# 情報処理技術の応用に関する 調査研究報告書

一人工知能技術を応用したデータベース利用技術に関する調査—

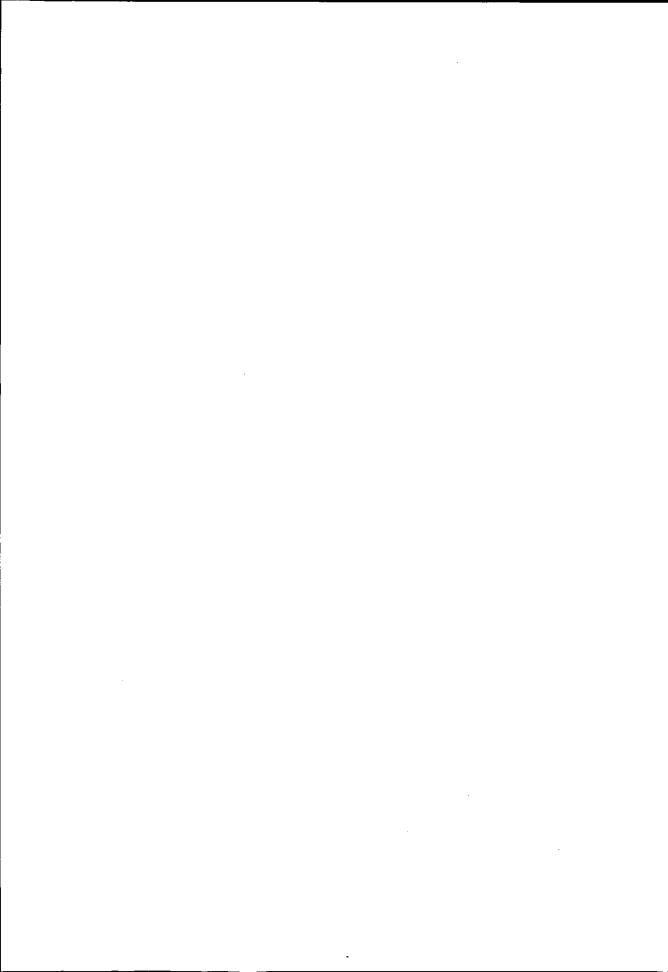
平成元年 3 月

# JIPDEC

財団法人 日本情報処理開発協会



この報告書は、日本自転車振興会から競輪収益の一部である機械工業振興資金の補助を受けて昭和63年度に実施した「情報処理技術の応用に関する調査研究」の成果をとりまとめたものであります。



高度情報化社会という言葉が実感を伴って受け入れられる時代となり、特に経済活動に おいては、有益な情報を入手し、いかに活用するか・・・ ということが課題となってい る。また、このようなニーズを持つ人々の層が拡大することによって、データベースの利 用者の層が拡張し、データベースや、統計に詳しくないユーザの利用が増加してきた。

現在、経済情報のデータベース化は進んでおり、情報そのものの提供の条件は整いつつ あるが、「データベースや、統計の専門家でない利用者」に情報を提供する為の環境は未 整備な状態である。この環境を整備するために、情報の選択や、活用に関するノウハウが 付加された使いやすいデータベースが求められている。

当協会では、このような背景に鑑み、新しいソフトウェア技術のひとつである人工知能 技術を使って、データベースに上述の利用支援機能を持たせるための調査研究を3年計画 で実施することとした。

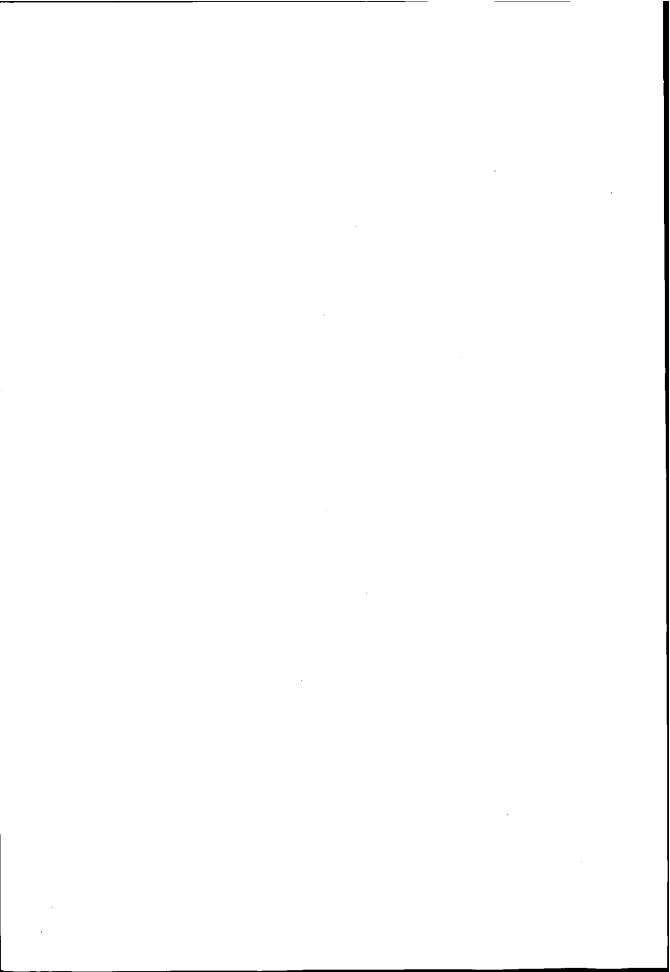
昭和63年度は、実際の経済統計データベース利用の現状調査、人工知能技術の現状調査、海外におけるデータベース検索の技術動向調査を中心に、利用支援のための知識、利用支援システムを実現するための技術に関する検討を行い、実用システムのための目標設定を行うことを活動内容とした。

平成元年度は、この成果をうけて、知識の整備及び実用システムに必要なマンマシン・ インターフェースに関する検討を重点的におこない、実験システムを構築する。

平成2年度は、実験システムの評価をもとに、実用システムの機能を再検討し、設計・ 開発に着手することを目標とする。

本報告書は、この3年計画の1年目に当たる調査研究の成果をとりまとめたものであり 今後の調査研究のための基礎資料として、利用活用されることはもちろんであるが、関心 のある方々に読まれ、データベースを利用した情報活用の一助になれば幸いである。

尚、本調査は株式会社CSK総合研究所に委託して実施した。



# 「人工知能技術を応用したデータベース利用技術調査」委員名簿

(順不同、敬称略)

# ( 本委員会 )

委員長 渕 一博 (財) 新世代コンピュータ技術開発機構

委 員 鵝野 公郎 筑波大学 社会工学系

委 員 石川 徹也 図書館情報大学

委 員 黒川 恒雄 工学院大学

委 員 佐藤 英人 大阪大学社会経済研究所

委 員 武藤 博道 (社) 日本経済研究センター

委 員 桐山 正敏 通商産業省 大臣官房 情報管理課

委 員 近藤 正幸 通商産業省 大臣官房 情報管理課

委 員 藤森 聿子 通商産業省 通商産業研究所 政策情報システム部

委 員 矢田 光治 (株) CSK総合研究所

# ( 専門委員会 )

委員長 鵜野 公郎 筑波大学 社会工学系

(兼本委員会)

委 員 石川 徹也 図書館情報大学

(兼本委員会)

委 員 田丸 征克 経済企画庁

委員長田博 アジア経済研究所

委 員 坂本 英陽 アジア経済研究所

委 員 増田耕太郎 日本貿易振興会

委 員 篠原 昇 (株)アイエヌ情報センター

委 員 佐藤 安夫 (財) 日本エネルギー経済研究所 エネルギー計量分析センター

委員 中島 淳元 (株)QUICK

委 員 仲田 雄作 通商産業省 大臣官房調査統計部 統計解析課

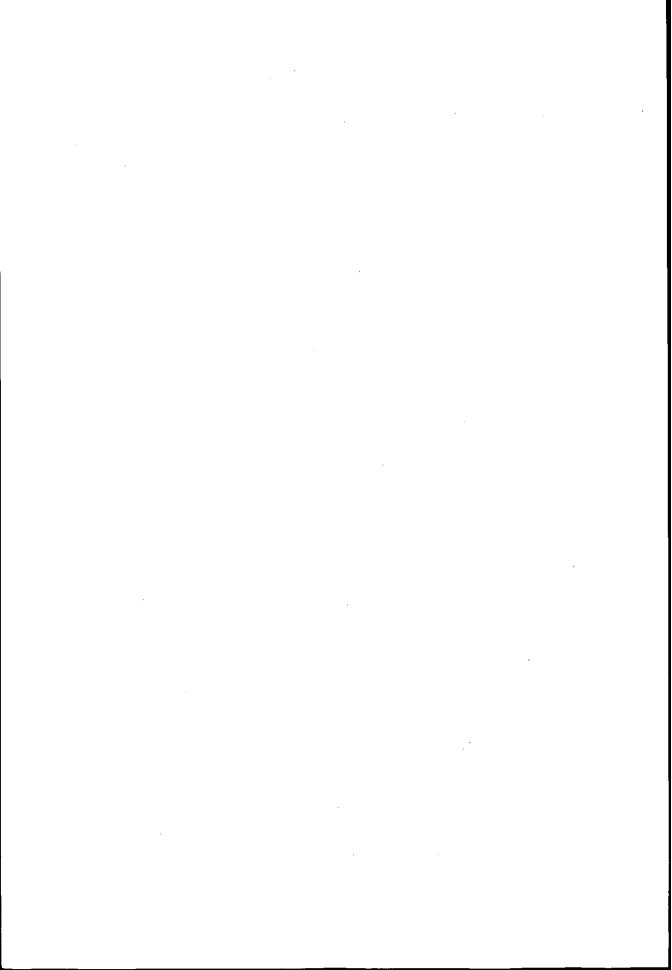
委 員 林 幹浩 通商産業省 産業政策局 調査課

委 員 藤森 聿子 通商産業省 通商産業研究所 政策情報システム部 (兼本委員会)

委 員 内田 和義 通商産業省 通商産業研究所 政策情報システム部

委 員 水島 豊 通商産業省 通商産業研究所 政策情報システム部

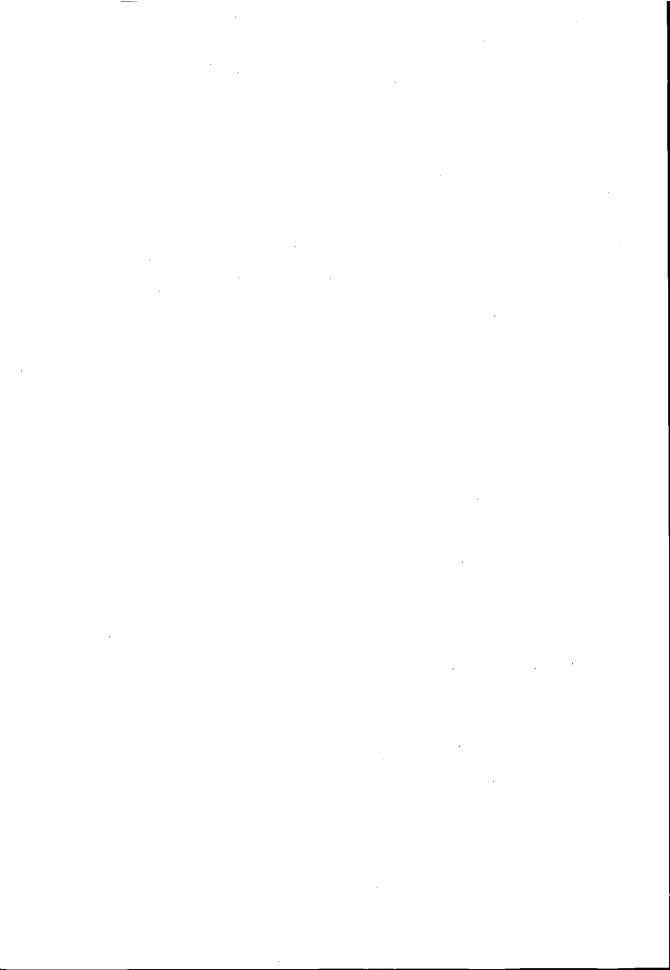
委 員 杉原井康男 通商産業省 通商産業研究所 政策情報システム部

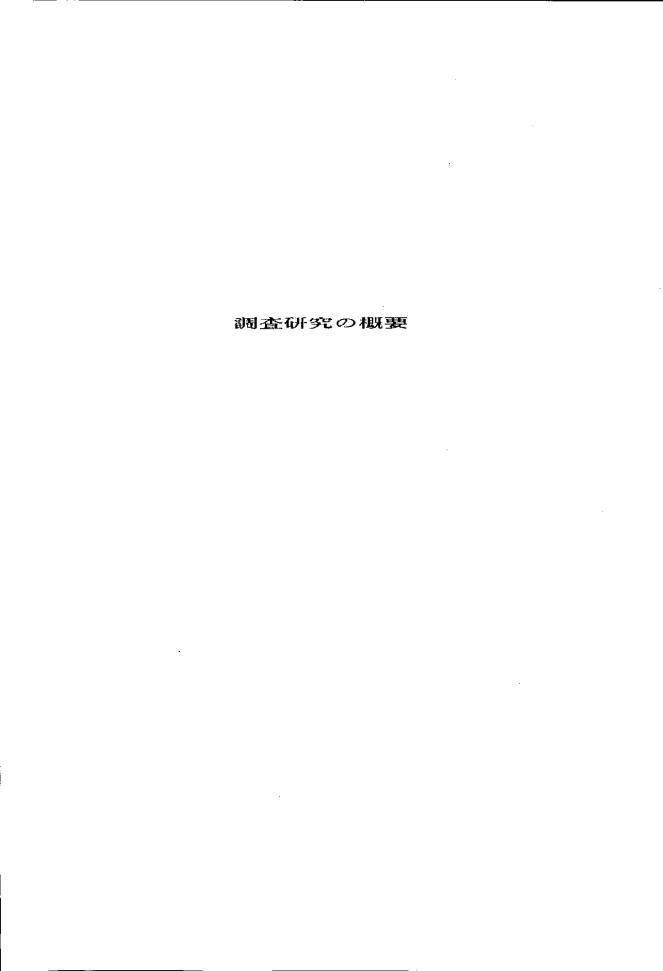


調査研究の概要	
調査の目的 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	• 3
実施計画 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	• 3
実施体制 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	• 3
本年度調査研究の内容・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	• 4
第1章 データベース利用の現状と人工知能技術	
1. 1 データベース利用の現状と課題 ・・・・・・・・・・・・	• 7
1.1.1 データベース利用の現状 ・・・・・・・・・・・・	. 7
1.1.2 経済統計データベース利用技術の高度化 ・・・・・・・	• 10
1. 2 人工知能技術の特徴とデータベース利用技術の高度化 ・・・・	• 13
1.2.1 人工知能技術の特徴と利用分野 ・・・・・・・・・・・	• 13
1.2.2 人工知能技術を応用したデータベースの高度利用 ・・・・	• 19
第2章 経済関連データベースの構築・利用における課題	
2. 1「PPIS」の概要と課題 · · · · · · · · · · · · · · · · · · ·	27
2.1.1「PPIS」の概要 · · · · · · · · · · · · · · · · · · ·	• 27
2.1.2 「PPIS」の利用における課題 ・・・・・・・・・・・・	• 34
2.2 「INDB」の概要と課題 ····································	• 35
2.2.1「INDB」の概要 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	• 35
2.2.2「INDB」の課題 · · · · · · · · · · · · · · · · · · ·	• 38
2.3 「 AID・XT」の概要と課題 · · · · · · · · · · · · · · · · · · ·	
2.3.1「 AID・XT」の概要 ・・・・・・・・・・・・・・・	
2.3.2「 AID・XT」の課題 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	
2.3.3 人工知能技術の利用可能性 ・・・・・・・・・・・・・	
2.4 ジェトロのデータベースシステムの概要と課題 ・・・・・・・	
2.4.1 ジェトロのデータベースシステムの概要 ・・・・・・・	
2.4.2 ジェトロのデータベースの内容とシステムの特長 ・・・・	

2.4.	3 ジェトロの貿易統計処理システムの利用状況 ・・・・・・ 71
2.4.	4 ジェトロにおける貿易統計データベースの工夫 ・・・・・ 74
2.5 ED	MC エネルギーデータバンクの概要と課題 ・・・・・・・ 78
2.5.	1 EDMCエネルギーデータバンクの概要 ・・・・・・・ 78
2.5.	2 データ利用方法からみた課題 ・・・・・・・ 86
2.5.	3 今後の課題 ・・・・・・・・・・・・ 87
第3章 デー	ータベース検索への人工知能技術利用
3. 1 3	文献情報データベースと人工知能技術 ・・・・・・・・・ 91
3.1.	1 文献情報検索への人工知能技術利用の動向 ・・・・・・ 91
3.1.	<b>2</b> データベース検索への人工知能技術利用事例 ・・・・・・ 92
3. 2	データベース検索と自然言語処理技術 ・・・・・・・・ 95
3.2.	1 自然言語処理インターフェイス「DB/I」の概要 ・・・・ 95
3.2.	2 自然言語処理技術と統計データベースの関連について ・・・100
3.3 #	統計データベースの検索と人工知能 ・・・・・・・・・106
3.3.	1 経済統計DB検索利用のための知識ベース化 ・・・・・・106
3.3.	2 知識ベース化の利点 ・・・・・・・・・・・・・ 111
第4章 海绵	外研究動向(統計情報とAI)
4. 1 i	毎外動向調査の概要 ・・・・・・・・・・・・・・・115
4.1.	.1 会議の構成 ・・・・・・・・・・・・・・・・116
4. 2 *	経済統計データベースの構築・利用の動向 ・・・・・・・ 119
4.2.	.1 統計的コンサルテーション-誰がエキスパートか? ・・・・ 119
4.2.	.2 既存の統計ソフトウェアに知識ベース・モジュールを付加する
	手段 ・・・・・・・・・・・・・・123
4.2.	<b>.3 統計的問題に対する協力的解法</b> ・・・・・・・・・・ 125
4.2.	.4 データ分析のためのコンサルテーション・システム
	NONPARE · · · · · · · · · · · · · · · · · 127
4.2.	.5 時系列分析のための統計エキスパート・システム
	ESTES

	4.2.6	データ分析のためのインタアクティブ・ツール
		${\tt KnowledgeSeeker} \qquad \cdot $
	4.2.7	知識獲得の方法としての計量心理学的手法 ・・・・・・ 135
	4.2.8	エキスパート・システムを利用者指向にするための統計利用 ・ 135
	4.2.9	エキスパート・システムMUSEに於ける統計手法表示
		統計的手法の選択 ・・・・・・・・・・・・ 136
	4.2.10	MULREGの構造、設計方法、及び戦略・・・・・・・ 136
	4.2.11	統計的知識エンハンス・システム・・・・・・・・・ 139
	4.2.12	統計データベースに統計を教える・・・・・・・・ 141
4.	3 海外	の於ける動向とわが国の研究へのインプリケーション ・・・143
第5章	· 人工知	能技術を応用した経済統計データベース利用支援システムの構築
5.	1 経済	「統計データベースの利用支援システムに関する提案 ・・・・ 147
		経済統計データベース利用の現状 ・・・・・・・・147
	5.1.2	知識ベースの構築におけるケーススタディ ・・・・・・150
	5.1.3	システム開発へのアプローチ ・・・・・・・・・ 153
5.	2 シス	テムの概要 ・・・・・・・・・・・・・・・ 156
	5.2.1	システムの特徴 ・・・・・・・・・・・・・ 156
	5.2.2	エンドユーザから見た機能概要(外部仕様) ・・・・・・159
	5.2.3	システム機能構成 ・・・・・・・・・・・・・・ 162
5.	3 シス	テムに必要な知識 ・・・・・・・・・・・・・ 169
		検索支援に必要な知識 ・・・・・・・・・・・169
		知識ベースの具体例 ・・・・・・・・・・・・・ 172
5.	4 シス	テムス実現のための方向性・・・・・・・・・・・173
		知識ベース獲得の方法 ・・・・・・・・・・・・・173
		初期試作システムの稼働環境(案) ・・・・・・・・174
		システム構築への課題 ・・・・・・・・・・・175
付録		
<del>张龙</del>	<del>√あ</del> ♪	





## 調査研究の概要

#### 【調査の目的】

近年における情報化の進展に伴い、データベースの構築が急速に進められており、その対象分野の拡大、利用可能地域の広域化と共に、利用者も専門分野を中心とした層から一般化している。

しかし、利用者がデータベースを真に活用するためには、データベースの所在、その構造、収録レコード/キーワード体系、検索コマンド体系等のデータベースそのものに関する知識と共に、収録されているデータの意味・内容や利用上の制約事項、分析のノウハウ等のデータ利用に関する知識が不可欠である。これらの知識は、データの分野によって異なっているため、その広範囲な習得は困難であり、データベースの利用普及を妨げる大きな要因となっている。この解決には、分野ごとに必要な知識を整理し、データベースの利用者が、収録されているデータと共に活用できる環境を整備する必要がある。

このため、本事業では、経済統計に関するデータベースの利用に焦点を絞り、データベースに関する知識、データベース利用に関する知識等を調査し、人工知能技術を応用して知識ベース化の検討を行い、利用者とデータベースの効率的なインターフェースの向上を図ることを目的として調査研究を行う。

#### 【実施計画】

昭和63年度は、データベースの構造、収録コード/キーワード体系、検索コマンド等のデータベースに関する知識・ノウハウと、統計に関する利用上の制約、分析等、データ利用に関する知識・ノウハウ等について海外の事例を含めた調査分析と共に、これらの知識のデータベース化について設計を行った。

平成元年度は、今年度の調査結果に基づき、知識情報の整備を行うと共にインターフェース機能の概念設計及びシステム設計を行い、実験システムの構築を行う。

平成2年度は、前年度までの調査結果に基づき、インターフェース機能の充実を図り、 実用システムとしての適用可能性について評価検討を行う。

#### 【実施体制】

「データベース利用技術調査本委員会」を設置し、基本的方針及び実施計画等を検討し

「同専門委員会」において詳細な調査及び検討を行った。

## 【本年度翻査研究の内容】

本年度は以下のような項目について調査研究を行った。

- ①経済統計に関するデータベースの課題
- ②構築するシステムの内容に関する検討
- ③海外における事例調査

但し、調査研究にあたっては、具体的な事例を取り上げて行った。

以下に内容を示す。

①経済統計に関するデータベースの課題

経済統計データベースのあり方について、経済統計データベースの構築側、利用者側からの多角的な検討を行った。

②構築するシステムの内容に関する検討

使い易い経済情報データベース及びデータ活用システムにもたせるべき機能、システム が供給するべき知識についての調査研究を行った。

## ③海外における事例調査

海外のシステム例、システムの運用形態及びその他のデータベースへの人工知能技術の 応用に関する研究開発についての調査を実施した。この調査では、文献調査及び海外のフィールドサーベイを並行して実施した。

# 第1章 データベース利用の現状と 人工知能技術

			ï

# 第1章 データベース利用の現状と人工知能技術

本章では、データベース利用の高度化と人工知能技術の関係を捉えるため、データベースの利用の現状と人工知能技術のデータベースへの適用について概説する。

尚、データベース利用及び人工知能技術の現状に関する各論については、第2章、第3章で述べることにする。

#### 1. 1 データベース利用の現状と課題

本節では、特に我が国のデータベース構築の経緯とその利用の現状と問題点について述べる。

#### 1.1.1 データベース利用の現状

# (1) データベース構築と利用の経緯

我が国のデータベースの構築とその利用は、1950年代より、文献情報を中心として開始 され、現在では、統計情報(数値情報)の利用もかなり一般化してきている。

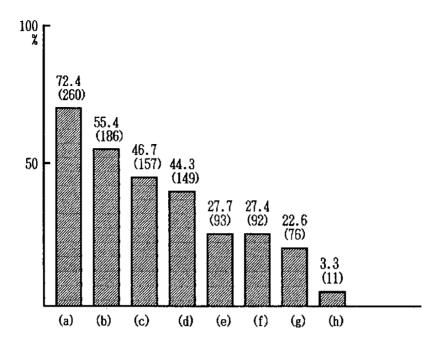
しかし、さまざまなデータベースが構築されるにつれ、その利用における問題点が多く 発生してきていることも事実である。

このような情報システム化におけるコンピュータの利用に関しては、業務処理をコンピュータ化する側面と、情報利用をコンピュータ化する側面の2つがあるといえるが、業務で得られたデータの利用体制を整えていくことが重要である。

こうした業務処理と利用面のインテグレーションを今後推進してゆくためには、データ ベースの構築とその利用技術の改善が重要となってきている。

#### (2) データベース利用における課題

図1-1は、データベースの利用上の問題点について、特に機能・操作性に関する指摘について示したものだが、この図からわかるように、指摘が一番多かったのは、「シソーラスやキーワードの標準化や拡充不足」が全体の77.4%、また「検索機能が不十分」についても55.4%と高い比率を示している。



- (a) シソーラスやキーワードの標準化や拡充の不足
- (b)検索機能が不十分
- (c)検索速度が遅い
- (d)接続手順が煩わしい
- (e)ファンクションキーが不十分
- (f)交換機能(カナ漢字等) が不足
- (g)キーボード配列が悪い
- (h) その他

