計

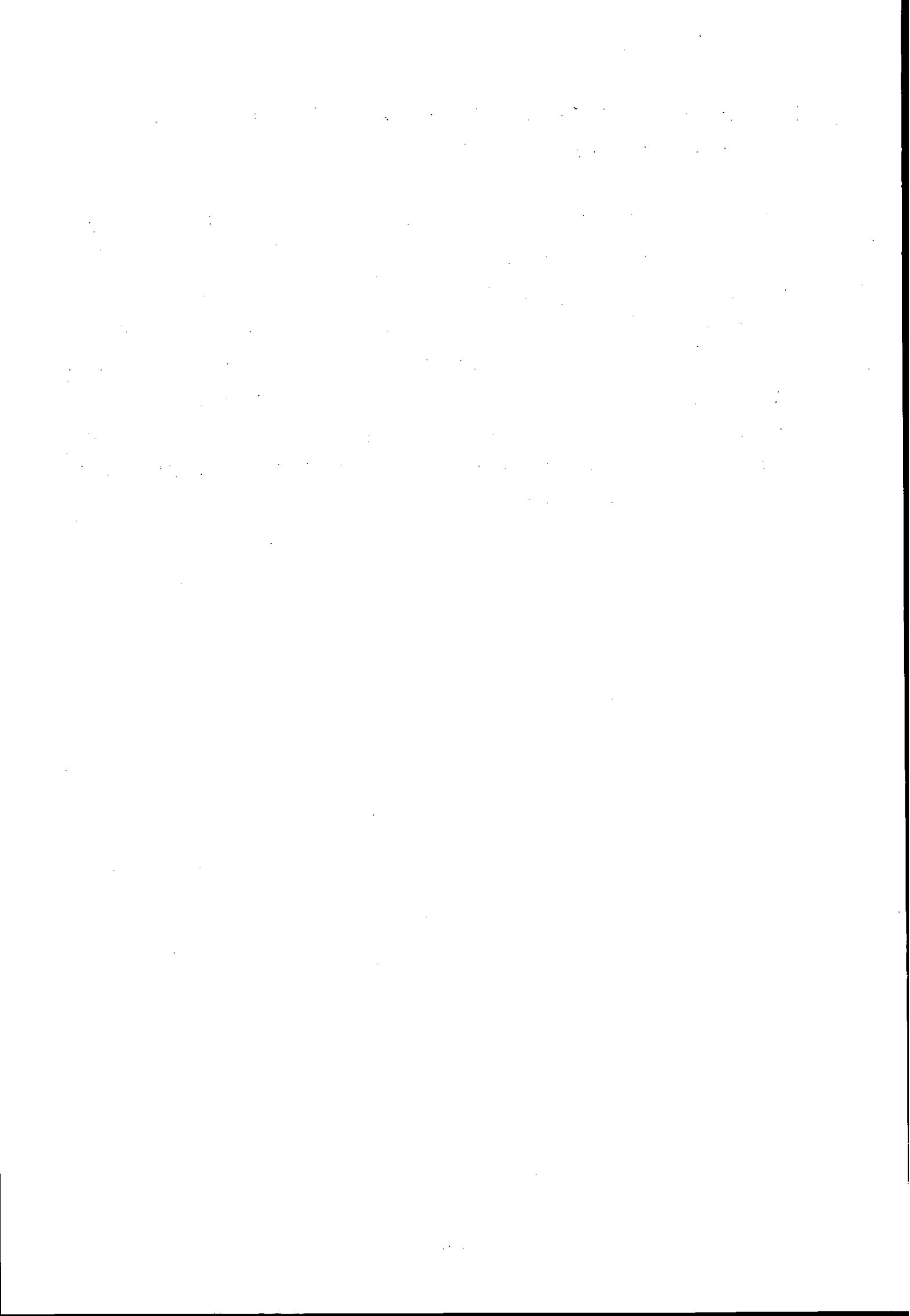
分散型データベースシステムの設計方法には、ボトムアップ方式とトップダウン方式の2種類がある。このため、データベースシステム内のスキーマを定義し、スキーマ変換等について考察した。更に交渉プロトコル、スキーマ統合等グループ通信方式について調査研究を行い、代表的な分散型データベースを例示した。

(3) グループウェアデータベースシステム

従来のリレーショナルモデルやネットワーク型モデルのデータベースシステムは、事務処理等の定型的な業務には適しているが、グループワークのような複数の利用者が協同して行う作業を支援することには適していない。つまり、グループウェアを遂行するためには図形、マルチメディア情報、ソフトウェアモジュール、スケジュールといった種々の情報が必要となり、これらの相互の関連構造を扱えるデータベースシステムが必要となる。このためには、オブジェクト指向データベースシステム、分散型データベースシステム、利用者インタフェースが重要である。

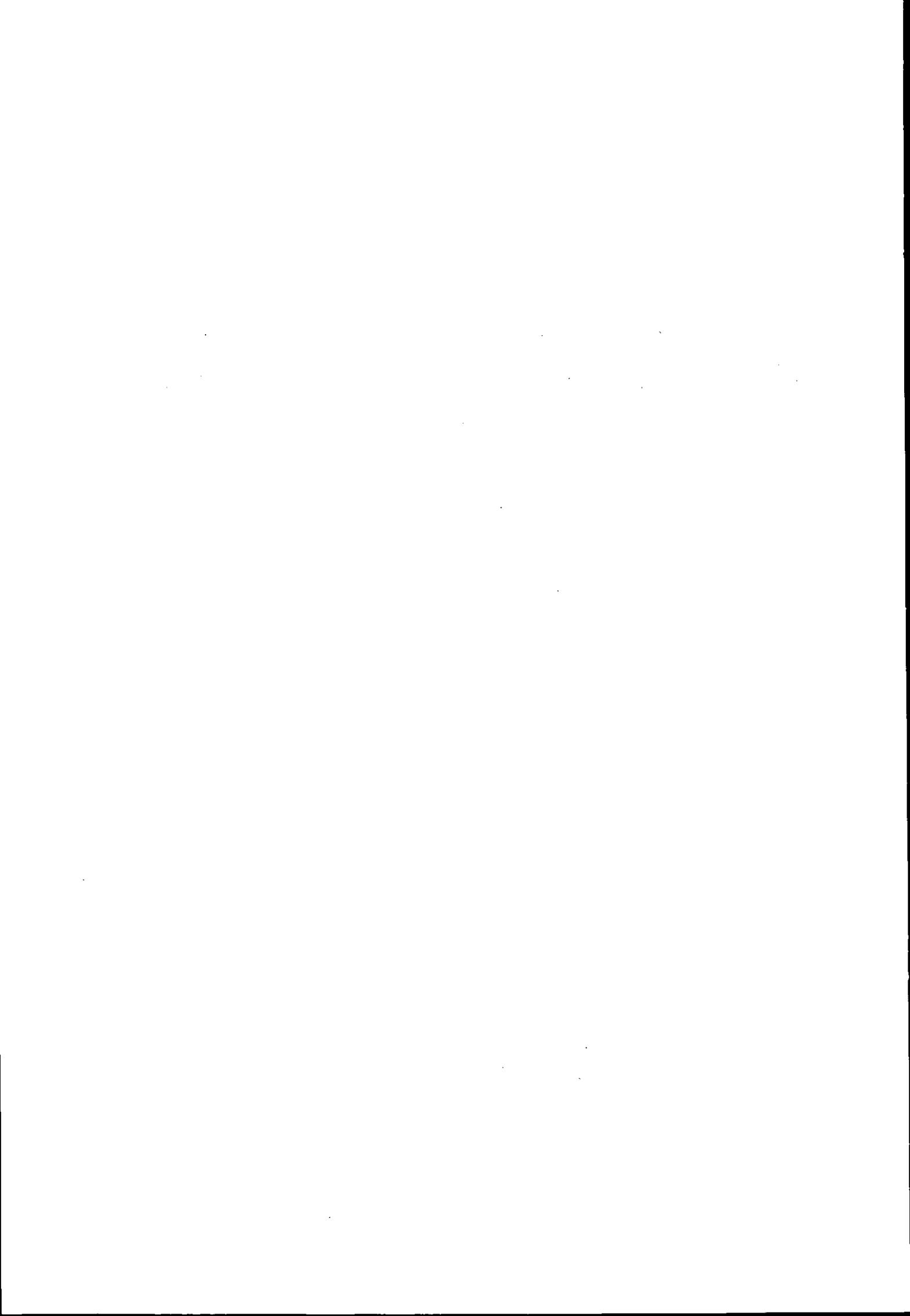
21.3 グループウェア発展に伴う情報システムの高度化並びにデータベース構築技術の進展とその課題

