

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)

昭和63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1/5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 東京工科大学教授 西野博二氏）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）」は昭和62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、昭和62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。

昭和63年3月

財団法人データベース振興センター

昭和62年度 データベース構築促進・技術開発委託課題

分野	課題名	会社名
社 会	1 産業技術のニュース・リリースのデータベースフィージビリティ調査	(株)エス・アール・シー
	2 保安情報データベース構築に関する調査研究	高圧ガス保安協会
	3 健康増進データベース構築	鹿児島県ニューメディア振興財団設立準備委員会
	4 法令・判例等ソースの構築(法律エキスパートシステムに応用)	(株)日本法律情報センター
	5 ハンディキャップパーソンの情報ニーズに即したライフサポートデータベースの設計	ダイヤル・サービス(株)
地域活性化中小企業振興	6 中国地域における技術情報データベース構築に関する研究調査	(財)中国地域技術振興センター
	7 公設試験研究機関データベースネットワーク構築調査	関西データベース協議会
	8 東北型地域産業情報データベース・モデルシステムの拡張	(社)東北機械電子工業会
	9 郡山地域商工業活動データベース構築調査	ふくしまテレメディアサービス(株)
産 業 別 拠 点	10 世界鉄鋼企業財務データベース	(社)日本鉄鋼連盟
	11 データベース構築を前提とした印刷情報の展開に関する調査研究	(社)日本印刷産業連合会
	12 鉄鋼関連流通データベース構築	(社)鋼材倶楽部
地 図	13 地図データベースを利用した商圈分析システムの開発	アジア航測(株)
	14 地域分析・計画のための地図データベースの構築	(株)日本総合技術研究所
素 材	15 実用化合物データベースのCD-ROM用実験システムの開発	日本電子計算(株)
	16 SiC系ファインセラミックスの成分特許データベースの構築	神鋼リサーチ(株)
	17 化審法既存化学物質リストのデータベースの構築	(社)日本化学物質安全・情報センター
エネルギー・資源	18 地熱資源利用基本データベースの構築調査	(財)新エネルギー財団
	19 省エネルギー技術に関するデータベース構築の調査研究	(財)省エネルギーセンター
部 品 ・ 材 料	20 電子部品データベース実験システムの構築	三菱商事(株)
標 準 化	21 漢字データベース構築	(株)理想社印刷所
海 外	22 機械翻訳システム用辞書データベースの調査	(株)サイコム カテナ(株)
	23 国内外に対応する建築文献データベースの構築に関する調査	(財)日本建築センター
	24 翻訳支援ツール(英・和)としての専門用語辞書データベースの構築	日外アソシエーツ(株)
利 用 技 術	25 集合論に基づく知的情報検索システムの開発	シャープ(株)
	26 エキスパートシステム技術を用いたデータベース知的インタフェースの試作	(株)シー・エス・ケイ総合研究所
	27 イメージ・データのコード・データ変換に関する調査研究	(株)エービーシ
そ の 他	28 国際ビジネスマン情報データベースの構築	(財)国際ビジネスコミュニケーション協会
	29 建設関連技術用語データベース構築の調査研究	(株)建設情報センター

目 次

1. 産業技術のニュース・リリースのデータベースフィージビリティ調査	1
1.1 調査の目的・方法	1
1.2 ニュース・リリースの現状	1
1.3 収集メカニズムの検討	3
1.4 データの入力形態の検討	4
1.5 技術分類体系の検討	5
1.6 検索項目の検討	6
1.7 今後の課題と展望	7
2. 保安情報データベース構築に関する調査研究	9
2.1 目 的	9
2.2 調査検討事項	9
2.3 保安情報データベースの必要性	11
2.4 既存データベースの利用の可能性	11
2.5 基本方針	12
2.6 保安情報データベースの構築の基本的な考え方	13
2.6.1 システムの基本概念	13
2.6.2 使用目的	13
2.6.3 利用方法	13
2.6.4 保守、運営	14
2.7 今後の検討課題	14
2.8 まとめ	15
3. 健康増進データベース構築	17
3.1 構築の目的	17
3.2 データベースの内容	17
3.3 データベースの基本構造	18
3.4 課 題	21

6.1	目的	37
6.2	成果の概要	38
6.2.1	中国地域におけるデータベース	38
6.2.2	データベースの構造研究	39
6.2.3	データベース構築のための今後の検討課題	40
7.	公設試験研究機関データベースネットワーク構築調査	45
7.1	調査の目的	45
7.2	公設試の情報化の現状	45
7.3	公設試における情報化ニーズ	48
7.4	データベースネットワーク構築の必要性と実現に向けての課題	49
7.5	データベースネットワーク構築ビジョンとステージプラン	51
7.6	当面とるべき対策	52
8.	東北型地域産業情報データベース・モデルシステムの拡張	55
8.1	モデル・データベース（東北地域企業情報データベース：TRIP） 内容の拡充と情報検索システムの開発	55
8.1.1	モデル・データベース構築の目的と経緯	55
8.1.2	データ拡充とシステム拡張の内容	56
8.2	モデル・データベースのシステム運用方策の検討	57
9.	郡山地域商工業活動データベース構築調査	59
9.1	データベース構築の意義	59
9.1.1	データベース構築の意義	59
9.1.2	ビデオテックスシステムによるデータベースサービスの特質	60
9.2	データベース構築の全体計画（郡山地域における商工業活動データベース化）	61
9.3	情報産業の現状と課題	64
9.3.1	企業における情報化の現状	64
9.3.2	企業における情報への対応	65
9.4	F T S ネットワーク機構の現況と展望	66
9.4.1	F T S の会社概況と設立経過	66

9.4.2	端末設置状況と事業展開	66
9.4.3	チャンネルメニューと検索システム	67
9.4.4	F T S の配信システム	67
9.4.5	F T S のネットワーク構想	68
9.5	調査用モデル画面の構築(モデル画面の制作に当たって)	69
1 0.	世界鉄鋼企業財務データベース	71
1 0.1	目的・背景	71
1 0.2	実施内容	71
1 0.2.1	データの作成	72
1 0.2.2	情報処理の内容	72
1 0.2.3	データベースの利用	73
1 0.3	今後の課題と展望	75
1 0.3.1	データの追加・補充	75
1 0.3.2	データ利用の体制整備	76
1 1.	データベース構築を前提とした印刷情報の展開に関する調査研究	77
1 1.1	調査研究の経緯	77
1 1.1.1	調査研究の背景と目的	77
1 1.1.2	調査研究に当たっての前提条件	78
1 1.1.3	調査研究内容	78
1 1.2	印刷情報の特色と汎用データベースの位置付け	78
1 1.2.1	印刷情報の特色	78
1 1.2.2	汎用データベースの構成	79
1 1.3	モデルタイプ・システムの基本設計	80
1 1.3.1	システムの要件および設計方針	80
1 1.3.2	印刷情報入力機能	80
1 1.3.3	データベース管理・運用支援機能	81
1 1.3.4	メディア展開・出力機能	81
1 1.3.5	データベースの構造と既存文章モデルとの関連	81
1 1.4	机上シミュレーション	82

1 1.4.1	シミュレーション方法	82
1 1.4.2	内部構造の検討	82
1 1.4.3	入力・登録処理におけるシミュレーション	83
1 1.4.4	メディア展開・出力処理におけるシミュレーション	83
1 1.5	汎用データベースに対する評価と今後の課題	84
1 1.5.1	評価	84
1 1.5.2	今後の課題	84
1 1.6	結 び	85
1 1.6.1	法制面への対応	85
1 1.6.2	ビジネス化の課題	85
1 1.6.3	今後の方向	86
1 2.	鉄鋼関連流通データベース構築	87
1 2.1	目 的	87
1 2.2	実施体制	87
1 2.3	実施内容	88
1 2.4	ま と め	90
1 3.	地図データベースを利用した商圈分析システムの開発	91
1 3.1	背景と目的	91
1 3.2	商圈分析データベースの概念	92
1 3.3	課 題	95
1 4.	地域分析・計画のための地図データベースの構築	97
1 4.1	構築の背景と目的	97
1 4.1.1	背 景	97
1 4.1.2	開発経緯	97
1 4.1.3	開発の概要	98
1 4.1.4	システムの基本的視点	98
1 4.2	システムの概要	99
1 4.2.1	システム設計上の考え方	99

1 4.2.2	システム構成の考え方	99
1 4.2.3	システムの開発計画	100
1 4.3	システムの内容	100
1 4.3.1	対象データ	100
1 4.3.2	ファイル構成	101
1 4.3.3	システム構成	102
1 4.3.4	ハードウェア構成	104
1 4.4	システムの利用例	104
1 4.5	システムの評価と今後の課題	106
1 4.5.1	システムの評価	106
1 4.5.2	今後の課題	106
1 5.	実用化合物データベースのCD-ROM用実験システムの開発	107
1 5.1	実用化合物データベースの全体像	107
1 5.2	CD-ROM用実験システムの開発	107
1 5.2.1	CD-ROM用実験システムの概要	107
1 5.2.2	システムの構成	109
1 5.2.3	データの構成	110
1 5.2.4	CD-ROM検索システム	111
1 5.2.5	実験システムの評価	114
1 5.3	今後の展望	114
1 6.	SiC系ファインセラミックスの成分特許データベースの構築	117
1 6.1	目的	117
1 6.2	概要	117
1 6.3	実施内容	118
1 6.3.1	情報収集	118
1 6.3.2	データ作成	119
1 6.3.3	データベース構築	120
1 6.4	検索例	123
1 6.4.1	一般検索例	123

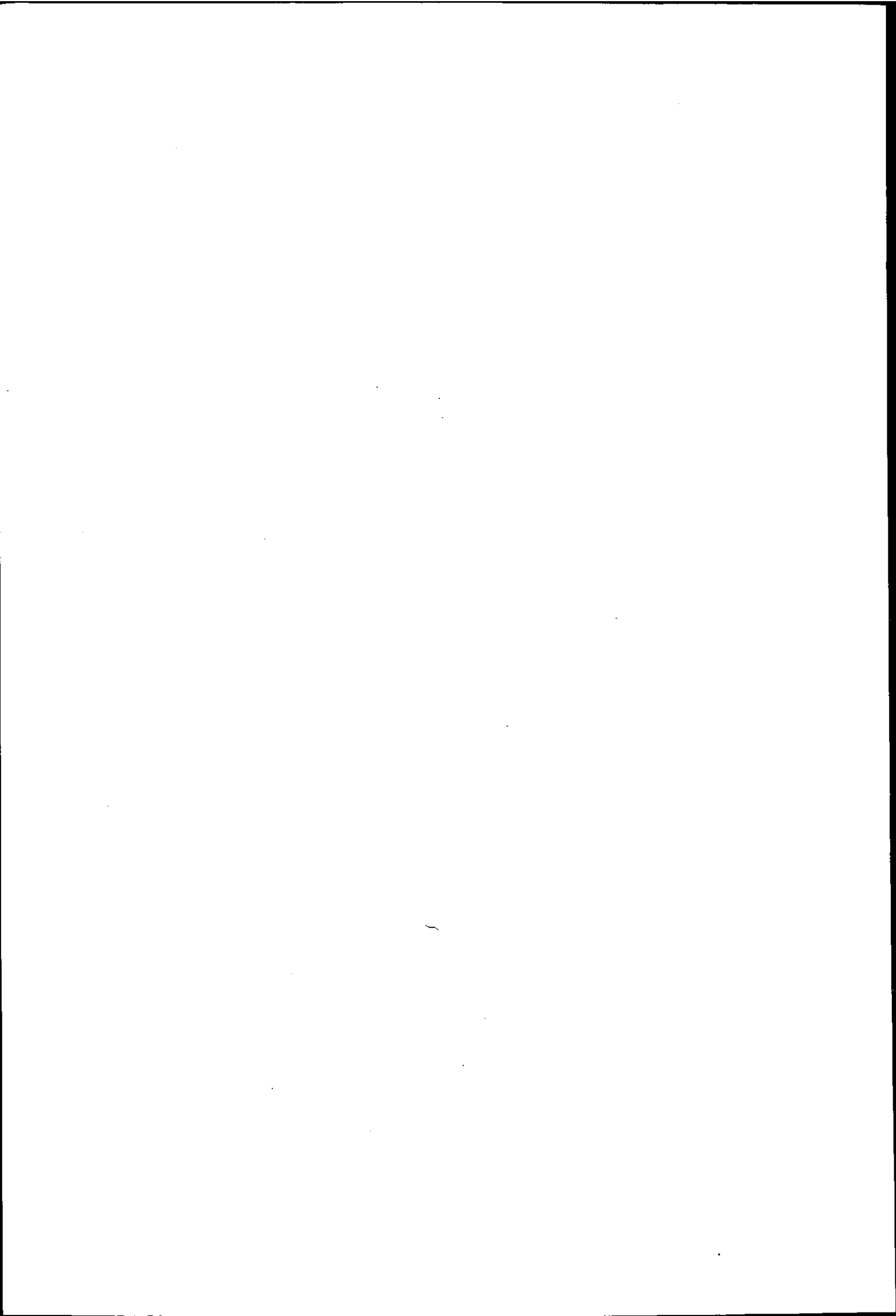
1 6. 4. 2	化合物検索例	123
1 6. 5	本データベース利用の効果	124
1 6. 6	今後の課題と展望	125
1 7.	化審法既存化学物質リストのデータベースの構築	127
1 7. 1	構築の目的と背景	127
1 7. 2	予想される効果	128
1 7. 3	本データベース構築の基本方針	128
1 7. 4	構築経過	129
1 7. 5	今後の課題と展望	129
1 8.	地熱資源利用基本データベースの構築調査	135
1 8. 1	調査の目的	135
1 8. 2	調査の方法	135
1 8. 3	調査の成果	136
1 8. 4	課題と展望	138
1 9.	省エネルギー技術に関するデータベース構築の調査研究	141
1 9. 1	背景	141
1 9. 2	省エネルギーセンターの情報サービス活動の問題点	141
1 9. 3	省エネルギー技術データベース構想	142
1 9. 3. 1	省エネルギーセンターの問題点の解決策	142
1 9. 3. 2	省エネルギー技術データベース構築の基本的考え方	143
1 9. 4	省エネルギー技術データベース構築案	144
1 9. 4. 1	登録対象情報	144
1 9. 4. 2	省エネルギー技術データベースシステムの展開	144
1 9. 4. 3	情報登録用紙のフォーマットとインデクシング	148
1 9. 5	省エネルギー技術データベースシステムの運用体制	149
1 9. 6	省エネルギー技術データベース構築の予想効果	150
1 9. 7	今後の課題	151
1 9. 7. 1	情報の質の向上	151

1 9. 7. 2	国 際 化	151
1 9. 7. 3	情報関連技術の進歩	151
2 0.	電子部品データベース実験システムの構築	153
2 0. 1	目 的	153
2 0. 2	実験システム構築体制	154
2 0. 3	実験システムの概要	155
2 0. 3. 1	提供情報の範囲	155
2 0. 3. 2	提供情報の内容	155
2 0. 3. 3	提供情報数	156
2 0. 3. 4	対象ユーザ	156
2 0. 4	実験システム構築・評価フロー	157
2 0. 5	実験システムの評価	158
2 0. 5. 1	評価のポイント	158
2 0. 5. 2	評価の結果（ユーザ20社の意見の集計および分析）	158
2 0. 5. 3	今後の課題	159
2 1.	漢字データベース構築	161
2 1. 1	概 説	161
2 1. 2	作業方針	161
2 1. 3	データ整備作業	161
2 1. 3. 1	収録文字種の選定	161
2 1. 3. 2	文字フォントのデジタル化	162
2 1. 3. 3	その他関連する各種情報の入力	162
2 1. 4	校閲作業	164
2 1. 5	検索システム研究	164
2 1. 5. 1	検索機能の検証	164
2 1. 5. 2	模擬システムの評化	166
2 1. 6	今後の開発計画	167
2 2.	機械翻訳システム用辞書データベースの調査	169

2 2.1	調査の目的	169
2 2.2	調査の内容	169
2 2.2.1	機械翻訳システムにおける辞書の分類	169
2 2.2.2	システム標準辞書の分類	172
2 2.2.3	専門辞書, ユーザ辞書の分野	173
2 2.3	今後の課題	174
2 3.	国内外に対応する建築文献データベースの構築に関する調査	175
2 3.1	目的	175
2 3.2	実施内容・方法	176
2 3.2.1	調査の内容	176
2 3.2.2	調査の実施体制	176
2 3.2.3	実施期間	177
2 3.2.4	予想される波及効果	177
2 3.3	成果と今後の課題	177
2 3.3.1	調査の成果	177
2 3.3.2	今後の検討課題	178
2 4.	翻訳支援ツール(英・和)としての専門用語辞書データベースの構築	179
2 4.1	開発の目的	179
2 4.2	開発の概要	179
2 4.3	用語の選定	180
2 4.4	英和辞書型の原稿作成とデータファイル作成	181
2 4.5	英和・和英型データベースソースファイルの作成	183
2 4.6	開発の成果と今後の課題	185
2 5.	集合論に基づく知的情報検索システムの開発	187
2 5.1	目的と内容	187
2 5.1.1	目的と予想される効果	187
2 5.1.2	内容と開発体制	188
2 5.2	コンピュータアシスト文書検索システム	189

2 5. 2. 1	システム構成	189
2 5. 2. 2	対象業務	190
2 5. 2. 3	導入効果	191
2 5. 3	データベース管理システム	192
2 5. 3. 1	入力	192
2 5. 3. 2	表示	192
2 5. 3. 3	操作	192
2 5. 3. 4	制御	193
2 5. 3. 5	出力	194
2 5. 4	今後の展開	194
2 5. 4. 1	集合論データベースシステム	194
2 5. 4. 2	文書検索システム	194
2 5. 4. 3	マンマシンインタフェース	194
2 6.	エキスパートシステム技術を用いたデータベース知的インタフェースの試作	197
2 6. 1	目的	197
2 6. 2	内容	197
2 6. 3	今後の課題	200
2 7.	イメージ・データのコード・データ変換に関する調査研究	203
2 7. 1	はじめに	203
2 7. 2	調査研究の範囲	204
2 7. 3	解決すべき技術的課題	204
2 7. 4	イメージ情報からの文字の抽出	206
2 7. 5	イメージ情報からの文字情報の抽出における問題点	207
2 7. 6	まとめ	208
2 8.	国際ビジネスマン情報データベースの構築	209
2 8. 1	概要	209
2 8. 1. 1	構築の背景	209
2 8. 1. 2	これまでの取り組み	209

2 8.2	実施体制	210
2 8.3	検討内容	210
2 8.3.1	チェック項目	210
2 8.3.2	「ISBISフォーム」の作成	213
2 8.3.3	「ISBISフォーム」から質問表へのワーディング	213
2 8.4	パイロットスタディ	214
2 8.4.1	パイロットスタディ実施の背景	214
2 8.4.2	パイロットスタディの実施と分析	214
2 8.5	今後の課題	215
2 9.	建設関連技術用語データベース構築の調査研究	217
2 9.1	目的	217
2 9.2	実施内容と方法	217
2 9.2.1	調査の内容	217
2 9.2.2	調査の実施体制	218
2 9.2.3	実施期間	218
2 9.2.4	予想される波及効果	218
2 9.3	用語集および技術用語集の現状分析のまとめ	219
2 9.4	成果と今後の課題	220
2 9.4.1	主要成果	220
2 9.4.2	今後の検討課題	220



1 産業技術のニュース・リリースのデータベース
フィージビリティ調査

株式会社 エス・アール・シー



1. 産業技術のニュース・リリースのデータベースフィージビリティ調査

1.1 調査の目的・方法

(1) 調査目的

現在、産業技術に関するデータベースで、即時性、トピック性という点で優れているのは、新聞の記事情報データベースであると思われる。しかしながら、新聞の記事情報はニュース・リリースの要約であり、かつニュース・リリース企業への連絡先が不明であるという欠点を持っている。また、記事情報はすべてのニュース・リリースを掲載しているわけではなく、編集者の価値判断による選択がなされた情報であるという特質を持っている。

利用者はこれらの特質を知りつつも利用しているが、利用者ニーズとしては、要約ではなく本文であり、かつ編集者の価値判断が入っていない生の情報を望んでいると思われる。

本調査はこれらの観点から利用者ニーズの高い産業技術のニュース・リリースのデータベース構築のため、その構築可能性を検討する事を目的とする。

(2) 調査内容と方法

産業技術のニュース・リリース・データベース構築は3カ年を計画している。

本年度は産業技術のニュース・リリースのデータベース化を検討するため、ニュース・リリースの収集メカニズム、技術分類コード、検索項目を研究し、そのフィージビリティを調査する。

1.2 ニュース・リリースの現状

(1) 新聞社によるニュース・リリースの利用実態

現在、ニュース・リリースの受信量が一番多いのは経団連記者クラブであり、多い日で20件/日、少ない日でも3件/日程度である。

経団連記者クラブのニュース・リリース申込方法は、発信者が48時間前にクラブの幹事会社に配布・申込みを行い、幹事会社はそれを受け、クラブのノートにメモし、黒板に予定表を記入する。企業はソース専用の箱にニュース・リリースを入れていく。

ニュース・リリースの提供の方法は3種類あり、一つは一般公開のレクチャー付きのもの

の、他は対象を限定したレクチャー付きのもの、ニュース・リリースのみのものである。

受信したニュース・リリースは不足部分と多過ぎる部分を編集者が検討し、不足している部分については取材をして追加する場合もある。特に、ベンチャー企業、中小企業でPR会社に依頼して作成したものはファクトが不足している事があり、報道側の取材が必要な事が多い。近年、技術物が多くなっているためニュース・リリースのみで価値判断をする事は難しい状況にあり、理解し易い物は大きく取扱い、理解しにくい物はボツになる事もある。また、価格表示のないものはボツになり易い。

記者クラブは、関係者以外立入禁止であり、クラブ員以外は直接企業の広報に行かなければニュース・リリースを手に入れる事はできない。記者クラブからニュース・リリースを持出すのは信義上の問題が生じると思われるが、企業から直接もらう分には問題は生じない。ただし、リークして先にどこかの記事になったものは報道しない。

最近の動きとして、電子媒体を利用してニュース・リリースを提供する実験がある事と、英文のニュース・リリースに対するニーズが非常に高くなっている。

経団連記者クラブ以外の主な記者クラブを以下に示す。

- ・ 重工クラブ（化学と鉄）
- ・ エネルギー研究会（電力・エネルギー）
- ・ 繊維クラブ（繊維）
- ・ 農林クラブ（食品）
- ・ その他業種別（運輸、建設、銀行、商社）
- ・ その他地方（大阪・名古屋のクラブ、県政クラブ等）

(2) 民間企業からみたニュース・リリースの特質

ニュース・リリースを活用する場合の前提として、ニュース・リリースは純技術的なニーズに基づいて発信している情報でないという事に注意しなければならない。

企業が発信する情報は、発表すべきタイミング、ライバル企業との状況を考慮して、ニュースバリューがあると判断されたものである。したがって、必ずしも純技術的なニーズと一致しているとは限らない。この事を念頭に置いて情報を活用する必要がある。

ニュース・リリースの本質は企業を知ってもらうためのものであり、技術情報を提供するためのものではないので過度の期待を持つべきでない。

(3) アンケート結果からみたデータベース構築の可能性

本アンケートの対象企業は一部上場企業でかつ広報担当部署があるという事で、ニュー

ス・リリースに対する意識が高い企業である。

したがって、情報提供者の母集団としては適当であると思われる。しかし、利用者という観点からみると、片寄った意見となっている。

① ニュース・リリースの情報量

ニュース・リリースの提供頻度を1社平均3回/月と考え、1,000社から情報収集を行うと3,000件/月の情報量となり、1日当たり100件のデータ量となる。

なお、情報内容を産業技術情報に特定化すると情報量は更に減少する。

② ニュース・リリースの収集方法

現在約50%の企業は原稿をワープロ入力し、フロッピーに登録しているが、その提供については消極的である。ただし、原稿そのものの提供については積極的であり、十分提供を受けられると思われる。

③ 固定フォームによる記載

固定フォームによる記載についてはかなり積極的であり、適切なフォームを設計する事によってかなりの企業で記載してもらえ事が予想される。ただし、要約については、原稿に線を入れてもらう等の方法を取る事も考えられる。

④ 検討課題

- ④ 既存のニュース・リリース提供システムを上回るサービス内容の検討
- ④ 検索項目としての企業名、地域等の導入検討
- ④ 情報の信頼度、裏付け等の必要性
- ④ 新聞に掲載されたデータ、されなかったデータとの仕分け

1.3 収集メカニズムの検討

(1) 対象機関

データ収集の対象機関は記者クラブと民間企業の広報担当の2者択一である。

記者クラブについては、業種が片寄る不安がある事、報道以外には本来活用できない事等の問題があり、理論的に不可能と思われる。また、無理に収集すると法律的問題が生じる恐れもある。それに対し、民間企業の広報担当の場合は、郵送先が1社増えるという労力負担で済む事から、データベースの有効性を打出せばかなりの協力が得られるものと思われる。また、広報担当からの入手は専門紙発表のデータも取れるという利点がある。

以上の点からデータ収集の対象機関は民間企業の広報担当が現実的である。対象業種は当初絞った形で進め、順次拡大していく方法が考えられる。

(2) 対象情報の分類

収集するニュース・リリースを産業技術情報とそれ以外の情報に仕分けする必要がある。提供サイドの労力を考えると仕分けをせずにすべてのデータを送付する事が最も簡単と思われるので、情報の分類は運営サイドとする事となる。ニュース・リリースの種類は以下の分類が考えられる。

- | | |
|------------|-----------------|
| ① 新製品のPR | ④ 技術提携等他社との関係 |
| ② 新技術開発の発表 | ⑤ 人事・決算等の自社内の情報 |
| ③ 新規事業の発表 | ⑥ その他 |

(3) データ収集の形態

アンケート調査の結果から判断すると、ニュース・リリース原稿での収集が最も可能性が高い。ワープロメディアについては、メディア化していない企業もあり、かつ送付システムを作る必要も生じるため、現在の時点では難しい。

また、固定フォーム用紙への記入は、ニュース・リリース提供頻度の高い企業では、そのような労力負担を避ける傾向にあるため、何らかの優遇措置を講じなければ難しいと思われる。したがって、固定フォームへの記入は運営者サイドで実施する事となる。

1.4 データの入力形態の検討

民間企業より入手したニュース・リリースを基に検索項目を考慮しながら固定フォームを設計した。文章については端末画面の一行当たり文字制限を考慮して、一行30文字とした。また、実際の入力機器はワードプロセッサを想定している。図面データの取扱いはイメージリーダーの利用が考えられる。

(1) 表 題

最大全角30文字(半角60文字)以内でニュース・リリースのタイトルを記入する。検索時にディスプレイに表示される。

(2) 副 題

最大全角30文字(半角60文字)以内で表題を補足する内容を記入する。表題メニュ

ーの表示には示されないが要約には表示される。

(3) 要 約

最大全角 300 文字 (30 文字 × 10 行) 以内で示される本文の要約を記入する。

(4) 本 文

最大全角 1,350 文字 (30 文字 × 15 行 × 3 画面) 以内で示されるニュース・リリースの本文を記入する。

(5) 登録日付

ニュース・リリースの提供年月日を記入し、日付けによる検索 (○年○月○日以後の情報) に利用する。

(6) 分類メニュー

別途検索する技術分類メニューのメニュー番号を記入し、技術分類による検索に利用する。

(7) 金 額

新製品の価格を記入し、価格による検索 (○万円以上) に利用する。

(8) キーワード

最大 16 個のキーワードを記入し、キーワード検索に利用する。

(9) 問い合わせ先

ニュース・リリースの内容に関する質問を可能にするため、住所、電話番号、ファクシミリ番号、担当部署、担当者を記入する。

(10) 有効期限

ニュース・リリース登録の期限年月日を記入し、登録データの自動削除に利用する。

1.5 技術分類体系の検討

ニュース・リリースのデータベースに登録される情報は、新製品、新技術等の情報を中心として、常に最先端の技術に関するものが多く、分野的にも全産業技術に至る幅広いものが期待される。また、この登録された技術や新製品の利用は日本のみならず、世界的な広がりを持つことが予想される。

このため、産業技術情報を分類するに当たり、国際的に通用し、かつ最新の技術を十分に取り込むことができるような広範な枠組みとして、国際特許分類の最新版 (第 4 版 Int.Cl⁴)

を参考にした体系を作成した。

この技術分類体系は、全体を9つの大項目（生物科学系，化学系，物理学系，機械工学・熱機械・化学工学，鉱業・金属，制御工学・物理工学・情報工学，電気・電子，土木・建設，その他）に分け，それぞれを9項目以内の中分類（合計64項目）と，各中分類を更に9項目以内の小分類（合計388項目）へと分類した3階層体系となっている。

分類表のタイトル，内容説明については次の点に留意する必要がある。

技術の主題は「ものの開発」，「ものの使用法」あるいは「ものの応用の仕方」に関するものであるが，この「もの」というのは『方法』，『生産物』，『装置』のように有形，無形を問わず，すべての技術事項を意味して使用している。

この分類表において，分類項目のタイトルや説明にはこの『方法』，『生産物』および『装置』が混在して用いられているが，その中にはそれぞれの用語が厳密に意味する技術内容だけでなく，その関連事項も技術的内容として含まれている。例えば，タイトル用語や内容の説明が『装置』と表現されている場合でも，そこには『方法』も含み得るし，逆に『方法』と表現されていても『装置』も含まれる。また，『生産物』と表現されている場合でも，その生産『方法』，生産『装置』も含まれる。

1.6 検索項目の検討

検索項目は入力データで設定した下記の項目について検討する。

- ① 産業技術分類検索
- ② キーワード検索
- ③ 企業名検索
- ④ 金額検索
- ⑤ 登録日検索

(1) 産業技術分類検索

「1.5 技術分類体系の検討」で述べた産業技術分類体系に従い，メニュー方式による産業技術分野別のトリー検索を行う。対象とする産業技術を大分類，中分類，小分類の順で絞り込んで行き，欲しいデータを探す事ができる。

(2) キーワード検索

入力データで設定したキーワードによる検索を行う。検索方法は利用者がキーワードを入力するのでなく、登録されたキーワード一覧表の画面よりキーワードを選択する方式を採る。これによって、端末操作の簡便化が図れるとともに、検索結果が0件という事が生じなくなる。

キーワード一覧表は五十音順にソートされ、かつ漢字表示も含む事によって利用効率が高くなる。

(3) 企業名検索

キーワード検索と同様に登録企業名一覧表の表示による企業選択方式を用い、得ようとする企業の情報を検索する。

この検索項目はライバル企業の新製品発表の動向、新技術発表の動向を調べるのに非常に有効であると思われる。

(4) 金額検索

対象とする新製品を検索する時、〇〇円以下、〇〇円以上、〇〇円～〇〇円の間等の利用方法が可能であり、例えば一般の人がワープロの新製品を価格設定をして検索したい時に有効である。

(5) 登録日検索

対象とする新製品の中から6カ月前から現在までに発表されたものをみたい、あるいは特定業種の1年間の技術動向をみたいという時にこの機能が有効である。

指定方法は、〇〇年〇〇月以後あるいは〇〇年〇〇月～〇〇年〇〇月の間、〇〇年〇〇月以前等の方法が考えられる。

1.7 今後の課題と展望

(1) ニーズ調査の必要性

本調査はニュース・リリースの構築可能性を探るため、ニュース・リリースの収集という観点から検討を行った。そのため、調査対象機関も民間企業の広報室、記者クラブ等、実際にニュース・リリースに関係している機関のみであった。今後、データベース構築の事業性という観点から、利用者となる一般企業の経営者あるいは技術担当を対象とした利用意向、利用形態等を検討する必要がある。

(2) ニュース・リリースの信頼性

登録したニュース・リリースに誤りが発見される、あるいは訂正が発生する等の事態に対する対応を検討する必要がある。新聞記事の場合は一度編集者のフィルターを通過しているため、一定の信頼性が保証されている。本システムにおいても運営上何らかのチェック機能を備える事が考えられる。例えば、新聞記事になった情報とならなかった情報との分類等利用者にとっては一つの判断材料となる。

(3) 入力方式の簡便化

収集したニュース・リリースを固定フォームに変更するためには、かなりのマンパワーを必要とし、システム運営上の負担となる事が予想される。そのため、データ提供者に優遇措置（検索料の割引等）を与える事によって提供者記入を進める方法が考えられる。また、メディア化している企業もかなりあるため、今後の情報通信の発達に併せて電子メールの活用等の導入を検討する必要がある。

(4) システムサービスの充実

既存の新聞記事データベースにはない特徴を強調する必要があり、システムサービスとしてデータベースの有効利用案内、付加サービスとして企業概要サービス等を検討し、ニュース・リリースのみが持つ特殊性を前面に打出すシステムサービスを充実させる。

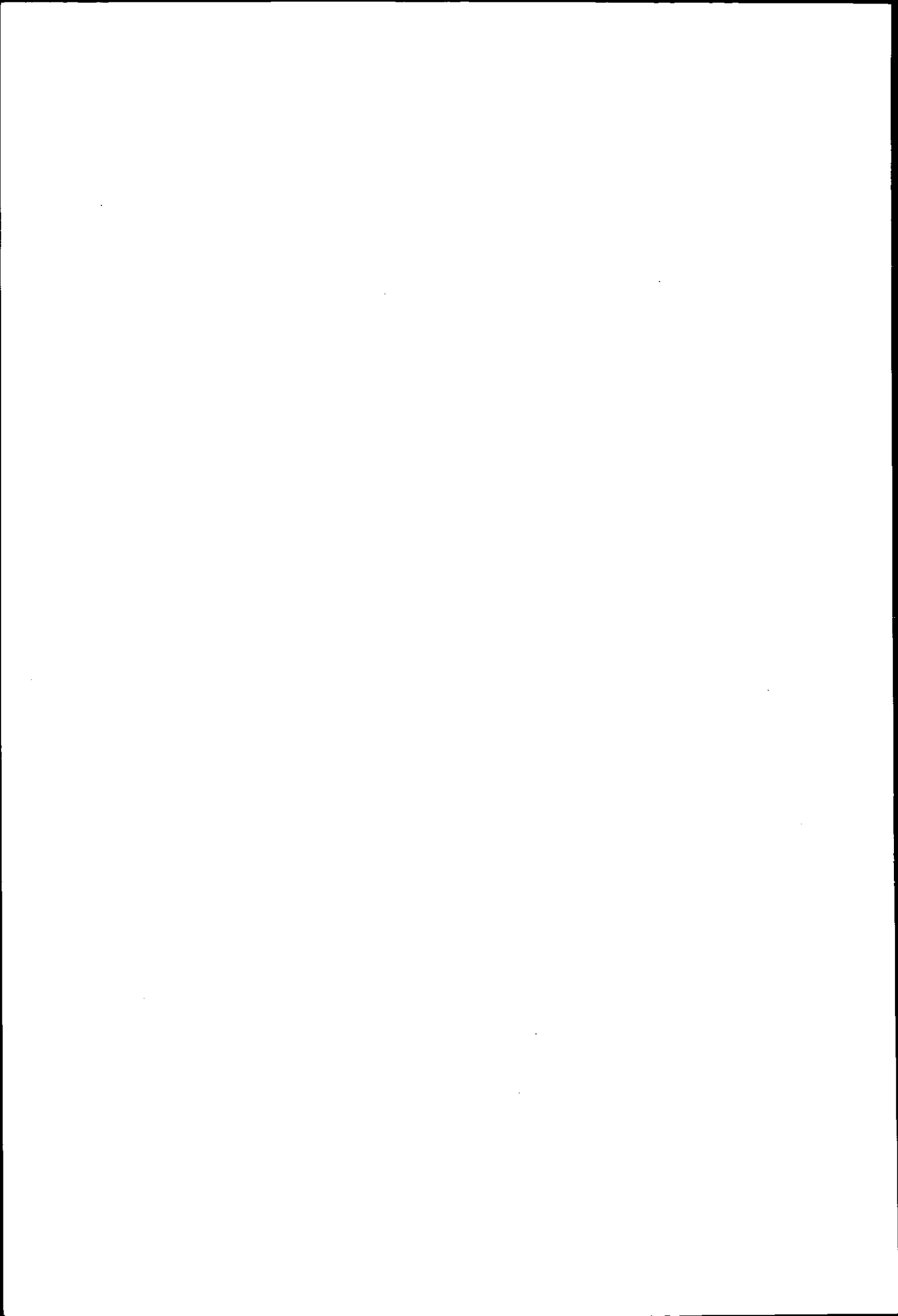
(5) 既存コンピュータソースの有効利用

現在、各種公益法人において汎用コンピュータの導入が進んでいる。しかし、これらのコンピュータは必ずしもフル稼働しているとは限らない。本データベースを構築するためにはこれらのコンピュータソースを有効利用した構築費用の負担軽減を検討する事が重要である。そのためには、特にニュース・リリースを公益的な活用のできるシステム内容にする事が必要である。

2 保安情報データベース構築

に関する調査研究

高圧ガス保安協会



2 保安情報データベース構築に関する調査研究

2.1 目的

現在わが国で利用出来る商用データベースのうち、保安情報に関するデータベースは目的および趣旨が様々であり、高圧ガス関連の保安情報データベースの目的には必ずしも合致していない。

高圧ガス保安協会が長年にわたり蓄積し、所有している各種情報は質的にも重要かつ有益なものが多いので、これをデータベース化し、関係方面の利用に供することは有意義である。

協会が蓄積している各種情報は、大別すると法規、技術基準、事故、検査データ、技術文献および図書等であり、これらのうち利用度の高いと考えられるものは、法令、物性、技術および事故の情報である。

高圧ガス保安協会では高圧ガス取締法等高圧ガスに関係する内外の法令、高圧ガスに関連する物質の物性データ、高圧ガスの保安に係る研究開発情報および高圧ガスに係る事故情報についてデータベースを構築する計画である。

このデータベースを構築するに当たって内外の既存データベースのシステムの内容を適切に把握し、データベース構築の方向、手順、データベースの内容、データの提供方法等について調査検討を行い、保安情報データベース構築に資することを目的とする。

2.2 調査検討事項

今年度は保安情報データベースに関連する内外の既存データベース、保安情報データベースに対するニーズ、各データベースに盛込むべき内容および利用形態について調査、検討を行った。

(1) 保安情報データベースに関連する既存データベースの調査検討

① 法令データベース

高圧ガス取締法をデータベース化するに当たり、通商産業省、総務庁、日本化学物質安全・情報センター等で保有している法令検索システムの入力データ、検索機能、出力

形式の調査を行った。

② 物性データベース

高圧ガス取締法に関連する物質約 100 数十種について、日本科学技術情報センターの熱物性データベース、DIPPR (Design institute for Physical Property Data) OHMTADS (Oil and Hazardous Materials Technical Assistance Data System)、日本化学物質安全・情報センターの安全性データベース等に入力されている物質の種類、物性値および出力形式について調査した。

③ 技術データベース

協会が保有すべき文献情報に関する日本科学技術情報センター (JICST) の JOIS データベース、CAS 等のデータの入出力形式および内容を調査した。

④ 事故データベース

協会が保有すべき高圧ガス取締法、液化石油ガス法等の事故に関連する事故データベースとして住化情報センターが保有する安全情報を始めとして商用データベースの入出力形式および内容の調査を行った。

(2) 保安情報データベースのニーズ (使われ方) の調査

保安情報データベースは

- ① 法令データベース
- ② 物性データベース
- ③ 技術データベース
- ④ 事故データベース

から構成される予定であるが、各データベースを構築するに当たり、本データベースに対する協会の内部および会員等外部のニーズを調査した。

保安情報データベースの 4 種のデータベースに求められるものはそれぞれにその内容や要求の度合いも異なっている。

それは、本データベースのユーザになる者の所属する分野 (業界) によって自ずと異なってくるものと考えられる。しかし、総じて求められる度合いや内容は大方決まって来る。つまり、いずれの分野 (業界) の者であっても高圧ガス取締法令の内容および事件事例については相当関心が高い。しかし、物性および技術文献については、その関心の度合いもまちまちである。

(3) 保安情報データベースに盛り込むべき内容

保安情報データベースに対するニーズと深くかかわりを持つデータベースの内容について

ては本データベースの使われる場面をニーズ調査等により想定し、データベース毎に検討を行った。

2.3 保安情報データベースの必要性

協会の日常業務および外部からの問合せ、回答等をベースにデータベースの必要性を検討すると、従来は専門的知識を持った人や個人的に蓄積した独自の資料等をもとに回答処理をしている現状から、データベースがあれば相当有効に役立つものとの意見が多い。協会として今後その規模の大小は別として、データベースシステムがあれば、相当利用されるものと判断されるので構築調査に踏み切った。

2.4 既存データベースの利用の可能性

法令、物性、技術および事故の4種のデータベースに関連する既存データベースの調査を行ったところ次のような結果であった。

(1) 法令データベース

総務庁および通商産業省の法令検索システムは行政側が保有しているもので、行政ニーズに対応したものとなっているため、民間利用には難がある。

また、それらベースは主として法律のみが対象となっているが、一般的な利用では下位の政令・省令更には告示・通達まで検索範囲を拡げることが要求される。

このため、既存のシステムは協会の利用方法に合致し難い。

(2) 物性データベース

海外に大規模なシステムはみられ、高圧ガス取締関連の物質も含まれているが、収録物質数が多過ぎる反面物性値の方は十分でないものが多く、協会の考えているようなデータベースとしては適切とはいえない。

(3) 技術データベース

国内にはJ I C S Tの大規模な科学技術データベースがあるが、収録範囲が余りに広く、かつ検索システムが特定の分野の情報を求める協会の利用方法に合致していない。

(4) 事故データベース

商用データベースとしては新聞データベース程度のものしかなく、事故の概要しかわからないので、協会が求める情報としては不十分である。

以上のように、既存データベースはその収録範囲が余りにも一般的すぎて、内容や利用法が協会のニーズに合っていない。

2.5 基本方針

高圧ガスはコンビナート等の石油化学工業を始め、生活用家庭燃料としての液化石油ガスまであらゆる分野で広く利用されている。しかし、高圧ガスは非常に便利なものである反面、ひとたびその取扱いを間違えると大きな惨事につながる危険性を含んだものである。このため、各種の法規制や技術基準があり、これらを裏付ける物性データ、検査データ、事故データ等も必要となって来る。これらのデータは行政的には担保として必要となるばかりでなく、企業側も操業する上で必要となるものばかりである。これらのデータは現在、行政サイド、企業サイド、協会とバラバラに保有されており、それぞれ過不足のあるものと考えられる。このような状況から協会が中心となって、散在する情報を保安情報として総合的に管理するシステムを構築し、行政を始め、メーカ、ユーザ等に広く提供するよう指向すべきである。

この総合保安情報センターの一環として今回保安情報データベース構築を打出したものである。

保安情報データベース構築の意義としては、次のものがあげられる。

- ① 高圧ガス保安情報を集大成し、事故、災害等の防止に役立てること。
- ② 難解でかつ複雑な法令の理解を促進すること。
- ③ 高圧ガス関連の物性データとしてまとまったものがなく、利用する場合散在するデータを探す必要があるが、データ検索の効率化をはかること。
- ④ 高圧ガスという特定の分野で専門家の手により選別され、まとめられた技術文献情報をより多くの関係者の利用に供すること。
- ⑤ 高圧ガス事故関連の資料情報を多面的にとらえることにより事故防止対策の確立をはかること。

2.6 保安情報データベースの構築の基本的な考え方

2.6.1 システムの基本概念

保安情報データベースは基本的には法令データベース、物性データベース、技術データベースおよび事故データベースから構成される。これらはデータベース管理システムとして位置付けられるが、この他に全体システムを構築するには、用語（辞書）機能を含む一般情報データベース、データ入力システム4種のデータベースの検索システム、管理システムおよびメンテナンスシステムが必要となる。

これらデータベースやシステムの構築に当たっては外部のデータベースの内有用なもので、かつ、利用できるものについてはその導入も考慮する。

2.6.2 使用目的

保安情報データベースに含まれる4種のデータベースはそれぞれ利用目的が異なるものと考えられる。法令データベースについては、複雑な文書構成になっている法令について上下、左右の相関関係を効率よく検索し、迅速かつ的確に内容を把握するために使用される。

物性データベースについては、ある物質が法令の規制対象に該当するか否かの判定に利用されるのみならず保安対策や保安教育に必要な物性データを効率よく引き出すのにも使用される。

技術データベースについては、世の中に存在する一般的な広範囲のデータベースでは検索し難いような文献情報等を限定されたデータベース化を図ることにより、その検索が効率化されるものと期待される。

事故データベースについては、高圧ガスの保安（安全）を確保する観点から類似事故のケーススタディ等を行う際に、広くより詳細なデータを収集する手段として利用されるものと考えられる。

2.6.3 利用方法

保安情報データベースは協会の会員サービスの一環としても位置付けられるため利用者の主体は協会の会員を想定している。一方協会内部や中央、地方の行政の利用も考えている。

各データベースの利用方法については、協会内部と会員等外部の利用方法が異なるが、

代表的な例について述べる。

(1) 法令データベース

- ① 許認可手続き
- ② 資格制度
- ③ 適用除外
- ④ 高圧ガス設備，特定設備，ボンベ等の輸入手続き
- ⑤ 解釈事例
- ⑥ 関係諸法規との関係
- ⑦ 法令改正の内容
- ⑧ 告示，通達の内容

協会内部では法令改正の照会回答に利用することが多く，外部では各種告示，通達の照会を利用することが多いものとみられる。

(2) 物性データベース

物質名および物性値名を指定し，求める物性値の数値を求める。

例えば「ベンゼン」の「沸点」と検索し，沸点の数値を求める。

(3) 技術データベース

著者名，タイトル，キーワード等により検索する。或いは，データの出典名を検索することにより，データの入手先を紹介するような利用方法もある。

(4) 事故データベース

事故防止対策を講じるため，海外事故を含めた類似事故および事故原因を検索する。

特殊な例としては，高圧ガス関連のヒヤリ・ハット事故を必要とする場合もある。

2.6.4 保守，運営

保安情報データベースの構築後の運営，保守については，本システムの管理を行うのは協会とし，現行の協会の会員制度とは別に，本システムを利用する会員制度を新たに設立し，その会費および利用時間，回数による経費の設定を行い運用管理を行うものとする。

2.7 今後の検討課題

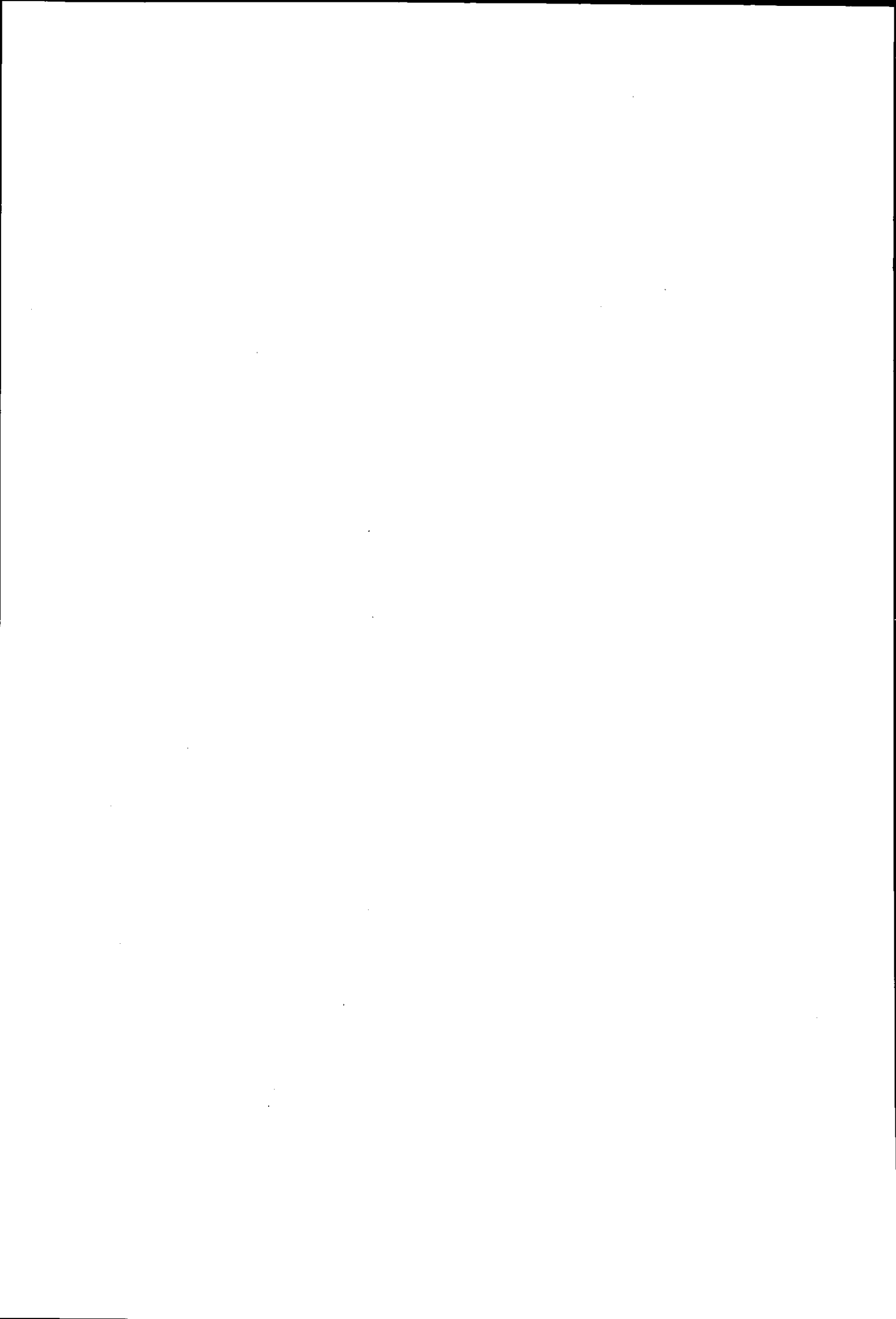
- (1) 4種のデータベース間で使用する用語等を管理する一般情報データベース中の辞書デー

データベースのあり方および作成。

- (2) 本データベースシステムの経済性の検討を行う必要がある。
 - ① 構築に用いる費用
 - ② 利用料金の設定
 - ③ ユーザ数
 - ④ ユーザ数と利用料金の関係
 - ⑤ データベースの運営体制
 - ⑥ コンピュータのハードウェアの規模と人員の配置
- (3) ユーザニーズおよび利用方法については、より詳細な調査を行う必要がある。
- (4) 保安情報データベースのシステムの基本計画について、より具体的な検討を行う必要がある。
- (5) 4種の個々のデータベースの問題点の検討を行う必要がある。
- (6) データの収集および評価体制作りの検討を行う必要がある。
- (7) 原報の保管体制作りの検討を行う必要がある。

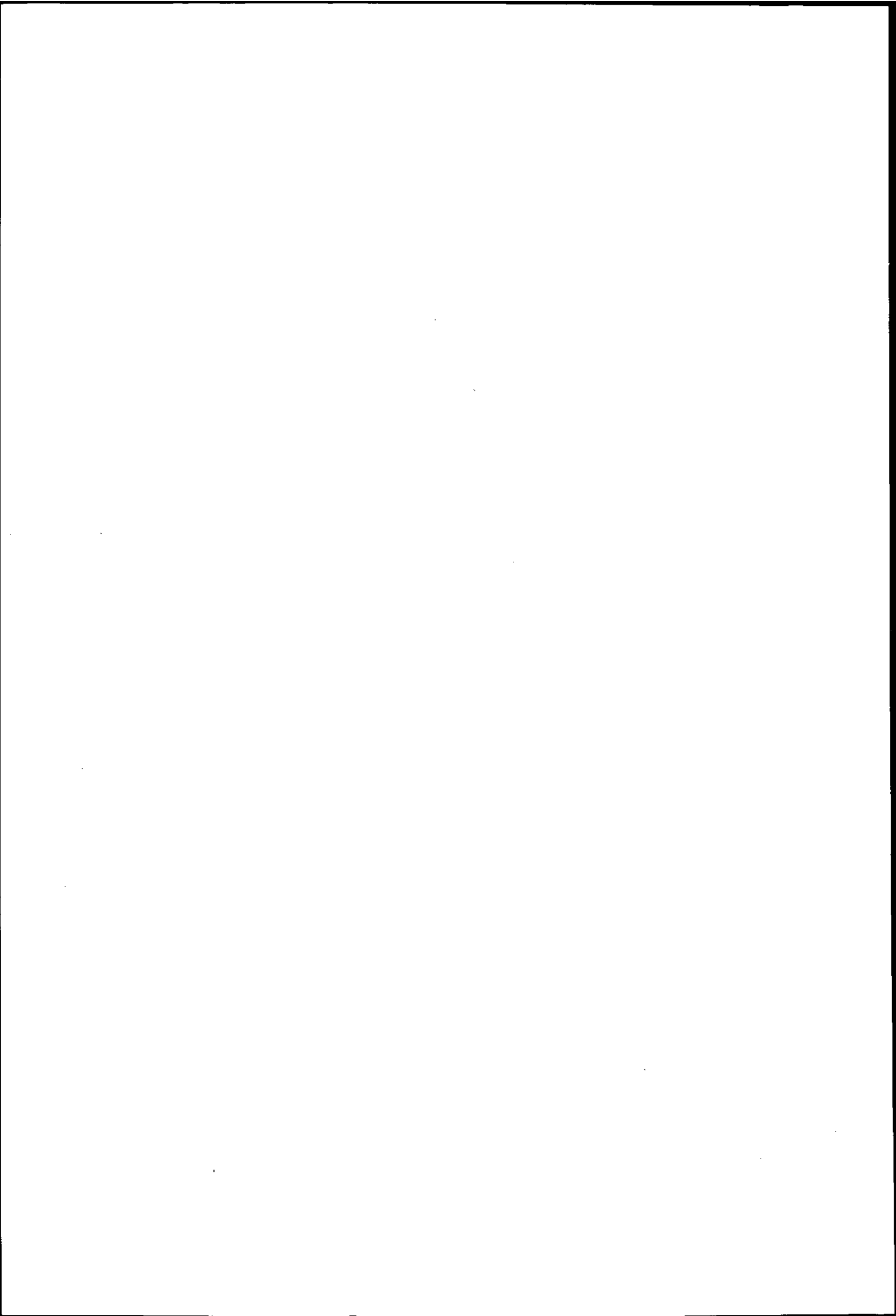
2.8 ま と め

調査、検討の結果協会に蓄積された知見やデータ並びに一部の既存データベースを利用するにしても独自のデータベースを指向する方が利用価値は高いことを確認した。



3 健康増進データベース構築

鹿児島県ニューメディア振興財団設立
準備委員会



3 健康増進データベース構築

3.1 構築の目的

わが国の社会，経済の高度化，成熟化に伴い，健康余暇に対する全国的なニーズの高まる中で，昭和60年8月通商産業省のニューメディア・コミュニティ構想のモデル地域の指定を受けた鹿屋市において，鹿児島県ニューメディア振興財団設立準備委員会は，「国際健康余暇開発システム」の構築に向け，各種の事業に取り組んでいるところである。

このシステムは，健康，余暇，スポーツ・レクリエーションに係る情報サービスを行い，地域社会の健康づくりに役立てようとするものである。

この分野におけるデータベースの構築は，対象領域が広く，未開発の分野であるが，今後，社会的需要は，ますます高まるものと思慮される。

本データベースは，鹿屋システムの一環として，個々人に適した体力づくり，健康増進に活用するスポーツ・レクリエーション分野を対象として構築したものである。

3.2 データベースの内容

データベースの内容としては，健康増進データベース開発委員会を編成し，この開発委員会により策定した健康づくりのための運動プログラム構築の基本方針に基づき，健康づくりに役立つスポーツ・レクリエーションの9分野，各分野5種目ずつの45種目について，基本的には，1種目5画像で構成し，内容的には下記のとおりとした。

なお，1種目5画像については，運動のフローチャートとして5画像で説明出来るものと判断し編集した。

画像No.	内 容 (小分類)
1. 画 像	種目概要，イメージ図
2. "	用具，費用
3. "	実施方法，ルール(注意事項)
4. "	効果，対象者
5. "	関係団体，連絡先，関係書籍

画像内容については、当初、個人の健康づくりに役立つスポーツ・レクリエーションを個人の健康状態や運動能力・趣向を判断してパーソナルな運動処方をアドバイスするコンピュータシステム（別途アプリケーションソフトを開発）と連携したデータベースを構築していくことを目標としていたが、資金、体制等の面で運動処方との連携は次のステップでの検討課題とした。

効果については、カロリー表示、目標心拍数等のパーソナルな内容について、いろいろ検討を加えた結果ケースバイケースで情報の内容が異なるため（幅がありすぎる）、危険性があると判断され、画像情報として取り込むことは難しかった。むしろ、健康づくりの啓発・動機付けという観点からアウト・ドアの種目を中心に、楽しくて、継続性、社交性（交流の場）がある種目にウェイトを置いた。

3. 3 データベースの基本構造

(1) データベースに利用されている主なデータ構造には、次のような3つの形態がある。

- ① 階層構造
- ② リレーショナル構造
- ③ ネットワーク構造

この3つの形態のうち、階層構造はその構造がツリー状であることから、一般にツリー構造と呼ばれている。

ツリー型データベースが多くのビデオテックスシステムで採用されており、鹿児島県ニューメディア振興財団設立準備委員会においても、これまでプレ実験用あるいはニーズ調査用のビデオテックス画像に採用し、実験を実施したところである。この実用実験の結果ツリー型データベースは、目的の情報に至るまで手間、時間を要するという問題点があるものの、一般の人が専門的知識あるいは特別な訓練を必要とせず容易に利用できるという評価が得られた。

このような結果から、データベースの基本構造としては、階層構造方式を採用した。

(2) データベースの基本構成

- ① ツリー構造

本データベースは、ツリー構造による（10進分類型）情報を整理・統合してある。

分類については前述のとおり健康づくりに役立つ運動種目を9分野に区分した。

9分野についてはそれぞれ1分野毎に中分類, 小分類に区分され, メインメニューを含め5階層で構成され, 小分類は基本的には5画像で構成されている。

このため, このデータベースは, 市販されているビデオテックスソフトで対応できる。

なお, 本データベースは, スタンドアローン型システムでも対応可能な5インチフロッピーディスク1枚(1MB)に収録することができるよう設計している。これは, 健康づくりに役立つプログラムとして生涯スポーツの指導者をメインにサブとして個人を普及のターゲットに想定しているため, 啓発, 動機づけに必要な情報量として, 小分類の情報を5画像程度の画像数におさえて編集している。

また, 場合によっては, 階層数を増やすことによって多くの情報を収録することは可能である。

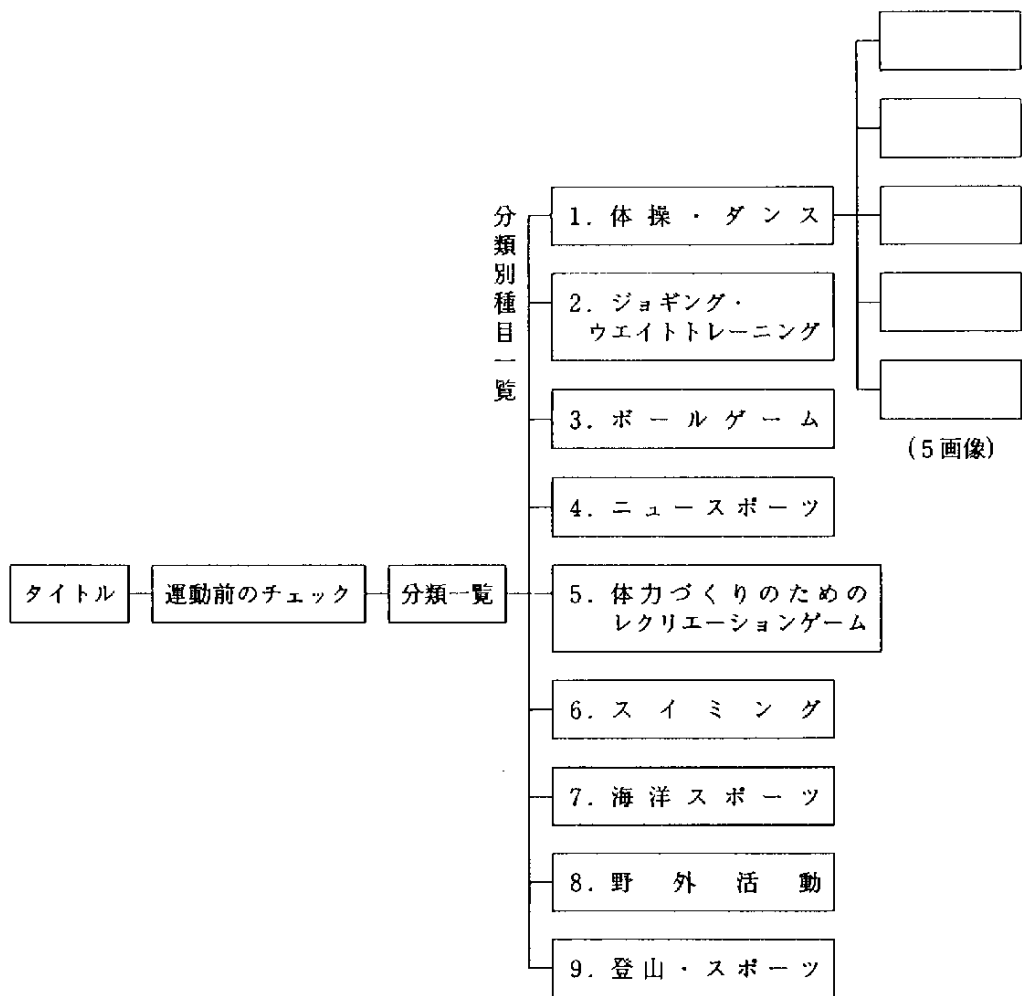


図3-1 ツリー構造イメージ図

② 検索フロー

本データベースは以下の表示順序を基本としている。

- a 起 動
- b メインメニュー
- c 運動前チェック (次画面)
- d 分類一覧(1～9) (選 択)
- e 分類別種目一覧(1～5) (選 択)
- f 種目画面(1) (次画面)
- g 種目画面(5) (選 択)
- h 分類一覧 or 分類別種目一覧

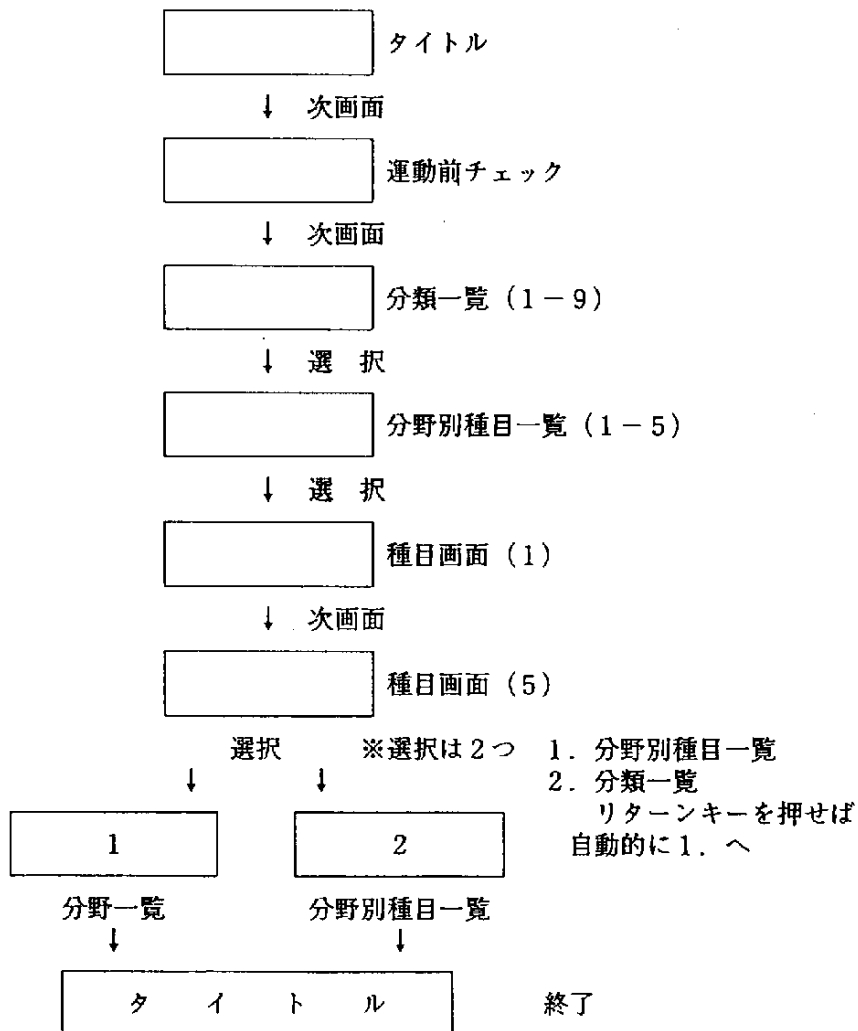


図 3-2 検索フロー図

3. 4 課 題

(1) 本データベースは、企業、自治体等の健康管理室、地域の健康増進施設での情報サービスシステムの一環として、それらの施設の管理者、インストラクター、企業の従業員、地域住民等が気軽に検索することにより、「健康づくりのためのスポーツプログラム」の啓発および動機付けという目的が達成できるよう、N A P L P S 方式のビデオテックスシステムの情報メディアとしての特性を生かして設計している。

このため、健康づくりにはどのようなスポーツ・レクリエーションが適しているかという観点から、中高年の生涯スポーツとして手軽に実行できるユニークな種目に限定して作成した。

今後、健康づくりに役立つスポーツ・レクリエーションデータベースとして利用者の多様なニーズ、ライフステージに応じた種目を拡大していくことが必要である。

(2) 検索システムの充実

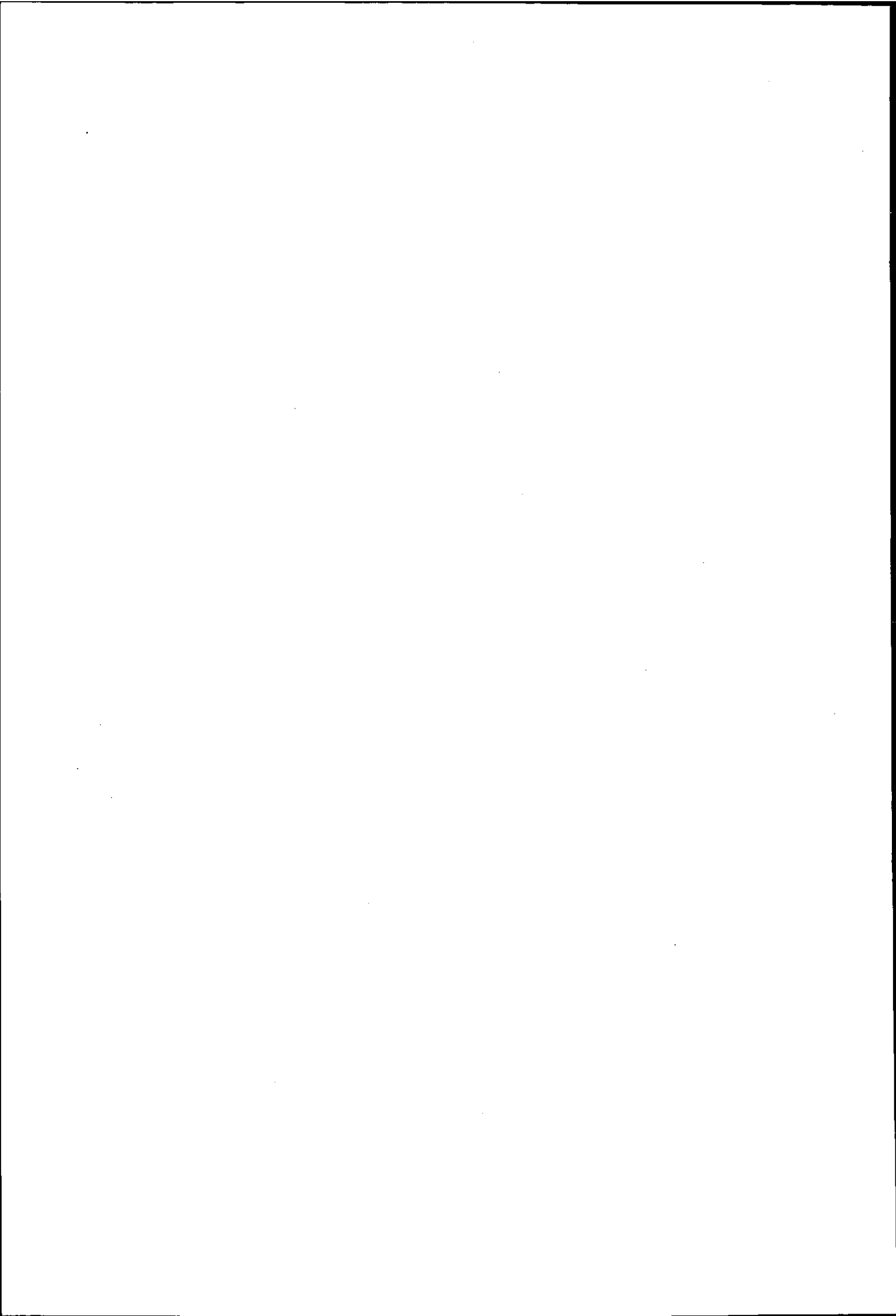
本データベースの検索システムについては、データベース検索の専門的知識を有しない者でも平易に利用出来るよう、ツリー構造によるメニュー検索方式を採用すると共に、利用者の興味の深化に対応し、「ゲームのやり方」「ルール」「用具」「対象者」「効果」「関係団体の連絡先」「参考文献、資料」という順序で情報内容を深化させ、啓発、動機付けという目的が達成できるよう設計している。