図1-1 機能・操作性に関する指摘(N=336,複数回答)

出典:DPC「データベース・サービスに関するユーザ意識調査」昭和63年3月

また、データベースの利用上の問題点としては、「検索用のコマンドがシステムごとに 異なり、統一されていないために利用しにくい」といった点について、以下のような調査 結果が出ている。

- ・マニュアルを見て不自由なから検索している・・・・66.1%(255社) ・不自由は感じない・・・・・・・・・・・・・・15.4%(57社) ・検索する気がしない・・・・・・・・・・・・・・・10.8%(40社) ・その他・・・・・・・・・・・・・・・・・・・・7.6%(28社)
  - 出典:同上

この調査からわかるように、「データベースを検索する気がしない」と「マニュアルを見て不自由ながら検索している」を合わせると、76.9%のユーザがコマンドの不統一に対して何らかの不満を持っているといえる。一方、「不自由は感じられない」が15%程度の比率を示しているが、これは、この問題が検索の専門家にとってはそれ程大きな問題に

なっていないためである。しかし、この比率は、ユーザ層が広がることによって、低くなる可能性を内包している。

本調査は一般的な調査ではないがユーザから見た利用上の問題点が反映されていると思われる。

#### (3) データベース利用の改善

このような利用上の不満が生じてきた背景として、以下のような理由が上げられる。

- ①データベース構築においては、利用者側に立ったシステム作りが難しい
- ②利用者が必要とする情報の詳細性のレベルと利用者のデータベースに関する知識の レベルに差がある

①への現在の対応は、人海戦術的なものであり、「どのようなデータが必要で、どのようなものをデータベース化すればよいか」に関する検討はされても、利用者の常識のレベル、目的の違いなどの付随情報に関しては未検討な点が多い。

また、②についても、特に経済統計データベースに見られるように、利用者の「常識」 は千差万別であり、初心者の利用に耐えられるようなデータベースの構築が難しいのが現 状である。

このような現状を打開する一つの方法として、データベースの利用者をサポートするシステムの構築が考えられる。即ち、利用者の不足している「知識」をシステム側で補うことによって、データベースの利用価値を高めようとする考え方である。この「知識」のシステム化についての技術的な実現方法として、今日、様々な分野で利用が進められている人工知能技術の適用が考えられる。特に、知識ベース(Knowledge Base)による「専門的知識」のシステム化を可能にするエキスパートシステム(Expert System)の適用が期待される。

以上のような点を踏まえながら、本調査研究では、「データベース利用技術の高度化」 に関して、特に「経済統計データベースの利用」に的を絞ったケーススタディを実施し、 「人工知能技術」の応用について検討を行った。尚、ケーススタディの結果と人工知能技 術の適用等の詳細については第5章で述べる。

# 1.1.2 経済統計データベース利用技術の高度化

データベース利用の高度化について、本調査研究では、特に「経済統計データベース」 の利用の高度化についての検討を行った。ここで、経済統計データベースを取り上げた理 由としては、以下の 2 点が上げられる。

- ①経済統計データベースの利用は、他のデータベースの利用と比較し、マクロ的な利用が多く、利用者のニーズも多様であることから、適切な情報検索を行うためには 広範な専門的知識が必要とされている。
- ②経済統計及び経済情報に関するデータベースの利用が拡大され、その利用者層が多様化することによって、検索したいデータの所在、出力されるデータの意味等に関する知識のサポートが重要となってきている。

本節では、この経済情報に関するデータベースの課題について言及する。

(1) 経済統計データベースの課題

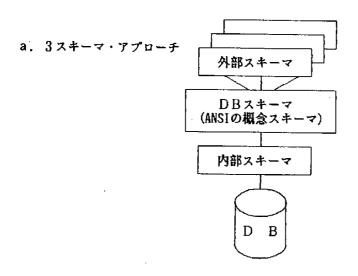
経済統計データベースの特徴として

- ①利用可能な統計データの案内については若干の印刷物があるものの極めて不備であり類似統計、関連統計、相互の相違などの比較は専ら利用者の経験に依存していること
- ②統計データに関する書誌情報や定義情報などのいわゆるメタデータが、データベース 化の際に、統計データ自体と分離されている為に、データベースユーザのデータの評 価等に困難が生ずること
- ③統計データに不可欠な時系列データとクロスセクション・データを総合的に管理する 必要があること
- ④種々の統計を横断的及び学際的に利用するニーズが高いこと等が挙げられる

このように経済統計データベースの利用には、統計及びデータベースに関する専門的な 知識が必要であり、習熟に時間がかかる。これらの特徴ふまえ、蓄積された統計情報を活 用促進するためには、以上の知識を整理し、専門知識のない利用者を支援する機能を構築 することが課題である。

# (2)データベース利用支援システム構築の方策

既に多くの経済統計データベースが利用に供されているが、この分野におけるデータベース形成では、一般のデータベース理論の単純な適用では不十分であることが明らかにされている。以下の図1-2にデータベース・スキーマの比較を示す。



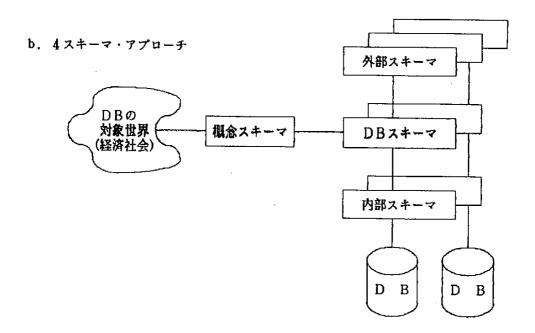


図1-2 データベース・スキーマの比較

出典:国土庁・調査局編「国土情報の知識管理」

図1-2のaが、通常のデータベースのスキーマである。この場合のデータベースの設計・作成は、次のような過程で行われる。

- ①データに対する「ユーザの要求仕様」の調査 (「外部スキーマ」の決定)。
- ②外部スキーマに対応しうるデータベースの「理論的構成」を検討する (「DBスキーマーの決定)。
- ③データの物理編成(データの物理的位置、データ圧縮法、データへのアクセス経路など)を考える(「内部スキーマ」の決定)。
- ④実際にDBスキーマに基づいてデータを収集し、内部スキーマに従ってデータベース 化する。

このような 3 スキーマ・アプローチによる管理方式に対して、最近、図1 - 2 の b のような 4 スキーマ・アプローチが提案されている。この提案では、データベースに関する論理的記述としての「DBスキーマ」とは別途に、対象世界を記述するモデルを構想し、これを「概念スキーマ」と称している。これを統計に即していえば、現実のデータ入手可能性とは別に、対象世界において概念上考えられることができる統計データをモデル化したものである。現実の統計調査は対象世界である経済社会をそのまま捉えたのではなく、特定の視点から一部を切り取っているに過ぎない。そのため現実の統計調査から出発するのではなく、まず対象である経済社会そのもののモデル化から始めて全体を体系的に捉えることが必要となる。

このようなデータベース理論における提案は、統計データベース利用者の立場からも支持されるものである。つまり、概念スキーマが統計データとデータベース利用者とをつなぐ架橋となるからである。このようなコンセプトが定着し、統計の具体的な体系化と情報整備が行われることが必要である。

# 1. 2 人工知能技術の特徴とデータベース利用技術の高度化

本節では、技術革新の著しい人工知能技術の側面から、データベース利用の高度化についての適用性について検討する。尚、人工知能技術のデータベースへの利用については、第3章で事例を取上げ説明する。また、経済統計データベースへの具体的な応用に関しては第5章で詳細に述べる。

# 1.2.1 人工知能技術の特徴と利用分野

# (1) 人工知能技術への期待

人工知能(Artificial Intelligence)技術の普及は、最近かなり急速に進められているが、「人工知能」の概念自体の研究は既に20年以上も前から行われている。その実用化を可能にしたのは、コンピュータのソフトウエアにおけるAI構築用ツールやAI用言語の発達もあるが、ハードウエアの発達によるところがより大きく起因している。

特に最近では、32bit パーソナルコンピュータの登場により、スピード、メモリーの制 約が大きく改善され、かつ身近なシステム構築が可能になったといえよう(図1-3)。

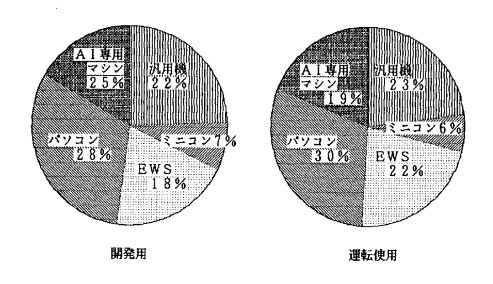


図1-3 人工知能システム使用ハードウエアの分類(1988年の各マシンの比率)

出典:「人工知能の技術と需要」(財)日本情報処理開発協会

こうした人工知能技術を取り巻く環境の変化の中で、特に実用化が期待され、実際に実 用システムが稼働している分野として、エキスパートシステムが上げられる。このエキス パートシステムの特徴は、専門的知識のシステム化を可能にした点にあり、この知識の格 納方法として、「知識ベース」が存在する。

エキスパートシステムは、当初は医療診断、機械の故障診断等の分野で多くの研究開発 が行われて成果を上げ、最近では、あらゆる分野でその利用が検討されてきている。この 背景には、エキスパートシステムへの以下のような期待が内包されている。

- ①業務の専門的知識の保存、管理が可能になる。
- ②業務の初心者でも、ある程度のレベルの仕事が可能になる。
- ③専門家自身が従来業務から開放されることで、より知的な作業が可能になる。
- ④専門的知識の支援を多くの素人が利用できる。

次にこうした期待に対して、技術的な実現を可能にするエキスパートシステムの構築の 基本となる知識ベースについて簡単に説明する。

#### (2) 知識ベースの特徴

知識ベースとは、その開発言語が Lisp, Prolog等の従来言語とは違った言語によって構築されているところに一つの特徴がある。知識構築するときにまず、事実と規則という 2 つの概念が必要である。事実とは、文字どおり事実のことであり、事物の性質及び特徴である。規則とは、推論するときの必要な判断基準である。事実はフレームの形で、規則はプロダクションルールという形で表現し知識ベースは構築される。

フレームについて「犬」と「猫」を例に挙げ説明する。一般に、動物は「足を持つ」かつ「鳴く」ので、その性格として「足を持つ」かつ「鳴く」ものだと定義できる。そして「犬」の概念は、動物に属しなおかつその特徴として「ワンワン」と鳴くと捉えられる。また「猫」の概念は、同様にしてその特徴として「ニャーニャー」と鳴くと捉えられる。ここで犬の「ハチ公」は「駅で待つ」習慣を持ち、猫の「マイケル」は「踊る」習慣があるとすると、これをフレーム表現で表せば、図1-4のようになる。

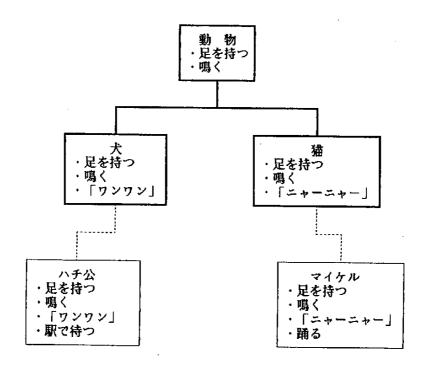


図1-4 フレーム表現

プロダクションルールとは、「IF (条件) THEN (結果) DO (アクション) 」といった形式で表現される規則である。専門家の知識は、このような規則の集合として表現され、このルールに基づいて推論が行われる。

# (3) エキスパートシステムの構成

エキスパートシステムは、人間の専門家の持っている知識をデータとして蓄え、これを 使い推論及び問題解決を行う。エキスパートシステムの基本構成は、図1-5のようなも のである。先に述べた専門家の知識は、知識ベースとして蓄積される。推論エンジンは、 断片的に蓄えられている知識を基に「だから~である」といった結論にたどり着く仕組み になっている。また、エキスパートシステムの扱う知識は、専門家の行う経験的な判断を 何らかの形でコンピュータに理解できる形にした知識であり、作り始める時点ではまだ言 葉で表現できないような未整理の状態の場合もありうる。

高度なエキスパートシステムになると、結論を出すまでの推論過程を示す解説サブシステムや専門家と対話しながら知識ベースの拡充を行う知識獲得サブシステム等を備えている。インタフェースとは、専門家が知識ベースを作成するためのエディターやユーザーのための簡易入力等のことである。

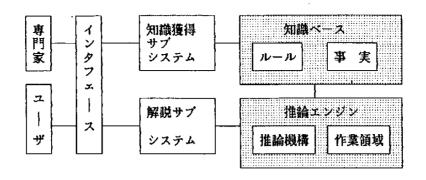


図1-5 エキスパートシステムの構成

#### (4) その他の人工知能技術

以上、人工知能の基本技術である知識ベース及びエキスパートシステムについて簡単に 説明したが、人工知能技術にはこのほかにも自然言語処理、パターン認識、知能ロボット 等の要素技術が多く存在している。その中でも、自然言語処理技術の一つの応用として自 然言語による情報検索システムの開発は、データベースの利用技術の面からも期待されて いる。(図1-6参照)

この自然言語による検索システムでは、日常的な言葉によって、データベースとインタラクティブに会話することによって、必要な情報を検索するシステムである。そのメリットは、専門的なコマンド等の知識がなくても、データベースにアクセスできること、自然言語というインタフェースによって各種のアプリケーションが利用できることなどが考えられている。

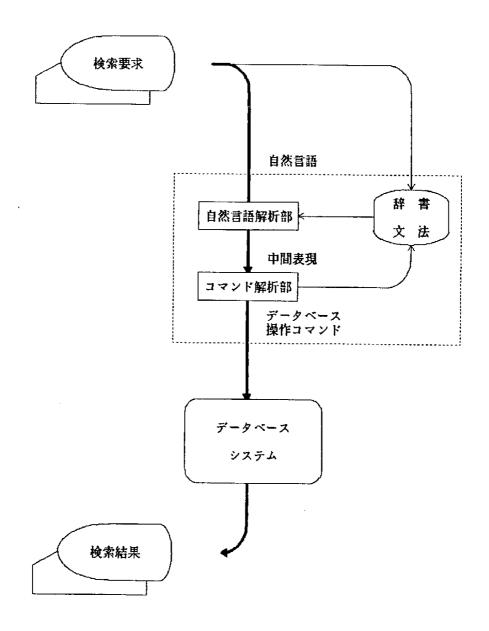


図1-6 自然言語インタフェースのシステム構成

人工知能技術の応用分野は多岐にわたっているが、最近の技術動向として、現在、導入 ・稼働中の各AIシステムについての比較を図1-7に示す。

また、今後の実用化の時期についての予測については、図1-8に示すように、エキスパートシステム、自然言語理解、マンマシン・インタフェイス等が早い時期に実用に供されるものと期待される。

尚、自然言語理解を利用したデータベースの検索事例については、第3章で詳しく述べる。

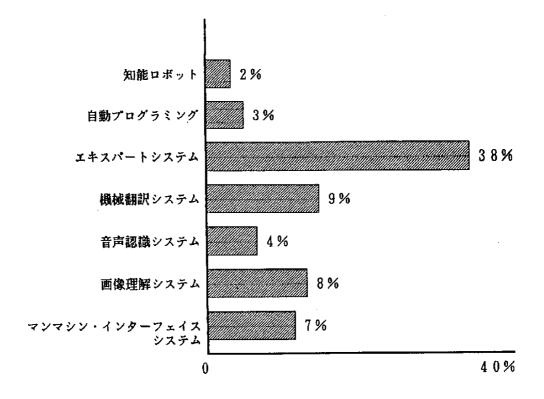
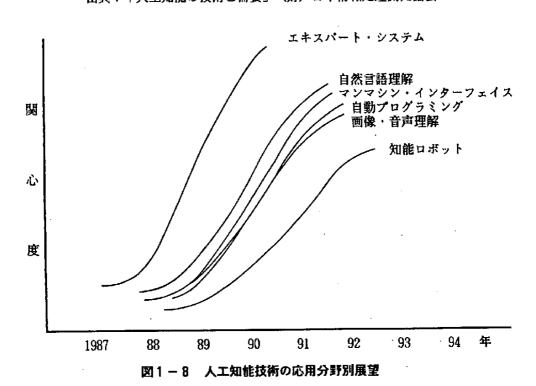


図1-7 1987年時点の導入のAIシステム 出典:「人工知能の技術と需要」(財)日本情報処理開発協会



# 1.2.2 人工知能技術を応用したデータベースの高度利用

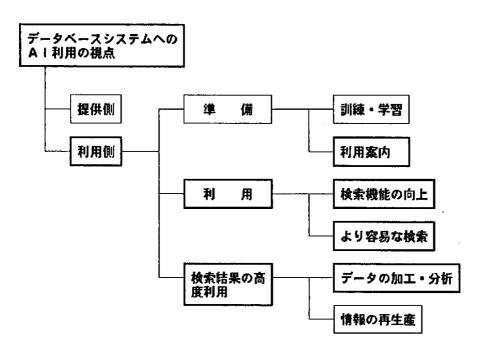
# (1) データベースへの人工知能技術の適用範囲

既に述べたように、人工知能技術の特徴とその応用分野の最近の動向からも分かるように、人工知能技術はあらゆるシステム分野に利用され始めている。その中でも、最近、特にデータベースと人工知能の関係についての調査研究及び開発が進んでいる。

そこで、本節の終わりでは、人工知能技術が、データベースの中でどのように利用されるかについての検討を行う。

本調査研究の目的は、データベースの「利用技術」の高度化にあり、誰にでも使いやすいデータベースシステムの構築のために、どのような開発方法が可能であるかについて、特に「経済統計データベース」を対象に、また、技術的には、人工知能技術の応用から接近を試みることにある。

そこで、データベース利用において、人工知能技術はどのような適用範囲をもっているのかを明確にしておきたい。図1-9は、データベース利用における人工知能技術利用の視点である。この図は、データベースの利用に関してシステム化が望まれている項目をまとめたものである。



注)太枠:本調査での検討項目

図1-9 データベースへの人工知能技術利用の視点

データベースに携わる人には、データを提供する「提供者」とそのデータを利用する「 利用者」が考えられるが、本調査研究ではデータベース利用の高度化を考えており「提供 者」側はその範疇から除いている。データベースの利用は3段階に分けることができる。

検索に先立ち、その「準備」があり、次にデータベースの「利用」、さらにその「検索 結果の高度利用」を考えることができる。

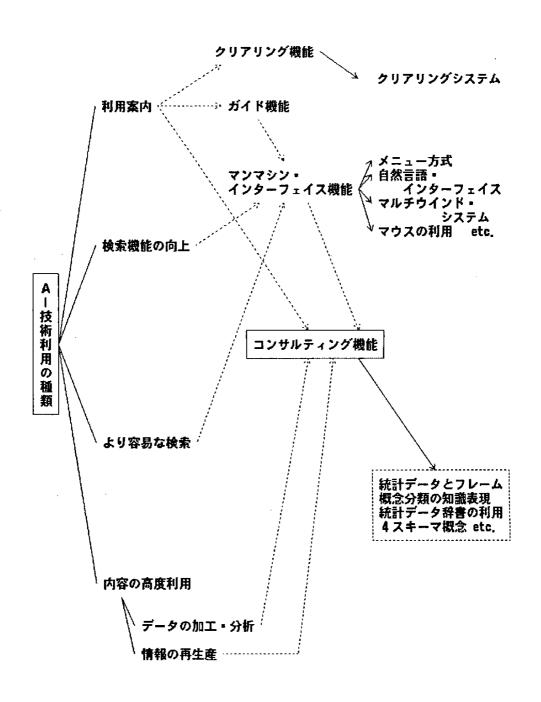
「準備」にはデータベースの利用に先立っての「訓練・準備」とその「利用案内」が含まれるが、「訓練・準備」は本研究の対象とはしない。「利用案内」は、データベースにある具体的データについて「これこれこういうものがある」などと教えるもので、人工知能技術の活用は有効と考えられる。

データベースの「利用」では、より効率的な検索のための「検索機能の向上」と自然言語などを用いた「より容易な検索」に人工知能技術の活用が有効と考えられる。

さらに「検索結果の髙度利用」では、「データの加工・分析」や「情報の再生産」を考えることができる。「データの加工・分析」はある限られた専門知識があれば人工知能技術の活用が有効と考えられる。「情報の再生産」は、結果の評価が定まらないため知識ペース化が困難と思われる。

従って、図1-9の太枠で示された部分を本調査研究での検討項目とする。データベースの利用における人工知能技術の応用については、このような視点からのアプローチを行う。

次に、これらの適用項目に対して、期待される機能とそれに対応する技術的項目について整理し、データベースへの人工知能利用についてまとめたものを図1-10に示す。



← 利用目的レベル → 期待される機能レベル → 具体的 A ! 技術 → 統計データベースの 新しい概念の適用 レベル

図1-10 データベースへの人工知能技術利用の関連図

このデータベース利用における人工知能技術分野からの接近では、知識ベースを利用した検索機能の向上に視点が置かれている。また、内容の高度利用についてもより利用者の立場にたったシステムとして「コンサルティング機能」のレベルアップを図ることが目標として考えられる。

さらにデータベースと知識ベースの親和性からのアプローチも重要であり、「概念スキーマ」を知識ベース化して構築することも可能性としては充分に検討する価値がある。 以上の点に関する詳細なケーススタディの結果は、第5章で述べる。

(2) データベース利用における人工知能技術への期待

経済情報に関するデータベースの利用の高度化にあたり、人工知能技術にどのような期 待があるのかについて検討した事項を以下に示す。

- ○レポート作成過程で過去に使用されたノウハウを知識ベースとして活用したい。
- ○データベースのデータは、数値の羅列であり、その意味を読み取ることは難しいことから、本にあるような注釈的なものがデータベースにも欲しい。
- ○データの分析のためのエコノミストとしての常識(基礎知識)がデータベースに 入っていると便利である。
- ○データベースを利用する際、前任者の使用例が利用情報として参照できるようなシステムがあると良い。
- ○利用者が、システムからのコンサルテーションを受けながら利用できるデータベースがあると良い。

また、これらの期待に対して人工知能技術を利用して実現するための**留意**点としては、 以下のような点が上げられた。

- ○人工知能技術を利用して構築されたシステムを「誰が」利用するかを、明確にしておくことが必要である。つまり、ユーザのレベルを専門家にするのか、初心者にするのかをはっきりさせることが重要である。
- ○システム開発を進める上での優先順位をはっきりさせることが必要であり、その 方法としては、パイロットシステムを短期的に構築し、ユーザのニーズを必ずフィードバックすることが重要である。
- ○経済統計データベースの利用では、複数のデータベースの利用や海外のデータの 利用等が多いことからネットワークを介して行われているが、このネットワーク

実現の為には通信機器間での通信手順やメッセージ内容等の取決めである通信規 約(プロトコル)の国際的標準化やコード体系全体の国際的標準化についての検 討が必要である。

第2章 経済関連データベースの構築・利用 における課題

_	 			
				!
			•	
				;
				1
		•		
				I
				I
		·		
		,		

### 第2章 経済関連データベースの構築・利用における課題

本章では、経済関連データベースの構築及び利用上の課題について、各経済関連データベースの構築及び利用の立場からのその概要と課題について報告を行う。

### 2. 1 「PPIS」の概要と課題

本節では、通商産業省における経済統計データベースである「政策情報データベース (PPIS:Policy Planning Information System)の概要と課題について、特に、利用における問題点を中心に報告する。

#### 2.1.1 「PPIS」の概要

政策情報システムとは、政策立案に必要な情報を提供して、行政のプランニング機能の 高度化を支援することを目的とする情報システムである。

この情報システムは、わが国の企業情報、貿易情報をはじめとし、各国別、産業別、地域別等の経済、貿易、エネルギー情報等の共用データベース群を有しており、ユーザはこれらのデータを用いて、検索、製表、加工、分析、グラフ出力等の作業を行うことができる。また、ユーザは自ら収集したユーザ独自のファイルを持ち同様の作業を行うことができる。

次に、このPPISの中のデータベースシステムの利用事例として以下の3つの事例に ついてユーザガイド、クリアリングシステムの機能を中心に説明する。

- ① 国内マクロ経済情報 (SDB)
- ② 日本貿易情報(TRADE)
- ③ 大規模企業情報 (KIMS)

#### ①国内マクロ経済情報 (SDB)

省内のデータベースの中で最も古いもので、30~40の統計より時系列データを収録 している。このためデータソースとしてはかなり複雑であるが、構造的にはシンプルなデータベースシステムである。

データの内容としては、マクロ経済データ、産業・企業活動データ、物価関係データ、

労働関係データ、国際収支統計データ等約30,000系列が蓄積されている。

図2-1~2-4は、このSDBを、ガイダンス機能を用いて使用した場合の検索状況を示している。

図2-1は、クリアリングシステムに相当し、政策情報システムの入口である。また、 図2-2に示されるようにSDBの場合、検索コードは4区分に分かれており、コードが 不明の場合は「\*」を入力することによってシステム側がサポートしてくれる。図2-3 は、部門とデータの出所から選択する場合を示しており、図2-4は種類の区分(月次、 年次、オリジナル、指数等)の選択を表している。

このように、SDBを「ユーザガイド」と呼ばれる機能を利用して検索した場合には、 システム側がかなりの部分をサポートしており、この機能は、直接コマンドによって検索 できない人のためのサポートシステムと言える。

図2-5は、この「ユーザガイド」を使用せずに、直接コマンドを入力した場合の例である。コマンド入力による検索はこのように短いものであるが、このような検索を行うためには、コードブックが必要であり、さらにそれを正確に理解する知識とコマンドを覚えることが必要となるため、使い慣れない人にとっては、ガイド機能による検索の方が便利である。

\*\* データベース・ガイド \*\*

PPIS

商用データベースの紹介

表示は?