情報技術の進展に伴い、現在、高度情報化社会を迎えつつあるが、事務分野では、依然、労働集約型であり、多くの人間が介在した会議、打ち合わせ等に時間がさかれ、このためこの分野の生産性の向上が求められている。このため、B-I S D Nの普及等通信技術の高度化等ともあいまって、今後は、オフィスの分散も積極的に行われるようになるだろう。このためには、分散オフィスにおけるグループウェアを支援するグループウェア及びこれのベースとなる情報を管理するためのデータベースシステムが極めて重要となる。しかし、こうしたデータベースシステムを構築していくには、解決すべきさまざまな課題がある。このため、グループウェア設計上の問題点、グループウェア利用上の問題点、グループウェアの統合化等の課題について調査研究を行い、その成果を記述した。



22 パーソナルコンピュータとLANの利用による
非定形データベースのプロトタイプ作成

株式会社 メイテック



22. パーソナルコンピュータとLANの利用による 非定形データベースのプロトタイプ作成

22.1 概 要

近年におけるパーソナルコンピュータ（パソコン）の普及率にはめざましいものがある。特にノートパソコンの発売は、パソコンの普及率を格段に高めたことは否めない事実である。このような現状の中で、文書及び資料は、そのほとんどがパソコンを利用した市販のアプリケーションソフトで作成されるようになった。

しかしながら、作成された文書及び資料（以下これらを情報という）は、そのほとんどが個人のコピーディスクで管理されているのが実状である。このことは、情報の管理体系と必要とする情報の検索に大きな障害となっている。したがって、情報の有効な利活用が行われなため、これら情報が死蔵化してしまい、業務の効率化と質的向上に対する大きな阻害要因となっているのが現状である。

本事業は、このような実状をふまえ、社内において非定形データベース検索システムを構築し、これによって、社内における情報の管理手法を規定するとともに、非定型データベース構築に係わるさまざまな技術的問題点を解決し、これらノウハウを広く社外に普及させることとする。

なお、プロトタイプの実成に当たっては、社内にて「統合情報管理システム（仮称）」という名称をつけた。

22.1.1 目 的

パソコンで作成された社内文書及び資料等のデータベースを構築するため、これに伴う技術開発を実施し、社内情報の管理、検索における一貫したデータベース構築技術の確立及び社内データベース構築の促進を図る。

22.1.2 基本方針

- (1) ハードウェアシステムは、社内でも利用されているものを活用する。
- (2) ソフトウェア環境は、汎用的なOSを使用する。
- (3) データベースに登録する情報は、現在コピーディスク等で個々に管理されている書類や図面とする。
- (4) 情報は、インデックスにて管理する。

- (5) 操作環境は、マルチウィンドウにてマウスを用いるようにし、キーボードからの操作を極力減らすようにする。
- (6) 開発するシステムは、移植性、拡張性を考慮する。
- (7) プログラム言語は、C言語とする。これに伴うコンパイラは、汎用的なものを使用する。

22.2 システム構成

22.2.1 ハードウェアシステム

プロトタイプを搭載するハードウェアシステムは図22-1のとおりである。

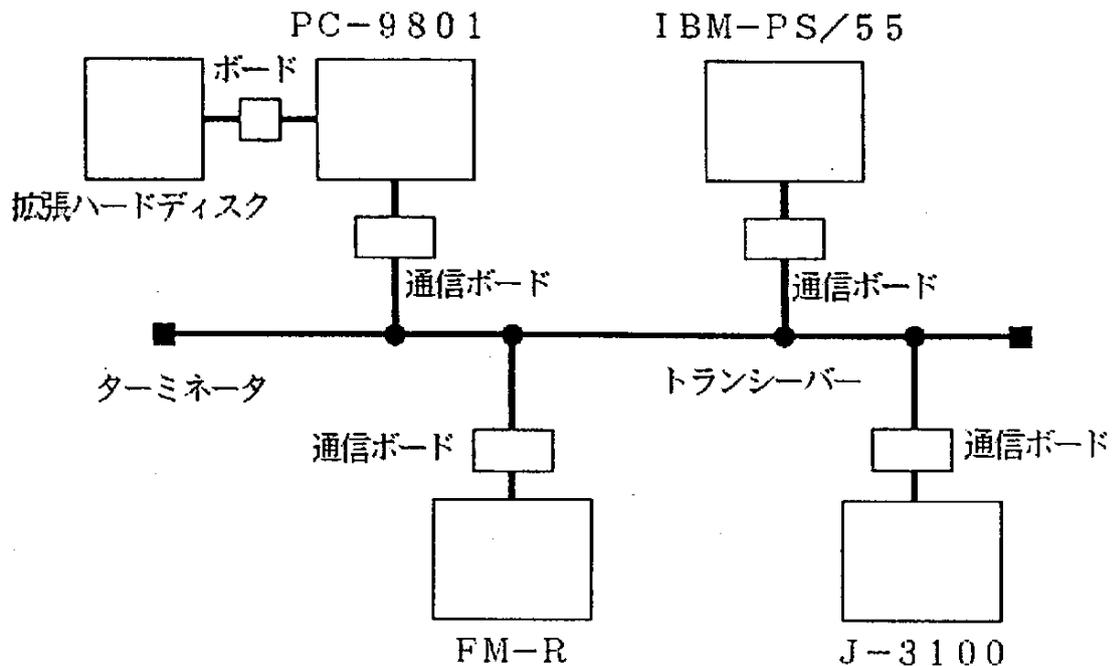


図22-1 ハードウェアシステム

ハードウェアは、社内で利用されているパソコンを活用することを基本とする。

本体 NEC, 富士通, 東芝, IBMから発売されている, 汎用的なパソコンを対象とする。

拡張機器 ネットワークボード
 拡張ハードディスク

光磁気ディスクユニット
 トランシーバー
 ターミナータ
 ネットワークケーブル

22.2.2 ソフトウェアシステム

プロトタイプを作成するに当たってのOS及び開発環境は、次のとおりである。

OS OS/2 拡張版
 IBM-DOS J4.05
 MS-DOS 3.3以上
 ネットワークOS 自社開発

22.3 運用概念

プロトタイプ運用の基本的概念は、作成された情報にそれぞれインデックスを付与し、この付与されたインデックスをデータベース化し、直接の検索対象とする。インデックスのイメージは、図22-2のとおりである。