今後、対象領域が拡大されるに伴い、簡単なキーワードによる検索システムを付加することが必要である。

(3) 他のコンピュータシステムおよび動画、音声メディアとの連携

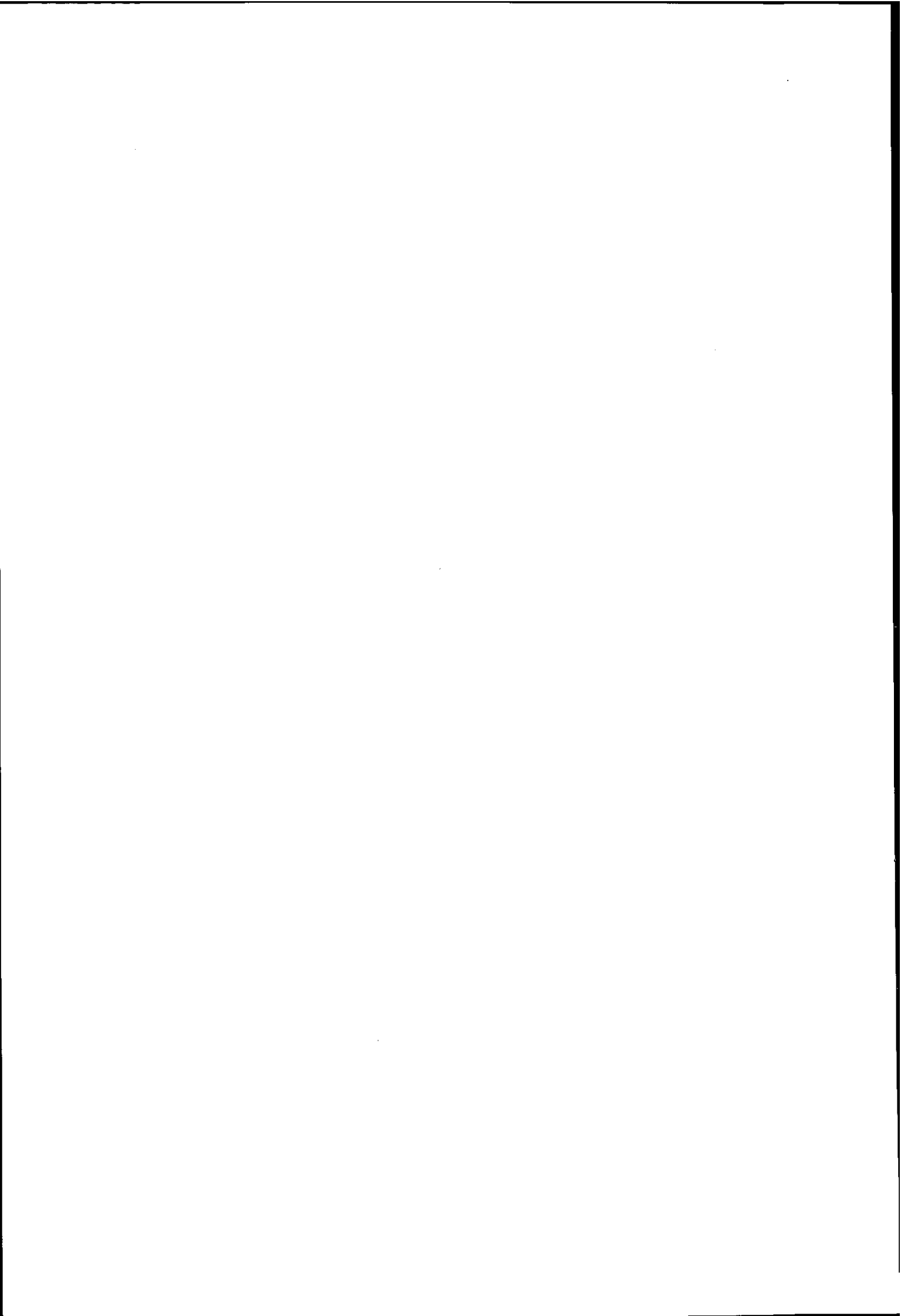
本データベースは、基本的には独立のシステムとして設計しているが、今後、運動処方システムや、更新性の高い施設、指導者等の地域情報データベース等とのネットワークが必要である。

また、運動プログラムの実行に当たっては、動き、音楽（BGM）といった静止画では表現できない情報が重要な要素を占めているためビデオ等の動画情報と音声メディアとの連携も考えられる。



4 法令・判例等シソーラスの構築
(法律エキスパートシステムに応用)

株式会社 日本法律情報センター



4 法令・判例等シソーラスの構築（法律エキスパートシステムに応用）

4.1 概要

4.1.1 目的

- (1) 株式会社日本法律情報センター（以下「当センター」という。）は、かねてから判例・関係論文等データベースを構築中であったところ、財団法人データベース振興センターから、昭和61年度データベースの構築促進及び技術開発事業の一環として「判例・関係論文検索システム」について事業委託を受けた。これにより判例データベース構築に当たっての、「ソフト」と「ハード」との関係、オフラインとしての「CD-ROM」とオンラインとしての「VAN」との関係、各領域の設定、キーワードの基本的な考え方、その内容・項目の検討などを行った結果、判例データベースの事業化が可能となった。「CD-ROM」「VAN」ともに実験段階を経て、昭和63年中にも本格的事業化を予定している。
- (2) 当センターは、裁判官・検察官・弁護士や大学教授などの「法律専門家」や中央・地方の官公庁、会社法務部関係者、図書館関係者、大学関係者などの「準法律専門家」に限らず、法律には素人の方でも気軽に利用できることを目標にデータベースの開発に当たった。ほぼ所期の目的を達したものの、素人が利用しても、法律専門家などが利用した場合と同等の結果を得るとともに、「漏れ」や「ノイズ」を防止するなど、より完全なものとし、また、多数の類義語などを体系的に分類・整理した類似語・関連語辞典、つまりシソーラス（thesaurus）構築の可否の検討は避けて通れないものとなった。
- (3) 更に、上記(2)の目的を徹底させるためには、AI（人工知能）を応用して、コンピュータに法律専門家の問題解決作業の肩代わりをさせたり、あるいは法律専門家の問題解決作業を支援させたりする「法律エキスパートシステム」を取り入れることが可能であるか否か、更に、可能であるとすればこれを判例等データベースのビジネスに取り入れることが望ましいかどうかを検討されるべきである。
- (4) このような理由で、「法令・判例等シソーラスの構築（法律エキスパートに応用）」について検討の必要がある。

4.1.2 実施内容

- (1) 判例等シソーラス・AI 応用検討委員会（以下「委員会」という）の開催
- (2) 各分野の判例約 1,000 件についての判例用語によるキーワードの切り出し
- (3) 法律辞書作成についての検討

4.2 委員会の概要

4.2.1 委員会の目的

- (1) 法令、判例についてシソーラスは必要か。法令と判例との使用用語の厳密性の差異。
- (2) シソーラスは、主としてどのような目的に利用するか。
- (3) シソーラスを作成する具体的な方法。
- (4) AI（人工知能）を判例シソーラスに応用することは理論的、ビジネス的に可能か。
可能であるとすれば、どのような事柄に対して応用すべきか。
- (5) 法律辞書作成の必要性の有無。
- (6) その他、上記に関連する事項。

4.2.2 検討結果

- (1) 法令については、その用語は厳密な規則性、統一性を持っているので、シソーラスは原則として必要はないと思われる。しかし、判例については、法令ほどの規則性、統一性は認められず、その他諸般の事情を考慮すると、原則としてシソーラスの必要性は否定できない。
- (2) シソーラス作成に当たっては、言葉特に日本語の曖昧さの問題、共同作業の場合の個人差、主観差の問題、同一人の場合であっても時間の経過等による変化の問題、グルーピング上の問題、グルーピングが困難な新しいグループの出現、頻度回数之多寡等からキーワードとするには不適切なものがあることの問題、法令と判例との関連付けの問題やメンテナンスの問題などがあり、過去に失敗、成功事例などがあるので、これらを参考とするとともに先達の意見も聞く必要がある。
- (3) 実験段階は別として、実用段階では「分野別」のシソーラスとすることが望ましいと考える。
- (4) 判例検索の具体的内容を検討してみると、AI（人工知能）を応用し、法律エキスパ

ートシステムを構築することは理論的にも実務的にも可能であると考える。

その目的は、過去に同種判例がなかったことを証明するいわゆる「零」証明のためではなく、「漏れ」や「ノイズ」がないことの証明に利用されるべきである。

- (5) IR（情報検索）とAI（人工知能）の機能の分離と限界ということ認識し、まずIRを固め、その上で、AIを構築すべきである。
- (6) IRを固める方法としては
 - ① キーワードの出現頻度が一度のもの、逆に極めて多数回に及ぶものなどは切り捨てる。
 - ② 同義語、類似語、関連語、反対語や上位概念、下位概念をどのようにすべきか、「ノイズ」の問題やアクセスタイムが長いなどの問題があるので、多角的に検討すべきである。
 - ③ IRシステムに入れて実証的に検討してみる必要がある。その結果、効率的なものだけ選ぶ方針とする。
- (7) 検索語の選定等でのノウハウは絶対に洩らしてはならない。また、その方策は可能である。
- (8) ソフトを中心に置き、ハードについてはどのメーカーの一定レベル以上の機種については適応が可能である、いわゆるマシンフリーの原則が貫ぬかれるべきである。

4.2.3 実験結果

(1) 実験の方向

当初は、実務分野中、法令・判例等データベースの利用者が最も多いと思われる「企業取引関係分野」と「科学技術・先端技術等関係分野」等について検討する予定であったが、委員会の大勢の意見等を基に、結局判例等データベースのアウトラインについて正確な知識等を得るとともに、判例等データベースについてのソースがビジネスとして成り立つかを幅広く知るために、全法令分野についての判例等ソースについて実験してみることにした。

(2) ソース作成の方法

原則として、最高裁判所判例は59年以降、高等裁判所判例は54年以降、地方裁判所判例は59年以降の分を対象とし、また、以下の公式判例集に登載されたものによった。

最高裁判所民事判例集

最高裁判所刑事判例集

高等裁判所民事判例集

高等裁判所刑事判例集

下級裁判所民事判例集

下級裁判所刑事判例集

各審級判例の内訳は以下のとおりである。

最高裁判所・・・・・・・・・・ 270件

高等裁判所・・・・・・・・・・ 546件

地方裁判所・・・・・・・・・・ 263件

簡易裁判所・・・・・・・・・・ 2件

合計 1,081件

(3) 実験に利用した法分野

憲法，国会法，裁判所法，行政組織法，行政救済法，財政法，警察・防衛法，土地法，環境保全法，教育法，民法，商法，民事訴訟法，刑法，刑事訴訟法，矯正保護法，社会法，経済法，無体財産法，条約の分野であって，全法律分野をカバーする。

(4) 各収録法の内容

収録した各法分野の内訳は，以下のとおりである。

憲法関係・・・・・・・・・・ 87件

国会法関係・・・・・・・・・・ 34件

裁判所法関係・・・・・・・・・・ 9件

行政組織法関係・・・・・・・・・・ 38件

行政救済法関係・・・・・・・・・・ 76件

財政法関係・・・・・・・・・・ 53件

警察・防衛法関係・・・・・・・・・・ 92件

土地法関係・・・・・・・・・・ 30件

環境保全法関係・・・・・・・・・・ 7件

教育法関係・・・・・・・・・・ 8件

民法関係・・・・・・・・・・ 266件

商法関係・・・・・・・・・・ 53件

民事訴訟法関係・・・・・・・・・・ 137件

刑法関係・・・・・・・・・・ 246件

刑事訴訟法関係	158件
矯正保護法関係	5件
社会法関係	153件
経済法関係	84件
無体財産法関係	22件
条約関係	4件
条例関係	10件

合計 1,572件

(5) 判例用語のキーワード付けの方法

前記判例集中の「判示事項」と「裁判要旨」部分について法律関係者が手分けしてキーワードを付し、これをコンピュータにキーワードとして切り出させた。これをソートし、これをもとに分類整理した。

(6) 人手によって付したキーワード(知能キーワード)の件数

合計 12,882個

(7) 特定キーワードの出現頻度別状況

① 1回	10,157
② 2回	1,335
③ 3回	474
④ 4回	263
⑤ 5回	149
⑥ 6回	87
⑦ 7回	85
⑧ 8回	43
⑨ 9回	41
⑩ 10~19回	167
⑪ 20~29回	46
⑫ 30~39回	12
⑬ 40~49回	15
⑭ 50~99回	5
⑮ 100回以上	3

仮に、1回のみ出現したキーワードと30回以上出現したキーワードの双方を、機械

的に切り捨てるとすると、その数は「2,690」となる。

(8) 類似語の状況

12,882語のキーワードは、基本ワードの形式で整理すると、合計1,417語となる。

一方、類似語を伴わないものであっても、それ自体複数回出現したものは、合計389語ある。

両者合計は、1,806語である。

若干無作為に抽出し、例示する。

(「頻度数」は、類似語を伴わないもので複数出現したものを示す。)

よ み	キ ー ワ ー ド	類似語	頻度数
あいて	相手	10	
あおり	あおり	3	
あかはた	赤旗		2
あきらか	明らか	4	
あくい	悪意	4	
あけわたし	明渡	4	
あたらないとされたじれい	あたらないとされた事例	1	
あっせん	あっせん	3	
あらた	新た	9	
ありばい	アリバイ	2	
あんぜん	安全	25	
い	医	25	
いいわたし	言渡	3	

4.2.4 AIの応用について

AIの応用については、特に権威ある専門家の意見として、判例ソースにAIを応用することは、可能であるのみならず、ビジネスとしても十分適応できる可能性があるとのことであった。

今後、上記その他IRの検討を待ち、これを踏まえて今後具体的・実践的に検討していくべきであるというのが、委員の一致した意見であった。

4.2.5 今後の対策

判例ソースの作成、AI（人工知能）応用の双方について、過去の例に学び、権威ある関係者の意見を徴し、また最大限コンピュータを利用するなどの方法で、今後の完成を期す。

4.3 成果と今後の課題

4.3.1 成 果

これまで、わが国には法令・判例関係のソースはまったく存在しなかった。

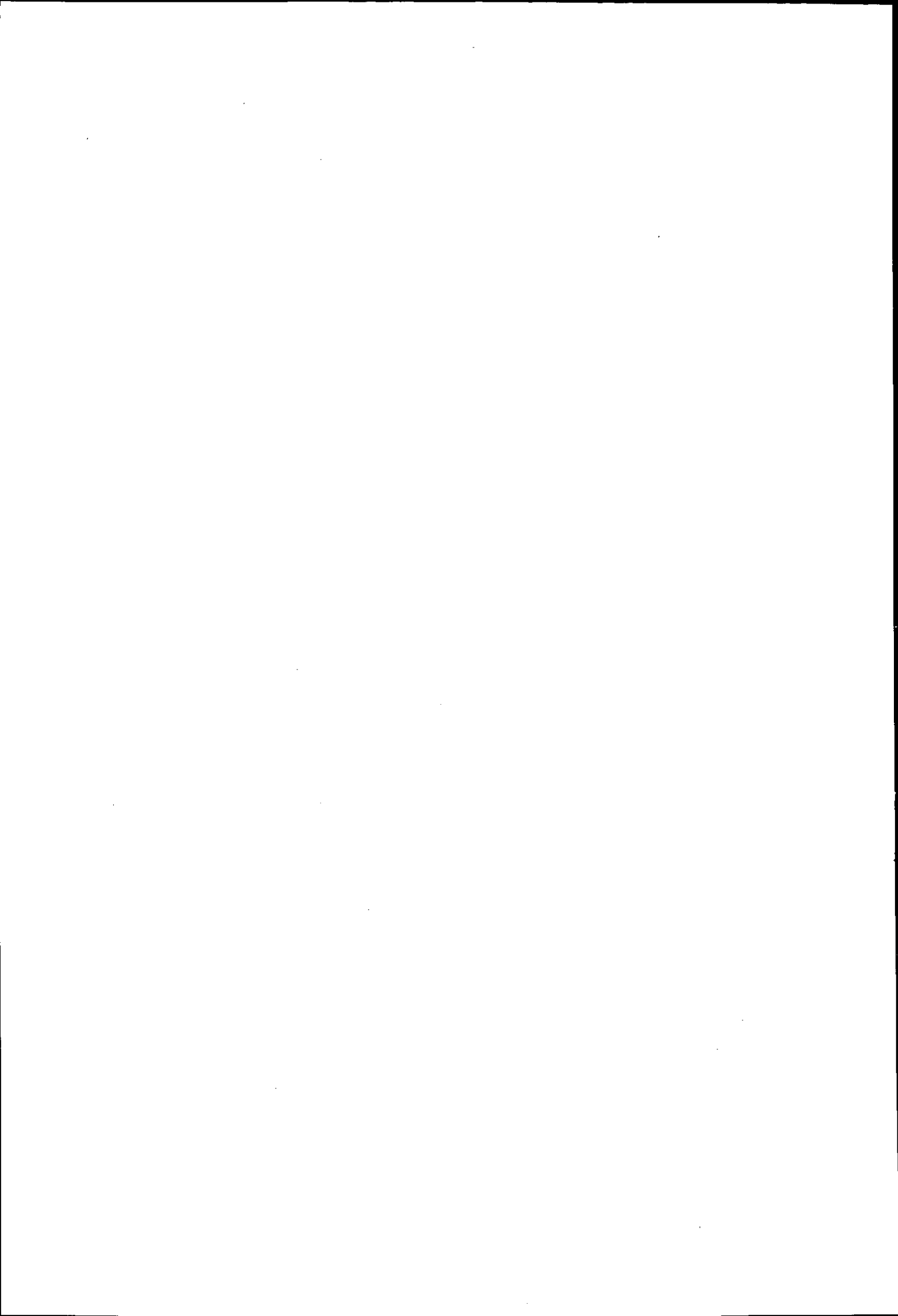
判例についてソースの必要があるかどうかについて検討の上、これが必要であるとの結論に達した上、法律の全分野の判例のソースにつき実験をし、一応の成果を上げたことは、今後判例等ソースの発展の上でもビジネスの面でも、その意義は大きいものがあると考ええる。

一方、判例ソースとAIとを結びつける検討をし、その方向付けを得たことも相当な意義があると考ええる。

4.3.2 今後の課題

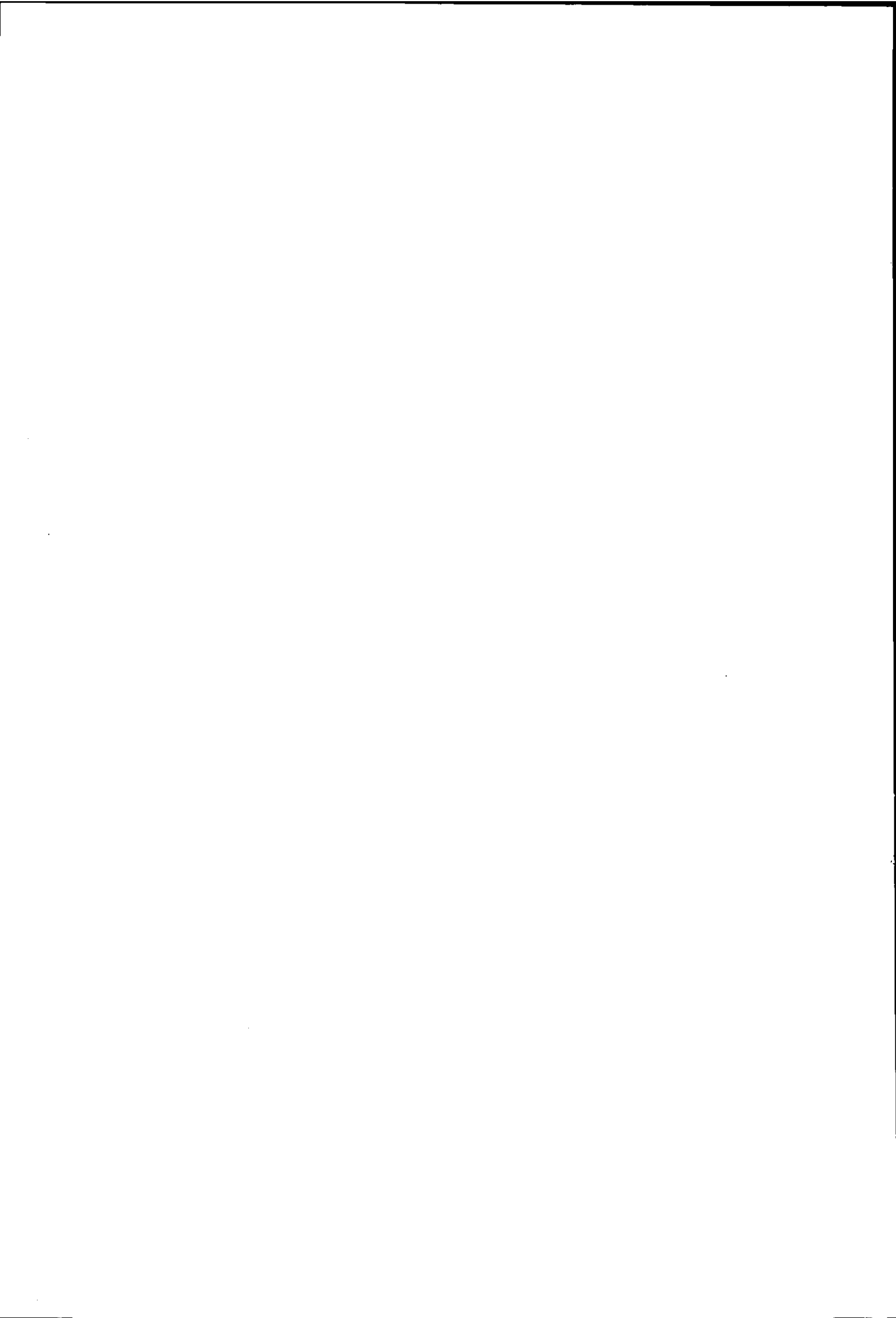
判例ソースの内容については、今後多数の判例を実証的に検討した上、商品としても通用する完璧なものとしていく必要がある。それには、少なくとも数万件の判例について検討していく必要があり、また、科学的、技術的な分析検討も不可欠である。

更に、AIを応用することについては、これからの課題であるので、その方法論を含め、今後検討していくべきであると考えている。



5 ハンディキャップパーソンの情報ニーズに即した
ライフサポートデータベースの設計

ダイヤル・サービス株式会社



5. ハンディキャップパーソンの情報ニーズに即したライフサポートデータベースの設計

5.1 構築の背景

5.1.1 わが国におけるハンディキャップパーソン（身体障害者）の実態

厚生省は、昭和62年11月4日付けで18歳以上の身体障害者と18歳未満の身体障害児の実態調査結果の概要を発表した。

それによると、身体障害者の総数は241万3千人で、在宅障害者の数は、前回の昭和55年に比べ20パーセント以上増加し、人口比も2.4パーセントから2.7パーセントとなった。

これは、高齢化に伴う脳卒中などの老人病による増加と、身体障害者福祉法の改正により新たに人工肛門、人工膀胱造設者や小腸機能障害者が加わったことなどが挙げられ、内部障害者は50パーセントも増えた。

5.1.2 ハンディキャップパーソン（身体障害者）の生活実態および情報ニーズに関するダイヤル・サービス（株）のヒヤリング調査結果のあらまし

(1) 調査目的

この調査は、限られた条件の中で、情報と接触している身体障害者の情報ニーズと情報行動を探ることを目的として実施した。特に情報を入手するメディアに対するニーズから新しい情報システムの方向性を探った。

(2) 調査対象

ハンディキャップパーソン（女性） 30名

(3) 調査方法

電話による調査 11名

面接による調査 19名

(4) 調査結果

1) ハンディキャップパーソンは、自らの障害を補うものとして、情報の新しいメディアを強く求めている。

2) ハンディキャップパーソンは、必要な情報を対価を支払っても獲得したいという意識が、きわめて強い。

5.2 構築の意義と目的

1981年の国際障害者年を機に、街には点字ブロック、車椅子用のスロープなど、身体の不自由な人々を受け入れるための「ハードウェア」の開発・普及が不十分ながらも徐々に進められてきた。

しかし、前述のとおり、1984年1月および1985年7月に弊社の「INSを考える女性委員会」が実施したヒヤリング調査等によると、身体の不自由な人々の多くが社会参加への強い意欲をいだきながらも、以下のような理由から、もう一步、足を踏み出せないでいる現実が明らかになった。

すなわち、ハンディキャップパーソンと「社会資源」を結び付けるための、情報ソフトの開発が、著しく立ち遅れているからに他ならないのである。

したがって、例えば、車椅子で生活する人々は、外出に欠かせない車椅子用トイレの設置場所を、事前に熟知することはなかなか困難なことである。

また、交通事故の激化、労働災害の多様化、人口の高齢化等によりハンディキャップパーソンの総数が激増をたどっている今日、ハンディキャップパーソンと、社会資源の架け橋となるような、まごころのこもった情報ネットワークシステムの構築は、緊急かつ、重要な社会的課題である。

また、弊社においては、1971年に開設した電話相談窓口「赤ちゃん110番」などを通じて、全盲の母親からの育児相談等を多数受け付けるなど、ハンディキャップパーソンの切実な情報ニーズとその対策の立ち遅れを日々の業務の中で痛感してきた。

そこで、国際障害者年が中間年を迎えた昨年、1987年より5ヶ年計画の予定で、全国241万3千人（1987年2月厚生省実態調査）のハンディキャップパーソンおよびその家族、ボランティア、社会福祉事業従事者などを対象に、弊社の20年近いテレコミュニケーションの実績とノウハウ、更に独自の研究活動（ダイヤル・サービスグループ内のシンクタンクとして生活科学研究所を併設している）をベースに、真にハンディキャップパーソンの情報ニーズに即した「ライフサポートインフォメーションデータベース」の構築と、それに基づく情報提供サービスに着手した。

5.3 概要

5.3.1 基本項目

- (1) ハンディキャップパーソンの情報ニーズに即した「ライフサポートデータベース」の

設計および情報の構築

- (2) 「ライフサポートデータベース」に基づき、電話による情報提供およびファクシミリ（主に耳と言葉の不自由な人々を対象）、パーソナルコンピュータ通信などを媒介として、個々のニーズにきめ細かく対応した生活情報の提供
- (3) 構築されたデータおよび利用者のニーズを適宜、報告書などの形にまとめ、社会還元していく

5.3.2 データベースの種類と量

1) 社会保障，社会資源など行政情報	1,000件
2) ボランティア情報	1,000件
3) 育児や料理など家庭生活に関する情報	25,000件
4) バーゲン情報，新商品情報，宅配サービス情報など買い物に関する情報	3,000件
5) 病院や補装具など医療に関する情報	5,000件
6) コンサート実施予定や車椅子で宿泊できるホテルなど趣味やレジャーに関する情報	5,000件
7) 親睦団体やサークルなどハンディキャップパーソンのネットワーク化に関する情報	2,000件
8) その他	多数

5.4 構築状況

昭和62年11月15日より5回のデータベース構築委員会の審議をふまえて、データベースの設計および作成に着手。

昭和63年1月末 設計終了

5.5 システム上の特色

本データベースは実施計画書と5回にわたって開催された委員会の成果を基に構築したものである。

本システムは、3期に及ぶ構想のうち第1期計画分、すなわちパーソナルコンピュータを使用してのデータベースメンテナンス並びに特定リサーチャー用検索システムの2点を目的として

構築したものである。

(1) データベースシステム

本システムの機能と将来のシステム拡張、変更を考慮した場合、データファイルはデータベースシステムで管理することが望ましい。

ファイルシステムに比較して、データをプログラムから独立させるためデータ、プログラムいずれかの変更に対し他方に比較的影響が少なく、またデータベース対話言語を通して、非定形照会にも対応できる等のメリットが考えられる。

既に導入済みのPC-9801（日電）上で稼働するRDB（リレーショナルデータベース）には、INFORMIX、dBASE II・III、RBASE 5000、SWING等があるが、業務の規模、言語・ユーザインタフェースおよび上位互換性からINFORMIXが最適である。

(2) OSとプログラミング言語

OS（基本ソフト）は、18bit PC上で汎用性の高いMS-DOS（マイクロソフト）を、またプログラミング言語としてはシステム記述性が高く、移植性の良いC言語が最適である。

5.6 今後の検討課題

5.6.1 点字および音声入出力システムの付加

本データベースは、3期にわたる構築計画の最終段階で、全国の身体の不自由な人々が必要な時、必要な情報に直接アクセスできるよう、パソコン通信の導入を計画している。

本データベースへのパソコン通信導入に当たっての最大の問題点は、目の不自由な人々からのアクセスを可能とするために、いかにして音声および点字による入出力を付加するかである。

本データベース構築委員会でも第2回、第3回会議でこの問題について討議した。本データベースへのパソコン通信導入に当たっては、

- (1) 音声および点字による入出力を容易にするために、漢字仮名混じりの情報に対応付けてすべて仮名文字のデータファイルの本データベース内部に併せ持つ。
- (2) いくつかの問題点は否定できないが、市販の盲人用パソコン通信ソフトを個々の利用者が活用することを前提とし、本データベース内部には特別なファイルを付加しない。

討議の結果、漢字仮名混じり文のデータに対応した仮名文字のみのファイルの本データベース内部に保有する場合、相当量の記憶領域が余分に必要となり、想定しているシステ

ム規模では技術的に不可能であることが明らかとなり、当面は「AOR テレワード」など、利用者が市販のパソコン通信ソフトを活用するとの前提の下に、本データベース内部に特別のファイルを保有することは行わない。

5.6.2 構築する情報の収集とその加工

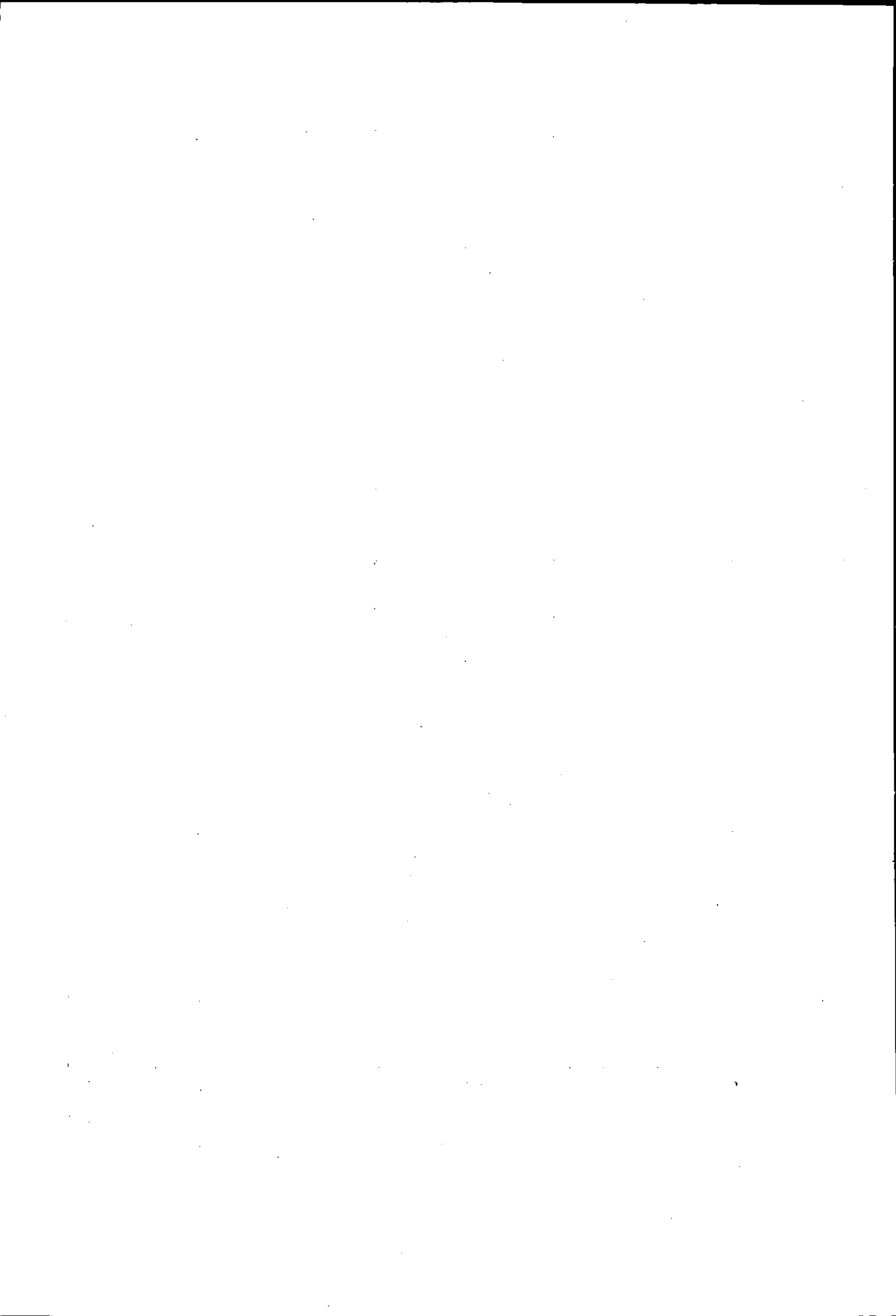
本データベースは第1期として、今年度は基本設計とその作成が一応の終了を見た。

次年度以降は本データベースの各カテゴリーについて、順次情報を入力・整備する第2期の段階を迎えることとなる。

言うまでもなく、本データベースは、高齢者や身体の不自由な人々のニーズにきめ細かく応じた情報を構築することに他に例を見ない独自性がある。しかし、現状ではこれらの目的をカバーするための出版物など既存のデータは皆無である。

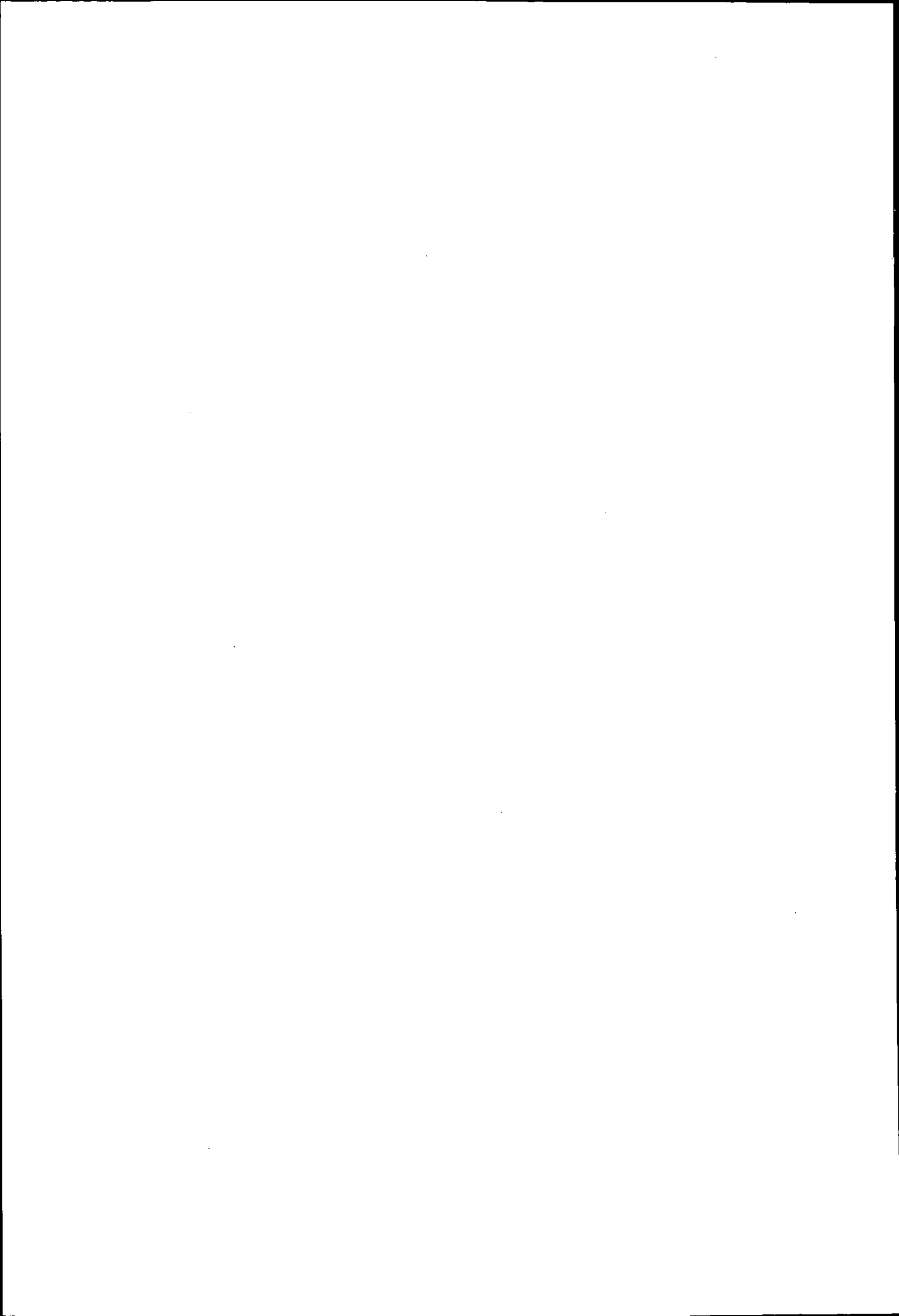
そこで、情報の収集とそのデータベース化については、一つ一つの情報内容に関して綿密な裏付け調査が必要となる。

しかも、本データベース構築に際しては、短期間に大量の情報収集とデータベースへの入力が求められており、人的・物的な情報収集態勢の整備と、それに要する費用の確保が最大の課題である。



6 中国地域における技術情報データベース
構築に関する研究調査

財団法人 中国地域技術振興センター



6 中国地域における技術情報データベース構築に関する研究調査

6.1 目 的

わが国の情報化は、従来コンピュータ・通信技術を中心としたハード技術を中心に進んできたが、社会の情報化の進展に伴い、このハードを用いた情報の入手や提供を主体とした情報そのものの重要性が高まりつつある。情報化が進展すれば、人は移動しなくても必要な情報を入手することが可能となり、情報を効率的に使用することは今後の産業活動、社会環境の整備、豊かな生活環境を構築する上で必要な要因である。また、中央と地方との間での情報ネットワークを作ることは地方の発展、活性化のための有益な手段と考えられる。しかしながら、情報化の進展は東京の一極集中化を更に加速化しているのが現実ではなかろうか。例えば、データベース利用者の75%程度は首都圏に集中しており、データベース・サービス業者の大部分は東京に立地している。

地域における産業・社会・生活等の情報化を進めるためには、各個人・各界の意識の高揚が基本的には望まれる。とくに、産学官が一体となったイメージ作り、個人に対する意識の高揚が必要であると考えられる。さもないと、わが国の産業構造が国際型へ転換する過程にあって、多様化する国民ニーズへの対応が遅れ、地域の企業、個人が取り残され、地域間の格差がますます拡大していく恐れがあるからである。中国地域が東京・大阪を中心とした日本経済から取り残されず、地域間競争に勝ち、地域活性化を図るためには、空路、道路を中心とした交通体系の整備と通信ネットワークの整備を進めることが緊急の課題となっている。また、高度の情報知識を有した人材育成、データベース産業の育成、ネットワーク化のためのプロトコルの統一も重要な課題である。

本事業はこうした背景から、中国地域の情報化の進展、とくに情報発信機能の役割を果し、地域の産業活性化を図るための一方策として、技術情報データベース構築の意義と効果を明らかにするとともに、具体的に提唱するデータベースの構造を示し、中国地域に蓄積されている技術情報をデータベース化するための基本的概念の提案を試みようとするものである。

6.2 成果の概要

本事業において検討・提案した内容を要約して示すと次のとおりである。

6.2.1 中国地域におけるデータベース

(1) 中国地域における産業の現状と課題

中国地方は、緑豊かな山と海に囲まれ、近畿地方と九州地方に接しており、古くから文化、技術の伝統があったが、物資の経由地、通過地という性格が強く、地域としての主体性が成立しにくい特性を有していた。瀬戸内海の恵まれた自然条件を活用した海運業、造船業が育ち、1960年代からは日本経済策の波に乗って、埋立てによる工業立地が進められ、大規模コンビナートの建設が行われ、瀬戸内海重化学工業地帯を形成した。現在では石油化学、造船、鉄鋼、自動車などの産業集積があり、裾野の広い関連企業群も育っている。

しかしながら、2度にわたるオイルショック、昭和60年9月からの急激な円高はこれらの産業に大きな衝撃を与えている。中国地方の主要産業である重厚長大型産業は付加価値が低下し、貿易摩擦にもより成長速度が減速化している。国際経済化が進む中で中国地域の産業構造を重厚長大型から、軽薄短小型、ハイテク化へと転換することや、新たな第3次産業への展開、地域に根ざした望まれる企業の育成をすることが求められるようになった。社会生活面の高度化を図ることによって、新たな雇用機会の創出や陰陽の格差是正、更には中央との格差の縮小を図ることが重要な課題となっている。

中国地域の工業製品の出荷額は全国の6～7%を占めるものが多く7%経済圏と言われている。このような産業構造においては、FA、OA、ネットワーク化が進められ情報化が進展している企業もあるが、全体的にはその進展度は低いものと考えられる。中国地域の情報サービス業の事業者数、従業員数、売上高は、各々全国比率では、5.3%、3.2%、2.2%となっており、すべての順位は東京、大阪、名古屋、福岡に次ぐものの、7%経済指標と比較すると大きく下回っており、中国地域の情報化の進展が進んでいないことが示されている。

(2) データベースの必要性

中国地域における企業をはじめとし、データベースに関する情報発信源の所在とニーズの調査を行い、地域に適した技術情報に関するデータベース構築の模索を行った。具体的には、アンケート調査により技術情報の現状を把握し、ニーズを調査分析している。

アンケートの回答のあった業種は製造業が大部分を占めており、現在利用しているデータベースの種類は研究開発、技術開発、特許、製品などの技術情報が大きなウェイトを占めている。最も欲しい情報としては、技術開発状況、製品開発、技術協力に関する情報が挙げられている。これらのことから技術情報データベース構築の必要性が裏付けられている。

(3) データベースの意義と効果

企業が必要としている「新素材」、「部品・材料」、「電気通信」等の分野におけるデータベースを構築することにより、中国地域の産業構造の転換を促進するとともに、中国地域産業活性化が図れるものと考えられる。また、製造業を中心とした各企業において、営業実績等の生の情報との有機的結合が迅速化され、経営戦略の策定と決断が容易になり、企業活動が前進する。更に、情報の地域発信が可能となり、東京等都市地域との格差是正に寄与するものと期待される。

(4) データベースの利用状況

中国地域におけるデータベース利用の現状について述べ、全国・他地域との比較を行い、中国地域独特のデータベースのあり方、発信源となる技術情報、必要とする技術情報は何かを明らかにしている。データベースを構築する技術については、通信映像技術の概略を述べた後、データベースの利用頻度を高めるための最新情報の入力確保、中国地域の特質を著しく表わした情報の整備、全国レベルでの地域データベースの相互連携、徹底したパーソナルニーズへの情報の提供を示している。

6.2.2 データベースの構造研究

(1) 構築するデータベースの基本方針

中国地域を情報源とする技術情報のみを収録し、地域外の情報については他地域に同一のデータベース構築の提言をし、それらのリンクにより実用性の高いデータベースとするため、基本方針は次のとおりとする。①データベースは「フルテキスト・データベース」の要素と「参考データベース」の両者を兼ねそなえたものとする。②データの提供方法としては、利用者の手元で活用できるよう「オンラインデータベース」とする。③データの収集とその入力には、イメージ情報が重要な要素をもち、またその抄録を編集する作業が発生するため、「集中入力方式」を採用する。④将来の便利性を考えて専用のハードウェアと独自のソフトウェアを開発する。⑤データの表現形式は、文字については「JISコード」による形式とし、イメージ情報については圧縮してファイルする。

⑥情報検索用端末は、パソコンとファクシミリを活用する。⑦他地域における同一データベースとリンクさせるための配慮を行う。

(2) 提唱するデータベースのハードとソフト

既存のデータベースで既に利用されている各種の機器・ソフトの比較検討をまず行っている。地域型データベースとして、技術情報をとりあげているため、かなりの規模の情報量となるとともに、ニーズに合った提供手段を研究する必要性を示している。ここで提唱するデータベースは、地域における技術情報を広範囲な分野にわたり収録するといった地域情報と、経済性を優先させる必要性、地域技術の振興を狙うために、データベースのハード、ソフトを具体的に開発することを前提としている。データベースに必要とされるものとしては、目的とするデータを回線を通じて遅滞なく、いかに高速度でデータを利用者に提供できるかということと、入力されたデータを有機的に格納保存出来るかを出発点としている。また、1次情報と2次情報の両者を同時に取り扱うこととし、大規模な記憶装置と同時処理可能な通信制御装置が必要であることを示している。データベース利用者は、パソコン通信により、情報検索をしてその結果提供された2次情報を見て、そこに収録された情報で不足な場合は1次情報が詳しく得られることを提案している。

(3) 他地域データベースとの連携

今後各地域の同種データベースを連携し、全国ネットワークのデータベースに拡大するため、データベース構造上の配慮や他地域接続インタフェースの配慮を行う。更にデータベース規格統一化のため、検索ツールの統一と通信モードの基準化の提案を行政に対して行い、今後日本のデータベースを欧米なみの水準にすることを提唱している。

(4) データベースの事業化

データベースの事業は、地域社会の官民が一体となって協力推進することが不可欠であるとの認識のもとに、第三セクター方式で発足することが望ましい。本研究のデータベース事業を実施に移した場合の事業規模と開発・運用体制についての検討を行い、事業面からデータベースシステムの評価をつけ加えている。

6.2.3 データベース構築のための今後の検討課題

中国地域において技術情報データベースを具体的に構築し、地域の情報化を進展させるための検討課題は、次のような観点からみることが出来る。まず第1は、データベースそのものの課題。第2は、地域と中央との情報化格差の是正。第3は、データベースを構

築するまでの具体的テストシステムの開発とその運用である。以下にそれぞれの課題について述べる。

(1) データベースそのものの課題

① 提供対象データの範囲

良きデータベースを構築するためには、質のよい、最新な情報を確保することである。情報を提供する立場と利用する立場では要求が異なる。いかにして良き情報提供者を探すかが重要である。例えば、ある技術分野では新規性がいくらか薄れた情報でも、他分野から見れば参考になる技術情報は、中国地域においても掘り起こせばいくらかは出てくるはずである。

② データ提供の体制

データ提供者が見つかったとしても、データ提供の体制が整備されないことには、そのデータはそのまま眠らざるを得ない。データ保持者が直接提供することは、データの形態もさることながら、立場、予算、手数、その他の制度上の制約が大きく極めて困難である。どのような体制によってデータ提供の道作りをするかが問題である。例えば、第3セクター方式、あるいは商工会議所の活用などによる提供体制についても検討することが必要である。

③ 提供の条件

データ提供者は多くの場合独立の経済主体であるから、データに対する対価は当然必要となる。現在のところデータに対する価値基準はあいまいである。これをどう明確化するかは重要な課題である。また、提供を受けた側のデータの取り扱いが問題となる。提供の条件として、そのまま収録するか、あるいはごく簡単な加工をして第三者に提供することが一般の商取引慣習からみて妥当である。ところが、利用者の立場からみると、データを要約化し、他のデータと融合させて新たな付加価値を加えることが望まれる。どのような提供の態様があり得るかを検討する必要がある。

(2) 地域と中央との情報化格差の是正

① 地域の情報化推進

産業高次機能の東京集中、産業構造の転換に伴う地域経済への影響等により、東京と地域の経済格差は拡大しつつある。地域における産業、雇用開発を促進し、地域経済社会の構築を図ることが求められている。これを実現するための手段として地域の情報化、とくに技術情報は重要な役割を果たす。今日のように国の施策が地域に目を向けている時代も少ないと言っても過言ではないであろう。ここで重要なことは、こ

これらの国の地域振興策を地域がどう活用して、先端産業の育成や情報関連分野の充実を図るかである。地域における情報化の効率的推進を図るためには、産・学・官のより一層の交流を促進し、地域の英知を結集することが必要となる。

② 情報化における東京と地域

東京には、経営戦略上の情報、国際動向の情報が集中しており、情報化における東京の優位性は揺るがない。また、国内的地位の向上のみならず、情報、金融、人の一層の集積が行われ、国際的な中枢都市へと成長しつつある。地域における情報化は、東京に集中した情報を活用することから進行しており、今後ともこの傾向は続くものと思われる。地域経済の活性化のためにはこのことは必要である。これと同時に、地域独自の情報発信を行い、地域の基盤の強化が求められる。東京の都市機能の制約と限界、危険分散という観点からみると、東京に集中している情報を可能な限りデータベース化して、地域で分有することによって、東京の一極集中を是正することが必要となる。また、地域の情報発信機能を高め、東京では出来ない機能を地域が独自性を活かしながら分担する必要がある。このためには情報通信ネットの全国的整備が急務となる。

③ 他地域データベースとの連携

地域が東京の単なる手足となるのではなく、東京との格差是正、地域からの情報発信を行って行くためには、次のような条件整備が必要である。まず、各地域と東京とのデータ伝送を直接行うことにより、効果的な情報の収集、伝達を行い、全国レベルの変化に即応し得ることである。各地域ごとに行われている社会活動や経済活動は、そのみならず、わが国全体の社会経済の発展にも必要である。このためには、情報ネットワークについても、各地域を単位としたネットワークを整備し、他地域データベースとの連携が重要となる。データベース相互間の交信を図るためには技術的側面からの問題解決が必要である。具体的には、データベース構造上の問題、データ伝送手順の問題、データ文形式などの問題解決を行わねばならない。これらの問題解決にはかなりの時間が必要である。

(3) 具体的テストシステムの開発とその運用

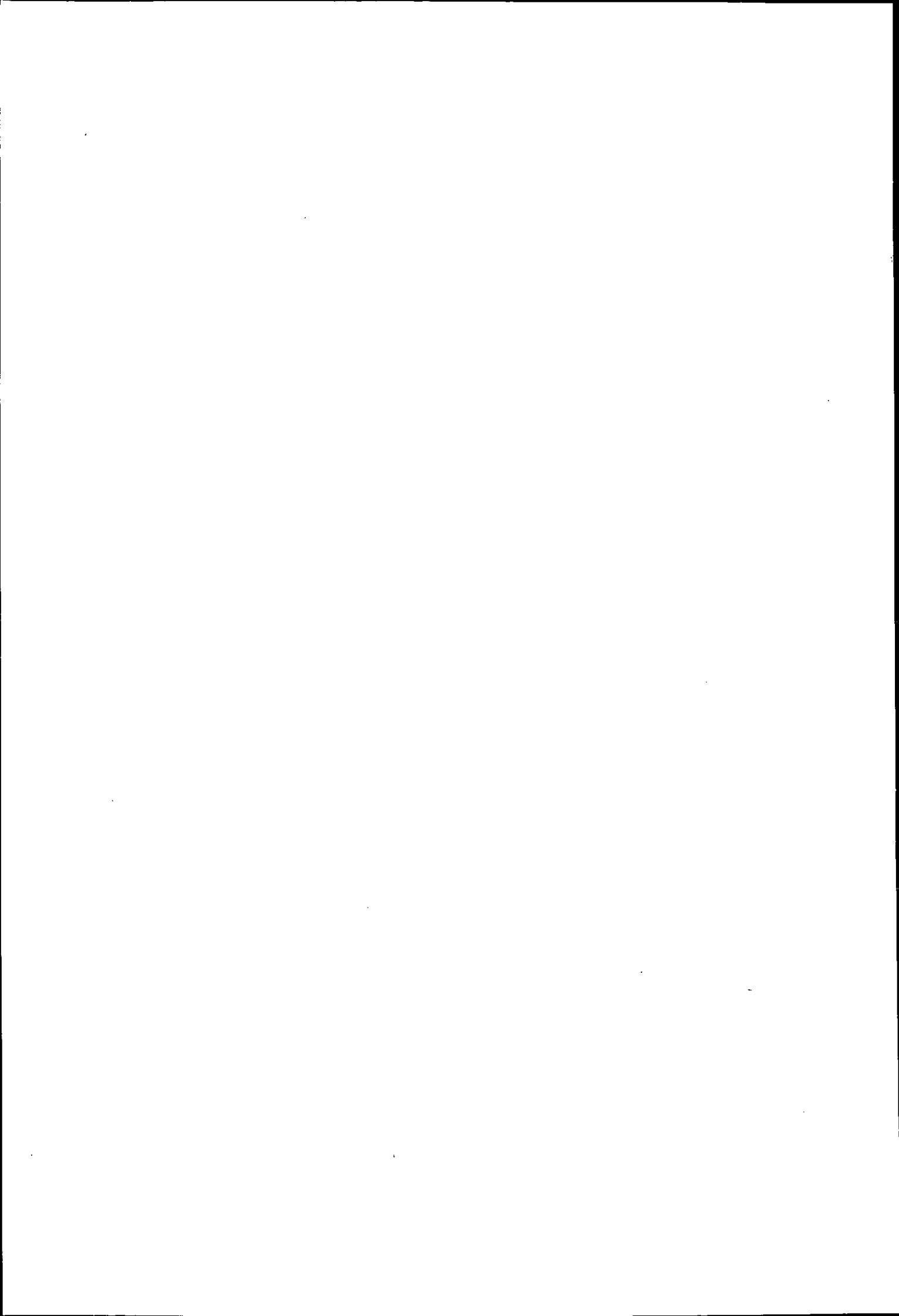
① テストシステムの開発

データベースを具体的に構築するためには、その前段階としてテストシステムを開発し、それを運用したうえで本格的な使用開始となる。まずテストシステムの設計が必要である。システムの概念設計に当たっては次のような点が検討課題である。取り扱うデータベースの量、データベース利用者の問題。具体的には、熟練者以外の者で

も活用を啓蒙することが狙いとなるから、検索の容易性、また、複数利用者が同時利用するための回線トラック数の検討などが必要である。運営収入の期待はあまり望めないで、設備コストは極力押えることが重要である。更に、データベースのシステム本体の他に、入出力装置の開発が重要な課題となる。キーボードによる文字情報の処理に加えて、図形をそのまま読み取り、データを圧縮して高密度記録ができるイメージ入力装置の開発は、対話機能を向上させる見地から早急に必要である。

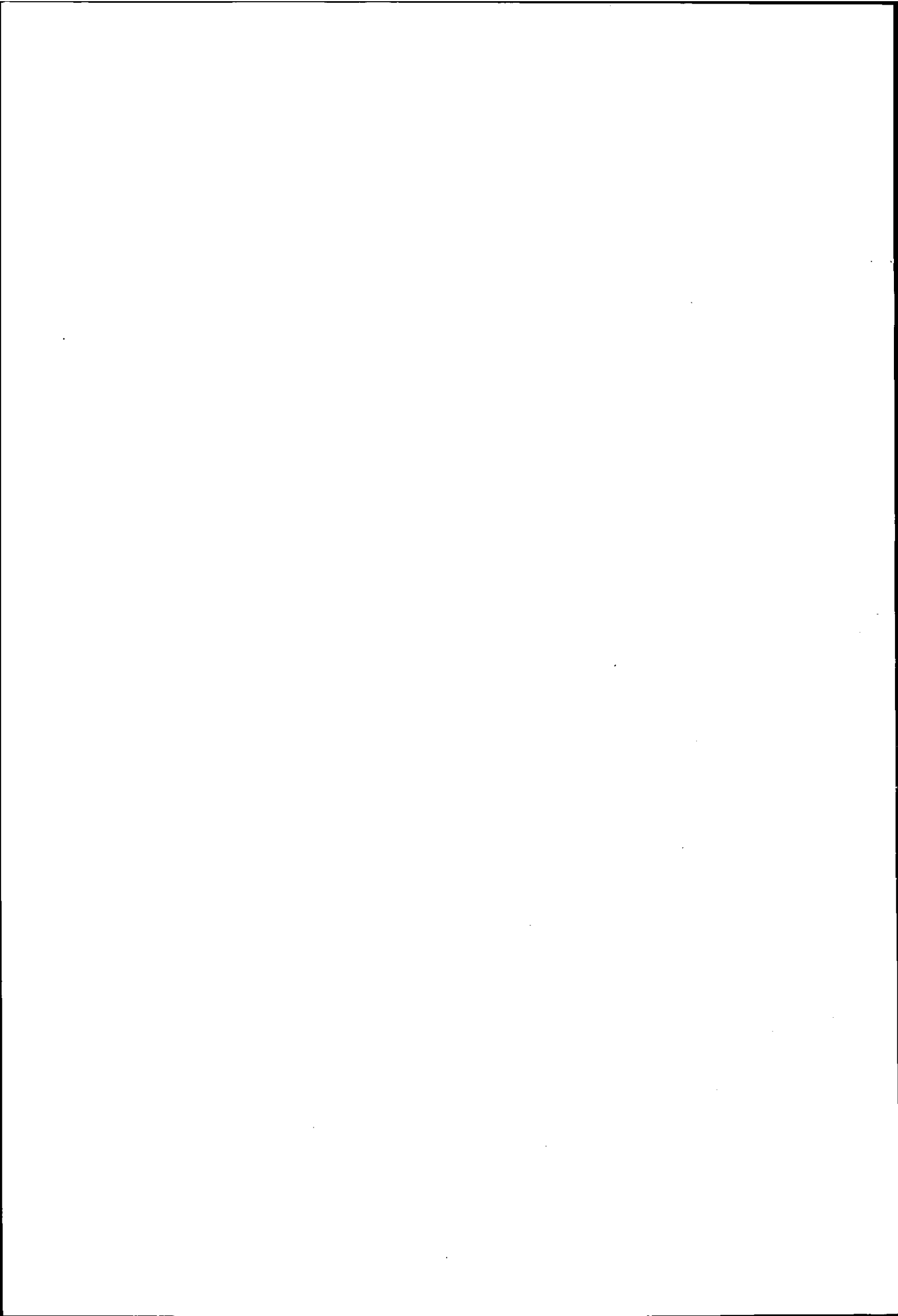
② 運用面

上記のようなテストシステムが出来上がったとしても、それをどう運用するかを検討しておくことが重要である。技術情報と一口に言ってもその種類は多く、内容的にも多種・多様である。当面のターゲットを何にするかを明確にしておくことが必要である。データベース構築上からみれば、データ収集の容易なものが優先されるはずであるが、これと利用者のニーズとがマッチしているかどうかを検討する必要がある。また、データ収集ではどこから、どんなデータをどんな体制で集めてくるかが問題となる。更に、そのデータをデータベースに入力するためのオペレータの確保をどうするかは重要な課題となる。



7 公設試験研究機関データベースネットワーク 構築調査

関西データベース協議会



7 公設試験研究機関データベースネットワーク構築調査

7.1 調査の目的

急速な技術革新と情報化が進展する中、中小企業においては、製品開発、事業転換を進める上で技術開発力の強化が一層必要となってきた。

公立で運営されている研究機関・公立試験場、いわゆる公設試験研究機関（公設試）ではこれら中小企業への技術指導・情報提供という使命に対応すべく、技術開発、技術指導、分析評価等の業務を積極的に行っており、それぞれの機関ごとに相当の技術情報のストックがなされてきている。

しかしながら、中小企業が欲する情報は、より高度化、専門化、あるいは融合化しており、一つの公設試の技術情報のストックでは、企業ニーズに対応できなくなりつつあるのが実情である。

こうした事態に対応するためには、複数の公設試にストックされた技術情報を共有し、相互に利用し得る体制を確立することが重要である。その共有、ネットワーク化の手段としてデータベース技術を利用し、府県間のネットワークを前提とした、広域利用を目指した公設試データベースの構築が急務である。

このため、本調査では、構築すべきデータベースの種類、ネットワークの形態等、近畿圏内公設試間のネットワーク化を前提とした公設試データベース構築のための調査を行ったものである。本年度は、近畿圏内公設試験研究機関で組織している近畿地方工業技術連絡会議（近工連）のメンバー機関によりモデル的に構想策定、実現の検討を行った。

本事業は、学識経験者、企業、官庁および本事業の当事者である公設試メンバーからなる委員会を組織、そのもとに公設試の実務メンバーによるワーキンググループ（以下WGという。）を設置し具体的な検討を行った。WGでは、各公設試の現状を明らかにしながら、データベース構築、情報ネットワーク形成のためのコンセンサスの形成を進めた。

7.2 公設試の情報化の現状

近畿圏内の工業関連の公設試験研究機関25機関を対象に、情報化の実態把握、課題抽出

のためのアンケート調査を行い、23機関から回答を得た。更にデータベースネットワークへの考え方、共同で構築すべきデータベースの内容等をヒアリング調査した。

調査の結果の概要は次のとおりである。

(1) 情報関連機器の導入状況

ほとんどの機関がパソコンを導入している。情報管理・情報提供手段としてでなく、技術開発・研究用として導入されている場合が多い。また、染織試験場などでは光ディスクを導入している。しかしながら、ネットワーク手段としてのファクシミリを導入している機関は少ない。

(2) 情報の創出・利用・提供の状況

独自のデータベースを構築している機関は、有効回答23機関のうち18機関。図書・文献DB(11機関)、実験・分析データ関連DB(11機関)、資料関連DB(8機関)、次いで企業等機関DB(6機関)、機器DB(5機関)となっている。これらデータベースは主にパソコンで管理され、内部の研究者および相談窓口が利用しているが、特定の業務目的のための断片的なものが多い。

各機関でのデータベース構築は比較的進んでおり、またコンピュータを利用したものであることを考えると、早急にデータベース構築の標準化を実施しないと、相互に利用したり、情報交換することが難しくなる恐れがある。

その他、外部への情報提供の手段としては、業務報告書、ニュース・レター、“～だより”というような印刷物が発行されている。これら情報の外部提供上の問題としては、作成費用、郵送費用の負担が大きいことが指摘されている。

(3) 情報の入手・利用の状況

外部データベースについては、有効回答23機関のうち13機関が導入・利用している。データベースの種類としては、技術・特許関連がほとんどであり、研究支援目的に利用されている場合が多い。次いで相談業務支援に利用されている。データベースについては、導入・利用のための予算がないという点が基本的な問題である。

その他、外部からの情報入手としては、雑誌類・図書類を始め、各公設試が発行している年報類、月報類を交換あるいは無償入手している。

(4) 情報の活用と流れの現状

公設試のデータベースネットワーク構想を考える上で、調査結果から公設試間および外部機関との間の情報のフローを整理する。

まず外部機関との情報フローについては、まず企業との情報フローという点で、これは

企業からの様々な相談・問い合わせに対して応えていく、いわゆる相談応答業務の遂行がある。ここでは、研究情報（既存の成果等）、研究員情報（こんな研究を誰がしているか等）、図書文献情報（こんな情報はどこに掲載されているか等）、設備情報（こんなことが出来る設備はあるか等）が流れ、具体的に相談業務・依頼試験等が行われる。

また、外部データベース、民間研究機関との情報の流れ、特許庁を始めとした官庁との情報の流れ、マスコミ、学会等との情報の流れが存在する。更に、国立試、他の公設試との間には、研究機関同士として相互に研究情報、研究員情報、図書文献情報、設備情報等の情報が流れる。

その実態からは、一つ一つの公設試の立場に立つと、同じ内容に関して多くのアクセスポイントにそれぞれ個別にアプローチしていかねばならず、人的、時間的、コスト的に負担が大きく、それ故十分なる情報活動が行い得ないという課題が指摘できる。

一方、企業など公設試の利用者の立場からしても一つの公設試で要領を得た回答が入手できにくいという不便さ、非効率性も発生することとなる。

このことから、各機関間の情報フローを整理し、相互に必要な情報を共有するための体制が不可欠である。

次に、公設試内部における各業務間の情報の流れを整理したのが、図7-1である。

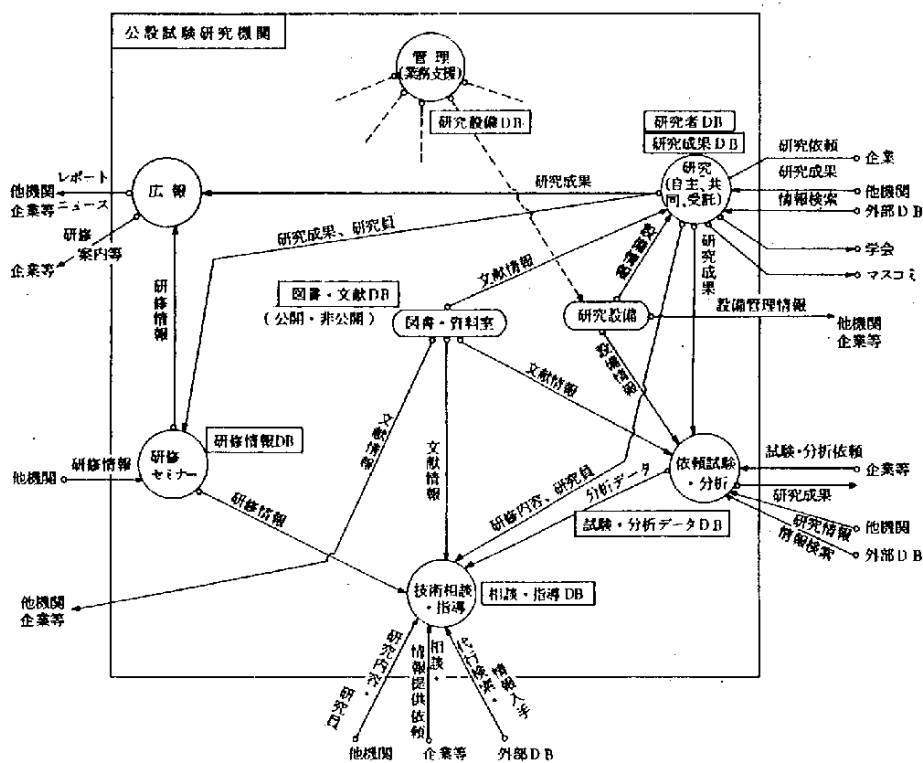


図7-1 公設試内部における各業務間の情報の流れ

ここでは、管理業務、研究業務、技術業務・指導業務、依頼試験・分析業務、図書・資料室、研修・セミナー業務、広報業務に分けてそれぞれの業務間の情報フローをみたが、それぞれの業務で必要とされる情報は共通するものが多いことが明らかとなった。また、管理業務更に技術相談・指導業務においては、機関内における情報を十分に把握しておく必要がある。特に、技術相談・指導業務にあつては、企業からの情報ニーズに的確に対応していくために、研究情報を始めとして研究者情報、図書・文献情報、設備情報等を一元的に把握しておくことが必要である。

(5) 情報化に関する予算の現状

情報化関連の予算では、コンピュータを保有している機関ではコンピュータの運営費用が大半を占めており、そのほか図書購入費、出版物の作成費がある。ただし、金額的にはあまり多くなく、データベース構築といった場合にはそのままでは難しい。

7.3 公設試における情報化ニーズ

公設試データベースネットワークを構築するためには、第3者からの要求を提示するのはもちろんであるが、基本的には公設試自身が技術情報のデータベース化を自らの問題と考え自助努力で遂行すべきものである。

その点、公設試自身は自らの情報化ニーズをどこにあると考えているのか、またそれぞれの機関が進めようとしているデータベースの構築計画はどんな内容か、本事業の構想策定のための当事者の考え方をアンケートにより整理した。

(1) 相談業務等からみた情報化ニーズ

相談業務等外部のニーズに対して応えていくために今後整備する必要がある情報としては、「特定の技術に関する情報」「各種資料」「企業情報」「研究員(人材)情報」「機器・設備情報」となっている。

この傾向は、既にデータベース化している情報内容とは異なっており、新たなデータベース整備が必要であることをうかがわせている。特に、中小企業等外部の様々な相談、問い合わせに答えていくためには、例えば企業カタログ、業界団体・製造業者の機関情報、また他の公設試の業務内容等の他機関の詳細な業務案内情報が求められている。これらの情報は、相談者等外部利用者の利便性を考えると、各公設試が個別に対応するだけでなく、機関相互にデータベースを共有化してどこからでも対応できるようにするのが望ましいで

あろう。

(2) データベース構築意向

データベース構築のニーズは各機関とも旺盛で、当面構築すべきデータベースがあると応えた機関は、有効回答23機関のうち21機関に及ぶ。内容的には、図書・文献データベース、実験・分析のデータベース、機器（設備）データベース等が優先すべきものとしてあげられている。

また、データベースの共有化問題については、共有すべきデータベースがあるとする機関は、有効回答23機関のうち21機関。独自にデータベースを構築する場合、予算、体制、人が問題になることも背景にあり、むしろ、データベースの内容によっては共有の資産として共同で構築するほうがよいと考えている。これら共有すべきデータベースの内容としては、機器（設備）データベース、研究者データベース、図書・文献データベース、実験・分析データのデータベースがあがっている。データベースを共有する場合の問題点としては、指導機関のリーダーシップ、各機関が同じレベルの情報を提供できるか、データベースの管理等の点が指摘されている。

7.4 データベースネットワーク構築の必要性和実現に向けての課題

(1) 必要性和期待効果

公設試データベースネットワークの実現のためには、外部からの要求に加え、当事者である公設試自らの切実なニーズが存在することが前提である。

先のアンケート調査からは、その切実なニーズはもとより、既に先行の芽がでていることが明らかとなった。現実的なデータベース構築の動きがあることは、その志向する実現形態がいかなるものであるにせよ、その原動力となるニーズが現実存在し、それが本事業にとって強い追い風になっていると認識できる。

データベースネットワーク構想の必要性については次の点に集約できる。

- ① 個別機関内部での業務の効率向上、利便性の向上等管理上のメリット
- ② 個別機関の対外部サービス実施上のメリット（サービス充実）
- ③ 近江連ネットに期待されるメリット

ここで、①②については現状調査の節で詳しく述べたところであるが、それにもまして③のメリットが大きいと考える。

③については、各機関のデータを共有化し、一元的な利用体制を構築した場合には、単に機関相互間のデータ交換、一元的なデータファイルの統合・ネットワーク化の実現というだけでなく、個々の機関が、より大きい仮想の「近畿圏公設試験研究機関連合」というべき機能体の一構成員としての位置付けを得ることができるというメリットが重要であろう。これは、もちろん各機関の主体性に制限を加えるような概念ではなく、むしろ逆に各機関のサービス機能における汎用性と専門性の両立を助けるものとなろう。これは、利用者のサービス要求に対して、自機関が自力で対応すべき部分とネットワークとして対応すべき部分との使い分けを可能とし、いわば質的な変化を持った情報ストックが期待できる。これは、まさに「情報資産の共有化による高度利用」ともいうべきものである。

(2) データベースネットワーク構築の課題

それでは、公設試験データベースネットワークを実現するために、どのように考えていけばよいのであろうか。現状調査からは、そのための重要な課題が明らかとなっている。ここに整理した課題は、データベースネットワーク構築の当事者である公設試験からの意見をまとめたものであり、極めて説得力のあるものである。

① コンセンサスの形成を図るべきである。

データベースネットワーク構築のためには、できるだけ多くの機関が足並みを揃えて標準化された基準に従って互いの情報を相互に利用し合うことが前提となる。そのために、各機関が共有のデータベースへの参加の意志決定が行えるよう近畿地域工業技術連絡会議の中に、例えば技術情報部会（仮称）を設置するなどして、種々の議論を経てコンセンサスを形成していくことが肝要である。

② 「標準化」を急ぐ必要がある。

個々の公設試験においては、既にデータベースの構築が進められている機関があるとともに、その他の機関においてもデータベース構築計画が検討されている。これらデータベースの構築が各機関ごとに進められるならば、今後公設試験間のネットワークを構築し、相互の情報交換を行う際に大きな支障となる。こういった問題を未然に防ぎ、全公設試験の情報資産を有効に活用していくために、今の段階で、物理的な統一ではなく、蓄積すべき情報、情報の内容、データのフォーマット等の論理的な標準化、すなわちビジネスプロトコルの標準化を急ぐ必要がある。

③ データベース構築の優先順位を決める。

地域の中小企業への情報提供サービス、きめ細かな相談業務の遂行といった対外サービスの質的な向上を図るため、また同時に各公設試験内の研究業務支援を強化すべく構築

すべきデータベースの内容を決定することが急務である。

- ④ 共有すべきデータベースと各機関が構築するデータベースの区分および構築の在り方を決める。

共有化すべきデータベースの内容と個々の機関が構築すべきデータベースの内容を明確に区分し、共有化すべきデータベースについて詳細に検討すべきである。

- ⑤ データベースネットワークの方向付けが必要である。

データベースの共有のためには、データベースの相互ネットワーク機能が不可欠である。ネットワーク機能を具体的にどういうふうに持っていくのか、また、近畿圏内のネットワークのみならず全国レベルのネットワークまでどういうふうに進展させていくのかを詳細に検討する必要がある。ここでは、当然のことながらVAN機能を介在したプロトコル変換等の作業を検討する必要がある。

- ⑥ 公設試の体制整備を図る

各公設試において、本構想を実現していくためのデータベースネットワーク構築の推進予算が組まれることが期待される。各機関において、情報関連人材の充実、情報化担当部署の設置、情報化関連予算の充実等、体制整備が不可欠である。

- ⑦ 助成制度の利用と確立

公設試は、本来各地の自治体に所属する機関であり、帰属する自治体からの情報化関連の予算が拠出される必要がある。しかしながら、本事業のような自治体を越えた共通の課題に向けての取り組みに対する予算の拠出には、その主体が異なることから、予算的には足並みが揃いにくい場合もあり得る。従って、国としての新政策の推進の観点から、特別な配慮をしていく必要がある。