分野番号は? 01 (企業財務、企業活動、企業プロフィール) 01 貿易 (我が国通関統計、DOT統計、世界貿易統計) 02 03 (マクロ経済統計、商品市況、金融為替) 経済 (地域基礎統計、工業統計、商業統計、県民経済計算、立地動向) 04 地域 05 (生產動態統計、大気汚染、鉄鋼産業) 06 資源 (IEAエネルギー統計、鉱物資源統計) (特許出願件数、研究活動動向、技術文献) 07 技術 (通産六法、国会議事録、法令) 08 法令 09 (マスコミ、出版、人物) 10 共用分析システム (EAST、MATE、PCNV、他) 他省庁、関係機関の紹介 11 商用データベースの紹介 (国内) 12

図2-1 SDBの検索例

1:続きを表示する

9:終了(開始画面へ)

2: 先頭から再表示する

(海外)

 PPIS \*\* 国内マクロ経済情報(検索コード) \*\*

 検索コード及び処理内容を指定して下さい

 検索コードは?

 統計コード
 \*

 分類コード
 種類コード

 極類コード
 1

 処理は?
 1

 1:検索を行う
 9:処理選択画面へ戻る

(注記)

SDB収録データを検索する為には検索コード(統計、分類、項目及び種類コードで構成)が必要です。既に御存知の方は入力を、不明の方は"\*"をコード欄に入力して下さい。検索コードのガイドを行います。

処理、検索コード、入力後に実行(送信)キーを押して下さい

# 図2-2 SDBの検索例

PPI	S **	国内マク	口経済	情報(	検索	コード)	**		
統計コ		: 統計名							l
分類コ	- F ( )	:					<del></del>		ſ
項目コ	ード ( )	:	<u> </u>						
種類コ	ード ( )	:	検索	コード	のガ・	イドを行	行います	。まず、	目
• • • •		•	的とす	るデー	タが、	、どの	流計に属	するかけ	展択
	<del></del>	1					ご御存知		
選択は	? (3						る統計の		
	Ĺ	J .	又は所	管の番	号を	異択し	て下さい	放出す	Ż
			統計を	表示し	まず			0 12 7	
番号	部	F		番号			lf	管	
1:	(国民経済	生産所得		$\bar{3}$ $\check{0}$		(総務)			
2:	(国民経済	国富	)	3 1		(経済)		,	
2 : 3 :	(生産指数等		)	3 2		(大蔵			
4:	(消費・家語	十二家計碼	査 )	$\bar{3}$ $\bar{3}$		(厚生)		情報部	
5:	(物価 卸		_ )	3 4		(通産		解析課	
			•				pun()	THE PERSON NAMED IN	
55:続	きを表示する	3		8 8	:前	断面へ関	<b>戻る</b>		
6 6 : 先	頭から再表え	示する		99			面へ戻	る	
表示して	いる部門、戸	F管に は続	きかあ	ります			50	_	

図2-3 SDBの検索例

	C 4.4.	国内マクロ	<b>以次科共</b> 制	/揺揺っニ	ድ <del>ፈ</del> ረ ዩነ	**
PPI					F # 1 F)	T T
	ード (B5) : ード (A1) :		<b>広工業指数</b>	<b>X</b>	)	,
項目コ種類コ	- F (0010) - F (	;	<b>拡工業</b>	,	)	)
選択は	? 1	下記		ディドを行いら目的と	います する種類の?	番号を選択
番号 種 1:( 2:( 3:( 4:(	MOI) ()		)(季節	市調整 市調整済 ※一夕	名目・1 )(指数・2 )(指数・2 )( )(	名目)
	きを表示す。 頭から再表		8 8 9 9	3:前画面 9:処理選	へ戻る 択画面へ戻	3 /

# 図2-4 SDBの検索例

	★ SD	Bシステム START ★
	・鉱 ・卸 ・消	ー タ 更 新 状 況 江工業指数 8804 85キンキジュン デス。 P売物価指数 8805 85キンキジュン デス。 付費者物価指数 8804 FNP(連報) 87 . 10-12
	問い	<b>合わせは内線 2 2 7 7</b>
	ETR CDP EQ B 件 見つかり	
コマンド ハ ?	DISP CDP TP(	(8601-8711) DT
コードバターン	ジテン	<del>ÿ−9</del>
B5A10010MAI	8601 8602	121.90 121.90

図2-5 SDBの検索例

### ②日本貿易情報システム (TRADE)

この日本貿易情報システムは、通関統計データの検索、加工、出力を行うシステムである。データはシンプルな構造であるが、品目分類等が細かく分かれている。

図2-6、図2-7に、このTRADEを使用した場合の検索例を示す。この場合も、 前述のSDBと同じようにユーザガイドを利用したものである。

PPIS *	* 日本貿易情報	最(検索条件の指定 1)	**
品目コードは?			
·品目1,品目2,·品目1-品目2,	···,品目N	;複数指定(カンマ 又はフ;連続範囲指定(ハイフ	ランクで選択区切る) フ ンで結ぶ)
国コードは?		***************************************	
·国1, 国2, ··· ·国1-国2,	·, 国N	;複数指定(カンマ 又はフ;連続範囲指定(ハイ:	」 ランクで選択区切る) フ ンで結ぶ)
		·方の指定欄の任意の位記 :方のコードから順にガ	

図2-6 TRADEの検索例

PPIS	** 日本貿易情報(検索	紫条件の指定 2) **
時点、	期間及び処理内容を選択し	<b>して下さい。</b>
時点区分は?	3 時点輸 1:会計年度 2:歷年 3:年月	随用は? 8707 ~ 8712 19(80 ) FY 19(86 ) FY 19(81 ) CY 19(87 ) CY 19(8501) 19(8712)
		(データ収録時点)
処理は?	1	
	1:処理を行う 8:前画面(検索 9:終了(処理》	常条件の指定1) へ戻る 選択画面へ戻る)

図2-7 TRADEの検索例

### ③大規模企業情報 (KIMS)

大規模企業情報は、全国8証券取引所上場企業及び店頭銘柄企業(但し、金融・証券保険業を除く)2,700 社の財務データ等を収録しており、項目(資本金、従業員等)の数が多いのが特徴である。財務データは、1社につき200 ~400 の項目数になるため、常によく使用される項目についてのガイドが作成されている。

図2-8は検索結果である57企業名の表示例であり、図2-9はその出力例である。 図2-10は同じ検索をコマンドで直接行った場合の例である。

PIS	* *	企業情報シ	ステム(検索条件の表示画	面) **
選択さ 企業	れた企業 集団の中	は以下のと から削除し	おりです。 企業 ( たい企業があれば、" N "	57)件 を入力して下さい。
). >	<企業	(コード>	<企 業 :	名>
]	(00	264)	(味の素(株)	)
	(00	3 4 6)	(東レ (株)	)
	(00	351)	(旭化成工業(株)	)
	(00	465)	(住友化学工業(株)	)
	(00	469)	(三菱化成(株)	)
【は?	1			r
				件の企業
	3: 先即	から再表示	する	
<u></u>	5:企業	と一覧をプリ		法の選択画面へ
	選 / 2 / 3 / 4 / 5 / 5 / 5 / 5 / 5 / 5 / 5 / 5 / 5	選択ま 全 本 を な ま ま を は で の の の の の の の の の の の の の	選択された企業は以下のとし 企業集団の中から削除> (00264) (00346) (00351) (00465) (00465) (00469) 理は? 1 1:データを表示再表に大 3:先頭かする。 4:ソートする	選択された企業は以下のとおりです。 企業 (企業集団の中から削除したい企業があれば、"N") (00264) (味の素(株) (00346) (東レ(株) (00351) (旭化成工業(株) (00465) (住友化学工業(株) (00469) (三菱化成(株) 1 : データを表示する・・・通し番号の頭から2:続きを表示する3: 先頭から再表示する4: ソートする(大きい順に並びかえる)5:企業一覧をプリントする9:検索方

図2-8 KIMSの検索例

# \*\*\* 損益計算書 (要約版) \*\*\*

企業コード (01353) 企業名 (トヨタ自動車工業 (株)

上場場所 (東1 阪1 名1 京都 広島 福岡 新潟 札幌 )

勘定項目	8706	伸び率	8606	8506	単位
売上高営業収益率 売上原価・営業原価	6,024,909 5,266,660	-4.44 -2.32	6,304,858 5,391,910	6,064,420 4,958,433	百万 百万
返品調整引当金差額 产上総利益計	758, 249	-16.94	912,948	1,105,987	百万 百万
販売費・一般管理費   営業利益   営業外収益合計	509, 884 248, 364	-12.63 -24.60	583, 560 329, 387	600,096 505,891	百万百万百万
営業外費用合計   営業外費用合計   経常利益合計	187,818 38,174 398,008	-1.82 18.20 -18.51	191,293   32,295   488,385	178,343 36,225 648,009	百万 百万 百万
特別利益合計 特別損失合計			400,000		百万百万
税引前当期純利益 旧特定引当金取崩額	398,008	- <u>18.5</u> 1	488,385	648,009	百万百万
旧特定引当金繰入額 税引前当期利益	398,008	-18.51	488, 385	648,009	百万百万

### 図2-9 KIMSの検索例

# ★ KIMSシステム スタート ★ 最新の財務データは 88年 2月です。

コマンドハ? RETR 8703 LE K003 LE 8802 AND K1411 GE 50000 830 件 見つかりました。

コマンドハ? SAVE 1 SAVE COMMAND END

コマンドハ? RETR 8703 LE KOO2 LE 8802 AND K132 GE 50000/1 57 件 見つかりました。

コマンドハ? DISP K012 K021 K0225 / L1

固定会社コード 会社名

上場場所

246 キリン (株)

東京1部

### 図2-10 KIMSの検索例

### 2.1.2 「PPIS」の利用における課題

以上、国内マクロ経済情報(SDB)、日本貿易情報(TRADE)、大規模企業情報(KIMS)の3つのデータベースを取り上げ、PPISの特徴について簡単に触れてきたが、本報告では、このような経済関連のデータベースを利用する場合の課題として以下のような事項が指摘された。

- ○データのコード体系などのガイドは行っているが、項目の定義や統計情報としての常 識等に関するガイドは、システム内では実現されていないのが現状である。
- ○「ユーザガイド」のようなサポートシステムを提供しているが、現在のような画面の 繰り返しで良いか、また、検索の速度に関しても課題が残っている。
- ○現在のデータベースの利用には、データベース及び経済統計に関して素人の人が利用するケースが多いが、その際、統計データベースの専門家によるコンサルテーションと同様なサポートを計算機側が行ってくれると良い。

### 2. 2 「INDB」の概要と課題

前節では、通商産業省の政策情報システムの利用とその課題について報告を行ったが、本節では、産業統計データベースの構築の現状とその課題について、㈱アイ・エヌ情報センターにおける経済情報のデータベースの構築例を取り上げ、その現状と課題について報告する。

#### 2.2.1 「INDB」の概要

アイ・エヌ「経済・産業」データベース(INDB)は、産業分野に重点を置いた経済 統計データベースであり、その第1次分として昭和61年4月に完成した。以後、漸次、 収録系列を拡張し現在では6万系列の規模に発展している。

収録分野については、表 2-1に示すように、マクロ経済、金融、財政、産業の 4 分野 を網羅している。

分 野	収録出典数	収録系列数
マクロ経済 金 融 財 政 産 業	2 4 2 6 1 2 7	19千 9千 1千 34千
合 計	159	63千

\*系列とは概ね項目と同義 例) 米国への自動車輸出台数で1系列 米国への自動車輸出金額で1系列

\*データ数 63/6末現在 約4,252 千データ データ数= 系列数×収録時間 更新データ入力毎にデータ数は増加 経済・産業データの多くは月次データ INDB収録データは毎月6 万ずつ増大

表 2-1 収錄分野

### (1) INDBの特徴

①産業分野では、質・量ともにわが国最大のデータベースであり、産業28業種のデータを横断的に収録している。28業種の内訳を表2-2に示す。

素材、エネルギー分野	組立加工、建設、不動産、商業分野
鉄鋼、非鉄金属、窯業土石 化学、電力、ガス、石炭、 石油、ゴム、食料品、繊維 木材、家具、紙・パルプ、 出版・印刷	一般機械、精密機械、輸送用機械、 電気・電子、その他製造、建設、 不動産、運輸、倉庫、商業、金融・ 保険、サービス

表 2-2 28業種の内訳

②マクロ経済、金融、財政分野は、「日経NEEDS」と重複しないデータを原則収録している。マクロ経済については"貿易・国際資本交流"統計に特色があり、金融・財政については従来未整備だったものを中心に収録している。

### ③統計異動処理の充実

経済・産業統計は統合、移設、分類変更等統計異動が頻繁に発生している。そこで、INDBでは、これら統計異動をきめ細かくフォローし、データの連続性の確保を実現している(図2-11参照)。

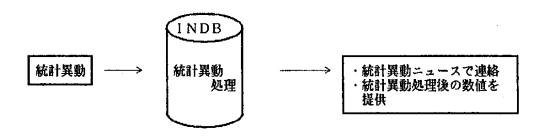


図2-11 統計異動処理への対応

### (2) INDBの利用法

#### ①バッチ帳表について

定期的資料、産業全体感の把握資料に最適であり、ユーザの指定するデータをユーザ が希望する型の帳表 (7種類の帳表パターンから選択) にアウトブットすることがで き、実数、前年比、単位等が表示される。欄外には出典及び数値状況が表示される。

	機械受注	大口電力使用量	建設工事受注額	新規住宅着工戸数	
63/4 5					
6				•	

〔機械統計月報〕〔電力調査統計月報〕〔・・・〕 P:連報 R:訂正・・・・・

表 2 - 3 国内景気指標の帳表

②フロッピーディスク (FD) 渡しによる利用

フロッピーディスク(FD)渡しは、簡単な分析、データのグラフ化に便利であり、 パソコンの機種はNEC98シリーズ、IBM5550、富士通FMR、ソフトは、 マルチプラン、ロータス1・2・3で使用可能である。

③磁気テープ (MT) の利用

社内データベース化、社内システム化により利用が可能である。

(3) INDBの主要データの関連(各サブシステムローカルなデータを除く)

INDBの主要データの関連を図2-12に示す。

①出典情報レコード:出典情報のマスタレコードのことで、発表機関、発表サイクル等 等を収録している。

②原票情報レコード:原票情報のマスタレコードのことで、調査機関、問い合わせ先、

期種、関連原票、データ特性のコメント等々を収録している。

③系列属性レコード:各系列の属性を定義したレコードで、単位、桁数、期種、期種の

変換時の計算方法、季節調整値の計算方法等々を収録している

④数値レコード :数値を収録したレコードである。

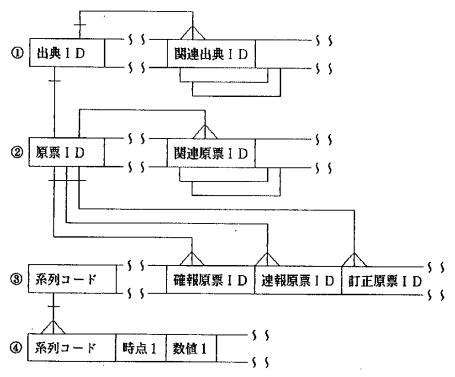


図2-12 INDBの主要データの関連(各サブシステムローカルなデータを除く)

### 2.2.2 「INDB」の課題

### (1) 要素管理システム開発の至る経緯

INDBの第1次のデータ整備において、系列コードは各整備担当者毎に系列の内容を基に付与されてはいたが、その付与基準はまちまちであり、また付与基準自体も考え方に 矛盾を含むものであった。更に系列コードを付与基準からチェックする機能はサポートされていなかった為、誤りも散見された。

また、名称による系列の内容表示も、その記述の形式を管理統制できないことから、系列の意味・内容情報をサポートする必要が意識された。しかし、コメント情報、外部コードについては、その時点でいずれも提供・使用形態を明確化できない上に、データ整備量も過大と判断された為、当面の措置として系列コード上に意味・内容を表現する方向が選択された。

### (2) 要素管理システムの特徴と現状

上記にような、問題意識から、主に、系列コードと意味・内容との対応関係を管理維持する為のチェック機能、デコード機能及びルール情報のメンテナンス機能として「要素管理システム」が企画された。図2-13に、系列コード要素管理の概念を示す。

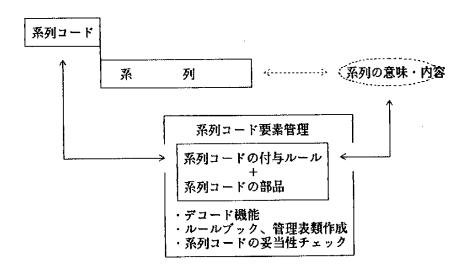


図2-13 系列コード管理要素の概念

また、この情報の検索等への応用については、後日の課題としたが、ユーザへは、「要素表示型コードブック」として、コードの情報を提供することとした。

この要素管理システムの開発では、系列コードの付与ルールを見直し、再構成する必要 があったが、この時の制約として極力、既存系列のコードを吸収する方向で開発が進めら れた。

#### ①要素管理の管理対象範囲

開発開始段階での既存系列については、極力吸収する方向で対応したが、やはり、吸収 できなかった系列もあり、現時点で要素管理の対象となっている系列は94%である。

②系列コードの付与ルール情報群の内容

系列コードの付与ルールとして以下の様なものがデータ化され、収録されている。

●カテゴリー・・・例えば、貿易月表の品目分類コードの集合や、国コードの集合 等々。SDM4Sの確定定義域に対応(概念定義域に該当する ものは無し)。

カテゴリーの固有の情報として以下の様なものをデータ化してある。

- (7)名称
- (イ)カテゴリーの内容についてのコメント
- (タ)所属するコード要素の一般形
- (エ)対応するコード要素のテーブルに関するディレクトリ情報
- ●コード要素・・・・各コード要素は、コード要素のテーブルにカテゴリー毎に登録される。系列コードを新規に付与する為には、ここに登録されているコードを使用しなければならない。

コード要素の固有の情報としては、対応する日本語名称のみ収録している。

●構文・・・・・個々の構文は、カテゴリーの結合規則として記述される。各カテゴリーに定義されコード要素の一般形を連ねることにより、当該 構文にて作成され得る系列コードの一般形が得られる。

この一般形同士が識別できない構文は登録できない。

- 一般形は、系列コードを要素に分解する際に、どの構文に準拠すべきかを判断する為に使用する。
- ●ルール・・・・ルールから貫えば、構文だけで系列コードの付与ルールとして完

結するが、実際に識別等の処理を行う為にはロジックが複雑になる。また、当社の規約としてアプリケーションの開発はコボルまたはアセンブラしか使用できない為、この様な処理はこのままでは開発できない。そこで、カテゴリーの切れ目の桁位置(桁割という)が等しいもの同士をルールという単位で括っている。ルール固有の情報としては、桁割情報と、含まれる構文の内容に関するコメント情報のみが収録されている。

<系列コードの考え方>

実際には、上記したデータの外に、メタルールと呼ぶべき考え方がる。原則としてIN の系列コードは次のような構成をとっている。

述語 + オペランド1 + オペランド2 + オペランド3.....

構文のデータはこれを捨象して、あたかも次の様な表として記録される。

構文コード	カテゴリ1	カテゴリ2	カテゴリ3	カテゴリ 4	カテゴリ5

しかし、事実上、カテゴリー1に来るのは経済活動コードと呼ばれる特殊なカテゴリー が殆どである。

カテゴリー2以降の間の格については活動コードが決定してしまう場合が多いが、その 方向でのシステム化は行っていない。

#### ③要素管理システムによる効果

この要素管理システムの開発による効果としては、以下のような点が挙げられる。

- ●社内的には、コード付与、メンテナンス時のサポートとして利用しており、恣意的なコードの付与や付与の誤りが減っている。
- ●要素表示型のコードブックは、マトリクス帳表等の定義時には若干の利用はされて いるようである。但し、マトリクス帳表のシステム自体は、要素管理のルール情報

を参照していないので、定義された帳表の整合性は人間で判断せざるを得ない。

●MT提供の利用者には、コードの付与ルールがデータ化されているので、直接的なメリットはあるようだが、この分野については試験的な提供にのみ止まっている。

### (3) 問題点と課題

「系列コード要素管理システム」の検討及び開発における問題点としては、以下のよう な点が指摘された。

①系列コードに意味を持たせること自体の問題

系列コードで意味を表現しようとした場合には、次のような問題がある。

- (7) 表現力に限界がある。
- (4) INDBでは、系列コードにアライアスを持てないので、表現は一通りに制限される。

これらの問題点は、構想当初から指摘されていたことであるが、この為にルール情報 が、

- (7) 系列の内容・意味情報の整理の仕方、経済統計に対する見方
- (イ) 系列コード自体の字形規則
- の2つを混在させることになり、いくつかの問題を生じさせている。

#### ②経済統計に対する見方の問題

経済統計に対する見方の整理・体系化が、系列の意味・内容のデータ化には必要ではあるが、系列コードとして表現するには、制約をおかざるを得ない為、本来の意味・ 内容の表現よりもむしろ系列間の識別に重要な意味・内容の表現に傾かざるを得なくなる。この事は、構文の切り方自体を恣意的なものにしてしまっている。

### ③字形規則と判別処理

逆に、コードを付与する為の基準としては、多様な内容を表現する必要から、どうしても複雑過ぎるものとなる。それ故、システム的なサポートが必要となったのだが、 処理の内容としては、COBOLで開発するにはやや複雑すぎたようである。

#### ④コード要素の付与規則

各カテゴリーに属するコードの一般形は、データ化し収録してあるが、実際に個々のコード要素を付与する基準はデータ化できなかった。この理由は、以下の問題が未整理であった為である。

- (7) コード自体が項目間の関連を表すべきであるかどうかの問題
  - 即ち、階層表現をとるべきかどうかについての問題である。これには以下のようなイレギュラーなケースがある。
  - ・分類の組替えが発生すると、上下関係は目茶苦茶に壊れてしまう。変更されな かった系列は存続させる必要があり、変更前のものについてもDB上に残す必要 がある場合には、コードに表現された階層は全く意味をなさなくなる。
  - ・純粋な階層関係ではない分類体系があり得る。つまり、複数の基準の直積で表現 された部分と階層の混在は、コードに各ケタの意味を不定にしてしまう。特に、 直積で表現されたものは、どのようにでも階層関係が組める。この解決には、こ れらを分解してしまうことが考えられるが、ユーザが使用している分類体系を無 闇に分解しても意味がない。
  - ・階層関係が厳密でない場合がある。つまり、合計値の厳密さを求めようにない 分類については、類似性、概念の包括性のみが階層を決定する。このケースは 歴史の浅い業界の統計に見られる。
- (4) コード要素の付与ルールを記述する型式が特定できない。

階層表現をそもそも必要としない分類もあり、一律、特定の記述型式でデータ化できる基準に限定することは難しい。勿論、コード自体で項目間の関連を表すべきではないとすれば簡単だが、系列間の関係を外に持たない以上、ある程度の表現をせざるを得ないのが現状である。

#### ⑤様文の表現力

系列コードの作成ルールである以上、意味・内容の充分な表現というよりは、系列の 識別へ傾きがちであることは、②で述べた通りである。しかし、捨象する内容がある 程度の分類項目であればともかく、各カテゴリーの格関係まで捨象している場合もあ る。特に、単文の表現ならまだしも、関係代名詞節、接続詞節のような節を伴った複 文で表される内容の表現ではかなり強引な省略を行わざるを得ない。しかも、その規 則すら、「識別する為に必要なものは残す」以外の規則は無い。また、「何と識別す る必要があるか」及び、「何と類似させる必要があるか」によって、構文の設定は様 々になる。

このような理由から、構文内でのカテゴリー間の意味的な関係まで特定することは、 このシステムでは断念した。

#### ⑥カテゴリー

- (7) カテゴリーの関係はデータとしては持っていない。 特に、類似の分類同士の際についての情報が必要と考えられたが、データ化する方 法を検討しきれなかった。
- (イ) 運用上は、更に異なるカテゴリーに含まれる類似概念のコードの異動情報が必要と なるのだが、これについては目処が立っていないのが現状である。
- (ウ) カテゴリーの設定についても恣意性が残る 分類体系自身に権威がある場合は別だが、工業会統計を収録するような場合、それ ぞれに統計毎にカテゴリーを設定して行くと、2~3ヶ程度のコード要素しかない カテゴリーが大量にできてしまうことから、利用上は、まとめることになり、恣意 性が残る。

### ⑦INDB自体の問題

(7) 上位系列(複数の系列の合計値)の収録を行うか否かについての基準が無く、また 系列の関連性を保存しないので、収録系列間で複雑な階層関係が発生することから コード要素の階層表示が難しくなっている。