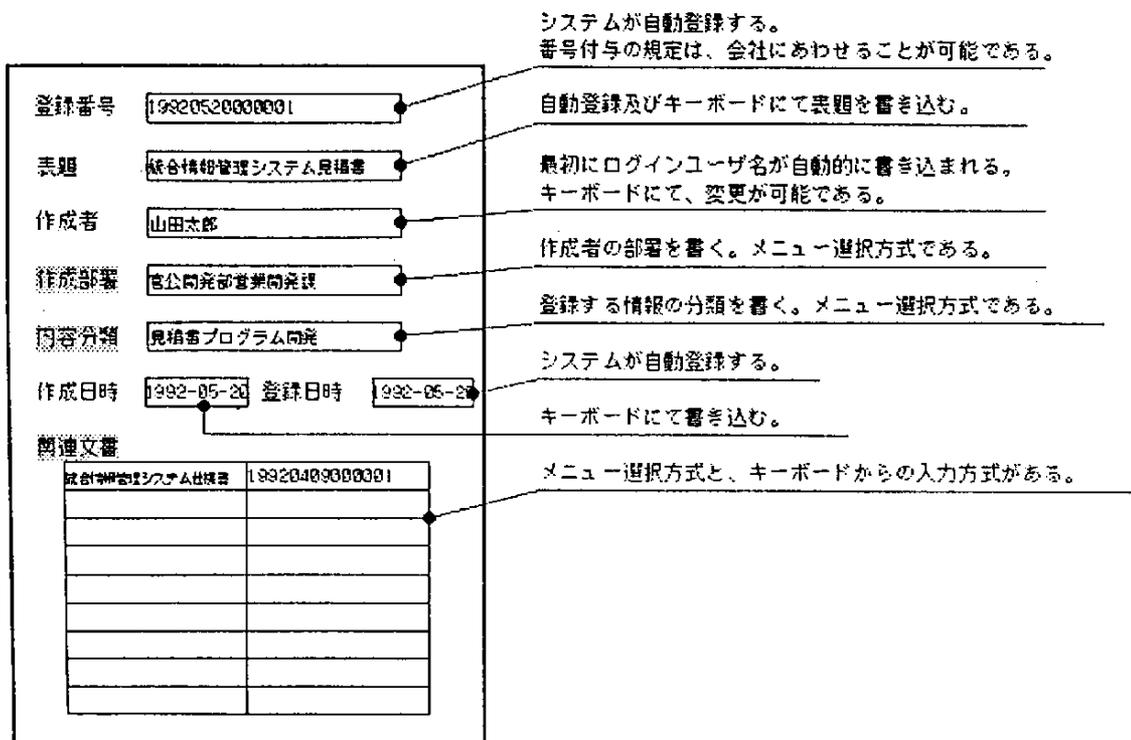


図22-2 インデックスのイメージ

プロトタイプの利用の基本方針は、情報の検索手法の簡易化及び豊富さと情報の登録の容易さである。これらについて以下に記述する。

22.3.1 情報の登録

通常のワープロソフトによる情報の作成から登録作業までの流れは図22-3のとおりである。

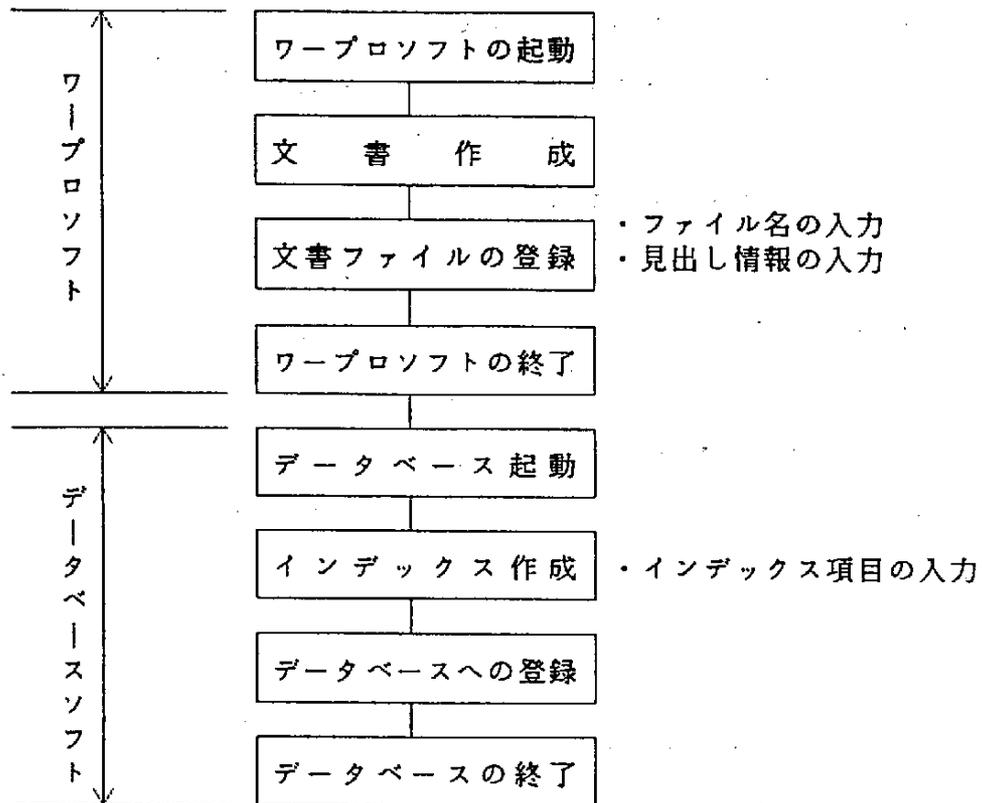


図22-3 ワープロソフトによる情報の作成から登録作業までの流れ

図22-3の流れからも分かるとおり、ワープロソフトの起動及び終了、データベースソフトの起動及び終了、更には、ワープロソフト側でのファイル名等の入力、データベースソフト側でのインデックス項目の入力等、さまざまな形で重複作業が発生していることに気づく。

プロトタイプの利用構想策定に当たって、この情報の作成から登録に至る一連の流れを簡素化することを第一の目標に定めた。即ち、重複作業の回避とインデックス項目の入力の自動化である。

重複作業の回避については、プロトタイプから、アプリケーションプログラムの起動を行うことによって、操作の簡略化を実現した。また、アプリケーションプログラム側で設定す

る見出し情報を自動的にインデックスの見出し情報に取り込むことにより、重複作業の回避を図った。この結果、プロトタイプにおける情報の作成からデータベースへの登録までの流れは、図22-4のように簡素化することに成功した。