7.5 データベースネットワーク構築ビジョンとステージプラン

委員会とWGを通じての具体的な検討により、少なくともWGメンバー、調査事務局の間においては、公設試データベースネットワーク構想についてのイメージ、前提条件として次のようなある程度のコンセンサスが形成できた。

- ① データベースの形態は分散型とする。すなわち、個々の機関で発生する情報については個々の機関でコンピューターリーダーダブル化し蓄積していく。
- ② その際、共有すべき情報については、あらかじめ定められた標準化の基準に従い整備す

る。標準化は、機種のご統一等物理的な統一を意味しない。

- ③ 機関相互のデータベースネットワークについては、中立の第3者機関（あるいは公設試の中心機関）がスイッチング機能を果たし、データベース共有の仕組みを運営していく。
- ④ データベースの利用者は、当面は各公設試の相談窓口の担当者とし、相談窓口を通じて企業への応答サービス、情報提供に供するものとする。これは、同時に内部業務の効率化に寄与するものとする。
- ⑤ ネットワークについては、近畿圏内のネットワークから着手し、近畿から全国的な標準モデルを提示する。順次全国レベルへと拡大していく。
- ⑥ データベース構築に伴う費用は、基本的には各公設試において負担する。従来、情報提供業務として実施していた刊行物の発行業務等が、データベースに置き換わったとして、それらの費用から拠出する考えもある。

次に、検討の結果整理できたそのプランを表7-1に提示する。本表データベースネットワーク構築ビジョンとステージプランは、先に述べた前提のもと、課題を順次解決していき、最終的には昭和68年に全国ネットワークを構築するというスケジュールでそのプロセスを描いた。

まず、この計画の初年度に当たる来年度には、最終的に構築すべきデータベース、また、それらのうち共有すべきデータベース、更には整備情報の優先順位をどうするか等の基本的な点について明らかにするとともに、全対象データベースの標準仕様を決定し、各公設試でのデータベース構築の条件整備を行う。更にモデルケースとして、優先順位の高い代表的データベースを選び、実際に構築してみることによってデータの収集・作成、維持・更新の方法についても実証的な検討を行い、その後の各種のデータベース構築についての指針の確立を図る。

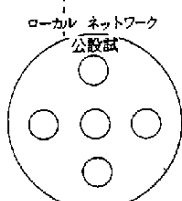
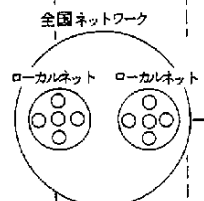
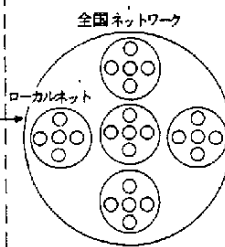
更に、地域ネットワークの核としてのローカルセンターおよびホスト機能について検討し、一部のデータベースについて地域ネットワークのモデル的展開を図る。

第2年度以降は順次地域を拡大し作業を進める。

7.6 当面とるべき対策

先に、ステージプランにおいて、来年度実施すべき作業を整理したが、具体的な来年度の業務としては、今年度の議論においてコンセンサスが得られた事項を具体化していくことで

表 7-1 データベースネットワーク構築ビジョンとステージプラン

年度	63	64	65	66	67	68
1. データベース及びネットワーク構築のための事前検討課題						
<ul style="list-style-type: none"> 対象データベースの決定 共有すべきデータベースと各機関個別構築データベースの区分 各区分毎の優先順位の決定 	→	→	→			
2. データベースの構築		① 1-st とは、最優先すべきデータベースに関するもの 2-nd とは、「1-st」以外の主要なデータベースに関するもの				
<ul style="list-style-type: none"> 標準化の検討 データの収集・作成の方法 データの維持・更新の方法 各機関でのデータベースの構築 共有データベースの構築 	1-st 2-nd 1-st 2-nd 1-st 2-nd					
	1-st	2-nd (順次 立上げ)	完	モデル機能を設定して構築する		
	1-st	2-nd (順次 立上げ)	完			
3. ネットワークの構築						
ローカルネットワークの核機能の整備とローカルネットワーク構築及び全国ネットワークの核機能の整備と全国ネットワークの構築		 ローカル ネットワーク 公設試	 全国ネットワーク ローカルネット ローカルネット	 全国ネットワーク ローカルネット		
<ul style="list-style-type: none"> ローカルネットワークの核機能の整備 ローカルネットワークの構築 全国ネットワークの核機能の整備 全国ネットワークの構築 (ローカルネットワーク間の接続) 	検討 モデル地域 以下 順位 整備 検討 モデル地域 以下 順位 構築	検討 整備 検討 接続	全国整備完了 全国のローカルネットワーク 実験段階 (モデルケース) 実用段階 (順次ローカルネットワーク間の接続)			
4. コンセンサスの形成	全国委員会設置	検討・実用化	(全国運営委員会)		NNの実現化	
5. 各機関の体制整備	検討	モデル機関実施	全機関で体制整備		完了	
6. 助成制度の確立	検討	実施				
7. その他(他ネットとの関係)					検討	接続

注 ① 上記3以下の件は、4の「全国委員会」の場で検討していく。

② ローカルネットワークの核機能とは、各通産局管内の代表機関又は第三セクターとし、管内各機関間のネットワークの核としてスイッチング機能等を果す。

③ ローカルネットワークとは、同管内の「核」を中心とした同管内の各機関を接続するネットワークである。

あると考える。

具体的には、次の事業を考えている。

① 主要データベースの標準化の検討

公設試が共同で推進すべき共有データベースについて、整備情報の種類の優先順位を決定し、各データベース構築の標準仕様を決定する必要がある。各機関における対応は、将来の機関相互のネットワーク化を前提として行われるべきである。

本年度の調査結果からは、まず相談業務支援、対外サービスの質的な向上のために、各公設試等機関概要情報、保有する技術・設備情報、依頼試験情報等、機関案内情報を中心として、優先的に共有データベースの標準化を検討すべきである。

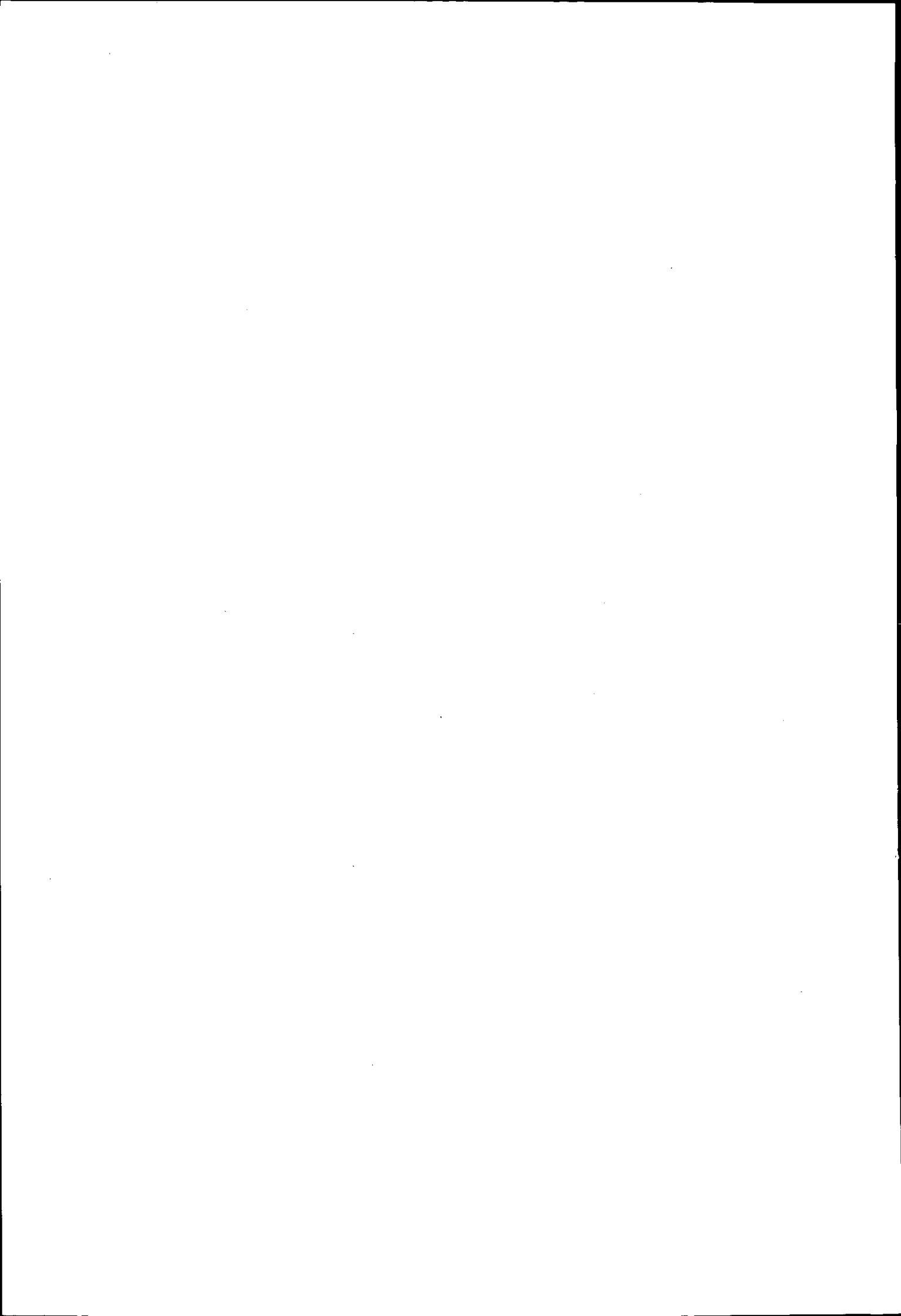
② 共有データベースのモデルシステム開発

共有データベースの構築のために、各機関がそれぞれどのようにデータベース構築に取り組んでいけばよいか、また、どの程度の予算、人の稼働を必要とするか、といった点を具体的に検討するために、特定機関において分散型のブランチとなるデータベースのモデルを開発する。

また、各機関相互のデータベースネットワークの方法について、パソコン通信、光ディスク、ファックス、VAN等のツールを活用したスイッチング機能を検討する。

8 東北型地域産業情報データベース・モデル
システムの拡張

社団法人 東北機械電子工業会



8 東北型地域産業情報データベース・モデルシステムの拡張

8.1 モデル・データベース（東北地域企業情報データベース：TRIP）内容の 拡充と情報検索システムの開発

8.1.1 モデル・データベース構築の目的と経緯

近年の情報関連技術の飛躍的向上等を背景に、これまで情報化への取り組みの相対的な立ち遅れが指摘されていた東北地域においても、域内各地で各種情報システムの開発検討が積極的に進められつつある。

こうした中で、今日の経済活動や社会生活活動の広域化に伴い、それらに係わる「情報活動」について東北地域をひとつのまとまりとして捉え、地域全体を包括したより広域的な情報ネットワーク形成がますます強く求められてきている。

殊に、今後、地域産業の自律ある発展を目指すためには、域内企業間の有機的結合や産学官の連携・促進を通じて技術開発基盤の強化あるいは域内外の市場拡大等を図ってゆくことが必要であり、その意味でも効果的な地域産業情報ネットワーク化が要請される。

本モデル・データベース（TRIP：Tohoku Region Industries' Promotion Data Base）は以上のような背景から

- ① 地域産業、地域企業間の有機的結合の支援
- ② 技術力高度化のための産学官連携の支援
- ③ 地域中小零細企業の情報化格差の是正と高度情報化の支援
- ④ 地域商工指導機関の指導力強化と相互連携の支援

を目途として、地域企業情報、技術開発情報、人材情報、施策情報等、東北地域産業振興に係わる総合的な情報を地域一元的に蓄積・提供する「東北地域産業情報データベースシステム」の構築を先導するモデル・システムとして地域企業情報のデータベース化を試みたものである。

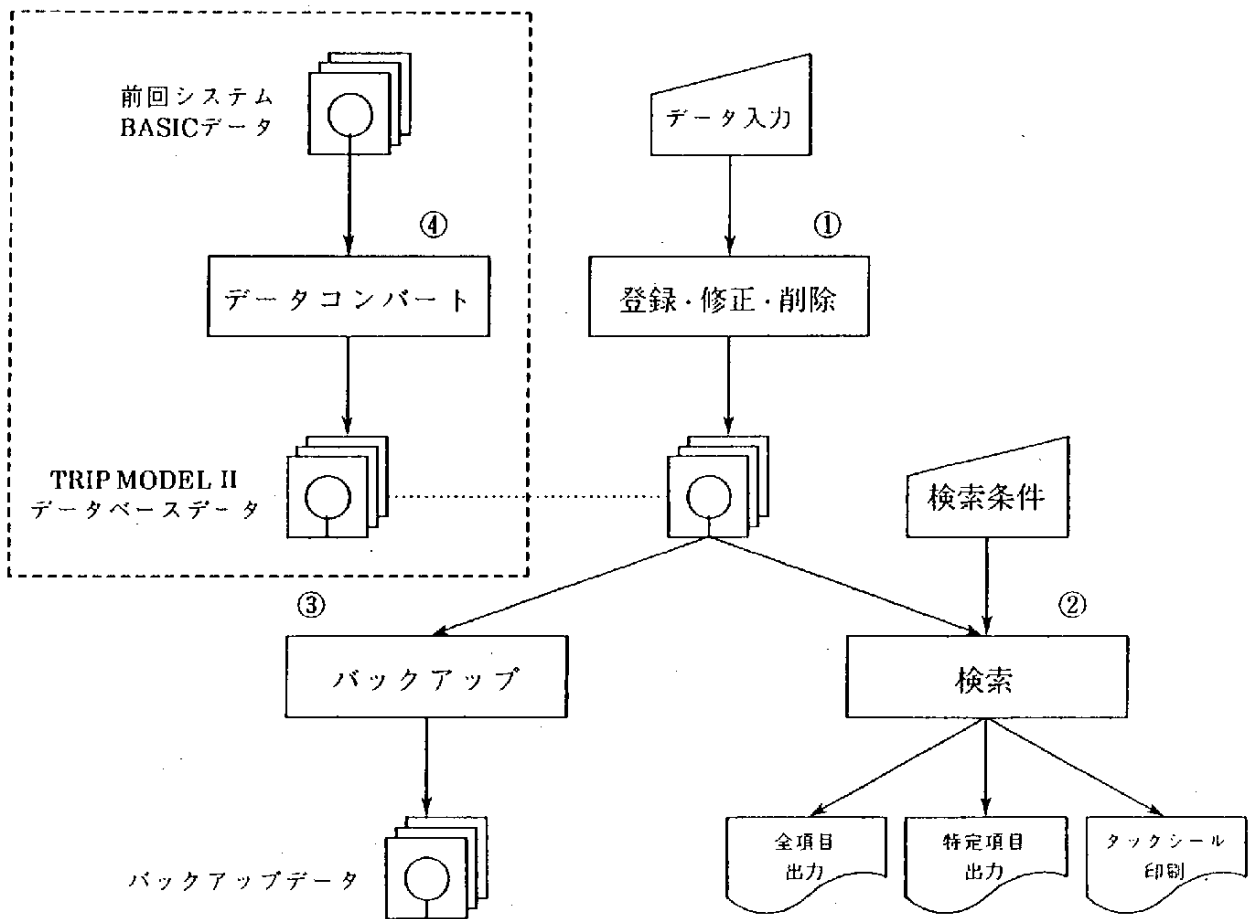
こうした考え方にに基づき、これまでに東北6県の約2,000社に係わる企業データをインプットしたTRIP MODEL-Iを構築したが、本年度はこれをベースに今後の具体的な運用に向けて、以下のとおり情報内容および機能の拡充を図った。

8.1.2 データ拡充とシステム拡張の内容

従来のデータ項目（企業概要，得意技術，主要製造品目，主要機械設備等 38 項目，768 バイト）に沿って，新たに約 600 企業のデータをインプットした。

また，システム機能面では，従来の TRIP MODEL - I で蓄積した企業データを生かし，より使いやすいデータベースとするため「検索機能」を充実した。

このモデルデータベース（TRIP MODEL - II）の全体フローを示すと図 8 - 1 のとおりである。



〔 〕内は前回システムから TRIP MODEL II への移行時に行う作業

図 8 - 1 モデルデータベース (TRIP MODEL-II) の全体フロー

TRIP MODEL-IIでは以下のファイルを使用した。

- | |
|--|
| ① 企業データファイル
企業情報の実データが登録されている。 |
| ② 企業データインデックスファイル
企業情報の更新のため、システム内部で使用する。 |
| ③ ワークファイル
企業情報の検索のため、システム内部で使用する。 |

また、プログラム構成は次のとおりである。

- ① メニュー (TK 2 1MN. PRG)
プログラムの選択実行を行う。
- ② 企業データ更新 (TK 2 2IN. PRG)
企業データの登録・修正・削除を行う。
- ③ 企業データ特定項目検索 (TK 2 3SC. PRG)
企業データの特定項目による検索を行う。
- ④ 企業データ任意項目検索 (TK 2 4SC. PRG)
企業データの任意項目による検索を行う。
- ⑤ 企業データ全項目出力 (TK 2 5PT. PRG)
企業データ特定・任意項目検索により検索したデータ的全項目出力を行う。
- ⑥ 企業データ特定項目出力 (TK 2 6PT. PRG)
企業データ特定・任意項目検索により検索したデータの特定項目出力を行う。
- ⑦ 企業データタックシール印刷 (TK 2 7PT. PRG)
企業データ特定・任意項目検索により検索したデータのタックシール印刷を行う。
- ⑧ 企業データバックアップ (TK 2 8BU. PRG) (ユーティリティプログラム)
企業データを別のフロッピーディスクにバックアップする。(データのコピーを行う)

8.2 モデル・データベースのシステム運用方策の検討

今回構築したモデル・データベースを運用するに際して当面解決すべき課題としては下記のようなものが挙げられる。

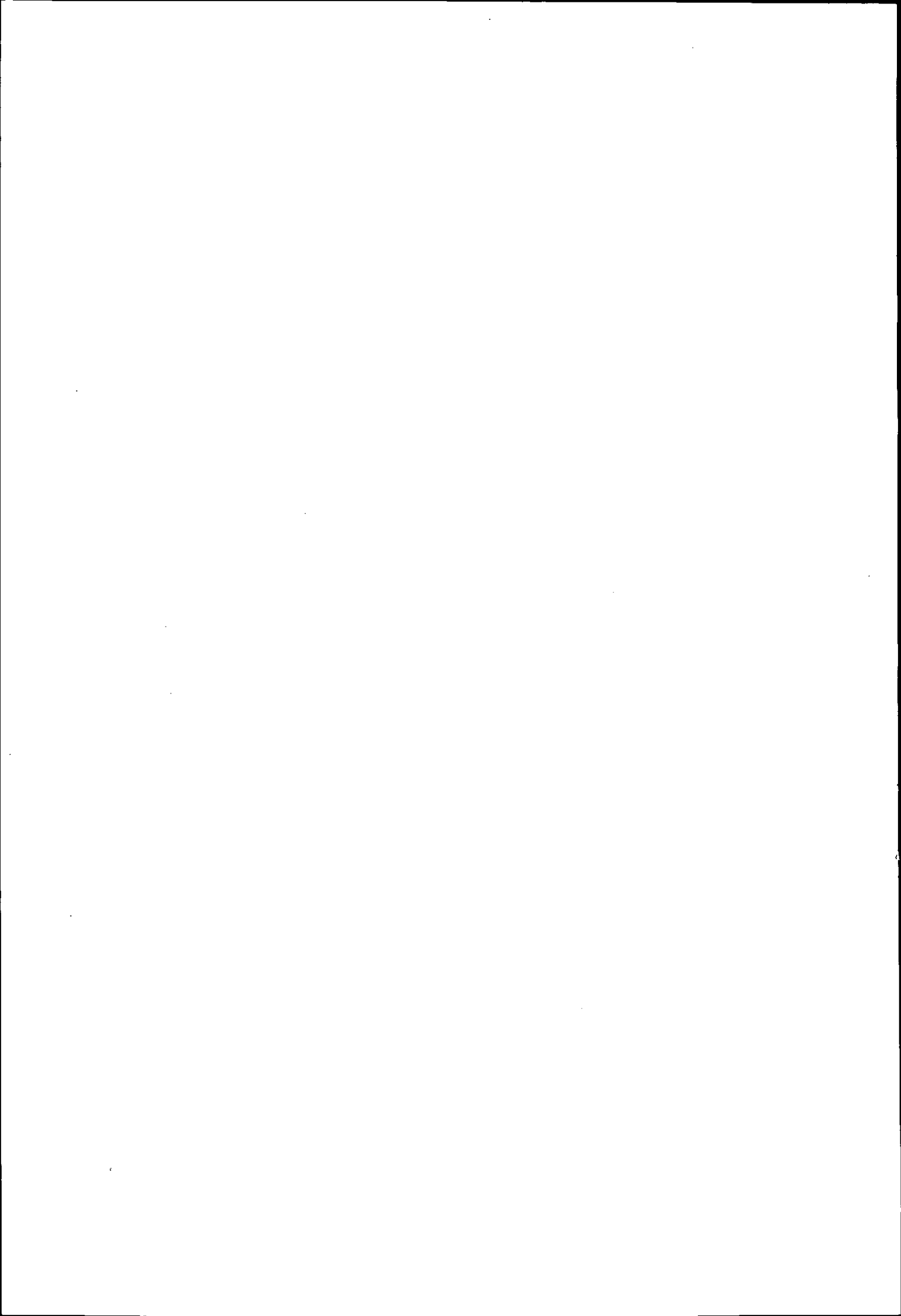
- ① 運用主体の確立
- ② 運用資金計画の設定
- ③ 地域への啓蒙，啓発
- ④ 人材の確保・育成
- ⑤ ソフトウェア・メンテナンス体制の整備
- ⑥ 企業秘密・プライバシーの保護

また，こうした課題を解決しつつ，中小企業関連団体等を通じ広く企業等へ情報を提供するためにオンライン・ネットワークの形成が必要であり，これにはVAN・パソコン通信ネットの有効利用を図ることが不可欠である。例えば，仙台コミネットの協力を要請することも検討に値しよう。

しかしながら，今後のモデルデータベース運用を展望した場合，その利用拡大や運用の健全性確保等の点では，単独の対応だけでなく他の地域データベース等との協同での対応が求められ，更に全国および国際的なデータベースへの容易なアクセス確保のためにも，東北地域独自の統括的な運用主体を構築し，利用ネットワークの形成・拡大に努めることが重要である。

9 郡山地域商工業活動データベース構築調査

ふくしまテレメディアサービス株式会社



9 郡山地域商工業活動データベース構築調査

9.1 データベース構築の意義

9.1.1 データベース構築の意義

郡山は人口307,745（昭和63年1月1日現在）、自然環境に恵まれた広大な市域と、交通の要衝として、地理的優位性に支えられ、南東北における産業経済の拠点都市として発展を続けている。

県の中央に位置する郡山の交通体系をみると、JR東北本線、磐越東線、磐越西線、水郡線が四方にのび、更に東北新幹線が走り、道路網も東京・青森を結ぶ国道4号線、いわき（太平洋）・新潟（日本海）を結ぶ国道49号線、更に国道288号線などがクロスする形で走っている。

高速道路網も、既に開通している東北自動車道のほかに東北横断自動車道（いわき・新潟線）の建設が進められているほか、将来の高速交通体系を展望するとき、昭和67年度開港を目ざす福島空港の整備も、郡山地域テクノポリス圏域内で進められている。

このような優位な条件の中にあって郡山市は、東北の中でも仙台市に次ぐ経済都市として商工業活動が活発に行われている。そして昭和61年12月3日に郡山市を母都市とする2市3町1村（郡山市・須賀川市・鏡石町・石川町・三春町・玉川村）で構成される地域が“郡山地域テクノポリス”の指定を受け、また昭和62年7月31日には、高度情報化時代をリードする都市としてニューメディア・コミュニティ構想応用発展地域の指定も受けた。郡山地域の発展をとげるための支援策も着々と進められてきている。

郡山市は、21世紀のまちづくりとして、近代福祉都市の創造をめざすため、郡山市第3次総合計画を査定している。この事業計画の中に、昭和63年度から市が持っている情報を市民に公開する制度がスタートする。情報化時代の中にあってその地方特有のデータが地域住民にサービスされるようになり、また外から来る人々にもサービス化されるようになってきつつある。

各種のデータをサービスする側は、データの材料選択はもとより、わかりやすく、使いやすいデータベースとして処理され加工されたものでなければならない。データベースの構築は、このようなことをふまえて行われるべきであり、地域データベースの開発は、これから積極的に推進されていくものと思われる。

データベースの構築は今までは中央での資料がデータベース化されてきたことから、中央で開発され、そして中央で使われてきたが、地方のデータもいよいよデータベース化される時代がきた。そして、地方のデータも中央で使われ、また中央のデータも地方で使われてきて、データベースの普及拡大がはかられてきている。これは、地方と中央との情報化の差がなくなり、高度情報化社会へと進んでいることを意味する。

経済県都として発展し続けている郡山には商・工業活動に関するデータがもっとも必要とされている。商業・工業に関するあらゆるデータが、県内はもとより、県外の人々からもすぐ取り出せる形で使用できるデータベースが求められている。

地方が持つ特有の資料を体系的にまとめあげたデータベース化は、使われるデータベースの構築につながる。高度情報化時代が進む中であって情報の多様化に応えたデータベースは地域の活性化を呼び、地域の発展を促すことになる。

9.1.2 ビデオテックスシステムによるデータベースサービスの特徴

ふくしまテレメディアサービス株式会社（以下F T S と略す）が、これから行うデータベースサービスは、ビデオテックスによるデータベースサービスで、その後ソフト、ハードの進展に合わせてサービスの方法を応用発展させ、そして将来パソコン通信等でもサービスを受けられるように発展させていく。

ビデオテックスの優位性は、電話と家庭用テレビという、どちらも一般家庭に普及率の高い電気製品を組み合わせでできているシステムで、初心者でも簡単に利用可能な情報検索機器として使用できることにある。すなわち、ビデオテックスは検索機能を持つ双方向のメディア媒体として、一般大衆の利用しやすいシステムであり、データベースサービスを簡単に行うことができる。

今までのデータベースは、プロフェッショナルな人々が使うシステムであった。端末操作により情報検索を可能にしてきたが、一般の人々は目的の情報を簡単に検索することが困難であった。そのためプロフェッショナルな使い方が主であり、少数の特定されたユーザだけにしか使われなかった。このことは、データベースの活用が一つの限られた人々によってしか使われず、利用頻度が少なく、またデータベースをサービスする企業化にも遅れをとってきた。これに対し、ビデオテックスは、大衆向け端末ということもあって、データベースの制作を最初からとってきており、文字と画像の組み合わせによる見やすく、わかりやすい情報提供を試みてきている。

またビデオテックスの媒体としての位置づけも、テレビ、ラジオ、新聞などのマスメデ

ィアに対し、パーソナルメディアでないかと言われてきているが、端末の普及が拡大し、端末のアクセス数が多くなった場合、また街頭端末の普及によりメモリーされた情報がエンドレスで何回でも大衆の目にとまるようになった時、マスメディアとしての潜在性も備えていることがわかる。また同時に、双方向性を利用する部分においては、パーソナルメディアとしての特性も持っており、メディアミックス時代の中であって、人々の生活も多様化している現在、大衆の求めている多様化の価値感に対応したキメ細かい情報のサービスはこれからの高度情報化時代に必要なことだと思う。大衆のニーズに応えたキメの細かいデータベースの構築は、更にわれわれビデオテックス業界の利用度を上げることになり、高度情報化をより一層進展させることにもつながる。

9.2 データベース構築の全体計画（郡山地域における商工業活動のデータベース化）

商工業に関するデータベース化についてデータベース委員会で話し合われた結果、各種資料のデータベース化に当たっては“使われるデータベース”の構築をめざさなければならないということが結論であった。どこにでもあるようなデータをまとめても利用価値がうすいのではないかとか、^④情動的なものをサービスしなければとか、いろいろ議論続出の感があった中で、各種データをどのような味付でユーザに利用してもらのかということであった。いわゆる一つのルールに従って体系化されたデータでなければならない。そして、これらを永続的にただ入力しておくのではなく、変化する新しい情報に対処してゆくものでなければならない。オンラインで刻々と入力していくようにして即時性があり、ニュースソース性を持たせた情報などが体系化されて入力されていなければならない。つきつめると、結局“使われるデータベース”の構築にたどりつくことになった。

このためには、今後ユーザ側にアンケート調査を行い、データベース構築に当たって企画から情報サービスまでの全体像をまとめることにした。データベース化するため、考えられる商工業活動の当面のデータ資料は次のとおりである。

・商業の概況

業種別データ

- ・卸売業（各種商品卸売業、繊維品卸売業、衣服、身の回り品卸売業、農畜産・水産物卸売業、食料・飲料卸売業、医薬品・化粧品卸売業、化学製品卸売業、鉱物・金属材料卸売業、機械器具卸売業、建築材料卸売業、家具

・ 建具・じゅう器卸売業，再生資源卸売業，その他の卸売業，代理商・仲立業）

・ 小売業（各種商品小売業，織物・衣服・身の回り品小売業，飲食料品小売業，自動車・自転車小売業，家具・建具・じゅう器小売業，その他の小売業）

商店数，従業員数，年間販売額，仲介手数料収入額，商品手持額，売場面積

◦ 商業分類別・規模別商店の状況

従業員数による区分別商店数

常時従業員数，売場面積，年間販売額，修理料・サービス料・仲立手数料の収入額，商品手持額

◦ 商業の推移

年次別データ

商店数（卸売・小売・飲食店）

従業者数（卸売・小売・飲食店）

年間販売額（卸売・小売・飲食店）

◦ 従業者規模別商店数

規模分類

1～4人，5～9人，10～29人，30～49人，50人以上

◦ 経営組織別商店数

法人，個人別の商店数と構成比

◦ 業種別売場面積

1店当たりの面積と商店数

業種別売場面積の構成比

◦ 飲食店の状況

業種別データ

商店数，従業員数，年間販売額，増減率

◦ 商業の状況

市内を地区別に分けたデータ

商店数，従業者，年間販売額，仲立手数料収入額，商品手持額，売場面積

◦ 全国および県内各市の推移

全国と県内を比較した商業データ

全国，福島県，郡山市，福島市，会津若松市，いわき市，白河市，原町市，須賀川市，

喜多方市，相馬市，二本松市

◦ 県内地区別の推移

地区別を比較した商業データ

県北地区，県中地区，県南地区，北会津地区，南会津地区，相双地区，いわき地区

◦ 工業の概況

業種別データ（食料，繊維，衣服，木材，家具，紙，出版，化学，石油，ゴム，皮革，
窯業，鉄鋼，非鉄，金属，機械，電気，輸送，精密，その他）

◦ 工業の概況および構成比

規模別・年度別

4～9人，10～19人，20～29人，30～49人，50～99人，100～299
人，300人以上

事業所数，従業員数，製造出荷額

◦ 産業別事業所数・従業者数

業種別データ（食料品製造業，繊維工業，衣服・その他の繊維製品製造業，木材・木
製品製造業，家具・装備品製造業，パルプ・紙・紙加工品製造業，出
版・印刷・同関連産業，化学工業，石油製品・石炭製品製造業，ゴム
製品製造業，なめしかわ・同製品毛皮製造業，窯業・土石製品製造業，
鉄鋼業，非鉄金属製造業，金属製品製造業，一般機械器具製造業，電
気機械器具製造業，輸送用機械器具製造業，精密機械器具製造業，そ
の他の製造業）

事業所数，従業者数，現金給与，原材料使用額，製造品出荷額

◦ 産業別支出額・製造品出荷額

支出額（現金給与総額・原材料使用額等総額）

製造品出荷額等（製造品出荷額・加工賃収入額・修理料収入額）

◦ 従業者規模別工業の現況

4～9人，10～19人，20～29人，30～49人，50～99人，100～299人，
300人以上

◦ 工業の推移

年度別データ

事業所数，従業者数，現金給与額，原材料等使用額，製造品出荷額等

◦ 地区別工業の現況

郡山，安積，三穂田，逢瀬，片平，喜久田，日和田，富久山，湖南，熱海，田村，西田，中田，

9.3 情報産業の現状と課題

9.3.1 企業における情報化の現状

郡山地域における汎用コンピュータの導入を業種別に算出して、全国数値と比較すると表9-1のようになる。(昭和62年調べ)

全国数値は昭和59年のもので、その後導入台数も当然増えていることから、単純比較は難かしいが、大体の傾向としては製造業，卸売業の利用が高く，建設業・サービス業の利用が低いことが判る。

また昭和59年に実施した「郡山市における情報化に関する調査」によると表9-2のようになっている。

表9-1 汎用コンピュータの導入状況

	郡山テクノポリス圏域内	全国平均値 (59.3)
製 造 業	43.7%	40.0%
建 設 業	3.1%	5.8%
卸 売 業	15.6%	12.3%
小 売 業	6.3%	4.8%
サービス業・その他	31.3%	37.0%

※ 郡山テクノポリス圏域とは郡山市，須賀川市，石川町，鏡石町，三春町，玉川村の2市3町1村の地域を指す。

表9-2 家庭・職場における情報通信機器の普及状況

	家庭にある	職場にある	家庭・職場に あ る	な い	無 回 答
ファクシミリ	2(0.8%)	51(21.5%)	—	118(49.8%)	66(27.8%)
文字多重放送用 テレビ	8(3.4%)	3(1.3%)	—	148(62.4%)	78(32.9%)
音声多重放送用 テレビ	59(24.9%)	13(5.5%)	—	104(43.9%)	61(25.7%)
パソコン	15(6.3%)	46(19.4%)	5(2.1%)	107(45.1%)	64(27.0%)
ワードプロセッサ		47(19.8%)	1(0.4%)	117(49.4%)	72(30.4%)

※「ない」場合にはほとんどの人が無回答であったので、無回答のほとんどは「ない」と考えてよい。

一方、情報化の現状を「ネットワーク視点」からみると、汎用コンピュータの75%がオンライン化されており、特に製造業・サービス業のオンライン化傾向が強い。

ネットワーク範囲は「本店、支店」間が多く、次いで「販売先」「仕入先」となっている。

情報内容としては「受発注情報」「販売情報」「請求、決済情報」が多く、今後主販売先との系列化も予想される。

ハードウェア的には、ファクシミリによる情報通信が有効活用されており、導入率も60%強と情報機器中では一番企業に浸透している。

コンピュータオンライン網との補完関係としてファクシミリ網が存在するとも考えられる。

9.3.2 企業における情報への対応

現在、入手している情報内容としては、全業種とも「業界・行政関係」と「経営関係」が多く、製造業では「技術関係」の需要が多い。

情報の拡がりを期するためには、外部データベースとのゲートウェイ機能を果たす機関が必要と思われる。

また、データバンクの活用意欲は高く、これらデータバンクの情報源情報を整備すれば、内容、利用頻度とも向上していくと思われる。

情報入手時、有償に対する反応は、有償でも欲しいという意向が強い。

9.4 F T S ネットワーク機構の現況と展望

9.4.1 F T S の会社概況と設立経過

ふくしまテレメディアサービス株式会社（略称：F T S ）は、郡山地域においてキャプテン方式による情報サービス事業を目的に昭和 6 1 年 1 月 3 1 日設立された。

郡山地区では、それまで郡山商工会議所を中心とした「郡山高度情報化システムの研究会」（会員約 1 2 0 社）において各種ニューメディアの研究が行われ、キャプテンシステムはキャプテン分科会で検討が続けられた。その中で、6 0 年 7 月以降約 8 回にわたる分科会の結果、N T S 方式によるキャプテンシステムが郡山地域にとって適切であると評価、日本テレメディア・サービス株式会社（略称：N T S ）の地方会社第 1 号として設立されたもの。いわば“子会社方式”によるプライベートキャプテンである。そして F T S は同年 7 月、試験放送を開始、同 9 月 1 日から営業を伴う本放送サービスをスタートさせた。端末利用は会員方式を採り、利用者負担は入会金（工事費、全国キャプテン登録料などを含む）15,000 円と月々の会費 1,500 円となっているが、試験・サービス放送の 2 カ月間は会費は無料とし、その間、情報内容の充実と端末の普及をはかり、本放送に踏み切った経過がある。

9.4.2 端末設置状況と事業展開

F T S は 6 1 年 9 月の本放送サービス開始時の端末利用契約台数は約 3 7 0 台となり、チャンネルメニューも 6 0 項目を超えた。現在の端末数は街頭用の大型 6 台を含めてざっと 7 0 0 台。設置か所の内訳は一般家庭が約 6 0 %、事業所・公共施設約 4 0 % の割合。大半が郡山市内に集中しており、郡山市役所では、本庁と各支所・公民館など市の公共施設に計 6 0 台を設置、市民の関心を集めている。これら端末普及は、当初、株主の事業所およびその従業員の家庭を対象として契約を結び、続いて一般への浸透をはかった。

普及作戦の具体例としては、6 2 年 4 月の統一地方選に際し県議選（郡山市選挙区）と郡山市議選で、各候補者の選挙事務所に端末を無償貸与、選管発表ごとの開票速報をサービス、また「ふくしま輸入博覧会」（6 1、6 2 年春）の会場にキャプテンコーナーを開設、郡山市の消費生活展でも街頭型を展示したほか市内のスーパー、デパートに常時、キャプテンコーナー（端末 1 0 台程度）を設けるなどキャプテンに対する地域住民の関心の掘り起こしに努めている。

9.4.3 チャンネルメニューと検索システム

F T S が提供する情報はニュース、天気予報、公共情報、ショッピング情報、株価情報など20チャンネルで、メニューはざっと100項目、画面数約1万。メニュー全体の50%強はF T S 独自の制作、残り50%弱が親会社N T S からの配信によるものである。このうちN T S のメニューは、読売新聞社、報知新聞社提供のトピック・総合ニュース、株価情報、一般経済情報など全国レベルの情報。これに対しF T S は、郡山市政広報を主体とする公共情報をはじめショッピング情報、タウンガイド、スポーツ、更に専門チャンネルの事業所案内、経営者情報など地元の新鮮な情報やデータベース的な情報をキメ細かく扱っている。

一方、情報画面の検索システムは、全国キャプテンと同じツリー方式と、N T S 独自のバック方式を併用しているのが特色。このN T S のバック方式は、情報を約40画面ごとにバック化して、これを端末機内のメモリーに記憶させてから電話回線を切断（3分間で自動的に切れる）することにより、回線利用料金を節減するとともに、記憶された各画面は何回でも繰り返し見ることができるといった利点を持つ。ツリー方式のように、一画面ずつ検索するわずらわしさがはぶけるうえ、回線利用料金が安いことから、この点を最大ポイントにP R を続けてきた。

情報サービス時間は、午前7時（冬期間は午前9時）～午後10時。情報画面の利用（アクセス）数は、株価情報がたえずトップを占め、以下トピックニュース、ゲーム、天気予報、総合ニュースの順（63年2月現在）となっている。

9.4.4 F T S の配信システム

F T S の情報提供システムは、N T S の中央センターと、郡山市内のF T S ローカルセンターを専用回線で結び、電話網を通じて家庭や事業所の端末に情報が送られる仕組み。東京にあるN T S の中央センターは、独自で情報を制作入力している他、外部情報センターに接続して情報提供しているものもある。例えば、読売総合ニュース等は読売新聞社のニューメディア編集配信システムから入力され、また、株価情報は東京証券取引所から、天気予報は気象協会からという具合にオンラインでN T S のC P U に入力され、その後地方（F T S ）に配信されサービスされるシステムになっている。

一方、両センターは、ビデオテックス通信網によってキャプテン情報センターとも結ばれている。このため、F T S の端末（N T S 式）からはF T S メニューのほかに全国キャプテンの豊富な情報を自由に見られるというメリットがある。

9.4.5 FTSのネットワーク構想

F T S は、全国ネット化を目指したN T S の地方会社第 1 号として設立、その後、長野、仙台などにもN T S 系列のローカルキャプテン会社が誕生した（長野は休止）。こうした事情を背景に、F T S はN T S と業務提携し、いわゆる“子会社方式”を採用、親会社からは端末リースのほかソフトウェアと共通する全国情報の供給を受けている。

F T S の端末設置は、現在郡山市内が大半を占めているものの、市外にも順次普及しつつあり、実質的なサービス範囲は県内全域に及んでいる。このため、F T S は将来、福島、いわき、会津若松の三市内にF T S 直営の「地域会社」を設立、運営する計画。県内の拠点であるこの三市内にアクセスポイントを設け、各地域ごとにアクセスすることにより、通信料が安くなるといった効果を狙ったものである。

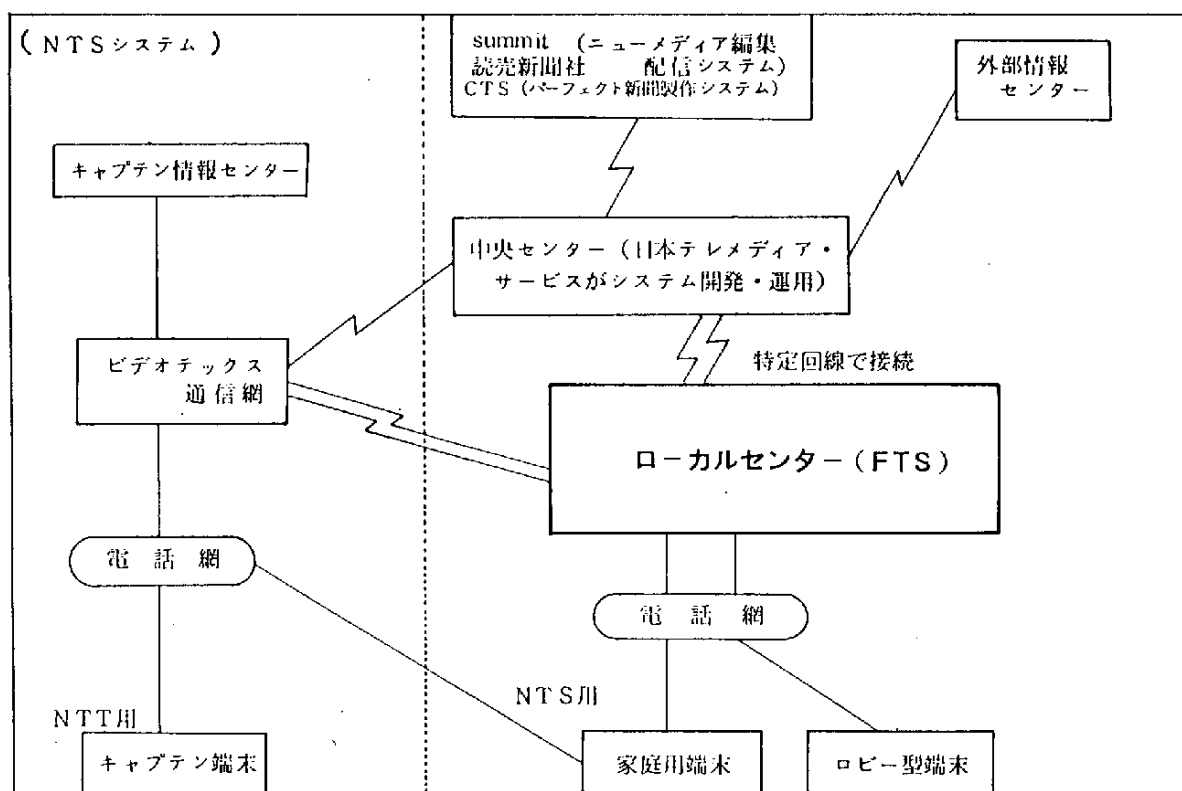


図 9 - 1 F T S ネットワーク構成図

9.5 調査用モデル画面の構築（モデル画面の制作に当って）

郡山地域の商工業活動をデータでとらえるデータベース化については、“使われるデータベース”を目指すことにし、調査モデル画面を制作、CPUに入力して、ユーザの批判をおぐこととした。キャプテンシステムによる情報の提供には、魅力ある画面展開を伴ったサービスメニューの開発や、データベース構築の設計などが、重要なポイントになってくる。

情報を的確に伝え、利用者の立場に立ったキャプテン情報の設計が強く望まれている現在、キャプテンの持つ特色を考えた場合、文字情報だけではなく、画像情報も考慮に入れた視覚に訴えるものもいれなければならない。そして、これ等の一体化された情報を制作するとき、情報量や内容を切り離して考えることはできない。

制作する側は、常に利用者にとって使い易い検索方法とは何か、適切な情報の量や内容とは、魅力ある画面のイメージとは何か、を追求してキャプテン画像の制作に当たらなければならない。

以上のことを考慮に入れ、今後のデータベース化を目指す。このためには、たたき台になる調査モデル画面が必要であり、データの種類によって、それぞれの企画・構成を行った。

(1) 文字情報でサービスするもの

□郡山市の事業所案内

市内の事業所を業種別に分け、会社概要をデータベース化する。

分類の方法は、製造業・建設業・卸売業・小売業・サービス業・その他に分け、これを細分化しその後バック化して情報を得るようにした。

（内容）

会社名・所在地・代表者名・創業年度・法人設立年度・資本金・従業員数

□郡山市内の証券会社案内

市内の証券会社一覧のあとに、各社ごとに詳細を説明。

（内容）

会社名・創業年度・法人設立年度・資本金・従業員数・取引銀行・会社の特色

(2) 文字+図形情報でサービスするもの

□郡山地域テクノポリスガイド

郡山地域テクノポリス推進機構についての案内。

（内容）

機構の概要・役員案内・業務案内。

(3) 文字+グラフ情報でサービスするもの

□郡山の産業パワー

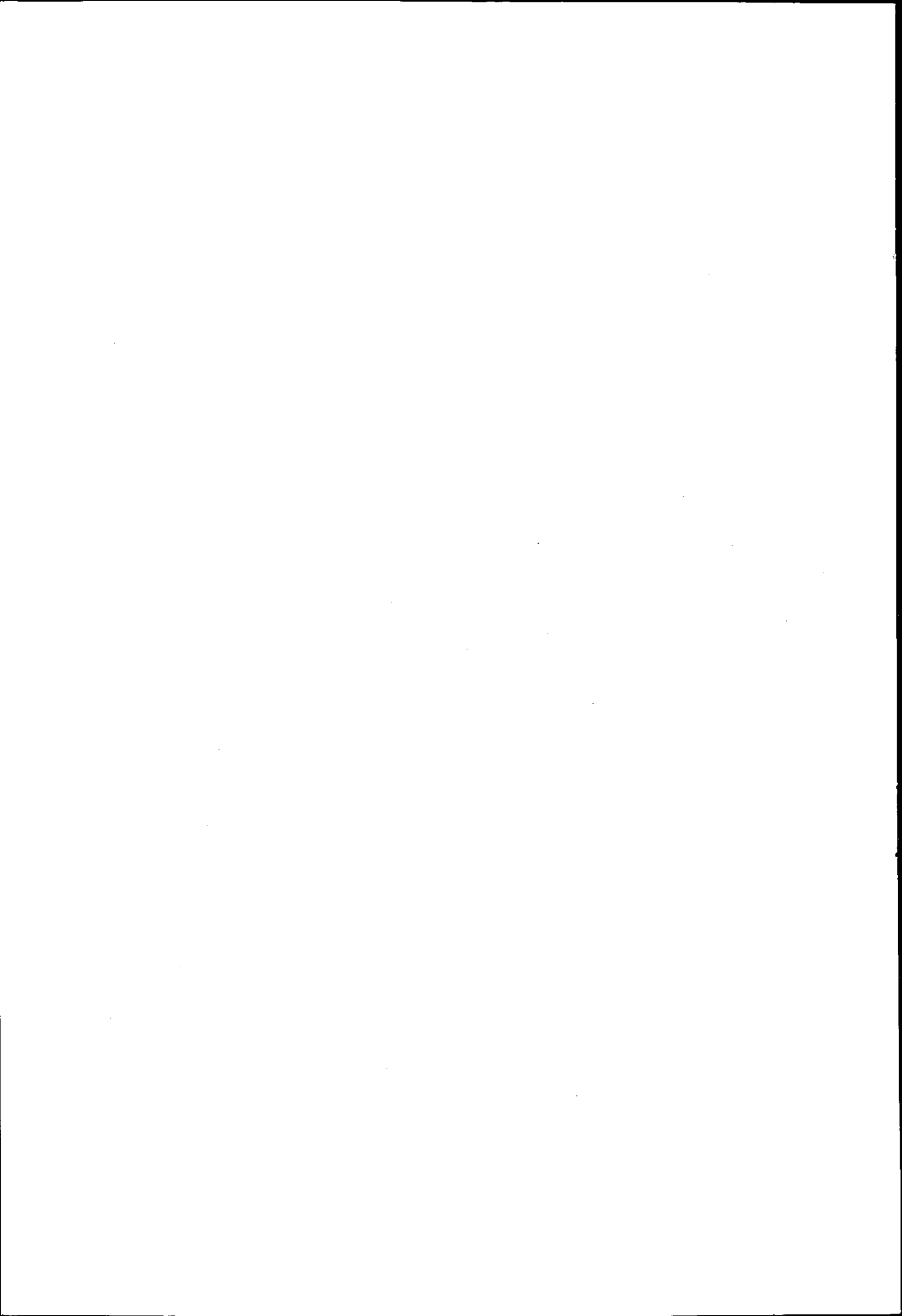
商圏人口・人口動態・事業所の推移

グラフで見る商業（事業所数・年間販売額・従業員数）

グラフで見る工業（事業所数・従業員数・製造品出荷額）

10 世界鉄鋼企業財務データベース

社団法人 日本鉄鋼連盟



10 世界鉄鋼企業財務データベース

10.1 目的・背景

鉄鋼業に関する政策課題の検討に際して、最近では内外の経済・通商情勢の変化を反映して、主要国鉄鋼業の比較分析のニーズが急速に高まっている。

即ち、日・米・欧の主要先進製鉄国は依然として世界の鉄鋼生産基地ではあるものの、需給のアンバランスから過剰能力の処理等の構造改善に直面している一方、NICSおよび共産圏等の鉄鋼業の急進展が世界の鉄鋼需給に大きな影響を与える状況になっている。

このような状況に対して、従来から行われてきた世界の鉄鋼需給に関する分析にとどまらず、設備投資動向、企業収益・財務体質、コスト競争力等の鉄鋼業の経営状況に関する分析が要請され、当連盟の調査スタッフの情報収集および分析能力をフルに活用して対応してきたところである。

データ整備に関しても、このような背景から「鉄鋼情報システム(SIS)」において鉄鋼の需給・貿易に関する分野を中心に海外データの収集・整備に力を注いできたが、今回のテーマである「世界鉄鋼企業財務データベース」の開発が残された課題とされてきたところである。

本事業は、内外の鉄鋼主要企業に関する財務および企業経営動向に関する財務データ等を広く収集し、データベースとして整備するとともに、その利用体制を充実することによって、鉄鋼業界内、関係官庁、関連業界のニーズに応えることを目的とするものである。

10.2 実施内容

当連盟では、国内の主要鉄鋼会社について有価証券報告書および営業報告書を収集、保管するとともに、「高炉6社決算状況」、「鉄鋼会社収益状況」および「鉄鋼会社資産・負債状況」を業務統計資料としてパソコンで集計している。また、海外の主要鉄鋼会社のアニュアル・レポートを収集、保管し、内外鉄鋼業の収益状況、財政状態の比較・分析等に活用している。

とりわけ、国際化を背景とした最近の国際経済・通商問題等に関して、内外鉄鋼業の比較

分析の重要性が高まっており、収益状況、財務特質、コスト競争力等の分析にこれらの財務データの利用が活発化しているものの、パソコンによる処理能力には限界があり、データベース化が待望されていたところである。

一方、当連盟では生産、受注、貿易、原料、労働等に関する鉄鋼統計について、集計作業からの一貫処理によるデータベース「鉄鋼情報システム（S I S）」を開発・運用してきた実績がある。

このような状況を勘案して、これまでのデータベースの構築および財務データの収集・利用の実績を生かして、「世界鉄鋼企業財務データベース」の整備を以下のとおり行った。

10.2.1 データの作成

(1) 対象会社およびソースデータ

内外の主要鉄鋼企業の財務データ等および企業業績評価等に必要な企業動向を中心とした会社概要データを対象とした。

・対象会社：本事業では次の公開情報によるデータ整備を行った。今後各種調査資料等によって情報公開を行っていない鉄鋼会社について逐次範囲を拡大することとしたい。

- ① 国内鉄鋼会社：当連盟および普通鋼電炉工業会の主要会員会社（45社）の有価証券報告書
- ② 海外鉄鋼会社：当連盟が収集済みの海外主要鉄鋼会社（11か国、23社）アニュアル・レポート

(2) データ項目：

(イ) 財務データ等

貸借対照表、損益計算書、製造原価明細書等および生産・販売等の営業データ並びにセグメント情報（アメリカのみ）

(ロ) 会社概要データ

事業分野、主要事業所、沿革、会社の現況等の企業動向に関する情報

(3) データ収録期間

1970年度～1986年度

10.2.2 情報処理の内容

内外財務データの一括管理という特性およびホストコンピュータの移設という事情を配

慮し、パソコン（IBM5550、財務データはマルチプラン、会社概要はワープロソフト）によるオフラインのデータエントリーを採用したが、財務データに精通した専門担当者による直接作業が可能となった。

なお、情報処理の手順を示せば、概略図10-1のとおりである。

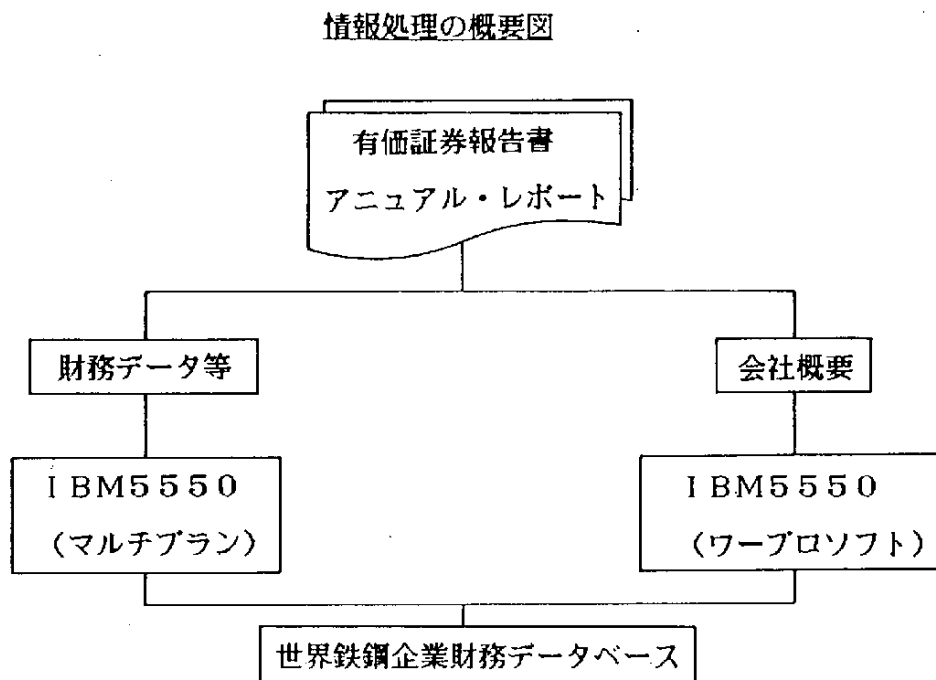


図10-1 情報処理の手順

10.2.3 データベースの利用

本事業においては、データベースの検索方法としてAS（汎用データ利用ソフトウェア）を利用することとしている。（ただし、当面は財務データ等に限定）

ASの機能面の特徴は、データ検索、データ解析・予測、表・グラフの作成やASの簡易言語によるプログラム開発等があげられる。また、操作面では、初心者から熟練者までレベルに応じた利用が可能である点があげられる。

本事業に則していえば、次のようなテストを実施して所要の結果を得ている。なお、テストデータにより機能テストの事例（グラフ）を示せば、以下のとおりである。

(1) 機能テストの内容

- (イ) データエントリーがマルチプランによる方法で支障がないか。
- (ロ) AS機能面の特徴であるデータ検索、データ解析、および表・グラフの作成が可能か。

即ち、会社別の財務データ項目の年次毎の検索、数社計や財務比率の計算、加工データの作成、分かり易い作表・作図等について、自由な出力が可能か。

(イ) コンピュータの専門知識を持たない業務担当者でも利用できる程度に操作方法が簡単か。

(2) 財務データ等の検索手順

財務データ（数字情報）については、ASによるデータの検索、データ解析、表・グラフの作成を行う。

ASによるデータの検索、データ解析、表・グラフの作成手順を示すと、以下のとおりである。

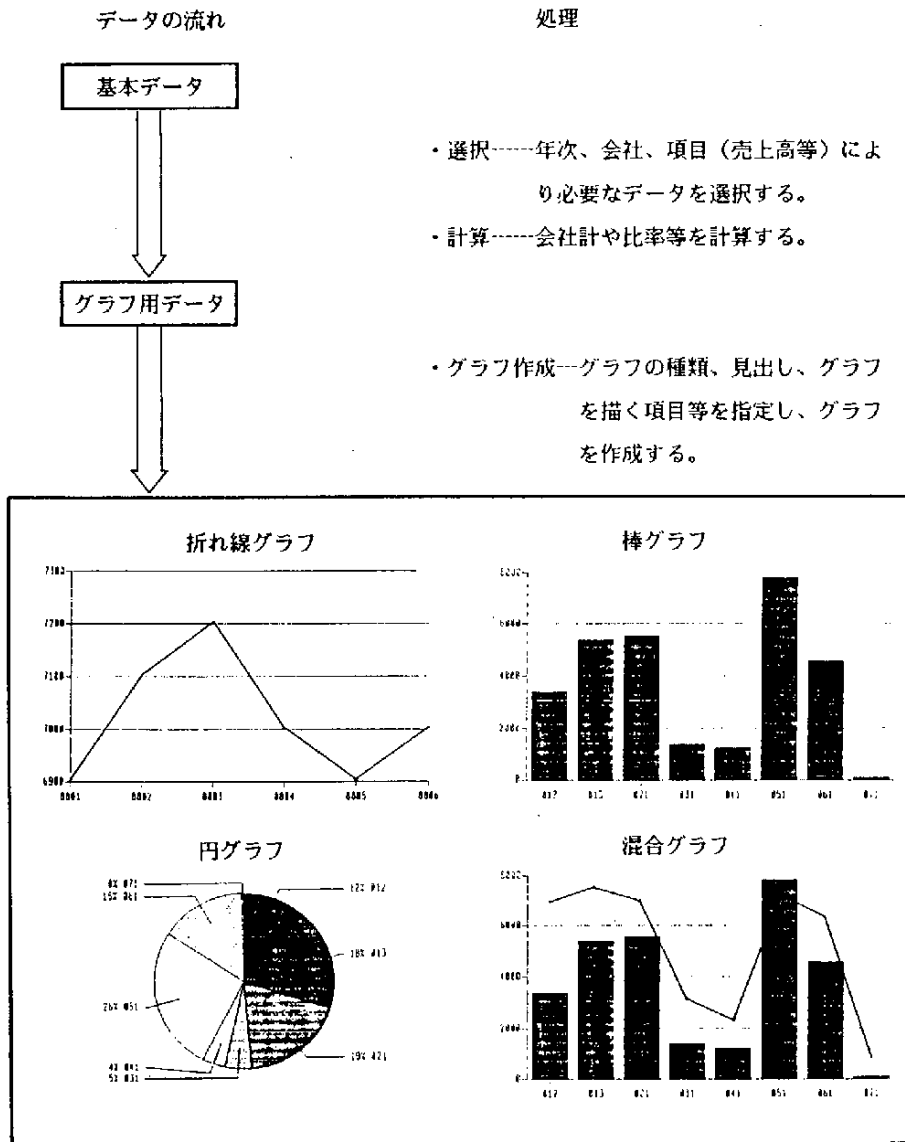


図 10 - 2 ASによるデータの検索、データ解析、表・グラフの作成手順

(3) サンプルによる出力事例

前述の検索手順に沿って、高炉5社をサンプルにデータの検索、データ解析およびグラフ作成の機能テストを行った事例を示すと以下のとおりであり、その有効性を確認した。

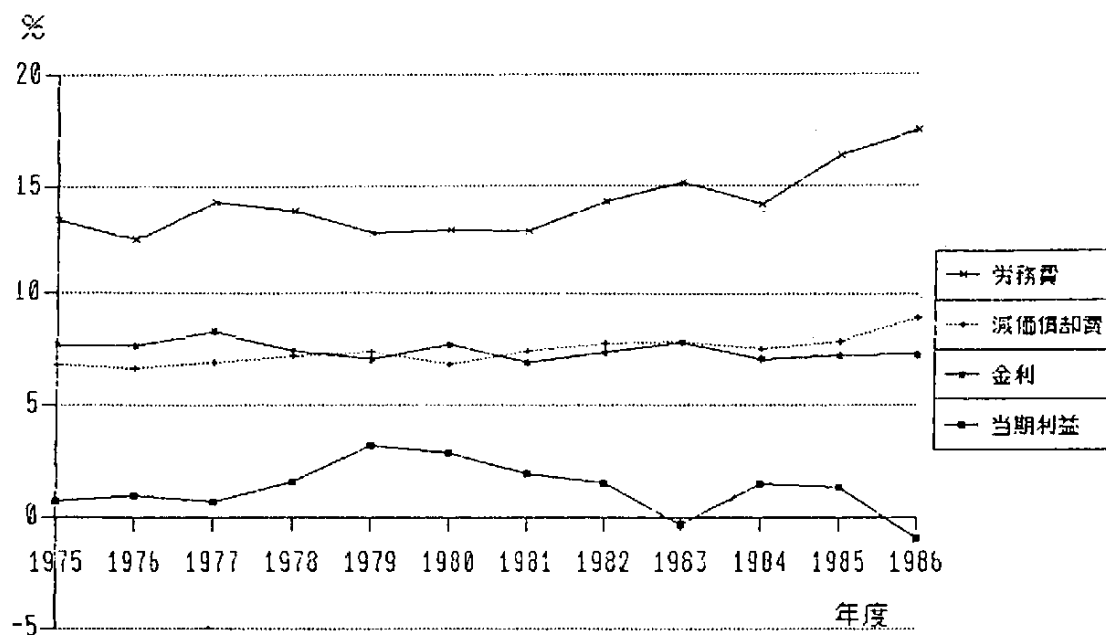


図10-3 機能テストの事例(売上高に対する主要コストおよび利益の推移)

10.3 今後の課題と展望

10.3.1 データの追加・補充

- (1) 本事業では、ソースデータの制約(有価証券報告書またはアニュアル・レポートに限定)、ソースデータの欠落、合併・分離および会計制度の変更等によるデータの不連続により、会社または年次によりデータの作成が不可能な部分があったため、今後可能な限り追加・補充の必要がある。
- (2) 今回はソースデータを公表データである有価証券報告書またはアニュアル・レポート

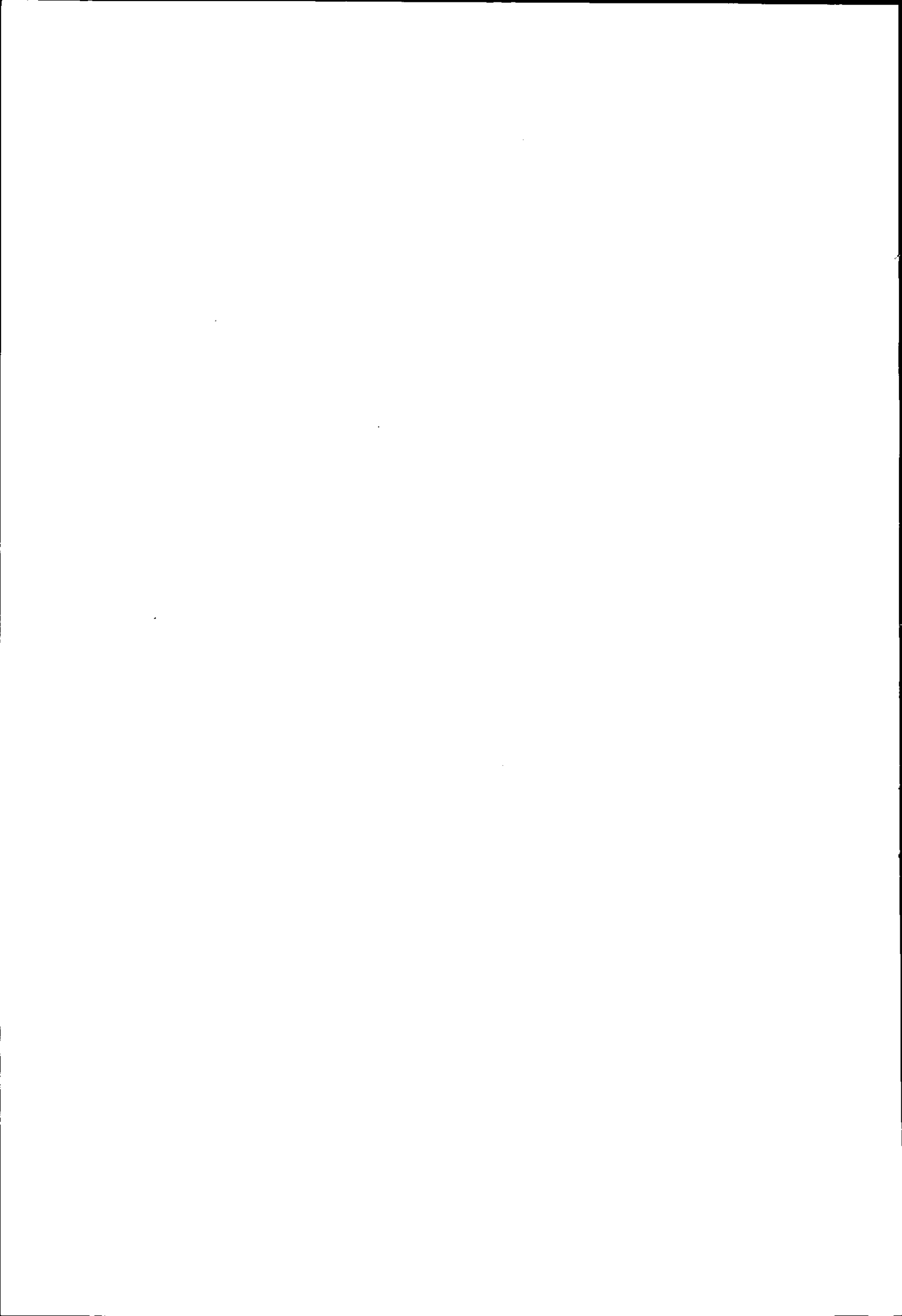
に限定したが、今後各種調査資料等によって情報開示を行っていない鉄鋼会社について逐次範囲を拡大することが必要である。

10.3.2 データベース利用の体制整備

今回は、財務データ等の数字情報については、ASの利用によるデータ検索、データ解析、表・グラフの作成が可能な体制ができたが、会社概要データ（文字情報）の検索を含め、総合的な検索利用体制の整備について、今後の検討課題として引き続き取り組んでいく必要がある。

11 データベース構築を前提とした印刷情報
の展開に関する調査研究

社団法人 日本印刷産業連合会



11 データベース構築を前提とした印刷情報の展開に関する調査研究

11.1 調査研究の経緯

11.1.1 調査研究の背景と目的

高度情報化の進展に伴い、ビデオテックスやCD-ROM等いわゆるニューメディアを用いた情報の利用・普及の裾野が広がりつつあり、「静的」な情報提供手段である印刷物では満たしきれない機能に対するニーズが著しく増加してきている。ニューメディア等の情報資源をいかに有効に利用、活用していくかに関心が集まる中、その扱う情報を収集・蓄積・管理・再利用するための有効な手段としてのデータベースが注目されているが、その構築には膨大な作業量・コスト・ノウハウが必要とされている。産業界において、このニーズに答える担い手をデータベースに輸入される大量のデータを取り扱う立場から探っていくと、情報の表現・加工に関して多くのノウハウを蓄積してきた印刷産業界が見出されてくる。こうした周囲の環境から、印刷情報を入力源とし複数メディア向け情報加工を前提とした「汎用データベース」の構築の検討が開始された。

印刷産業の持つポテンシャルに着目したとき、「汎用データベース」の基礎となるのはCTS (Computerized Typesetting System) 技術であるが、これはあくまで印刷工程の電子化という目的で構築されたものである。従って、原データを独立させデータベースとして取扱ったり、ニューメディア向けに加工するという面ではまだ工夫が必要とされる。

本年度は、昨年度の調査・研究結果を踏まえて、印刷物作成とデータベースの統合処理による汎用データベースシステムの実施設計と試用・運用に関する調査研究を行った。主な作業内容として昨年度の概念設計に基づいて、

- ・ モデルタイプ・システムの基本設計
- ・ 机上シミュレーション
- ・ モデルタイプ・システムの評価
- ・ 環境・周辺動向、影響要因の検討、モデルの見直し・拡張

を実施し、実システム構築のための知見を得ることを目的とした。

11.1.2 調査研究に当たっての前提条件

本調査研究では、以下に示す前提を置いて調査・研究を進めた。

- ・ 印刷情報の特徴・構成を踏まえたものとする
- ・ 汎用データベース化に際しての問題点を浮き彫りにする。
- ・ 一般的なデータベース構築との共通点、相違点を明らかにする。
- ・ 図形、画像データ内部の操作に重きを置かない。
- ・ 印刷情報を原始情報とすることがメインであるが、他のデータを元データとすることも妨げない。
- ・ C T S データには様々なものがあるが、本調査ではそれらをモデル化する。
- ・ 出力側のメディアをビデオテックス、テレテキスト、オンライン情報サービスおよびC D - R O M等いくつか事前に想定する。
- ・ 採算性や著作権問題等は汎用データベースを実際のビジネスとして展開する際の問題として位置づける。

11.1.3 調査研究内容

開発、実現レベルのフェーズ化を考慮し、実用化に向けて有益な知見を得るため机上シミュレーションを実施した。具体的な問題に対しては、委員会、および小委員会を設けての検討を踏まえて対処するものとし、システムの事例、シミュレーションのデータ等に関しては各委員の協力の下、収集・整理作業がなされた。

11.2 印刷情報の特色と汎用データベースの位置付け

11.2.1 印刷情報の特色

印刷情報の特色を明確化するため、その中心構成要素であるC T S 情報を検討した。

C T S 情報の内容は、

- ・ 内容を表す文字列であり、文字コード自身からなるテキスト情報
- ・ 印刷物としての文字の見えかた（属性）、印刷物上での文字の位置、文字間の関係（レイアウト）を表す（パラメータ）と考えられる制御情報

に大別される。

次に、C T S 情報の特徴を踏まえ、C T S 情報の処理上の特徴と問題点を検討した。

その結果、汎用データベースの構築に当たってはCTSデータの多義・多様性を考慮していく必要があると考えられた。CTSを利用し、印刷物を得るまでの作業形態はデータベース化の流れに則して、出版物作成を中心目的とした作業形態から、データベース作成を意識した形態に変わってきている。こうした作成環境の整備に加えて、CTS自身にもデータベース作成を支援する機能が実現されつつある。しかしながら、CTS情報を集積しただけではデータベースの構築は実現されない。これは、CTS情報においては、データベースとの比較でみたときの「構造を管理する」という視点がないことによる。データベース化において重要な役割を果たす、原情報に関する論理構造などの意味的な情報は、特別な処理を施さない限り入力時には把握されにくく、データベースとして利用していく上で、現在の処理作業に含まれない意味的な情報をどのような形で取り出し・付加していくかが課題となってきた。