### (イ) 1次統計、2次統計の扱い

- ・2次統計には、全く異質の統計を組み合わせて表示した統計表があり、表自体の制 約を外さない限り、無理なコードの付与方法にならざるを得ない。また、同じ系列 でも、選択した統計表により、構文、カテゴリーの設定が異なってしまうケースが ある。これは特に、産業分野の周辺的統計に著しい。
- ・また、系列の所属する統計表は、確報値については1つしか定義できないようになっており、同じ系列を含む複数の表を収録する場合、同じ内容の系列をコードを変えて収録しなければならない為、コード要素の一部には不整合が生じている。しかしながら、統計表については、いずれを選択するかはユーザの希望に依存し、統一的な扱いができないのが現状である。

#### (ウ) 類似統計の問題

内容だけ見ると全く同じだが、互いに数値が異なる統計がある。調査対象範囲が微 妙に異なる等が原因であるが、これをコード上に厳密に反映することは難しく、特 に、その違いが何に由来するかが統計ユーザに対して明確でない場合には、調査発 表者自体を明示する以外に対応策がないのが現状である。しかしながら、産業分野 の統計では、ユーザがこのような統計をいずれも必要とするケースが多く、対応に 苦慮することになる。

#### ⑧運用上の問題

#### (7) 通常運用

- ・コード要素の新規追加時に、コード要素の付与規則がデータ化されていない事が問 題点として挙げられている。
- ・また、階層表現をとらざるを得なかったコード要素については、体系を維持することが難しく、結局断念したケースもある。

### (イ) 収録範囲拡張時の問題

- ・上記⑦の問題は、常に考えられる事であるが、以前と比較すると系列コードの付与 の工程は短縮されている。
- ・一方、産業統計の場合には、その業種毎の特性に関する知識が整備上不可欠となる が、習得するための手段・時間は限られており、これが最大の問題点である。

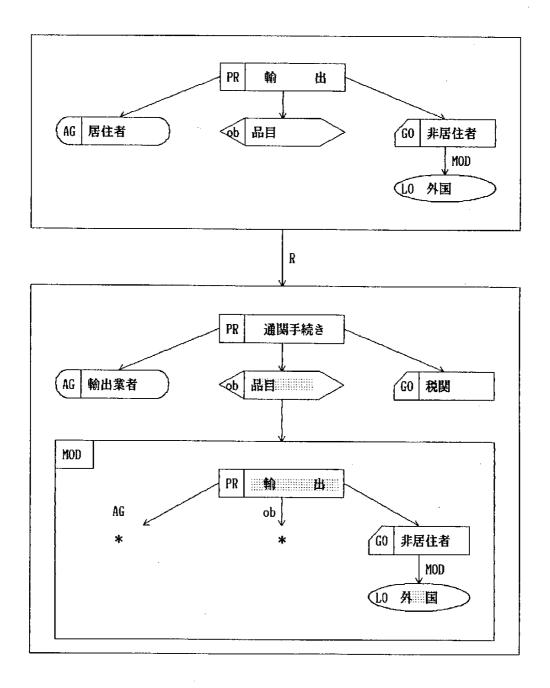
#### (4) 今後の対応

INDBの大きな特徴として、データベースを「事象」と「観測」の関係で捉え、そこに文法理論を適用していることが挙げられる。特に、「観測」、「観測する行為自体」、「統計」に関して以下のような項目を挙げている。

①観測者 ②観測結果の表示 ③観測範囲 ④観測方法 ⑤ネットワーク化 ⑥フレーム表現の利用

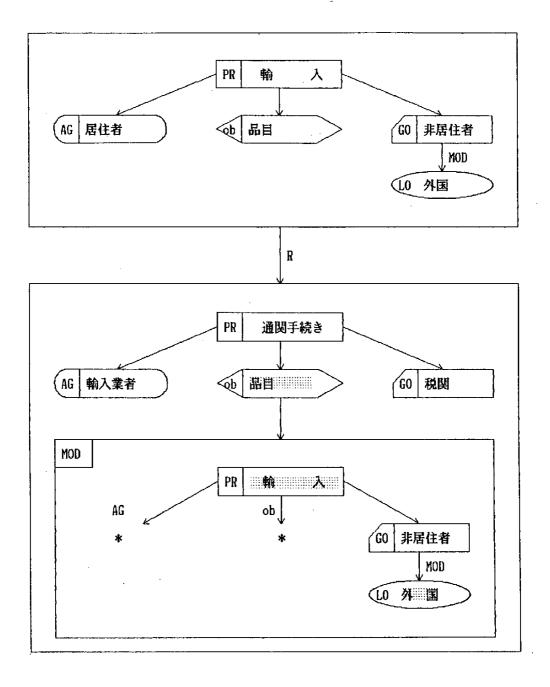
この中で注目されるのは、⑤のカテゴリーのネットワーク化を試みている点である。 2-14-1、2-14-2 は、経済統計の意味連間ネットワーク図である。

結果的には、産業、マクロ、金融、財政等を一括して抽象化することは、観測方法が不明であり、また定義がまちまちであることから難しいが、分野を限定することによって有効な手段となる可能性を含んでいる。このような、経済統計の意味連間ネットワークの作成は、⑤と共に人工知能技術の応用への示唆を提供してくれるものでる。



統	計	通関輸出
原	票	E41010CM , E4202DCM , E4202OCM , E4204DCM
		E42040CM
		•

図2-14-1 経済統計の意味連間ネットワーク図



統	<b>a</b> †	通関輸入	
原	票	E41020CM , E4203DCM , E4203OCM , E4205DCM	
		E42050CM	
		•	

図2-14-2 経済統計の意味連間ネットワーク図

### 2. 3 「AID·XT」の概要と課題

本節では、経済統計の中で、特に貿易統計データベースの開発と利用が進められている アジア経済研究所の貿易データ検索システム「AID・XT」について報告する。

アジア経済研究所では、長年にわたって国連およびOECD編集の約130ヵ国、1962年~最新年の世界貿易統計データベースを構築し、研究所内外の利用に供している。

従来は磁気テープによる検索を行っていたが、昭和57年度には、さらに一歩進めて磁気ディスクを利用し、アジア経済研究所独自のオンラインによる検索システム「AID・XT」を開発した。これにより、利用者はコンピュータの専門家の手を煩わすことなく、各自が端末装置の利用を介して、貿易データの検索、加工、計算、作表等を行うことが可能になった。

そこで、貿易統計データベースの利用の髙度化の視点から、「AID・XT」に於ける データベースの構築、システムの機能、利用傾向、課題等について述べることにする。

### 2.3.1 「AID·XT」の概要

# (1) AID·XTデータベースの構築

貿易統計は、一国の経済を知る上で基本的な統計の一つであるとともに、他の統計と比較して信頼性が高いことから、その利用価値は非常に大きいと言える。

アジア経済研究所では、以前より国連・OECD貿易統計を国際商品分類別(SITC R1またはSITC R2)、相手国別の詳細な年次データとして磁気テープの形で整理保管し、検索利用してきたが、利用の増大、データ量の膨張などのために、磁気テープの形では利用者の要望に簡単に答えることが難しくなってきた。

こうした状況のもとで、昭和57年度に開発されたのが、AID・XT ( AIDは、 Ajiken Indicators of Developing Economies の略、Xは原データ、TはTrade の意味 )である。

国連とOECDは、それぞれ異なる国コード、数量単位コードを使用しており、また、磁気テープに記録するときの形も異なっている。このため、コード等を整理・統一し、さらにオンライン処理を可能にするために、これらを圧縮した形に変換してディスクに蓄えている。AID・XTは、このデータをデータベースとして使用している。データベースの編成は、表2-4のとおりであり、このデータベースから、ユーザの要求に従って必要なデータを取り出すことができる。

保有状況・・・・報告国 130ヵ国 期間=1962年~1987年

データ量・・・・約1,400 万時系列レコード

データ項目・・・年次=西暦年

輸出入区分=AID・XTコード 相手国=AID・XTコード 報告国' = AID・XTコード

商品分類<sup>2</sup> = SITC·R1, SITC·R2コード

金額=1000USドル表示

数量=AID・XT数量単位コード表示

注) 1 報告国、相手国は単国のみで、地域計は含まれていない。

商品分類は、国連データはSITC・R1を、OECDデータは 1977年まではSITC・R1を、1978年以降はSITC・ R2を使用している。

#### 表2-4 AID・XTデータベースの編成

出典:アジア経済研究所「アジ研ニュース (NO.50)」

また、このAID・XTデータベースの構築は、図2-15のようなプロセスで行われている。

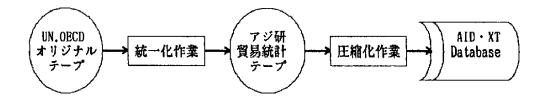


図2~15 A I D • X T データベースの構築

出典:アジア経済研究所「アジ研ニュース (NO.50)」

#### (2) システムの機能

AID・XTは、データベースから特定のデータを抽出する検索機能、抽出されたデータを加工計算する加工計算機能、抽出されたデータと加工計算されたデータを作表する作表機能を持っている。

これらの機能の選択は、AID・XTの「機能選択画面」を通じて実行する。

#### ①検索機能

利用者が指示した条件により、AID・XTデータベースから特定のデータを抽出する

機能で、検索条件は、「検索条件指示画面」を通じ、報告国・輸出入区分、商品分類、相手国、年次をそれぞれコードで指示する。

尚、この機能には、単にデータを抽出する以外に、次のような処理も同時に含まれている。

- (7) 報告国の統合
- (イ) 相手国の統合
- (ウ) 商品分類の統合

次に、以下のような要求内容に対して検索を行った場合の例を図2-16に示す。

### <要求内容>

「日本やアジア諸国から、衣料品が日本、アメリカ、EC、世界にどれだけ輸出されているかを調べる。輸出国として(I)日本、(2)香港、(3)台湾、(4)韓国、(5)アジアNIES(香港・台湾・韓国の計)及び(6)これらの諸国の合計、をとり、輸出相手国として、日本、アメリカEC(加盟9ヵ国)、世界をとる。対象期間は1965、70、75~80年とする。」

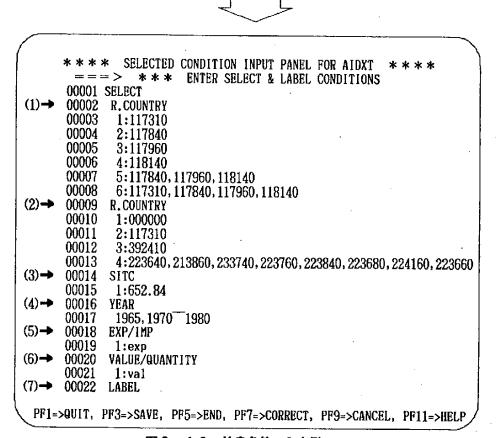


図2-16 検索条件の入力例

- (1) 報告国 (R.Country)では、上記、輸出国を(1)~(6)まで、国コードで入力する。
- (2) 相手国 (P.Country)に、相手国コードを入力する。
- (3) SITCは、R1で、652(綿織物) と 84(衣料) の合計を求める。
- (4) 年次を指定する。
- (5) ここでは輸出と指定する。
- (6) ここでは価額と指定する。
- (7) 統合した国と統合したSITCに対してのみ、ラベル指定する。

出典:アジア経済研究所「アジ研ニュース(NO.50)」

以上のように検索条件を指定し、検索が終了すると、次にユーザは、この結果をプリントアウトするか画面に表示することになる。

### ②作表機能

検索機能で抽出されたデータ、または加工計算機能により加工計算されたデータを作表 し、端末装置のブラウン管上に表示したり、計算機室の高速ラインプリンターにプリント したりする機能である。

ユーザは、貿易マトリクス、時系列表などを「作表条件指定画面」を通じて、各自必要 に応じて作表することができる。

#### ③加工計算機能

既に検索されているデータに対して次の計算を行う機能がある。

- ○年度比(前年度比、基準年度比)/報告国比/相手国比/商品比
  - ・「年度比」は、貿易取引(数量、価額)の増減の変化を時系列でみる場合に使用する。
  - ・「報告国比」、「相手国比」は、貿易取引(数量、価額)の国別シェアを見るため に使用する。
  - ・「商品比」は、商品シェアを見るために使用する。

これらの指示は、「加工計算条件入力画面」を通じて行う。

AID・XTは、現在標準的な機能しか備えていないが、さらに進んだ分析を行う人の ために、抽出したデータを、SASプログラム・パッケージで利用できるデータ形式に変 換するプログラムを用意している。

以上、AID・XTを使用することにより、過去25年間以上にわたる各国の商品ごとの貿易取引のデータが簡単に取り出せることから、貿易に関する実証研究をしたい人にとって非常に有効な武器となっている。

### 2.3.2 AID·XTの課題

#### (1) 利用のメリット

次のような形で貿易統計を利用する場合、出版物から手作業でデータ収集することは事 実上不可能であり、AID・XT利用のメリットがある。

- ①複数国のデータの国際比較 (ドル・ベース)
- ②同一分類 (SITC. R1) による長期時系列 (AID・XTには、SITC. R2 分類のデータをSITC. R1分類へと自動変換する機能がある。)
- ③独自の地域分類、品目分類
- ④貿易マトリクスの時系列表

### (2) 利用者層の分類

- ①国際標準貿易分類 (SITC) について全く知識がない人とある人。即ち、最近の利用者の傾向としては、貿易統計の専門家と一般ユーザへの二極分化が見られる。
- ②出版物を見る代わりに使用する人(少量データ)と大量データ処理をする人。

### (3) システム利用前の準備

システム利用の為の検索条件の設定に至るまでに次のような準備が必要となる(図2~17参照)。

①対象国の決定

地域計を作成する場合は、地域名と国名の対照表を作成(例:EC加盟国リスト)する。

#### ②対象品目の決定

SITC各桁以外の統合品目概念を使用する場合は、独自の分類とSITCコードの 対照表を作成する(例:製造業品のSITC分類による定義)

- ③検索年次等の決定
- ④「データ入手状況表」によって、各国データの最終年を確認。必要があれば検索条件を修正する。
- ⑤作表に於ける表のかたちをイメージし、表頭及び表側に対応する項目(報告国、品目等)を決定する。

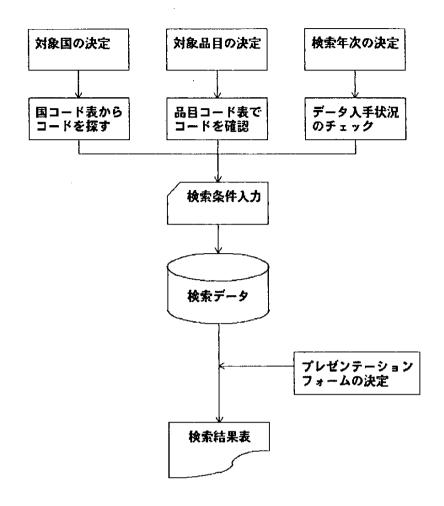


図2-17 検索システム利用の手順

#### (4) システム利用の問題点

- ①「データ入手状況表」が画面に表示できると便利である。
- ②検索条件の設定方法に関しては、専門的利用者については現在のシステムで特に不都 合はない。
- ③一般ユーザは、(3)の②、⑤に困難を感じ、専門家に相談せざるを得ない。特に②に関しては、人工知能技術を利用したシステム化が期待される。

### (5)検索データの加工

シェア計算等のほかは、現在SASフォーマットに落として各自計算している。今後は 標準的な「水平分業指数」、「RCA指数」、「集中度」等の計算機能が加わると便利で

### (6) データに関する問題点

AID-XTのデータに関して主な問題点は以下の通りである。

#### ①社会主義諸国等のデータの欠如

中国及びソ連・東欧等社会主義諸国のデータを含め全世界のデータを網羅している訳ではない。また、各国のデータが全て最新年までそろっている訳ではなく(例:マレーシアは1962年~1985年)、対象年の一部データが欠如している場合がある。

#### ②相手国の欠如

1970年以降、相手国台湾が国連貿易統計から除外されている。

上記の問題点は世界計及び地域計を計算するとき注意を要する。利用者に負担をかけず にすむシステムの構築が急務である。

#### 2.3.3 人工知能技術の利用可能性

上記のように、AID・XTシステムは、貿易統計の専門家にとっては非常に便利なシステムであり、検索機能、作表機能、加工計算機能など多機能システムである。しかし、最近の利用者の傾向から分かるように、貿易統計に関して一般ユーザの利用も増えてきたことから、2.3.2 の(4)の③のような問題点が指摘された。

そこで、これらの問題点から、貿易統計データベースの検索に対して、人工知能技術が どのように利用できるかについて、AID・XTシステムを例に整理してみよう。

- (1) AID・XTシステムにおける検索に必要な知識
- ①データの入手状況(対象国、対象年)に関する知識
- ②国別コードに関する知識

2.3.2 の(3) の検索手順で示したように、利用者は、AID・XTの国コードに関して知っている必要がある。これには例えば、ヨーロッパ・グループ、アジア・グループ、アフリカ・グループ、ECのみといった表現を理解する知識も必要となる。

- ③SITCの分類に関する知識
- ④プレゼンテーション・フォームに関する知識

表頭及び表側に対応する項目を指示することにより作表を行うが、この作表規則を

理解することが必要となる。

### (2) AID・XTシステムの改良と人工知能技術の利用

(1)に示したような知識に関して、貿易統計の専門家がどのように対応しているかについての分析が可能ならば、それらを知識ベース化することによって、貿易統計の一般ユーザをサポートするシステムを構築することも可能といえよう。但し、どのような範囲で支援するシステムかを明確にする必要がある。こうした点も含め、今後、AID・XTシステムに望まれる機能をまとめると以下のようになる。

- ①データ入手状況表が表示され、特に、「これは~年までしかデータがない」といった サポート機能
- ②検索条件の設定に関する知識、条件設定をサポートする機能
- ③一般の利用者にとって、商品の中身について説明するような機能
- ④頻繁に使用される国グループに対応した国コードの蓄積とそれに関連する情報の提供
- ⑤加工計算を支援する機能(通商白書、経済白書等に見られる一般的な指数表示機能)
- ⑥欠損データに関する知識の提供と欠損データの補充機能

### 2. 4 ジェトロのデータベースシステムの概要と課題

前節まで、経済関連データベースの現状について、主にデータベース構築の視点から報告を行ってきた。

そこで、本節では、経済関連情報の中で特に貿易関連のデータベースである日本貿易振 関会における貿易統計データベースの概要と利用を中心に報告を行う。

# 2.4.1 ジェトロのデータベースシステムの概要

ジェトロのシステムは、1969年コンピュータ導入以来、一貫して海外情報の処理と 利用を中心に行われてきた。この中で、昭和60年度の実績では、事務管理業務はわずか 約2%のシェアしか占めていない。

海外情報の処理と利用は、当初、個別処理が大部分であったが昭和57年共同利用オンラインデータベースシステム(WISE)開発以来、ジェトロ各事業実施に役立つ汎用システムを中心に運営されている。本報告では、このWISEを中心にシステムの特徴について報告する。

### (1) WISEの基本的な考え方

### ①全社的な情報利用

ジェトロの海外情報は、ジェトロ海外情報収集網から集積・蓄積したものと、各国の統計局などから入手しているコンピュータデータに分けることができる。いずれも特定部門・特定業務目的に使用するのではなく、全社的に使用できるものである。このため、システム開発は全社的に利用できるオンライン化が不可欠と考え、ジェトロの全端末からリアルタイムで検索できるようすすめられている。

#### ②全員が利用の対象

WISEのデータベース利用は特定の担当者を想定せず、役職員全員を利用対象者に考えられている。

#### ③外部サービス

WISEのデータベースは、ジェトロの内部利用に限られず、海外の経済貿易情報に関心を持つ者に有用な情報源となることを目標にしている。

通商産業省のACOSシステム1000と接続して相互に相手のデータベース利用を促進するため、昭和59年9月「通商産業省-JETROオンライン情報交換システム」を

稼働させた。これによりWISEのデータベースは通商産業省の全端末(約300台)からアクセス可能となった。

さらに、昭和61年4月よりWISE中のデータベースの一つ「国別経済貿易情報(JETRO ACE)」をデータベース・サービス会社からオンラインサービスとCD-ROMによる提供を開始した。

### ④OAの核

ジェトロでは、OA化推進の核をデータベースと考え、各種の事業運営に役立つ内外のデータベースを社内に構築することをめざしてしている。特に、海外事業所を活用して収集した海外情報をデータベース化して、情報の多角的、多目的利用をはかることを中心にしている。

### (2) 共同利用データベースの開発

WISEのデータベース構築はDBMSにINQを用い、検索・更新等を簡便に行うためのアプリケーションプログラムを作成した。

### ①マニュアルレスのシステム

WISEが全国的・全員を対象にする限り、マニュアルが不要なシステムを目指している。開発当初はブックタイプのキーボードによるシステムを考えたが、パソコン端末機の 導入時に、全てガイド画面方式に統一した。メニュー画面に実行可能なコマンドを表示し 表示されているものから選択すればよいので、マニュアルは殆ど使用されていない。

マニュアルレスのシステムの最大の長所は、利用体験のない者や使用頻度が少ない者でも気軽に操作できることである。

#### ②情報内容に即した処理

WISEのシステムでは、データの加工が殆どできない。まとまった形でアウトプットできるようにシステムがつくられているのが特徴である。マニュアルレスはコマンドレスにつながり、データベース内容を最大限に生かすように考えられているので、情報内容に精通していない者でも簡単に結果をえることができる。

#### ③多様なレポート出力

ディスプレイ画面で見るだけでなく、通常の使用方法ではハードコピーを作成することが多い。画面のコピーが社内の報告資料として、そのまま使用できるように画面設計している。また、プリント用のコマンドを用意して簡単にレポートが作れるようにしている。

ジェトロのコンピュータ情報について、2 - 18 - 1、2 - 18 - 18 - 18

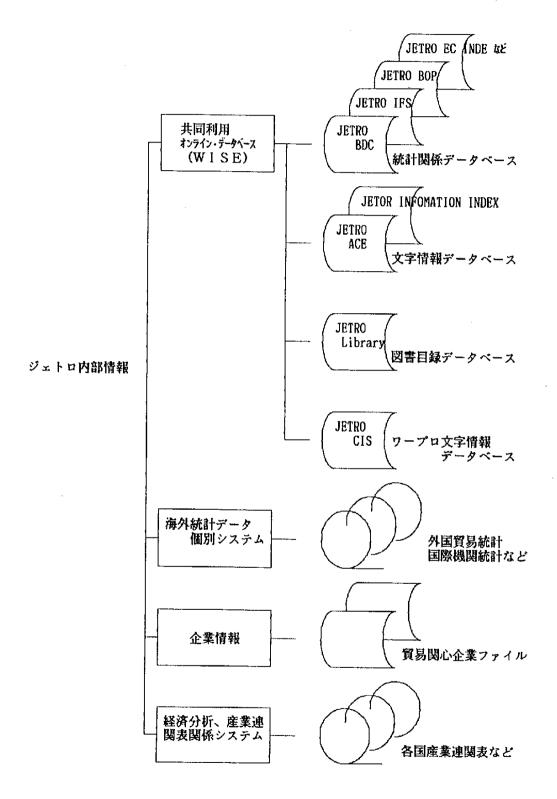


図2-18-1 ジェトロのコンピュータ情報

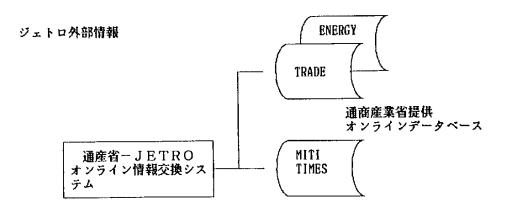


図2-18-2 ジェトロのコンピュータ情報

### 2.4.2 ジェトロのデータベースの内容とシステムの特長

図2-19に示すようにWISEのデータベース一覧はメニュー画面のとおりである。 その中から代表的なものとして、「貿易統計システム(JETRO BDC)、国別経済 貿易情報システム(JETRO ACE)の二つを取り上げ説明を行う。

# JETRO データベース利用ガイド

#### データベースの種類

01:貿易統計:BDCシステム(日、米、EC等の貿易統計検索)

02:1FS(時系列表·地域別国際比較表·国際収支表·商品価格表)

03:BOP (時系列表)

04:ジェトロ情報インデックス (ガイド画面を用いて検索)

05:情報伝達システム(ワープロ情報の掲示、メーリング)

06:EC国民経済計算(統合勘定表·付帯表·人工·価格指数·等)

07:EC事業所統計(NACE分類による製造業統計)

08:EC生産・輸出・輸入統計 (産業別)

09:定期刊行資料の検索(ジェトロ収集の統計書等定期刊行物受入れ状況)

10:国民経済・貿易情報・ACEシステム(各国の経済・貿易事情検索)

99:終了

選択は?