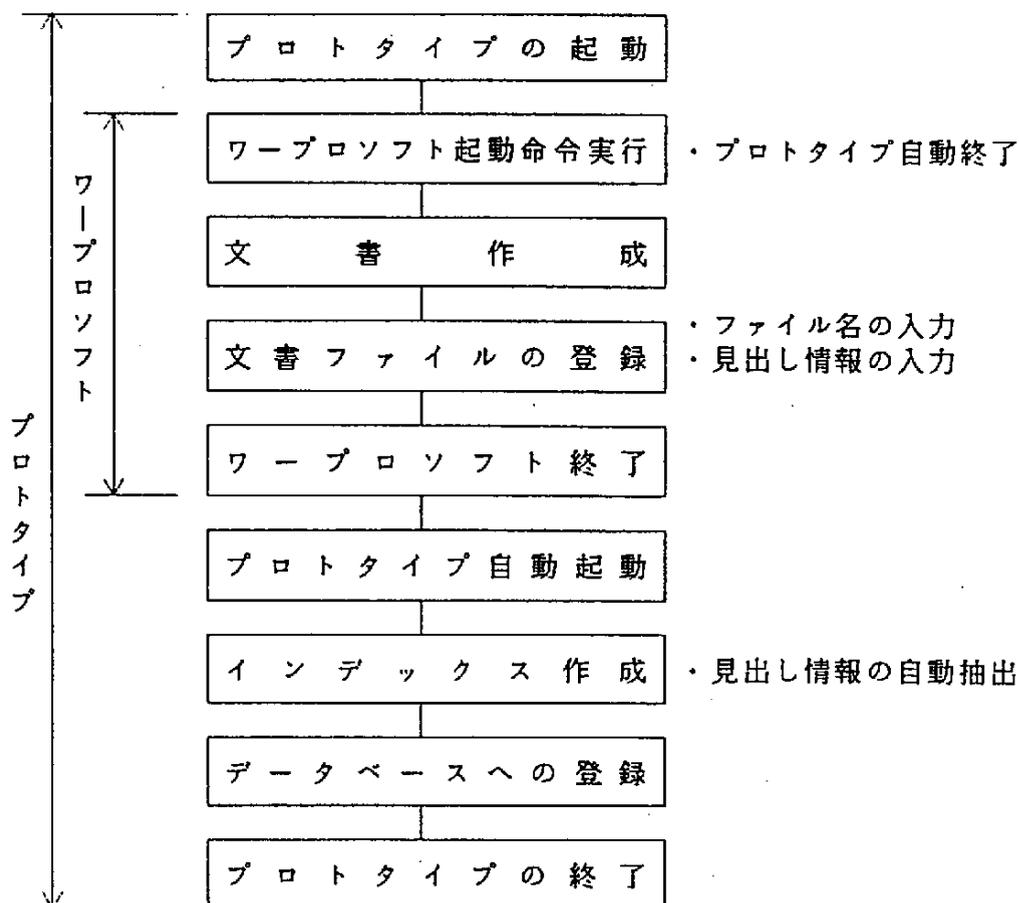


図22-4 プロトタイプにおける情報の作成からデータベースへの登録までの流れ

図22-4の流れからも分かるように、ワープロソフトの起動、プロトタイプの再起動の手順が自動化され、しかも文書ファイル登録時に入力する見出し情報をインデックスの表題に自動入力することが可能になったため、情報のデータベースへの登録がかなり簡素化される結果となった。

22.3.2 情報の検索

データベースの検索は、キーワードによる検索が一般的である。プロトタイプにおいては、このキーワード検索をあえて否定することを試みた。その理由として、第一に、キーワードファイルが大規模になるため、パソコンの磁気ディスク容量が圧迫され、本来の情報格納容量が十分にとれない恐れがある。第二に、キーワードのメンテナンスに多大な時間とコスト

が発生する。このため、プロトタイプにおいては、キーワード検索の機能をもたせなかった。これに代わる検索手法として、多観点分類別検索を考案し、これを採用することにした。この検索手法について説明をする。

一般に情報は、その見方を換えることによって、データベースの内容が全く異なったようにユーザには見えてくる。例えば、アパートの空室情報のデータベースがあったとする。このデータベースの見方を沿線別、間取り別、価格別というように3点の観点からみることによって、ユーザにとっては、一つのデータベースがあたかも三つ存在しているかのように見える。即ち、沿線別のデータベース、間取り別のデータベース、価格別のデータベースというようにである。

これによる利点は、第一に、データベースを検索するに当たって、複数の検索の糸口が与えられるということであり、第二には、この分類に要するメンテナンスが、キーワードに比較して、格段に少なくすむということである。

情報を複数の見方でみると、大きく三つの見方がある。即ち、作成部署別、内容別、顧客別である。これらを具体的に記したものが表22-1である。

このように情報を分類することによって、再現性が高く、かつ、網羅性に優れた検索が可能となる。

プロトタイプ作成において扱う情報は、フロッピーディスクにて管理されている書類、資料及び図面である。即ち、社内でアプリケーションソフトを使用して作成された情報である、という前提条件のもとに表22-1のごとく情報の分類を試みた。

分類の観点は、情報が作成された部署、配布された部署及び情報の内容についてである。これら分類を更に細分化することにより、きめ細かい、かつ、検索の糸口の豊富な「辞書」を作成することが可能となる。

22.4 機能概要

プロトタイプの機能は、図22-5のとおりである。

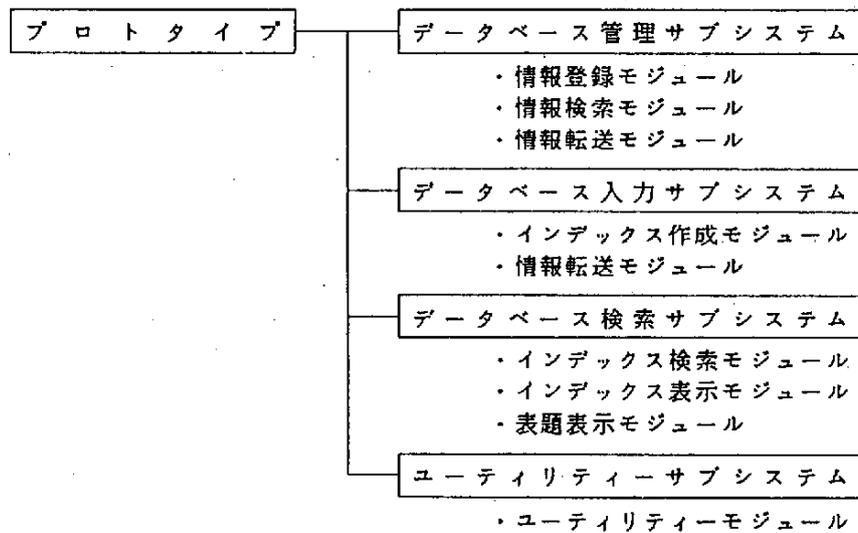


図22-5 プロトタイプの機能

細部機能は、以下のとおりである。

22.4.1 データベース管理サブシステム

データベース管理サブシステムは、データベースに登録されている情報を直接管理するサブシステムである。このサブシステムは一部のモジュールを除いて、ユーザが直接動作させることはできず、データベース入力サブシステム及びデータベース検索サブシステムからのコマンドを介して動作させることができる。即ち、これら二つのサブシステムから送られるコマンドを判断して、データベース管理サブシステムの機能が動作することとなる。