汎用データベースの実現イメージとしては、画像、図形も含めたマルチメディアと複数のニューメディアを対象とするマルチメディアといった「2つのマルチメディア的視点」からの構築イメージがある。その果たす役割として、

- ・ 印刷物固有の情報の分離・蓄積
- ・ メディア向け情報の抽出
- ・ 各メディア向け情報の作成
- ・ 検索しやすい形式

等が洗いだされた。

11.2.2 汎用データベースの構成

汎用データベースのコンセプトは以下のように整理できる。

(a) 格納本体

文章における節、段落といったなんらかの「情報のくくり、論理操作の最小単位」にCTSデータを切出し、格納する。

(b) 管理情報

本汎用データベースはCTS情報自体にデータベース的な利用を実現する。即ち、CTSデータに論理的な管理情報を格納する構造を取り入れた。論理関係構造、レイアウト、書誌・権利、作成・運用履歴、数値化、意味といった視点から管理情報を必要に応じて格納する。

(c) 運用支援情報

汎用データベースを管理運用するために必要な情報を運用支援情報として用意する。
更に、汎用印刷データベースマネジメントシステムにおける基本機能として、登録、変換、属性変更、メディア向けの変換、編集・加工、検索、運用・管理等を設定した。

11.3 モデルタイプ・システムの基本設計

11.3.1 システムの要件および設計方針

システムの要件および設計方針として、マルチメディアへの対応、標準化への対応、原情報独立、統合化、即応性、分散化・ネットワーク化への対応、再利用性、使いやすさ、変更容易性、セキュリティの実現等の項目を考慮した。更に、その他の観点として、汎用性、柔軟性、効率性、理解のしやすさ、可搬性、拡張性、信頼性、網羅性から必要なものを考慮しつつ、上記の要件を満足すべく設計を進めた。

11.3.2 印刷情報入力機能

入力するC T S データの選定においては、モデルタイプ・システムの性能、達成上の問題点等、実用化に対する知見を得る目的に照らして、表現形態、内容が一般的、網羅的で、かつ汎用的なものを選択するのが望ましいと考えられた。従って具体的選定の方法としては、構成、内容から印刷物を以下の3種類に分類し、各々の代表的サンプルを選択してデータベース化を試みた。

- ・ 論文系 (論文集、学界雑誌等、章節等の明確な区分が可能なもの)
- ・ 便覧系 (便覧、書誌情報、カタログ、辞典、事典、図鑑等見出しとそれに対する内容の説明から成り立つもの)
- ・ 統計系 (数値情報が構成内容の中心になる、統計表、時刻表、電話帳。ただし F A C T 情報・データベース化済みの情報は除外)

C T S 情報の汎用化により、汎用データベースにおける構造・表現形式を考えた。データベースに格納する際のC T S 情報からの抽出結果として、基本オブジェクト、論理構造、管理情報を想定し、切出し操作、マルチメディアへ展開するための情報付加、検索しやすい形式での蓄積等の入力処理を検討した。

入力データに関して、原情報としてのC T S 情報中の各制御パラメータは、目的や重点の置き方によって汎用データベースに残る場合と残らない場合が生じる。また、データベ

ースに登録される際の処理形態も異なる。

これらのことより、原情報としてのCTS情報を汎用データベースに格納するに際して想定される処理としては、

- ・ 追加情報を含めた情報加工を必要とする。
- ・ 値の対応テーブルを参照する程度の情報加工を必要とする。
- ・ そのままデータベースへ格納する。
- ・ データベースへは格納しない。

等の処理を組み合わせた操作が必要とされた。

11.3.3 データベース管理・運用支援機能

データベースを管理・運用支援する機能については、DBMS（データベース・マネジメント・システム）的視点から機能の洗いだしを行い、データ入力・登録支援、論理格納構造の定義、内容操作、編集・加工、検索、運用・管理等について機能・処理の具体化を行った。

11.3.4 メディア展開・出力機能

メディア展開・出力機能とは、登録格納されているデータを印刷物、ビデオテックス、CD-ROMといった各メディア向けに情報の追加、属性変更組み替えといった操作を施し、外部ファイルの形式で変更出力することを指す。出力されたファイルは個々のメディアにおける応用プログラムの入力ファイルとなる。

各メディアで独自に必要な操作処理は、このファイルをもとに各々必要な変換処理を行っている。各メディア向け加工の工程においては、

- ・ 印刷固有の情報分離を支援する。
- ・ 各メディア（アプリケーション）用の情報抽出を支援する。
- ・ 各メディア用情報（特に共通的な情報）の追加を支援する。

等の作業を中心に進めることとした。

11.3.5 データベースの構造と既存文章モデルとの関連

データベースの論理モデルを決定するに当たって、ODA（Office Document Architecture）、SGML（Standard Generarized Markup Language）等の文書標準化に関する動向・個別プロトコルを参照・検討した。これは、情報の持つ伝送、交換される

という性質から、特に互換性、インタフェースの確保という目的で標準化が重要視されているという背景による。これらの文章構造に対する標準化は、将来を踏まえた時、避けては通れない重要なものであり最新の動向に着目していく必要がある。

11.4 机上シミュレーション

11.4.1 シミュレーション方法

基本設計の結果を踏まえ、実用化に向けての論理構造、物理構造の設計等を行う詳細設計を進めた。この詳細設計は、モデルタイプ・システムの処理形態、性能、達成上の問題点等、実用化に対する知見をより具体的に獲得する事を目的としている。検討に当たっては、実際のシステム化を前提に机上シミュレーションによる方法により作業を進めた。机上シミュレーションにおいては、詳細設計を現実的なものとするために、実際にいくつかのCTS情報をサンプルデータとし、更に出力メディアにおけるサービスを想定し、CTSのデータをもとに情報作成、内容の変更を行いデータベースに登録する機能と、必要な情報をデータベースから取り出し、各メディア向けに変換・出力する機能を中心に、実際の入力・登録側と出力・展開側、両方の処理を試行（机上シミュレーション）した。

11.4.2 内部構造の検討

印刷情報の格納形態を示すものとして印刷情報の論理構造を検討した。本モデルタイプでは、以下の3種の構造情報を想定した。

- ・ データベースに格納されているCTS情報の実体である基本オブジェクト
- ・ CTS情報の持つ意味的な構造を表現する論理構造
(すべての印刷物に共通の構造を規定することは困難と考えられるため、論理構造を規定する手掛かりとして、印刷物を
 - ・ 論文系 — 論文、小説等
 - ・ 便覧系 — 便覧、字辞典等
 - ・ 統計系 — 表情報等に分類し、各系毎の論理構造を検討した。)
- ・ 基本オブジェクトに対する各種の属性・管理情報
(管理情報には定義情報、オブジェクト管理情報、入力データ管理情報、メディ

ア向け属性情報等がある。)

11.4.3 入力・登録処理におけるシミュレーション

C T S データの入力処理概要はC T S データの分割，分割された要素に対する基本オブジェクトの対応付け，オブジェクトの関係情報の生成，管理情報の生成等の手順を経るものとした。こうした手順により実施したシミュレーション例として，

- ・ 基本分割単位の抽出

基本分割単位の抽出に使用する組み合わせとしてゴシック文字部，ゴシック文字部で見出し記号で終わるもの，ゴシック文字部と段落記号に囲まれたテキスト列，段落記号に囲まれたテキスト列などを想定し抽出を行った。

- ・ 基本オブジェクトの抽出

基本オブジェクトは基本分割単位の列として抽出される。基本オブジェクトとしては，見出しブロック，見出しブロックには含まれているブロック列の2種が考えられた。

等を実施した。

11.4.4 メディア展開・出力処理におけるシミュレーション

出力系アプリケーションに対するシミュレーションは以下のように行った。

対象データとしては，

- ・ 論文系 印刷人のための画像通信の話（「プリンターズ・サークル」）
- ・ 便覧系 現代印刷人のWORD手帳 2（「プリンターズ・サークル」）
- ・ 統計系 1987年12月の株式・金融，商品，気温データ（「MRI-TMS（トップ・マネージメント・サービス）」）

を設定して，各系に対して，CD-ROM / オンライン，ビデオテックス，テレテキスト（文字放送）の3種類のメディア系への出力イメージをシミュレーション結果としてまとめた。こうした一連のマルチメディア向けの情報加工を行う上で，出力側メディアの特徴，扱うデータの内容および必要とされる付加情報を把握するため，対象出力メディアの機能を分析し，各メディアの情報構造に対する整理を試みた。

11.5 汎用データベースに対する評価と今後の課題

11.5.1 評価

入力・登録処理においては、制御情報を利用して、章名、節名等の見出し部分および章、節、パラグラフに相当する文章部分の識別を行った。この際、

- ・ 文字情報のみを対象とする場合でも、必ずしも本来文字情報であるものが、文字によって表現されているとは限らない。
- ・ 意味的な区切りと、制御情報による区切りとは一致しない。

等の問題があった。これらの例外ケースについては、基本的には、入力処理時の付加情報としてCTSデータとともに人間が入力したり、システムで一旦識別処理を行い、これを人間が修正するという2つの対応が考えられる。

論理構造の作成には、CTS情報自体には、見出し、説明文といったデータの意味構造に関する情報がなく、制御情報の組み合わせ等による抽出方法の指示が必要である。このように、入力・登録処理で使用する付加情報は、意味構造に関する定義情報であり、処理の効率化とともに、それらの管理が必要であることが判明した。

メディア展開処理においては、CTS情報のメディア展開を行うため、CTS情報に対応する印刷物から意味構造を抽出したが、これも、CTS情報から識別可能な意味構造とは開きがあった。汎用データベースに蓄積される情報においては、メディア展開を意識することにより、これらの相違を付加情報という形でうめられる。印刷物自体から抽出される意味構造は、各メディアに共通的に使用できる部分が多い。この意味で、汎用データベース中の情報は、印刷物の持つ意味構造を忠実に反映できることが望ましい。また、アプリケーションでは、対象情報の検索機能が、大きな役割を持っている。汎用データベース中の情報として、検索キー生成のための補助情報を持つことが望ましい。しかし、CTS情報の性質を考えた場合、本来検索キーが、意図する意味を自動的に付与することはむずかしいと考えられた。

更に、モデルタイプ・システムに対するシミュレーション結果を踏まえて、マルチメディアへの対応、標準化への対応、原情報独立、統合化・即応性、使いやすさ等の要件の達成状況を評価・検討した。

11.5.2 今後の課題

評価された結果を基に今後の汎用データベースの拡張、より高度な技術を取り入れた再

構築を検討した。本節においては、現在の技術レベルで対応が取れると判断される内容について検討を加えていった。

特にデータベースを構築する技術上の観点から、入力・登録作成作業を効率よく支援する目的として、キーワード自動付与技術を、システムの信頼性、安全性確保の観点からデータベースに対するセキュリティ対策を、更に標準化に対する新しい動向として、CD-ROMの標準化について検討した。

本汎用データベースを更に拡張発展させていくため、未解決の問題点に対する今後の課題項目として、データベース周辺技術を核とし、かつ今後の技術革新が注目されるシステム構築技術として、分散化、マルチメディア化(HDTV)、DTP(Desk Top Publishing)を取り上げ検討を加えた。

11.6 結 び

11.6.1 法制面への対応

汎用データベースを実際に事業化するに当たって、著作権を中心とした権利保護に関する問題は、十分なる検討・配慮が必要と思われる。本調査研究の対象であるマルチメディア対応汎用データベースにおいては、

- ・ データベースとして蓄積される個々の情報の権利
- ・ データベース自体の権利
- ・ 各メディアによって伝達されるデータベースの保護

等に関する検討が必要となろう。

現在、これらは未解決の部分が多く残っている。しかし、本調査研究において権利保護を実現する機構の実現性は技術的に可能であることが判明している。従って、システムの実用化という各局面において、前提条件を明確化した上、各方面の参画による十分な調整討議に問題の解決が委ねられると思われる。

11.6.2 ビジネス化の課題

汎用データベースの事業化に当たっては、法制面の問題の他に、

- ・ 市場性を考慮した応用分野、格納内容の選定
- ・ 需要調査等ユーザのニーズを反映したサービス内容の決定

- ・ 構築時のみならず、運用時にわたるコスト見積と採算性の検討
- ・ 各メディアの特徴，サービス事業の現状に基づく料金体系の選定
- ・ データベースの設置場所，専用センター化といった運用形態，体制の検討
- ・ 格納内容の機密保護，セキュリティ確保

などに対する評価・検討が十分実施されることが望まれる。

11.6.3 今後の方向

本調査研究によって、種々の検討が加えられた汎用データベースは、本格的実用化に向けていくつかの課題や検討の余地を残すものの、CTS技術、画像処理技術等印刷周辺分野の技術のみならず、情報通信・ネットワーク、人工知能等の技術革新の成果を取り込みつつ、

- ・ 印刷資源の分散・共有化の実現、
- ・ データベース構築時のコスト低減、
- ・ 既存データの利用機会の増大、
- ・ 新しい情報サービス・ビジネスチャンスの醸成

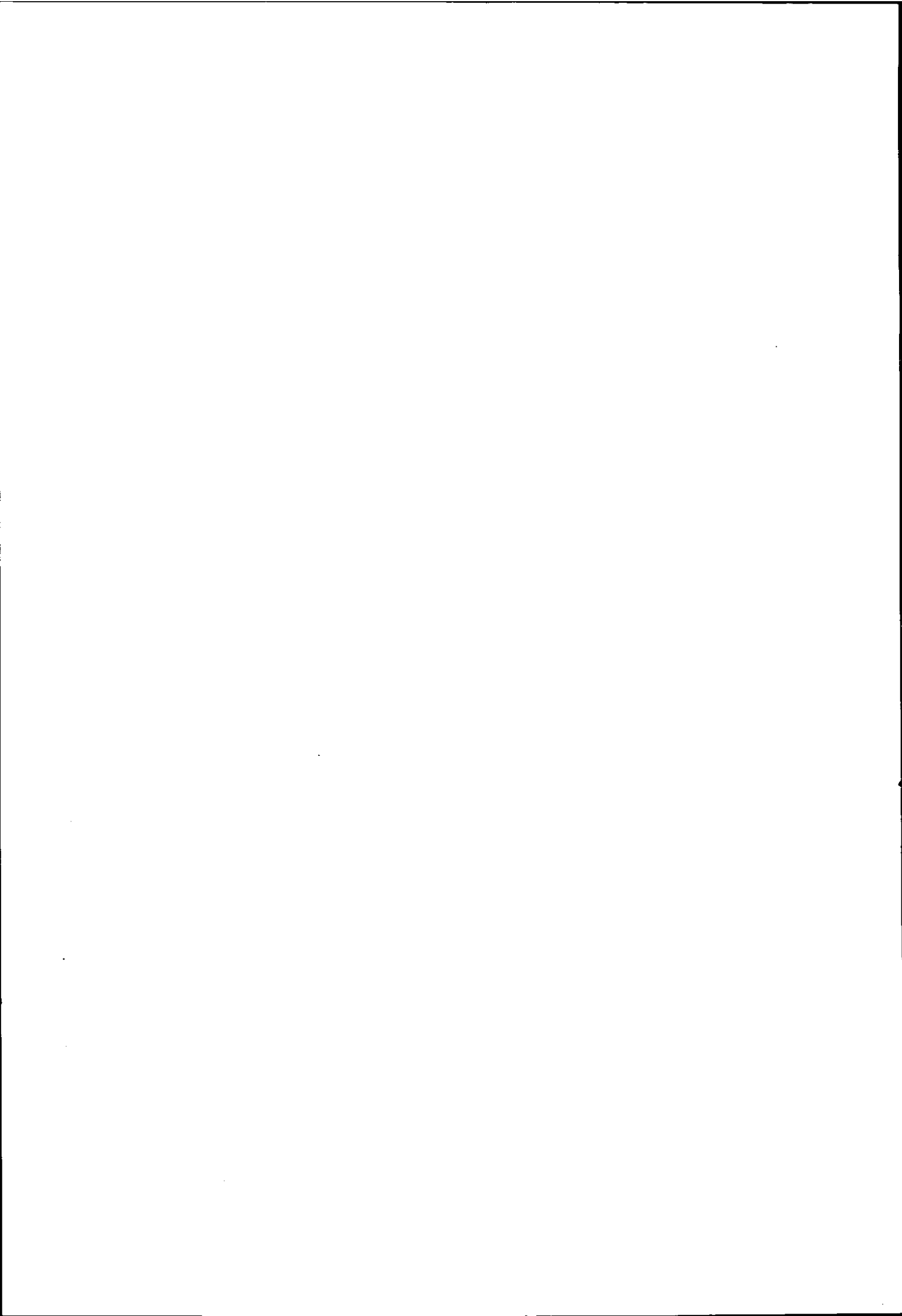
等の面で多大な効果が期待できる。

伝統的な統計分類では製造業に分類された印刷業は、従来からも情報の複製、加工に携わってきており、その業務の性格から、委託加工をこなすサービス産業であるといえる。従来からの文字情報から近年の図形・画像情報に関する情報処理技術を生かしたサービス産業への貢献が期待される。

こうした中、この分野に数多くのノウハウを有する印刷業を中心に、各方面の協力の下、プロトタイプを試作等を含め汎用データベースの実用化、普及に向けた研究の深まりと情報処理産業への貢献が望まれる。

12 鉄鋼関連流通データベース構築

社団法人 鋼材倶楽部



12 鉄鋼関連流通データベース構築

12.1 目 的

この事業は、内外鉄鋼流通分野のデータベースの構築を目的とするもので、これにより業界各社（鋼材倶楽部・日本鉄鋼連盟会員）が共通の情報を効率的に使用でき、鉄鋼業界全体の統計業務の合理化を目指すものである。

この鉄鋼関連流通データベースの構築に際しては、鋼材倶楽部に蓄積している約14万6千件の鉄鋼需給関連の文字情報とデータベース化される数値情報とを複合化したデータベースサービスシステムの開発についても併せて調査・研究を行う。

このデータベースは、鉄鋼業界における自家用データベースとして、業界各社に情報サービスの提供を行うものであるが（他の鉄鋼関連団体とは相互に情報を交換）、データサービスの公共性という観点から、データベースサービス業者ならびに一般ユーザに対しても一定の条件のもとに情報を提供するものである。

またわが国鉄鋼業が、国内鉄鋼流通分野のみならず世界の鉄鋼業に関する各種情報や海外市場の動向を網羅したデータベースを構築し国際的情報交換を可能とする体制を整えることが、「情報立国」をめざすわが国の国策に沿ったものと考えられる。

12.2 実施体制

この事業は、鋼材倶楽部事務局内に設置されている「データベース作業部会」において推進した。

この部会は鋼材倶楽部において、鉄鋼需給に関する調査、統計、資料、情報処理に精通している職員で構成しているが、必要に応じて外部の有識者を招聘して検討・作業を行った。

12.3 実施内容

前述のデータベース作業部会を中心に、早期構築が望まれる「鉄鋼関連流通データベース統計群」の選択を行った結果、表12-1のとおり、24,269系列の諸統計を抽出した。

これらの統計のうち、今年度事業として鉄鋼需要産業統計、鉄鋼市況情報統計、アジア鉄鋼需給情報（文字情報）の構築作業に取り組んだが、具体的な収録期間、系列数の内容は以下のとおりである。

○数値情報のデータベース化

① 鉄鋼需要産業統計

イ. 収録期間：昭和46年1月～62年12月

ロ. 統計系列数：503系列

ハ. 内容：建設等の活動水準，機械工業生産量，自動車生産・販売量，造船受注・起工量等

② 鉄鋼市況情報統計

イ. 収録期間：昭和47年1月～62年12月

ロ. 統計系列数：270系列

ハ. 内容：鉄鋼品種別・サイズ別・地域別市中価格

○文字情報のデータベース化

アジア鉄鋼需給情報

イ. 収録期間：昭和61年10月～昭和62年2月

ロ. 入力件数：450件

ハ. 内容：アジア鉄鋼需給情報（アジア各国の鉄鋼生産，設備，需要動向に関する現地新聞情報記事）

表 12 - 1 鉄鋼関連流通データベースの統計群総括表

統計内容		系列数	主な統計出所
鉄鋼需給	生産	148	通産省「鉄鋼統計月報」他
	出荷	1,271	同
	輸出	2,725	大蔵省「通関統計」他
	輸入	1,045	同
	在庫	242	通産省「鉄鋼統計月報」他
	価格	26	新聞市況他
(2) 鉄鋼次製品	生産	175	通産省「鉄鋼統計月報」他
	出荷	224	同
	輸出	234	大蔵省「通関統計」他
	輸入	20	同
	在庫	39	通産省「鉄鋼統計月報」他
	価格他	38	新聞市況他
(3) 海外鉄鋼統計	韓国	2,825	韓国鉄鋼協会「鉄鋼統計年報」他
	台湾	897	台湾「金属材料及製品工業市場簡訊」他
	中国	50	中国国家統計局「中国統計年鑑」他
	ブラジル	2,543	CONSIDER他
	メキシコ	103	ILAF A他
	アルゼンチン	110	同
	ベネズエラ	97	同
	米国	7,753	AISI他
EC12カ国	2,630	EUROSTAT他	
COM7カ国	500		
(4) 需要産業統計	建築・土木	137	建設省「建設統計月報」他
	産業機械	70	通産省「機械統計月報」他
	電気機械	42	同
	自動車	98	自動車工業会「自動車統計月報」他
	造船	29	運輸省「造船造機統計」他
	一般機械	91	通産省「機械統計月報」他
	容器	3	通産省「鉄鋼統計月報」他
	鉄道車両	10	運輸省「鉄道車両統計」他
(5) 一般経済統計	94	経企庁「国民所得統計年報」他	
合計		24,269	

12.4 ま と め

鉄鋼データバンク構想の一つの柱である鉄鋼関連流通データベースを構築するなかで、今年度は①鉄鋼需要産業統計、②鉄鋼市況情報統計、③アジア鉄鋼需給情報（文字情報）のデータベース化を図ることができた。特に鉄鋼需要産業統計、鉄鋼市況情報統計については鉄鋼需要の分析、予測をするうえで貴重なデータであり、加えてアジア鉄鋼需給情報についても厩大な潜在的鉄鋼需要の期待できる当地域における鉄鋼需給情報を的確に把握することにより、鉄鋼業界としての市場調査業務の一翼を担うとともに、鉄鋼の安定した供給体制の確立に資するものと期待されている。

これを踏まえて、昭和63年度以降においては、以下の課題を中心に具体的な作業を実施することとしている。

① 鉄鋼需給関連データ構築の拡充

現時点で構築されているデータでは必ずしも十分でないので、引き続き必要とされるデータを抽出し、順次入力作業を推進し、その拡充を図る。

② メンテナンス用ソフトの開発

データベースのメンテナンス作業の形態毎のソフトを作成する。

③ データベース対象統計の統合・整備

データベース用コードの作成を行い、データフォーマットの共通化のプログラムを開発したうえで、ファイル共通化ソフトを開発する。

以上の作業が完成すれば、

① 業界各社が個別に保有し、蓄積している基本統計類がデータベース化されることにより統計関連業務が合理化され、効率的利用が可能となり、鉄鋼データバンクの早期実現に近づく。

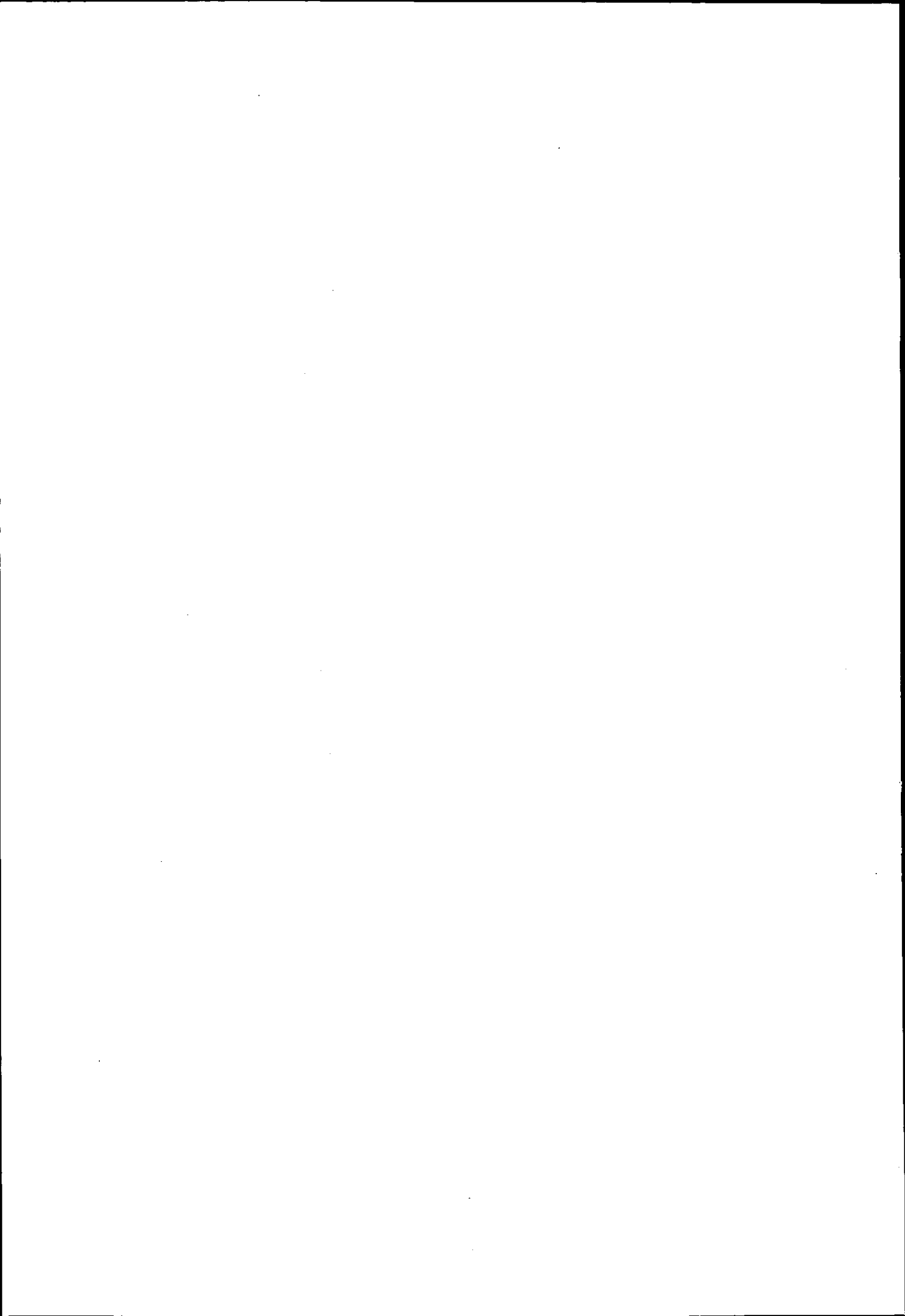
② 会員外にとっても、鉄鋼関連統計の利用が容易となる。

③ 鉄鋼業界の自家用データベースとして産業別情報拠点の一翼を担える。

13 地図データベースを利用した

商圈分析システムの開発

アジア航測株式会社



13. 地図データベースを利用した商圈分析システムの開発

13.1 背景と目的

本開発は、昭和60、61年度実施の「産業分類別にみた地図利用とデータベース化の研究」を受けて実施するものである。

昭和48年の、いわゆる大規模小売店舗法の制定により、大規模小売店舗の出店ないし増床に際しては、主として中小小売業の事業活動の確保という観点から、商業活動調整協議会での審査という形で調整が行われ、今日に至っている。特に、大都市圏においては、昭和53年の改正以来、実質的な出店規制の意味合いを濃くしている。

この商業活動協議会の場における審査は、大規模小売店舗の過密の程度、地域経済の動向、大規模小売店舗の増加の程度および小売商業施設充足の程度について行われ、それぞれについて類似都市比較指標、将来性指標、影響度指標、消費者便益指標が用意されている。

これらの各指標での審査は、原則として大規模小売店舗の出店等の行われる地域の含まれる市区町村単位での把握のもとになされているのが実情である。

このために取得するデータ上の制約がなければ、行政界を何ら意識しない形で、実質的な商勢圏や影響圏にブレイクダウンした小地域を対象として適宜の指標を得ることができれば望ましいこととなる。ところで、人口を始めとする地域の社会経済指標のデータ群は行政区単位に収集・整備されていることが多く、そのままの形や加工することなしに細分化された指標を得ることができにくい状況にある。

この様なことから、各指標を当該市区町村についてのみならず、実質的にはより狭域な広がりを持つであろう小売商業の商勢圏についても得ることができれば、商業活動調整協議会の場での資料として有効となるだけでなく、大規模小売店舗の出店を計画する側の立場においてもより適切な指針として利用できることとなる。また、出店時のみならず、既存店立地の見直しという要求の一部にも応えることができるものとなる。

これらの何れの場合でも、最も重要なことは、商勢圏の設定をいかに簡便かつ適切に行うことができるかにかかっている。このことは、商勢圏が固定的なものではなく、当該の地域の地域的制約に大きく依存しているからである。

他方、近年の計算機利用技術における図形処理技術の進展を背景に、地図情報をデータベース化し、種々の業務に役立ててゆこうとする動きが盛んである。商圈の分析や設定もその一つとして取り上げられ、また期待もされている。

しかしながら、地図情報のデータベース化は大量なデータを取り扱うため、入力に多大の時間を要したり、処理目的に応じたデータ構造の作成といった分析的業務以前の問題がある。地域を小売商圏としていかに適切にとらえ、地図情報に展開してゆくかといった問題がそれである。これは、本来的には買物行動をも加味した総合的なモデルの構築が理想ではあるが、現状は、その様な段階にない。(吸引力モデル、グラビティーポテンシャルモデル、地図一般化モデルの何れにおいても適用上の問題を残していながらも、これらを適用してゆかざるを得ない側面を持っている。同時に小売商圏の分析は、その結果が店舗立地や営業政策への指針となるべく要求が高いため安定性を持ったものでなければならない。)

そこで、本検討の目的として、ソフィスティケートなモデルではあるが、修正ハフモデルを取り上げ、これに250mメッシュによって地域を細分化したデータを用いることによって影響度指標の把握を始めとする各種指標を有効に取得できるためのデータベースの構築とそのアプリケーションソフトウェアを実験的に作成した。これは、過年度の検討においても250mメッシュを用いての適用により安定性や適合性に関し比較的良好な結果を得ていることの延長上にある。

13.2 商圏分析データベースの概念

(1) 特徴

用いるデータの質を向上させることにより、修正ハフモデルによる推定の安定性を求めることを目的としてケーススタディを行った。

これに用いたデータベースは、

- ① ランドサットのTM(サーマルマップ)データを用いて土地被覆区分を行い、町丁目字別人口(ないし国勢調査人口)、町丁目字界の界面データ(ないし、国勢調査の調査区)より同定した250mメッシュの人口データ
- ② 道路のネットワーク情報とバスルートおよびバス停位置の1次情報(何れも、ベクター型のリンク・ノード情報)より加工し同定した250mメッシュのルート接続データ
- ③ 商業集積地の外周範囲を面積同定した250mメッシュの集積地データ

より構成される。

これらの特徴は、250mメッシュであるので一様な単位面積を持ち、かつ、このメッシュサイズは全般に町丁目字界の大きさよりも小さく、従来、町丁目字界をそのまま用いるか、または、鉄道道路敷、河川等の地域分断線で適宜に分割して取り扱っていた各居住ゾーンの設定に対し、この個々のメッシュを居住ゾーンと位置づけることによって画一的な処理が可能となる。

また、各ゾーンの中心的位置を代表点として集積地への距離等を取り扱っていたことに対しても画一的な処理が可能となる。

人口については、町丁目人口をランドサットTMデータの土地被覆区分（25mピクセル毎の重みづけによる）により、常に新しい情報を得ることが可能となる。

道路ネットワークやバスルートのリンク・ノード情報は、250mメッシュ相互間での接続方向データとすることによって、地図情報としては比較的煩雑なネットワーク情報を一様な精度で要約できることに意義がある。このことによって、アルゴリズム的な面はともかくとして、処理時間的には、従来、大型汎用機によらなければ効率が悪かった最短経路問題に対しても、パーソナルコンピュータでのコストパフォーマンスを向上させることが可能となる。（最短経路問題として適宜のメッシュ間相互の時間距離を与えることは、従来、手作業等で行っていた距離のデータ取得を容易にかつ、画一化できる。）

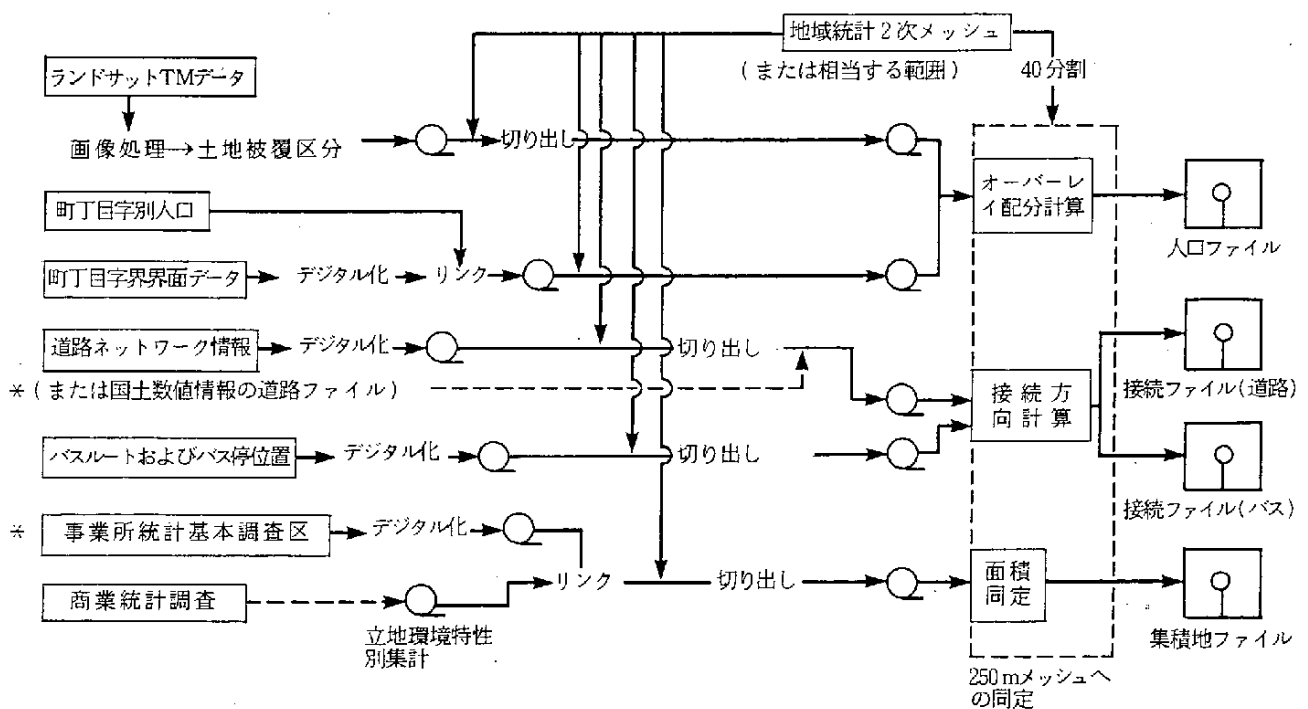


図13-1 商圈分析データベースの概念構成

- * 国土数値情報は、現在の所、官公庁・自治体の他、特に使用許可を得たものに限られたものとなっているため、ここでは類似データを別途作成している。（将来的には例えば、自動車のナビゲーションのためのデジタル道路地図の構想にみられる様に、一般に公開される可能性を含んでいる。
- * 事業所統計調査の基本調査区は、事業所統計調査の他に、商業統計調査、工業統計調査の調査区としても用いられている。ところで、商業統計調査の立地環境特性格集計の各区分は、この調査区を積み上げたものであるから、この調査区についても一般に公開されれば、商業統計調査の利用勝手はより高まるが、現在の所、その使用には制限がある。ここでは、実際の利用の方向をさぐるために参考として用いた。

(2) データ構造およびフォーマット

データベースとしての250mメッシュは、地域統計メッシュの2次メッシュ区画を縦横それぞれ40分割したものである。メッシュコードは、1次区画、2次区画は地域統計メッシュのコードをそのまま用い、その後4桁で250mメッシュコードを付与することによって体系付けられる。

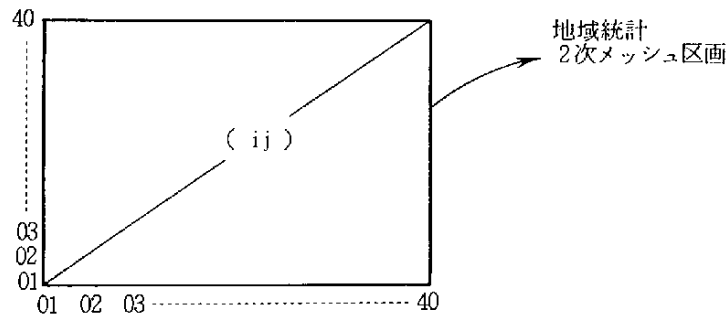


図13-2 250mメッシュ

メッシュデータであるので、従前の1kmメッシュや500mメッシュとの対応を図れるのは当然であるが、ここではむしろこれから分割してゆくのではなく、2次メッシュ区画から直接分割することとしている。これは、2次メッシュ区画が概ね10km×10kmの範囲に相当するためである。即ち、従来の市区町村を一括して扱う場合は別にして、例えば、地域を調査居住区に分割して影響度指標を求める様な場合は、一度に取り扱える範囲としてこの2次メッシュ区画のサイズが適当と考えたからである。

むしろ、適宜の範囲から、この2次メッシュ区画に相当する範囲を自由に切り出して処理できることの方が望ましいためである。(処理に際しても、1項目当たりのデータ数は最大で1600(40×40)であるので、パーソナルコンピュータでの取り扱いでも十分対応できることとなる。)

一例として、表13-1、表13-2に人口および進路の接続データのフォーマットを示した。

表13-1 データフォーマット(人口)

2次メッシュコード	250mメッシュコード	数 総 人	数 世	男女別数		都道府県 市区町村コード		
				男	女	位	位	位
						1 第	2 第	3 第
1~6	7~10	11~14	15~18	19~22	23~26	27~31	32~36	37~41

表13-2 データフォーマット(道路網)

ド ド 2次メッシュコード	ド ド 250mメッシュコード	接 続 情 報														
		道路規格 ㉔				道路規格 ㉕				道路規格 ㉖						
		非接続フラグ	接 続 方 向				非接続フラグ	接 続 方 向				非接続フラグ	接 続 方 向			
			1	2	3	4		1	2	3	4		1	2	3	4
1~6	7~10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

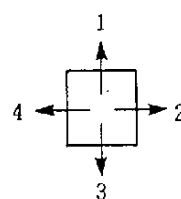
道路規格A；11m以上

B；5.5m以上11m未満

C；2.5m以上5.5m未満

非接続フラグ；同一メッシュ内で他の規格道路の何れにも接続しない場合に立てる。

接続方向；各メッシュの接続方向は下図の方向に反応する場合に立てる。

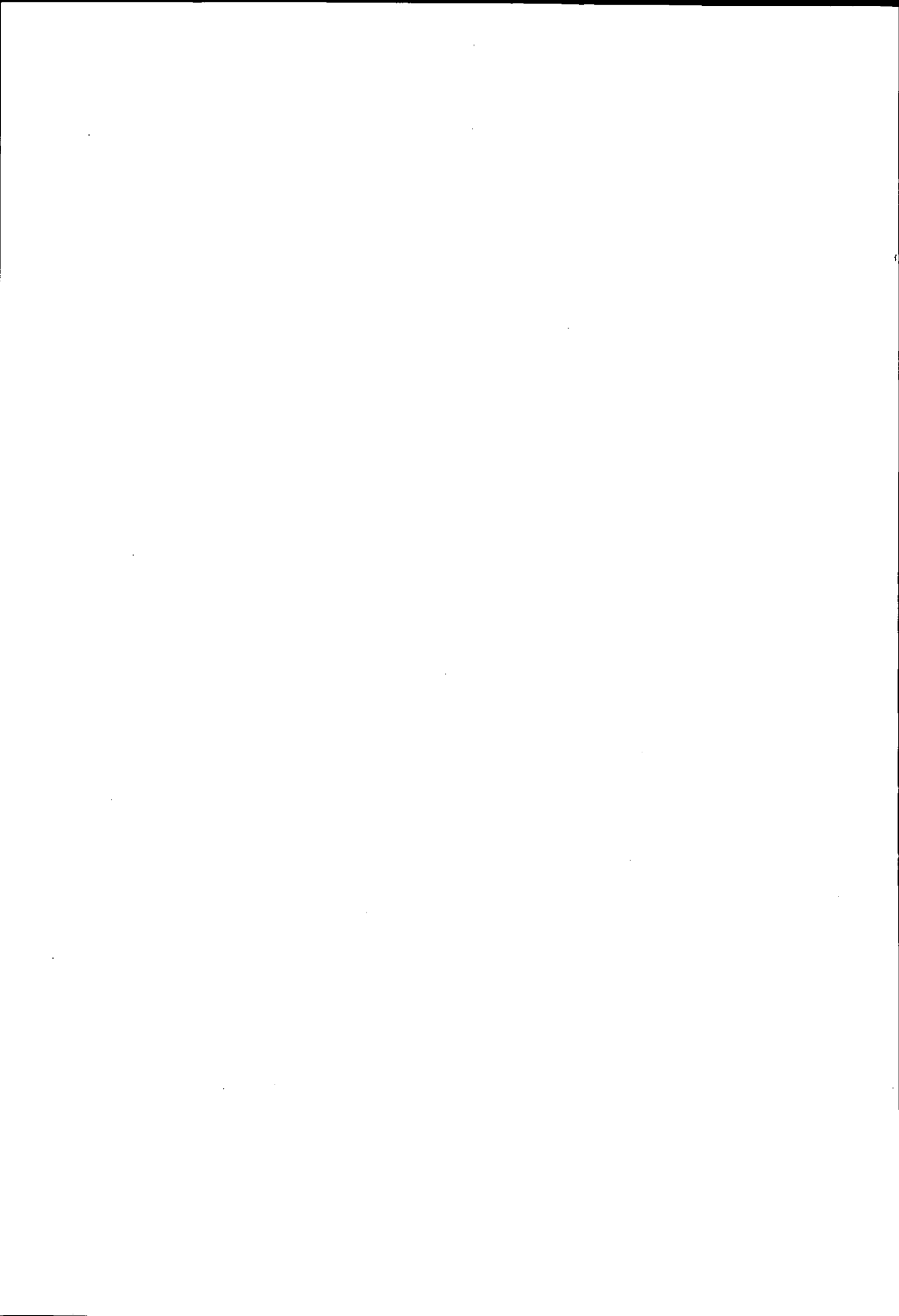


13.3 課 題

本開発のケーススタディに用いたデータベースは、ランドサットのTMデータを除き、何れも行政で保有しているデータがそのソースである。

今回の開発の一つの主眼である道路ネットワーク情報のメッシュ形式への表現とランドサットTMデータによる人口の適切な配分の何れも目的に応じた加工方法がその要点である。

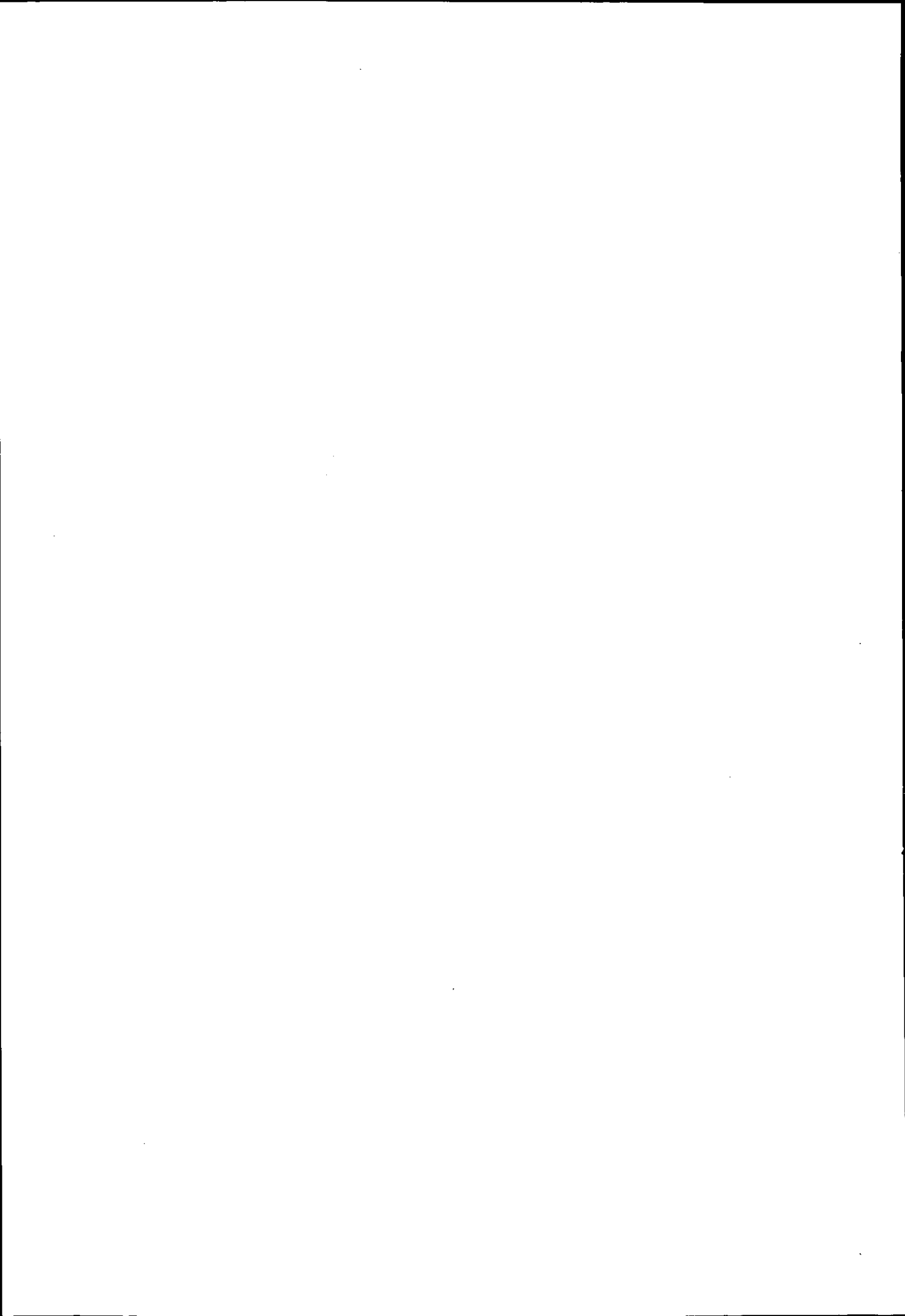
ところで、これらの前提としてはソースデータの使用に関しての規制緩和が必要であり、それを実現してゆくためにも本データベースの利用の有効性を強力に訴えてゆくことが必要である。



14 地域分析・計画のための

地図データベースの構築

株式会社 日本総合技術研究所



14 地域分析・計画のための地図データベースの構築

14.1 構築の背景と目的

14.1.1 背景

近年、コンピュータマッピングを利用した地理情報システムが注目を集めている。米国などにおいては、地図データベース、あるいは地理情報システムは、特定の地域を選別したり、市場戦略のための手段（エリアマーケティングツール）として一般に利用されている。しかし、わが国においては、あるまとまった広さで地域をくくり、それらを分類していろいろな基準でグループ化したり、ある情報項目に対する特化した地域を選別したりして、その地域にあったマーケティングを行うという、いわゆるエリアマーケティングの概念や、その重要性ということあまり論じられてこなかった。

ところが、ここ数年、「市場の成熟化」、「大衆から分衆へ」ということばに代表されるように、市場のセグメント化が重要視されるようになってきた。どの地域に対しても同じようなマーケティングを行うのではなく、世代構成、家族構成はもとより風土、環境などを考慮し、ある軸で市場をくくり、その地域の特性にあったマーケティングを行うほうが、効果が高く、また効率が良いか認識されるようになってきた。

ところで、「市場をとらえる目」、それはもちろん数値から導き出されるものだけではなく、ある切り口、評価軸を必要とするものである。これは個人の感性に頼るという性格をもつものであるが、データのプレゼンテーションの仕方によっては、誰もがその市場に関して着目すべき点を見いだせる可能性をもつ。この意味でひらめきを喚起させることができるデータの「見せ方」は存在する。

本システムは、エリアマーケティングに役立つ基礎的なデータの蓄積とともに、こうしたデータのプレゼンテーションの側面も検討しうるシステムを目的とした。

14.1.2 開発経緯

「地域分析・計画のための地図データベースの構築」事業は61年度に提案されたものであり、今年度は引き続き61年度までに作成された地図データベースの機能の拡充を図ることを主眼とした。62年度は具体的には、蓄積するデータの拡充を図るとともに、人口推計値の算出、エリア潜在需要力の算出など、地域分析に必要と考えられるアプリケ

ーションの開発を目標とした。

この目標を達成するため、本年度においては以下の3点を主体に事業をすすめた。

- (a) メッシュデータと行政区データのリンク
- (b) 60年メッシュデータの推計
- (c) エリア潜在需要力の算出

1.4.1.3 開発の概要

(1) メッシュと行政区の対応テーブルの作成

Area information Mapping System（以下aimと略する）は、当初はメッシュデータを任意のエリアで集計できるように開発された。そこで、いろいろな単位で集計されるデータを一旦メッシュデータに変換することで、aimを利用してそれらのデータを任意のエリアで集計することが可能である。

そこで、最終的には利用者側に蓄積されている顧客情報などの固有のデータの取込みを考えるが、その上でまず手始めに、行政区単位で収集されている統計データの取込みを行った。メッシュ・行政区の対応テーブルは、そのためのパラメータの役割を果すものである。

(2) 60年メッシュデータの推計

統計データの結果の公表までに時間がかかるため、リアルタイムなデータとしては利用しづらい。そこで、国勢調査に関しては、調査結果の速報あるいは行政区単位での調査結果が公表された時点で、メッシュデータの推計値を算出できるような推計値算出モデルの開発を行った。60年の国勢調査の結果が公開された時点では、データの整合性の分析、検討を行い、推計モデルをより正確なものへ近づけるための再検討を行い、次期に備える。

(3) エリア潜在需要力の算出

集計されたエリア内のデータを利用して、エリア潜在需要力を算出できるようなモデルの検討を行った。これにより、その地域内における商品の販売余力の推定が容易に行え、また出店規模をどのくらいにするか大まかな判断材料とすることができるようになった。

1.4.1.4 システムの基本的視点

本システムは次のような基本的視点をもつ。

(1) マーケティングデータの実用的なサービスシステム

任意のエリアのデータを不特定多数の利用者に、素早く提供できる。

(2) 地図データベースと統計データベースを連動させた総合的なデータベースシステム。

単にデータを収集するだけでなく、データを視覚化し、地域分析を容易に行える。

1.4.2 システムの概要

1.4.2.1 システム設計上の考え方

(1) 収録するデータについて

統計データ（メッシュデータ）だけではなく、地域分析に必要と思われる関連データを収録した。具体的には課税対象額や納税義務者数などの所得に関するデータの蓄積を行った。また、大型店の立地データやマンションなどの建設データの取込み方法を検討した。

(2) システムの機能面から

- ・ データ収集からサービスまで一環して行うシステムであること。
- ・ 数値データのグラフ化や、地図上でデータの分布や各地域のランクなどを把握できるようなマップ化を図った。
- ・ 統計データの結果公表までのタイムラグを埋めるようなシステムとする。
- ・ 宅地開発計画、造成計画、住宅分譲情報などを予測の要因として埋め込み、情報をフィードバックさせながら将来予測のシミュレーションができるようなシステムを目指す。

1.4.2.2 システム構成の考え方

aim におけるハードウェア構成としては、統計データはデータ量が多いため、データの蓄積あるいは集計という面ではミニコンクラスの計算機に頼らざるを得ない。しかし、集計されたデータをグラフにしたり、マップにしたりするような加工処理はパソコンの方が扱いやすい。

そこで、データ集などからのデータの抽出や集計結果の出力レイアウトの工夫はパソコンを利用し、実際のデータベースの構築と集計、検索といった処理はミニコンで行うという構成をとっている。

なお今後は、CD-ROMなどの大容量の記録媒体を利用して、地図とデータをユーザのパソコン上で利用できるマーケティングシステムの開発を計画している。

14.2.3 システムの開発計画

(1) データ収集計画

61年までに以下の統計データを取得済みである。

- ・ 国勢調査 55年
- ・ 事業所統計 56年
- ・ 商業統計 57年(売上別)

62年度においては、統計データとして以下のデータを蓄積した。

- ・ 課税対象額
- ・ 納税義務者数
- ・ 商業統計57年(規模別)
- ・ 家計調査年報55年、60年

また、地形データとしては、

- ・ 60年人口ドット

を新たに蓄積した。

(2) システム開発計画

62年度においては以下のようなモデルおよび機能の開発を検討した。

- ・ 推計値算出モデル
- ・ エリア潜在需要力算出モデル
- ・ エリア特性の時系列比較機能
- ・ エリア潜在需要力測定機能

14.3 システムの内容

14.3.1 対象データ

主なデータ収集源は国の指定統計データとした。その他行政区単位に収集されたデータの蓄積も行った。対象とするデータは以下のとおりである。

表14-1 対象とするデータ

<ul style="list-style-type: none"> ・ 人口関係 <ul style="list-style-type: none"> 年齢別 未婚・有配偶 通勤・通学地 従業上の地位 就業状況 ・ 商業関係 <ul style="list-style-type: none"> 売場面積別 売上別 従業員数別 業種別 ・ 所得 <ul style="list-style-type: none"> 課税対象額 納税義務者数 高額所得者数 	<ul style="list-style-type: none"> ・ 世帯関係 <ul style="list-style-type: none"> 家族構成 所有形態 居住形態 居住空間規模 ・ 事業所関係 <ul style="list-style-type: none"> 開設年度別 従業員数別 業種別
--	--

14.3.2 ファイル構成

本システムにおいては、原データとしては統計情報が主体である。この統計情報は、調査年度、調査周期がそれぞれ異なるため、データの収集・更新を重要視し、各統計ごとにファイルを細分化した。

また、地図データに関しては、行政界、鉄道、駅などのデータを属性別に蓄積し、必要なデータを選択して地図上に表現できるようにした。更に、ユーザの利用情報を管理するために利用者管理システムを用意し、利用情報ファイルおよび集計指示ファイルを管理している。

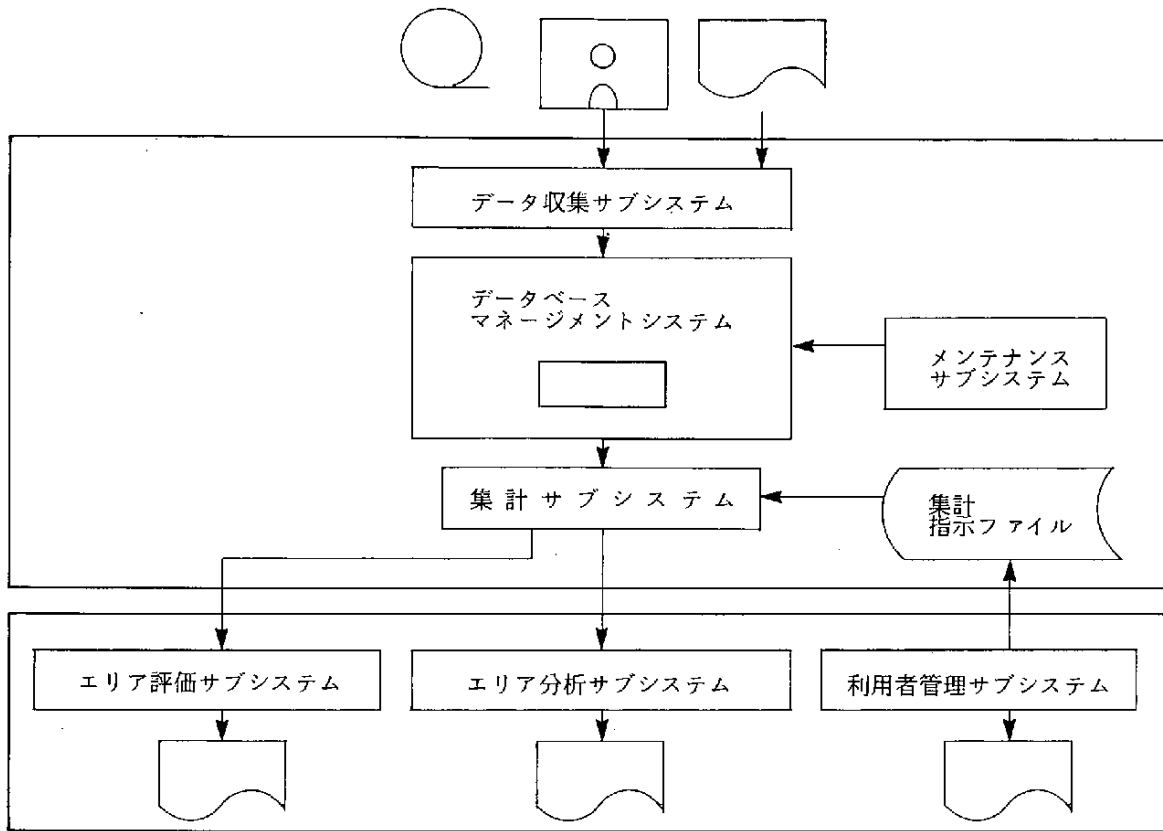


図14-1 システム構成

14.3.3 システム構成

(1) データ収集サブシステム

データ収集サブシステムは、統計データ取得部と地図データ取得部から成る。

(a) 統計データ取得部

1) ファイル作成モジュール

データ集などから作成されたデータシートに基づいて、統一形式の収集データファイルにデータの入力を行う。

2) 形式変換モジュール

統計データ (MT) の形式変換を行い、収集データファイルに蓄積する。

3) チェックリスト変換モジュール

収集データに蓄積されているデータのチェックやリストを出力する。

4) データベースロードモジュール

収集データファイルに蓄積されたデータをデータベースにロードするための形式変換を行う。

(b) 地図データ取得部

1) 地図データ収集モジュール

入力はパソコンとタブレットを利用して行う。地図データベースは基本的にはVAX上に蓄積される。

2) 地図データ出力モジュール

出力はプロッタにて行う。ただし、パソコン上でも画面に地図を表示することができる。

(2) データ更新サブシステム

(a) 60年人口ドット情報の作成

現在のシステムにおいては、人口ドットが集計の基本となっている。そのため、60年のデータを任意のエリアにおいて集計するためには、55年と同様、60年人口ドットを作成し、システムに蓄積する必要がある。

そこで、60年ドットの作成は次の手順で行う。

あるメッシュ内における人口の増加あるいは減少を、55年から60年にかけて集中的に住宅が建設、分譲された地域、例えば、〇〇ニュータウンと称されるようなある特定地域に注目して人口ドットを更新した。

(b) 推計値の作成

行政区単位で集計された60年国勢調査項目の公表結果（これは、報告書として公表されている）を参考として、55年の調査結果との比較、あるいは住宅情報を取り入れて当社オリジナルの人口推計モデルを作成した。このモデルに基づいて、60年国勢調査メッシュ推計値の作成を行った。

(3) データ集計サブシステム

集計指示ファイルに基づいて、実際に任意のエリアにおけるデータを集計するサブシステムである。

(4) 集計データ分析・評価サブシステム

ユーザから指定された任意の商圏内における集計されたデータからその地域がどのような特性を持つかという分析をし、指標などを使いエリア力の評価を行うシステムである。

(5) 利用者管理サブシステム

利用者の必要とするデータの情報をオーダとして入力し、任意の商圏内の集計指示フ

ファイルを作成する。また、任意の地点を中心とした縮尺2万5千分の1、5万分の1の地形図、人口分布図の出力指示ファイルを作成する。更に、システムの利用状況を管理する利用者管理ファイルを作成する。

14.3.4 ハードウェア構成

(1) 計算機本体

集計処理の効率アップのためVAX 11を使用している。

(2) 外部記憶

データ量はかなり膨大になるため処理速度の点からも磁気ディスクを利用している。

(3) 通信

オンラインサービスは行っていない。将来はサポートする予定である。

(4) パソコン

集計結果をビジュアルに表現するために、パソコンを利用している。これはミニコンクラスでグラフィック機能を扱おうとすると、どうしても処理が重たくなってしまうためである。パソコンベースの方が簡便に処理しやすいという利点がある。

また、利用者からのオーダのあった集計エリアに関する履歴や、利用者の履歴管理もパソコン上で行っている。

14.4 システムの利用例

aimにおいては、地域の区画にかかわらず、任意の広さで統計データの集計ができたため、独自の商圏内の人口や世帯数の量が把握でき、ビジュアルに地域分析できることが最大の特長となっている。従って本システムがターゲットとする利用者は、地域をある程度マクロでとらえて、地域分析や市場戦略立案などを行う企業や自治体である。

表14-2 aimの利用例

分野	業種	利用形態
金融・保険	銀行	出店計画
	保険	代理店出店計画 地域別顧客分析
大型小売店	百貨店	出店計画 事業計画
	スーパーマーケット	〃
小売店	コンビニエンスストア, パン, 化粧品, ガソリンスタンド, アイスクリームショップ, 郊外型大規模書店などチェーン展開を行っている所	出店計画 立地分析 事業計画
飲食店	コーヒー, 居酒屋, ファーストフードレストラン, 喫茶店, 惣菜屋 弁当屋, レストラン	出店計画 商圈調査 事業計画
製造業	食品, 衣服, 家電など消費者向け 製品を生産している製造業	地域計画策定 〃 販売店への営業資料作成
建築・不動産業	建設会社	宅地開発計画 建設計画 立替, 修理需要の把握
	不動産会社	立地調査 店舗誘致計画 ビル開発計画 宅地開発計画
調査・研究	シンク・タンク コンサルタント会社	各種調査分析
	広告業	地域別広告戦略策定
その他	レンタカー会社, ホテル, クリーニング, コインランドリー	出店計画 事業計画
	学校, 保育園, 病院, 学習塾	立地計画

14.5 システムの評価と今後の課題

14.5.1 システムの評価

本年度（62年度）の事業の成果は、60年国勢調査項目のメッシュ推計値の作成およびその蓄積と、商業統計の規模別集計データ、家計調査年報データ、所得データを取り込むことができたことである。これにより、時系列の分析やエリア潜在需要力算出が簡単に行えることができるようになった。すでに、いくつかの地域分析に本年度開発されたシステムを利用してみた。その結果に対する利用者の評判はかなり高いものであった。

これは、

- 1) 異なる様々なデータを任意の商圈で集計できる。
- 2) 集計コストが安い。
- 3) 集計時間が短い。
- 4) エリア力を把握するのに、aimの出力結果が非常に役にたつ。

などの利点を備えているためと考えられる。

14.5.2 今後の課題

それぞれの業務によって、同じデータをみてもその評価の仕方はまるで異なるものになる。また収集すべきデータは多岐にわたっている。しかしその中で本当に意味のあるもの（データではなく情報という）はごく限られている。この情報をいかに引き出すか、いかにその情報の持つ意味の重要さをみせるかがエリアマーケティング、ひいてはこうした分野での地図データベースの有効性を喚起させることになる。

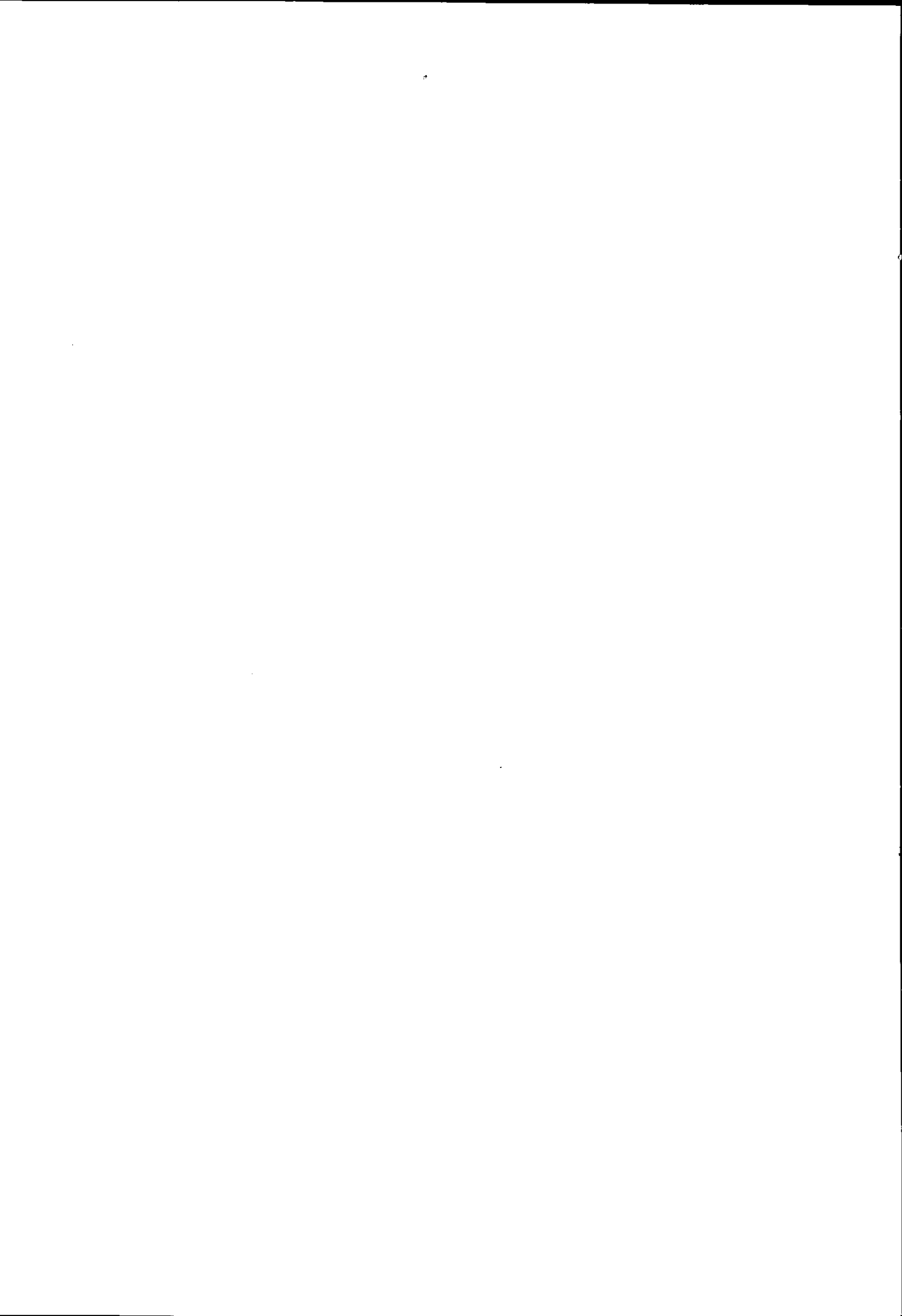
今後の課題として

- 1) 利用者の必要とする情報の分類
- 2) 効果的なデータの利用

などを具体的にとりまとめていく必要がある。aimの将来方向として、システムそのものをスタンドアローン型のワークステーションに移行し、各業種向けのマーケティングのためのエキスパートシステム構築も検討中である。

15 実用化合物データベースの
CD-ROM用実験システムの開発

日本電子計算株式会社



15 実用化合物データベースのCD-ROM用実験システムの開発

15.1 実用化合物データベースの全体像

化学分野において、各種材料、医薬、農薬等の新規化学物質の有効な製法等の研究開発が積極的に推し進められている中で、必要とされる化学物質について記載した商品カタログ的な情報を網羅した実用に即した化合物データベースを構築することは、研究開発の促進および流通の迅速化また国際的視野の面からも大いに寄与するものである。

実用化合物データベースの情報源となるのは化学物質に関する商品カタログである。商品に関する情報は、数値情報、文章情報が混在しており、更に情報量が膨大なこと、情報の更新が頻繁に発生すること等から、適切なDBMSの管理のもとにホストコンピュータ上にオンラインデータベースとして構築することが望まれる。

一方、商品カタログには化学構造式や図表データのような凶化情報が含まれ、また利用者は、商品カタログを生での入手を希望している。

このことより商品カタログを一次情報としてそのまま収録可能な媒体が必須となり、この要求に応えるものとして、CD-ROMが最も適切と考えられる。

CD-ROMに格納された情報はパソコンをベースとした入出力環境で出力表示が可能であり、また、ホストコンピュータとオンライン接続することにより、CD-ROMとオンラインと連動させた総合的な利用が図られる。このシステムを概念を図15-1に示す。

本システムにより、CAS登録番号等のリンク情報を介した商用データベースのアクセス、またプライベートなデータベースとの連動といった利用の高度化、また将来的にはネットワークを通じた商品のオーダエントリシステムへの拡張が期待される。

15.2 CD-ROM用実験システムの開発

15.2.1 CD-ROM用実験システムの概要

オンラインとCD-ROMを連動させた総合的な実用化合物データベースの構築に向けて、その第一歩としてCD-ROM用実験システムの開発を実施した。

本システムの開発の骨子は以下のとおりである。

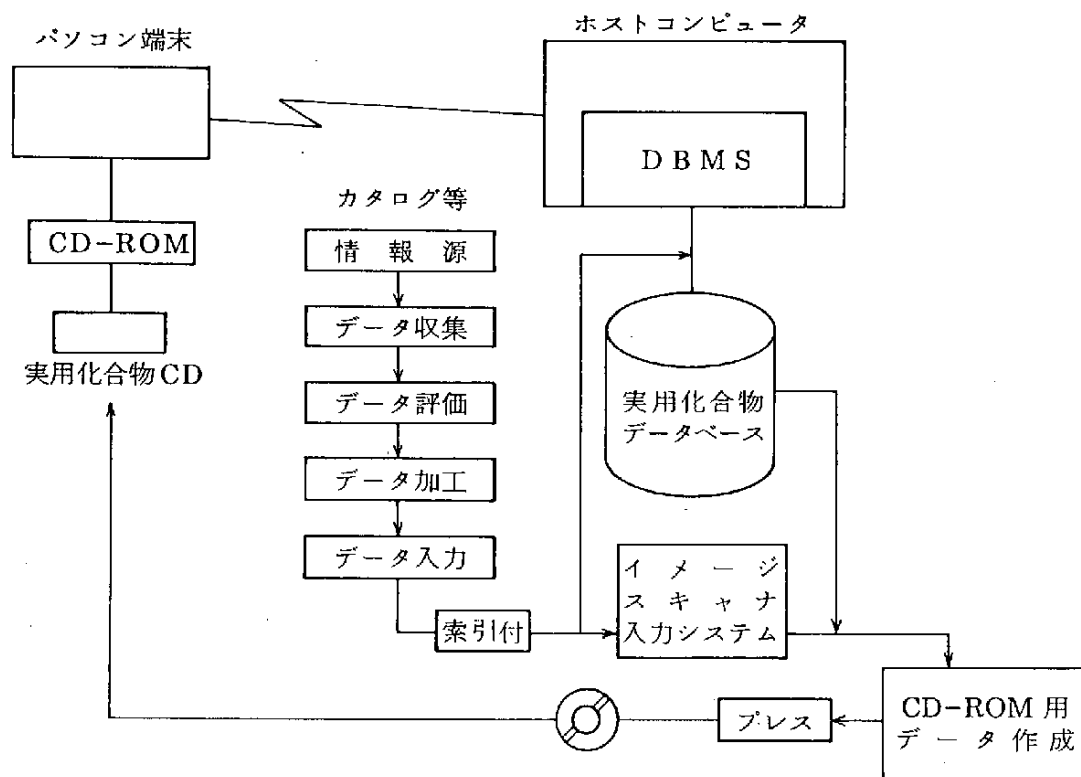


図15-1 实用化合物データベースのシステム概念図

- ① 本システムはパソコンをベースにCD-ROMドライバ、ビューコントローラ、高精度ディスプレイおよびLBP（レーザビームプリンタ）によって構成する。
- ② 商品カタログをイメージ情報としてCD-ROMに格納，CD-ROM一枚を作成する。
- ③ 商品に関するインデクス情報をパソコン上にデータベース化する。
- ④ 商品に関する情報を検索し，必要な商品カタログを出力表示するためのソフトウェアを開発する。