0 1

図2-19 WISEのメニュー画面

- (1) 貿易統計システム (JETRO BDC)

#### ①概要

ジェトロのコンピュータデータの大部分を占めるのが各国の貿易統計の磁気テープである。コンピュータ導入以来、各国の貿易統計に関して、最も詳細で最新の内容のものを、 磁気テープで収集することを目標にしている。

現在保有する主要なものは、

- (7) 日本通関統計(大蔵省)
- (4) 米国通関統計(商務省)
- (ウ) 英国通関統計(中央統計局)
- (1) フランス通関統計(大蔵省)

- (オ) 西独通関統計(中央統計局)
- (a) OECD貿易統計(全加盟国) (OECD統計局)
- (+) E C 貿易統計 (全加盟国) (E C統計局)
- (7) Direction of Trade (IMF)

などである。

これらの処理と利用はバッチ処理主体で運用されているが、バッチ処理では迅速な利用には限度がある。できれば、全データをオンラインデータベース化して、自由にアクセスできることが最終目標である。

#### ②システム開発の経緯

コンピュータ導入以来、貿易統計データのオンライン検索システムを開発することが大きな課題であった。障害となったのがデータ量と検索キーとなる商品分類である。

米国輸入統計を例にとると、品目数は約1万。1回の更新データは100万系列を超える。このようなデータを毎月、多数の国を対象に運用することは容易ではない。しかも、各国の品目分類は異なるうえに、毎年1~数回の改訂がある。利用者が自由に検索するには、商品分類や品目番号の知識なしには困難である。

そこで、JETRO BDCシステム(貿易統計データベース)は、上述の課題を解決 する一方法として開発された。その狙いは、次のとおりである。

- (7) 限られた磁気ディスク容量に多くの国のデータを収録することを最優先にする。現 在収録している通関統計は日本、米国、英国、フランス、西独、イタリアおよびE C(合計) 7ヵ国である。
- (イ) 各国の品目分類番号に精通している者はきわめて少ないことから、ジェトロ独自の コードを付して、各国の分類番号が分からなくても検索できるようにした。

このことから、重要な品目約2000品目を選んで、データベース化をはかっている。 さらに、利用者が分類番号の改訂の有無を意識することなく、利用できるように時系列的 なデータ編集を行っている。品目分類改訂の一例を米国の半導体の輸入分類で示したのが 図2-20である。

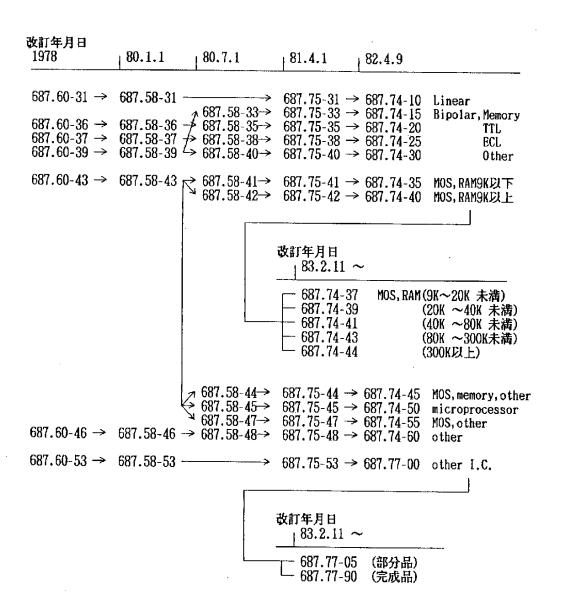


図2-20 | C (集積回路) の品目分類番号の変遷 (米国輸入統計)

(ウ) アウトブット形式を最も使い易いものと思われるものに予め決定した(図2-21 -1、図2-21-2 参照)。

R IDLE K	B0			079		R	5.04	OPEN	チュウ
	SPORTS -1-K	FOOTWE D200 ) Rank	AR	* * ( JAPAN RANK 7 1-06 SHAR		( V ( U Rank		K DZ	)
WORLD (0000)	N	.A.		N.A.	2	23,594 4,217	100.0 100.0	  -	
TAIWAN (1170)	N	.A.		N.A.		1 81,032 1,576	36.2		
R KOREA (1121)	N	.A.		N.A.		71,272 1,406	31.9		
CHINA (1130)	N	.A.		N.A.		23, 812 968		_	
ITALY (2350)	N	i.A.		N.A.		19,825 32	•		
REPORT CT	Y 1	туре	1	PARTNER CTY	0000	ITEM H	(D200	VALUE	1

図2-21-1 JETRO BDCシステム品目別国別表画面(例)

K DZ	N.A.	223,594 4,217	100.0	0.2 100.0	_
N .	N.A.	21,836 1,370	100.0	0.0 100.0	<u>-</u>
	E 1	E 1 PARTNER CTY	E 1 PARTNER CTY OO ITE	E 1 PARTNER CTY OO ITEM KD200	E 1 PARTNER CTY OO ITEM KD200 VALUE

図2-21-2 JETRO BDCシステム国別品目別表画面 (例)

#### ③システムの特徴

JETRO BDCの概要は、図2-22に示したとおりであり、以下のようなサブシステムから構成されている。

- (ア) オンライン検索サプシステム
- 二(イ) オンライン・プリント・サブシステム
  - (ウ) オフライン・プリント・サブシステム
  - (1) 国分類コード・マスタ管理サブシステム
  - (オ) 品目分類コード・マスタ管理サブシステム
  - (カ) データベース更新サブシステム

- (‡) 個別業務アウトプット・サブシステム
- (ク) 他のデータベースへのデータ加工サブシステム
- (ケ) その他

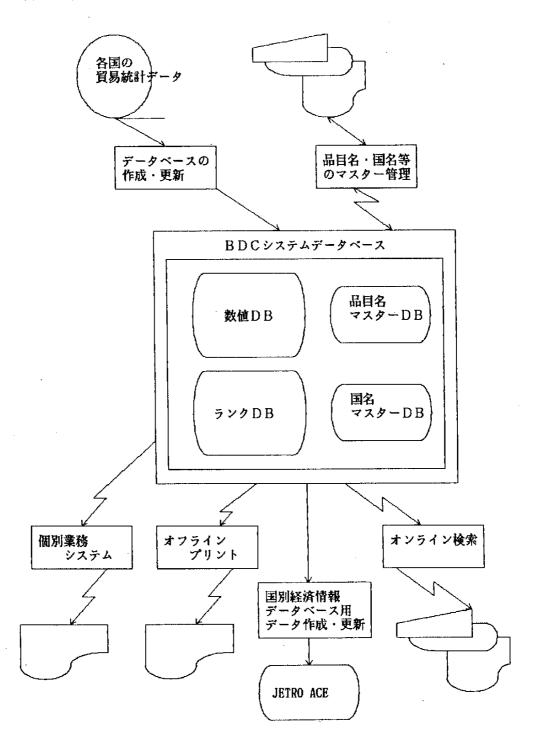


図2-22 JETRO BDC (貿易統計システム) の概要

本システムの特徴の一つは、海外事務所のスタッフあるいは海外諸国から受け入れた調査員にための英語のガイド画面を採用していることである。このため、簡単な説明で外国人でも使用できるデータベースとして好評であり、在日外国公館あるいは海外でオフラインブリントによるアウトブットも好評である。なかでも国別の貿易動向が分かるものは、他に殆どないため、概略から詳細まで自由に選択できる実用的なものとして評価が高い。

#### ④課題

コンピュータ資源等の制約から、保有データの一部をデータベース化しているが、理想は全データを対象にすべきである。実用上ビジネスに使用するには現在のシステムで十分であるが長期の時系列分析を行う場合や、対象としていない品目のデータを希望する場合には、バッチ処理で対応するしかないのが現状である。

JETRO BDCシステムの今後の課題は、全データを対象にできるシステムを目指すために、分類の改訂情報のデータベース化、分類コードを容易に検索できるための分類表データベースの開発が必要とされている。また、図表化してアウトブットできるようにすることも大きな課題といえる。

(2) 国別経済貿易情報システム (JETRO ACE)

#### ①概要

ジェトロは約60ヵ国に約80の海外事務所を持ち、情報の収集等にあたっている。この海外に事務所にネットワークを生かして開発されたものが国別経済貿易情報のデータベース(JETRO ACE)である。

JETRO ACEの特徴は次のとおりである。

- (7) 世界主要100ヵ国が対象である。
- (イ) 経済・貿易・投資に関するマクロ情報を網羅している。

#### 〔主要な項目〕

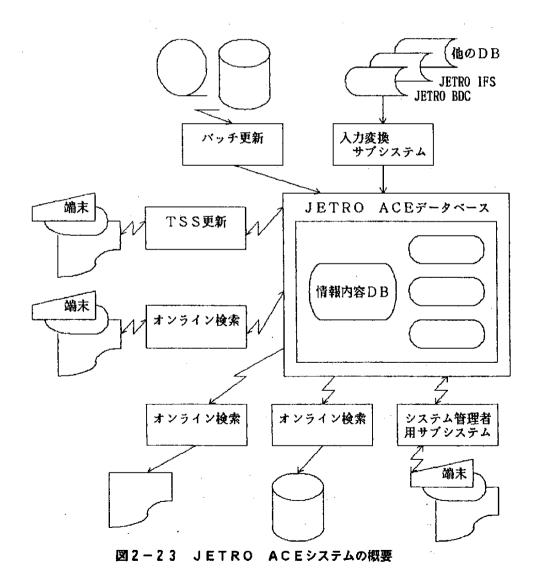
- ○経済動向(経済概要・動向、経済政策、開発計画、経済見通し)
- ○貿易概況(貿易構造・動向、総額推移、主要品目別輸出入、主要地域別輸出入)
- ○投資動向(対外・対内直接投資概況、投資環境概要、外貨政策)
- ○対日経済関係(日本との貿易概況、投資交流、経済協力、対日輸入規制)
- 〇主要経済指標(GDP(GNP)、消費者物価、卸売物価、鉱工業生産、労働指標 国際収支)

- ○各国にプロフィール(面積、人口、民族、言語、気候、政情、マスコミ、教育)
- ○主要官庁、貿易関係機関(主要機関の住所、電話番号)
- (ウ) 現地資料を原則として情報源にする。
- (エ) 日本語で利用できる。

海外情報源の多くは現地の公用語で書かれている。アジアの国だけみても、韓国語、中国語、タイ語、マレー語など、多様な言語が使用され、各国の資料を読むことは容易ではない。

- (オ) 簡潔にまとめた情報で提供する。
- (カ) 国と次項を指定すれば、必ず最新の情報内容が得られる。

このJETRO ACEのシステムの概要を図2-23に示す。



-66-

このシステムの開発では、検討段階から、対外サービスを想定した。これは、米国の状況から、わが国でも急速に商用データベースが普及することが予想されること、JETRO A-CEに類似したデータベースは皆無と考えられたこと、海外への関心の高まりと国際化の影響で、海外情報ニーズは飛躍的に高まると考えられたことなどの理由による。さらに、海外情報ニーズの高まりは、利用者層の拡大につながり、従来のような海外情報に精通しているもの以外の需要者にとっても使い易いシステムであることが必要となると予想され、このことから、中間管理職者が自ら端末機を操作できることを前提として、検索に習熟を必要としないシステムを目指した。

内容的な特徴と操作的特徴を以下に示す。

#### -内容的な特徴-

- ○簡潔に要旨のみを記述する。
- ○長文の場合、小見出しを付け、項目別に分けてデータ作成する。
- ○内容を刻々更新し、最新情報を提供する。
- ○概要、それを裏付ける統計データおよび出所・参考の3本立の構成にする。
- ○注釈をできるだけ記述し、利用者の情報内容理解に役立たせることを目指す。

#### 操作的な特徴-

- ○徹底的なメニュー方式を採用した。
- ○コマンドによりアウトブットができるようにした。利用者はハードコピーによる 使用が多いので、プリント用コマンドを入力することにより、まとまった情報内 容を一括し、用紙サイズに合わせてアウトプットできる。
- ○統計データは加工しないで利用できるようデータ入力した。
- ○5個以内のコマンドで検索できるよう簡素化をはかった。

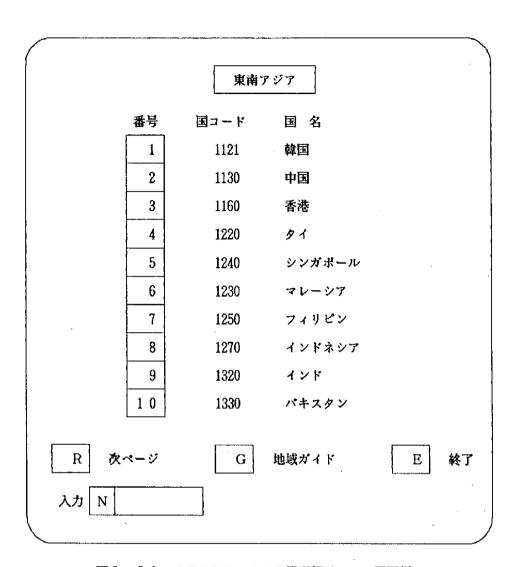


図2-24 JETRO ACE国選択メニュー画面例

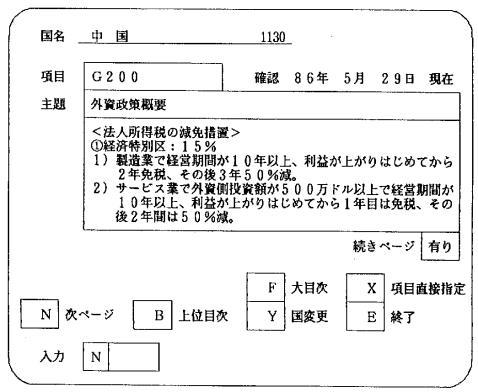


図2-25 JETRO ACE内容画面例

#### ②他のデータベースの活用

JETRO ACEには、他のデータベース(特に統計関係データベース)の内容に一部が簡単に分かるようになっている。例えば、BDCシステムの日本通関統計から、国別に次のような項目の内容を検索できる。

○国別貿易バランス、○国別輸出総額推移、○国別輸入総額推移、○国別部門別(大分類別)輸出額、国別部門別(大分類別)輸入額、○国別主要輸出商品、○国別主要輸入商品、○国別輸出増加種主要商品、○国別輸入増加種主要商品、○国別輸出減少種主要商品
○国別輸入減少種主要商品

このため、貿易動向を個別商品ごとに調べる場合には、BDCシステムを利用し、A-CEシステムではどのような商品の輸入が多いのか、輸出の伸びが著しい商品はどれかなどを簡単に知ることができる(図2-26参照)。

国名	名 中国 1130	
項目	A140-86	確認 86年 5日 現在
主題	日本の商品別輸入(86年)	
内容	日本の輸入 1986年1月- 価価 名 総額・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	2 100.5 -8.9 0 18.8 34.9 4 13.9 -1.6 6 33.5 -35.4 8 30.9 13.4 9 5.7 13.4 1 0.4 47.2 6 0.1 124.3 8 0.2 21.5
		続きページ 有り
S	出所・参考	F 大目次 X 項目直接指定
N	次ページ B 上位目次	Y 国変更 E 終了
入力	N	

図2-26 JETRO ACE内容画面例 (JETRO BDC より作成)

#### 2.4.3 ジェトロの貿易統計処理システムの利用状況

#### (1) 利用状況の特徴

前述のように、ジェトロにおけるデータベースシステムは、利用者の立場を考慮して開発されたシステムである。では、その利用状況はどうであろうか。

ジェトロにおける貿易統計データの利用状況について、特に、貿易統計データの収集方針、利用者、利用者の特徴と傾向、及び利用状況の特徴と傾向等についての最近の調査結果をまとめると表2-5~表2-8のようになる。

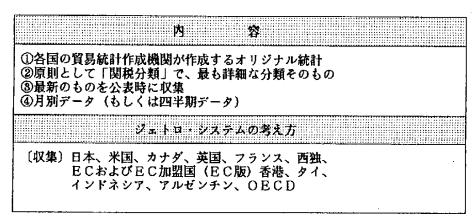


表2-5 貿易統計データの収集方法

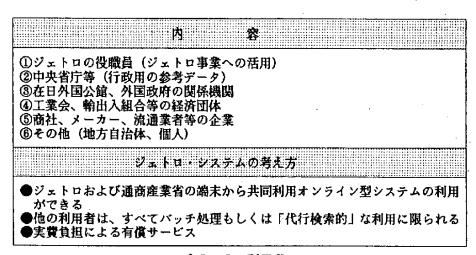


表2-6 利用者

# 内

①新しい情報需要層 国際化の進展により、従来「関心を持たなかった」者の利用が著しい

② 高年龄化傾向

特に、部下を持たない中間管理層からの需要が目立つ

③需要者の2分化

- ・統計(「海外情報」)に対する理解が充分でない層
- ・従来からの「情報に精誦」している層

## ジェトロ・システムの考え方

- ●ガイド画面/メニュー方式による検索利用オンライン型システムの利用
- ●メニュアルレスのシステム
- ■智勢を前提としない簡便な操作
- ●「ジェトロ」分類による統一管理、統一運用
- ©ガイダンス、コンサルティング的なレフェレンス ワークが不可欠

#### 表2-7 利用者の特徴と傾向等

#### 内 **2**

- ①ビジネス利用の必須条件
  - ・「最新」のデータ
  - ・「詳細(具体的)な品目分類」
- ②情報ニーズの背景
  - 「具体的な商談」「特定のプロジェクト」があり、そのための情報収集で ある場合が多い
- ③情報提供の有償化
  - ・使用目的に合致したデータを得るための経費負担
- ・加工度のレベル、アウトプットの様式に対する個別ニーズが強まる傾向 ④緊急ニーズへの対応(「即戦力」型)のアウトプット ⑤「情報発掘」型ニーズの急増
- - cf. 「データ検索」型ニーズ

#### ジェトロ・システムの考え方

- ●各国のオリジナル統計の収集
- ●各国ごと二個となる分類の管理
- ●SDI的な迅速なサービス
- ●利用者のパソコンででーた加工、グラフィック化できるように、電子メディ アで提供する
- ●個別ニーズに対応できるソフトウエアの整備
- ●「会議資料」、「稟議」「報告」等に直接使用できるプリント出力 ●「データ加工」のためのソフトウエア整備

表 2 - 8 利用状況の特徴と傾向

#### (2) 今後の課題

ジェトロでは、この外に通商産業省との間にオンライン情報交換システムにより、データベースの相互利用を図っている。

これにより、国内の情報をジェトロ自身で整備しないで利用できるメリットは極めて大きい。 さらに相互データベースの種類を増加させることを課題として取り組んでいる。

以上のことらジェトロにおける中長期的な主要課題をまとめると以下のようになる。

- ①海外事務所とオンライン化による情報ネットワークを整備すること。
- ②課外投に提供できるデータベースを開発すること。
- ③投資関連情報、企業情報、引合情報、海外生活関連情報などのデータベースを開発すること。
- ④OA化の進展の中で、データベース内容(検索結果)とOAソフトとの連携ができるシステムを早急に開発すること。
- ⑤商用データベースとして、社外に方に広く利用して貰える方法を確立すること。

#### 2.4.4 ジェトロにおける貿易統計データベースの工夫

ジェトロにおける貿易統計データの利用状況とその対応については、既に2.4.3 で述べたが、こうした個々の対策は、本調査研究の目的とかなり一致する点が多い。そこで、その具体的対応の例を最後に取り上げる。

#### (1)独自の商品分類の作成

貿易統計データベースの特徴として、各国ばらばらのデータに対応してゆかなければならない点が指摘されてきたが、これへの対応として、分類を探す仕組みとして、一種の細分類に関して特徴を抽出し、基準を作成し、分類表を整理した。

この例として商品分類名(履き物の例)を図2-27に示す。

#6 D A #5 #1		1	Τ_	T	(到品=如製品) No.
商品分類名	NBOC NO.	1	E	数昆耶位	Description
8. 第序120 - 各位形型名字 FT1 用11 4gg	KCOOO	0	0		Books Newspapers & Journals
—— <b>書師・</b> 小冊子	KC010	0	0		Printed Books.Brochures
新明、雑誌等の定期刊行物	KC020	0	0	NO	Newspapers. Journals and Periodicals
——散計图·图案	KC030	0	0	кс	Plans and Drawings
一広告、商業9901新	XC040	0	0	кс	Advertising Material.Commercial Catalog
	XC050	0	0	КG	Pictures, Designs and Photographs
d.181 & 40	KD000	0	0		Footveer
スポーツ用の履き物	K0100	0	0	DZ	Sports Footwear
スキー和	K0101	0	0	D Z	Ski-boots
― その他のもの	KD102	0	٥	DΖ	Other Sports Footwear
-ゴム,ブラスチョウ製(テニス川を除く)	XD103	0		DZ(XC)	Other Sports Footwear of Rubber Plastic
- 新製(テニス・ハスナット用を除く)	XD104	0		DZ(KC)	Other Sports Footwear of Leather
しずこれがわた。体操用のもの	KD105	0	×	DZ(XC)	Footwear for Tennis,Basket,Gymnastics
<b>一単刻のもの</b>	KD106	0	×	DZ(KG)	Footweer for Tennis,Basket of Leather
一助級用機維製のもの	KD107	0		DZ(KG)	Footwear for Tennis.Basket of Textile
<u></u> ++ンパスシューズ	K0108	0	×	DZ(KG)	Canvas Shoes for Tennis.Basket
一般き物(スポーツ川を除く)	KD200	o	0	D Z	Footwear (excl.Sports Footwear)
- ゴム・ブラスチック製のもの	X0201	0	0	D Z	Footwear of Rubber, Plastics
防水性長税	K0202	0	0	DΖ	Books of Rubber, Plastics
□□短靴	KD203	0	×	DZ ·	Shoes of Rubber, Plastics
<u></u> →サンダル	KD204	0	×	DZ	Sandals of Rubber.Plastics
	K0205	0	0	DZ	Footmear of Leather
ストラッフが指にかかるもの	kD206	0	0	D Z	Footwear of Leather, Strapped (Sandals)
□□短脫	10207	0	×	DΖ	Shoes of Leather
――- 紡織用繊維製のもの	RD208	0	0	DΖ	Foolwear of Textile
ーキャンパスシューズ 	kD209	0	×	DΖ	Canvas Shoes
L—-地下足袋	XD210	0.	×	DZ	"Jikalebi"
中敷き,ヒールクション等の物品。部分品	KD300	0	0	кс	Parts of Footwear, In-soles, Heel Cushions
· .			_		

【注】 KD100,KD200: 精出人の定義が一部異なる。 テニス、メスタラット、体境用の茶質の称は輸入ではスポーツ用風き物に(KD100) に輸出ではスポーツ用を除く風き物 (KD200) に分類されている。

図2-27 商品分類名(履き物の例)

また、こうした分類の定義をデーターベースに追加してゆくためのユーザインタフェースとして、BDCシステムでは、分類定義画面が用意されている。図2-28に示した分類定義画面例から分かるように、履き物の定義に関して「PARENT」、「BROT-HER」、「CHILDREN」といった概念により構造化を行っている。特に、分類に意味は持たせていないが、最初に分類を定義し、そこから構造化を行うアプローチをとっている。「PARENT」、「CHILDREN」といった分類関係は、例えば集合関係で含まれていれば、「子供」であり、また、相反する商品があり、その上に同じカテゴリーが存在すれば、「兄弟」と呼ぶといった考え方である。

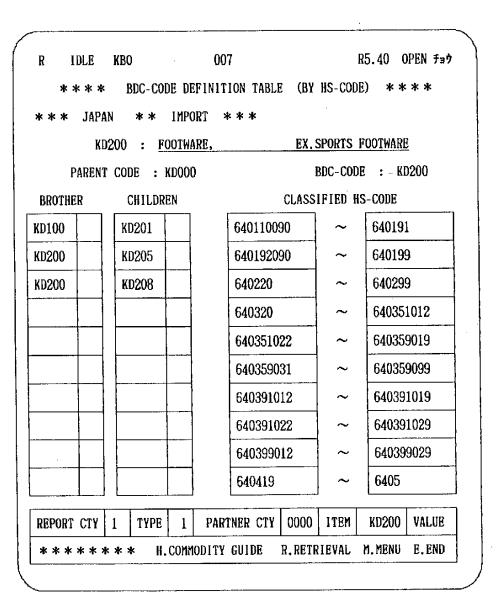


図2-28 BDCシステム 分類定義画面

このように、ジェトロの貿易統計データベースでは、貿易統計の特徴の一つである商品 分類の複雑さを、ユーザの立場に立った分類定義を用意し、また、先に述べたように使い 易いメニューを用意することによって、使用者層の「素人化」に対応している。

以上、本節では、ジェトロの貿易統計データベースについて、特に、「JETRO B DC」と「JETRO ACE」を中心に報告したが、今後のデータベース利用への人工知能技術利用にとって貴重な示唆を提供するものと言えよう。

## 2.5 EDMCエネルギーデータバンクの概要と課題

本節では、(財)日本エネルギー経済研究所エネルギー計量分析センターの「EDMC エネルギーデータバンク」について取り上げ、データバンクの概要、データ利用面からみ た課題等について報告を行う。

#### 2.5.1 EDMCエネルギーデータバンクの概要

#### (1) 表形式によるデータ収録

現在、EDMCエネルギーデータバンクには、統計情報を主とした約5万系列のデータ が表の形式で収録されている。データは国内と海外に大別され、国内データは650表、 海外データは550表である。

表のタイプとしては、時系列表、マトリックス表、文字表に分類される。

時系列表の例を表 2 - 9 に示す。検索結果の画面表示では、最新時点が表示されるよう 工夫されている。各表は、第1行に表名、出所、単位が示され、必要に応じて 2 行目以降 にコメント (この例では 2 ~ 5 行目) が付けられている。統計データを時系列に扱ってい ると、しばしば発生する統計の変更が共通の悩みである。そのような場合に、表形式でデ ータを収録していると比較的容易に対処できるものと思われる。