22.4.2 データベース入力サブシステム

データベース入力サブシステムは、情報をデータベースへ登録するための前処理をするサブシステムである。データベース入力サブシステムの基本的概念を図22-6に示す。

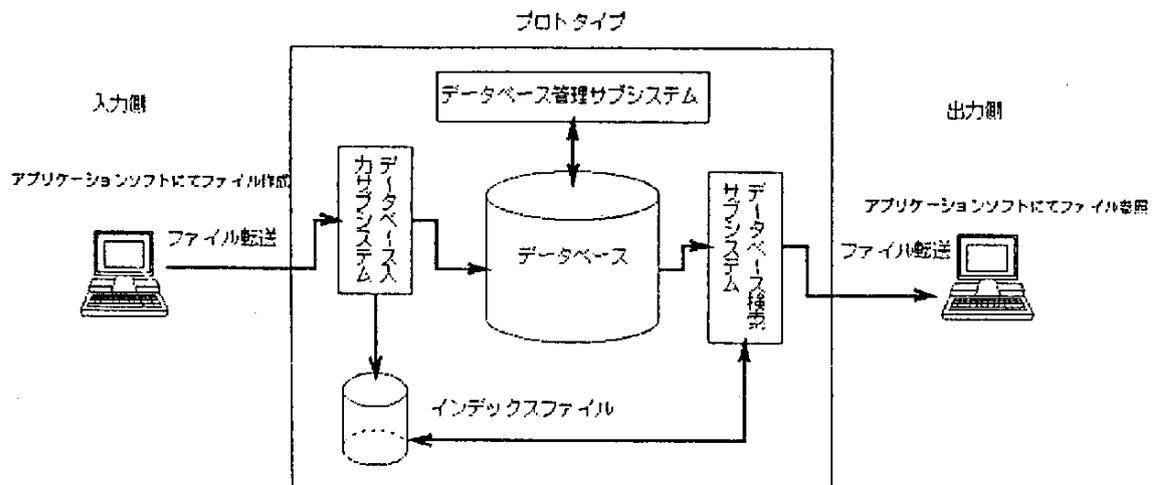


図22-6 データベース入力サブシステムの基本的概念

22.4.3 データベース検索サブシステム

データベース検索サブシステムは、情報の検索を行うサブシステムである。

The screenshot shows a search interface with the following elements:

- 登録番号 (Registration Number):
- 表題 (Title):
- 作成者 (Author):
- 作成部署 (Department):
- 内容分類 (Content Classification):
- 作成日時 (Creation Date/Time):
- 登録日時 (Registration Date/Time):
- 検索実行 (Execute Search):
- 表題表示 (Display Title):
- インデックス表示 (Display Index):

インデックス検索を実行すると、左の図がウィンドウに表示される。この図の項目に検索の情報を入力し、「検索実行」のコマンドをマウスでクリックすると、検索が実行される。

登録番号、表題、作成者は、キーボードからの入力となる。ワイルドカードの使用ができる。

作成部署は、この項をマウスでクリックすると、データベース入力サブシステムと同様に、作成部署一覧のウィンドウが表示される。この中からマウスで選択してこの項目を入力する。この検索の特徴は、作成部署が階層関係になっていることから、例えば「官公開發部」と入力すると、官公開發部で登録されている情報が全て検索できる。

内容分類は、この項をマウスでクリックすると、データベース入力サブシステムと同様に、内容分類一覧のウィンドウが表示される。この中からマウスで選択してこの項目を入力する。この検索の特徴は、内容分類が階層関係になっていることから、例えば「見積書」と入力すると、見積書が全て検索できる。

作成日時、登録日時はキーボードからの入力となる。ワイルドカードの使用ができる。

検索方法は全て「AND」検索である。検索情報は、全て記入する必要はないが、全く記入しないと、検索は行われない。検索情報の記入が終わったら、検索実行コマンドをマウスでクリックする。

検索結果の表示は、「表題表示」「インデックス表示」の2種類ある。表題表示は、表題と、登録番号のみの表示である。インデックス表示は、インデックスの表示である。

図22-7 インデックス検索画面とその説明

22.5 プロトタイプの評価

プロトタイプの評価を実施するに当たって、操作性、応答性及び拡張性を評価の項目とした。評価の基準としては、はなはだ主観的ではあるが、社内にてデモを繰り返すことによって、それぞれの意見をプロトタイプの評価とした。

22.5.1 操作性に関して

情報の登録作業から検索に至る一連の流れの中で、特に情報の登録に関しては、まだまだ違和感を与える操作が存在するようである。

プロトタイプを今後運用することによって、情報の登録作業に関する適切なマニュアルが作成され、この作業がルーチン化されれば、評価も自ずと変わってくるように思える。

情報の検索については、ある程度の評価を得ることができた。この評価のポイントとしては、やはり、キーボード操作を極力減らしたのが第一の要因であり、第二は、前述したような検索手法をとったことが要因として挙げられるのではないかと考える。

操作性に関する総合的な評価は、おおむね良好と考えるが、個々の点を見て行けばまだまだ改良を加えるところがある。

22.5.2 応答性に関して

応答性に関する評価については、その主体になるものとしては、情報の検索速度である。現在のところデータ数がまだ乏しいため、明確に評価する訳には行かないが、今後データ数を増やして行ったときに、かなり検索速度が悪くなるように思われる。画面の推移等に関しては、満足の行く応答速度が得られた。

応答性に関する総合的な評価は、今後の試験運用を待たなければならないが、仮に応答性に関する評価が悪くても、それに対する対象は容易にたてられる。よって、応答性についても、それなりに評価は期待できるものとする。

22.5.3 拡張性に関して

拡張性に関する評価については、適用機種の種類内であれば、問題なく動作する。台数の拡張については、イーサネットケーブルの物理的諸元の種類内であれば、動作させることは可能であるが、応答性に問題が発生する可能性がある。

拡張性を確保できた大きな要因としては、画面作成ツールを採用した結果による。画面作成ツールがそれぞれの機種の種類内のプラットフォームの役割をはたした。また、ネットワークOSについても、基本的な部分については、メーカー提供のOSを採用したが、それより上部のアプリケーション層については、当社で開発したものを採用したため、メーカーにとらわれることなくプロトタイプが開発が行えた。