本システムはシステム自体として独立した構成と機能を持ち，手軽なパソコン環境で，化学物質商品の検索，カタログ情報の即時提供を可能とするものである。

開発に当たって，本システムが実験システムとして，实用化へ向けての基本的な条件，機能を提供することにより，評価が可能となることを基本要件とした。

15.2.2 システムの構成

本システムのハードウェア構成を図15-2に示す。

図に示すように本システムは16ビットパソコンをベースにCD-ROMドライバ、ビデオコントローラ、高精細ディスプレイ、LBPを組み合わせたニューメディアシステムとなっている。

今回の実験システムでは、これらの構成機器は日立製の以下のものを採用した。

- ・ パーソナルコンピュータ：B-16EXIIモデルJ
- ・ CD-ROMドライブ：CDR-1003S
- ・ ビデオコントローラ：RVC-MAP-01
- ・ 高精細ディスプレイ：K15-9120PH
- ・ レーザビームプリンタ：R31-0015

高精細ディスプレイはたて型A4のサイズで8ドット/mmの分解能、またLBPは16ドット/mmの解像度をもち、カタログそのものの原寸表示、印刷が可能となっている。

また、本システムのソフトウェアはオペレーティングシステムとしてのMS-DOS、イメージ情報の入出力制御を行うインタフェースプログラムおよびインデクス情報の検索出力制御を行うアプリケーションソフトウェアによって構成される。

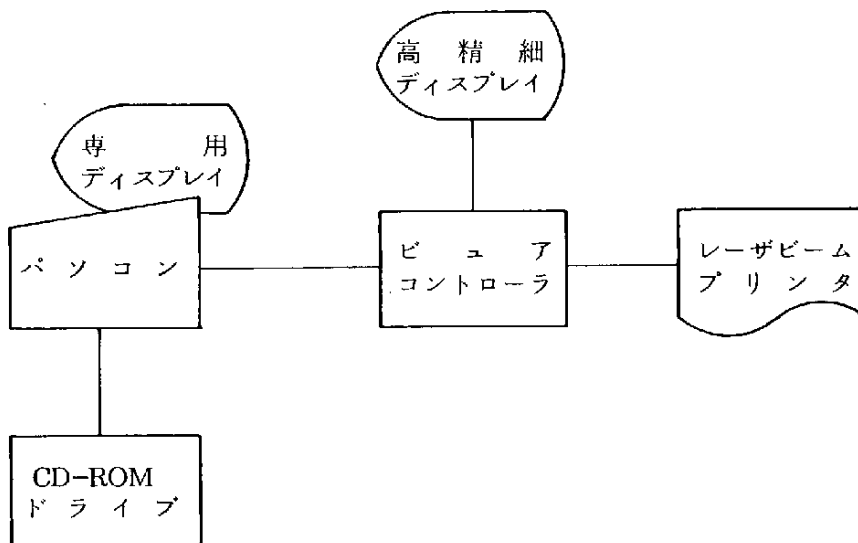


図15-2 CD-ROM実験システムのハードウェア構成

15.2.3 データの構成

本システムで用意されるデータの種類は、パソコン上にデータベース化される商品に関するインデクス情報とCD-ROMにイメージ情報として格納される商品カタログそのものに分けられる。

(1) インデクス情報

インデクス情報としてデータベース化されるべき情報は、

- ・ 成分名等化学物質の特性を表わす情報
- ・ 化学製品メーカー名等流通に関する情報
- ・ CAS登録番号等の付加価値情報
- ・ CD-ROMとのインタフェースをとるための情報

がある。

本実験システムではデータ項目として以下のものを設定した。

- ① 商品名
- ② 会社コード
- ③ 分野コード
- ④ 成分化学分類コード
- ⑤ 成分名
- ⑥ 成分分子式
- ⑦ CAS登録番号
- ⑧ カタログ先頭ページ番号
- ⑨ カタログページ数

①～⑦のデータ項目は検索の対象になる他、付加情報として出力の対象ともなるものである。従って商品名、成分名、成分分子式の文字情報については検索用の英数カタカナ文字によるデータと出力用の漢字データの2種類を用意した。

(2) イメージ情報

流通している商品カタログは単品カタログから総合カタログまで多種にわたっている。

CD-ROMに inputs するカタログの単位としては各商品カタログをそれぞれで閉じた構成とさせるために総合カタログ的なものについては、各商品に着目し、関連する記述部分と表紙、裏表紙でまとめたものとした。

このようにカタログを再整理した上で、全体に通し番号を付与し、このページ情報をカタログと共にCD-ROMに格納することにより、インデクス情報とのリンクを可能

とした。

カタログのCD-ROMへの入力条件としては、サイズはA4、データ記録密度は8本/mm、データ圧縮方式はMMR(MRI)圧縮方式を採用した。

本実験では、入手した商品カタログの中から、単品カタログおよび総合カタログをサンプリングし、化学製品メーカー7社について、商品の数として600件、全体として3,000頁をCD-ROMに入力した。

今回、作成したCD-ROMディスクのラベルを図15-3に示す。



図15-3 今回作成したCD-ROMディスク

15.2.4 CD-ROM検索システム

CD-ROM検索システムはパソコン上に蓄積されたインデクス情報を検索し、必要な商品カタログを高精細ディスプレイまたはLBPにイメージ出力表示するためのアプリケーションソフトウェアである。

本システムは以下の特徴と機能を持つ。

- ① 統一された画面構成でメニュー展開による利用方式をとり、またファンクションキーによるワンタッチ操作で前後方への画面遷移が可能で操作性を高めている。
- ② 検索機能を一画面に集約することにより、手続きをスッキリさせると共に、ヘルプ画面を用意し、利用者を支援する。
- ③ 検索項目は7項目で項目内はOR条件、項目間はAND条件の論理演算による絞り込みが可能である。また文字情報の語句指定は完全一致、中間一致、前後方一致のいずれの方式もとれる。
- ④ 検索された商品の情報について、一覧表の形式での簡易表示と各商品についての詳細表示の2段階の機能をもっている。
- ⑤ 検索手続きと連動して必要な商品カタログを即時に出力表示する。

画面は初期画面の操作後

- ① 検索画面
- ② 簡易表示画面
- ③ 詳細表示画面
- ④ 高精細ディスプレイコントロール画面

の順で制御される。

検索画面においては、コードまたは文字列により、検索指定を行うが、名検索項目に対してヘルプ画面が用意されており、ヘルプ画面での入力も可能となっている。

検索された商品について簡易表示、詳細表示の後、高精細ディスプレイコントロール画面において、高精細ディスプレイへの商品カタログの表示、ページめくり、またLBPへの印刷を指示する。

①～④の画面を図15-4に示す。

検索条件入力画面

(1) 商品名 [] [] []

(2) 会社コード [] [] []

(3) 分野コード [] [] []

(4) 成分化学分類コード [] [] []

(5) 成分名 [] [] []

(6) 成分分子式 [] [] []

(7) CAS-REGISTRY-NO [] - [] - [] [] - [] - []

検索条件を入力してください。

検索件数： 件 処理状況：検索実行中

F1: 簡易画面 F2: 検索実行 F3: 条件クリア F4: ヘルプ F5: 終了

検索結果一覧画面

検索件数： 頁数： /

番号	商品名	分野	頁数
[]			
[]			
[]			
[]			
[]			
[]			
[]			

詳細表示する商品

F1: 詳細画面 F2: 前頁

検索結果詳細画面

商品名： CAS-REGISTRY-NO:

会社名： TEL FAX

所在地： 〒

分野：

成分名： 分類：

分子式：

成分名： 分類：

分子式：

成分名： 分類：

分子式：

F1: 前頁表示 F2: 簡易画面 F3: 検索画面 F4: 次頁 F5: 先頭頁

F6: 前頁 F7: 次頁 F8: 印刷 F9: 印刷 F10: 印刷

高細格DSP表示画面

商品名： 頁数 (/)

高細格DSPに出力中です。

印刷中

F1: DSP表示 F2: 詳細画面 F3: 検索画面 F4: 次頁 F5: 先頭頁

F6: 前頁 F7: 次頁 F8: 印刷 F9: 印刷 F10: 印刷

図 15-4 CD-ROM 検索システムの画面構成

15.2.5 実験システムの評価

今回、開発したCD-ROM実験システムについて、検索システム、イメージ情報、インデクス情報の面より検討を加えた。

以下にその要点を示す。

(1) 検索システム

- ① メニュー展開による利用方法の他に、コマンド形式等を併用し、慣れたユーザが簡単な手続きで、複雑な論理演算操作等ができるよう検討する。
- ② 利用者に固有な質問の登録、分野、成分分類等のユーザ独自の定義等ユーザに密着した機能をもたせる。
- ③ 分子式等の表現方法に自由度をもたせること、また成分名についてシステムを検討するといった使い勝手の良いシステムとする。

(2) イメージ情報

- ① 写真等のデータ圧縮効率を高めるための入力方法を更に検討する。
- ② 表紙、裏表紙の削除や収録内容を構造式や図化情報等イメージ情報としてのメリットを発揮できるものに限定して、収録商品の数を増やすことが考えられるが、カタログを生形の形で提供するという要求と合わせ、最適な方法を選択する。
- ③ 商品カタログは一般にカラー印刷であり、鮮度を上げるために工夫が必要である。
- ④ CD-ROMの提供に際して、メーカーまたは分野等、CD-ROM一枚の収録単位をどのようにするかを検討が重要である。

(3) インデクス情報

- ① インデクス情報としてデータベース化するデータ項目は利用者の利用目的を規定する重要な要素となるものであり、今後より付加価値の高いものにしていく。
- ② CD-ROM検索に適した有用なDBMSを総合的な見地より決定していく。

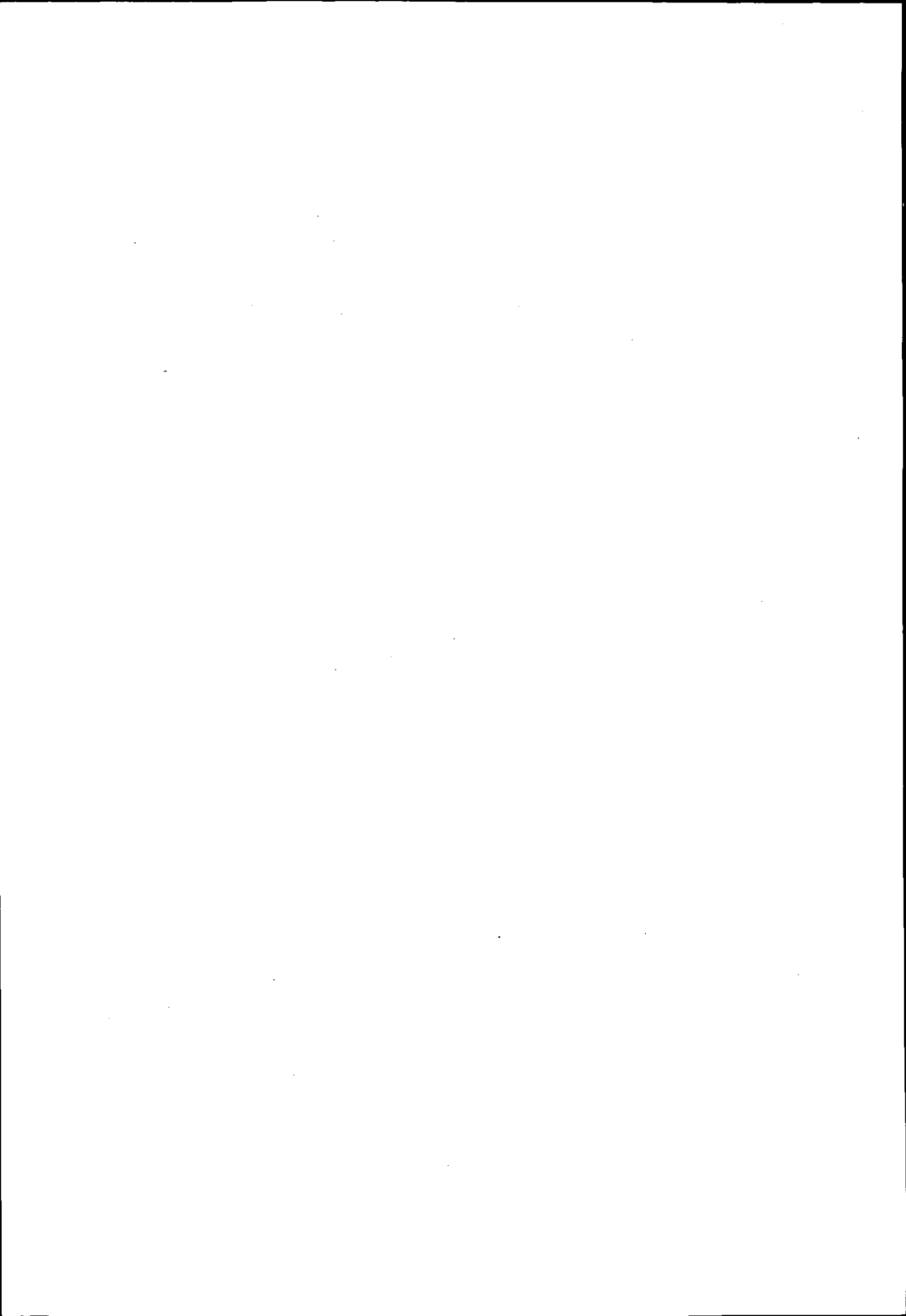
15.3 今後の展望

総合的な実用化合物データベースシステムの構築に向けて、その第一歩としてCD-ROM用実験システムの開発を実施した。

今後、本システムの継続的試用・評価、実験サービスを行い、CD-ROMサービスの商用化を図っていきたい。

また、一方でCD-ROMとの整合性をとりながら、オンラインデータベースの構築を目指し、総合的なサービス形態の提供を追求する。

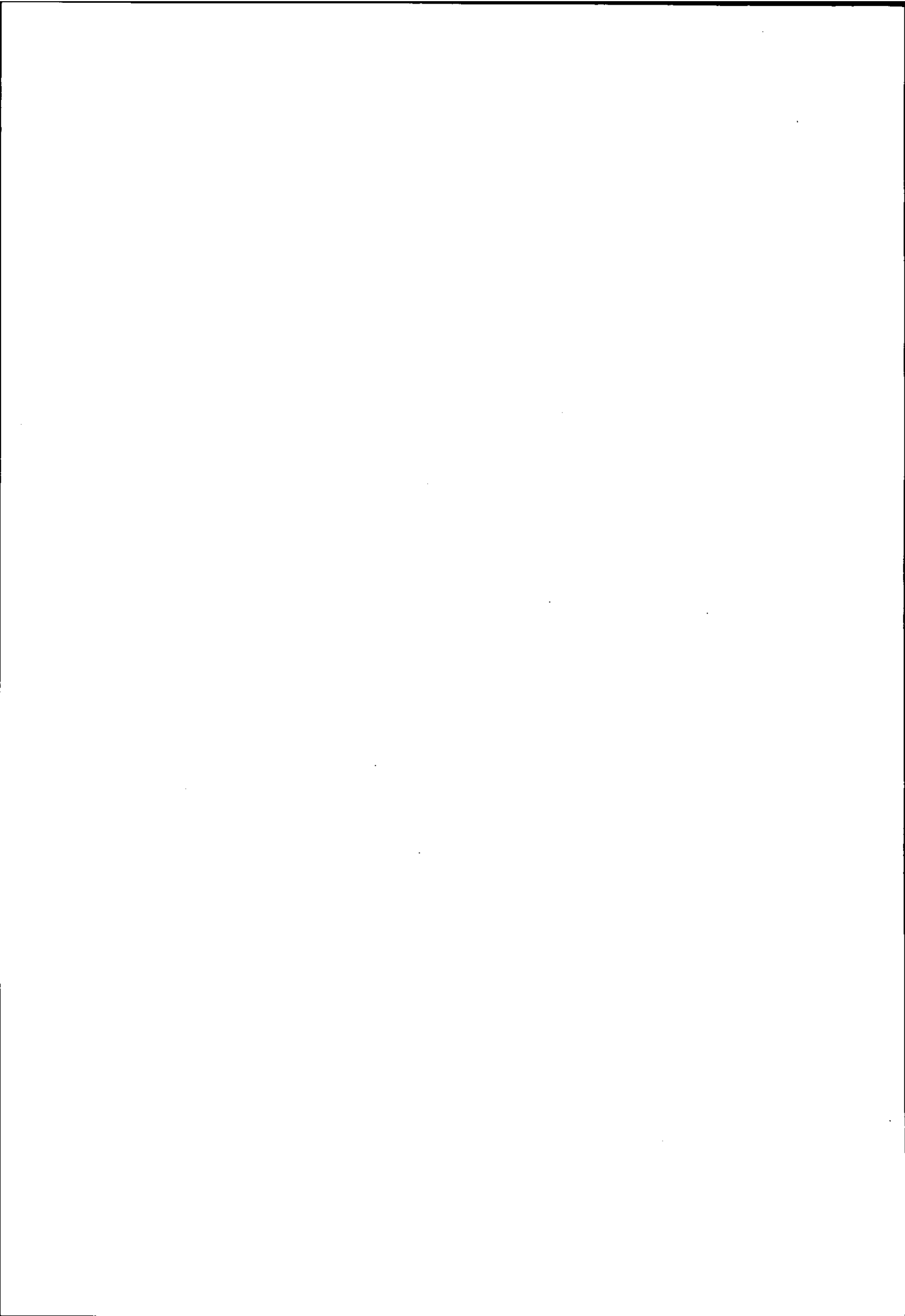
本構想は化学文献等の商用データベース、プライベートなデータベースとの連動、また商品の直接的なオーダーサービスの可能性をも内包しており、ユーザーズに大きく応えられるものである。



16 SiC系ファインセラミックスの

成分特許データベースの構築

神鋼リサーチ株式会社



16. SiC系ファインセラミックスの成分特許データベースの構築

16.1 目 的

ファインセラミックスについては、通商産業省の次世代産業大型プロジェクトテーマに指定されており、参加企業数も著しく多く、将来の市場拡大が予想されている。

ファインセラミックスの成分特許については、特許件数および添加化合物の種類が多く、また、添加化合物の組合せによる特許出願も多く、特許公報の調査・作表のみでは、現在、研究開発中または着手予定のテーマが特許侵害をしているか否かの調査が極めて困難であるので、成分特許データベースの構築が必要である。

昨年、弊社では、特許情報の有効活用および研究開発効率化の支援を目的として、SiN系ファインセラミックスの成分特許データベースを構築したので、本年度は、SiC系ファインセラミックスの成分特許データベースの構築を行った。

16.2 概 要

日本特許情報機構(JAPIO)のデータベース(PATOLIS)を情報源とし、公開特許公報の収録が開始された昭和46年7月以降の公開特許公報および公告特許公報よりデータ収集を行った。

データ収集の方法は、PATOLISの機械検索を中心とし、検索もれの有無については、特許公報1年分のマニュアル検索によりチェックを行い、その結果により検索式を補正し、特許検索を実施した。

SiC系ファインセラミックス成分特許データベースは、データベースソフトとしてd-BASE IIを使用し、特許番号(公開または公告)、発明の名称、出願人、出願日、特許日(公開または公告)、添加化合物名称、成分添加範囲を収録した。

検索ソフトは、d-BASE IIまたは市販の流通ソフトを使用することにより、AND、OR、NOTなどの論理演算子、>、<、=などの比較演算子、+、-、*、/などの算術演算子、文字列などの部分検索が利用でき、更に、出力様式および出力項目は自由に設定できる。

SiC系ファインセラミックス成分特許データベースのデータ数は、現在特許件数で630件、特許請求範囲項目件数で1,377件、化合物の組合せ件数(データ数)で5,837件を収録している。

16.3 実施内容

16.3.1 情報収集

(1) PATOLIS 検索式の作成

① 予備検索

昭和60年1月から昭和61年12月迄を検索範囲とし、下記の予備検索を行い、SiC系ファインセラミックスに関するIPCおよびフリーキーワードのもれを調査した。

- 01, F514517 + F514535 + F514539
(SiC) (SiC質) (SiC粉)
- 02, (F542294 + F542328) * F530951
(シリコン) (シリコン) (カーバイド)
- 03, F638342 * F733061
(珪素) (炭化物)
- 04, (F638342 + F638344 + F638357 + F638358
(珪素) (珪素化) (珪素質) (珪素製)
+ 638360 + F638363) * F733035
(珪素体) (珪素粉) (炭化)
- 05, C04B35/52 + C04B35/54 + C04B35/56 + C04B35/64
(炭素または (グラファイト基) (炭化物基) (焼結方法)
炭化物基)
- 06, C04B35/66 + C04B35/70 + C04B35/80 + C04B35/58,106
(耐火材料 (金属質を (繊維、 (けい化物基)
又は固め材) 含むセラミック) ウィスカ-類)
- 07, C04B35/18 + \$C04B35/00
(珪酸塩基を (組成に特徴をもつ
アルミに富むもの) セラミックで下位分類を含む)
- 08, &01 + &02 + &03 + &04 (式01~式04の合計)
- 09, &05 + &06 + &07 (式05~式07の合計)
- 10, &08 # &09 (式09以外の式08のもの) ——> オラインにより抄録文手記
- 11, &09 # &08 (式08以外の式09のもの) ——> オラインにより要約文手記
- 12, &08 * &09 (式08と式09を同時に
満足するもの) ——> オラインにより抄録文手記

② SiC系ファインセラミックスPATOLIS検索式の作成

予備検索により入手した特許公報抄録文、特許公報要約文およびマニュアル検索の結果より、SiC系ファインセラミックス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。

(2) 特許検索および特許公報の選択

作成した検索式により、昭和46年7月から昭和62年3月迄の特許検索を実施した結果該当特許1,961件に対し、SiC系ファインセラミックス成分特許に関する公報は、512件（公開公報：432件、公告公報：80件）で、適合率は26.7%であった。

SiC系ファインセラミックスに関する特許公報のうち、明らかに製法特許に属するものは収録対象外としたが、製法特許で少しでも成分の規制が行われている特許公報については収録対象とした。

16.3.2 データ作成

(1) 成分特許表の作成

成分特許請求の範囲記載の方法としては、例えば、化合物A, B, C, D, Eの1種または2種以上を1~3%と、化合物F, G, H, I, Jの1種または2種以上を5~10%と残部が炭化珪素からなることを特徴とする焼結体材料などと表現されている。

また、特許請求範囲が2項以上で表現される場合は、第1項の規制内容に追加して規制する場合と第1項の一部を入れかえる場合と第1項の規制範囲を狭く特定する場合がある。これらの方法で表現されている特許公報の内容を表形式で表現するために成分特許表を作成した。

(2) 化合物名称変換記号表の作成

特許公報では、化合物の名称の表現方法に各種のものがある。例えば、アルミナは、 Al_2O_3 、Al酸化物、酸化アルミニウム、アルミニウム酸化物などと表現されている。化合物検索を行う場合、これらの表現をそのまま使用すると検索式が著しく長くなり、また、検索もれの可能性も大きくなる。従って化合物検索のために作成した化合物名称変換記号表では、例えば、上記のアルミナの場合はAl酸化物という表現を用いることにした。例えば、Al酸化物を検索する場合は、変換記号“AAO”を用いる。化合物名称変換記号表は、A-HYOからZ-HYOまでで構成されており、それぞれの表は記号フィールドと化合物名称フィールドからなっている。検索の場合は上記の例の如く記号は3桁で使用するが、データ更新時などのデータ処理を容易にするため、化合物名

称変換記号表の記号は各表とも2桁で表現した。

16.3.3 データベース構築

(1) 収録データ項目と収録法

公開・公告の区別，特許番号（公開・公告），特許請求範囲項目番号，特許日（公開・公告），出願人，出願日，発明の名称，添加化合物の名称および成分添加範囲を収録した。

① 公開・公告の区別

公開特許公報は特開昭，公告特許公報は特公昭とし，それぞれ記号A，Bを並記した。

② 特許番号（公開，公告）

例えば，S 60 - 11255 などと表示し，左づめで入力した。

③ 特許請求範囲項目番号は，2桁で表示した。

④ 特許年月日（公開，公告）および出願日

例えば，S 60 - 11 - 12 などと表示し，左づめで入力した。

⑤ 出願人

例えば，株式会社神戸製鋼所は神戸製鋼所とし，株式会社を省略し，その他は公報記載どおりとした。（共同出願の場合もすべて収録した。）

ただし，外国出願人で非常に長いものは一部省略した。

尚，出願人については，データ入力と検索を容易にするためにPATOLISの出願人コードとは別に出願人コードを設定した。

⑥ 発明の名称は公報記載のとおりとした。

⑦ 添加化合物の名称

公報記載のとおりとした。ただし，添加化合物の種類が著しく多い場合，例えば，IVa，Va，VIa 族の元素およびそれらの酸化物，窒化物，炭化物の場合，Me，Me酸化物，Me窒化物，Me炭化物，Me = Ti，Zr，Hf，V，Nb，Ta，Cr，Mo，W などと表現した。これにより，ディスクの記憶容量を節約し，出力様式をコンパクトにした。

⑧ 成分添加範囲

モル％はM，体積％はV，反応前の原料組成％はBなどと添加成分数値範囲に記号を付けた。重量％は記号を付けないことにした。例えば，モル％で1～5％は1～5

Mとした。

(2) ファイル構造

SiC系ファインセラミックス成分特許データベースは特許書誌事項に関するファイル、SiC1Bと、添加化合物の名称および成分添加範囲を収録したファイル、SiC2Bから構成されている。

フィールド名で、NO、PATNO、SHUCD、SHUNINCDについてはファイルSiC1B、SiC2Bともに収録している。これは、ファイルの結合処理により収録したものであり、ディスクの記憶容量を著しく増大させない項目を選定し、ファイルSiC1Bの検索結果からファイルSiC2Bの検索を実施するために収録した。

公開・公告の区別であるSHUMS（特開昭，特公昭）は、SHUCD（A，B）のデータを変換することにより収録した。

出願人名称であるSHUNINMSはSHUNICD（出願人コード）のデータを変換することにより収録した。

なお、ファイルSiC2BのSEB9およびSEB10については46桁とし、化合物名称の長いものに対応するようにした。

(3) データシート作成

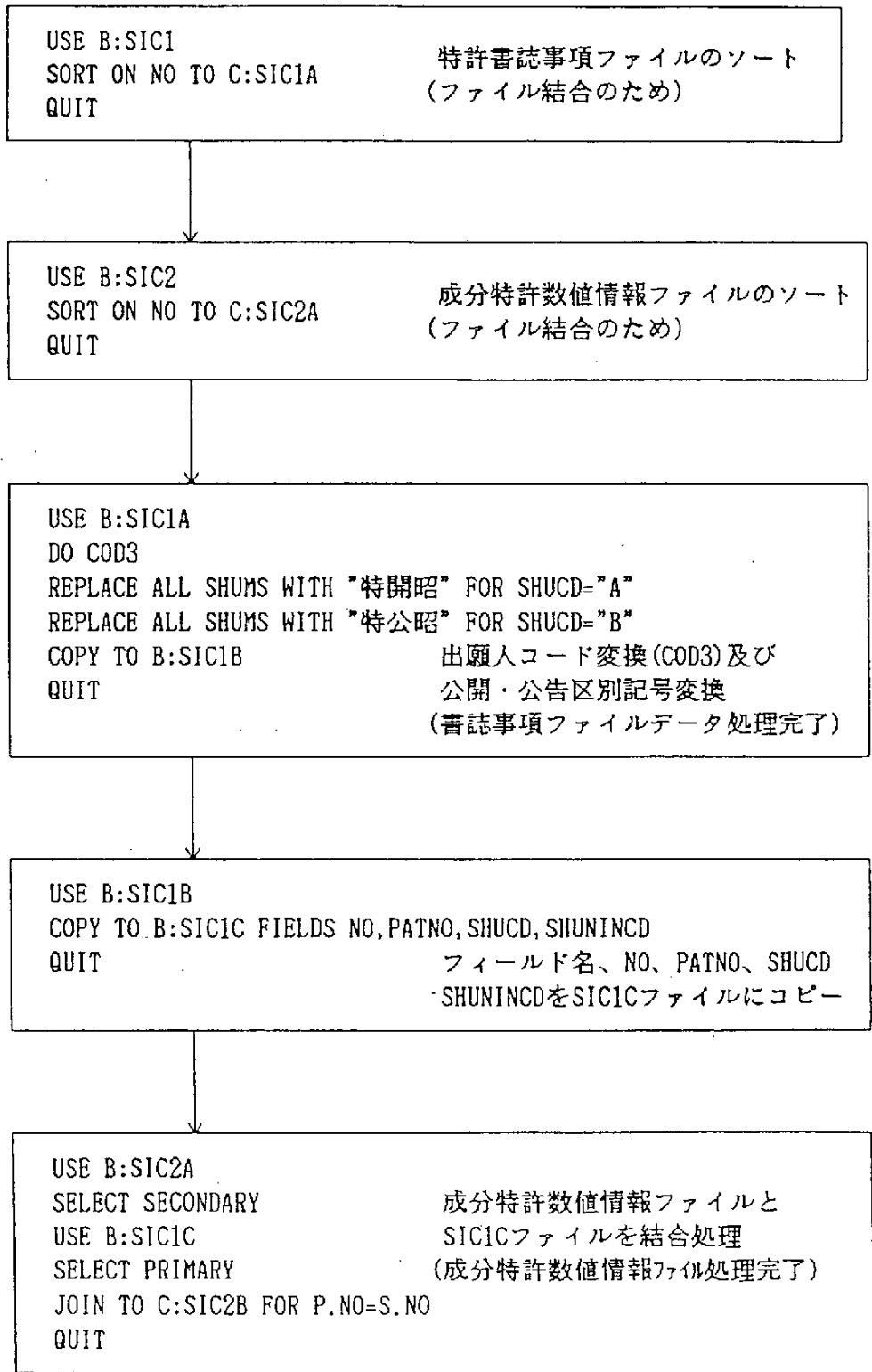
成分特許表のデータを正確にインプットできるように、2種のデータシートを作成した。データシート様式1は特許書誌事項ファイルにデータ入力するためのものであり、データシート様式2は成分特許数値情報ファイルにデータ入力するためのものである。

(4) データ入力

データベースソフトd-BASEⅡを使用してデータ入力を行った。

(5) データ処理

データベースソフトd-BASEⅡを使用し、下記手順によりデータ処理を行った。



16.4 検索例

d-BASE II を使用して入力したデータの検索ソフトとしては数社より発売されているが、今回は、(株)ユーザーズソフト社の dREPORT GENERATOR を使用して、SiC 系ファインセラミックスの成分特許データベースについて、各種検索を実施した。(詳細は報告書を参照)

16.4.1 一般検索例

- (1) ファイル SIC1B で出願人コード K05 のものを検索した。SiC 系ファインセラミックスの成分特許については、神戸製鋼所から 2 件出願され公開されていることが判る。その 2 件の特許の内容について、ファイル SIC2B でフィールド名、NO により検索を実施した。
ファイル SIC1B およびファイル SIC2B の検索に要した時間は、合計で約 5 分であった。
- (2) ファイル SIC1B で、フィールド間の AND 条件、OR 条件を用いた検索を実施した。検索条件は、特許日が昭和 62 年 1 月以降で、出願人コードが K05 または S06 のものである。
- (3) ファイル SIC1B で、完全指定部分一致の検索を実施した。検索条件は発明の名称の中で 4 桁目から 2 桁に "炭" という字を含んでいるものである。
- (4) ファイル SIC1B で 4 つの条件を指定した検索を実施した。検索条件は、特許日が昭和 62 年 1 月以降で、発明の名称の中に "誘電" と "炭化" と "絶縁" という単語を同時に含んでいるものである。
- (5) ファイル SIC2B のフィールド名、SEB1 または SEB2 または SEB3 の中に "アルミナ" という単語を含んでいる特許の検索を実施した。その検索に要した時間は約 2 分であった。なお、"アルミナ" を含むレコード NO のもののみが検索されたが、特許の内容を調査する場合は、フィールド名、NO により検索することにより数分でその内容を知ることができる。

16.4.2 化合物検索例

- (1) 化合物検索例 1.

検索条件、BAG (B 炭化物) または BAF (B 酸化物) または TBE (Ti 硼化物) を含む特許は特許請求範囲項目数で 238 件であった。更に ZAQ (Zr 硼化物) および MAD (Mg 酸化物) を同時に含む特許については、特許請求範囲項目数で 6 件、特許

件数で4件であった。更にFAA(Fe)を同時に含む特許は特許請求範囲項目数で4件、特許件数で2件であった。その内で、昭和56年以降の特許については、特許請求範囲項目数で2件、特許件数で1件であった。また、その内容について調査するためファイルSIC2Bで検索を実施した。ファイルSIC1BおよびファイルSIC2Bの検索に要した時間は合計で約5分であった。

(2) 化合物検索例2.

検索条件、CAA(炭素)と、BAO(Be)またはBAP(Be化合物)またはBAQ(Be珪化物)またはBAR(Be酸化物)またはBAS(Be硝酸化物)またはBAT(Be酢酸化物)またはBAU(Be水酸化物)またはBAV(Be炭化物)またはBAW(Be炭酸化物)またはBAX(Be炭酸水素化物)またはBAY(Be窒化物)またはBAZ(Be硫酸化物)を含む特許は特許請求範囲項目数で89件であった。更にSAV(Si窒化物)およびZAA(Zn)を同時に含む特許は特許請求範囲項目数で2件、特許件数で1件であった。また、その内容について調査するためファイルSIC2Bで検索を実施した。

(3) 化合物検索例3.

検索条件、CAA(炭素)とAAH(Al)と、BAA(B)またはBAB(B化合物)またはBAC(B珪化物)またはBAD(B珪酸化物)またはBAE(B合金)またはBAF(B酸化物)またはBAG(B炭化物)またはBAH(B窒化物)またはBAI(B燐化物)を含む特許は特許請求範囲項目数で89件であった。更にAAU(Al炭化物)を同時に含む特許は特許請求範囲項目数で5件、特許件数で3件であった。更にAAO(Al酸化物)を同時に含む特許は1件であった。またその内容について調査するためファイルSIC2Bで検索を実施した。ファイルSIC1Bの検索に要した時間は約2分で、ファイルSIC2Bの検索に要した時間は約6分であった。

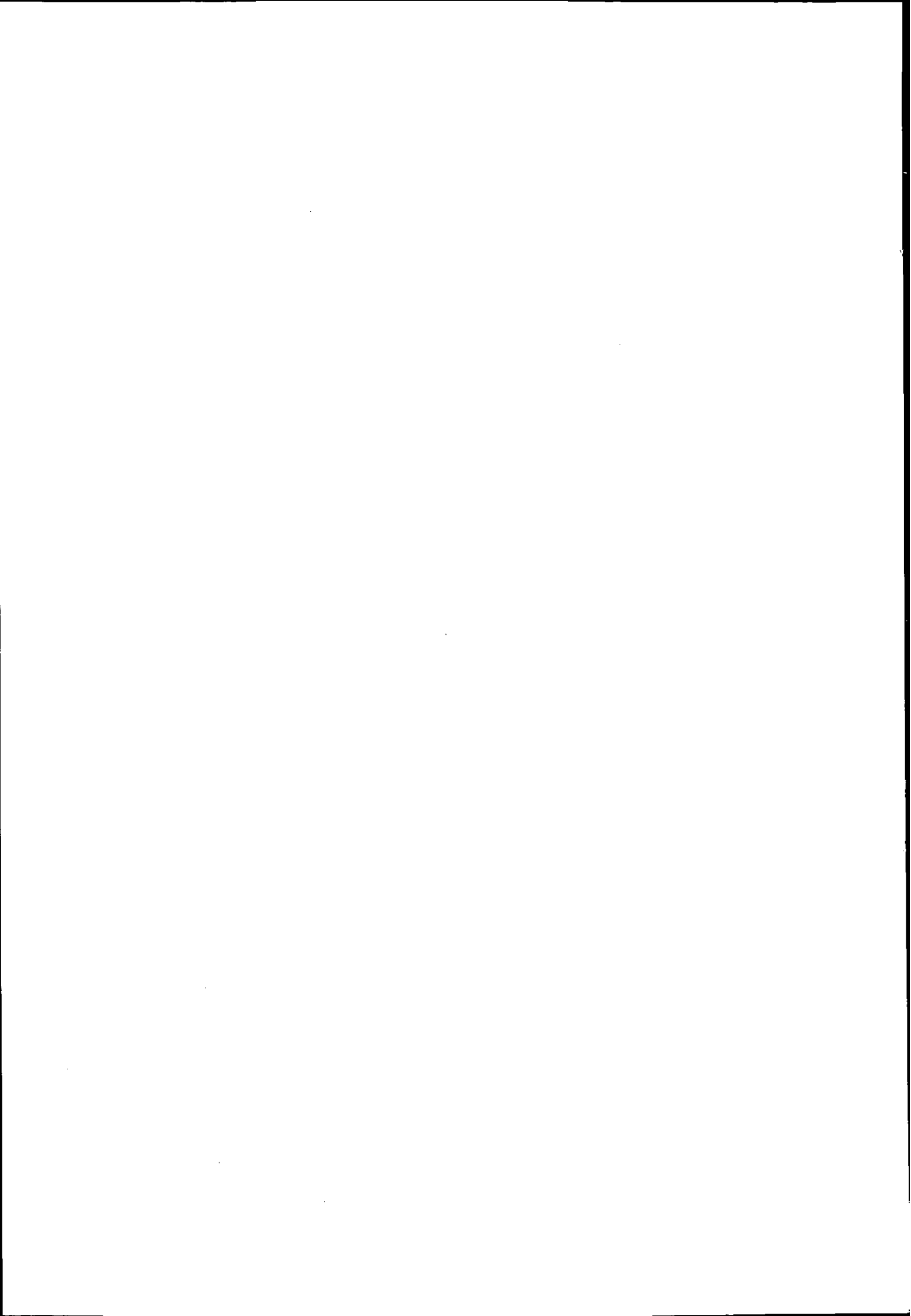
16.5 本データベース利用の効果

- (1) SiC系ファインセラミックスの成分特許については、特許件数および添加化合物の組合せによる特許出願が多く、特許公報の調査・作表のみでは、現在、研究開発中または着手予定のテーマが特許侵害をしているか否かの調査が極めて困難であるが、本データベースを活用することによりSiC系ファインセラミックス成分特許の調査が精度よく容易にかつ迅速にできる。
- (2) SiC系ファインセラミックスの研究開発を実施する場合、焼結助剤などの化合物を添加し、各種の製法で実験を行うことが多い。他社特許を侵害している組成で実験を行った場合には膨

大な費用と時間の無駄になる。実験を行う場合、製法については開発の狙いの変更を行うことが比較的容易であるが、成分については開発の狙いの変更に伴い原料配合の変更が必要であり、それまでの実験が無駄になる。本データベースを活用することにより、このような無駄を避けることができ、研究開発の効率化に大きく寄与できるものである。

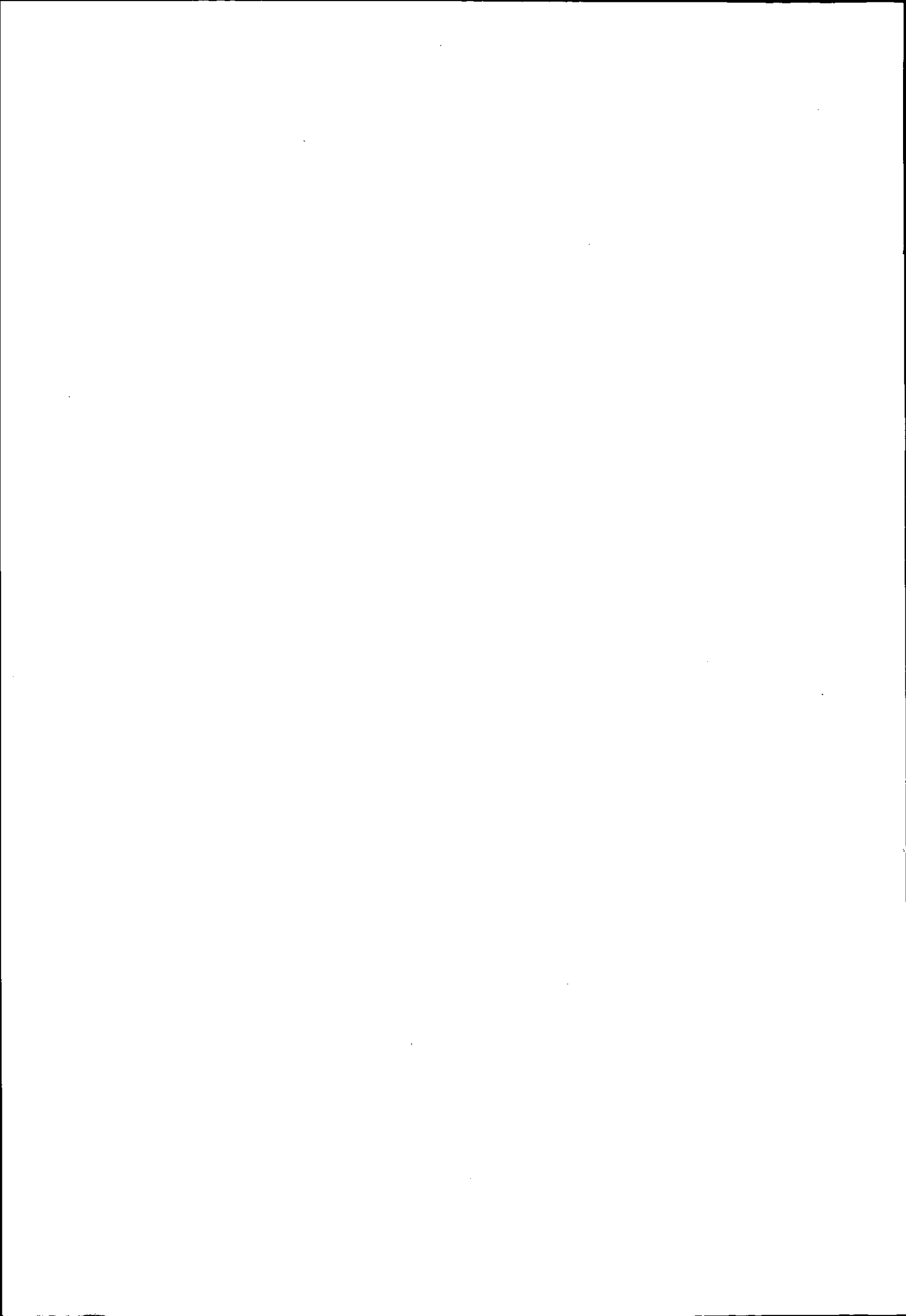
16.6 今後の課題と展望

- (1) エンジニアリングファインセラミックスは、SiN系、SiC系およびジルコニア系が主要なものである。SiN系およびSiC系の成分特許データベース構築は完了したので、ジルコニア系についても成分特許データベースを構築すること。
- (2) ファインセラミックスについては添加成分のみでなく製法についても研究開発が盛んであり、上記3品種の製法特許データベースを構築し、エンジニアリングファインセラミックス特許データベースとして集大成すること。
- (3) ファインセラミックスについては通商産業省の次世代産業大型プロジェクトテーマにも指定されており、将来とも長期にわたって研究開発が継続されるテーマであり、参加企業数も多く、本データベースは広く活用されるものと考えられる。



17 化審法既存化学物質リストの データベースの構築

社団法人 日本化学物質安全・情報センター



17. 化審法既存化学物質リストのデータベースの構築

17.1 構築の目的と背景

PCB等の問題を契機として環境を通じての人の健康に対する悪影響を防止するため、昭和48年10月16日付で「化学物質の審査及び製造等の規制に関する法律」（通称、化審法）が世界に先がけて公布され、新たに化学物質を製造または輸入しようとする場合、事前に所定の試験データとともに国に届けて安全性の確認を受けなければならないとする、いわゆる「事前審査制度」が確立された。

本法公布の際、現に業として製造または輸入されていた物質は事前審査の対象とならない。そして、これらの物質は官報に公示され「既存化学物質名簿」としてまとめられている。また本法施行後、審査を受け安全性が確認された化学物質は毎年数回官報にいわゆる「白」告示物質として公示されている。これらの物質は先の既存化学物質に対して「新規化学物質」と称され、リストアップされてまとめられている。既存化学物質名簿中の化学物質の数は、固定されて増加することはないが、新規化学物質としてリストアップされる物質の数は官報に公示される毎に増加する。

既存化学物質名簿には約20,000の物質名が記載されている。また新規化学物質として、これまでにリストアップされた物質数は約3,000物質に達し、従来は毎年約300以上の物質が新たに加えられている。これらの物質名簿は通商産業省の監修のもとに整理され冊子体として刊行されている。

この冊子体には既存化学物質として約20,000の物質名称が記載されているが、一つの化合物を表わす名称だけでなく、多数の種類化合物を表わす、多重表現化合物の名称（我々はこのような物質を“総称名物質”と定義する）も含まれており実質的には数百万の単一化合物を含んでいると見てよい。

冊子体に記載されているデータは、日本語の名称、簡単な構造分類と官報公示整理番号のみであり、冊子体中の目的とする物質を検索するには、慣れた人が行ってもかなりの時間を必要としている。

一方、海外の米国、EC諸国でも化審法に相当する法律がその後制定され、既存化学物質名簿が作成されている。これらの名簿では、物質名称に加えて、化合物を同定するために国際的に使用されているChemical Abstracts (CAS)登録番号、分子式が記載され、更に検索に便利のように編集の上、収載されている。またデータベースも作成されている。従って海外からも日本

の名簿に分子式，CAS登録番号を付与することが強く要望されている。またEC委員会の既存化学物質名簿作成当事者からも要望が出されている状況である。

また，昭和60年に実施した168社に対するアンケート調査で，本データベース構築への要望が強く示されていた。即ち，冊子体の「既存化学物質名簿」の使用頻度は年間1万件あり，利用目的は届出に使う（36%），研究開発の段階で検討する（58%），という傾向にあった。そしてその大半がデータベースが完成した時点で利用したいとしている。

上記のような背景から，本データベースのニーズは高いと判断し，昭和60年度より，構築に着手した。

なお，本データベースでは，「既存化学物質名簿（約2万件固定）」と「新規“白”告示物質名簿」（約3千件毎年追加）とを合せたものを構築した。そして，両者を合せたものを「既存化学物質リスト」と称した。

17.2 予想される効果

- (1) 化審法届出の要否の判断が容易
- (2) 新素材等の開発計画に寄与
- (3) 安全性試験結果と物質情報（分子式）の相関
- (4) 海外からの調査の要望に対応
- (5) CAS登録番号の調査が可能

17.3 本データベース構築の基本方針

- (1) 通常の化学的知識があれば利用出来る
- (2) 検索モレが無いように設計
- (3) 分子式の検索では部分検索も出来る
- (4) 構造式検索は行わない
- (5) CAS登録番号を付与する
- (6) 海外からの利用も可能とする
- (7) 維持，運用コストの経済性を考慮
- (8) 既存データベースマネジメントシステムを利用

17.4 構築経過

本データベースは、図17-1のような経過で構築される。このうち、各年度における構築内容は次のとおりである。

昭和60年度

- (1) システム基本設計
- (2) 物質名(日本語)ファイル
- (3) 物質名(英語)ファイル

昭和61年度

- (1) システム詳細設計
- (2) プログラム・コーディング
- (3) 2,800物質のデータベースの構築

本年度は、(1)物質名(カタカナ)ファイルを完成させるとともに、2,800物質データベースの検索テストをくりかえし、問題点をシステム担当者(NTT)とともに解決した。

そして、本年度までに調査完了したデータをパンチ入力し、昭和63年度からの一般公開準備にとりかかっている。

17.5 今後の課題と展望

本データベースの構築は、昭和60年度から開始され、昭和62年度で完成、終了した。

構築完了後における運営方式およびサービス形態等を調査するため、昭和62年7月、この分野の権威者および(社)JETOC(社団法人日本化学物質安全・情報センター)有志会員の社内専門家から成るワーキンググループ(WG)を結成し、種々の角度から検討した。

また、このWG検討会の中では、期間中に実施したアンケート調査(201社回答)の結果にもとづき、需要予測、公開方法、公開方針などを検討した。

その結果、以下のように、本データベースの運営を行うこととなった。

- (1) 63年10月より64年3月までを試行期間とし、センターが問い合わせを受けてオンライン検索結果を回答する。併せてオンライン検索サービス開始の準備を進める。
- (2) 64年4月よりオンラインサービスを開始する。
- (3) 当分の間、(営業収支の見通し等のつくまで)印刷物、テープ等の販売形での提供は行わな

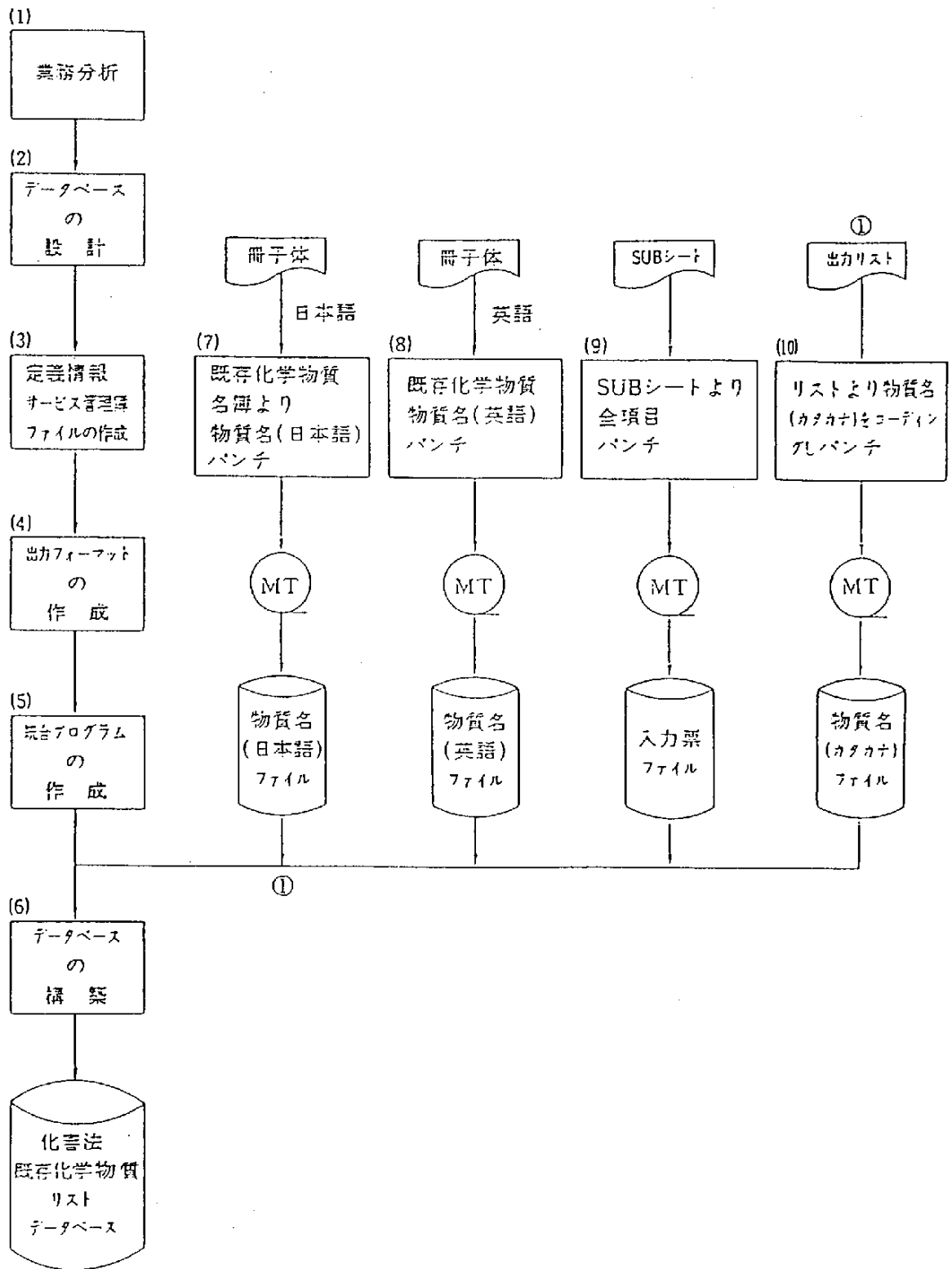


図17-1 データベース作成の流れ

い。

- (4) 料金は、会員、非会員の2本建てとする。
- (5) 2, 3, 4, 5類を優先して公開し、続いて6, 7類を公開する。
- (6) 新規物質の入力も行い、同時にサービスする。

尚、一般公開時における JETOC と NTT および利用者との関係は、図17-2のとおりである。

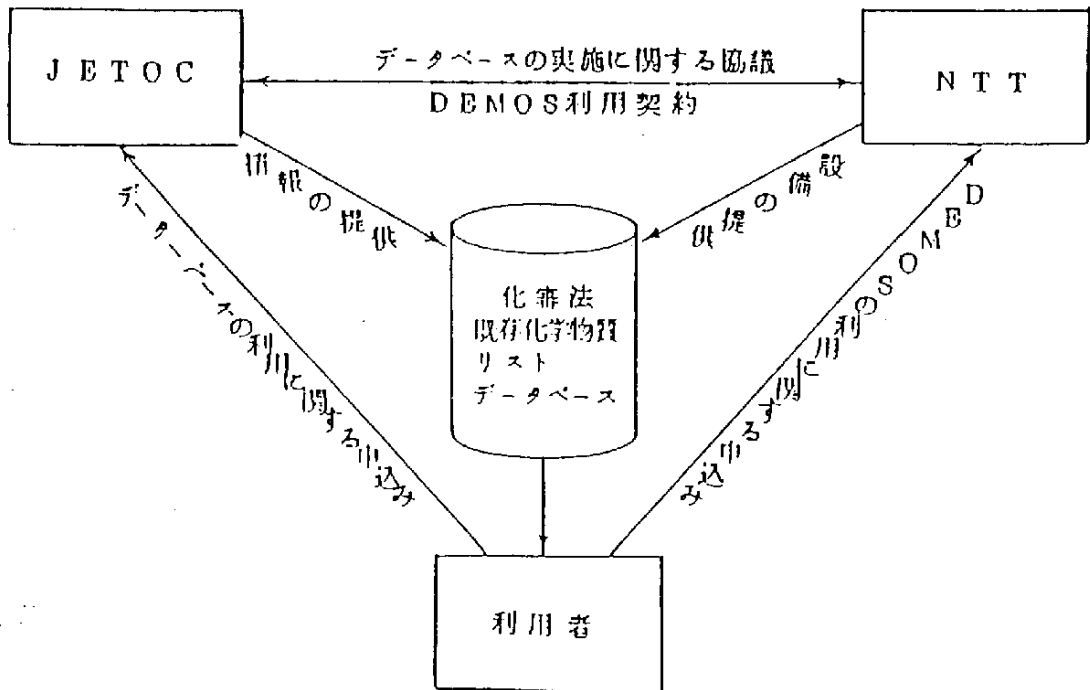


図17-2 JETOCとNTTおよび利用者との関係

検索の実例

(1) 分子式から

*** RET 2 開始 87-08-28

サービス提供者名は? JETOC

サービス名は? DATABASE

*** DATABASE 開始

#: /LOG

[01]: (MF)C3H16N2O

*** 1 件です。

[02]: /KPR F5

JETOC / KASHIN

87年08月28日

1

【物質番号: AN】02-0381-000

【物質名: NJ】

1,3-ジアミノ-2-プロパノール

【物質名: NE】

1,3-Diamino-2-propanol

【分子式: MF】C3H16N2O

【CASN: RN】616-29-5

【公示番号: NG】2-315

【構造分類: SS】02

【物質分類: SC】S

【WLN: WL】Z1YQ1Z

【TSCA: TS】TS 【ENC S: EC】EC

【分解性: BD】 【濃縮性: BA】 【毒性: TX】 【物質規制: C】

【物質情報: CI】

【告示年月: DE】

【全環数: TR】

【環系数: RS】

【成分環数: NR】

【環大きさ: SR】

【環組成: AR】

【名称KW: KW】プロパノール, アミノ, シ

【総称名称: GN】

*** 出力件数 - 1

#: /EOF

*** DATABASE 終了

*** RET 2 終了

シッコウ オフリ. シッコウ ユニット 1.75 (F)

(2) CAS 番号から

サービス名は? DATABASE

*** DATABASE 開始

※: /LOG

[01]: RN-596-01-0

*** 1 件です。

[02]: /KPR F4

J E T O C / K A S H I N

8 7 年 0 8 月 2 8 日

1

1

【物質番号: AN】 05-0047-000

【物質名: NJ】

3,3-ビス(4-ヒドロキシ-1-ナフチル)-フタリド

【物質名: NE】

3,3-Bis(4-hydroxy-1-naphthyl)-phthalide

【分子式: MF】 C28H18O4

【CASN: RN】 596-01-0

【公示番号: NG】 5-3350

【構造分類: SS】 05

【物質分類: SC】 S

【WLN: WL】

【TSCA: TS】 NTS【ENCs: EC】 EC

【分解性: BD】 【濃縮性: BA】 【毒性: TX】 【物質規制: CS】

【告示年月: DE】

【総称名称: GN】

*** 出力件数 - 1

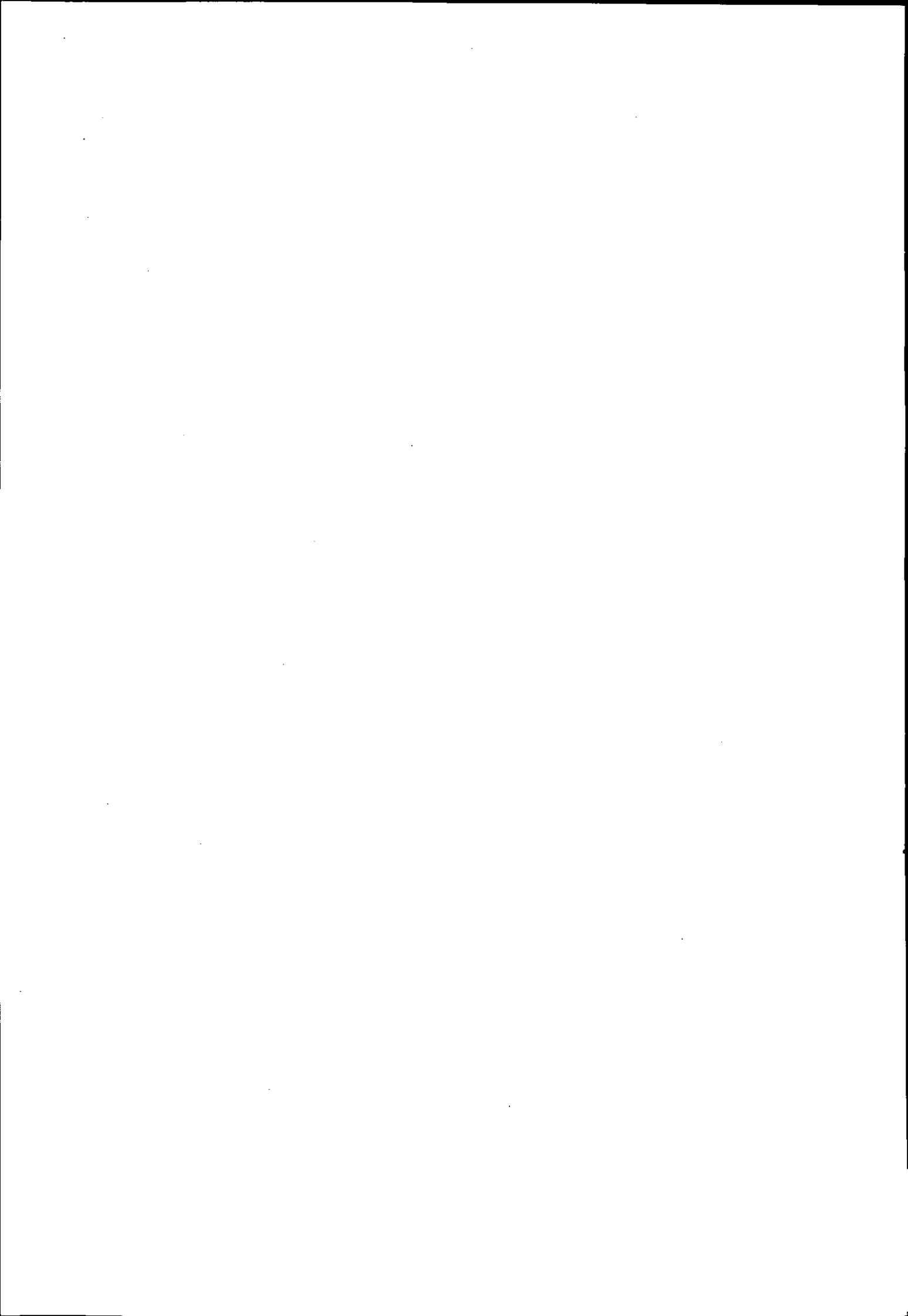
※: /EOF

*** DATABASE 終了

*** RET 2 終了

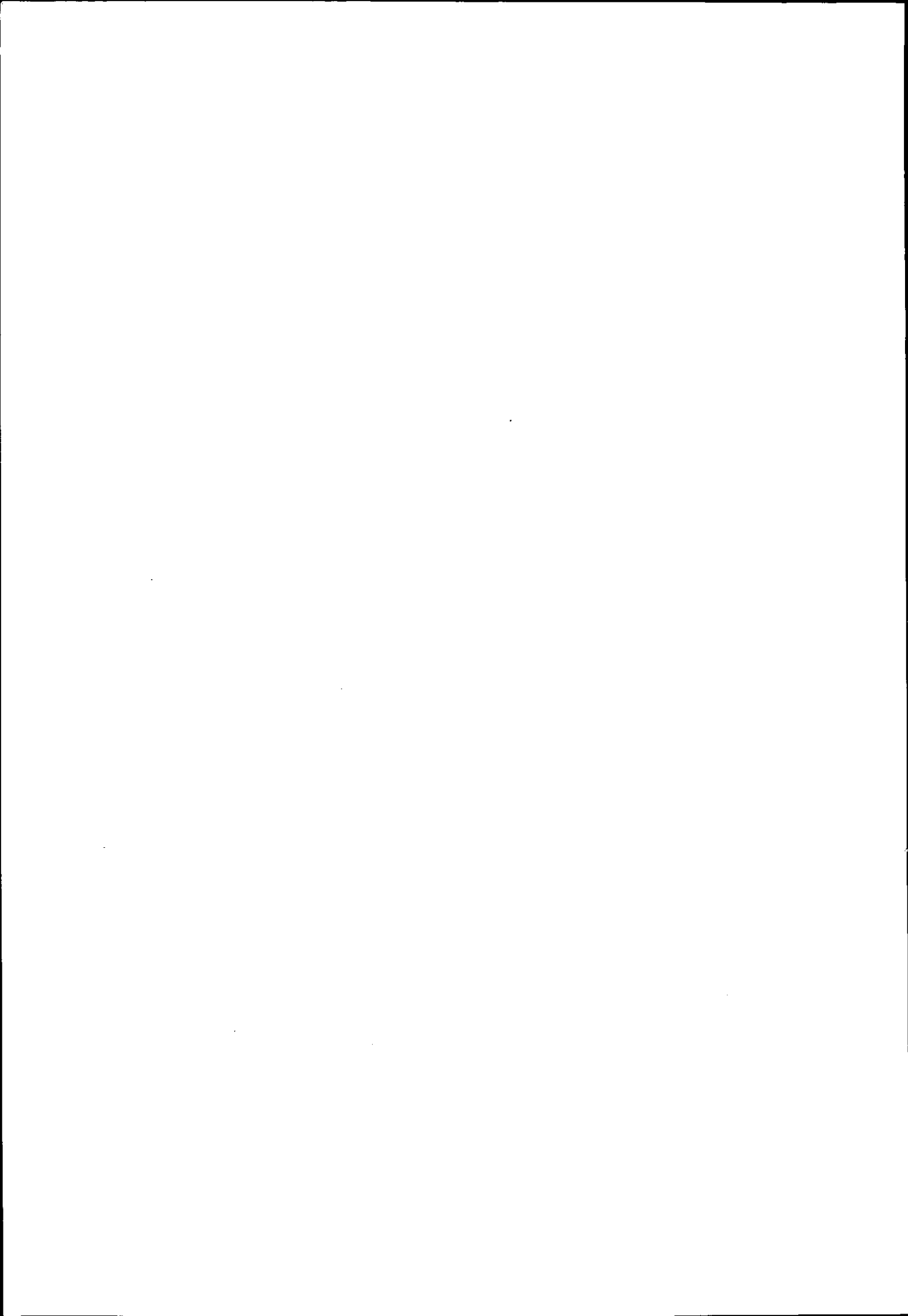
ジッコウ オフリ. ジッコウ ユニット 1.75(F)

/007/ LIB RET2



18 地熱資源利用基本データベースの構築調査

財団法人 新エネルギー財団



18. 地熱資源利用基本データベースの構築調査

18.1 調査の目的

わが国において、地熱エネルギー資源は温泉や火山を通して広く人々に知られている。また、国の政策としても各種地熱資源調査、技術開発を通して、その利用・活用の促進が図られているが、地熱エネルギー資源の有効利用という点ではその特性が十分に発揮されるところまでは至っていない。地熱エネルギー資源の最適利用を推進していくためには、各利用者がその特性を十分に理解してプロジェクトを進める他、地方自治体、学校、報道機関等に対しても広くその特性、有効性、利用方法を啓蒙していくことが必要である。

本調査研究は、関係各官庁、地方自治体、新規参入計画者、研究機関等が使い易い地熱資源利用に着目したデータベースの研究・開発を行い、地熱エネルギー資源の利用促進に資することを目的として行われた。

18.2 調査の方法

本調査研究は新エネルギー財団が、学識経験者、関係事業者等からなる「地熱資源利用基本データベース構築検討委員会」を設置し、その協力の下に行った。

文献調査（財団所有の文献他）および各検討委員の質問票による意見聴取等により基礎的な次にあげる調査を実施した。

(1) データベース構築のニーズ調査

地熱資源利用データベースの構築に関するニーズに関して各委員（地熱資源利用基本データベース構築検討委員）に対し、質問票にて調査した。

(2) 国内外における地熱関係データベースの所在調査

国内および国外のデータベースに関して文献などにより調査を実施した。

(3) データの所在調査

質問票および文献調査により地熱資源利用データベースのデータソースとなる文献の所在を調査した。また、未公開事例データについて、そのデータベースへの取入れ方および取扱いに関して検討を行った。

(4) データの分類・整理方式の検討、要素抽出

質問票および文献により調査し、地熱資源利用データベースに適した分類・整理方式を策定

した。

要素の抽出についても質問票および文献により調査し、事例別に「地熱資源利用データベース」に必要な要素を抽出網羅した。この中から利用頻度、重要度等を考慮して入力対象項目を絞り込んで選定し、「地熱資源利用基本データベース」の項目とした。

(5) 入力・検索方式の検討

データの準備、入力作業および検索作業についての各方法を検討し、実施に当たっての留意点の整理を行った。

18.3 調査の成果

これらの基礎的調査を通じ、次のような成果が得られた。

(1) データベース構築のニーズ調査

地熱資源利用データベースのニーズとして、データに対するものとデータベース機能に対するものがあることが分かり、それぞれの内容は次のとおりである。

1) データに対するニーズ

① 地熱資源利用事例に関するもの

- ・利用施設の所在地、諸元、設計の技術基準、建設費等
- ・利用施設の利用状況
- ・利用地の自然環境（気温、降水量、地形等）
- ・利用地の経済環境（水道料金、電力料金、燃料価格等）
- ・利用地の社会環境（法規制、人口密度、産業構造等）
- ・利用施設の使用地熱流体性状（地熱熱水および地熱蒸気の噴出量、化学成分等）

② 地熱坑井、温泉の位置、性状

- ・所在地、湧出量、化学成分等

③ その他

- ・地熱探査・開発に関するもの

2) データベース機能に関するニーズ

① 地熱資源利用関係の資料の網羅

② キーワードによる容易な検索

③ データの更新、追加による最新データの入手

④ 石油代替効果の算定

- ⑤ 統計処理機能
- ⑥ 図形・グラフ処理機能
- ⑦ 通信ネットワーク化

以上のように、地熱資源利用事例データの他、地熱資源量や地熱資源探査などに関するデータなど多岐多用のニーズがあることが分った。しかし、すべての要望に答えられるデータベースを構築することは現実的には難しく、意見の集中したニーズに絞り、データベース化しようという方針が確認され、地熱資源利用データベースの入力対象は、地熱熱水および地熱蒸気を利用する過程で得られる次の項目の情報とすることとなった。

① 地熱流体に関するもの

地熱熱水および地熱蒸気の噴出量（湧出量）、温度等の物性および化学成分等

② 利用事例に関するもの

利用施設：採取・輸送・利用・熱交換等施設の諸元・特性・材質・建設費等

利用状況：用途別の利用温度、稼働率、経費、利用料金等

③ 利用地に関するもの

自然環境：気象（気温、降水量（降雪量））、地形等

経済環境：公共料金（水道料金、電力料金）エネルギー消費量等

社会環境：法規制、人口密度、産業構造等

④ 関連するもの

研究成果：既刊行物、報告書、雑誌等に発表された地熱資源利用に役立つ成果

その他：補助制度、一般情報等

(2) 国内外における地熱関係データベースの所在調査

地熱関係のデータベースは国内ではSIGMA（System for Interactive Geothermal Mapping and Assessment）、アメリカにおいてはEDB（Energy Data Base）などといった地熱資源の探査データや技術情報などに関するものはあるが、地熱資源利用に着目したデータベースはないことが分かった。

(3) データの所在調査

地熱資源利用データベースのデータソースとしての、地熱資源利用に関する報告書、研究所報告、大学紀要、学会誌など約300文献の所在および現地の事例データの収集方法について整理ができた。

(4) データの分類・整理方式の検討、要素抽出

データの分類・整理の方針が決まり、事業内容、熱水成分、利用規模、利用施設、利用内容、

運用実績，コスト，文献などに関する約 500 項目の要素の抽出ができた。

(5) 入力・検索方式の検討

入力を行う場合の要因の解析，稼働率の取扱い，データの質と量，供給熱量・利用熱量および代替熱量の関係，供給施設と利用施設別入力，使用ソフト，データベースマネジメントシステムなどに関する留意点および検索を行う場合の要因の解析，検索方法の選択，使用ソフト，機密保護，グラフ表示の有無などに関する留意点の整理ができた。