世界の石油生産量の推移/A A:原油・シェールオイル・オー B:1985年はユーゴスラビア C:1976年以前はその他西フ	「ルサンド・N C	BP/WE GLを含む。	10*3	B/D ·
C:1976年以前はその他西	アフリカに含まれ	<b>いる。</b>		
リング 1976年以前はその他アミ WLP04 1 アメリカ計 2 アメリカ(原油)	ファに含まれる。 1984 10505	1 9 8 5 10545	1 9 8 6 10250	1 9 8 7 9905
3 アメリカ(NGL) 4 カナダ 5 北アメリカ計 6 アルゼンチン	1645 12150 465	1815 12360 460	1840 12090 430	1910 11815 425
6 アルゼンチン 7 ブラジル 8 コロンビア 9 エメキシコ 10 メキシコ	460 170 255 3015	585 175 275 3015	595 300 280 2750	585 385 170 2870
11 トリニダード・トパコ 12 ベネズエラ 13 その他 14 ラテンアメリカ計	170 1875 270 6680	175 1730 270 5685	165 1845 250 6615	165 1765 235 6600
15 オーストリア 16 デンマーク 17 フランス 18 イタリア	25 45 60 40	25 60 70 45	20 75 75 50	20 95 75 60
19 ノルウェー 20 トルコ 21 イギリス	755 40 2580	835 40 2655	910 45 2665	1070 50 2605
22 西ドイツ 23 その他	80 140 .	80 150	80 155	75 145
24 西ヨーロッパ計 25 アプダビ 25 ドバク	3765 825 375	3960 890 390	4075 1090 390	4195 1200 420
28 イラク 29 クウェート 30 中立地帯	2195 1225 985 405	2215 1440 920 340	1905 1745 1250 345	2310 2090 1080 385
31 オマーン 32 カタール 33 サウジアラビア 34 その他	420 425 4760	505 340 3565	560 355 5150	570 345 4365
35 中東計 36 アルジェリア 37 アンゴラノC	240 11855 1075 185	295 10900 970 230	315 13105 1060 280	350 13115 1040 360
39 リピア 40 その他北アフリカ 41 ガボン	860 1105 Y 150	905 1105 Y	835 1045 Y	920 1000 Y
42 ナイジェリア 43 その他西アフリカ 44 アフリカ計	1385 440 5200	155 1475 450 5290	160 1465 465 5310	155 1290 460 5225
46 プルネイ 47 インド/D 48 インドネシア	10 160 560 1410	10 150 605 1335	15 165 625 1400	10 140 610 1345
50 その他東南アジア 51 その他アジア 52 オーストラリア	440 Y 85 550	435 Y 100 635	540 Y 115 610	510 Y 105 650
53 アジア・オセアニア計 54 自由世界計 55 中国 56 ソ連	3215 42865 2300	3270 42465 2515	3470 44665 2630	3370 44320 2680
57 57 58 58 59 共産圏計 60 その他東半球計	12450 Y 485 15235	12150 Y 465 15130	12560 Y 470 15660	12740 Y 460 15880
61 62 世界計 63 OPEC計	58100 18470	57595 17215	60325 19555	60200 19030

表2-9 時系列表の例

石油	製品の産業別販売量(FY19	987)	通産/エネ生需	10#3KL			
J1 1 2 3 4 5 6 7 8 9 10	20387 20387 元本 元本 元本 元本 元本 元本 元本 元本 元本 元本	灯油 26411 2303 115 42 1274 4221	軽油 29588 1552 492 238 3221 2080	A 重油 22341 1668 3976 676 6845 1431 533 197 984 1113	B重油 1817 59 7 12 651 199 90 25 56 78	C重油 41283 87 104 47 25 17709 1200 2434 4303 5037 1838 1598	潤滑油 2131 32 38 4 45 1130
12	(納納金) (納納金) (納納金) (納納金) (納納金) (納納金) (納納金) (納納金) (納納金) (納金) (納金) (納金) (納金) (納金) (納金) (加金) (	764 234 280 19 230 182	12832 666 11598 155 413 213 8961	552 357 1679 2859 132 210 2352 165 244 6027	77 28 99 587 16 17 541 13 65 430	1598 363 937 2514 365 119 1994 37 18486 2311	256 9 81 159 8 9 617

表2-10 マトリックス表の例

石油代替エネルギーの供給 開発及び導入を行うべき 及びその種類ごとの供給 JY15	数量の目標(1987年1	〇月16日 閣議決定) の供給数量の目標(万k	(1) 備考	
石炭	10,100	34.2%	石炭の供給数量は、13,6	00万トンである。
原子力	8,600	29.1%	原子力の供給数量は、原子力数量であり、原子力発電に係る50万キロワット、年間発 億キロワット時である。	発電による電気の供給 る施設の出力は、5, 電電力量は3,510
天然ガス	5,800	19.6%	天然ガスの供給数量は、輸入 供給数量3,800万トンと る天然ガスの供給数量42億 数量である。	本邦において生産され
水力	2,400	8.1%	水力の供給数量は、一般水力 数量であり、一般水力発電に 450万キロワット、年間発 ロワット時である。(なおれ 施設の出力は2,100万キ	切がるがりがりをおいなる
地熱	4 4 0	1.5%	地熱の供給数量のうち、地熱 係る施設の出力は260万キ 力量は160億キロワット時	ロリット 任間数係録
その他の石油代替 エネルギー	2,200	7.4%	その他の石油代替エネルギー	とは、太陽エネルギー
(参考) 合計	石油換算 3.0億k	1 100.0%	、石炭液化油等をいう。	:
			•	

#### (2) 簡単な操作で必要な表の検索

EDMCエネルギーデータバンクにおける検索は、メニュー選択方式で簡単に行うことができる。図2-29はメニュー画面の例を示したものである。

	中二・ナッツ しょと 海和し ツエキハ ・・・・	{	
	表示するデータを選択して下さい。・・・		
	PF12キー を押すと、海外・国内選択2	ィニューに戻りま	す。
	PF3キー を押すと、前の画面に戻りる	<b>きす。</b>	
	. •	出所	期間
01			
02	世界の原油確認埋蔵量の推移	OGJ	70 - 87
03		WO	70 - 86
04		wo	60 - 87
05		ogJ	76-86
06		OGJ	$\frac{7}{10} - 8\frac{7}{10}$
07		PIW	78 - 87
08		WO	70 - 86
09		ВP	$38 \cdot 87$
10	世界のNGL生産量の推移	UN	62 - 85
11		UN	62 - 85
12		UN	62 - 85
14		ЙЙ	62 - 85
15		ВP	5.7 - 8.7
16		UN	82 - 85
17		BP	57 - 87
18	3 世界の石油製品生産量(合計)の推移	UN	62 - 85

図2-29 データバンクを検索するためのメニュー画面例

このようなメニュー画面の中から必要とする表の番号を選択するだけで検索ができるようになっている。また、画面に表示されていない部分のデータをスクローリング機能により容易に見ることができる。

#### (3) 単位換算、前年比、期別変換等の加工

図2-30と図2-31は、単位換算の例である。エネルギーデータの場合、種々の単位が用いられ、重量間での換算、容積間の換算、さらには重量・容量から発熱量への換算というニーズがきわめて高い。そのため、簡単な操作で単位換算を行う機能が用意されている(この例では8のKLへの換算を指示している)。

		•			
	単位	位換算/加〕	Lの種類は		8
	1 2 3 4	Kcal BTU Theri kwh	換算	1 6 1 7 1 8	オリジナル表示 前期比 指定行に対する構成比
	4 5 6 7 8	Joul 石油lT 石炭lT KL	換算 換算 換算 換算	1 9	100%とする行は 最右列に対する構成比
	9 1 0 1 1	CF 米ガロン <b>英ガロ</b> ン	換算 換算 換算	2 0	を乗じる
	1 2 1 3 1 4 1 5	BL B/D CF/D CM/D	換算 換算 換算 換算		
. ]		-		PF.12%	は海外・国内選択メニューに戻ります。

図2-30 単位換算の例

			· · · · · · · · · · · · · · · · · · ·	
世界の原油生産量の推移		OGJ/		
WPL01	1984	1985	1986	1987
1 アフガニスタン				X
2 オーストラリア	27912	32264	27390	3284
3 バングラディショ				5
4 ブルネイ	9285	8704	9865	777
5 ビルマ	1741	1741	1741	139
6 インド	31510	35630	36152	3534
7 インドネシア	77295	70796	72188	6882
8 日本	348	580	754	69
9 マレーシア	26809	25127	29189	27 <u>10</u>
10 ニュージーラント	1045	1045	1683	162
11 パキスタン	1045	1973	2379	243
12 フィリピン	696	522	348	29
13 タイ	1103	1973	2089	179
14 台湾	174	174	116	17
15 アシア・オセアニア計	178962	180471	183894	18029
16 オーストリア	1335	1277	1277	127
17 デンマーク	2611	3366	4352	545
18 フランス	2031	2843	3424	371
19 西ドイツ	4584	4700	4642	429
20 ギリシャ	1562	1509	15 <b>09</b>	139
21 イタリア・	2611	2437	2843	290

図2-31 単位換算の例(換算後のデータ表示)

また、時系列短期データの場合、データの期種統合が可能である。月次および四半期データは、期種統合により、四半期、半期、年/年度データに加工することができる。データが表示されている画面でPF12キーを押すと前述の単位換算/加工メニューと同一画面に期種統合メニューが表示され、ユーザはそこで選択する。この図2-32と図2-33では、月次データの表に対して、21を指定して四半期データに統合した例を示している。

単位	立換算/加.	この種類は	t · · ·	2 1
1	Kcal	換算	16	オリジナル表示
2 3	BTU	換算	1 7	前期比
3	Theri		18	前年同期比
	k w h	換算	Ĩ Š	指定行に対する構成比
<b>4</b> 5	Joul	換算		100%とする行は
Ğ	石油1丁	換算		100,000
Ž	石炭1丁	換算	2 0	最右列に対する構成比
Ŕ	Κ̈́L	換算	2.1	月から歴年四半期への統合
8 9	CF	換算	2 2 2 3 2 4 2 5	月から年度四半期への統合
1 Ŏ	米ガロン	換算	23	月から年度半期への統合
īĭ	英ガロン	換算	2 4	月から年度半期への統合
î ē	BL	換算	25	月から歴年への統合
$\tilde{1}$ $\tilde{3}$	$\tilde{B}/D$	換算	26	月から年度への統合
1 4	CF/D	換算	20	717.0 - G
$\hat{1}$	CM/D	換算	2 7	を乗じる

図2-32 データ期種統合の例

石油製品の生産量	通産/エネ生需				
JSL02 1 燃料油計 (千KL)	8 7 Q 3	8 7 Q 4	8 8 Q 1	8 8 Q 2	
2 揮発油 (千KL)	36459	39977	41637	35941	
	9078	9078	8113	8994	
3 ナフサ(千KL)	1878	2169	2844	1839	
4 ジェット燃料(千KL)	1130	1001	943	920	
5 灯油 (千KL)	3843	5245	6490	4281	
6 軽油 (千KL)	6249	6724	5730	6167	
7 重油計 (千KL)	14280	15758	17516	13733	
8 A重油 (千KL)	4307	5468	6853	5028	
9 B重油 (千KL)	415	478	481	344	
10 C重油 (千KL)	9558	9812	10183	8361	
11 潤滑油(千KL)	577	599	589	606	
12 アスファルト(千KL)	1432	1592	1555	1317	
13 グリース(千KL)	16891	18713	17332	18591	
14 パラフィン(千KL)	33844	32946	34169	31913	

図2-33 データ期種統合の例(統合後のデータ表示)

## (4) ダウンロードによるデータ利用の多様化

日本エネルギー経済研究所エネルギー計量分析センターでは、利用者のニーズの多様化 に対応するために、88年6月より、ダウンロード機能を設けることにより、パソコンを 利用した機能拡張を実現している。このパソコン機能の利用により、グラフ作成、公表さ れた情報と自社情報の加工等への利用が可能となっている。主な機能を整理すると以下よ うになる。

- ●グラフ作成
- ●独自データとの組み合わせ、再編、加工
- ●回帰分析等の統計解析

#### 2.5.2 データ利用方法からみた課題

図2-34は、EDMCエネルギーデータバンク利用の基本的な操作手順を示したものである。

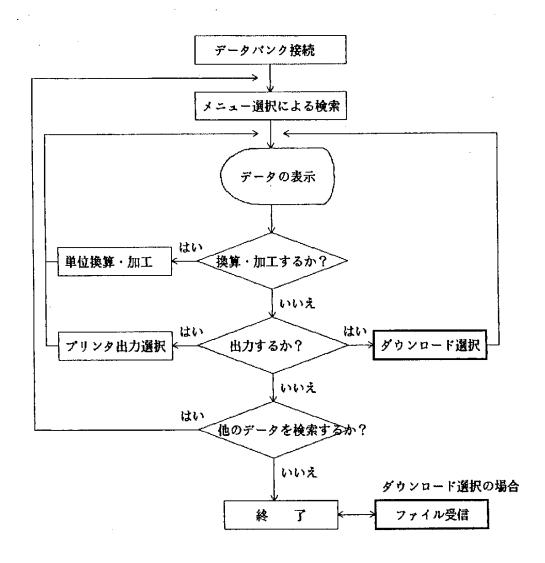


図2-34 EDMCエネルギーデータバンク利用の概念図

操作性を上げるためのメニュー化、表イメージのハンドリングは、一方で新たな課題を 引き起こしており、主な課題を上げると以下のようになる。

- ●必要とする項目(系列)がどの表に収録されているか分かりにくい
  - ・表の形でデータが入っていることから操作性は良い。しかし、必要とする表がどこ にあるか、また、どの表に必要とする情報が入っているかが分かりにくい。

## 2.5.3 今後の課題

以上、EDMCエネルギーデータバンクの特徴について述べたが、前述したようにEDMCエネルギーデータバンクの大きな利点は、データが表イメージで収録されていることから、専門的に利用する人にとっては非常に使い易いデータベースであること、また、ダウンロードによりパーソナルコンピュータ用の市販統計ソフト(ロータス1・2・3やマルチプラン等)を有効に活用できる点などが指摘されよう。

また、ユーザインタフェースとして、メニュー選択方式を採用し、操作を簡易化している点も利用者サイドに立ったシステムと言えよう。

最後に、人工知能技術の適用の視点から考察すると、データが表イメージで入っている 点が注目され、この表の構成を決定しているところに利用者に有効な知識が隠されている ことが指摘できよう。そして、必要とするデータと収録されている表との関連性を明確に する知識をシステム化し利用者をサポートすることも、今後の機能向上につながると言え よう。



# 第3章 データベース検索への人工知能 技術利用

#### 第3章 データベース検索への人工知能技術利用

本章では、データベースの利用における人工知能技術の利用について、特に検索システムの高度利用を中心に取り上げ、応用事例としてAI化の研究が盛んな文献情報検索の事例と、検索システムへの自然言語処理の応用等について報告する。

また、経済統計データベースへの人工知能技術の利用について、本調査研究の中で検討した「知識ベース化」について報告する。

#### 3. 1 文献情報データベースと人工知能技術

データベース利用技術の高度化は、経済情報データベースといった数値情報のデータベース利用に限らず、文献情報の検索システムの分野でも進められている。そこで、本節では、文献情報検索への人工知能技術の利用動向と事例について報告する。

#### 3.1.1 文献情報検索への人工知能技術利用の動向

現代の図書館の利用は、ユーザニーズが文献情報調査からファクト情報要求へと多様化してきていることから、単なる書庫という機能から知識情報を提供する巨大な文献情報提供システムへと変貌し始めている。こうした現状に対して、図書館には、単に図書を収集するだけでなく、情報を加工して二次情報及びファクト情報を提供することが期待されている。

また、図書館自体の役割も、レファレンサーにより活性化された資料情報と外部の専門知識とを結合し、ユーザ向けの情報利用環境を提供することが期待されている。

このような状況に対処するため、情報処理技術の利用は必要不可欠となっており、特に 最近のデータベースとネットワークの普及は、大きなインパクトとなっている。例えば、 図書検索用カードは、書誌データベースに変わり、図書館以外の事務所、書店、街頭、さ らには家庭においてオンライン・データベース検索を可能にしてきていることがあげられ る。さらに、この検索システムからは、書誌情報だけでなく、文献の要約やフルテキスト 等の出力を可能にしてきている。それ故、検索システムに、図書館の窓口、レファレンサ ー、読書相談者の役割を機能化する要求がされるのである。

このように、文献情報の検索における多様なシステムの機能化が盛んに進められているが、最近の傾向として次項のような問題点が指摘されている。

検索システムは専門家用に開発されていることから、一般のユーザが操作するには、

#### 問題点一

- ① 検索システムの使い方を熟知しなければならない。
- ② 自分の検索目的にあった検索式を表現することができない。
- ③ 検索したい対象分野の知識がないと検索できない。
- ④ 利用者の知識水準、興味の範囲、目的等が無視されてしまう。

#### と言った点があげられる。

以上の問題点を解決する方法として、検索システムの研究分野では、「カードイメージ 画面」、「メニュー方式による入力支援方式の研究」、「シソーラスの改良」、「類義語 の処理能力の向上」等の研究が進められている。さらに人工知能技術の利用の観点から、 情報検索に必要な知識を一種のエキスパートシステムとして捉え、検索のための知的コン サルテーション・システムの構築により、誰にでも情報検索が容易にできるシステム作り が進められている。

#### 3.1.2 データベース検索への人工知能技術利用事例

人工知能技術の応用事例として、知的図書館システム「ILIAS」の事例を取りあげる。

このILIASは、一般の図書館ユーザを対象として開発された文献検索システムであり、ユーザとしては、大学の教養課程の学生を想定し、レーポートテーマ決定のためのコンサルテーション機能と検索機能とを結び付けることによって一般の図書館ユーザ向けの検索システムのあり方を探るために開発されたプロトタイプシステムである。

ILIASの特徴は、ユーザの個人情報(利用目的、興味、知識水準)の利用、ユーザに対する検索過程での情報提供を重視している点にある。具体的には、以下のようなところにAIの要素技術が活用されている。

①ユーザの興味、知識水準を判断し、検索戦略、検索の範囲、検索結果のサービス方法 を支援する。

- ②対象分野の知識に関する質問応答サービスを行う。
- ③ユーザが入力した日本語から、検索情報としてキーワードを自動的に抽出し、概念木 を作成する。
- ④不十分な検索情報については、不足している情報は、推論、連想して推測し、入力された情報が相互に矛盾していたり、意味的におかしい場合は排除するといったことを自動的に行い補足する。
- ⑤検索のための最適な入力情報をユーザの興味、知識水準を考慮し、キーワード間の意味チェックを行うことにより推薦する。

ILIASの利用によって、ユーザはレポートテーマを入力するだけで良く、ユーザの 興味や目的はレポートテーマという具体的な形で表現され、ユーザはシステムの助言によってテーマを選択し、参考文献に到達できる。これは従来型検索システムのキーワード入力に対応している。

従来の検索システムとILIASの概念的な相違は図3-1のようになる。

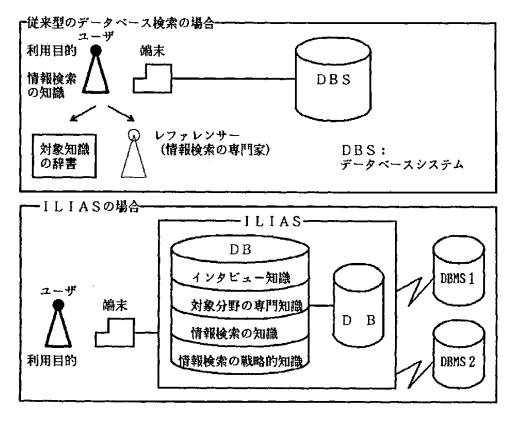


図3-1 従来システムとILIAS

ILIASは、人工知能の要素技術であるルールやフレームを用い、情報検索に必要とされる諸知識を知識ベース化することによって、ユーザは情報検索の知識に煩わされることなく検索を実行できる利点がある。また、このシステムでは、将来的に、音声認識や映像によるイメージ出力等のマン・マシン・インタフェースを装備することによって、よりユーザフレンドリーなシステムを目指しておりその実現が期待される。図3-2に1L-IASの実行例を示す。

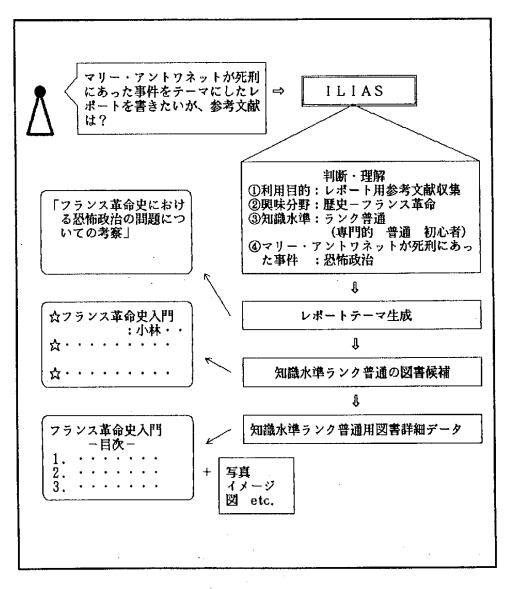


図3-2 ILIASの実行例

#### 3. 2 データベース検索と自然言語処理技術

データベース利用への人工知能技術の応用において、エキスパートシステムといった知識ベースの利用以外に、検索時のマン・マシン・インタフェースの改善に、自然言語を利用する方法の期待が高くなっている。この自然言語入力を可能にする人工知能技術として「自然言語処理技術」がある。

本節では、このデータベースへの自然言語処理技術の利用例として、「DB/I」(DB mate)という自然言語インターフェースを取り上げる。さらに、自然言語による問い合わせと統計データベースの事例について報告する。

#### 3.2.1 自然言語処理インタフェース「DB/I」の概要

#### (1) 自然言語処理インタフェース

データベースを操作するためには、そのためのコマンドを知らなければならないことから、初心者や普段データベースを操作していない人にとってはマニュアルと格闘することになる。また、コマンドは通常、英語の構文に従っていることから、それだけで嫌悪感を抱く人も多い(以上のことは、データベースに限らず、多くのコンピュータシステムに共通したことである)。

これらの問題を解決し、より多くの人がデータベースを利用できるために考えられたのが、「自然言語(日常の日本語)」を使う方法である。自然言語を利用すれば、マニュアルを読んで覚える必要もなく、簡単にデータベースを操作することができる。また、情報要求内容を思ったまま(コマンドに変換することなく)入力すればよいので、利用者の思考に合った操作が可能になる。

ここで、問題になる点として、「日本語文をキーボードで入力するのは大変だ。」といった指摘もある。これは、自然言語入力では実際に手間がかかるからである。これに対して、確かに一般的なメニュー方式などの方が優れている面も多いことは事実であろう。入力が簡素であるからである。しかし、メニュー方式では決められた手順に従わなければならず、時には回り道をしたり、イライラさせられるケースも多い。つまり、自然言語を利用する理由としてここで言及しておきたいことは、「単にコマンドを日本語文に変えるだけでなく、人間の思考に合った操作性を提供できる」という点である。

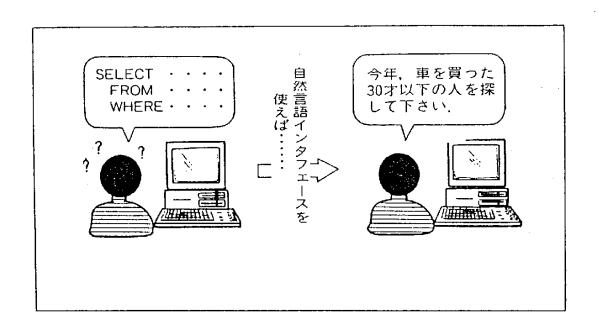


図3-3 自然言語処理インタフェースとは

#### (2) 「DB/I」の仕組み

上記のような理由から、開発された「DB/I」の仕組みについて説明する。図3-4 は、自然言語でデータベースを操作するための処理概要の例である。この図の中で、点線で囲まれた部分が「自然言語処理インタフェース部」である。入力された自然言語文は、自然言語解析部で、構文解析、意味解析が行われ中間表現に変換される。その後、中間表現からデータベースの操作コマンドが生成される。このコマンドはデータベースシステムに渡され、実際の操作が実行される。

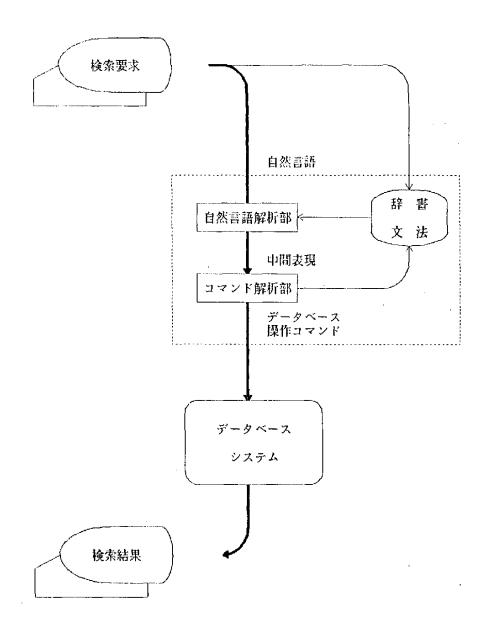


図3-4 自然言語処理インターフェースの概念図

この「DB/I」では、対象データベースとして、パソコン上のデータベースシステムとして利用されている「dBASEII(日本アシュトンテイト社)」に対応している。しかし、DB/Iは、データベースシステムと自然言語インタフェース部を切り離していることから、他のデータベースシステムへの対応も可能である。