よって、今後の拡張については、ほとんど問題なく行えるものとする。

22.6 今後の課題

今後の課題として考えなければならないのは、実用システムへの移行についてである。実用システムの移行に際して、プロトタイプの評価を考慮した。

プロトタイプで得られた評価について、第一に考慮すべきことは、操作性の更なる向上と扱える情報の範囲の拡大である。具体的には、イメージ情報等である。

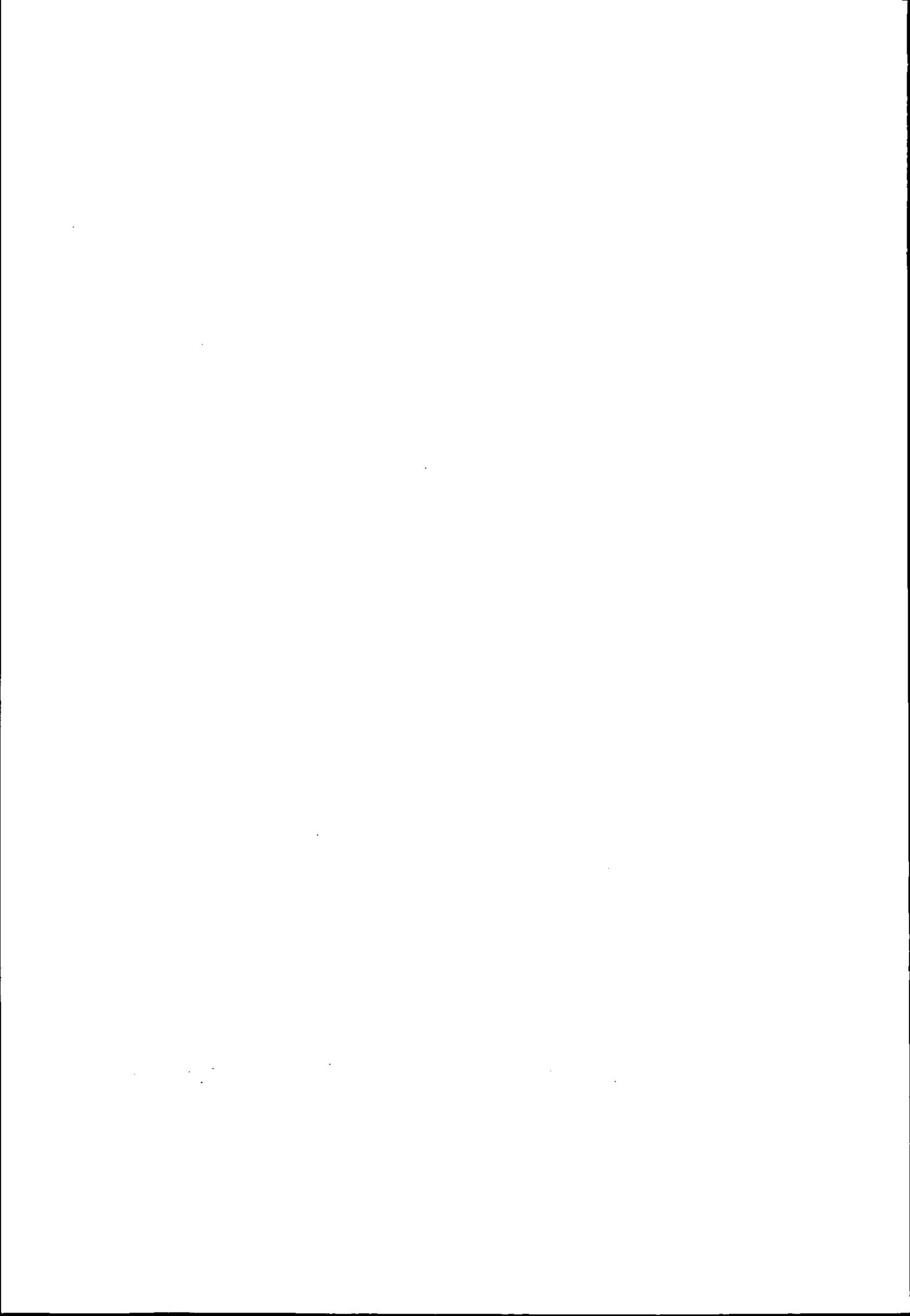
第二に考慮すべきことは、システムが遊びをもつことである。これによって、楽しみながらシステムを運用することが可能となり、幅広く運用されることが予想される。プロトタイプにおいては、機能及び仕様を満足させるためにこのような余裕がなく、無味乾燥なものになってしまったが、実用システムに移行するときは、大いに遊び心をもったシステムを構築することを心がける必要がある。

表22-1 情報の見方の例

作成部署別	内容別
本社	一般文書
総務部	通達
総務第1課	案内
総務第2課	見積書
人事部	提案書
人事第1課	提案書
人事第2課	会議資料
人材開発課	営業会議
経理部	議事録
経理課	技術会議
財務課	議事録
広報部	部課長会議
広報第1課	議事録
広報第2課	打ち合わせ議事録
経営企画部	技術打ち合わせ
経営企画課	営業打ち合わせ
経営管理課	設計書
情報システム部	基本設計書
情報システム課	概要設計書
システム管理課	仕様書
研究開発部	顧客別
技術管理室	A社
研究開発室	技術部
商品開発室	B社
東京本社	総務部
総務部	C社
営業企画部	設計部
営業企画課	D社
コンサルティング課	テクニカルセンター
官公開発部	
業務管理課	
営業開発課	
特機営業課	
特機システム技術室	
品質保証課	
技術開発部	

23 知的資源型データベースの調査研究

株式会社 ジャパンコミュニケーションズインスティテュート



23. 知的資源型データベースの調査研究

23.1 調査研究の目的

これまでの情報化社会は、情報処理技術及び情報伝達手段の発展と相まって、情報量の飛躍的増大により地球規模のネットワークを築き上げ、これにより情報利用者は時空間の観念の枠を大きく広げてきた。

特に、データベースはこのネットワークの中核として位置づけられ、情報の源流としての役割を担ってきたのは事実である。しかしながら、これらのデータベースで蓄積・流通されている情報の形態はテキスト情報が主体であり、最近イメージ情報の一部が加わっているものの、その情報をもつ意味や表現形式は自ずから限定されている。

一方、情報技術と情報利用の多様化が更に進展しつつあるなかで、仮想現実の具現化、感性情報の収集や利用等が進みつつあり、情報利用と、情報そのものの質的変換が著しい。これらに利用される情報は「知的資源」と総称され、「情報」、「情報を生み出すための知識、知恵」及び「創造性等を育む社会的環境」が代表される。

このような知識資源利用の実現には、利用できる情報（知的資源）の収集・蓄積・流通及びユーザ側のニーズが十分に検討されなくてはならない。ここに「知的技術」に代表されるような新技術の研究開発や製品化が求められているのに止まらず、この知的資源の活用は現在のデータベースやシステム概念を大きく変えることも意味するものである。

以上のような状況に鑑み、情報利用者の視点に立って、知識資源となりうべき情報の所在、特徴及び利用形態などを分析・把握し、社会の発展に沿ったデータベースの高度利用及びシステムの有りうべき姿を、近未来、遠未来にわたって予測し、必要な方策を検討し、新しい情報需要及び利用形成についての検討の資に供することを目的として、調査研究を実施した。

23.2 調査研究体制

この調査研究を実施するに当たり、幅広くユーザの意見や要望を集めるため、(財)データベース振興センターと、電子ライブラリーコンソーシアムとが協力して「知的資源利用調査研究委員会」、「電子ライブラリーコンソーシアム特別研究委員会」を設置し、この委員会の提案、意見、要望等を収集・整理した。また、これらの委員会とは独立して(財)データベース振興センターに設けられた「知的資源利用調査研究に関するアドバイザー会議」からの意見をまとめた。

23.3 検討項目

(1) 知的資源として利用される情報の概念形成について

新しい情報需要となりうるものを掘り起こし、その需要と供給の可能性及び発展性を検討し、またこれらを利用可能ならしめる技術についての検討を行う。

- ・ 知的資源の概念、及び種別
- ・ 知的資源の現状、将来展望
- ・ ユーザ側からみた需要の態様と範囲
- ・ 市場形成に必要な要素技術の検討及び課題

(2) 知的資源の利用方法等の実態調査の実施

既に知的活動に利用されている知的資源利用システムの応用例や知的資源の所在に関する実態調査を行い、それぞれの動向、問題点、課題等を探る。

- ・ 博物館、美術館、図書館等の資源所在及び利用形態
- ・ 感性情報、カオス情報等の意味情報の利用の動向

(3) 知的資源利用のための方策の検討

上記の検討をふまえ、知的資源の利用を促進するための方策を検討し、提言の形で報告書としてまとめる。

- ・ 知的資源利用に必要とされるユーザのあり方
- ・ 望まれる知的資源提供・利用のためのデータベース、システムのあり方

23.4 検討内容

平成4年度は、設置された委員会及びアドバイザー会議において、以下の項目についての検討を行った。

- (1) 社会的なインフラストラクチャとしての知的資源の必要性（データ開発）
- (2) 知的資源の収集・蓄積・流通の各レベルでの新しい技術やメディアの必要性
- (3) 情報表現の多様化や時空間を超えた情報の利用とその価値の認識
- (4) 知的資源を利用した場合の個人の意識変革や社会的変化の特長の認識