18.4 課題と展望

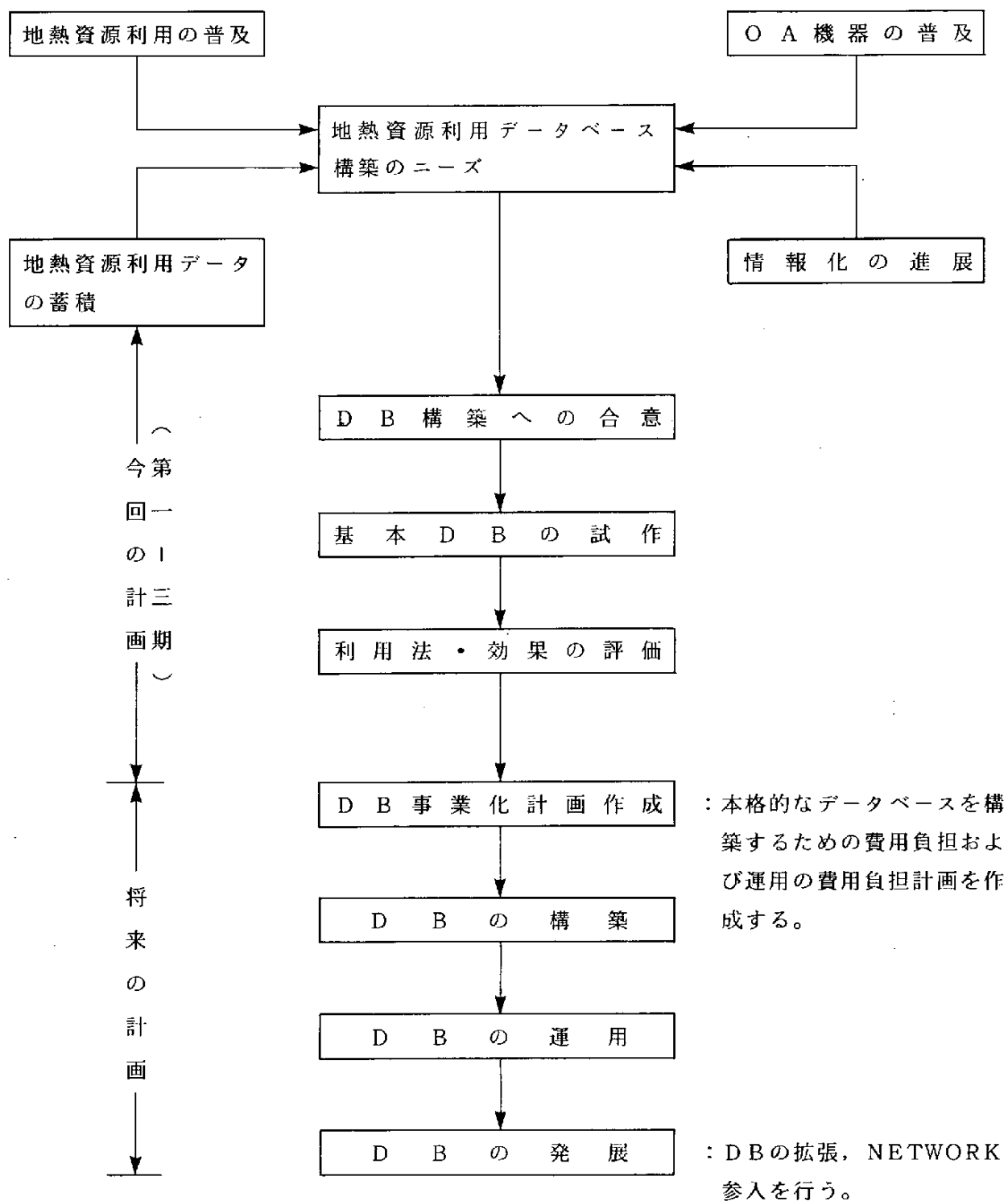
データベース構築のニーズ調査，国内外における地熱関係データベースの所在調査，データの所在調査，データの分類・整理方式の検討，要素の抽出および入力検索方式の検討を実施し，「地熱資源利用データベース」構築の方向付けとデータベース構築に関する基礎的考察を行った。

この結果として，地熱資源利用データベース構築のための基本的合意とその実現の手順（図18-1）が得られた外，入力対象の明確化，文献所在リスト，要素抽出表の作成および入力・検索方式検討上の留意点の整理を行うことができた。

今後，これらの成果を引継いで，入力・検索についての細部検討と，他のデータベースとの相互乗入れを含む運用方式の基礎調査を行うこととなる。

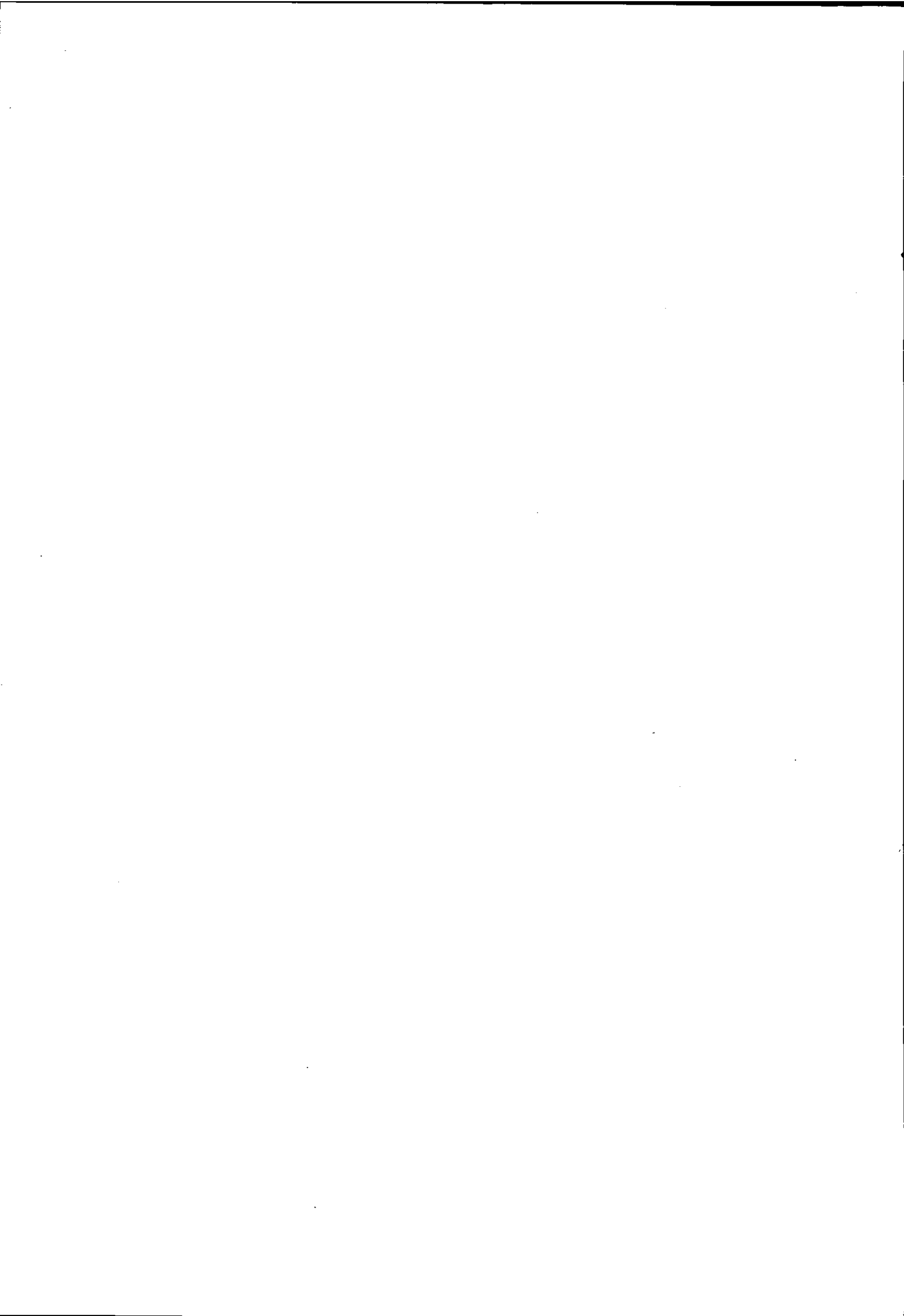
熱水利用に着目したデータベースのニーズと今回の基礎的考察を通じ，次の波及効果が期待され，地熱資源利用の計画の促進に役立つものができるとの確信を得たので，その実現に向けて努力して行きたい。

- ① 地熱資源利用計画の経済性を含めた策定を容易，迅速，精度高く実施できる。
- ② 需要量に対し過大な地熱資源利用の計画を抑制できる。
- ③ 地熱資源利用の普及・啓蒙に役立つ。



(注) 基本 D B : 地熱資源利用基本データベース
 D B : 地熱資源利用データベース
 O A : オフィスオートメーション

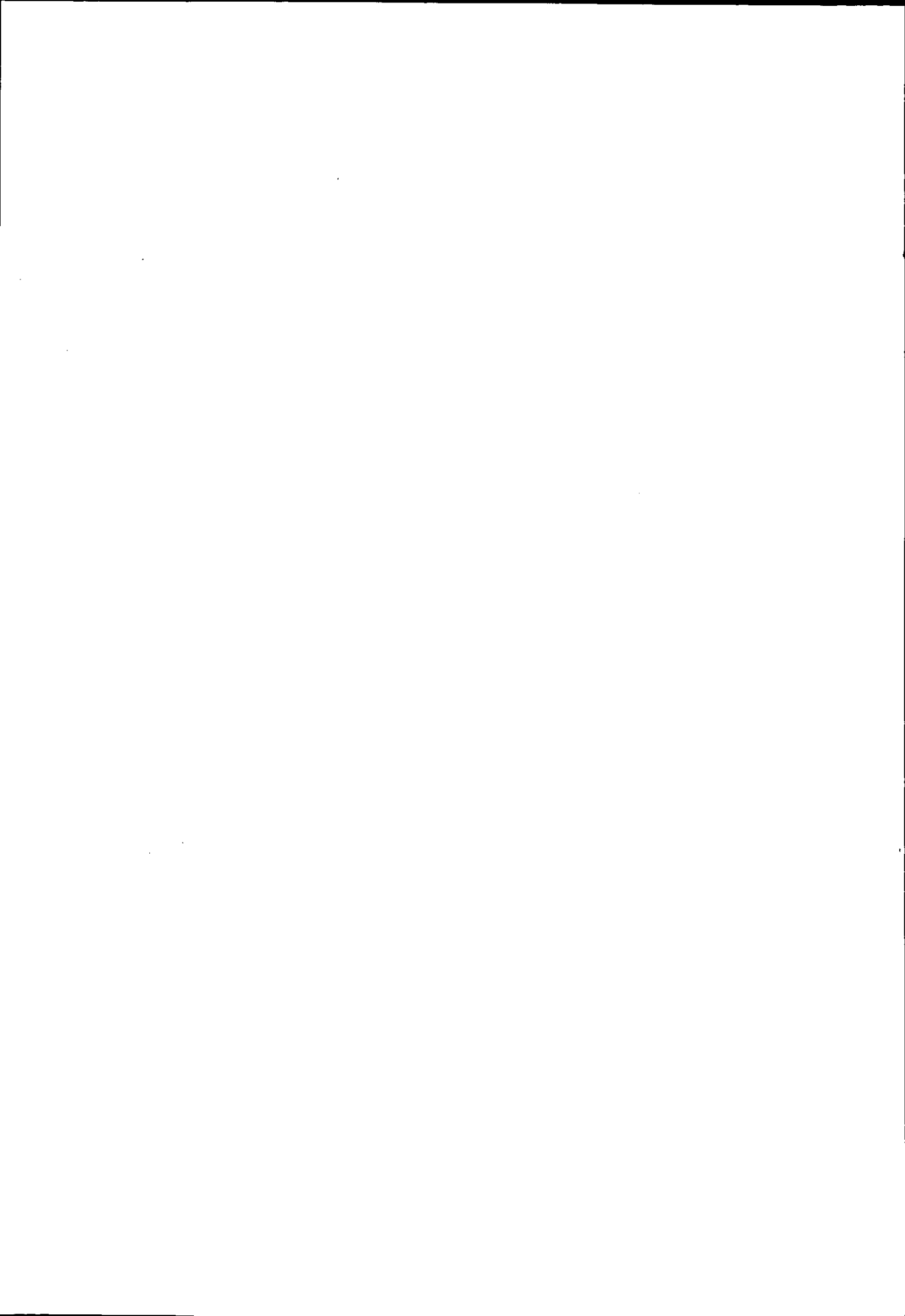
図18-1 地熱資源利用データベース実現の手順



19 省エネルギー技術に関する

データベース構築の調査研究

財団法人 省エネルギーセンター



19. 省エネルギー技術に関するデータベース構築の調査研究

19.1 背景

脆弱なエネルギー供給構造を有するわが国が、今後とも経済の持続的成長、国民生活の安定・向上を図っていくためには、柔軟かつ強靱なエネルギー需給体制を形成することが必要である。このため石油代替エネルギーの開発導入などの供給面はもとより、需要面における実効性のある省エネルギー対策を引き続き推進していく必要がある。

これまで着実に推進してきた省エネルギー政策を、より一層推し進めるためには、産業部門を始め、民生、輸送部門という各分野に対し、それぞれのニーズに合ったきめの細かい省エネルギー対策が必要であり、このためには先進的な省エネルギー技術から身近な省エネルギー機器に至るまで幅広い技術情報を著積し、それぞれのニーズに応じた情報の選択が可能となるデータベースが不可欠である。

また、発展途上国からの省エネルギー協力要請に対し、わが国の省エネルギー技術の発展により得られた成果を提供し協力していくという国際的な責務もあり、そのためにも国内の省エネルギー技術のデータベース化が必要である。

このため、これまで省エネルギーの技術、知識の普及・促進に努めてきた(財)省エネルギーセンターにおいて省エネルギー技術に関するデータベース構築の調査研究を行うことになり、3回にわたる委員会およびワーキンググループ8回を開催し、検討を行った。

19.2 省エネルギーセンターの情報サービス活動の問題点

情報サービス活動の問題点を次の(1)から(5)までに分けて整理する。

- (1) 既存のデータベースでも一部省エネルギーという分類をもっているが、省エネルギー関連の文献をすべては網羅していない。また、わが国のエネルギー技術情報については、数も内容も充実している海外の商用データベースに依存する面が大きいようである。

現在、省エネルギーセンターとしては年1回「省エネルギー情報ガイドブック」を発行し情報提供を行っているが、省エネルギー管理技術そのものは機械工学や化学工学分野の中のひとつの技術として独立したものではなく、あらゆる分野にわたって共通的な管理技術と見るべきであり、その面からの情報が少ないと考えられる。

- (2) 省エネルギーセンターの情報収集・提供は現在、主に印刷物(刊行物)によっているので、

時間的にも質的・量的にもユーザの要求に応えることが難しい状況になってきている。

これらの点は独自に実施した利用者に対するアンケート結果にも示されており、情報入手状態が不十分な理由として、①情報源が限られている、②他にもっと良い情報があるのではないかと思う、などがあげられている。

- (3) 省エネルギー情報は関連分野が多岐にわたるため収集する資料が論文、報告書、事例集、データ集などと多種類になり、手作業やパソコンでは共通の方式で整理・加工、蓄積、提供することはできなくなってきた。
- (4) 情報量が増加するにつれて省エネルギーセンターが扱う資料の量がふえ、資料室における整理・保管が満足に行えなくなっている。また、原資料の整理あるいは二次資料の作成のために、一応、分類およびキーワードを制定してあるが完全な体系になっていない。収集から検索・提供に至る情報管理体制の整備が必要である。
- (5) 海外機関からの省エネルギー技術情報の入手については、既存の商用データベースを利用してもまだまだ不十分であり、欧米先進国への情報提供については省エネルギーセンターとしてはこれから着手する問題である。一方、発展途上国への省エネルギー技術情報の提供は年々増加している状況であり、今後更に強力に行うことが要請されるであろう。

19.3 省エネルギー技術データベース構想

19.3.1 省エネルギーセンターの問題点の解決策

19.2 であげた問題点(1)～(5)の解決策を検討した。

問題点(1)は独り省エネルギーセンターの問題点というよりは日本のエネルギー関連部門すべてに関係することである。エネルギー情報もしくは省エネルギー技術情報を海外の商用オンラインデータベースサービスに依存する面が強いことは、日本で独自の省エネルギーデータベースを構築すべきことを示唆している。また、日本を代表する「JICST科学技術文献データベース」の省エネルギーの分類に含まれる関連の文献が少ないことから、情報源を強化することも含めて、やはり省エネルギーデータベースの構築を行うとともに、適切に、オンライン情報検索を行ってJICSTデータベースにある省エネルギー関連文献の状態を把握しておくことが必要となろう。

問題点(2)の情報収集・提供を時間的・質的・量的に向上するためには、(1)と同様に省エネルギー技術データベース構築とオンライン情報サービスの実施により解決の方向に向か

うことができる。

問題(3)の多種類の資料，増えていく技術分野に対応するための対策は，情報・資料のデータベース化しか考えられない。

問題(4)の分類・キーワードの整備は，データベース構築に際し必然的に行うことになるが，データベース構築と情報システム運営の基礎として，省エネルギーセンターの情報管理体制の確立が極めて重要となる。

問題点(5)の海外機関との情報交流促進，および発展途上国への省エネルギー技術情報の提供促進は，省エネルギー技術データベース構築により発展の方向に進むことは明らかである。

以上のように省エネルギーセンターにおいて，

- ① 省エネルギー技術データベースの構築
- ② オンライン情報検索サービス
- ③ 情報管理体制の確立

を実施することにより諸問題は解決され，または，少なくともその方向へ向かうと考えられる。

19.3.2 省エネルギー技術データベース構築の基本的考え方

次のような基本方針でデータベース構築を進めていく。

- ① データベースの構築は省エネルギーセンターの会員にサービスすること，およびセンター内の業務の効率化を目的に行う。
- ② データベースの利用者は，当初は各産業のエネルギー管理技術者を対象とする。
- ③ データベースの型は文献型とし，ファクト型については将来の検討事項とする。
- ④ データベース化する情報はセンター外部（会員）へ提供できるものを対象とする。
- ⑤ 対象情報は省エネルギーセンターでなければ収集できないものを中心とする。このような情報，資料については過去にさかのぼってデータベース化を行う。
- ⑥ 将来の情報サービスの拡大を考慮して，民生用の省エネルギー技術情報も対象情報の範囲に入れておく。
- ⑦ データベースシステム構築に際し，コンピュータとネットワークを利用してユーザに情報（二次情報）をリアルタイムに提供できるように計画する。
- ⑧ 一次情報（資料）の蓄積・保管にはリアルタイムに資料提供できる方式を計画する。
- ⑨ 省エネルギーセンターがこのデータベースサービスの拠点となる。データベースサー

ビス運用には専門の職員（情報専門員）を置く。

- ⑩ データベースシステムは段階的に支部（所），会員へと増設して行く。

19.4 省エネルギー技術データベース構築案

19.4.1 登録対象情報

省エネルギーセンターが扱っている情報を図19-1に示す。

これらの対象情報を過去の経験から，省エネルギーセンターの独自性が強く必要度の高い順にあげると，次のようになる。

- ① 省エネルギー工場事列表
- ② 省エネルギー基盤高度化設備投資促進税制対象設備一覧
- ③ 省エネルギー技術文献資料（図書・報告書）
- ④ 省エネルギー技術雑誌記事資料
- ⑤ 省エネルギー機器・資材カタログ
- ⑥ 省エネルギー技術データシート集

19.4.2 省エネルギー技術データベースシステムの展開

ホストコンピュータと省エネルギーセンターを接続した第1段階から，次は各地の支部（所）に端末機を設置する第2段階，更には地方を含む会員会社，団体などに端末機を設置する第3段階までの発展の過程が考えられる。構築段階を図19-2に示す。

注1) 内部管理型情報であっても、情報・資料の内容によっては外部提供可能型情報となりえる。
 注2) 情報・資料の分類は、一部相互に重複する内容もある。

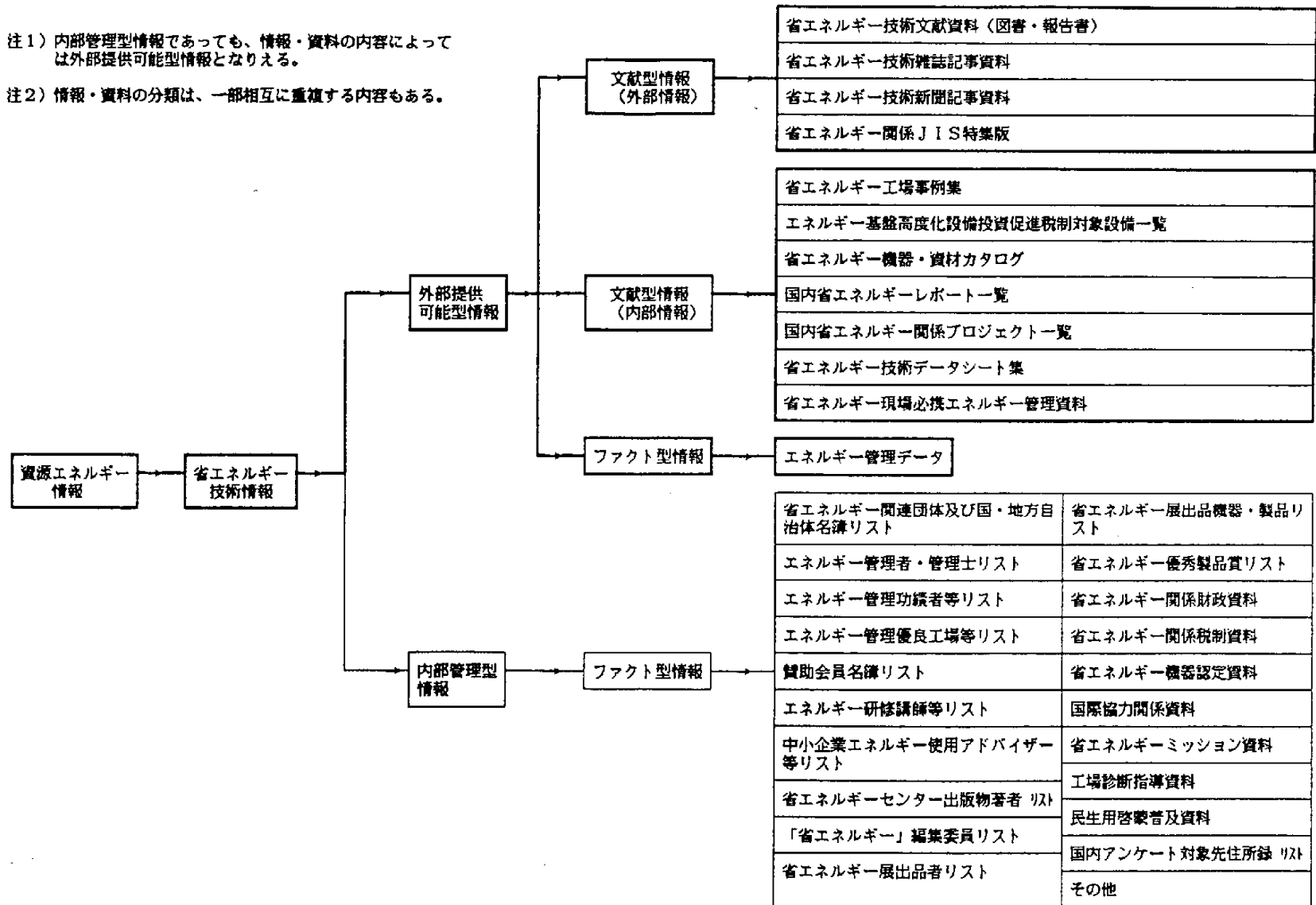


図19-1 省エネルギー技術情報・資料の位置付け

図19-2 省エネルギー技術データベースシステムの構築段階

計画段階	対象情報(審積)	システム利用対象者	システム構成と検索方法				一次資料			システム費用		備考	
			※1コンピュータ	省IT244-センタ	支部(所)※	会員会社	審積方法	依頼方法	送付方法	開発費	運用費		
第1段階	①省エネ工事事例集 ②省エネ基礎設備一覧※ ③省エネ技術文献資料 ④省エネ技術雑誌記事 ⑤省エネ技術データシート集 など	1.省エネルギーセンター職員 2.会員会社	※1コンピュータ	入力用 検索	① ② ③ ④ ⑤ (電話・FAX・文書)		1.原文 2.コピー 3.光ディスクファイル	1.電話 2.FAX 3.文書	1.コピーの郵送 2.FAX 3.電話	必要	1.以外費(コンピュータ、回線費) 2.審積データの入力 3.新規登録情報の入力 4.光ディスクへの入力	省エネルギーセンター本部でのみ使用可 支部(所)との連携を強化	
			レンタル	端末機2台									
			ネットワークシステム	専用回線									
第2段階	第1段階に加えて、 ①国内省エネルギーレポート一覧 ②国内省エネルギー関係プロジェクト一覧 ③エネルギー現場必須管理資料 ④エネルギー管理データ など	1.省エネルギーセンター職員 2.会員会社	※1コンピュータ	入力用 検索用	① ② ③ ④ ⑤ (電話・FAX・文書)	端末機 端末機 端末機 端末機	1.原文 2.コピー 3.光ディスクファイル	1.電話 2.FAX 3.文書	1.コピーの郵送 2.FAX 3.電話 4.光ディスク、FAX自動送付	僅少	1.以外費(コンピュータ、回線費) 2.新規登録情報の入力 3.光ディスクへの入力	会員会社で希望があれば端末機の使用可	
			レンタル	端末機2台	順次端末機設置	順次端末機設置							
			ネットワークシステム	専用回線	公衆回線	公衆回線							
第3段階	実情に応じて対象情報を追加入力していく。	1.省エネルギーセンター職員 2.会員会社	※1コンピュータ	入力用 検索用	① ② ③ ④ ⑤ (電話・FAX・文書)	端末機 端末機 端末機 端末機 端末機 端末機	1.原文 2.コピー 3.光ディスクファイル	1.電話 2.FAX 3.文書	1.コピーの郵送 2.FAX 3.電話 4.光ディスク、FAX自動送付	僅少	1.以外費(コンピュータ、回線費) 2.新規登録情報の入力 3.光ディスクへの入力	会員会社の積極的な利用を期待	
			レンタル	端末機2台	順次端末機設置	順次端末機設置							
			ネットワークシステム	専用回線	公衆回線	公衆回線							

注) ※ 省エネルギー基礎高度化設備投資促進税制対象設備一覧、 ** 省エネルギーセンターの支部(所)

最終的に光ディスクファイル装置を連動したデータベースシステムを想定すると、図19-3に示すようなハード装置となる。

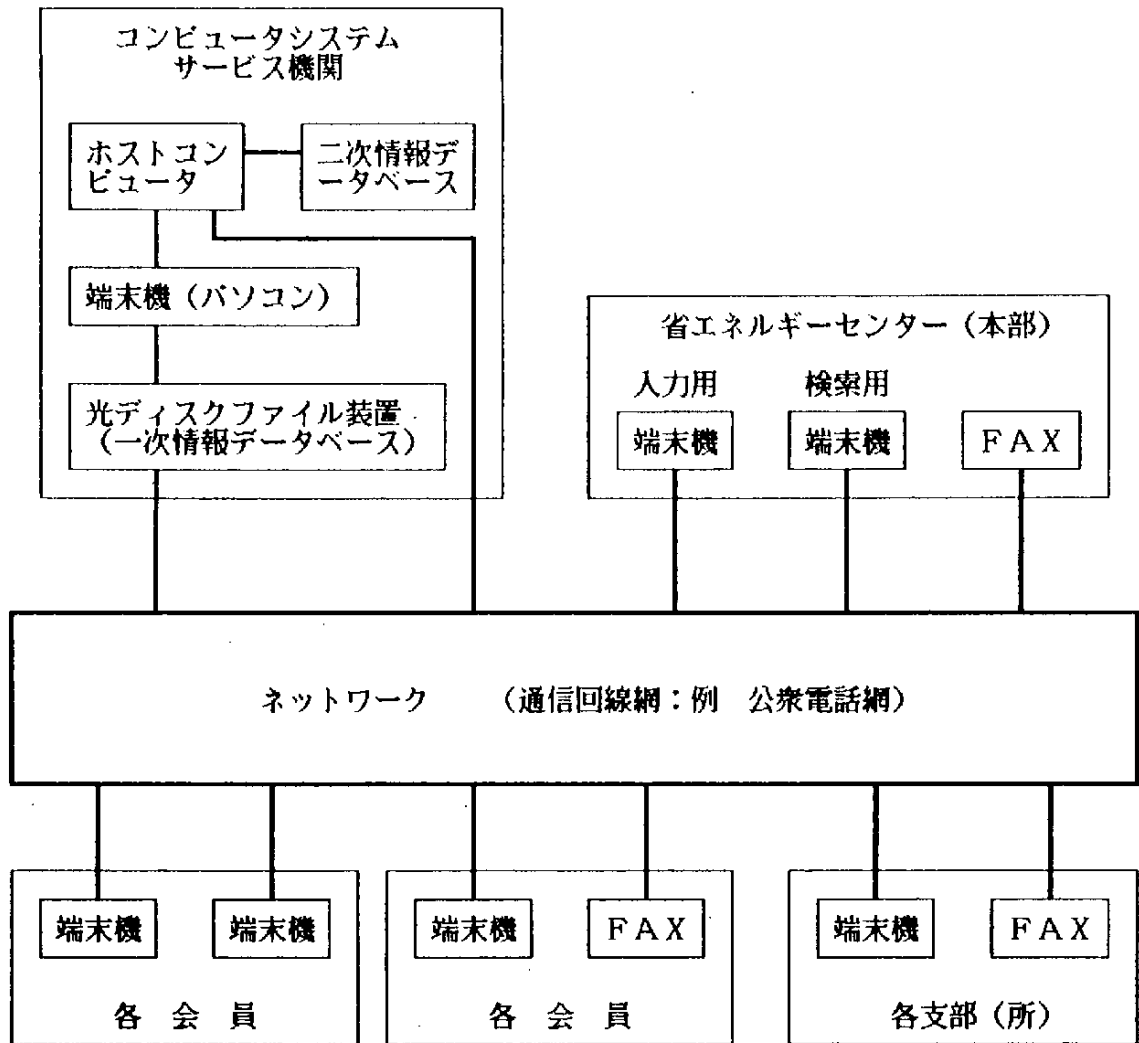


図19-3 省エネルギー技術データベースシステム

19.4.3 情報登録用紙のフォーマットとインデクシング

フォーマットの項目は対象とする情報の種類により異なるが、できるだけ統一した項目とする。省エネルギー事例全集を表19-1に示す。

表19-1 登録用紙のフォーマットの項目の例

①	入力年月日	⑩	事例分類
②	光ディスク内番号	⑪	業種名(統制語)
③	現物保管番号	⑫	業種名(自然語)
④	省エネルギー推進全国大会参加年	⑬	機器・設備・資材分類(予め設定)
⑤	事例タイトル名	⑭	機器・設備・資材分類
⑥	工場名	⑮	製品名(統制語)
⑦	ページ	⑯	製品名(自然語)
⑧	抄録	⑰	キーワード(統制語)
⑨	事例分類(予め設定)	⑱	キーワード(自然語)

注) ここでの統制語の意味は予め決めた用語を使用することを指し、自然語の意味は必要に応じて自由に用語やコードを使用することを指す。

分類については、JICSTの主題別カテゴリー、DOE(アメリカエネルギー省; Department of Energy)のEDBカテゴリー、あるいはエネルギー関連各種基準・制度などの分類表(例、表19-2)を参考にしながら作成しなければならない。

分類に当たっては大分類、中分類にとどめ、小分類は設けないようにする。

表19-2 工場に係わる事業者のエネルギー使用合理化判断基準の概要

①	燃料の燃焼の合理化	⑤	熱の動力等への変換の合理化
②	加熱及び冷却並びに伝熱の合理化	⑥	抵抗等による電気の損失の防止
③	放射・伝導等による熱の損失の防止	⑦	電気の動力、熱等への変換の合理化
④	廃熱の回収利用		

索引語の作成にはJICSTシソーラスを参考にすることは勿論であるが、シソーラス体系をとるか、キーワード集の形式とするかはDBMSとも関連してくるので十分な検討が必要である。検索の効率と作成の経済性を考えると、検索に用いる索引語は予め選択して定めた統制語で構成されたキーワード集と、必要に応じて自然語を使用する併用方式がよい。

19.5 省エネルギー技術データベースシステムの運用体制

適切な運用を行うために、情報専門員をおいた省エネルギー技術情報センター（仮称）を設置する必要があり、省エネルギーセンター、その支部(所)、およびユーザ(会員)に分けて主な業務と位置付けを図19-4に示す。

	省エネルギーセンター	支部(所)	ユーザー(会員)
情報処理のフロー			
情報収集 登録受付	情報、資料の収集	<ul style="list-style-type: none"> 情報、資料の収集 本部、会員への連絡 	<ul style="list-style-type: none"> 情報提供 論文、報告書提供
分析、加工 ・整理	<ul style="list-style-type: none"> 対象資料ごとに登録用紙に記入 登録用紙によりオンライン入力（二次情報） 原資料の整理 		
受入資料 案内	<ul style="list-style-type: none"> 既刊の刊行物に掲載 新たに受入資料案内を作成 	<ul style="list-style-type: none"> 本部を支援 	<ul style="list-style-type: none"> 刊行物、受入資料案内から情報を受け取る 一部は端末機から
蓄積・保管	<ul style="list-style-type: none"> 二次情報はコンピュータに入力し、データベース化 原資料（一次情報）をそのまま保管、または光ディスクに入力 	<ul style="list-style-type: none"> 本部を支援 	
検索・利用	<ul style="list-style-type: none"> ユーザーの質問についてのオンライン検索 原資料をコピーしてユーザーへ送付、または光ディスクからコピーしてユーザーへ送付 センター業務のためにオンライン検索 	<ul style="list-style-type: none"> 情報、資料の提供 本部、会員との連絡 	<ul style="list-style-type: none"> センターまたは支部(所)に質問、または欲しい情報を要求 一部は端末機から

図19-4 情報処理フローにおける業務分担

19.6 省エネルギー技術データベース構築の予想効果

省エネルギー技術データベース構築による効果は大きく分ければ次の3項目となる。

① 省エネルギー技術の普及促進

先進的な省エネルギー技術から身近な省エネルギー機器に至る幅広い技術情報の効果的な収集、分析、整理・加工、提供手法が確立され、それぞれのニーズに応じた情報が迅速、的確に処理されることにより、「エネルギーの使用の合理化」、「省エネルギー関連の技術開発」および「広報活動」など、省エネルギー技術の普及促進が一層効果的に行える。

② 発展途上国の省エネルギー推進

省エネルギー技術に関しては、発展途上国からの協力要請が多数寄せられている。

データベース化により、各国の事情に合わせた省エネルギー技術情報の提供が容易となり、発展途上国の省エネルギーの推進に寄与する。

③ 先進諸国との省エネルギー技術交流

データベースを構築し、将来、IEAなど国際的な組織と技術情報の交換ができれば、先進諸国との迅速な省エネルギー技術交流が可能となる。すなわち、海外の先進的な省エネルギー技術の情報を国内に普及するとともに、国内の優れた技術の情報を海外に提供できることになる。

データベースの活用については、省エネルギーセンターが行っている広報活動、調査研究、エネルギー管理者の養成、技術講習会、中小企業向けエネルギー使用診断指導、発展途上国に対する省エネルギー技術の普及促進などの事業の中で効果的に利用することが可能になる。データベース構築後の利用対象者は次のように広範囲にわたることが予想され、利用手段は段階的に発展すると思われる。

- ① 利用対象者： 関係省庁，業界団体，調査研究機関，各国政府機関，地方自治体，図書館，民間企業，国際機関
- ② 利用目的： 省エネルギー技術の一層の普及のため，ニーズに応じた省エネルギー技術情報の入手と活用。
- ③ 利用手段： 次のように展開すると予想される。

第1段階 省エネルギーセンターとホストコンピュータ間のオンライン検索

第2段階 省エネルギーセンターと各支部(所)間のオンライン検索

第3段階 省エネルギーセンターと各支部(所)，会員間のオンライン検索

省エネルギーセンターの賛助会員数は熱および電気のエネルギー管理指定工場はもとより、自治体・公共企業体、銀行・信託業などを含む3,300社に達するので、オンラインによるデータベ

ースの利用頻度は飛躍的に伸びると考えられる。

19.7 今後の課題

今後、実際にデータベースを構築・運用し、データベースシステムを維持・発展して行く場合に常に念頭において置くべき3点を提言する。

19.7.1 情報の質の向上

「省エネルギー技術データベース」構築における抄録の作成やインデクシングなどに対する価値が、国民的な立場から認められるような社会的仕組みが出来上がるように努力することが重要であり、そのことが、情報の質の向上につながる。情報サービスが社会一般に広く利用されるようになった時、情報は社会の財産となるのである。

19.7.2 国際化

科学技術情報の国際化はことさら新規なテーマでなく、あらゆる場所で実施されている。省エネルギー技術の各テーマは非常に国際的なテーマであるから、今後、海外のエネルギー関連機関との国際的交流がますます盛んになるであろう。機会あるごとに外国語によるデータベース構築の必要性を訴えていくことが、情報管理部門の責務である。

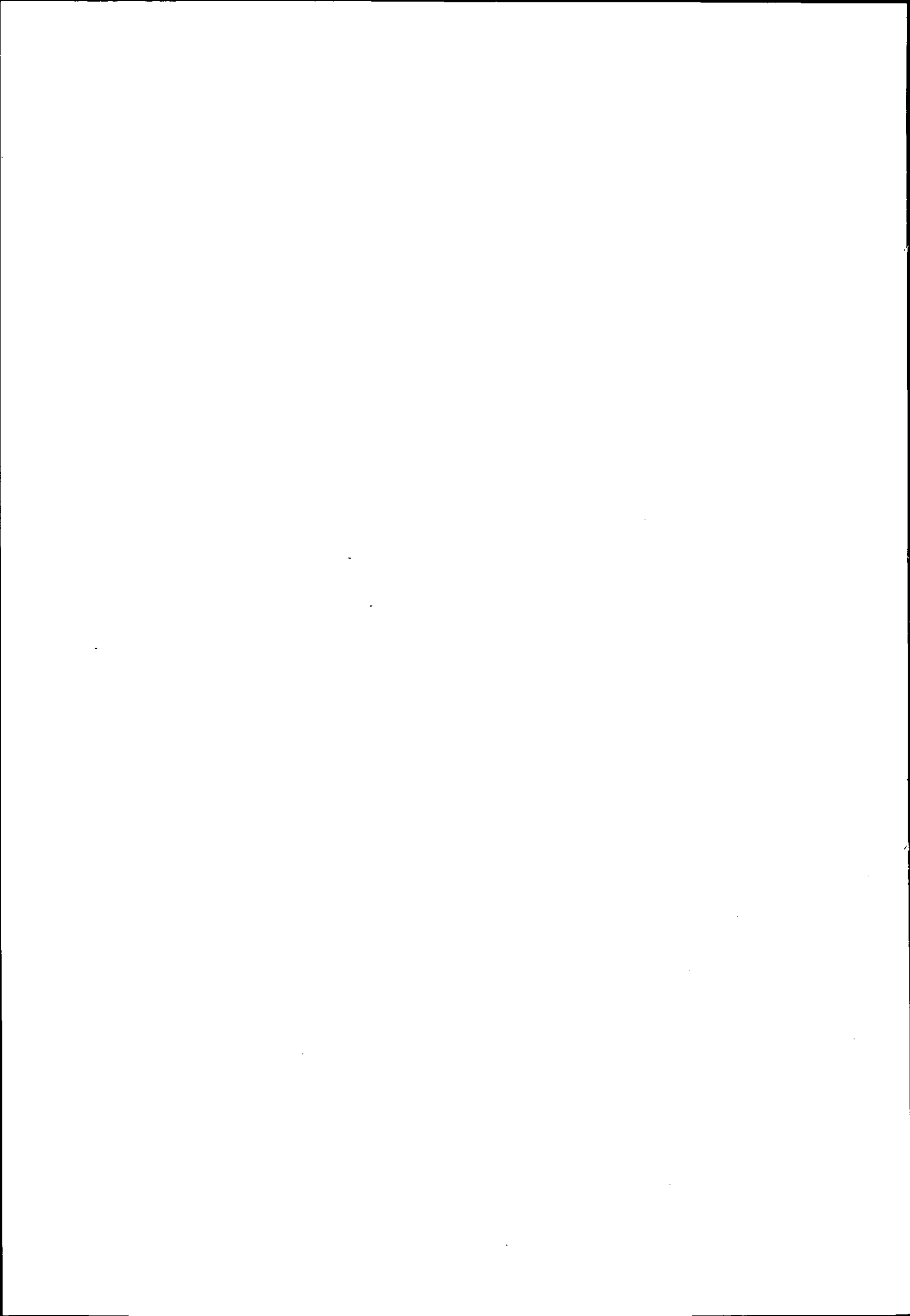
19.7.3 情報関連技術の進歩

省エネルギーデータベースを構築した後、これを維持・発展して行くためには、情報処理関連技術の進歩に絶えず注目する必要がある。省エネルギー技術データベースシステムについて日頃から考慮すべき点をあげておき、時期を決めて検討し、積極的に採用して行くことが望ましい。



20 電子部品データベース実験システムの構築

三菱商事株式会社



20. 電子部品データベース実験システムの構築

20.1 目的

半導体素子を中心とする電子部品は技術的に高度であるばかりでなく多様なユーザーニーズに対応して、多くのメーカーが多種多様の品種をもつため、関連する製品・技術情報が次々と生み出されている。

これらの電子部品のユーザーは必要とする最新データを入手することに多くの時間を費やしており、より効率の良いデータ提供システムの出現を望んでいる。

更に半導体・電子部品メーカーにとっても相次いで投入される新製品の情報などを迅速且つ効率良く多くのユーザーに伝達するために、新しいメディアの利用が重要になってきていると思われる。

これらの事情に鑑み、NTT並びに大手半導体メーカーとともに「電子部品データベース研究会」を発足させ、わが国で流通している電子部品の製品・技術情報をデータベース化し、公衆通信網を介して広くユーザーに提供するオンライン・データベースの実現を企画・研究してきた。

しかしながらこのデータベース化は相当の規模になることが予想され、いかに整理して提供していくか、ユーザーのニーズはどのようなものか、最適な提供システムは何かなど、様々な角度からの調査・検討が更に必要である。

このような調査・検討を行うには小規模の実験データベースを構築し、様々な角度から評価することが最適の方法と判断された。

20.2 実験システム構築体制

実験システムの構築とその評価のため、下記3グループを組織し、三菱商事が全体の取りまとめを行った。

(1) システムグループ

メンバー 担当業務	<ul style="list-style-type: none">• NTT（データ通信事業本部） 三菱商事• 実験データベースにおけるホストシステムの設計およびデータ入出力の設計ユーザ端末のマンマシンインタフェースの開発およびメンテナンスユーザ指導などの技術サポート
--------------	--

(2) 情報提供・編集グループ

メンバー 担当業務	<ul style="list-style-type: none">• 日本電気 日立製作所 東芝 富士通 松下電子工業 三菱電機沖電気 日本T I 日本モトローラインテル・ジャパンNTT（データ通信事業本部） 三菱商事• 実験データベースで使用される情報の提供、収集、編集検索語付けおよび入力作業を行った。
--------------	---

(3) 評価グループ

メンバー 担当業務	<ul style="list-style-type: none">• 電子機器メーカー他 20社NTT（データ通信事業本部） 三菱商事• 実験データベースに対するユーザとしての意見の提出およびその収集・分析を行った。
--------------	--

20.3 実験システムの概要

20.3.1 提供情報の範囲

電子部品	半導体素子	メモリ	マイクロコントローラ	MPU
		MPU周辺IC	リニアIC	
		ボルテージレギュレータ	トランジスタ	
	受動部品	アルミ電解コンデンサ		

20.3.2 提供情報の内容

カタログ情報 半導体メーカーおよびコンデンサメーカーのデータブックより製品毎に各々の特性情報を抽出してデータベース化した(抄録)。
ユーザは多様な属性にもとづき自由に検索できる。

ニュースサービス カタログ情報に加え実験的に下記のニュースを提供した。

新製品ニュース……………メーカー各社の新製品速報

広報室……………メーカーからの様々なお知らせ・案内

海外ニュース・イベント…海外の半導体・電子産業関連ニュース・国内のイベント・
学会予定の案内

ハイテク新製品ニュース…産業情報マガジン「Infordia」から関連記事を抄録

電子メール ユーザ・メーカー各社間のための双方向通信

20.3.3 提供情報数

カタログ情報	製 品	品種数 (点)
	メモリ	435
	マイクロコントローラ	50
	MPU	100
	MPU周辺LSI	40
	リニアIC	100
	ボルテージレギュレータ	60
	トランジスタ	800
	アルミ電解コンデンサ	1500
	計	3085点
新製品ニュース	半導体メーカーの新製品情報を新聞発表と同時に掲載 ニュース点数 116点	

20.3.4 対象ユーザ

エレクトロニクス産業に従事する回路設計技術者・部品管理・品質管理担当者・部品購買担当者

対象企業数 約12,000(事業所ベース)

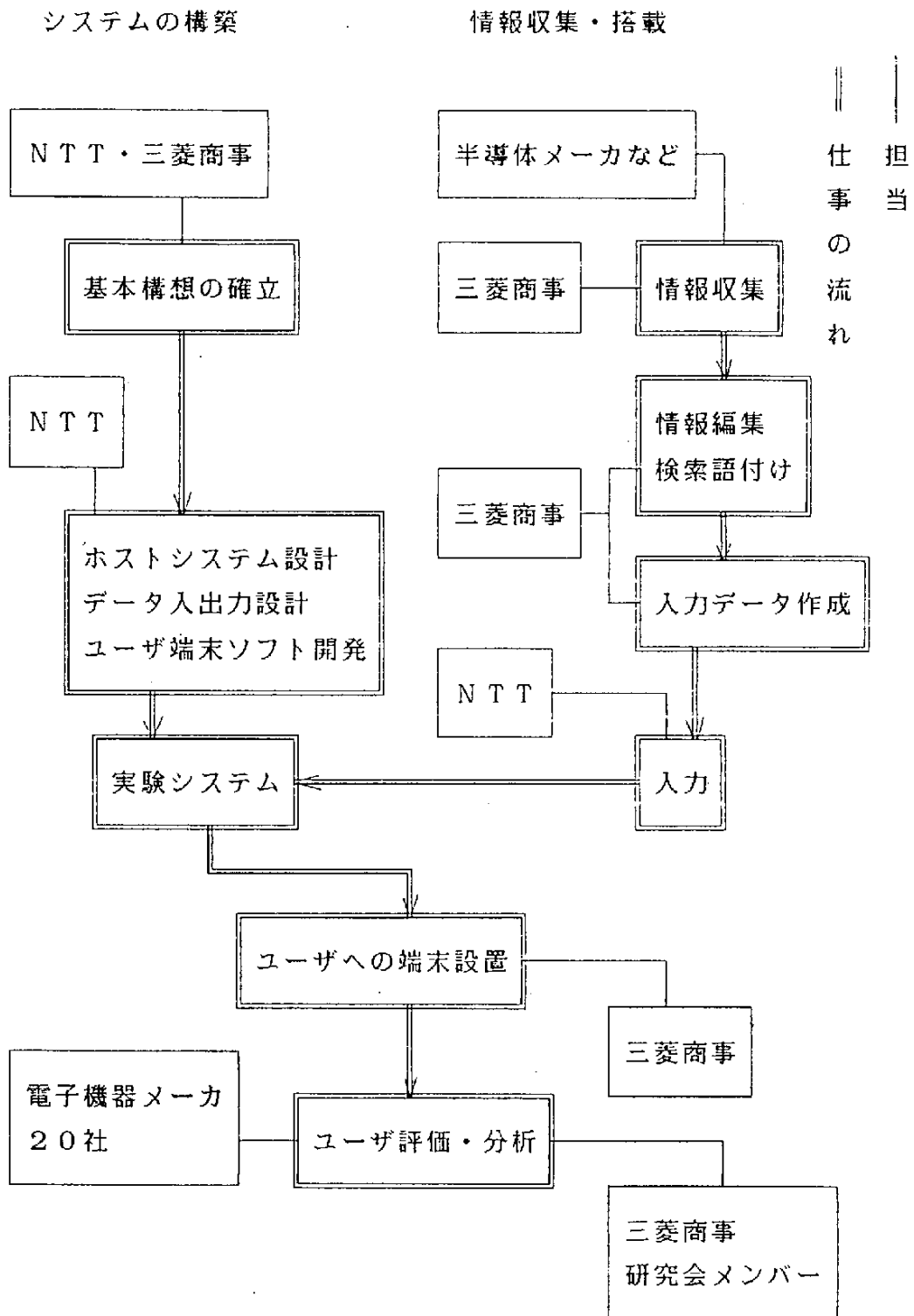
利用の目的

多数の部品の中から横断的に最適な部品を選択する。

新製品情報の速やかな入手

その他半導体・電子産業関連ニュースの速やかな入手

20.4 実験システム構築・評価フロー



20.5 実験システムの評価

20.5.1 評価のポイント

- (a) システムの使いがって
システムの立ち上げ，検索手順など利用者の立場から使いがってを評価
- (b) 属性の種類と検索項目
製品毎の属性の種類と検索項目数の過不足につき調査
- (c) カタログデータベース（抄録）の情報価値の評価
抄録として提供するカタログデータの価値について評価
- (d) ニュースサービスの種類と内容
副次的サービスである各種ニュースの利用価値の評価

20.5.2 評価の結果（ユーザ 20 社の意見の集計および分析）

(1) 評価に参加したユーザ 20 社

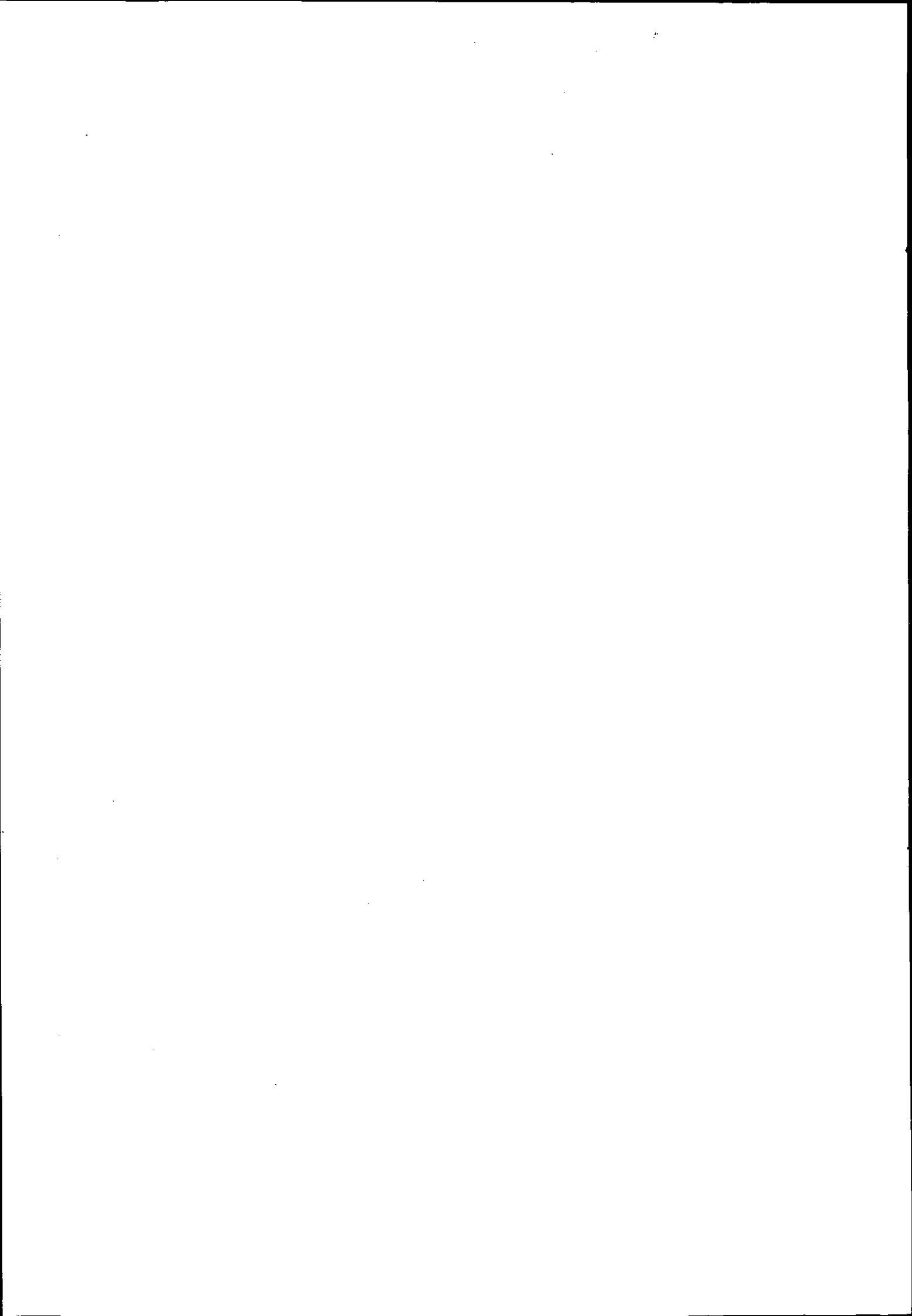
赤井電機	天昇電気工業
アンリツ	日通工
NTT（LSI研究所）	日本光学工業
オリンパス光学工業	日本ビクター
キャノン	日立製作所
シーティーシー	三菱重工業
シントム	三菱電機
セイコーエプソン	ミツミ電機
セイコー電子工業	ユニデン
セガ・エンタープライゼス	ワコム

(2) 評価結果のまとめ

- | | | |
|------------------|---|-----|
| (a) システムの使いがって | ・検索テーブルを使って新規の部品を選択するシステムはすべてのユーザに高い評価を受けた。 | |
| | ・アクセス時間の短縮を希望するユーザ | 60% |
| | ・検索手順の簡素化を希望するユーザ | 50% |
| (b) 特性情報の種類と検索項目 | ・満足であるとするユーザ | 60% |
| | ・種類の増加を希望するユーザ | 35% |
| | ・その他 | 5% |
| (c) 情報価値の評価 | ・情報内容に満足とするユーザ | 60% |
| | ・更に充実を希望するユーザ | 40% |
| (d) 副次的サービス | ・実験システムのメニュー例に満足するユーザ | 65% |
| | ・更にメニューの増加を希望するユーザ | 35% |

20.5.3 今後の課題

- ・NTTデータ通信事業本部および電子部品データベース研究会メンバーほか関係各グループの熱心な協力のもと、実験システムの構築については十分な成果が得られた。
- ・しかしながら今後商用化を行うためには評価結果を十分に検討し、システムのバージョンアップおよび提供情報の特に質的向上に一層の研究が必要と判断される。



21 漢字データベース構築

株式会社 理想社印刷所



21. 漢字データベース構築

21.1 概 説

漢字データベース構築に関しては、昭和61年度データベースの構築促進及び技術開発促進事業として委託を受けて実施された事業（「漢字データベース構築」（61-開-17）参照）において、その基本方針が定められており、当期事業はその成果に基づいて継続的に実施したものである。

21.2 作業方針について

漢字データベースの構築には、大きく分類して次の作業を必要とする。

- ① 漢字データベースに収録すべき字種の決定
- ② 選定された字種の文字フォント作成
- ③ 選定された字種に関する意味・読み方・熟語・その他の情報収録
- ④ 収録された関連情報の検収
- ⑤ 漢字データベース独自の検索システムの開発

上記の、それぞれの作業の関連性に留意しながら、次のとおり当期作業を実施する方針とした。

- ① 漢字データベース構築作業として
 - (1) 第一次収録対象文字種 21,000 文字の選定およびデジタル化作業
 - (2) 上記文字種のうち、JISを中心とした6,000字に関連する意味・熟語その他の内容入力
 - (3) 入力内容についての校閲作業
- ② 検索システム開発について
 - (1) システム仕様と開発方針の検討
 - (2) 模擬システムの開発

21.3 データ整備作業

21.3.1 収録文字種の選定

当期間内において収録を決定した字種ならびに収録を保留した字種の概要は次表のとおりである。

文字の分類		収録決定した数	収録を保留した数
①	JIS 78 制定字種に含まれる漢字	6, 349	—
②	JIS 83 で新たに採用された字形	230	—
③	①②以外でCTS 利用頻度の高い字形	2, 268	約 4, 200
④	その他	4, 624	約 3, 300
合計		13, 471	約 7, 500

21.3.2 文字フォントのデジタル化

漢字データベースに収録すべき文字のイメージをデジタル化し、100 × 200 ドット（左右 100, 天地 200 ドット）の高品質出力用のものと、その他の一般的な利用に供することを目的とする 48 × 48 ドットのフォントを作成収録した。

なお、100 × 200 の高密度文字から、プログラム処理によって自動的に密度変更を行う機能について別途開発を行い、中間的ないくつかの密度（例：64 × 64 など）の文字フォントを自動的に作成することを可能とした。

デジタル化の作業手順は以下のとおりである。

- (1) 文字デザイン作業（デザインが既存の文字については不要）
- (2) 文字デジタル化装置により 100 × 200 の文字フォント作成
- (3) ディスプレイ画面および出力テストによる文字精度確認 → 一部文字については修正
- (4) フォント密度変更システムにより 48 × 48 文字フォントへの加工を行う
- (5) ディスプレイ画面上で精度確認 → 一部文字については修正
- (6) 検字番号設定
- (7) 登録完了

21.3.3 その他関連する各種情報の入力

漢字データベースの最も重要な要素である「文字に関連する情報」の入力作業について以下に概要を述べる。

- (1) 入力原票の作成
 - ① 原票記入の負荷を軽減するために、資料として採用した辞典類に、入力が必要な部分をマークする。
 - ② 当該辞典類に不足する情報について、記入できる範囲の情報を記入する。

- ③ 追加修正情報の入力容易な入力機器を選定する。
- ④ ある程度、入力オペレータの判断に依存できる部分は原票整理を省略する。
- ⑤ 入力結果のモニタ・リスト上に、不足情報を追加記入する。
- ⑥ 文字固有のコードなどをキーとして別に入力された情報をマッチング処理で併合できるものは機械的処理でカバーする。

以上の方針に基づいて検討の結果、入力原票作成を効率良く行うための方法として以下の3種のファイルを作成することとし、管理番号をキーとして併合する方針とした。

① 第1ファイル(辞典類を原票として入力するファイル)

(管理番号、熟語例、略字の管理番号、正字の管理番号、同字の管理番号)

② 第2ファイル

(管理番号、当該文字に対応するJISコード、属性、代表総画数、その他画数、代表部首分類、その他部首分類、代表音、その他音、字訓、字義など)

③ 第3ファイル

(管理番号、文字イメージ、簡体文字管理番号)

以上のほか、検字番号を決定することが必要であるが、検字番号は極めて重要であり、入力および校閲作業を進めた上で決定することが必要となったため、当面の文字管理は暫定的に定めた「管理番号」によって行う方針とし、別途「管理番号と検字番号対応表」を作成してセッティングする方針とした。

(2) 入力作業

前述の入力原票作成方針に基づいて入力を行う場合、原票そのものが必ずしも完全ではなく、逐次入力内容の改善を図る必要性が高いため、極力追加修正の容易な入力システムを選定する必要がある。

また、ピーク時には大量の入力能力が必要となることが予測されるために、入力可能な機械の固体数が多い機種であることが望ましい。

以上の要件を満たす機種の選定を行った結果、OASYS 100シリーズの日本語ワードプロセッサが、最も適正が高いと判断し、入力機種として採用した。

入力に際しては、必ずしも原票どおりに入力するだけでなく、部首分類表などの資料を参照しながらの入力や、その他辞典類を参照しながらの、判断を必要とする難度の高い入力作業となった。

当期の入力実績は次表のとおりである。

入力した文字種数	6, 579
入力熟語数	約 155, 210
入力文字の総字数	約 2, 873, 000

21.4 校閲作業

入力された情報の校閲には、原則として次のような作業が必要であり、現在、継続的に作業をすすめる、情報ファイルの精度を高める作業を行っている。

- ① 入力原票として利用した辞書類との照合による入力ミスのチェック
- ② 入力原票に不足する情報の追加記入
- ③ 追加記入にしたがって更新された入力結果の確認

21.5 検索システム研究

漢字データベースに関連する各種ソフトウェアのうち、データベース構築ならびに検索機能上の問題点を把握・解決することを目的とする次の活動を行った。

- (1) 検索機能のシミュレーションと性能予測
- (2) プログラミング言語の調査
- (3) システム・ソフトの調査
- (4) 大容量記憶装置，ディスプレイ装置等の調査・テスト
- (5) 既存の汎用パッケージ利用の可能性

21.5.1 検索機能の検証

漢字データベースを検索利用するための、画面編集，検索方式，操作性などを検証することを目的とするために、パソコンによる「検索模擬システム」の開発を行った。

- (1) 模擬システムの概要

使用機種： FMR-60

文字範囲： JIS 全文字種

OS 名： MS-DOS

主要機能： データ検索，画面編集・表示

使用言語： C言語

開発画面： 62年度報告書に定義されたものから10画面

(2) 模擬システムの概略図

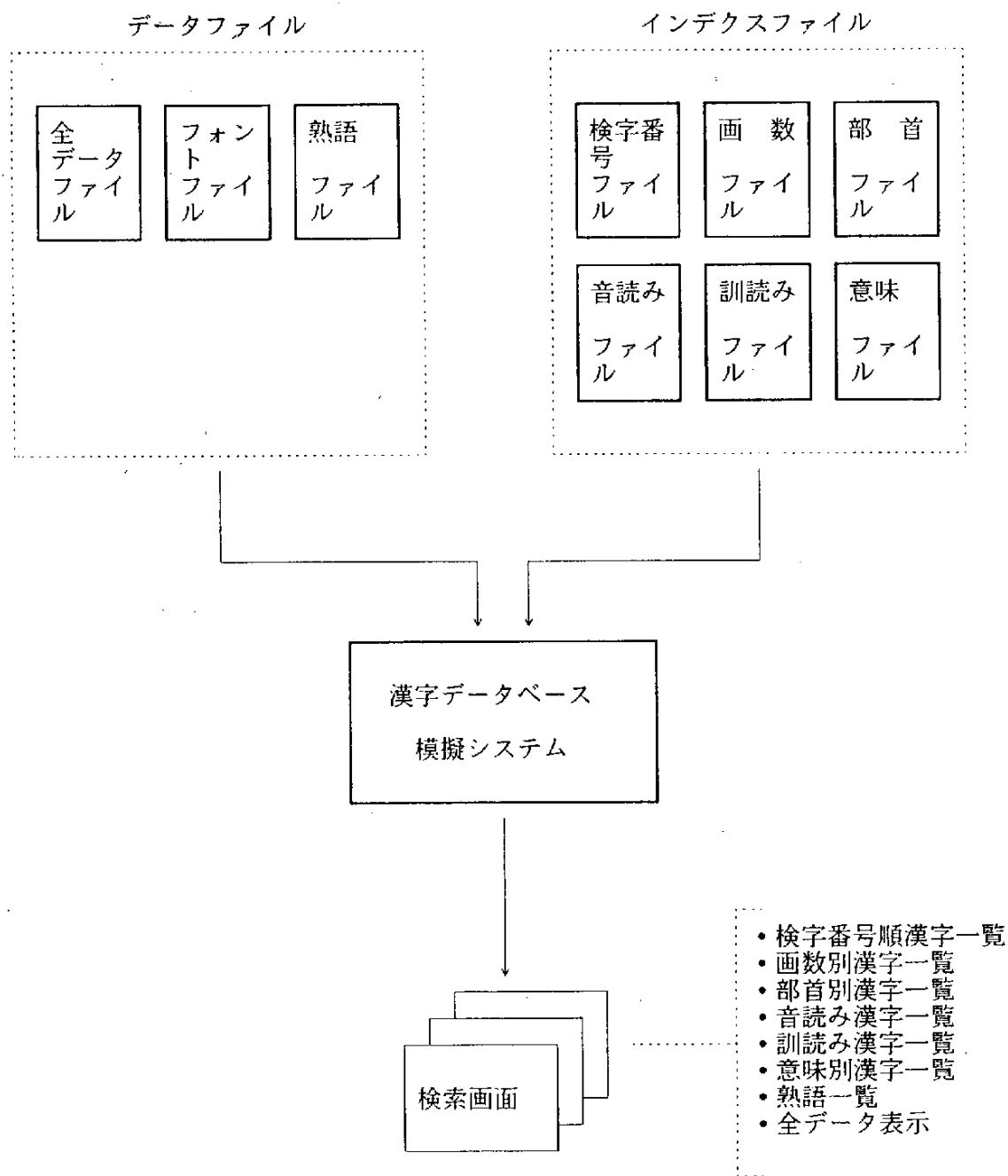


図21-1 模擬システムの概略図

21.5.2 模擬システムの評価

模擬システム開発の各段階ごとにプロジェクト内でのテストと評価を行い、各機能についての確認を行ってきたが、開発予定機能のほぼ完成した時点でのプロジェクト内での評価は次のとおりである。

当初、模擬システムをパソコン上で開発することで、最も懸念したのは実用的な検索速度の可能性についてであったが、実際に検索テストを行った結果、殆どの検索画面が約1秒以内で表示されるなど、予想以上の検索速度を得ることができた。

模擬システムでは、収録した情報量が実用レベルのものと比較して著しく少ないため、検索速度については必ずしも適切な評価ができない一面はあるが、漢字データベースの情報量を、20メガバイト程度のハードディスクに収録できる程度に内容を限定した一般利用者向けのサービスについては、ほぼパソコン機能だけで十分な性能が確保できるものと考えられる。

また、更に情報量を圧縮した限定利用（例えば小・中学校の学年単位の教育用利用）の場合には、フロッピーディスクだけで利用することも可能であり、模擬システムそのものを少し整備するだけで、ある程度の実用化のメドが立ったとの評価ができる。

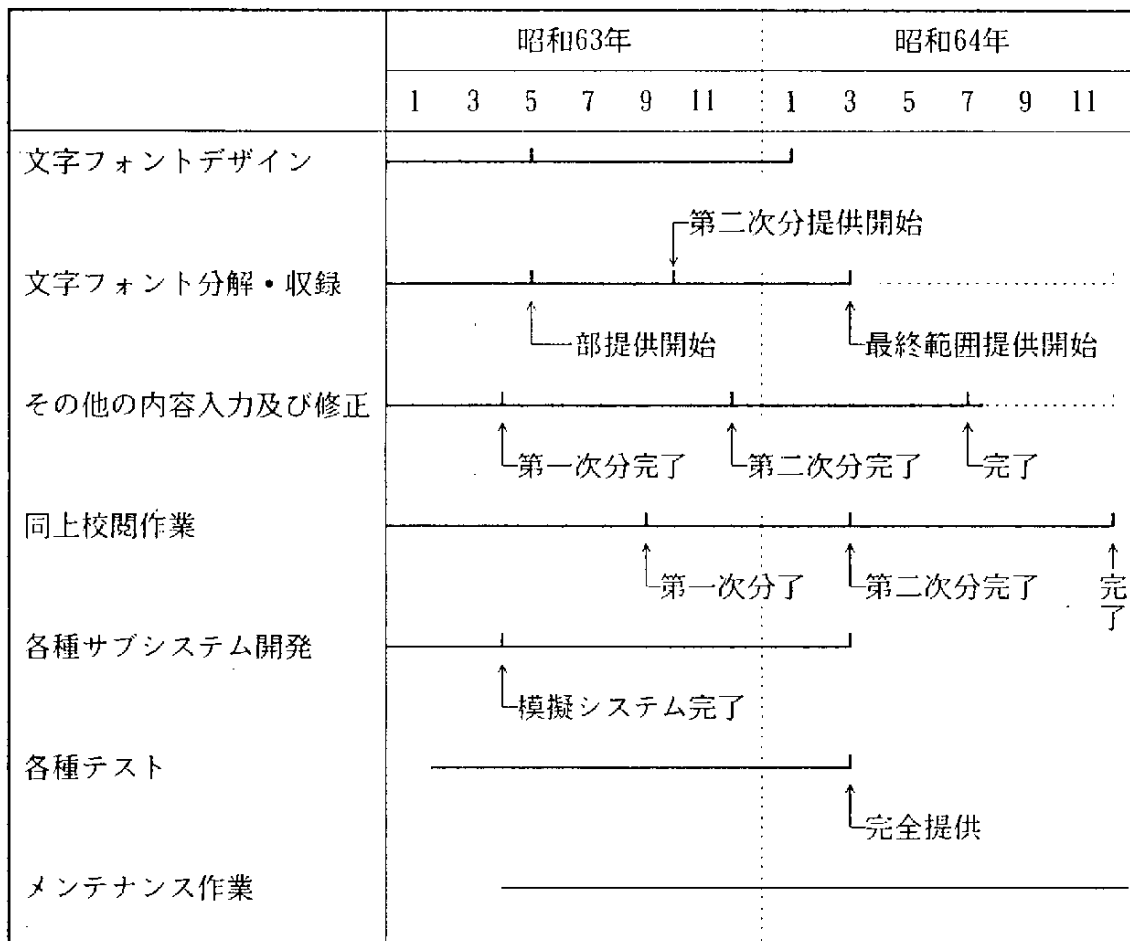
今後のデータ整備の進捗に伴って、CD-ROM利用の可能性、オンライン型検索機能の調査・研究などを行い、前述のような段階的なサービス提供を実現するための初期開発の目的は達成できたものと思われる。

なお、各種のテストの段階におけるプロジェクト内の意見として、実際の利用を想定した検索方法や画面構成に、まだ、多小改良の余地があるとの見解があり、今後の予定としては、この模擬検索システムを一部利用者に提供し、テスト的な利用を通じて得られた各種の意見を分析した上で、一般の利用者にとって漢字データベースのもつ情報内容を、どのように検索できるようにすることが最も魅力があり、有意義であるかについて調査を行い、より実用性の高いシステムとして育成したいと考えている。

21.6 今後の開発計画

当データベース構築の作業計画概要は次表のとおりである。

今後の作業計画概要



※ データメンテナンスは継続的に必要である。

漢字データベースの開発計画は上記のとおりであるが、その他下記の実用化計画を実施する。

- (1) 文字フォントの実用化は他のサブシステムとは別に、昭和63年4月に権利登録を行い、63年10月に実用登録を行う。
- (2) 文字フォントの実用化に当たっては、各種コード体系に対応する情報を完備し、汎用性を持たせて多くのOA機器に活用できる機能を完備させる。
- (3) 電子出版等の新しい利用形態に積極的に対応させるため、CD-ROM等のニューメディアへの適用機能を開発する。

なお、利用者への提供サービス開始は下記のとおり予定している。

(1) 文字フォントの提供

- | | | |
|--------------|------------|-----------|
| ① 第一次構築範囲のうち | 約 13,000 字 | 昭和63年 5 月 |
| ② 第一次構築範囲 | 約 21,000 字 | 昭和63年10月 |
| ③ 最終収録文字範囲 | 約 30,000 字 | 昭和64年 3 月 |