## (3) 「DB/I」の特徴

次に、「DB/I」の特徴について説明する。「DB/I」の最大の特徴は、日常の日本語でデータベースの操作が行えることである。もちろん、現在の技術では、日常語といっても構文的に制限されたものになってしまうが、コマンドを覚えるほどの苦労はしないで習得できる。

一つの日本語文は、単に一つのコマンドに変換されるのではなく、コマンドの列に変換される(図3-5参照)。

dBASEでは、一連の操作が複数のコマンドになる場合が多い。例えば、最も基本的な操作である検索にしても、FINDコマンドだけでは単にポインタを移動するだけであり、画面に表示させるにはDISPLAYコマンドを使わなければならない。素人にとってこのことは、マニュアルを一読しただけではなかなか分からない。しかし、このような処理も「DB/I」を利用することにより、意識することなく実行できる。

```
1: SET TALK OFF
 2: SET DELETE ON
 3: ?"****** 東京在庫 ********
 4: SELECT 1
 5: USE 製品台帳
 6: INDEX ON 製品番号 TO 製品台01
 7: SELECT 2
 8: USE 東京在庫
 9: SET RELATION TO 製品番号 INTO 製品台帳
10: LOCATE FOR (製品台帳-)製品名 = 「ホッチキス")
11: IF . NOT. EOF()
12: LIST 製品台帳-)製品名 , 在庫 FOR (製品台帳-)製品名 = *ホッチキス*)
13: ELSE
14: ?****** 条件に台うデータはありません。 *******
15: ENDIF
16: SET RELATION TO
17: ?"******* 大阪在庫 ********
18: SELECT 1
 19: SELECT 3
20: USE 大阪在庫
21: SET RELATION TO 製品番号 INTO 製品台帳
22: LOCATE FOR (製品台帳-)製品名 = "ホッチキス")
 23: IF . NOT. EOF()
 24: LIST 製品台帳->製品名 、在庫 FOR (製品台帳-)製品名 = "ホッチキス")
 25: ELSE
 26: ? ***** 条件に合うデータはありません。 *******
 27: ENDIF
 28: SET RELATION TO
```

図3-5 変換結果の例

29: SET INDEX TO 30: SET TALK ON

## (4) 曖昧な検索表現への対応

一般に、検索を行うためには、ファイル名、フィールド名(項目名)、条件(キーワード)等を指定しなければならない。しかし、「DB/I」では、ファイル名やフィールド名を省略した検索も可能である。例えば「○○名簿の中から所属部門が営業の人を探して下さい。」という代わりに、「営業の人は?」という入力で十分である。この時、「DB/I」はファイル名やフィールド名を自動的に導き出してくれる。また、一つに決定できない時は候補を表示してくれることから、後は指示された通りに答えていけばよい。

さらに、システムが提供している以上に柔軟な検索を行いたい時は、辞書登録の機能を 利用できる。この機能を用いて同義語を多数登録しておけば、同じ意味の検索をいろいろ な形で入力できるようになる。

以上、「DB/I」の機能について簡単に説明してきたように、「DB/I」の開発目的は、計算機にあまり馴染みのない人にとって使い易い利用環境を実現していることである。そして、システム全体としては、まだまだプリミティブなものであり、改良点も多いと思われるが、こうした「自然言語処理技術」の応用によるデータベース利用の高度化は今後最も期待される技術の一つであろう。

#### 3.2.2 自然言語処理技術と統計データベースの関連について

人工知能技術のデータベースへの利用について、特に文献情報の検索への利用、自然言語処理技術の事例等について報告してきたわけだか、ここで、本調査研究の課題である統計データベースの利用に関して視点を戻し、大阪大学の佐藤助教授らの開発した自然言語で統計データ検索を行うシステムー国土情報辞書システム (Land Information Dictionary System)「LIDS86」について簡単に紹介する。

## (1) 自然言語による統計データベース要求文

自然言語によるデータベース検索の入力文は、用語の説明の要求などの補足的な入力文 もあるが、その中心となるものは、検索すべきデータを指定する文(以下、「データ要求 文」と呼ぶ)である。

例えば、「昭和50年の大阪と京都の性別人口データが欲しい」という文章は、データ 要求文の一例である。 (図3-6(a)参照)

図3-6(b)は、(a)で要求しているデータの構造をスキームの形式で表したものである。 これは、「時点カテゴリが(昭和50年)に制約され、地域カテゴリが {大阪 京都} に 制約された性別の人口」という要求内容を表している。

#### (a) データ要求文

## (b) 要求データのスキーム

\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	時点	地域	性	人口
要求データ	{昭和50年}	{大阪 京都}	d.性	0

図3-6 データ要求文と要求データのスキーム

出典元:統計データベースの設計と開発

このようなデータ要求文を解析する目的は、(a)のようなデータ要求の入力文を解析し、(b)のような要求データのスキーム情報を抽出することにある。

## (2) 統計データ要求文の解析

図3-6(a)の問い合わせ文は、自然言語処理による構文解析によって、図3-7のように解析される。この解析結果は、図3-6(b)の要求スキーマに直接対応している。図3-7の3つの<分類属性句>は、図3-6(b)の {昭和50年} , {大阪,京都} , {d.性} にそれぞれ対応し、図3-7の<統計属性リスト>は、図3-6(b)の「人口」の要求に対応している。従って、「昭和50年」というカテゴリが「時点」カテゴリであり、「大阪」と「京都」というカテゴリが「地域」カテゴリであり、「人口」という統計属性が「人て」という記述対象型に関するものであることがわかれば、図3-6(b)の要求データのスキーム情報を生成することができる。

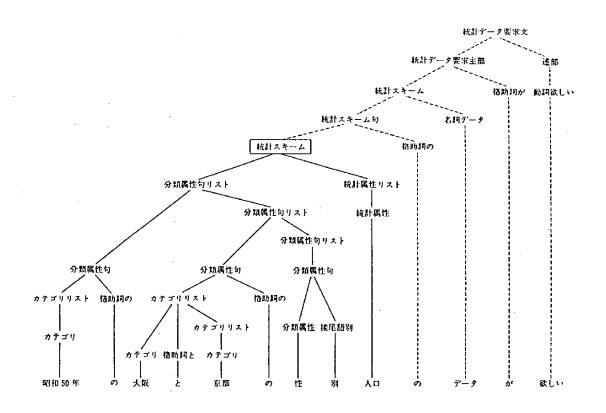


図3-7 統計データ要求文の解析木

出典元:統計データベースの設計と開発

## (3) システムの構成

図3-8は、国土情報辞書システムLIDS86のシステムの全体構成である。以下、 簡単にこの構成について説明する。

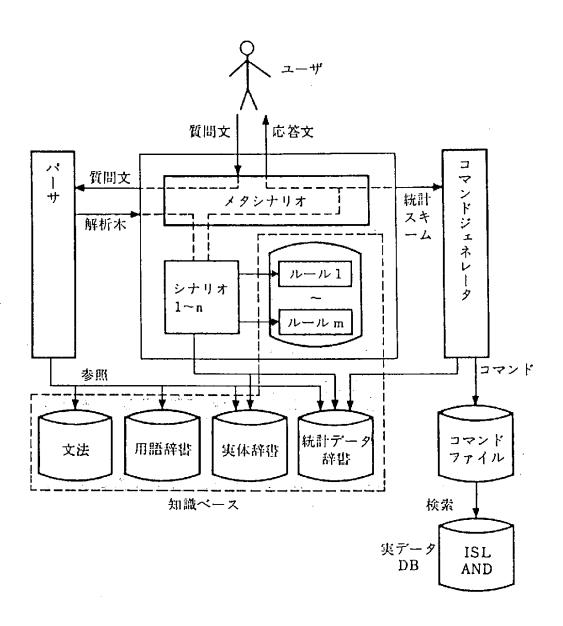


図3-8 国土情報辞書システムの構成

出典:統計データベースの設計と開発

- ①メタシナリオ・・・・メタシナリオは、システムの全体の制御を行うプログラムである。メタシナリオは、ユーザから入力を受け取ると、それをパーサに渡し、その解析結果を基にシナリオセットから適当なシナリオを選択し、起動する。
- ②パーサ・・・・・パーサは、入力文の自然言語解析を受け持つプログラムである。このシステムのパーサでは、意味文法が使用され、形態素解析、構文解析、意味解析に3者を同時に行うようになっている。入力文は、この意味文法に従ってボトムアップに解析され、その結果は解析木の形でメタシナリオに渡される。
- ③シナリオセット・・・シナリオセットは、シナリオと呼ばれるプログラムの集合体である。シナリオは、質問応答の場面ごとに作成されており、ユーザからの入力に対してはシナリオがどのように振る舞うべきかが記述されている。シナリオの選択は、メタシナリオが解析木と文脈とから判断して決定される。
- ④知識ベース・・・・知識ベースは、文法知識、辞書、ルールで構成されている。
- **⑤**コマンドジェネレータ
  - ・・・コマンドジェネレータは、統計データベース検索シナリオから 呼ばれる。このシナリオにより、ユーザが求めているデータに 関する情報を受け取り、それに対応するデータをデータベース より検索するためのコマンドに生成する。

## (4) 統計データの自然言語検索の例

上述の機能によってシステムが実行された場合の検索例の一部を図3-9に示す。この図は、ユーザが「昭和50年の大阪と京都の性別の就業者数」というデータの要求を入力した場合の例である。入力文は、パーサにより解析され、図3-7のような解析木がメタシナリオに返される。メタシナリオは、解析木の構造から、統計データ要求文の入力であることを判断し、統計データ辞書検索を行うシナリオを起動する。

このシナリオでは、まず、構文解析結果の中にパーサでは解消できないような文意の曖昧さが残されていないかどうかを確かめ、残されていた場合には、その曖昧さを解消するためにユーザに質問を行う。図の中の「大阪は大阪市のことですか?」という質問が、そ

の例である。

このような曖昧さを解消した後、シナリオは、図3-6(b)のような要求データのスキーマを作成し、統計データ辞書に収録されている概念ファイルのスキーマ情報とマッチングを行う。次に、マッチした概念ファイルに属するDBファイルのスキーマ情報を取り出しユーザの要求データスキームとの一致具合を調べる。候補にあったDBファイルの全てについてスコアリングを施した後、スコアの高い順にDBファイルの内容を日本語に変換してユーザに示す。図中の「<<就業者数に関するデータ>>」以下はその例である。

ここでは、「ただし、性別分類別には観察されていません」と表示されているように、 ユーザ要求の「昭和50年の大阪の性別の就業者数」のうち、「昭和50年の大阪の就業 者数」は求められるが、「性別」は得られないことを示している。

このように要求に合致するデータが存在しない場合でも、概念ファイルーDBファイルの区別と上記のスコアリングとにより、近傍の解を示すことができる。

ユーザがそのDBファイルに満足すれば、コマンドジェネレータにそのDBファイルの情報を送り、DB検索コマンドを生成する。図中の下端は、生成されたリレーショナルDB検索コマンドの例である。

LIDS: ようこそし I D S 8 6 へ。

必要なデータは何ですか?

USER:昭和 5.0 年の大阪の性別の就業者数のデータが欲しい

LIDS: 大阪は、大阪府のことですか?

USER:いいえ

LIDS: 大阪は、大阪市のことですか?

USER: Ill.

L105: <<< 就業者数に関するデータ >>>

調査時点 1 0 別市区町村(会県合計)別産業大分類 A 別就業者数というデータ があります。但し、性別分類別には観察されていません。

このデータでよろしいですか?

USER:他には

LIDS: <<< 就業者数に関するデータ >>>

調査時点 1 0 別市区町村 (含県合計) 別職業大分類別款業者数というデータがあります。但し、性別分類別には観察されていません。

このデータでよろしいですか?

USER:産業大分類Aとは何か

LIDS: 産業大分類Aは次の様なカテゴリを含みます。

- 全産業
- 農業
  - 林粜狩猟粜

- **集 版 \***
- 拡棄
- \* 建股泵
- 製造業
- \* 卸売・小売業
- 金融保険業
- \* 不動産業
- \* 運輸通信業
- \* 電気ガス水道
- サービス菜
- \* 公務 。
- \* 分類不能の産業

USER:はじめのデータが欲しい

LIDS: <<< 就業者数に関するデータ >>>

調査時点 1 0 別市区町村(含県合計)別産業大分類 A 別航業者数というデータ があります。但し、性別分類別には収察されていません。

\_ このデータでよろしいですか?

USER: IL IN

GET INTO T164644

(時点,地域区分,産業分類,就業者数)

JITEM.TIIKI.SANGYO,SHUGYOSH

FROM LIDS#LIDS154

AND TIIKI=27100 ··········「大阪市」のコード

— 生成されたコマンド ——

# 図3-9 統計データの自然言語検索の例

出典:統計データベースの設計と構築

以上簡単にシステムの説明を行ったが、本システムは統計データの検索において、世界 が限定されているので自然言語処理技術が応用可能である貴重な事例であると言える。

### 3. 3 統計データベースの検索と人工知能

前節まで、データベースと人工知能技術の事例について自然言語処理技術を中心に述べてきたが、、本節では、今回の調査研究の目的である経済統計データベース利用の高度化というテーマに基づき、既に第2章で報告されたデータベースの事例に対して、人工知能技術がどのようにアプローチ可能かを中心に報告する。

特に、本節では、「政策情報システム (PPIS)」の検索利用を例に人工知能技術の 適用についての検討事項について説明する(また、その概念をさらに詳細に検討し、ケー ススタディを行った結果については、第5章で述べる)。

### 3.3.1 経済統計DB検索利用のための知識ベース化

## (1) 作業の目的

経済統計DB検索利用のための知識ベース化の作業の目的は、経済統計の検索に対し、 人工知能技術の一つである知識ベースの構築に的を絞り、既に第2章で報告した「政策情報システム(PPIS)」の中の「国内マクロ経済情報(SDB)」を参考に、既存の統計DBのIndexファイルの知識ベース化を試みることである。

## (2) 考察の方法

ここでの知識ベース化の試行における方法としては、以下の2点がある。

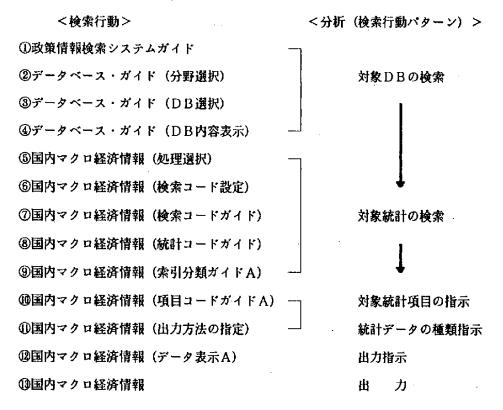
- ① 「政策情報システム (PPIS) 」の利用例を基に、統計DBに対する利用者の検 索行動を一例として整理する。
- ② ①の検索行動に対して、第2章で報告された「INDB」における「経済統計の意味連関ネットワーク」等の事例を基に、統計DBの、特にIndexファイルの知識ベース化についてシステム設計の視点から考察する。

#### (3)検索行動の分析

第2章の「政策情報システム (PPIS)」の「国内マクロ経済情報 (SDB)」について、メニュー選択によるガイドシステムを利用した場合の検索行動と、その検索行動パターンを分析した結果を図3-10に示す。

この分析から分かるように、検索行動は、「対象DBの検索」、「対象統計の検索」、 「対象統計項目の指示」、「統計データの種類指示」、「出力指示」、「出力」といった

### 6つのパターンに分類できる。



## 図3-10 検索行動の分析(SDBのガイドシステム利用の場合)

#### (4) カテゴリー化 (Categorizaton)の例

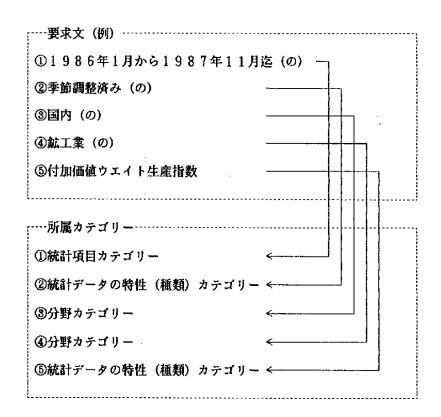
ここで、検索例として、「1986年1月から1987年11月迄の季節調整済みの国内の鉱工業の付加価値ウエイト生産指数が知りたい」といった場合を分析し、カテゴリー化を試みる。

カテゴリーとして

分野カテゴリー
 統計項目カテゴリー
 ボナ言葉の集まり
 統計データの特性カテゴリー
 データベースのデータを構成する項目を示す言葉の集まり
 統計データの特性カテゴリー
 データベースのデータの種類を示す言葉の

集まり

の3種類を設定し、上記要求文を構成する単語と各カテゴリーとの対応付けを行う。



このように、要求文は、各カテゴリーに分類することができる。そこで、これらの各カテゴリーごとに知識ベース化を行うことによって、統計DBに対する Indexファイルの知識ベース化が可能になると考えられる。

## (5) 統計DBに対するIndexファイルの知識ベース化

図3-11は、統計DBに対するIndexファイルの知識ベース化に関する概念図である。図中の番号が、上記の各カテゴリーの内容を示している。

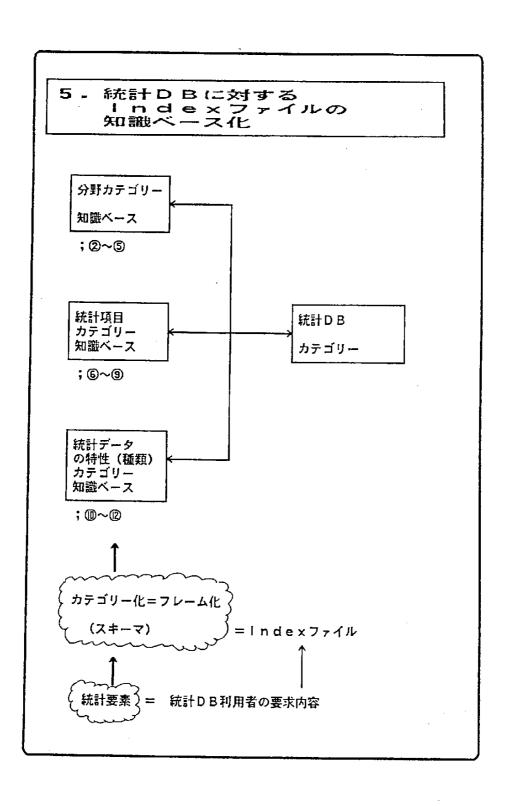


図3-11 統計DBに対するIndex ファイルの知識ベース化

### (6) 知識ベース化について

#### ①分野カテゴリー

SDBでの検索における要求文を分析することによって、分野カテゴリーは、「どうなる」「どうする」という現象及び「何を」にあたる対象の2種類のスキーマに分けることができる。

(a)現象スキーマ・・・・動詞のフレーム化 (例) 取得、処分 etc. (b)対象スキーマ・・・・名詞のフレーム化 (例) 株式、債券 etc.

さらに、これらによって構成される知識ペースの構造に関して、対象となる株式及び債券は取得と処分という2つの現象に関連づけられる項目であるので、その関連を知識ペース化することが可能である。それ故、以下のような例が考えられる(図3-12参照)。

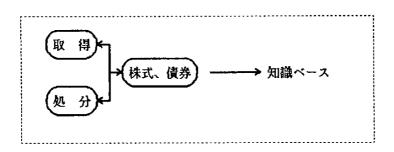


図3-12 知識ベース構造例

#### ②統計項目カテゴリー

統計項目カテゴリーに関しても知識ベース化が可能であると考えられ、例えば、



といった知識を知識ベース化する場合を意味する。

③統計データの特性 (種類) カテゴリー

統計データの特性カテゴリーとは、

上記DB は、□、□特性から成る ➡ 知識ベース

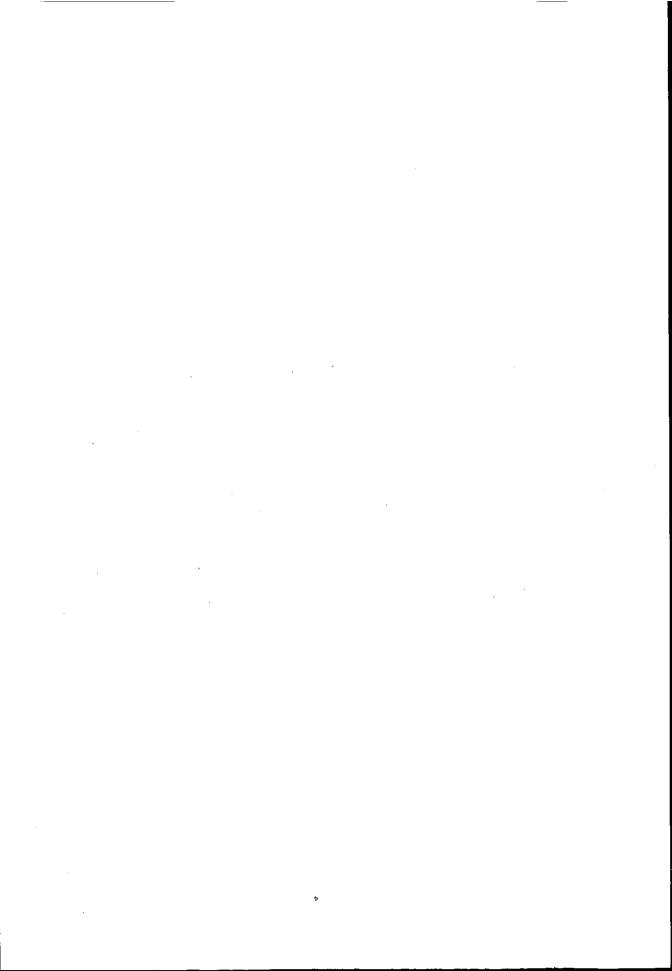
といったDBの持っている特性あるいは種類に関する知識をまとめて知識ベース化するものである。

## 3.3.2 知識ベース化の利点

以上のような方法論に基づいて、検索行動を支援するための知識ベースを構築することの利点は次のようになると考えられる。

- ① データベース利用の素人の検索行動をシステムがサポートすることから、誰でも標準的な検索が可能となる。
- ② 要求を満たすデータの特性、加工、表示に関する知識がサポートされることで 統計データの意味解釈の誤解が減少する。
- ③ ①の実現により、経済統計データベースおよびシステムの専門家が、初心者へのコンサルテーションに時間を費やされることがなくなる。

以上、本節では、「政策情報システム(PPIS)」の中のSDBの検索事例を基に、 その検索行動を分析し、知識ベース化する方法論の一つについて説明した。



第4章 海外研究動向 (統計情報とAI)

•		

## 4.1 海外動向調査の概要

本章では、海外における研究動向を紹介する。統計の分野にAI技術を持ち込もうと する努力は、国際的にも最先端の研究分野の一つとして関心を集めている。

例えば、EC統計局では従来からCRONOSの名称をもつ統計データベースを構築 しているが、最近になって「ECC DOSESプロジェクト」の名称をもとに人工知 能との結合を図っている。このプロジェクトはまだ研究の初期段階にあり、EC加盟各 国の研究者に研究プロポーザルの提出を呼び掛けている段階にある。

継続的な国際会議としては、「統計・科学データベース管理に関する国際作業会議」(International Working Conference on Statistical and Scientific Database Management,略称SSDBM)がある。この会議はカリフォルニア大学ローレンス・バークレー研究所のショシャニ教授が中心になり、1981年に第1回を開催(カリフォルニア)して以来、第2回を1983年(カリフォルニア)、第3回を1986年(ルクセンブルグ)と継続してきた。第4回は1988年6月21日から23日にわたり、イタリー・ローマで開催された。また、第5回が1990年4月にアメリカ・ノースカロラ(注1)イナで開催されることになっている。この会議はEUROSTATも後援者の中に名を連ねており、参加者にも各国統計作成部局のコンピュータ技術者が含まれるなどの特徴をもっている。

また、「AIと統計に関する国際ワークショップ」もその第2回の会議が1989年 1月4日から7日にわたりフロリダで開催されている。この会議はAT&Tベル研のゲール博士を中心とするグループが推進母体となっているが、参加者は国際的な広がりをもっている。以下では、この会議に提出された論文に基づき、海外における研究動向を探ることにする。

#### (注1)

第4回の会議については、大阪大学社会経済研究所の佐藤英人助教授による紹介がある(大阪大学学報No.416)。また第5回会議の開催要綱は「データベース・フォーラム」誌(Vol.2 No.3) に掲載されている。