これらの項目についての主要なキーワードは以下のとおりである。

23.5 主要キーワード

(1) 社会的なインフラストラクチャとしての知的資源の必要性（データ開発）

① 従来の枠組みとは異なった新しい情報化投資

- ・情報化の飽和感があるが、これは現在の情報化には充足感がないことの現われ
- ・情報利用の進展により、新たな情報化投資が可能となる

② 脳とニューラルネットワーク — 知的資源利用の二つの意味

- ・膨大な情報を使いやすいものに転換するシステムの構築
- ・脳の知的活動により情報蓄積と情報整理の方法

③ 電子ライブラリーコンソーシアム

電子ライブラリーとは、「電子図書館」という考え方ではない。図書館、美術館、博物館等に所蔵されているすべてのデータを、知的資源として活用するという従来の枠組を超えた考えが電子ライブラリー。

④ 当委員会テーマに対する二つのキーワード

- ・知的資源は、情報の元 — クリエイトしていくもの
- ・未来を志向した知的資源の活用

⑤ データの拡大とバーチャルリアリティ

- ・ビジュアルデータとしての恐竜の骨、人間がどう感じているかという公害データ等、人間の五感に対応する感性データまでの拡大。テキストデータ中心の現在のデータベースを超えて、人類や時代が求める課題に対応する新しいデータの拡充や新技術の開発によるバーチャルリアリティの実現。

(2) 知的資源の収集・蓄積・流通の各レベルでの新しい技術やメディアの必要性

① 情報技術の発展と社会への応用

- ・ユーザ側の観点からは、多くの取り組みテーマがある。（供給側では行き詰まり感＝普及しない。）
- ・ユーザインタフェースを考慮した情報機器の提供の必要性

② 意味ある情報のデータ化と情報の復元の方法

- ・音楽CDにおけるデータ化と復元は、一定のロジックによるフィルターによって行われる。
- ・バーチャルリアリティでは、フィルターのあり方は重要テーマ
- ・電子デバイス、カオス、マイクロマシンなど新しい観点からのフィルターのかけ方の吟味

③ コミュニケーションギャップを埋めるインタラクティブシステムが具体的な領域で、AIなどで相手のレベルに対応するようなものとして、つくられたら良い。

- ④ ものごとの生成過程を具現化するインタラクティブシステムとそのためのデータベース
- ・情報のやりとりをしながら相互に情報が編集加工され、結果が出てくる。
 - ・創造的なプロセスで利用できるデータベース — 関心のあり方に対応してオーガナイズされた情報が重要。
- ⑤ 知識データベースと認識データベース
- ・従来のデータベースは知識データベース
訴訟などで争うときは有効
 - ・認識データベース
快適な人生や新しい何かを生み出すために必要。これは、ソフトかもしれないが、カオス、ニューロ、アナロジーというもので大変重要。
 - ・AIでは「演繹」についてはある程度できるが、「帰納」は相当困難といわれている。しかし、各委員の話ではこれへの期待が大きい。
- ⑥ 価値ある情報は、創造力をつくり出す
- 新しい技術による新しい創造
- ・「映画」という新しい芸術の誕生
 - ・ニューメディア、情報機器による新しい芸術の創出。その一つとしてのバーチャルリアリティ。
- ⑦ 枠組変更の必要性と電子化辞書研究所
- ・電子化辞書 — 200億円の資金を集めた同床異夢
 - ・枠組内での壮大な無駄遣い — 5年後の結果
40万語の語彙集、数10万語の日英対照表
 - ・枠組変更なしに、金と汗だけでは何も生まれない。
 - ・枠組変更による電子化辞書のテーマ
ニューロ、ファジー、AIなどにより、人間という有機物が母親の胎内から蓄積してきた常識や永い歴史の中で捨象され変化してきた伝統などを無機物であるコンピュータにどう移植するかの研究。
- ⑧ 合財袋と情報活用の豊かさ
- ・個人の生き方は、生活のベースとなる知識を未整理のまま詰め込んだ「合財袋」から生き生きした状態を取り出して使うことで成立している。
 - a ; 「合財袋」というわれわれのもつ内側の技術としてのシステム — 未検証ではあるが、AIなどによってどこまでマシンに預けられるか。
 - b ; 古文書判別での「墨の色」というような表現技術の豊かさや、「合財袋」という検

索・思考技術の豊かさの実現は、人間とマシンの間のユースウェアの豊かさとそれに対応する質の高いデータベースづくりが課題。

- ⑨ スーパーコンピュータ — シミュレーションの段階を超えた第一原理からの原理計算
- ・仮定を一切おかないで、分子レベルから計算したときに、材料の性質などすべてをコンディションできる。
 - ・ケミストリーでは、ケミカルリアクションを全部スーパーコンピュータでやり、新しい物質などすべてつくってしまう。コンピュータのつくった材料の特許は、誰が取るか。
 - ・スーパーコンピュータが、理論科学、実験科学の他にもう一つ違う科学をつくった。
 - ・乱流方程式は今でも解けないが、こういう世界でスーパーコンピュータが利用できれば、知的資源のジャンルが一つ増える。
 - ・原理計算は科学の世界のバーチャルリアリティである。スーパーコンピュータが材料をつくり、製法、性質も解明。
 - ・原理計算は、基本原理が明確に分かっていること。第一原理からの出発。
 - ・シミュレーションは、原理は分からないが経験的に判明しているパラメータよっての推定。
- ⑩ スーパーコンピュータのシミュレーションの限界と実験の重要性
- ・スーパーコンピュータでシステム的设计をしたが、試作で最も難しかったのは、あるインタフェースの装置をつくることだった。こういうことは、やってみて分かること。コンピュータは、判明した条件しか与えられていないということを人が認識していることで安心して使える。

(3) 情報表現の多様化や時空間を超えた情報の利用とその価値の認識

① 情報の区別

- ・客観的事実としての情報
- ・提供側のメッセージ、価値判断を加味した情報

② 情報は既存のものという概念

- ・知的資源はイマジネーションや仮説も含む。

③ 一次情報から高次情報のデータ活用へ。感性データは高次情報

a ; 情報の階層構造

- ・甲が「AはBである」と言った ——— 一次情報
- ・甲はウソをついている ——— 二次情報
- ・甲がウソをついた理由 ——— 三次情報

・甲は肝心なときにウソをつくので信用できない —— 四次情報

・甲はムシの好かない奴だ —— 五次情報

b ; 今までは、一次情報のデータを沢山集めてデータベース化し、検索することにより高次情報は利用者の領域になっていた。

c ; 役に立つのは高次な情報であり、それは感性データである。

d ; 高次情報を如何にうまく組み上げてつくるか。それは個人ではなく、多くの人々の種々な判断を経て、社会的に形成されるものである。

e ; コンピュータやマルチメディアによって、情報をいろいろな方向から感性や直感なども取り込んで、高次情報を形成し利用できれば良い

④ データベースは人間の能力を高めるためにある

・一次から四次、五次までのデータベースがあったとしても、人間の思考に合わせて三次から一次、四次から五次へというようにどの方向へもダイナミックに対応するインタフェースをもつメカニズムが重要

⑤ 情報は刺激。問題意識がないと只の文字

・情報に注目したり、活用したりは、“我が意を得たり”があればこそで、それは人によって違う。
・その意味で一次情報は重要であり、編集・加工された二次以降の情報は誰かの色つきになるのが問題。
・どんな情報をどう使うかということがないと、情報を扱うのは困難

⑥ 学術文献情報は、よく使うが非一次データ

・論文の前の一次データは、統計・実験・観測・調査等のデータ
・これらの一次データの中に、失敗実験データ（これは重要）を含め、一次データ集をつくっても、論文データはやはり必要