(2) O A 機器用熟語ファイルの提供

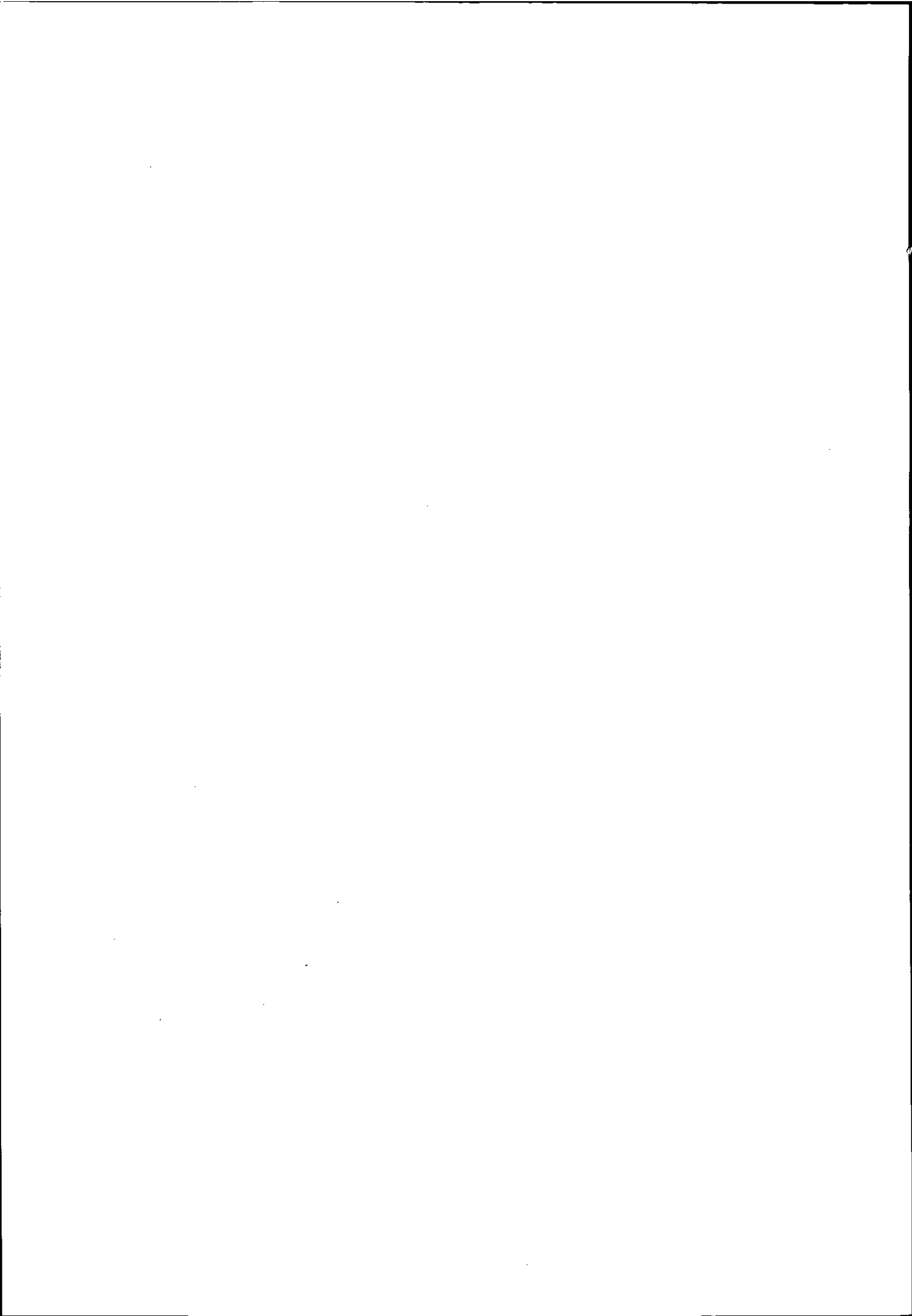
- | | | |
|--------------|------------|-----------|
| ① 第一次構築範囲のうち | 約 13,000 字 | 昭和63年 9 月 |
| ② 第一次構築範囲 | 約 21,000 字 | 昭和64年 3 月 |
| ③ 最終収録文字範囲 | 約 30,000 字 | 昭和64年12月 |

(3) データベース構築および検索システムの提供

- | | | |
|--------------|--|-----------|
| ① 一部機能の提供 | | 昭和63年 4 月 |
| ② 予定機能すべての提供 | | 昭和64年 3 月 |

22 機械翻訳システム用辞書データベースの調査

カテナ株式会社
株式会社 サイコム



22. 機械翻訳システム用辞書データベースの調査

22.1 調査の目的

わが国を取り巻く環境の複雑化にともない、国際的情報流通への対応の高度化を更に図る必要が迫っている。わが国が言語の障壁を乗り越えて、情報のコミュニケーションを展開して行くためには、機械翻訳技術の利用が大きく望まれるところである。そして、そのようなニーズの高まりに対応して、いくつかの商用機械翻訳システムも発売されてきている。しかしながら、一部の例外を除いて、一般の利用者の多くがこれを直ちに導入し、本格的に運用させる状況には未だ至っていない。この最大の理由は、機械翻訳用辞書作成・利用技術の未整備に求められる。また、そのために必要な機械翻訳システム用辞書と、そこに係わるデータベース技術に関する実体に即した研究も、未だ十分になされているとはいえない。そこで本技術研究では、機械翻訳システムのための翻訳辞書に係わる諸々の課題を検討し、機械翻訳システム用辞書データベースに関する作成・利用技術の確立を図ることを目的として調査した。

機械翻訳システム用辞書の中でも、システム辞書、文法辞書、基本語辞書といわれるものは、自然言語処理の中枢に係わるものであり、システムの開発者側のみならず、諸々のレベルでの利用者にとっての便益という面からも、それらの辞書に関する技術の成熟が強く望まれる処である。しかし、そうしたシステムの内部に係わるような辞書に関する諸問題については、十分な分析がなされているとはいえないし、利用者あるいは開発者相互がそれぞれの技術を確立するに必要な情報の交流が、十分になされているともいえない。

更に、機械翻訳システムの効果的な利用をめざす立場から、専門分野辞書、ユーザ辞書の存在も欠くべからざるものである。

機械翻訳システム用の各辞書に関する課題と問題点を明らかにすべく調査を行った。

22.2 調査の内容

22.2.1 機械翻訳システムにおける辞書の分類

利用者の側からみたとき、機械翻訳システムがユーザに提供する辞書構成の枠組みは、以下である。

(1) システム標準辞書

機械翻訳システムに標準装備されて提供される辞書データをシステム標準辞書と呼ぶことにする。ここには、翻訳処理に必要な文法的内容や、基本語彙としての一般語が含まれる。

利用者にとって、システム標準辞書は、翻訳システムと一体になっているとみることでもできるが、一方で、利用者がその内容に深く係わる場合もある。また、翻訳システムの基本エンジンや処理アルゴリズムに比べて、対象分野や、利用形態などに応じて変更する可能性が高く、基本システム完成後も、たびたび変更される可能性があるので、一部の実験システムを除けば、プログラムとは独立の要素として分離してある。

システム標準辞書の規模を見ると、多くのシステムでは、特にその多くを占める基本語について、5万語程度の語彙数が収納されている。ただし、専門語については、ここでは別に扱う。

商用の機械翻訳システムにおけるシステム標準辞書の内容について見てみると、特許・科学文献の翻訳を目的とするなど一定の限定を行っている場合も見られるが、おおむねはある程度広い適用分野を想定したものになっており、一般的な用法を中心として収納される。利用目的に応じた精密化は、専門語辞書やユーザ辞書を利用することが想定されている。

他方、機械翻訳をある程度高度に行うためには、翻訳処理系はかなり複雑な情報を要求するが、こういった情報を利用者側に要求するのには無理があるので、あらかじめシステム標準辞書に準備しておく必要がある。そこでシステム標準辞書には、構文・意味情報など、翻訳アルゴリズムの中核に係わる情報が、豊富にしかも注意深く格納されている。

システム標準辞書では、このように、翻訳アルゴリズムに起因する精密な情報の要求と、利用者の側からの特殊性と理解容易性の要求を調整しつつ、大量の言語データを扱わなければならないという課題を背負っている。

(2) ユーザ辞書

現在の機械翻訳システムでは、ユーザが辞書を変更して、自らの目的に応じて翻訳結果を改良するための機能が準備されている。そこでは、翻訳システムに関する詳細な知識を持たないユーザにもこうした機能が利用できるという要件が考慮されている。その反面、機械翻訳システムにユーザが介入できる範囲（現状の機械翻訳の技術水準ではこの範囲が相当に広い）が、かなり限定される結果になる。単語の辞書については、商用

の機械翻訳システムにおける辞書変更内容が見られるが、そこで見ることの出来る項目としては、品詞、訳語、意味大分類などに限られる場合が多い。ここには、これらの項目だけでは、利用者が必要とする翻訳精度の向上が十分に達成できないということと、逆に、意味分類をはじめとする情報を、翻訳システムの内部についての知識を持たないユーザが処理することが困難であるという、相反する問題が生まれる。

従って、システム標準辞書との係わりで特に問題になるのは、利用者による翻訳機能の改良が、現在の機械翻訳システムの仕組みだけでは十分ではないということである。実際、機械翻訳システムの本格的な実稼働を実現するためのいくつかのプロジェクトでは、機械翻訳システムに関する知識を持たない利用者だけでは実稼働は困難であるとして、ユーザ側と開発が一体となった、辞書や文法の大規模な改良作業が行われているのが実状である。

(3) 専門辞書

各分野で用いられる専門語は、一般語に比べはるかに多くの語彙があり（数百万語ともいわれる）、しかも日々に新語が生まれてくるという状況がある。分野毎での語彙の分布や用法にも、顕著なばらつきがある。そのために、機械翻訳システムでは、専門分野に依存した専門辞書を、システム標準辞書とは別にユーザに提供したり、ユーザがそのような辞書を利用するのを保証するといった考慮を払っている。専門語（あるいは専門語としての用法）には、名詞の比率が圧倒的に多く、対訳情報程度で十分にその目的を達成できる場合がほとんどであるなど、翻訳アルゴリズムに依存する情報が一般語の場合に比べて少ないので、特定の機械翻訳システムとは独立にデータを作成・管理できる部分が多いという特徴がある。ただし、適用分野や、適用業務によっては、特定の文体や書法を用いるといったことがある。例えば、製品パンフレットや規格文書などで、主語のない英文を扱う必要があるとか、プログラム・マニュアルなどで、例示や書式指定などの翻訳処理の対象にしにくい形式があるなどの事がある。これは、ユーザ辞書に係わる問題でもあるが、開発者側からは余り議論されていない。

専門辞書は、システム標準辞書と共に開発者側から利用者へ供給される場合が多いが、専門分野への依存性から開発者側独自の作成には限界があり、内容としては、その他の機関が作成したデータにはほぼ準じているようである。しかし、現在利用可能な専門辞書はきわめて限られており、また、データ内容に翻訳アルゴリズムへの依存性が薄いにもかかわらず、各機械翻訳システムで共通の専門辞書を用いることが出来るといった状況にもなっていない。

22.2.2 システム標準辞書の分類

システム標準辞書は、システム開発者側が翻訳プログラムと一体となって、利用者に供給するデータである。システム標準辞書は、その情報の性格に応じて、文法辞書、システム辞書（カーネル語辞書）、基本語辞書の3種類に大別することが出来る。ただし、文法辞書と基本語辞書とは、ほぼ完全に独立した体系ということが出来るが、システム辞書（カーネル語辞書）は、文法辞書と基本語辞書の両者の性格を合わせ持ったものを取り出した概念であり、そのインプリメントの方法は様々である。

(1) 文法辞書

翻訳処理に用いる文法ルール等のデータが記載される。機械翻訳システムのエンジンとは出来る限り独立させて、一般的な言語現象を記述できるように考慮してあるが、各機械翻訳システムのアルゴリズムへの依存性はきわめて高い。GRADEシステムなどが、機械翻訳システム用の文法辞書（文法記述言語）の洗練された例であろうが、そこでは、翻訳処理の中核的なメカニズムを記述するような内容になっている。従って文法辞書は、ほとんどの場合、ユーザからはまったく触れることの出来ない部分になっている。また、ユーザからの要求に対し、この文法辞書に関するエキスパートがいかにして改良を行っていくかという問題についての議論も見られない。

(2) システム辞書（カーネル語辞書）

助詞、助動詞、前置詞、冠詞などの「機能語」、あるいは、「する」、「なる」とか、“have”、“get”といった単語（カーネル語と呼ぶ）は、通常の単語の場合に比べて、非常に特殊な働きをする。こういった単語については、単語としての一般化した情報の他に、様々なアルゴリズムに係わる情報を与える必要がある。

これらは、単語辞書として、基本語辞書と同じ枠組みで扱う場合も多いが、カーネル語のためだけに必要な記述項目が、一般の単語についても必要になってしまうというようなことを避けなければならない。そこで、カーネル語に関しては、単語辞書の中に文法記述を行うことの出来る仕組みを実現する場合もある。また、文法辞書の中に、直接カーネル語の記述を導入する場合もある。しかし、こうした場合にも、ユーザや辞書要員がカーネル語に係わるような修正を行う時の一貫性の保証などの問題を考慮する必要がある。

(3) 基本語辞書

システム標準辞書の中の基本語辞書（一般語辞書ともいい替えられよう）は、分野を問わず用いられる単語およびその単語に関する情報を扱うものである。現在の機械翻訳

システムでは、一部には意味処理まで取り込んだ高度な自然言語処理技術を用いたアルゴリズムが採用されてきており、基本語辞書においても高度な内容の情報が要求されるようになってきている。例えば、動詞などについては、色々の用法に対応した文型や格パターン、共起する名詞などに関する意味的な制約情報などを記述する必要がある。一方、これに対応して、膨大な数の名詞などには、意味素性などの情報を記述しなければならない大量の語彙を対象として、これらを、矛盾なく、均質的に実現することは、非常に困難な課題である。現在、㈱日本電子化辞書研究所（略称：EDR）において、自然言語処理のための基礎データとして利用することを目指した電子化辞書の開発が進められているが、現在の処、（一部は機械可読になっている）冊子体辞書（人間用に出版されている辞書）などのデータと、辞書作成者の知識とに頼らざるを得ない。

22.2.3 専門辞書、ユーザ辞書の分野

専門辞書の作成・管理には、その規模の大きさや分野毎での専門的知識の必要性、時間的な変動要因など多くの問題を抱えている。

ユーザはシステム標準の辞書だけで満足するとは限らない。むしろ、ユーザ独自の必要性や知識に基づいて、独自のユーザ辞書や専門辞書を開発することを考慮する必要がある。そうした辞書作成を、円滑に、効率よく、そして翻訳システム全体の機能との整合性を保って行うことを保証する、本格的な支援環境システムの開発も要求される。

機械翻訳用の辞書では、改定、修正といったことが頻繁に発生すると考えなければならない。また、利用目的毎に内容に差異のある辞書が提供されるといった場合もある。こうしたときに、データやその修正内容をどう管理していくかは、開発者、利用者の双方にとって重大な問題となる。

従って、ユーザ辞書と専門辞書においては、或る意味で分けられにくい部分があることを承知する必要がある。

ここでは、専門分野を情報処理分野に限って、或る機械翻訳システムで試みた一部を述べた。

22.3 今後の課題

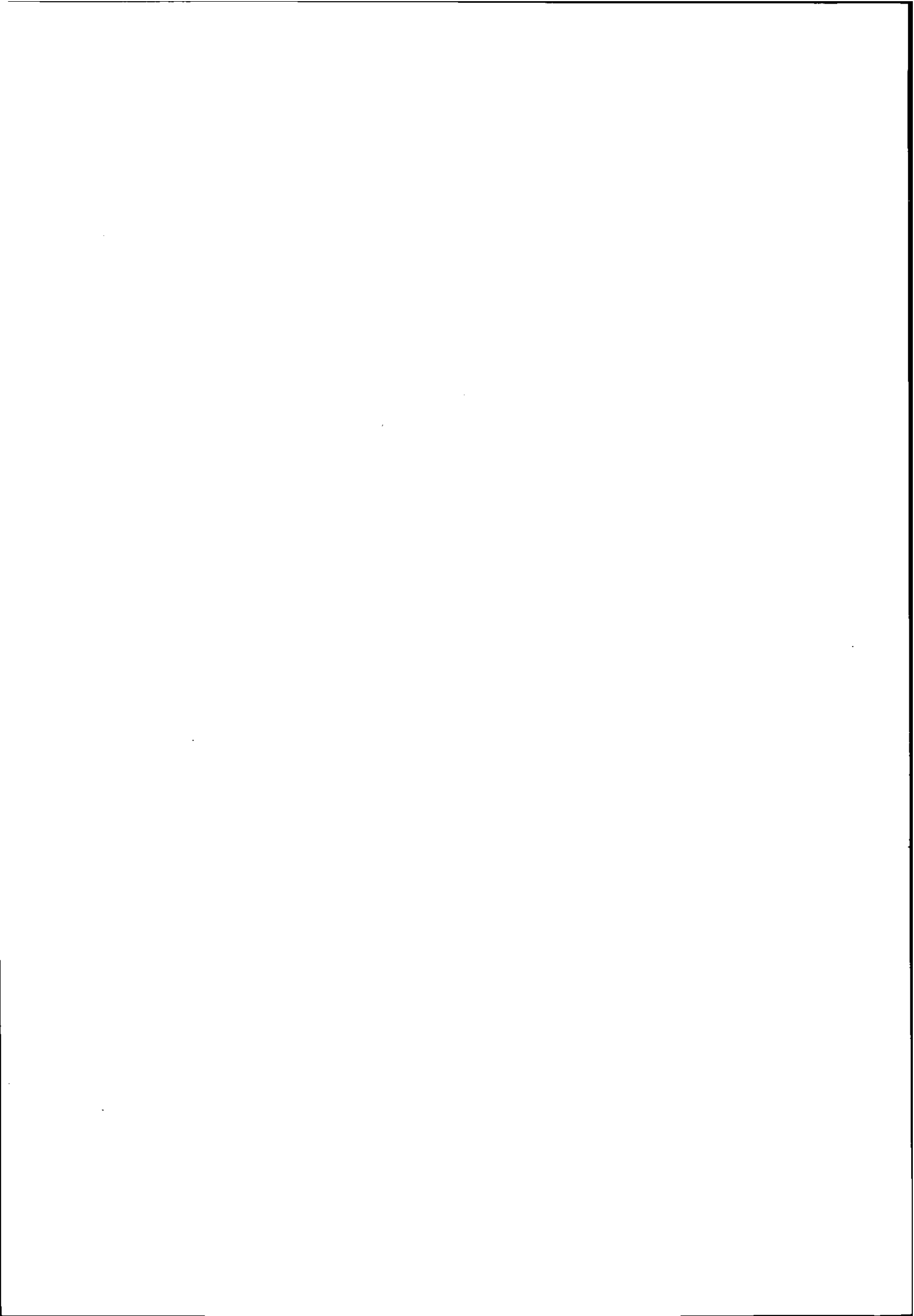
今後の課題として、機械翻訳システム用辞書データベースの提供サービスを実施するために、以下のことがあげられる。

- 1) 機械翻訳システム導入時のユーザ側負担の軽減として、経費の削減および作業時間の短縮。
- 2) 辞書データベースの更新の容易性と迅速性を図り、機械翻訳システムの質的向上、サービスの充実。
- 3) 辞書の一元化により、登録・更新される語等の概念の統一と利用上の解釈の混乱の防止。
- 4) 専門辞書データベース提供サービスによる、各分野への機械翻訳システム利用の容易性。

こうしたことが実現することは非常に難しいことであり、一朝にてできるとはいえないが、決して不可能ではないと考える。また、これらがうまく機能すれば、その相乗効果により、機械翻訳システムの利用が更に促進され、社会的にも大きな効果が期待できる。

23 国内外に対応する建築文献データベース
の構築に関する調査

財団法人 日本建築センター



23. 国内外に対応する建築文献データベースの構築に関する調査

23.1 目的

近年、建築分野における国際取引の拡大およびハイテク技術の利用増大という世界的傾向を背景に、国際的な文献情報交換へのニーズが急速に高まりつつある。しかし、わが国の建築技術情報は、ごく一部の研究論文が英文で著作されるくらいで、日本語で著作された文献の大部分は海外に紹介されていないのが現状であり、日本の建築技術情報に対しては諸外国からの強いニーズが予想されているにも拘らず、この問題に対応する国内体制が整備されていない状況にある。

一方、海外では、CIB/ICONDA という建築文献に関する国際的なデータベースネットワークが既存のSTN(科学技術情報ネットワーク)により地球規模の情報サービス体制が敷かれている。日本でもJICSTがSTNの構成員となり情報サービスを展開している。しかし、このICONDAに対する日本からの建築文献のデータ提供が強く求められているにもかかわらず、日本側からのデータ提供の体制は十分に整備されているとは言えない。

このような背景を基に前年度より国内外に開かれた建築文献の情報サービスを展開するため建築文献データベースの構築および効率的運営体制に関する調査を実施している。

前年度調査では、建築文献データベースの事業化方策としては、関連して先行している政府系のデータベースサービス機関(JICST, JACIC等)および民間資本によるデータベースサービス企業などとの役割分担、協力体制を調整して推進して行くことが重要であるという前提の基に、国内においても諸外国並みのデータベースの普及・活用されるような環境整備を行っていくことが最も重要な課題であるという結論を得た。

しかしながら、建築関連業界は大企業から個人経営の零細な設計事務所といった裾野の広い産業形態を有するという特殊性と建築生産や技術、学術的な面での分類方式、技術用語の不統一さなど他産業に比べデータベースサービスが普及するために解決すべき基盤となる問題は多いのが現状である。そのため、今年度の調査においては、これらの問題を念頭におき、データベース利用の普及と環境整備のためには、幅広い建築関連技術者が簡便に、かつ、低価格で利用するために望ましいサービス形態とはどういうものかについて検討を行うことにした。つまり全国ネットによるオンラインサービスということだけでなく、新たな情報提供媒体、コンピュータ機器の利用により、ユーザの資本投下能力、現有のシステム環境に応じた情報提供ができるよう多様なサービス方法の検討を行うことにより、国内における建築文献データベースサービスの需要喚起をまず図ることを今年度調査の主たる目的とした。

23.2 実施内容・方法

23.2.1 調査の内容

今回実施した調査の主な内容は次のとおりである。

- (1) 建築関連業界における定期刊行物の文献の収集・管理状況の調査
 - 建設会社の各部門（設計，開発，工務）における文献の収集・管理状況の調査
 - 建築設計事務所における文献の収集・管理状況の調査
- (2) 建築関連文献のデータベースサービスに対するニーズ調査
 - 建設会社の各部門における文献情報のニーズと外部に希望するサービス内容の調査
 - 建築設計事務所における文献情報のニーズと外部に希望するサービス内容の調査
 - 建築材料，部品メーカー，プレハブ住宅メーカー，建築設備関係のサブコントラクターに対する文献情報ニーズと希望するサービス内容の調査
- (3) 定期刊行物の建築文献を主体としたデータベースのサービス内容，方法の検討
 - 扱う文献の範囲の検討
 - データベースに盛り込む内容の検討
 - ユーザへのサービス形態の検討
- (4) プロトタイプの検討
 - データベースフォーマットの検討
- (5) フィージビリティスタディ
 - 各種のサービス形態による事業化比較
- (6) 建築文献データベースの国際化対応の検討
 - ICONDA サービスの現状調査
 - ICONDA へのインプット方法の検討

23.2.2 調査の実施体制

(1) 実施機関

財団法人日本建築センターが財団法人データベース振興センターから調査委託を受け、建築文献データベース構築運営調査委員会を設け調査を推進した。

(2) 協力機関

既存のCIB国内連絡協議会（CIBの正会員2機関，準会員5機関）を中心に関係する機関および学識経験者を含めて作業体制を構成した。

23.2.3 実施期間

1987年7月から、1988年3月まで

23.2.4 予想される波及効果

- (1) 国内における建築文献情報の交流が活発化される。
- (2) 国内における建築文献情報利用の環境が整備され、需要喚起に結びつく。
- (3) 建築文献情報の海外への情報提供ルートが確立される。
- (4) 建築産業の国際交流に対する摩擦解消の一助になる。

23.3 成果と今後の課題

23.3.1 調査の成果

海外の主要国では、建築関係の文献データベースが広範に整備・活用されており、他の産業分野とともに、全世界をネットワークとしたオンラインデータベースサービスが稼働するようになった。わが国においては産業構造の転換期に当たり、建築界においても建設会社を始めとして、経営戦略の転換、事業のソフト化といったことに対する動きも多い。

本調査の主体である建築におけるデータベースの分野においても、JICSTやJACICといった国がバックアップするデータベース提供機関が整備され、サービスが充実していくとともに、民間企業においても本分野で積極的な事業展開をめざしているところもあるといった状況であり、今後のデータベースサービスについてのそれぞれの機関での役割分担と、ユーザにとって使いやすいデータベースサービスを展開していくことについての相互調整の必要性は非常に高いことがわかる。

本調査で明らかになった主要な成果をつぎに示す。

- ・ 建築関連業界における文献の収集・管理・利用の状況が明らかになった。
- ・ 建築関連業界では、文献情報として、(1) 図、写真等のイメージ情報、(2) 建築経済指標、統計値などの数値情報、(3) 技術文献等の文字情報といった3つの分類ができ、それぞれに高いニーズとそれぞれの情報提供形態の独自性が考えられることが明らかになった。
- ・ 建築生産関連の業界構造はすそ野のひろがったピラミッド型の形態をしており、情報サービスとしては、幅広いユーザが利用できるような形態を取ることが、このデータベースを推進普及させるための絶対条件であることがわかった。

- 既存の建築関連文献情報サービスは、発生している文献情報のほんの一部であり、また、取り扱っているテーマも学術的なもの、建築作品などといった内容が特化しているものが多く、建築に関係ある文献すべてが対象とされているわけではない。建築関連業界は業種も多様で幅広く、それぞれの分野で多くの文献が発生しており、ねむっている貴重な文献について積極的に活用できるようなサービスについてのニーズが高いことがわかった。
- 日本は、技術力、経済活動の面で近年先進国として世界的に認知されているが、社会資本の蓄積はまだ不十分であり、日本で発生している建築技術文献を世界に広く流していく体制が確立されていない。その原因としては国際的な情報交流に対する意識がまだ低いこと、また、産業自体がそこまで資本投下能力がまだまだないこと、日本語・英語のランゲージギャップが存在していることなどである。

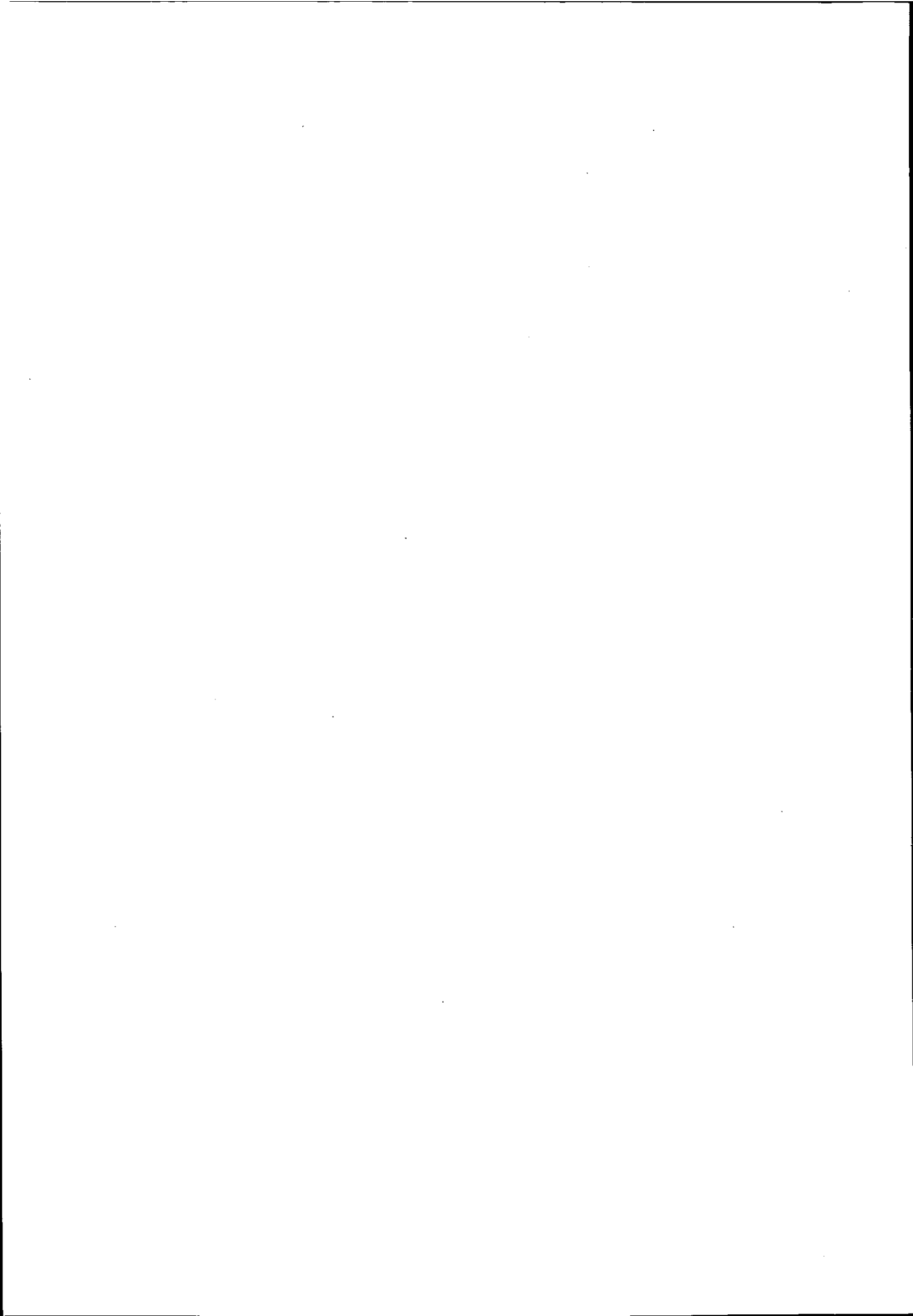
23.3.2 今後の検討課題

以上の調査結果から、今後の検討課題として下記の事項があげられる。

- (1) 前年度調査では、建築に関連する定期刊行物を約800冊抽出しているが、各業界とのヒアリングにおいて提案されたことであるが、設備・衛生といった分野における文献の抽出が少ないことなど、取り込む文献の選択について、幅広い業種を擁する建築関連業界の状況を考慮して、本文献データベースで収録しようとしていた文献の選択については更に詳細な調査が必要である。
- (2) 今年度調査における提案として、幅広いユーザに利用してもらえるようなデータベースサービスを展開していくことにしているため、それぞれのユーザの利用形態に即した情報伝達メディアとして、オンライン、パソコン通信、CD-ROMなどのデータ提供のメディアの技術的な面を含めた検討が必要である。
- (3) 建築文献データベースのニーズとして、将来的には、文献の原文を文字情報だけでなく、図形等を扱う図面情報や写真・絵等を扱うイメージ情報としての提供が建築関連技術者のニーズとして高いことがわかっており、どうデータベース化するかについての技術開発上の問題、データベース化する上での問題を検討していく必要がある。

24 翻訳支援ツール(英・和)としての
専門用語辞書データベースの構築

日外アソシエーツ株式会社



24. 翻訳支援ツール(英・和)としての専門用語辞書データベースの構築

24.1 開発の目的

近年、わが国の貿易黒字に起因する国際摩擦が増大している一方、情報に関しては、逆にわが国の輸入超過が非難的となっている。日本語の国際非流通性その原因であることは明らかであり、摩擦回避の目的ばかりではなく、製品輸出に伴うドキュメント、マニュアル類の要求も含め、技術資料等の翻訳(特に和文英訳)の必要性がこれまで以上に高まってきていると言えよう。

このような背景から、近年、機械翻訳の技術は急速な進展を見せている。しかしながら、それらの翻訳システムはいずれもプレエディット(前処理)、ポストエディット(後処理)を必要としており、人手を介さずに機械翻訳がなされているわけではない。人工知能的な推論機能を持ち、知能的な翻訳を実現するためには、相当数の研究者が長い年月をかけて開発することが必要であると予測されており、一切の人的関与を必要とされないレベルに達するのは、まだかなり先のことと見なければならぬ。

本開発は、機械翻訳における人的作業、または機械に支援された翻訳作業における支援ツールとして、特に一般の翻訳者には苦手な専門用語を中心として、豊富な文例を含む辞書データベースを構築するものである。更に、このデータベースを利用して、翻訳文書を効率的に作成するため、CD-ROM上に辞書データベースを格納し、パソコンで作動する翻訳文書編集プログラムの開発へ発展させようとするものである。

24.2 開発の概要

専門用語といってもその分野は多種多様であり、すべての分野を網羅する辞書を短期間に作成するのは困難なので、一分野毎に完成して行く方法を取った。

翻訳の現状をとらえてみると、技術資料関係が最も多い。技術関係分野の翻訳者は約1万5千人と推定され、そのほとんどの翻訳者が何らかの形でコンピュータ関係の翻訳に携わっていると伝えられている。このような現状を踏まえ、まず第一段階としてコンピュータ用語をデータベース化することにした。

辞書データベースに収録するデータ要素は、基幹語(英語見出し、品詞、日本語訳、発音記号、一般訳語)、解説、参考語、文例、複合語(必要に応じて解説付与)、複合語の参考語、文例とした。

本課題では、発音記号および通常語としての(一般辞書上の)意味も含んだうえで、特に基幹

語のオリジナル解説，文例，複合語に重点を置いた。

24.3 用語の選定

コンピュータ分野の辞典・用語集は出版社，学会，協会，新聞社等から数多く刊行されている。これらの辞典・用語集を調査し，本課題のデータベースを構築するために，ソースとして適している辞典・用語集を選び出した。

これらのソースを丹念に調べた結果，翻訳作業のために必要とされる基本的な用語（英語）は複合語を含め，約 5,000～6,000 語であることが推定できた。基本的な用語を単語レベルに分解すると，約 2,000 語になる。この単語レベルから，辞書データベースを構築するうえで重要な 1,496 語を「基幹語」として選定した。

用語の重要度を計る尺度は種々考えられるが，本データベースでは内容の記述量，複合語の語数により，基幹語の重要度を 4 ランクに分け「*」を付すことによって重要度の目安とした。

次に，前述したソースおよびその他の技術資料を参考に，基幹語を含む複合語をできるかぎり採集した。採集した複合語の数は 16,767 件を数え，コンピュータ用語辞書データベースとして十分な質と量を確保し得ると判断された。

複合語の採集後，基幹語に重要度のランク付けを行った結果，次の数字が得られた。

重要度	基準	基幹語数	構成比
***	特に重要で複合語が 20 以上	47 語	3.2%
**	重要で複合語が 15 以上	228 語	15.2%
*	重要で複合語が 10 以上	508 語	33.9%
無印	複合語が 5 以上	715 語	47.7%
合 計		1,498 語	100.0%

翻訳を行う場合，用語の対訳・定義がわかるだけでは十分でない。その用語がどうゆう場合に使用されるかを多く知り得た方が正確で，かつ，効率の良い翻訳を行うことができる。また，技術文書は文学関係とちがって専門用語（術語）を多用するものの，構文自体はさほど複雑でなく，文例の中から相当部分を引用できるものも数多く見受けられるので，本辞書データベースでは，できるだけ多くの文例を収録することにした。収録に当たっては，本辞書データベース構築の基礎資料とした典拠資料およびその他の技術資料を参考とし，文例を作成した。

24.4 英和辞書型の原稿作成とデータファイル作成

今年度は期間、予算の関係から実用タイプのデータベースを構築するのがむづかしく、選定した基幹語 1,496 語の中から 201 語を選び、プロトタイプのコンピュータ用語データベースを作成することとした。

本辞書データベースに収録するデータ要素は合計 16 になった。そのデータ要素に付した TAG を表-1 に示す。

表 24-1 データ要素と TAG 一覧

TAG	内 容	TAG	内 容
(n)	語番号	(c)	オリジナル解説 / 2 章目以降
(1)	見出し語	(8)	参考語
(2)	品詞名	(b)	文例
(3)	カタカナ読み	(9)	複合語
(4)	訳語, 日本語表記	(d)	解説
(5)	発音記号	(e)	参考語
(6)	一般的な意味	(f)	文例
(a)	オリジナル解説 / 冒頭部分	(x)	典拠コード

原稿作成作業は 3 つのグループによって進められ、①品詞, カタカナ読み, 訳語, 発音記号, 一般的な意味を記入, ②基幹語, 複合語の解説およびそれぞれの参考語を記述, ③文例, 複合語の整理・標準訳付与とし, これらを基幹語の語番号によって管理した。作業の概略フローを図-1 に示す。

作成された入力原稿をもとに, 入力・校正を行い英和辞書型のデータファイルを完成させ, 英和・和英型データベースソースファイル作成のための入力データとした。

入力で考慮しなければならないのは, 主に発音記号および上付・下付数字の処理の仕方である。発音記号の表記は表-2 のとおり対応文字を割当て, 対応文字のないものにはカタカナを割当てた。アクセント記号は該当文字の前に置く入力方式を取った。また, 上付および下付数字はそれぞれ, オーバーライン, アンダーラインを該当文字の前後に置いた。

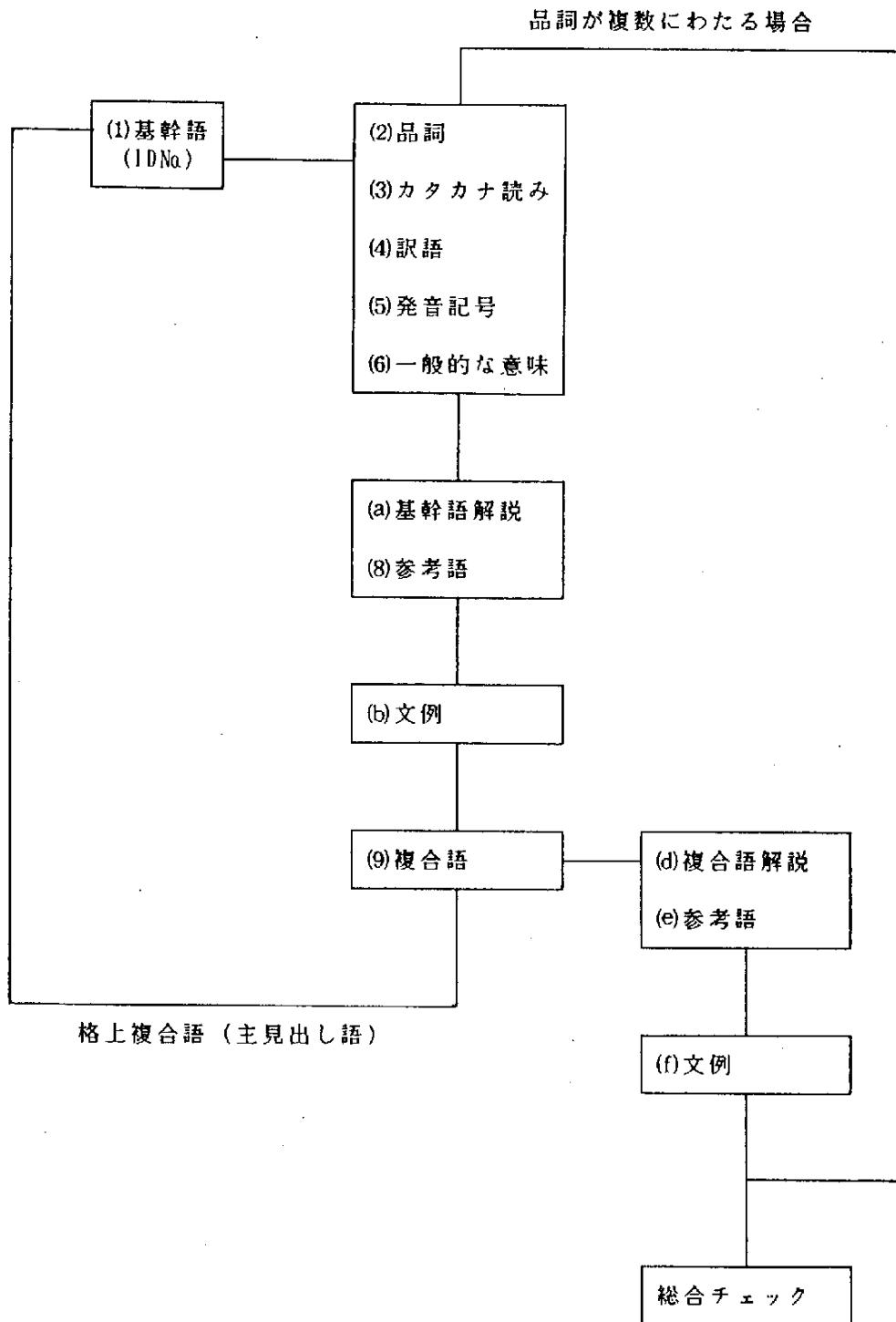


図 24 - 1 原稿の作成ステップ

表 24 - 2 発音記号の対応文字

記号	対応文字	記号	対応文字	記号	対応文字	記号	対応文字
i	i	y	y	l	l	w	w
e	e	u	ウ	f	f	r	r
ε	ε	p	p	v	v	a	a
æ	エ	b	b	θ	θ	x	x
ə	ヨ	t	t	ð	ズ	Φ	Φ
æ	ユ	d	d	s	s	;	ト
Λ	Λ	k	k	z	z		
a	ア	g	g	ʃ	シ		
o	o	m	m	ʒ	ジ	~	~
ɔ	オ	n	n	h	h	.	.
u	u	ŋ	グ	j	j	.	.

- (注) 1. アクセント記号は該当文字の前に置く。 ex, e → 'e
 2. 対応文字が、英字・ギリシア文字にないものにはカタカナを割当ててゐる。

24.5 英和・和英型データベースソースファイルの作成

英和型データファイルは作業用のソースファイルであり、このままではデータベースとはなり得ず、下記の各処理を必要とする。図-2に概略のプロセスフローを示す。

データファイル中では前述のとおり各種の用語がTAGで識別され、あるいは解説文中にマークされている。

主見出し語：TAG (1)に英文，TAG (4)に和文。

副見出し語：TAG (9)に「英=和」の形式，TAG (8)，(e)に「英(和)」の形式

および解説文(TAG (a)，(c)，(d))中に▲英(和)▼の形式

これらをすべて抽出し、「英=和」の形式に統一して、英文部分のコードをEBCDIC大文字に変換してソートキーとし、ソートしてキー部同一の用語を集め、和文部分の異なるものを連結して「英=和1；和2；……………」の形式に統合する。

この段階でリストをチェックし、校正もれ或いは用語の不統一等を発見し、原データファイルの該当箇所を修正する。

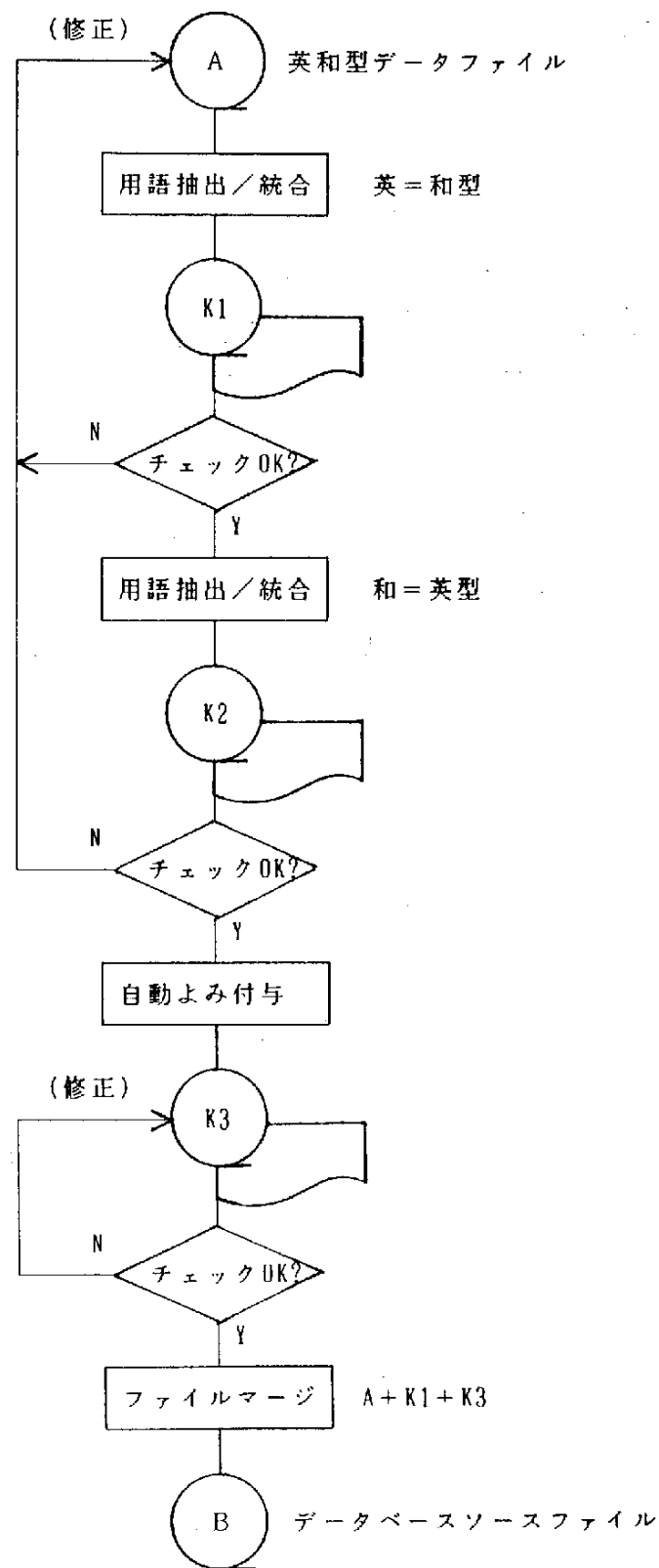


図 24 - 2 英和・和英型 DB ソースファイル作成プロセスフロー

英和型用語ファイルのチェックが完了したら、「英＝和 1；和 2；……………」型のレコードを「和＝英」型に展開し、和文部分をキーとしてソートし、キー部同一の用語を集め、英文部分の異なるものを連結して「和＝英 1；英 2；……………」の形式に統合する。

英文用語のアクセスキーとしては、用語抽出／統合過程でソートキーとして付与される EBCDIC 大文字列をそのまま利用する。

和文用語のアクセスキーとしては、当社開発の日本語データベース構築支援システム“NICE”の自動よみ付与機能を用いてかなよみを付与し、更に EBCDIC カナ文字に変換したものを用いる。

自動よみ付与機能は、必ずしも 100% 完全なよみを与えられるとは限らないので、和文の文字コード順または付与したよみの 50 音順のリストを出力して、付与したよみの適・不適をチェックする。

上記の過程で作成された用語ファイルの内容（英数カナアクセス付き）を、原データファイルの該当論理レコード中に併合することによって、データベースソースファイルが完成する。

このファイルはアクセス要素を含んでいるが、狭い意味でのデータベース（ランダムアクセス可能）ではないので、データベースソースファイルと呼んでいる。

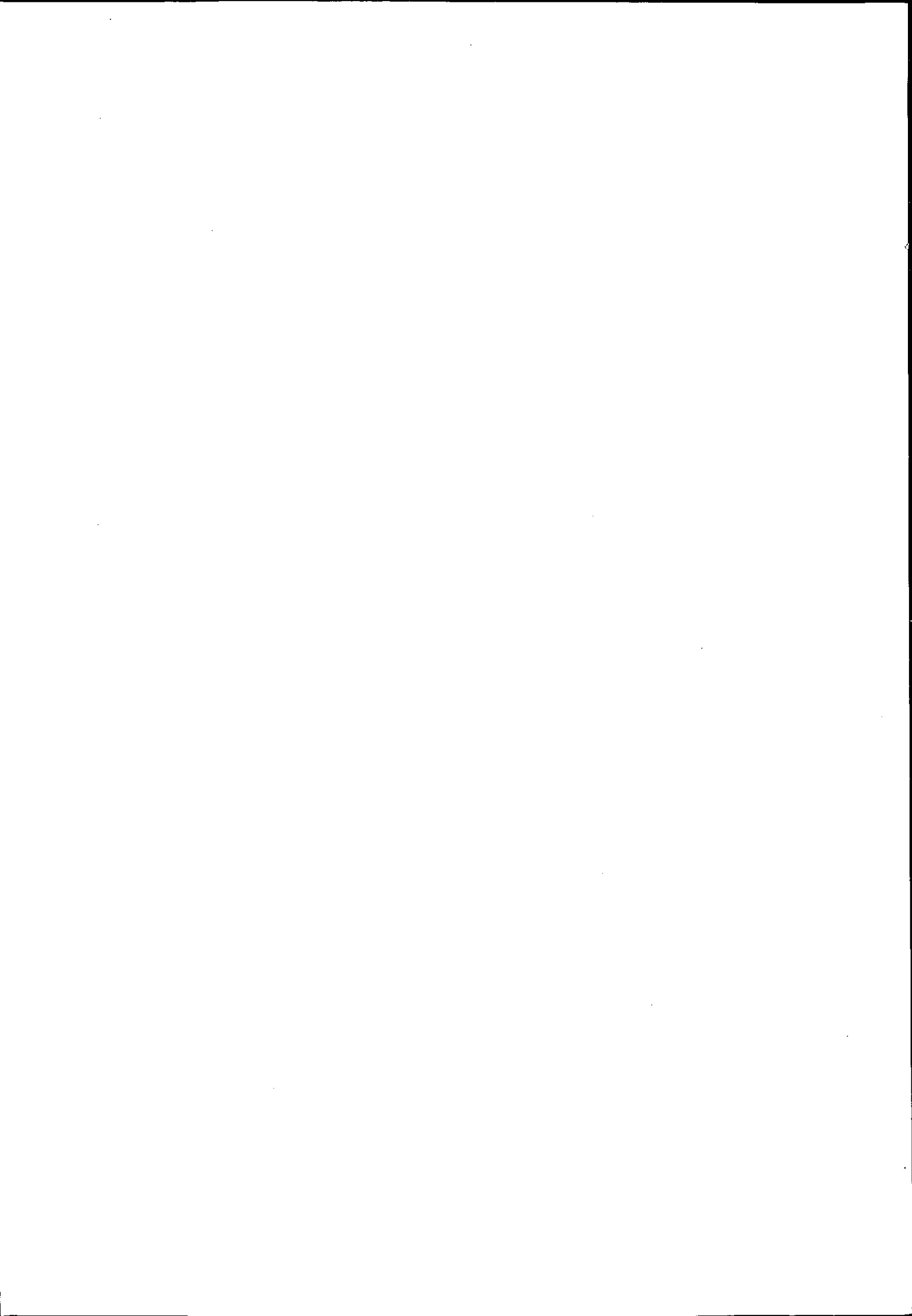
24.6 開発の成界と今後の課題

本開発は、翻訳支援ツールとしての英和・和英辞書型データベースのプロトタイプを構築するものであった。

まず対象分野としては、ハイテク産業として輸出入およびそれに伴う翻訳需要の多いコンピュータ関連分野を選択し、文例を採集した。基幹語の 201 語はもちろん不十分な数であり、実用上は選定した 1,500 語程度は必要である。しかしこの 201 語には約 7,800 語の副見出し語（複合語）および参考語が包含されることを忘れてはならない。参考語もまた辞書の見出しとなり得るので、僅か 201 語の基幹語から 8,000 語の辞書を作成し得たという言い方も可能である。

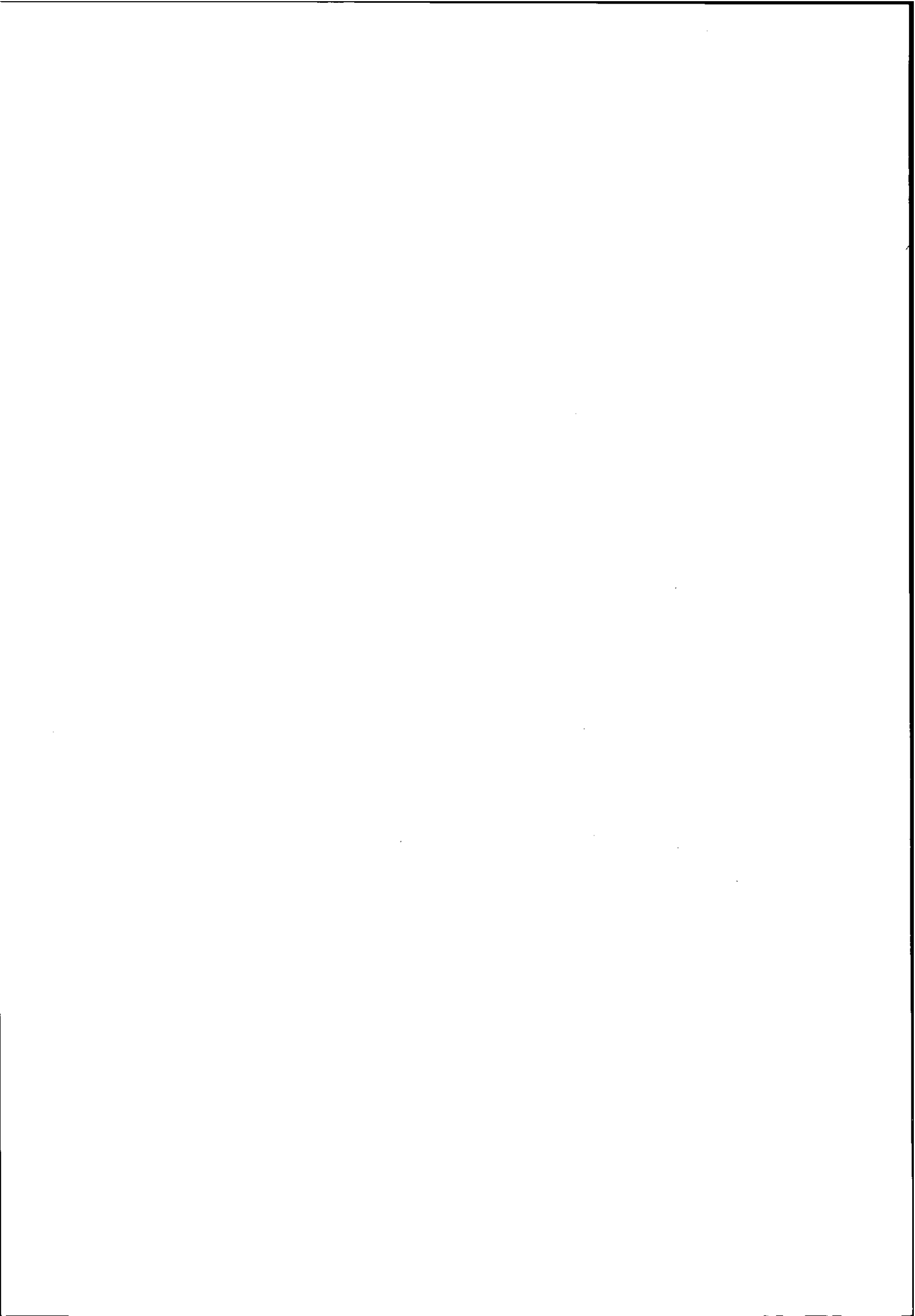
今後の課題としては、第一に基幹語の数を 1,500 語程度に増強し、実用上十分なデータベースを CD-ROM 上に構築すると共に、翻訳支援ソフトウェアを開発して、パソコン上で自在に翻訳作業が行える環境を開発することである。

以降は、段階的に対象分野を拡大し、産業・経済・技術分野のみならず、学術・人文・社会あらゆる分野の翻訳支援を可能にする大規模用語データベースを構築するのが目標である。



25 集合論に基づく知的情報検索システムの開発

シャープ株式会社



25. 集合論に基づく知的情報検索システムの開発

25.1 目的と内容

25.1.1 目的と予想される効果

ほとんどの人は自分の頭の中に記憶している以上の情報を必要としている。また、いかに有能な頭脳でも必要な情報をすべて記憶しておき、必要に応じて取り出すことは極めて困難であろう。しかも現代では必要とする情報は刻々増大している。

事務所のキャビネットの中に整然と保管してある書類や資料から情報を取り出すのは容易である。大量の情報を保管し必要な情報を柔軟に取り出すために、コンピュータをより積極的に利用することが不可欠であろう。コンピュータの中にさまざまなデータを記憶させる場合も整理して記憶させていればデータを再利用出来る。役立つ情報を必要なときに得るには、実世界に近いコンピュータモデルが必要である。

文書データは、情報を表現するものとしては最も一般的な形式である。しかし、文書データを構造的に管理する方法については統一的な一般的解法はない。個人の整理のノウハウにより管理が行われているのが現状である。文書データに対して、ユーザが自分の好みで構造化を行えば、検索は極めて容易になる。また、構造を利用し連想出来れば、検索は更に容易になる。同一のデータベースに対し、ユーザに合わせてユーザごとに構造化を行うことも必要である。

柔軟に対応可能なデータ構造の開発は、汎用的なシステムを構築する上で必要不可欠なものである。本開発は集合という基本概念に戻り、より基本的なデータのモデルを作成することを目指すものである。集合という他の多くのデータモデルにも共通する基本概念の長所を基盤として、新しいデータベース管理システムを開発し、プロトタイプとして実現することを目的とする。

本開発システムは、次のような効果を期待し開発する。

まず、文書データを容易に検索出来るようになり、文書は書類や紙面ではなく、コンピュータに記憶できるようになろう。また、ワードプロセッサやコンピュータで作成した文書の内容の確認が容易になり、文書名に気を遣うことなく様々な内容の文書を登録できるようになる。記憶された文書から自分の希望する文書が簡単に取り出せるようになり、コンピュータやワードプロセッサのユーザはデータベースの利用へと進む。

既存のデータベースシステムは構造化されたデータを管理している。一方、本システムは、構造化されたデータをフラットなものとして扱える。従って、既存のデータも活用出来る。つまり、既存のデータベースとの共存も可能である。しかもスキーマ変更等の作業は一切必要ないため、システムの導入は非常に容易であり、十分な普及が見込まれる。

ユーザが自分のデータ構造を容易に作成できる本集合論データベース管理システムは、既存のデータベースのみならず人工知能研究の問題となっている知識ベース管理にも影響をあたえよう。また、ユーザの利用状況を管理する機構は、機械翻訳システムやワードプロセッサにおける知的文書処理を実現するための辞書管理機構や、マンマシンインタフェースにおけるユーザモデルの認識への応用が期待される。

25.1.2 内容と開発体制

本開発は、まず最先端のデータベース管理システムの研究について広く調査することから始める。既存のデータモデル（ネットワークモデル、リレーショナルモデル、エンティティセットモデル等）を再確認し、我々が想定している集合論データベース管理システムの開発方針を決定する。

本開発のプロトタイプシステムにおいて、集合概念に基づく柔軟なデータ構造表現の実現とともに、データ構造の容易な生成と変更を実現する。つまり、ユーザの利用状況に従ってデータ構造が生成され、ユーザが希望する構造へと変化していくシステムを構築する。ユーザが希望するデータ構造を陽に定義出来るなら、ユーザは自分のソースを定義できることになり、最適なマンマシンインタフェースを持つことになる。

本開発では、データ構造に関する知識の表現方法に、新しいネットワークモデルを採用する。データ構造の入力システムは、素人のオペレータでも使える使いやすいインタフェースを持つ。データベース管理システムを実現する手段として、UNIXが持つ一般的なコマンド群およびシステムコールを駆使してプログラミングを行う。UNIXは様々なパターンマッチング機能を持ちデータベースの検索向きであり、プロトタイピングに適している。また、このためにユーザが希望する高速応答が実現できる。

IX-11をプログラム開発に用いる。IX-11はUNIXコンピュータであり、UNIXを高速に実行する。このため、プログラム開発に最適である。

開発システムはIX-11上でテストランする。ワークステーションやパソコン等のハードウェアとソフトウェアを利用し文書検索システムを構成する。

25.2 コンピュータアシスト文書検索システム

25.2.1 システム構成

本委託事業で開発したシステムをSTDB (Set Theory Data Base : 集合論データベース) と呼ぶ。STDBは開発マシンIX-11で実行可能である。STDBを業務に適用し文書検索システムとして利用した。

始めに、文書検索システムの機器構成について述べる。IX-11に接続しているターミナル、ワークステーション、パソコンからIX-11にログインしSTDBを実行することができる。これらの機器はIX-11に直接、あるいはIX-11が接続しているネットワークシステムに接続されている。機器毎の構成について述べる。

(1) IX-11

IX-11のOSはUNIXである。STDBはUNIXのコマンドとしてIX-11上にある。IX-11には、標準UNIXコマンド(パークレイ版、AT&T版、パブリックドメインコマンドを含む)のほか、JWPと呼ばれるワープロソフトとワープロユーティリティがある。

文書ファイルはIX-11上にある。文書ファイルはUNIXが管理する。STDB使用者はファイルを用途別や内容別に分けて蓄積する。

IX-11には外部入出力装置として磁気テープドライブが接続されている。ネットワークシステムとターミナルがターミナルドライバへ接続している。

(2) ターミナル

ターミナルはIX-11用標準漢字ターミナルである。画面上に漢字を40文字×24文字表示できる。表示文字は高輝度表示もできる。キーボードにはJIS標準キーの他、かな漢字変換やワープロソフトで使用する特殊キーがある。

ターミナルはネットワークシステムやIX-11に接続している。

(3) ワークステーション

ワークステーションのOSはUNIXである。UNIXの"cu"コマンドを使い、ワークステーションはターミナルとして使用できる。この場合、ワークステーションがIX-11用標準漢字ターミナルのコードを正しく解釈するように設定する。UNIXの"uucp"コマンドを使い、ワークステーションとIX-11間でファイルを送受信できる。

ワークステーションには外部入出力装置としてフロッピィディスクドライブとプリンタが接続されている。ネットワークシステムやIX-11がターミナルドライバへ接続し

ている。

(4) パソコン

パソコンのOSはMS-DOSである。ターミナルエミュレータソフトを使い、パソコンをターミナルとして使用する。この場合、エミュレータソフトがIX-11用標準漢字ターミナルのコードを正しく解釈するように設定する。ファイル転送ソフトを使い、パソコンとIX-11間でファイルを送受信できる。IX-11側からファイルを送信する時は、IX-11上にある送信用コマンドを実行しファイルを送らせる。UNIXファイル変換ソフトを使い、UNIX形式フロッピーディスクからファイルを読み込み、逆にMS-DOSファイルをUNIX形式でフロッピーディスクに書き込む。ワープロファイル変換ソフトを使い、ワープロ形式フロッピーディスクからファイルを読み込み、逆にMS-DOSファイルをワープロ形式でフロッピーディスクに書き込む。

パソコンには外部入出力装置としてフロッピーディスクドライブとプリンタが接続されている。ネットワークシステムやIX-11がターミナルドライバに接続している。

(5) ネットワークシステム

ネットワークシステムにIX-11、ワークステーション、パソコン、ターミナルが接続されている。ネットワークシステムに接続されているマシンは他のマシンに接続できる。ネットワークシステムはおおのこの回線を接続するだけで、STDB使用に影響しない。

25.2.2 対象業務

下記文書を文書検索システムで検索し、業務に利用した。

(1) 部門でワープロ文書を管理する。

担当者がIX-11のワープロで作成した文書を内容別に登録する。年度報告書や稟議書といった部門発行文書ワープロファイルを部全体で検索し再利用する。出張報告書や提案書といった個人発行ワープロ文書を部員が検索し再利用する。STDB利用者は用途別にワープロファイルをグループわけする。新たに作成した文書は用途にあったグループへ追加登録する。以前作ったワープロファイルを検索し、その場で再利用できる。

(2) プログラム、仕様書、マニュアルを管理する。

プログラムライブラリ、仕様書、マニュアルを文書ファイル形式で登録する。STDB使用者は目的別にグループわけをする。ファイル内容が改正されたとき、古い情報を新しい情報に入れ換え内容変更処理を行う。内容を知らない人も検索できる。

- (3) 磁気メディアで購入した資料を管理する。

図書目録，人名録，辞書といった情報を文書ファイル形式で登録する。専用検索ソフトとは違った検索ができる。新聞，論文集，資料集といった情報を登録する。内容で検索し，必要な所だけを選んで読む。STDB使用者は資料毎にグループわけをする。内容が改正されたとき，古い情報を新しい情報に入れ換え内容変更処理を行う。情報が多用途で利用できる。

- (4) 電子メールやニュース記事を管理する。

個人宛に送られてきた電子メールや電子掲示板で見つけたニュース記事を登録する。必要かどうかわからなくても文書ファイルとして登録しておけば，キーワードで検索できる。ファイル名で検索する必要がなく，ファイル名に工夫がいらないため登録の手間が減る。

25.2.3 導入効果

文書検索システムを導入し，下記の効果が得られた。

- (1) ワープロファイル検索時間の短縮

1回のワープロファイル検索に30分～1時間かかっていたのが，STDBを使い30秒になった。ファイル管理に1日トータルで3時間かけていたのが10分になった。紙ファイルで検索したあと，そのワープロファイルを探すといった2度手間がなくなった。

- (2) 作成文書の内容統一

稟議書や出張報告書といった定形文書を再利用し文書が統一された。提案書や報告書に以前のを引用でき文書内容が統一された。

- (3) 保管効率の向上

ファイルを登録するときに，細かく分類したり工夫したファイル名を付ける必要がなくなった。紙ファイルは探した後，元あった場所に戻さなくてはならないが，その必要がなくなった。

- (4) 検索ミスの減少

内容見落としがなくなり検索が正確になった。検索しようとした文書と同じような内容の文書は一度に検索でき，探し間違いが減った。ワープロファイルが存在すれば，必ず見つかるようになった。

- (5) 利用回数の増加

コンピュータ内に貯えられている様々な情報を，同じ操作で検索できるようになった。

購入した情報を検索、利用する回数が増えた。

25.3 データベース管理システム

25.3.1 入力

STDB が検索するファイルを文書ファイルと呼ぶ。文書ファイルは UNIX 上のファイルであり、UNIX のファイル管理機能で管理される。文書ファイルの先頭のコードから最初の改行コードまでを 1 行目、改行コードの次のコードから次の改行コードまでを 2 行目（以下続く）とする。最後の改行コードから最終コードまでを最終行とする。文書ファイルの最終コードは改行コードである必要はない。

ワープロ別に用意した専用ツールを使用し、ワープロファイルからアンダーラインや倍角といった制御コードを除き文書ファイルとする。また、テープやフロッピィディスクを媒体として発売されている情報から文字情報を抜き出し文書ファイルとする。

STDB 利用者は、文書ファイルを内容別でディレクトリにわけ登録する。文書ファイルが登録されているディレクトリを文書ファイルディレクトリと呼ぶ。

25.3.2 表示

STDB は注目している文書ファイルディレクトリ名を表示する。このディレクトリをカレントディレクトリと呼ぶ。カレントディレクトリに登録されている文書ファイルに存在しているかどうか調べる文字列をキーワードと呼ぶ。また、キーワードがどの文書ファイルの何行目にあるかあらわしたものをインデックスと呼ぶ。カレントディレクトリに登録されているキーワードやインデックスを表示する。また、登録されているキーワードについてそのキーワードを含む文書ファイルからキーワードを含む行を表示する。

25.3.3 操作

STDB は stdbsh と stdbcreat の 2 つのコマンドを持つ。

stdbsh は引数に文書ファイルディレクトリ名を与え実行する。指定した文書ファイルディレクトリがカレントディレクトリになる。stdbsh はコマンドインタプリタであり、文書検索に必要な機能を持つコマンドが実行できる。このコマンドを stdbsh コマンドと呼ぶ。

stdbcreat は文書ファイルディレクトリにキーワードとインデックスを登録する。また、文書ファイルを追加登録する。

STDB は UNIX のファイル管理機能をそのまま使用しており、ファイルとディレクトリへの操作は UNIX コマンドを使用する。また、STDB は文書ファイルディレクトリ内の文書ファイルや登録されたキーワードを削除できない。削除が必要な場合は新たに登録し直す。

表 25-1 stdbsh コマンド一覧

コマンド名	機能
stgrep	インデックスを生成し表示する。
stegrep	インデックスを表示する。
stls	キーワードを表示する。
stmore	検索文書を表示する。
stman	stdbsh コマンドを説明する。
stcd	文書ファイルディレクトリを変更する。
sh	UNIX シェルを実行する。
help, ?	stdbsh コマンド一覧を表示する。
end, halt, exit	キーワードとインデックスを登録して実行を終了する。
quit, q	キーワードとインデックスを登録しないで終了する。

表 25-2 stdbcreat 使用法一覧

使用法	機能
stdbcreat ファイルディレクトリ	ファイルに記述されている文字列をキーワードとし、ディレクトリに登録する。
stdbcreat コードファイルディレクトリ	ファイルに記述されている文字列をキーワードとし、ディレクトリに登録する
stdbcreat ディレクトリ	ディレクトリにファイルを追加登録する

25.3.4 制御

STDB はキーワードとインデックスを文書ファイルについて生成し管理する。キーワードはアスキーキャラクターコード（改行コードとスペースを除く）、半角カタカナコー

ド、シフト J I S 漢字コードの例である。キーワードへは管理番号が与えられる。キーワードは登録番号とともにキーワードファイルに登録される。インデックスは対応するキーワードの管理番号で指示されるインデックスファイルに登録する。キーワードファイルとインデックスファイルは文書ファイルディレクトリに一对一に対応するインデックスディレクトリに登録する。

25.3.5 出力

検索した文書ファイルは、UNIX コマンドを使って文書全体を表示する。UNIX のエディットコマンドや清書コマンドで編集する。変換ツールを用いてワープロファイルに変換しワープロで編集する。印刷の場合は、文書ファイルを IX-11 からワークステーションやパソコンに転送し、転送先のプリンタで印刷する。

25.4 今後の展開

本課題で開発した STDB は文書検索システムとして利用できる。STDB の機能と特徴を生かして、今後は次のように展開していく。

25.4.1 集合論データベースシステム

文書検索知識はキーワードとその組み合わせであるといえる。キーワードの組み合わせを登録、利用する機能を実現する。これは文書検索知識を表現し利用するシステムであり、STDB はエキスパートシステムとして機能する。キーワード作成と、その組み合わせを効率よく検証し登録するためのツールを開発する。

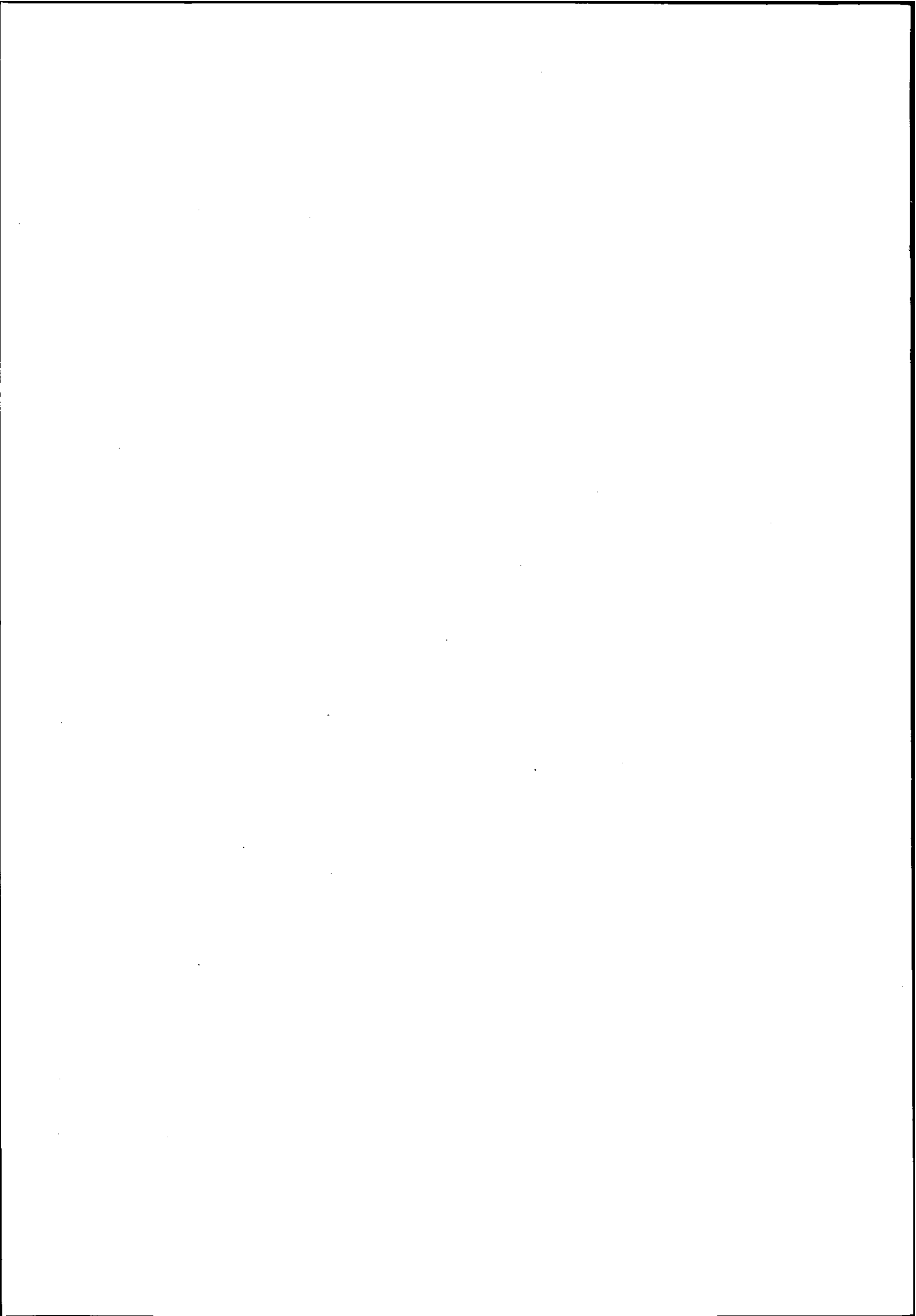
25.4.2 文書検索システム

STDB はキーワードを条件として文書を検索する。文書検索では、このほかの条件として日付やページ数といった量であらわされた条件を使う。日付やページ数についての記憶はあいまいなものであり、あいまいな情報を具体化させ条件をしばりこむシステムを実現する。

25.4.3 マンマシンインタフェース

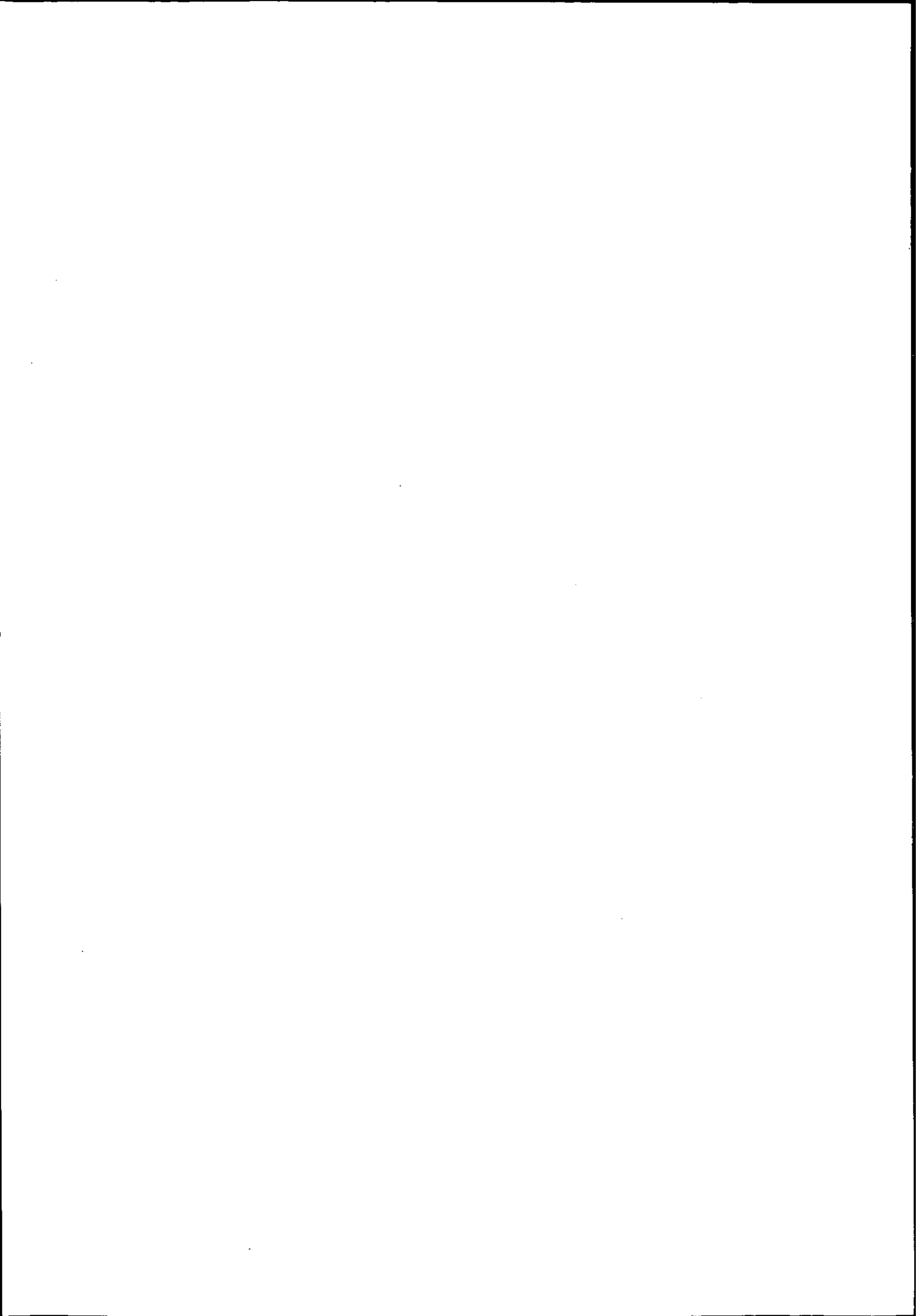
システムで重要なのは、インタフェース部分である。STDB は UNIX シェルと UN

I X コマンドをまねたインタフェースを実現し、UNIX 利用者になじみやすいインタフェースを実現した。条件式とその検索結果をベン図で画面表示するなど、使用者に視覚情報を提供するライブラリを開発する。また、だれでもが使用できるように、アイコンやスプレッドシートで操作するインタフェースシステムを開発する。



26 エキスパートシステム技術を用いたデータベース
知的インタフェースの試作

株式会社 シー・エス・ケイ総合研究所



26. エキスパートシステム技術を用いた

データベース知的インタフェースの試作

26.1 目的

情報化社会の急激な発展に伴い、情報量はますます増大している。こうした状況の中で、必要な情報すべてを自ら収集し管理することは極めて困難である。

このような大量の情報を体系的に整理・蓄積し、必要な時に必要な情報をタイムリーに供給できるデータベースの構築とその利用は、欧米各国においてもその重要性が高く認められており、実用例が相次いで公表され、利用されている。わが国においても、様々なニーズに基づいたデータベースシステムが開発されている。しかし、こうしたシステムは、誰もが簡便にかつ低コストで使う訳にはいかない。そのため、わが国におけるデータベースは利用者側の積極的な利用機運が今一步であることは否めない。

このようにデータベースは、その重要性や利用ニーズに対して、十分にこれを充足しているとは言えない。その理由は以下の2点に集約されよう。

第一は、データベースの検索等の煩雑さ、いわゆるマンマシンインタフェースの問題である。

第二は、データベース関連のサービスを享受するためのインフラストラクチャの問題である。

言い換えると、データベース利用者に対する友好的かつ魅力あるインタフェース機能とデータベースを利用するための端末設備のコストパフォーマンスとの二者を、いかにして同時に達成するかという問題である。

そこで、エキスパートシステム技術を用いてデータベースのインタフェースを構築し、データベース利用者に対する友好的かつ魅力あるインタフェース機能とデータベースを利用するための端末設備のコストパフォーマンスの二者を同時に達成することを目的として、開発を行った。

26.2 内容

エキスパートシステムとは、「専門知識を蓄積した知識ベースと、それを利用して問題解決を行うための推論機構を備えているものである。」とまとめられよう。(図26-1参照)

ところが、従来のデータベースシステムでは、データ検索をする場合にシステムは単にデータの有無を返すだけで、ユーザとの対話がほとんどなく、検索対象への接近過程が極めて限定されている。利用者がデータベースシステムを使ってデータを検索する場合、ユーザが希望するデー

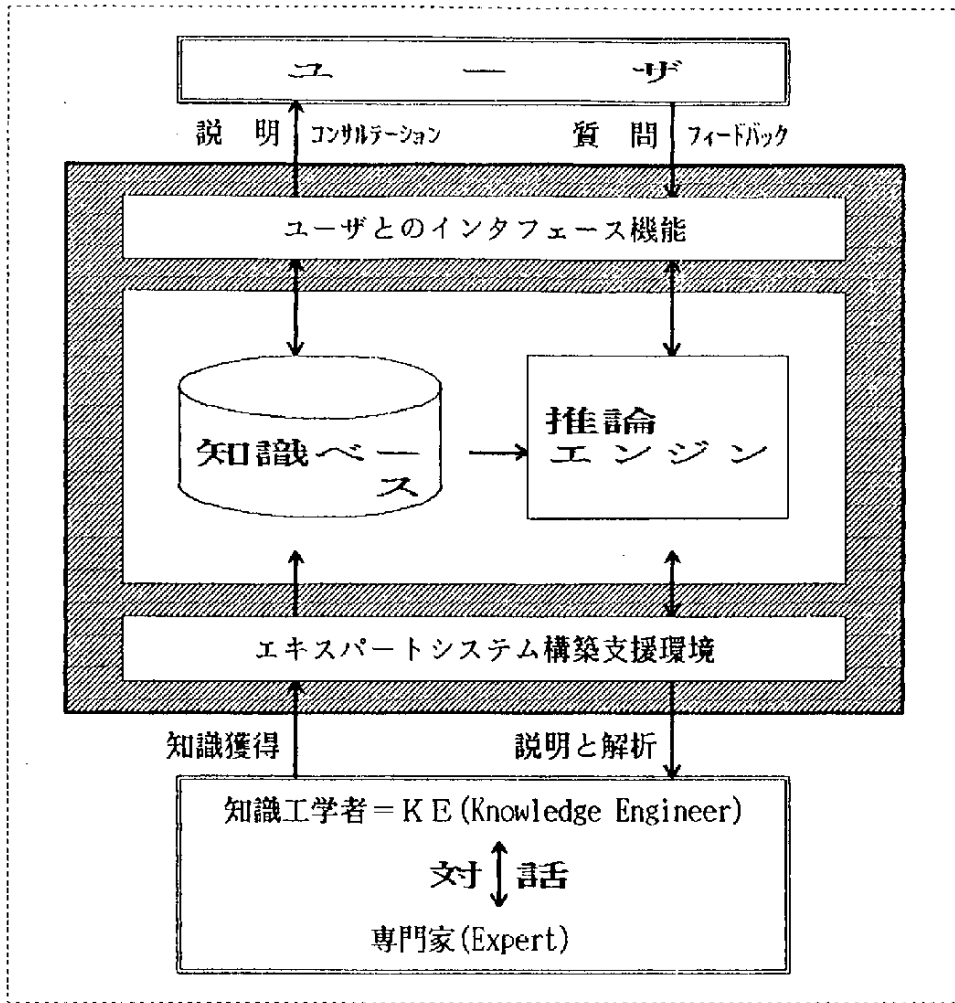


図26-1 エキスパートシステム

タがデータベースにない時、利用者は類似データ検索を試みる。また、仮に従来のシステムがデータを検索してきても、そのデータがユーザの望んでいるものとは違う場合もある。それは、システムとユーザとの間の対話がほとんどないため、システムがユーザのニーズを把握できないのが原因であると考えられる。そのため、利用者は自分の希望するデータが検索されるまで入力・検索失敗を繰り返さなければならない。(図26-2, 26-3 参照)

そこで、本システムでは、エキスパートシステムの特徴を生かして対話型のインタフェースを実現し、検索対象への無理のない到達を可能にしている。(図26-4 参照)

このデータベース知的インタフェースの試作において、不動産(分譲マンション)紹介業務を選択した。その理由は、以下のとおりである。

分譲マンションというものは、購入者の希望に合う物件が見つかることの方が少なく、大抵の

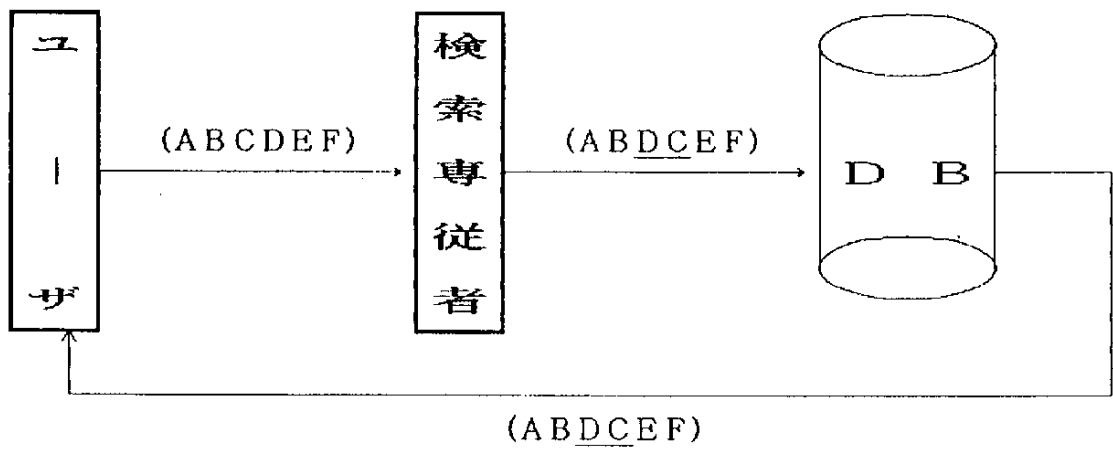


図26-2 検索専従者を介しての検索

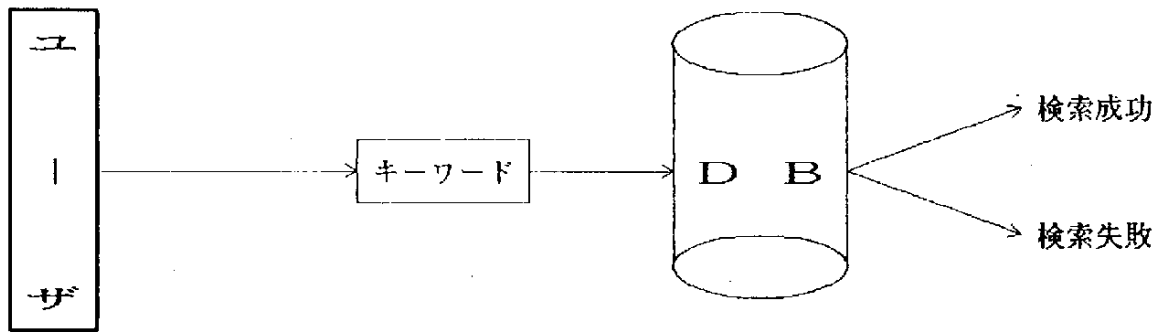


図26-3 従来型の検索

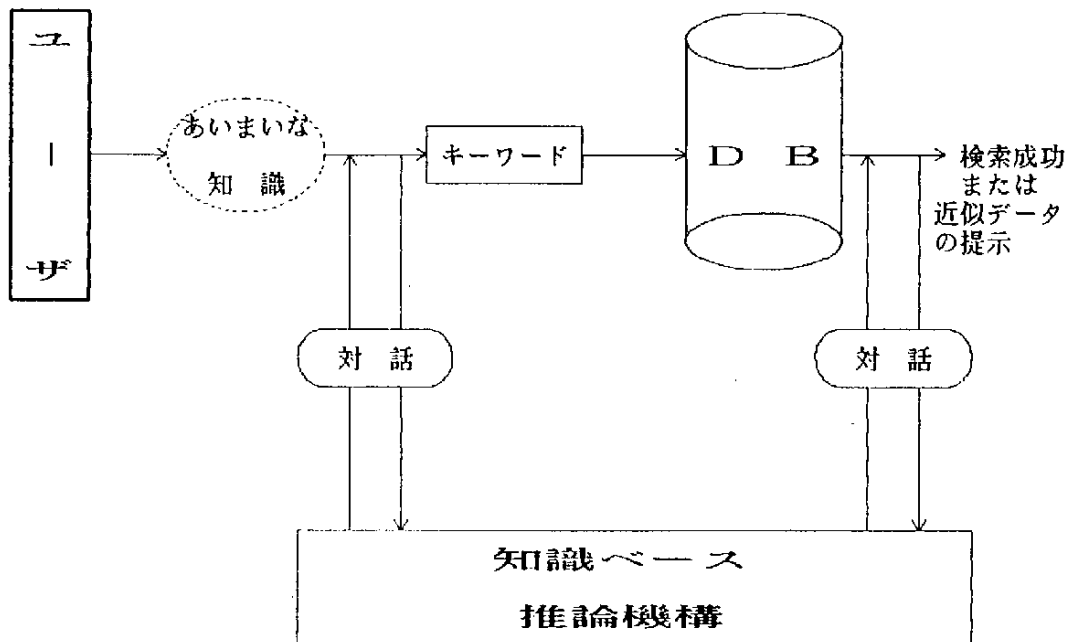


図26-4 知的検索

場合不動産業者は、購入者の希望条件に類似した物件を紹介することが多い。その観点から、従来システムでは、購入者の希望する物件をうまく検索することが難しいと考えられる。

以上のことより、対話形式によりユーザの希望条件を的確に把握（推論）し、データ検索を行う知的インタフェースのサンプルデータとして不動産（分譲マンション）紹介業務を選択し、インプリメントを行った。

本プロジェクトでは、人間中心的側面から検討を加えた形でマンマシンインタフェースの簡便化を図り、入力時の煩雑さを取り除く。また対話型を基本とし、利用者との緊密な情報交換を重ねつつ、場合によってはアドバイスも交えながら、検索対象への無理のない到達を可能にするものになっている。

インプリメントに当たっては、使う側の立場に立ち、できるかぎり簡便な入力方法を実現するために、タッチタブレットを利用したメニュー選択を採用した。

そのような簡便な入力方法を利用し、初期データとして、価格、広さ（間取り）、沿線、地区の4条件を調べ、検索を行う。そのデータによりある程度物件を絞り込み、管理費、駅からの距離、設備等の条件により最終的に物件を絞り込む。検索の基準となる価格、広さ（間取り）、沿線、地区の4条件の優先順位は、システム内で固定しておき、条件に見合う物件がない場合、優先度の低いものから変更・追加も可能である。

このことにより、あたかも人間（不動産業者）が対応しているようなきめの細かい検索を実現している。

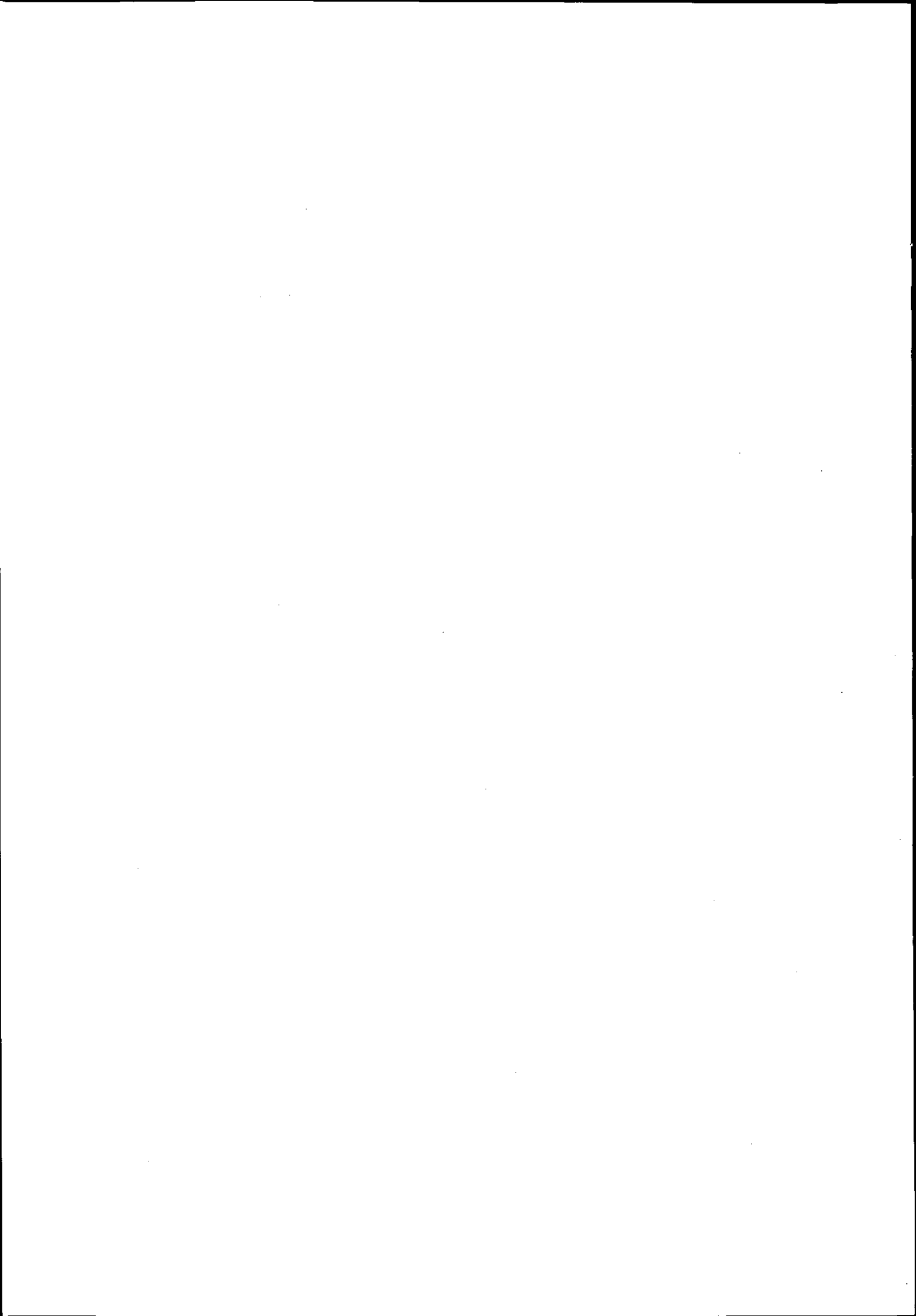
26.3 今後の課題

今後は、音声・画像利用技術などとデータベースとの融合化や、またデータベース利用機運向上のためにデータベースの内容の充実を図り、データベースの利用機運の向上に努める必要がある。そのためには、①政府からの融出資制度、政府所有情報を利用したデータベースの構築、②データベースに関する技術の共同開発等の促進等を行う必要がある。

昭和62年に各省庁から一部提供の始まった政府所有情報は、膨大であり、かつ内容的にも非常に充実している。この情報を利用し、日本開発銀行の融出資制度などを有効に使って利用価値の高いデータベースを構築することは、データベースの利用機運向上にとって急務であろう。

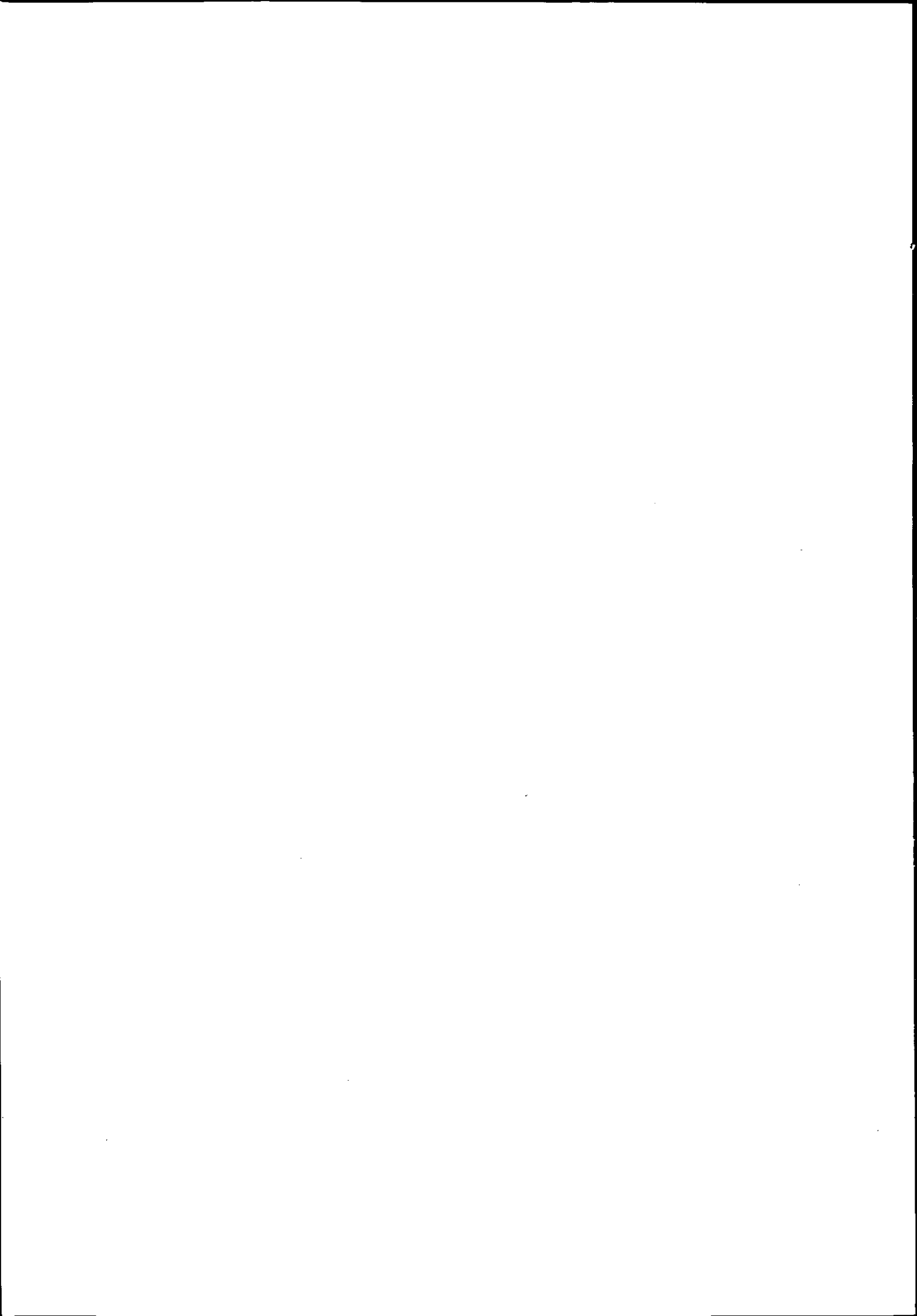
また、現在のデータベース・システムは、特にインタフェース部分に問題があり、非常に使いにくいシステムになっている。また、検索コマンドが煩雑な上、入力デバイスが不統一であるということも、その原因の一つであると考えられる。

最近、人工知能技術などの先端技術開発が、盛んに行われている。そのような先端技術とデータベースとの連結は、データベースの利用機運の向上には必要不可欠である。このためには、官民一体となったデータベース関連技術の開発体制の確立が望まれる。



27 イメージ・データのコード・データ変換
に関する調査研究

株式会社 エービーシ



27. イメージ・データのコード・データ変換に関する調査研究

27.1 はじめに

書籍等に収録されている膨大な情報を、コンピュータ・システムにおいて容易に検索できるようにすることは、情報処理の分野として非常に重要な意味を持つ。しかし、これら紙面などに記録されている情報の検索は、コンピュータで処理することができないため、人手による作業に委ねるしかなくいわゆるデータベースとしての情報とはならない。従って、これまでに作り出されてきた貴重な情報を有効に活用するためにも、これらの情報をコンピュータで処理できる形に変換する技術の開発が早急に実現されることが望まれる。

情報にはその重要性が、時間とともに減少するものばかりではなく、論文、判例などのように、時間の経過とともに増加するものもある。これらをデータベース化しようとする試みにおいては、印刷された情報は、光学的な変換を経て光ディスクなどの媒体に収録される。その際、現在の技術レベルでは紙面上の情報を、無条件に光ディスクなどの媒体に収録することしかできない。当然、検索を目的としていながら、キーワードを付与することはできない。

このようにして光ディスクなどに収録された情報を、コンピュータによる処理の対象にするためには、多くの課題を解決しなければならない。また、光ディスク媒体上に収録されている情報から、検索を目的としたキーワードを抽出するためには、パターン認識、自然言語処理、データベースなどの分野における非常に高度な課題も山積している。

このようないくつかの重要な問題を解決できるならば、重要論文、歴史的資料、法令、判例などの整理における入力およびデータベース化が促進されることになる。これらの情報のデータベース化は非常に強く要請されているものであり、国家レベルでの研究推進のためにも早期実現が期待される。また、ここで議論するイメージ情報からの文字情報の切り出しをベースとする新しい技術が実用化されれば、その他取り残されてきた領域のデータベース化も急進し、その研究の進展に貢献することが可能である。更に印刷情報の読み取りから文字の切り出しに至る過程の後に、自然言語処理の段階で音声合成に連結すれば、朗読システムへと発展することになる。

尚、本調査研究の実施に当たっては、社内委員会を設置して行うと共に、外部の識者からも専門的な意見を聴取し、参考とした。

27.2 調査研究の範囲

本調査研究の対象となる技術要素の関連を図示すると図27-1のようになる。

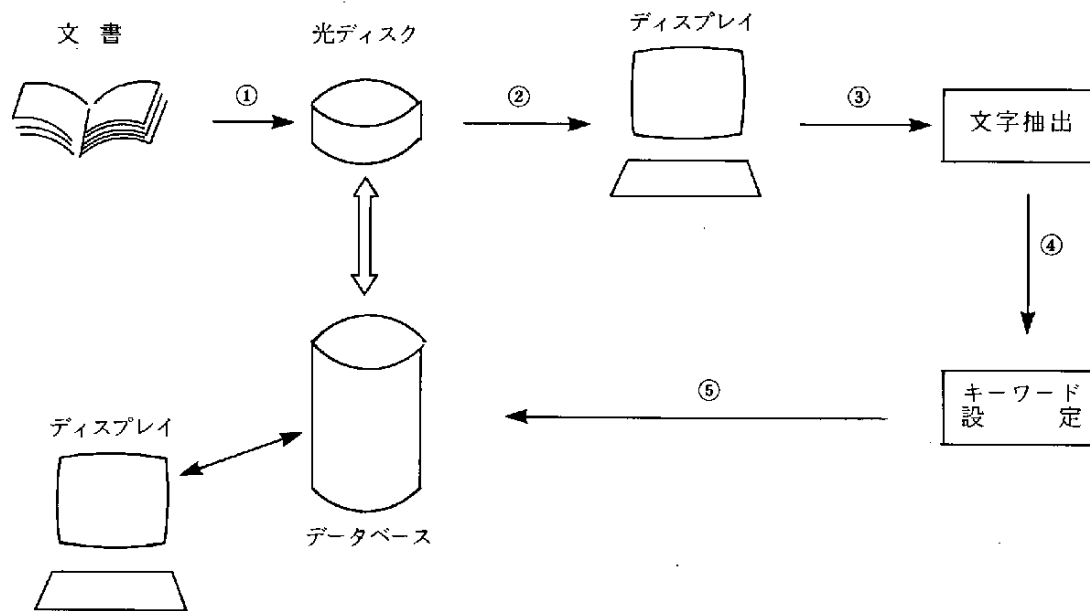


図27-1 関連する技術要素

すなわち、本調査研究がカバーする範囲は、次の5分野に大別できる。

- ① すでに紙等の媒体に印刷されている情報を、大容量記憶媒体である光ディスクに記憶しなおす（媒体変換）。
- ② 光ディスクに記憶されている情報をディスプレイ上に表示する（情報表示）。
- ③ ディスプレイ上に表示された情報に、高度なイメージ処理技術を施して、その中から文字情報を抽出する（文字抽出）。
- ④ 抽出された文字情報に自然言語処理技術を適用して、有意のキーワードを抽出する（キーワード設定）。
- ⑤ 抽出されたキーワードを使用して、光ディスクに記録されている情報との連結（リンク）を設定し、データベース化を図る（データベース構築）。

27.3 解決すべき技術的課題

前節において定義した、本調査研究の網羅すべき5分野の範囲について、各々の問題点を明確にしておいた上で、本質的に本調査研究が対象としなければならないものを見極め、その限定された範囲について具体的な問題設定を行うこととする。

(1) 媒体変換における課題

まず、情報を読み取る技術の観点から現状を確認すると、情報の読み取り装置の代表としてのOCR（光学文字読取り装置）も限定されたアプリケーションにおいては適用されだしている。郵便番号のように数字に限定するとか、数字の形を決めるなどの非常に強い制限の下では、手書きの文字の読み取りも可能になってきている。

一方、読み取られた情報を記憶する技術については、従来からの高密度磁気大容量記憶装置を始めとして、光ディスク装置、CD-ROM、光磁気ディスク装置と、かなりの種類の媒体に対する可能性が考えられるが、イメージ情報の記録という観点からは、現実的には光ディスク装置が最適である。しかし、本調査研究の成果が生かされる世界においては、現状の光ディスク装置の性能では不足していて、更に高度な技術開発が要求されるところである。

(2) 情報表示における課題

光ディスク装置に蓄積された情報を、ディスプレイに表示する技術そのものは、現在提供しているレベルで十分である。

(3) 文字抽出における課題

光ディスク装置などに蓄積されているイメージ情報から、有意の文字情報を切り出すことが本調査研究の本質的な目標である。イメージ情報から文字情報を切り出す技術は、いかにソース・データが印刷物からとったものであっても、手書きOCRにおける漢字の識別に適用されるものを導入しなければならない。現状では、任意の手書き漢字を識別できる技術はまだ未完成であり、今後かなりの精力的な研究・開発を待たなければ実用に耐え得るものは提供されないであろう。

(4) キーワード設定における課題

抽出された文字から有意の文字列を構成して、イメージ情報に対するキーワードを設定していかなければならない。このための技術は自然言語処理の一環をなすものであって、これまた現状の技術レベルは実用の域に達していない。しかし、昨今の自然言語処理の中における日本語処理は急速に進展しており、そう遠くない将来に文字列を適当に切り出してキーワードに変換できるようになることが期待できる。

(5) データベース構築における課題

イメージ情報から切り出された文字列のコード情報をもとにキーワードが抽出できれば、あとは光ディスク装置に記録されたオリジナルのイメージ情報とキーワードを使用して、検索の手段を提供すればよいことになる。

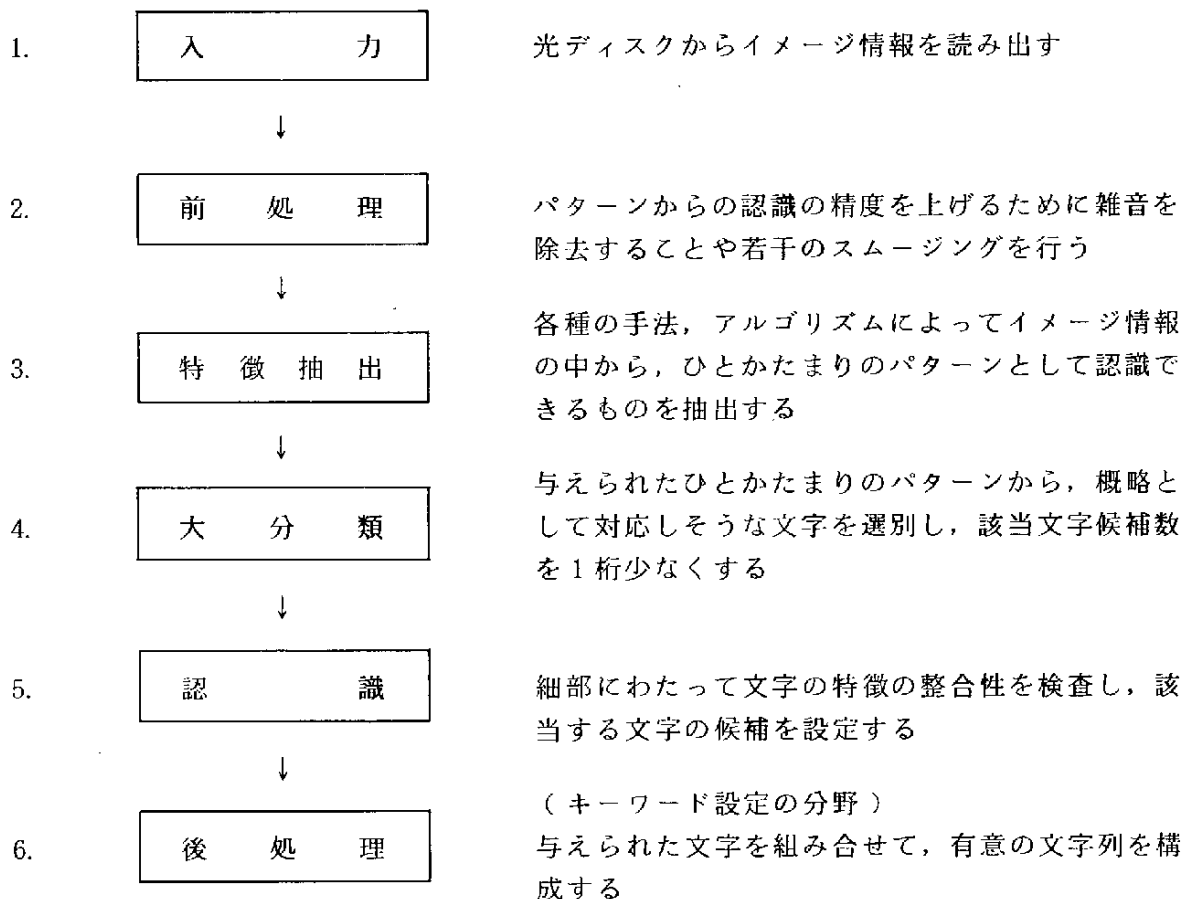
(6) 本調査研究における課題の焦点

5分野に大別した技術的課題を概観した上で、これらの課題の中から、中心的な意味を持つイメージ情報からの文字列の抽出に焦点をあてて、以降の議論を進める。

27.4 イメージ情報からの文字の抽出

(1) イメージ情報の処理の流れ

イメージ情報として光ディスクに蓄積された情報から、文字情報を切り出すためには、次のような6段階に分けて処理をする。



(2) 文字抽出における前処理

イメージ情報は白と黒との2値で表現されたものとして処理されるが、各々が鮮明に分離されていない場合や、用紙の汚れ、活字の欠け、また、ドット印字方式における中間色の取り扱い等々があり、前処理において汚れなどによる不要な情報の除去、位置・大きさの正規化などを行うことが必須となってくる。

(3) 特徴抽出

① 文字領域の切り出し

前処理の施されたイメージ情報に対して文字の切り出しをするためには、そのイメージ情報が持っている性質が重要なポイントになる。

現状では、任意のイメージ情報から文字領域を自動的に切り出すことは、不可能であると言って差し支えない。しかし、この技術は文字抽出のためには不可欠であるので、当面はマンマシンインタフェースとして、Computer Assisted (Aided) に文字領域の切り出しの機能を提供することが現実的な解となっている。

② 文字の特徴抽出

文字の識別に当たっては、イメージとして蓄積されている情報の中から、あるひとかたまりの情報を取り出しその特徴を抽出して、予め登録してある文字の特徴との比較を行うことになる。単一の書体で、均一の大きさの活字といっても、人間の目で見るときは同一に見える文字が、コンピュータを利用したパターン認識の上では異なったものと識別されるのは珍しいことではない。

いずれの形態の印刷物をソースとするにしても、解像度の点からみれば、同一の文字はほとんど存在しないといわざるをえなくなる。すなわち、活字の認識でありながら、手書き漢字の識別と同等の困難さを取り込まなければならない。

③ 大分類

概略文字としての認識がなされた情報について、最終的にコードを与えなければならない。この際、候補として考えられるものの中から、更に絞り込んだ選択が必要となる。2000～3000文字の対象から200～300程度の数に限定することが実際的である。

このための種々の手法が提案されているが、未だ研究室の研究の域を出ていない。

④ 認識

大分類の過程を経て絞り込まれた候補の中から、最適と思われる文字に対するコードを割り当てるのが認識のフェーズである。類似文字の中から、イメージとして捉らえた情報に照合して最適な文字を選別するために、パターン整合法、定点サンプリング法、単純2値マッチング法などの手法が主に採用される。

27.5 イメージ情報からの文字情報の抽出における問題点

a) 文字認識のためには、記憶されている情報そのものが漢字の複雑さに対応できるだけの解像

度で蓄積されていなければならない。

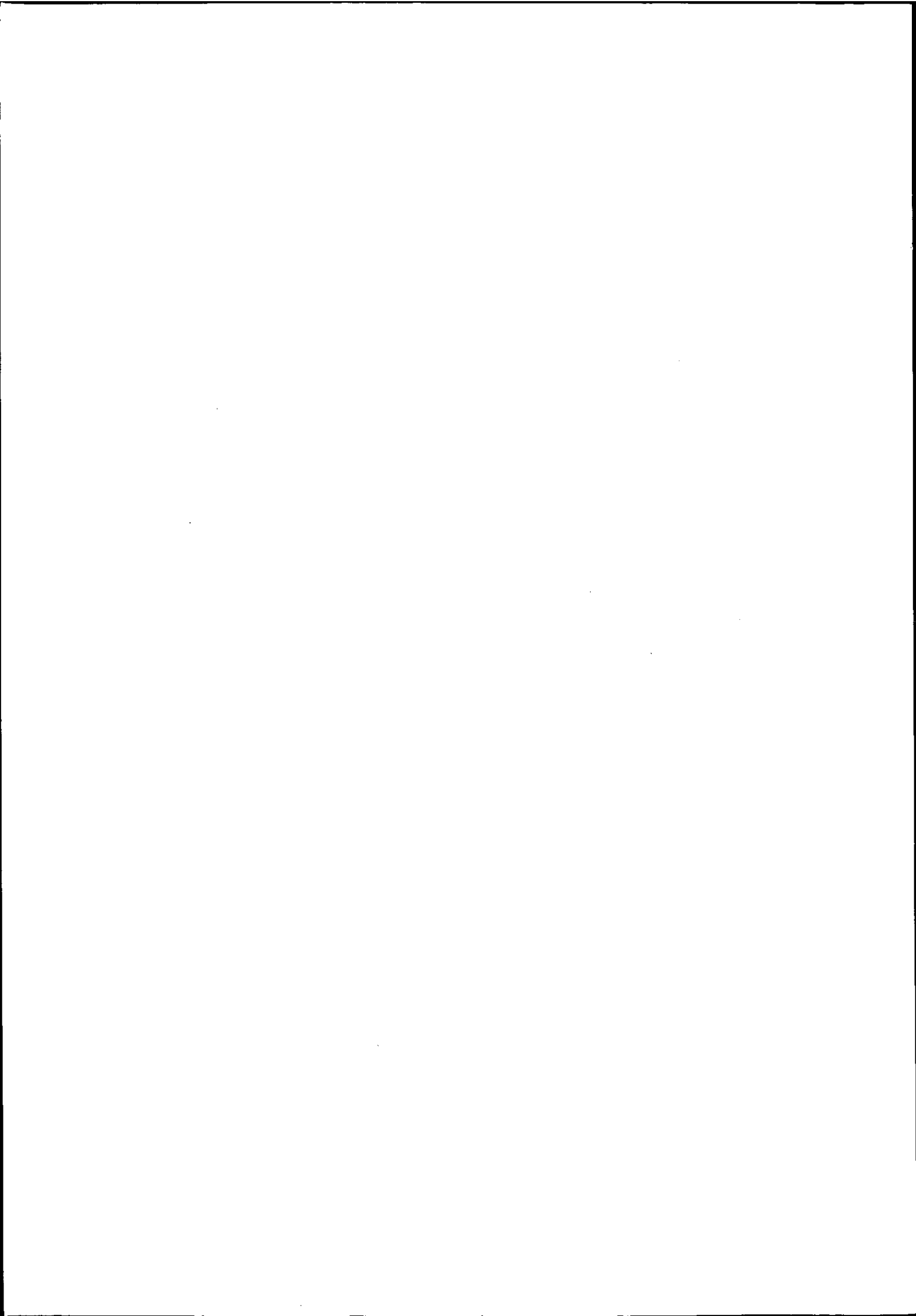
- b) 一般的なイメージ情報から文字パターンに相当する情報を切り出すためには、現在のコンピュータのパワーでは不足している。
- c) 一般的なイメージ情報が扱えるようになるためには、並列処理が要求される。このためのハードウェアを含めた開発が待たれる。
- d) コンピュータにおける文字認識の過程は、人間のそれとは全く異なっている。

27.6 ま と め

本調査研究を通じて、パターン認識技術の適用について議論することができた。書籍の情報を検索するアプリケーションにおいて、この技術の適用は検索範囲を大幅に広げることが可能になる。また、法曹界では判例の検索がデータベースによって可能になることが期待されている。更に、音声自動合成の技術との結合は、身体障害者に対する福音となりうる。その他まだまだ気付かない適用分野があると思うが、技術の確立を待って大いに適用範囲の拡大を考えることが望まれる。

28 国際ビジネスマン情報データベースの構築

財団法人 国際ビジネスコミュニケーション協会



28. 国際ビジネスマン情報データベースの構築

28.1 概要

28.1.1 構築の背景

わが国企業の急速かつ広範な国際化の進展に伴って、新しい国際化時代にふさわしい国際ビジネスマンを求める声は従来にも増して高まっている。

それに加え、最近では「海外不適応」といった問題もクローズアップされるに至り、各企業では幅広い視点に立って国際要員を選抜・育成し、海外に派遣することにきめ細かく気を配る必要に迫られている。

もちろん企業においては、国際化研修などを通じて積極的に国際要員の育成が行われているが、その前提とも言うべき国際ビジネスマンとしての要件について、それが実際にはどのようなものであるかという点になると、きわめて曖昧で、確固たるイメージやモデルは未だ出来ていないのが現状であろう。

当財団法人では、国際ビジネスマンとしての要件の一つとなっている外国語、そのなかでも国際共通語とされる英語のコミュニケーション能力については、既に8年以上、25回に及ぶ、TOEIC (Test of English for International Communication) の実施により共通尺度化を達成し、TOEICは企業、団体ならびに学校などにおいて幅広く活用されるに至っている。

しかし、国際ビジネスマンとしての要件を考える場合、語学力だけに留まらないことは言うまでもなく、もし、国際ビジネスマンとしての要件が明らかにされ、チェック方法などが確立され、それが国際要員選定や育成、派遣に反映されるとすれば、わが国企業のみならず、日本経済の進展、広くは世界経済の進展にさえ寄与するものと思われる。

28.1.2 これまでの取り組み

上記のような理由から当財団法人ではかねてより、国際ビジネスマンとしての要件とはどのようなものか、そして、それを共通尺度化することはできないかという視点から研究を重ねてきた。

1986年12月から昨1987年4月にかけて、5回にわたって開催した「国際コミュニケーション能力研究会」もその一環であるが、当該研究会では、このテーマに関連ある専門家や、

海外派遣、国際化研修に携わる企業関係者を招き、国際ビジネスマンとしての要件を知る背景、チェックすべきポイント、チェックの方法、共通尺度化の可能性、利用・活用の方法などについて、基礎的な研究を実施した。

その結果、研究会として、共通尺度化に対する一応の可能性を見出したため、国際ビジネスマンの要件を備えた人材登録データベース構築に取り組むこととなり、これを「国際ビジネスマン登録制度」= ISBIS (アイビス: International Student Businessman Identification System) と名付け、その準備活動をスタートした。

28.2 実施体制

1987年9月に至り、上記の目的を具体化するため、「国際ビジネスマン情報委員会」(通称 ISBIS 委員会) を設立した。構成メンバーは次のとおりである。

- 主 査 小林 薫氏 (産業能率大学)
- 委 員 占部 吉真氏 (日産自動車)
- 委 員 川嶋 弘尚氏 (慶応義塾大学)
- 委 員 高橋 武男氏 (キャノン)
- 委 員 西村 捷敏氏 (日本電気総合経営研修所)

また、次の方をアドバイザーとして意見を伺った。

- 稲村 博氏 (筑波大学)

なお、委員会と並行して、これまでの TOEIC 受験者の中から ISBIS 登録対象適格者の抽出を開始した。

28.3 検討内容

28.3.1 チェック項目

ISBIS 委員会は、1987年9月から12月まで月1回、計4回の会合を予定していたが、一部検討事項が残ったために、1988年1月、5回目の会合を追加開催してひとまず終了した。

その間、国際ビジネスマンの資質、適性、能力、またアセスメント方法などについて検討を重ねた結果、まず国際ビジネスマンとして登録する際に必要と思われる個人データと

して、次のような項目が挙げられた。

氏名

生年月日

年齢

性別

最終学歴

国際関連学歴

海外留学・研修・勤務・海外生活（単身を含む）の経験

海外要員トレーニング受講経験

職歴

職務上の専門分野

資格・技能・特技

逆境体験

結婚歴・離婚歴

家族

高齢者の有無

健康状態や疾病経験

スポーツ活動

文化活動・趣味

ボランティア活動

希望職種

海外勤務の可否

言語運用能力（英語＝TOEICスコア，他の言語に対する自己評価）

また，国際ビジネスマンとしての適性，能力を知るための基礎的なデータとしては，次のような項目が挙げられた。

（なお，このチェック方法を将来は国際スタンダードとしても利用することを考えているため，参考として英語も付記する。）

【精神・心（MIND AND MENTALITY）】

意欲（Will）

積極性（Positiveness）

献身性 (Devotedness)

好奇心 (Curiosity)

心情 (Emotion)

感受性 (Sensitivity)

共感性 (Sympathy)

親切心 (Kindness)

知力 (Intelligence)

理解力 (Comprehension)

判断力 (Judgement)

独創力 (Creativity)

【 態度・行動 (ATTITUDE AND BEHAVIOR) 】

人間関係 (Interpersonal Relationships)

指導性 (Leadership)

協調性 (Collaborativeness Cooperation)

社交性 (Sociability)

行動力 (Performance)

実行力 (Execution/Implementation)

調整力 (Coordination)

対応力 (Responsiveness)

管理能力 (Managerial competence)

企画立案 (Planning)

意思決定 (Decision-making)

問題解決 (Problem-solving)

【 国際適性 (INTERNATIONAL APTITUDE) 】

異文化適応力 (Cross-cultural Adaptability)

異文化理解 (Understanding of host country)

自文化理解とID (Understanding of home country and Identification)

文化差適応 (Cross-cultural adjustment)

プレゼンテーション能力 (Presentation)

自己主張 (Self-assertion)

交渉力 (Negotiation)

コミュニケーション力 (Communication skills)

国際経営能力 (International Management)

企業環境理解 (Understanding of business environment)

業務運営力 (Managerial experience & practice)

事務遂行力 (Operational experience & practice)

28.3.2 「ISBISフォーム」の作成

上記の項目は、あくまでも国際ビジネスマンとしての要件に必要と思われる項目を列記したにすぎず、これらの項目によって即座に、国際ビジネスマンとしての適性、あるいは不適性が判定できるわけではない。

アセスメントはこのシステムの中でも最も重要な分野だが、数多くのデータを蓄積し、そのデータを基にして統計的、あるいは時系列的に分析する過程において、国際ビジネスマンとしての要件をクローズアップするプロセスをとることになる。

もちろん、企業担当者や専門家による分析などを必要とすることは言うまでもないが、いずれにせよその前提となるのは、同じチェック項目によるデータ蓄積である。異なったチェック項目では、データ間の比較は不可能であるし、また、チェックのたびに項目が動くようでは、統計的、時系列的追跡は不可能となる。

したがって、既に示した個人データおよび基礎データを基にデータベースを構築するためには、まず一定の様式を持ったチェック・フォームが必要となる。

また、この統一フォームを各企業や団体、学校などが利用するようになれば、データベースは更に充実したものとなることは明らかであるとの観点から、ISBIS委員会では、統一フォームたりうるチェック・フォームを作成し、これを「ISBISフォーム」と名づけた。

28.3.3 「ISBISフォーム」から質問表へのワーディング

「ISBISフォーム」は基本的に、個人データについては該当番号の選択と自由記入、基礎データについては5点法によるマークシート方式をとっている。そのチェックは、本人が行う場合もあれば、企業や団体、学校によっては、上司あるいは教師などの第三者が

行う場合も考えられる。

第三者チェックの場合にはそれほど問題はないが、本人チェックの場合は、特に基礎データについて、恣意性や中央効果（5点法の場合には回答が中央の3に寄ってしまう傾向）が出るのが考えられるため、基礎データの各項目を三つの要素に分け、更にそれぞれを相反する要素と対比してスコアをつける方法を取り入れることとした。

つまり、そうして得た三つの要素のスコアを平均し、その平均値を各項目に再びコンバートしてスコアとすることにより、できるだけ恣意性や中央効果を排除する方法をとることとした。

そのため、基礎データの各項目を三つの要素に分ける作業を行い、更にそれぞれを相反するワーディングに落とし込む作業を行って、質問表を完成した。

28.4 パイロットスタディ

28.4.1 パイロットスタディ実施の背景

当初の計画では、「ISBISフォーム」および質問表の完成に伴い、一般ビジネスマンおよびTOEIC受験者のうちの適格者を対象としてISBIS登録を呼びかける活動を開始する予定であった。が、先のISBIS委員会において、チェック項目の妥当性や、質問表による回答と「ISBISフォーム」の相関などについて、事前にサンプル調査を行って検証する必要があるとの答申があったため、質問表の制作、印刷をはじめとするパイロットスタディの全容を計画、実施することとなった。

パイロットスタディのサンプルは、当財団法人に関係する法人（TOEICユーザ企業、団体など）を中心に、現在および将来の国際要員となる多様な母集団を対象とすることとしたが、一方において、国際要員の現状を多面的に分析するために、過去のTOEICデータを用い、企業の国際ビジネスマンの属性別英語力の測定を行った。

また、パイロットスタディによって得られるサンプルをまずデータベース化するため、「ISBISフォーム」の入出力、更改プログラムの設計、製作をスタートした。

28.4.2 パイロットスタディの実施と分析

パイロットスタディでは、上記対象者から1000余名のサンプルを得た。その回答を、28.3.3で述べたように「ISBISフォーム」へコンバートし、各種属性分析を行った結

果、チェック項目の妥当性、質問表と「ISBISフォーム」との相関が検証された。

同時に、このパイロットスタディを参考に「ISBISフォーム」の入出力、更改プログラムの設計、製作を完成させ、パイロットスタディの1000余名のサンプルがデータベース化された。

その結果、国際ビジネスマンとしての資質、適性、能力について、具体的なイメージを想定しうるマトリックスが抽出された。

またパイロットスタディによって、「ISBISフォーム」を構成する二つの側面、つまり「国際要員としての資質、適性」（性格や性向に左右される面が大きいもの）と「国際要員としての能力」（トレーニングなどによって獲得しうるもの）との比較も行われ、その差異と相関についての計数的な把握もなされた。これにより、企業が求める「国際要員養成システム」への連動の手掛かりが得られた。

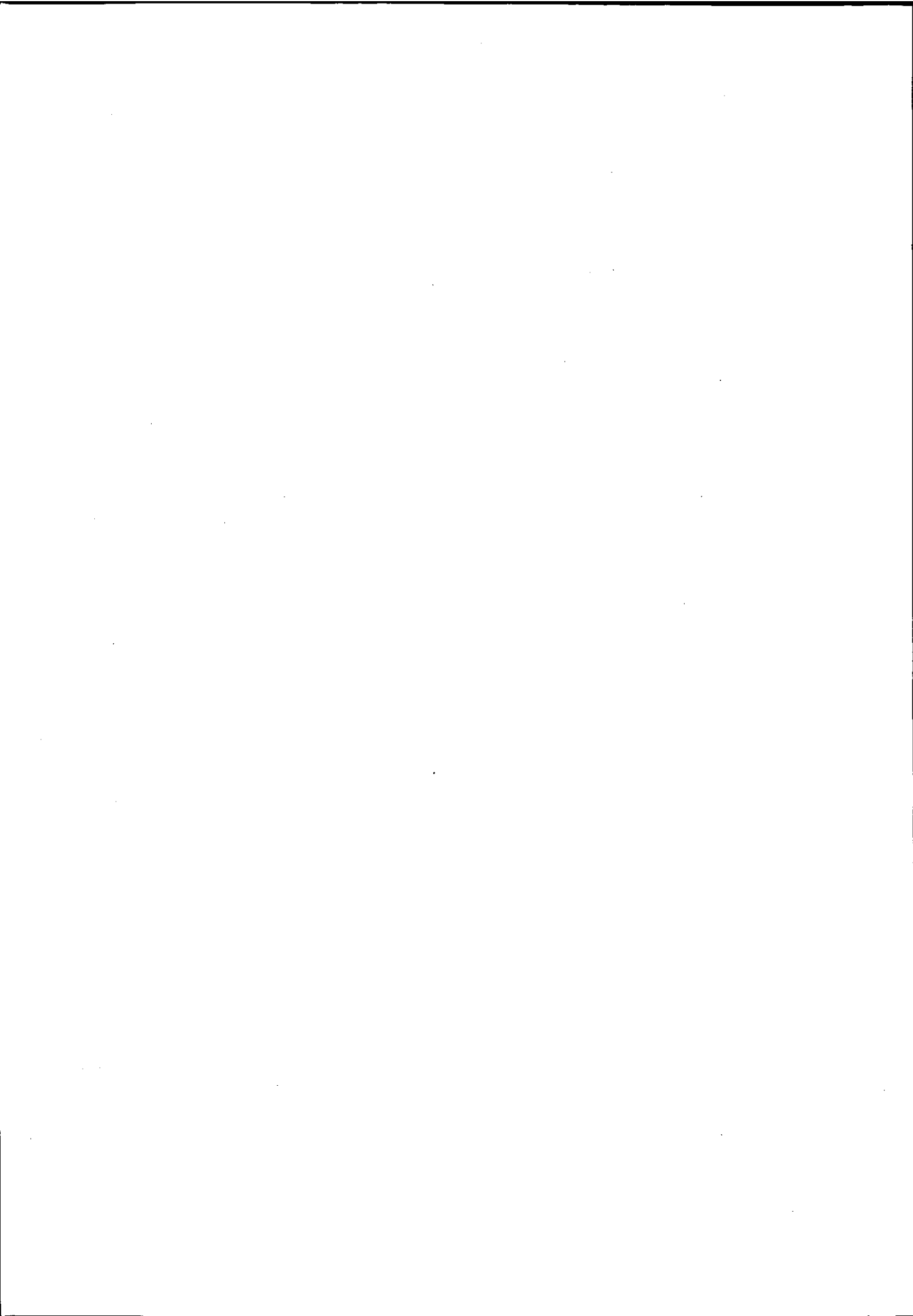
このように見えてくると、このデータベースは、これまでまったく手掛かりのなかった、あったとしてもきわめて曖昧だった国際ビジネスマンのイメージやモデルを確立する近道になることは確信しうる。

28.5 今後の課題

「国際ビジネスマン登録制度」= ISBISを事業として展開するに当たっては、更にリサーチが必要になってこよう。専門的分析や追跡調査、時系列的な分析に加え、国際ビジネスマンとしての要件を備えた人材と、それを求める企業側との需要供給のリサーチも欠かせない側面である。

また、ISBISのデータベースをどのような形で公開していくか、統一のフォーマットとして「ISBISフォーム」をどう浸透させていくか、企業や団体、学校などが「ISBISフォーム」を利用した場合、そのデータをどうISBISのデータベースに取り込んでいくかなどの課題もある。

しかし、前にも述べたように、これまでは手探りであった国際ビジネスマン、国際要員の要件を考えるうえで、このデータベースがきわめて大きな力を持つことと確信する。



29 建設関連技術用語データベース構築の調査研究

株式会社 建設情報センター



29. 建設関連技術用語データベース構築の調査研究

29.1 目的

用語に対する認識は、個人個人によりまちまちである。つまり、同じ構造を持った用語であっても、人・時代・場所・分野などによって受け取る概念が違っていることがある。

しかし、情報の流通の場においては、このような状況は望ましくなく、用語に対する共通の概念が要求される。

そのため、一般に用語に対し適切な定義解説を用意することが行われている。

情報流通の重要性が益々増大していく今後の社会においては、この用語の使われ方解説などについてはユーザに適宜提供され、かつ改定されるべきである。

今回、建設分野の技術用語を対象に用語の概念、認識を代替する各種利用事例などを調査すると共に、具体的に選定した分野の用語について利用事例等をリンクした建設関連用語データベースのプロトタイプを構築することとする。

プロトタイプ化されたデータベースをもとに逐次用語の追加、利用事例の追加を行うことにより将来的にはデータベース社会の基盤データベースに発展するものと予想される。

一方、建設産業は社会にとって基盤を形成する産業であり、またここへの投資はその波及効果が大きいことが知られている。更にこの建設産業をささえる建設技術はそのカバーする範囲が広いことも特徴である。

このような産業で使用されている技術用語のデータベースを構築することは、他産業にも大きな影響を与え得る。

29.2 実施内容と方法

29.2.1 調査の内容

今回実施した調査およびプロトタイプ構築の主な内容は下記のとおりである。

(1) 技術用語データベースの必要機能の洗い出し、利用方法の検討

このレベルでは、単に建設関連技術用語にこだわることなく、技術用語データベースとして汎用的に持たなくてはならない必要機能をピックアップし、それぞれの機能に対し様々な視点から検討を行った。

また、技術用語データベースの利用形態についても検討を加えた。

(2) 建設関連技術用語の現状調査と継続的な収集方法

建設関連技術用語の現状調査を実施した。

また、用語データベースを考慮する場合、用語そのものの収集方法が、継続的に行えるものでないと、時代の変化についていけないデータベースとなってしまう。どのような原始データから用語を収集するかは、用語データベースの性格を決定する重要な要因であるので、この継続的な収集方法について考察検討を行った。

(3) 建設関連技術用語データベース構築コスト調査およびニーズの調査

データベースの構築コストと、データベースのニーズは構築計画を立てる時、重視すべき要因であり、これを建設関連技術用語データベースに絞って検討を実施した。

(4) 建設関連技術用語データベースの提供媒体、提供方法の検討

データベースは、最終的にユーザに提供されないと、その価値がなくなる。

上記各項目の調査・検討結果を踏まえ、提供媒体、提供方法の検討を行った。

(5) 建設関連技術用語データベースの構築計画

上記各検討結果を踏まえ、建設関連技術用語データベースの構築計画を立案した。

また、収集した建設関連技術用語の一部についてはデータベースのプロトタイプ化を計画し、課題の抽出を行った。

29.2.2 調査の実施体制

調査研究に当たっては、(株)建設情報センターでプロジェクトチームを編成し、外部の関係する機関および学職経験者の協力を得て実施した。

29.2.3 実施期間

1987年7月から、1988年3月まで。

29.2.4 予想される波及効果

- (1) データベース整備の基盤整備であり、電子化辞書などに展開できる。
- (2) シソーラスに展開可能である。
- (3) 建設関連産業の情報流通がやりやすくなる。
- (4) 他の分野の技術用語データベース構築の参考になる。

29.3 用語集および技術用語集の現状分析のまとめ

実例	必要機能 度数	見出し語表記				見出し語副表記				並べ方			分類		関係				解説			索引						使用頻度・出現頻度	出典・用例									
		かな漢字(英字・カナ混り)	ひらがな・カタカナ	ローマ字	英語	フリガナ(キーワード)	ふりがな(読み方)	かな漢字(英字・カナ混り)	対訳			全用語を(読みの)五十音順	分類別に五十音順	アルファベット順	階層分類(コード)	並列分類(略記・記号)	普通名詞・人名・地名	同意語		階層	関連後		解説		総索引(全用語)五十音順	規格別(分類別)	英和(対訳)索引			独語	仏語	露語	略語一覧(英字)	補図				
									英語	独語	仏語							露語	USE(AND)(を見よ)		UF	BT	NT	対語・反対語											短文	スコープノット	限定句	
広辞苑		○				○										○				○	○														○	出典 用例		
現代用語の基礎知識	○					○						分類別	○			列記						○																
マグローヒル 科学技術用語大辞典	○								○					○		○																						
JIS工業用語大辞典	○					○		○	○				○			○											○	○										
岩波 理化学辞典	○							○	○	○	○					○																						
JICST 科学技術用語ソーラス	○					○							○			○	○	○	○	○			○	○														
建設省 建設技術用語ソーラス	○					○							○			○	○	○	○																			
土木学会監修 土木用語辞典			○					○	○	○			○			○							○				○	○	○									
彰国社 建築大辞典		○						○	○					○		○																						出典
CORNET 文献キーワード集	○					○								○		○	○																					
度数(計)		7	2	1		3	2	3	5	3	3	1	6	2	1	4	3	1	9	3	3	2	7	2	7	1	3	2	2	5	1	1	1	2	6	1	2	
コメント		有意の語を最初に表記する例が多い				ソーラス(キーワード集)ではフリガナが重要 対訳表記では英語が半数				読みの50音順がほとんど			分類で用語の範囲を制限		同義語(USE)は不可欠 階層で用語の範囲を制限				解説がなく分類毎に分けていない場合用語を制限する限定句が必要			見出し語表記が分類別の時全用語索引が必要 英和対訳索引 半数にあり						ブック形式では図の 使用多い										

29.4 成果と今後の課題

29.4.1 主要成果

- (1) 技術用語類の国内の現状と建設分野における技術用語集・シソーラスの現状について概略が明らかになった。
- (2) (社)土木学会, (社)日本建築学会, (財)日本建設情報総合センター等の建設関連機関における建設技術用語・シソーラス開発の計画予定が調査できた。
- (3) 技術用語データベースの必要機能の洗い出しと, データの継続的な収集方法について調査できた。
- (4) 技術用語データベースのプロトタイプ化を試行し, データベース構築方法・提供・媒体等の検討すべき課題を得た。
- (5) 上記をふまえ, 今後建設関連技術用語データベース構築と利用方法の確立・方策についての示唆を得ることができた。

29.4.2 今後の検討課題

技術は絶え間なく進歩し, 新しい技術と共に発生する用語は, 生き物のごとく日々新たに増える。継続的に新しい用語を採集し, 定義を下しメンテナンスを行わなければ, データベースの価値はなくなる。

技術用語データベース・メンテナンスのための機能は, 次のようなものである。

(a) 新規用語集のルーチン化

データベースのメンテナンスは, データの収集をいかにスムーズに行うかにかかっている。新しい用語が発生の都度, どんどん収集できるような仕組みが必要となる。

具体的には, 新規用語の発生源との共同作業などがあげられる。

イ. 学協会など公的機関の用語委員会とのリンク

ロ. データベース・プロデューサとのリンク

ハ. 用語集, 辞書類出版作業とのリンク

(b) 用語の定義・判断

言葉はどんどん発生し, 利用されるので, 明確な定義, 関連などが後追いになる場合すらある。処理過程でも, 言葉の利用をガイドすることも必要となる。

イ. 暫定的な定義, 位置付けの明確化(分からない場合の相談窓口)

ロ. 定義の確定

(c) 定期的な処理

メンテナンスされたものは、その都度利用に供されなければならない。

(d) 技術用語データベース利用の普及、指導

用語を自由に、正しく使用するためには、道具だて（ハードウェア、ソフトウェア）の整備だけでなく、どう使いこなすかのヒューマンウェアの技術が重要になる。

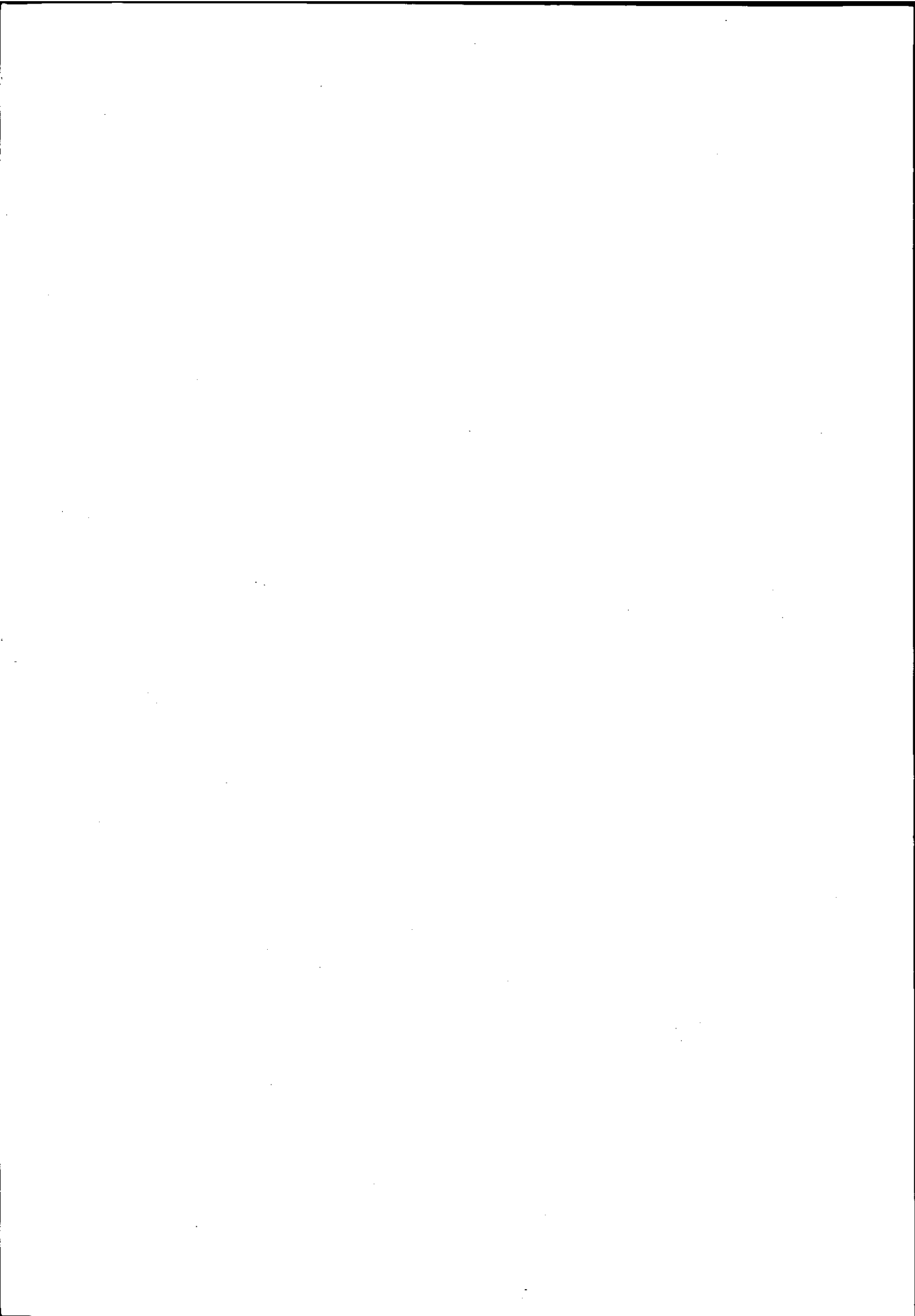
これら収集から普及・指導に至るまでを継続的に管理するためには

(a) 定常業務として権限をもって行う強力な事務局が必要

(b) 専門識者からなる常設の委員会などの組織化が不可欠

となる。

データベースは、初期の立ち上がりはもちろん、その後の運用しだいで成果に大きな差が出てくる。(社)情報科学技術協会の国際十進分類や情報処理にたいする取り組みのごとく、管理体制をいかに整備するかが初期構築よりむしろ重要課題といえる。



禁 無 断 転 載

昭 和 63 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
TEL 03-459-8581

印刷所 株式会社 正文社
〒113 文京区本郷3丁目38番14号
TEL 815-7271~3