⑦ 常識は高次の情報

・その形状はインフォーマルで、新しい事実により変化して行く。これが分野別にまとめられたら面白い

⑧ 使用目的に合致する組み方のデータでないと無意味

・料理は栄養素データでは不可。
・医者とは物質の化学構造式データでは配剤不可。

⑨ 医者とは患者のコミュニケーションギャップ

・医者とは患者の各々のもつ一次情報のレベルが違う。このギャップを埋めるのが二次三次情報であれば有益

- ・医者と患者でいうと、患者の一次情報をもとにシステムの方からのアプローチで一次情報を集め、有効な二次三次情報を引き出して適応するというようなのが望ましい。

⑩ 政策目的のためのデータベース

- ・一次情報としてのそのこと自体よりも、ある場合は人々がどう思っているのかの方が重要。

⑪ 情報利用における情報発生源の条件の重要性

- ・基盤技術促進センターの情報の利用状況。
 - a ; 物性から情報を引く人はいない。
 - b ; 誰がどういう装置でやったか。発生源に関心がある。
- ・マイクロメカ等の新技術の情報も、それが医学でどう使われるのか、何が期待できるのかに興味がある。
- ・データを創り出した人の動機とか目的が重要。

(4) 知的資源を利用した場合の個人意識変革や社会的変化の特長の認識

① 知的資源をクリエイトする風土。スケールの大きな問題の議論が必要。

- ・欧米の先進性とスケールの大きさ
- ・日本の考える情報を超えたところで知的資源をクリエイト
(例：宇宙居住を想定した総合研究)

② バーチャルリアリティのテーマ

- ・今は「臨場感」に偏っている。
- ・西洋音楽と東洋音楽の違い — 西洋人、東洋人の感性の違い

{	西洋音楽 — 楽譜あり、複雑なメロディの組み合わせ
}	東洋音楽 — 楽譜なし、単純なメロディの時差演奏
- ・「音」のもたらす情報の重要性。バーチャルリアリティのポイント
 - プラントの先端での故障を「音」で知る。
- ・音楽に対する感覚 — 地域間ギャップ、世代間ギャップ
 - ロックは洋の東西を問わず、若い世代に流行
 - 中年以上には騒音。ロックはリズム

③ 文化ギャップ

- ・国際通信
 - 高速通信回線とスーパーコンピュータだけでは解決しない。

・翻訳

言葉には“表面”と“裏の約束ごと”がある。

言葉の“表面”に忠実に訳すと，“真意”が伝わらない。

自動翻訳での課題。

・統計データ

「定義」以外にデータの背景にある国の社会状況の違いを考慮した評価が必要。

・記号

日本人は記号を多用。「ワッ!」「ギャー!」などの表現も多い。英米人はすべてを文字、言葉で表す。

・知的資源を面、音、記号というように、扱いにくい情報に拡大して行くと国際的な感覚の違いも問題。

・表現のルールの必要性

情報は、コミュニケーションによる相互理解が必要。独りよがりの表現の氾濫は日本文化の欠点になる恐れ。

④ 普遍的に感動を呼ぶ文学・絵画など

情報の形式、与え方による人間の受取り方

⑤ コンピュータグラフィック（CG）によるルノアールの絵

・CGの絵にルノアールの絵と同じ感銘は受けない。

・CGにより、ルノアールの技法で、自分の肖像画が描かれても意味がないか。

昔の貴族の気分？

・ルノアールが描いたということに価値がある。

・本物と偽物 — 偽物は無価値か？

○区別不明ならブランド信仰打破に有効。

産地ブランド故に、山梨の牛が松阪、神戸で肉になる。

○バイオは、人のつくっている遺伝子を大腸菌に作らせている。偽物であるが大いに役立っている。

・絵は本物でないと感動しない。CGによって絵をいろいろ変化させて、人間の感じ方がどう変わるかによって何故絵に感動するのかを究明すれば、コンピュータで感性を扱う上で参考になる。 — これからのCGの方向

⑥ 感性情報のポイントは“なじみ”と“意外性”

絵や音楽は繰り返し鑑賞することによる“なじみ”と、前と違った何かを感じることで感動を呼ぶ。

⑦ 音楽の演奏

人間の演奏に感動するのか、コンピュータが難曲を演奏しても駄目か。

- ・人が超人的なことをやるから感動がある。

飛行機が空を飛んでも当たり前

- ・ホールとテレビの音楽鑑賞の感銘の差

ホールに一緒にいるという環境と部屋でテレビを見ているという環境の差

⑧ 音楽のリズムと言葉のリズム — リズムとは？

- ・英語がうまくならないのは、日本語のリズムから英語のリズムになれないから

- ・シュプレヒコールのリズムで、学生運動の派閥が分かる。

感性的思考パターンにはリズムが関係

⑨ 教育により人間が果たした役をシステムの処理能力に頼る危惧

- ・人間の本質的活動に係わる問題

a ; 今の高校生以下のゲーム世代に知的創造力があるかどうか心配。

b ; コンピュータどっぷり型でない今の大学生でも、好奇心や人のやらないやり方で考えるという思考性がなくなっている。

- ・機械でやれることをやり、更にその上のことを人間があるという考え方が大切

23.6 成 果

知的資源の普及や利用については、総合的な政策や利用できる技術開発が必要となるばかりではなく、社会的な構造や個人の意識を大きく変える必要がある。そのため、以下の項目についてその方策や具体的な実現に向けた検討が望まれることが分かった。

(社会、個人からのニーズ)

- ① 個人のもつポテンシャルが知的資源の利用により大きく広がる。
- ② 知的資源を利用する専門家の尖鋭的な集団とそれを統率や管理する機能が発生してくる。
- ③ ボーダーレス化により文化や情報が融合化した知的資源が発生することで、価値観や行動が大きく変化する。

(技術)

- ① 技術の進歩により、平面的な情報利用から三次元、四次元の世界の擬似的な表現が可能となる。
- ② 単体の機能をもつ機器を複合結合して利用することにより、更に高度利用が可能となる。
- ③ 新技術の導入により生活空間や都市空間の変化が顕在化し、又は新技術と既存技術の相互

依存による活性化が図れる。

(関連産業等への影響)

- ① 知的資源とその利用システムは産業構造を変え、新規需要創造を形成するインフラストラクチャとなりうるものである。
- ② 知的資源に必要なデータは開発が必要であり、またユーザにあまねくこれらが利用できる機器が必要となる。
- ③ 知的資源を利用できる能力を涵養するため、現在始まりつつあるコンピュータリテラシーの成果が重要な鍵となる。

(予想される効果)

- ① 既存データベースの将来展望の検討に資することができる。
- ② ユーザのニーズが明確となり、技術開発や新製品開発の方向性を見い出せる。
- ③ 新しい情報需要の形成の一つとなりうる可能性が高い。

23.7 今後の課題

- ① 知的資源の概念の形成
- ② 知的資源利用に必要とされるユーザのあり方
- ③ 知的資源の需要の態様とデータ開発の展望
- ④ 望まれる知的資源提供・利用のためのデータベース、システムのあり方
- ⑤ 市場形成に必要な要素技術の検討及び課題

禁 無 断 転 載

平 成 5 年 3 月 発 行

発 行 財団法人 データベース振興センター

東京都港区浜松町二丁目4番1号

世界貿易センタービル7階

TBL 03-3459-8581

印刷所 株式会社 正文社

〒113 文京区本郷3丁目12番2号

TBL 3815-7271~3