## 4.1.1 会議の構成

「第2回 AIと統計に関する国際ワークショップ」に提出された論文は、テーマ別に見ると次のとおりであった。

### 1. Consultation Systems

- Anderson, K., Coleman, D., Hill, C., Jaworski, A., Love, P., Simaan, M., Special Cause Management: A Knowledge-Based Approach to Statistical Process Control
- Clayden, A., Croft, M., Statistical Consultation-Who is the Expert?
- Darius, P., A Toolbox for Adding Knowledge-Based Modules to Existing Statistical
  Software
- de Greef, P., Cooperative Statistical Problem Solving
- Dumer, J., Hanratty, T., Taylor, M., Nonpare, a Cosultation System for Analysis of Data
- Gale, W., Microstrategies: A Test Bed for Statistical Knowledge Acquisition and Representation
- Haaland, P., Liddle, R., Lusth, J., Wilson, D., Dexter: an Expert System for Evaluating Design Alternatives
- Hietala, P., ESTES: A Statistical Expert System for Time Series Analysis

  Irani, E., Slagle, J., Long, J., Matts, J., Formulating an Approach to Develop a
  - System for the Temporal Analysis of Clinical Data: the POSCH AI Project

## 2. Inference and Machine Learning

from Satellite Images

- Abramson, B., The Statistics of State-Spaces
- Brown, D., Pittard, C., A Parallel Tree Structured Classifier for Induction Systems
- Chan, P., Fisher, D., Statistical Guidance in Symbolic Concept Learning Collard, P., Gautier, C., Al and Statistics: A Case Study: Cloud Fields Analysis
- de Ville, B., Suen, B., Knowledge Seeker: an Interactive Tool for Data Analysis Esposito, F., Lanza, A., Malerba, D., Roselli, T., Integrating Statistical and Conceptual Learning Approaches for Knowledge Acquisiton

Hansson, O., Mayer, A., Probabilistic Heuristic Estimates

Kandt, K., Yuenger, P., Learning Seismic Event Concepts from Observation

McLeish, M., Cecile, M., Enhancing Medical Expert Systems with Knowledge Obtained from Statistical Data

Miller, N., Melton, R., A Tree-Based Approach to Intelligent Clustering for Waveform Data

Tan, M., Schlimmer, J., Extending a Bayes Classifier with IWN

Zhou, X., Dillon, T., Combining Artificial Intelligence with Statistical Methods for Machine Learning in the Real World

## 3. Knowledge Acquisition

Cholawsky, E., Psychometric Techniques as a Method for Acquiring Knowledge

Corter, J., Carroll, J., Potential Applications of Three-Way Multidimensional Scaling and Related Techniques to Integrate Knowledge from Multiple Experts

Cox, L., Incorporating Statistical Information into Expert Classification Schemes to Reduce Classification Costs

Ford, K., Petry, F., Adams-Webber, J., Chang, P., A Probabilistic Approach to the Automated Acquisition of Production Rules from Repertory Grid Data

Sauter, V., Madeo, L., Using Statistics to Make Expert Systems "User-Acquainted"

## 4. Knowledge Representation

Dambroise, E., Massotte, P., Statistical Methods Representation in the Expert System Muse:Choice of Statistical Methods

DuMouchel, W., The Structure, Design Principles, and Strategies of Mulreg

Hand, D., Statistical Knowledge Enhancement Systems

Lubinsky, D., Teaching Statistical Databases about Statistics

Oldford, W., Software Abstraction of Elements of Statistical Strategy

Shibata, R., Sibuya, M., Structured Statistical Data Description

Spirtes, P., Representation of Causal Structure in TETRAD II

## 5. Uncertainty Representation

Geiger, D., Pearl, J., Logical and Algorithmic Properties of Conditional Independence

Neufeld, E., Using a Theorem Prover to Compute Reference Classes and Build Provisional Models

Shafer, G., Shenoy, P., Probability Propagation

Smyth, P., Goodman, R., On Rule-Based Probabilistic Inference: Theoretical Principles and Practical Techniques

Spiegelhalter, D., Lauritzen, S., Techniques for Bayesian Analysis in Expert Systems

Wen, W., Information and Reasoning under Uncertainty

Yeh, A., Finding the Likely Behaviors of Static Continuous Nonlinear Systems

## 4.2 経済統計データベースの構築・利用の動向

研究のアプローチを2つに大別すると、統計の利用にエキスパート・システムを持ち込もうとするものと、エキスパート・システムの中に統計的知識を組み込もうとするものということが出来る。

このうちから経済統計データベースの構築と利用に関連するものについて概要を紹介 する。

第1のテーマであるコンサルテーション・システムは、ユーザ・インターフェースに 関連している。

## 4.2.1 「統計的コンサルテーションー誰がエキスパートか?」

A. Claydon and M. Croft (U.K.)

この論文は英国のメディカル・スクールにおける医療統計に関するコンサルテーション活動を背景として書かれたものである。統計の専門家ではない医療研究者に対する統計コンサルテーションを録音し、或いはビデオ録画して、そのコンサルテーションの過程、論理、内容を分析した。その結果明らかになったことは、統計に関する直接的な知識よりも、コンサルタント自身の貢献が大きいことが判明した。すなわち、コンサルタントは研究分野に関して広範な知識を有しており、その知識を利用して特定の研究を広範な研究計画の中に位置付けようとするのである。

こうした知見に基づいて、統計学に関するルールを組み込んだエキスパート・システムでは研究上のコンサルテーションには適さないとの判断が下された。

論文は、研究分野に関するエキスパートの知識をプロダクション・ルール化したシステムを紹介している。このシステムはハイパーテキストを用いてヘルプ機能及び照会機能を提供している。エキスパート・システムと統計的知識ならびにリサーチ・デザイン上の知識を組み合わせることによって、統計の専門家でない人々に対し、有効なコンサルテーションを行うことが出来る。

このシステムは "KnowledgePro"と呼ばれるエキスパート・システム・バッケージを 用いている。このパッケージはエキスパート・システムとハイパーテキストを組み合わ せている。

システムの目的は研究計画を完成させることである(図4-1参照)。システムが用いるルールの例を図4-2に示す。アンダーライン部分はハイパーテキストとのつながりを

示す。図4-3はハイバーテキストの例である。

このシステムでは、実際の統計データセットに直接アクセスして情報を取得するルールを幾つかもっており、ユーザが知らないような純粋に統計的な事実はシステム自身が確認する方法を取っている。

```
RESEARCH PROTOCOL
Name
                  : Dr Smith
                                      Position: Research Assistant
Place of work
                  : Medical Department
Title of study
                  : Comparison of two treatments for malnutrition
Purpose
 Type of study
                  : Comparative experiment
 Study Population:
 Major Hypothesis: Malnutrition less severe with Treat. A cf. Treat B
 Ethics
Methods
 Design of Study:
 Sample selection:
 Sampling frame
 Sample size : 55 (Large)
 Variables
Results
Analysis of Study:
Data summary
Computation
 Interpretation
```

#### 図4-1

```
topic 'TYPE OF STUDY'.
  If?STUDY_MODE is Experimental
    write(is this experiment a Comparative experiment,
          or is it a Complex research study?
    and menu ('TYPE OF STUDY',
             L'Comparativeexperiment',
              'Complex research' ]).
  If?STUDY_MODE is Observational
  then
    write(is your observational study what we can call a
          Simple descriptive research study, or is
          if a comparative Observational study?
          Alternatively, it may be a study which is more
          correctly termed Complex research.
    and menu ('TYPE OF STUDY',
             ['Simple descriptive',
               Observational study',
              'Complex research' ]).
  IF?STUDY_MODE is Help then do(DESIGN_HELP).
end.
```

..Observational study

An observational study is a <u>research design</u> in which observations are made without the allocation of subjects to any specific conditions under test. As a result, though <u>descriptive</u> and <u>comparative</u> statistics are still calculated, any <u>inferences</u> drawn are more tentative, because of the effect of the intervening <u>uncontrolled</u> variables.

⊠4-3

4.2.2 「既存の統計ソフトウェアに知識ベース・モジュールを付加する手段」

P. L. Darius (Belgium)

この論文は、SASに組み込まれたコンサルテーション・システムを紹介している。 このシステムはTAXSYという名称で呼ばれる。

既存のコンサルテーション・システムではユーザ自身が知識ベースに変更を加える可能性には関心が払われていないが、知識ベースを容易に変更し、或いは知識を追加出来る機能は、極めて必要性が高い。

また、多くの場合、知識工学的ツールと統計パッケージは分離しているか、或いはOSを介してコミュニケートしている。こうした分離構成は、統計に馴染みの深いユーザ (言い換えれば統計に関する部分を知っているユーザ) がシンボリック・プロセッシングを行う知識工学部分にアクセスすることを難しくしている。後者の環境は前者と全く異なるからである。

TAXSYは、図4-4に示すように、SAS環境の中で、ユーザが入力するルール (同じ) くSASプログラム) を扱うシステムであり、統計パッケージの中に知識工学 的ツールを組み込む方法を取っている。

TAXSYの中心部分は推論エンジンであり、

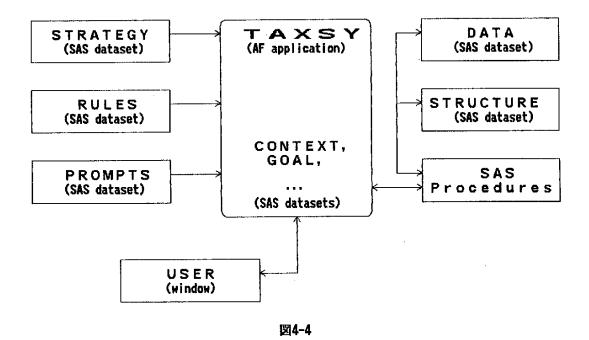
I F · · ·

 $AND \cdot \cdot \cdot$ 

THEN · · ·

形式のフォーマットを有する。STRATEGYデータセットは全体プロセスを繰り返し行うガイドとなるスーパーストラクチャーである。GOALを指定し、用いるRUーLEとPROMOPTSデータセットを指定する。PROMPTSはルールにない値をユーザに問い合わせるか、統計計算の結果として、入手する。STRUCTUREデータセットはデータセットに関する種々の情報のうち、SASフォーマット化出来ないもの(例えば実験デザインなど)を扱う。

このシステムは、論文の著者によれば、ユーザのイニシアティブを大幅に認め、また 知識ベースが可能な限り簡易化されている。しかし一方、システム全体が特定の統計パッケージに依存しているという限界を有する。



### 4.2.3 「統計的問題に対する協力的解法」

P. de Greef (The Netherlands)

ユーザとシステムの協力関係を分析した論文である。過去のエキスパート・システム において、ユーザがシステムを使いこなせない例があった。そこでタスクを分担することについて十分な検討を加える必要性が感じられたのである。この論文はそうした分析 例であって、具体的なシステムとして完成したものの紹介ではない。

タスクを分解するという概念は図4-5に示される。例えば、「データ収集の計画」といっても、変数を具体化する上で確信が持てないこともあろうし、或いはサンブル・ササズに関して助言が求められる場合もあろう。こうしたサブ・タスクを識別し、それをシステムの機能とするか(アドバイス機能)、或いはユーザとシステムの認識を照合するか(チェック機能)、といった検討を行う。

統計コンサルテーション・システムは問題解決力を持つだけでは不十分であり、ユーザとの協力関係を保てることが重要である。著者はコンサルテーション・システムの開発にあたってユーザを分析することを奨励するとしている。

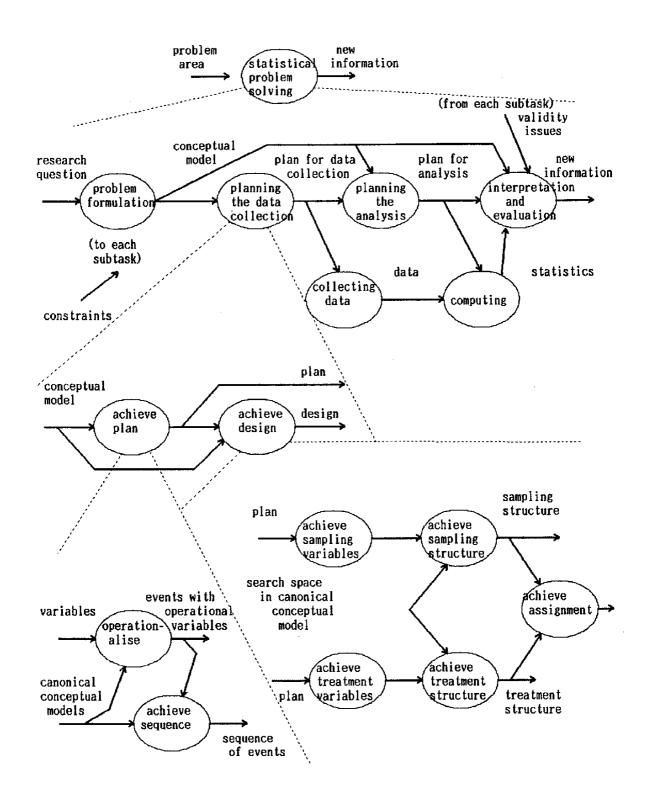


図4-5: The task-decomposition of statistical problem solving(basis task model)

4.2.4 「データ分析のためのコンサルテーション・システム NONPARE」

J. C. Dumer Ⅲ, et AL (USA)

統計データが適切であるか、分析方法が目的やデータに適合しているか、などをチェックし、分析結果の解釈を呈示する機能を提供することを目的とする。システムはNON-PAREという名称を持ち、ノンパラメトリック統計分析を行う上でのインテリジェント・インターフェイスを目指している。プロトタイプが完成しており、1989年には実地のテストが行われる段階である。

NONPAREはGenieと称する米国陸軍開発のエキスパート・システムを用いている。

論文の大部分はNONPAREの実際のセッションを追う形で書かれている。これについては図4-6、図4-7、図4-8、図4-9を参照するにとどめる。

Do you have a sample X1, ..., Xn? y
Are you interested in whether the data conform to a specified distribution? n
Are you interested in the probability of occurrence of a particular category or event? y
Enter the name of the category of interest -> perforation
Are the n trials producing the values X1, ..., Xn independent?

### 図4-6 セッションの開始

Are the n trials producing the values X1,  $\dots$ , Xn independent? what is in dependent

Independence relates to freedom from external influence or control-here, the reference is to measurements (data) being free to assume values without regard to other measurements that may be made.

Are the n trials producing the values X1, ..., Xn independent? y Does each trial have the same probability p of producing the perforation ? y

Are you interested in considering whether the probability of occurrence of the perforation equals or is bounded by some specified value p\*?

#### 図4-7 ユーザによる問合せ例

The binomial test is an appropriate procedure, To execute the binomial test, use the menu to complete this statement:

I am interested in testing the null hypothesis that: The probability of occurrence of the perforation

#### Pick One

equals some value p\*
does not exceed p\*
is at least p\*

図4-8 ユーザによる検定の選択

I am interested in testing the null hypothesis that:The probability of occurrence of the perforation does not exceed  $\mathbf{p}^{*}$ 

Specify the sample size  $n \rightarrow 14$ 

Specify a value for  $p^* \rightarrow .80$ 

Specify the number of datum values assigned to the perforation  $\rightarrow$  11

# 図4-9 仮説の確認とパラメータの指定

### 4.2.5 「時系列分析のための統計エキスパート・システム ESTES」

P. Hietala (Finland)

ESTESは熱練していない分析者に対して時系列データの予備分析を行う上でのガイダンスを行うことを目的としている。ESTESは、an Expert System for Time Series analysisを略した名称である。ここでいう予備的分析とは、時系列データを扱う上で、季節性、トレンド、異常値、シフト、その他を検出し処理するための操作をいい、モデル構築の前段階、或いはその過程で行われるものである。

論文は、まずシステム構築にあたっての原則、例えばシステムが判定を下すのではなく 助言を与えるにとどめるといったことを簡単に述べる。

次いで、システムの構成について説明する(図4-10参照)。システムはPC(アップル・マッキントシュ)の上でPrologとPascalによって書かれている。

- メイン・モジュールはモジュール間のコミュニケーションを行う。
- ・ 統計知識ベースは時系列分析に関する知識から構成されている。
- ・ 推論エンジンは知識ベースを利用して駆動される。
- ユーザ・インターフェイスはユーザとのインターアクションを担当する。
- ・ グラフィック・モジュールは結果の画像表示を行う。
- ・ 時系列発生モジュールはサンプル系列をジェネレートする。
- ・ 数値計算モジュールは各モジュールに対して計算を行う。Prologはシンボリック・ プロセシング用であって数値計算用ではないので、数値計算は全てPascalを用いて 書かれている。

論文はその大部分をESTESのセッションの提示にあてている。その一例を図4-11に示す。画面には時系列データが提示されている。ボックスの中はユーザとシステムの間の対話であり、ユーザが4期ごとの季節性であるとのインプットを行い、OKにより対話を終了したところである。後にシステムは季節性が実は12期ごとであることを感知し、ユーザに確認を求める。また、トレンド除去のために対数変換を行うことをアドバイスし、ユーザの要求に答えてその理由を説明する。(図4-12参照)

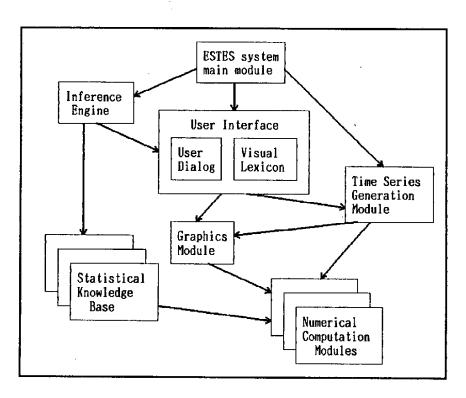


図4-11 ESTESのシステム構成

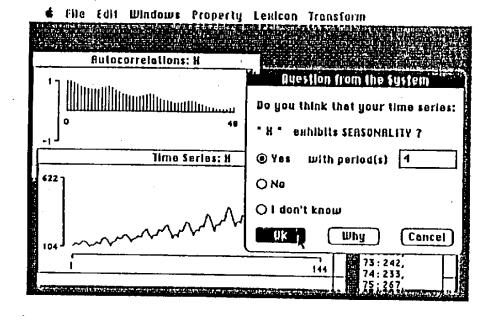


図4-11 セッション例:ユーザとシステムの応答

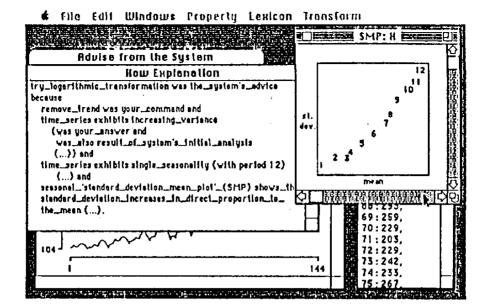


図4-12 セッション例:システムからのアドバイス

第2のテーマであるマシン・ラーニングに関する論文は我々の関心に直接関連するもの は少ないが、そのうちで次の論文に触れておこう。

4.2.6 「データ分析のためのインタアクティブ・ツール KnowledgeSeeker 」

B. de Ville and Ed Suen (Canada)

この論文はKnowledgeSeeker (KS) の名称の基に市販されているソフトウェアの紹介である。このソフトは帰納推論に関する成果をマシン・ラーニング及び統計の双方から取り入れたとしている。KSの目的は大量の多変量データセットを分析する上でのツールである。いくつかの特徴を挙げれば次の通りである。

- ・ 類似したルールを統合するプロセスを持ち、知識の冗長性を減じる様にしている。 これにより分析結果の理解可能性が向上している。
- ・ ツリー型のブラウザーにより分析経路をモニターすることが出来る。
- · 欠損値の扱いに関しては各種の方法が組み込まれている。
- ・ あるデータセットについて、Y=f (X1, X2, ・・・・, Xn) の様な関数形が当てはまる時、Yに対する影響の上で類似した独立変数を同じカテゴリーにクラス分けすることにより、相互に独立なサブセットを形成するアルゴリズムを有する。 図4-13参照)

KSはIBM PC上にインプリメントされている。

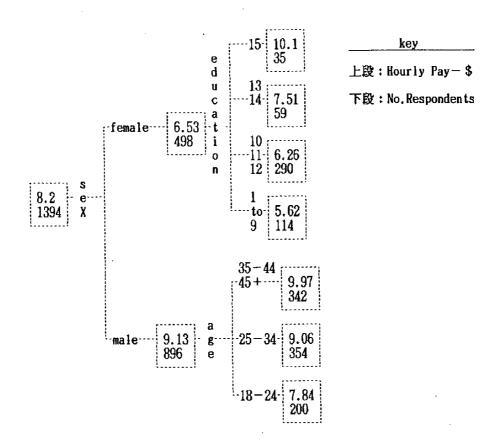


図4-13 時間当たり給与に関する主要説明変数

第3のテーマは、知識獲得である。エキスパートから知識を獲得するプロセスはエキスパート・システム構築の上で最も困難な問題である。

## 4.2.7 「知識獲得の方法としての計量心理学的手法」

E. Cholawsky (USA)

ナレッジ・エンジニアがエキスパートから知識を獲得する方法として、計量心理学の経験が有効であるとを論じた論文である。論文はこの分野に於けるこれまでの経過をサーベイし、さらに計量心理学の手法自体を簡単にサーベイしている。その上で知識獲得への計量心理学的手法の応用可能性として次の点について論じている。

第1に、最も直接的な応用としては、エキスパートの知識をスケール化(何段階かへの位置付けを行うこと)ないし順序付けすることである。これはマッピングと呼ばれるが、 1対の比較を行うことや、質問票による回答を得るが含まれる。エキスパートはそうした 過程に直接関わると同時に、原データを提供する。

第2に、数値や図により問題点をエキスパートに示し、討論の出発点とすることが出来 る。そうした整理はエキスパートの協力による場合もあるし、そうでない場合もあり得る。

第3は、多次元のスケールやクラスター分析の結果として得られた数値をもって、エキスパートの代理とすることが出来る。

論文は次いで、1次元のスケールによるエキスパートの評価例を示している。

# 4.2.8 「エキスパート・システムを利用者指向にするための統計利用」

V. L. Sauter and L. A. Madeo (USA)

エキスパート・システムに於ける統計利用は通常の場合、問題の分野に関する知識獲得の手段としてのそれである。この論文は異なる視点に立ち、システムのユーザに関する知識及びそのニーズに関する知識の獲得に統計を利用することを主張している。ユーザに関する知識を持たないエキスパート・システムは常に「平均的」な回答を行う。論文は商品の欠陥発見を例に挙げているが、より一般的にいっても個別ユーザに関する統計をエキスパート・システムが持つことにより、問題の識別や対処法のアドバイスにあたり一層的確なレスポンスをすることが出来るとされる。

(ファイル:MITI-CSK2)

第4のテーマは知識表現である。

4.2.9 「エキスパート・システムMUSEに於ける統計手法表示・・・統計的手法の選択」
E. Dambroise and P. Massotte (France)

産業分野に於ける実用エキスパート・システムの場合、ますます複雑性が増しており、 意思決定サポートが重要である。本論文は産業分野におけるデシジョン・サポート・シス テムの中にデータに即した統計手法の選択機能を持たせることを論じている。

MUSEのアーキテクチャは次の5つからなる。

- ・ ドメイン・モジュール:エキスパート・システムの立ちあげ
- ・ データ・モジュール:データ・タイプの識別、欠損値や異常値の検出など
- ・ 目的モジュール:ユーザの要求
- ・ マッチング・モジュール:統計手法、データ・タイプ、目的のマッチング。統計手 法はフレームの形で表現されている。
- ・ 加工モジュール:統計分析の実施。

#### 4.2.10 「MULREGの構造、設計方針、及び戦略」

W. DuMouchel (USA)

MULREGはデータ分析、回帰モデル構築、及び結果の表示・利用のためのインターアクティブな環境である。既に400箇所で利用されている。対象は統計を専門としない工業技術者及び技術者である。

その特徴は次の通りである。

- モデル構築のあらゆる過程にグラフィックを組み込んでいる。
- ・ 特に新規に開発したものとして、多重回帰の独立変数間に見られる相互依存関係の 視覚化がある。
- データ変換及び残差の処理などにグラフを組み込んだ。
- ・ モデルが確定すると予測値の信頼区間及びフィットされた面のプロットを作成出来 る。

MULREGはルールによるプログラムではなく、またオブジェクト推向の言語による

ものでもないが、多重回帰分析を行う場合にこれまでよりも「馬鹿げていない」環境を提 供している。

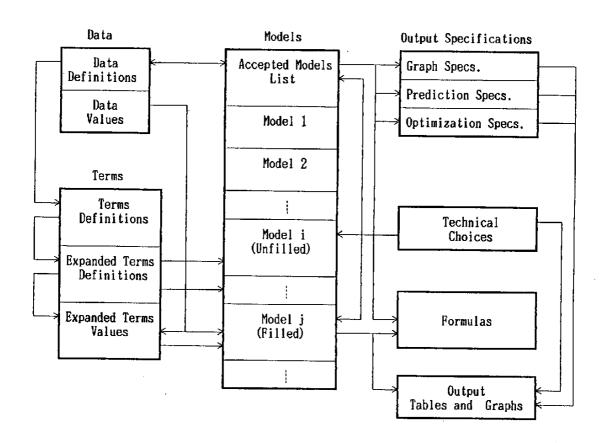


図4-14 MULREGオブジェクトのデータ構造

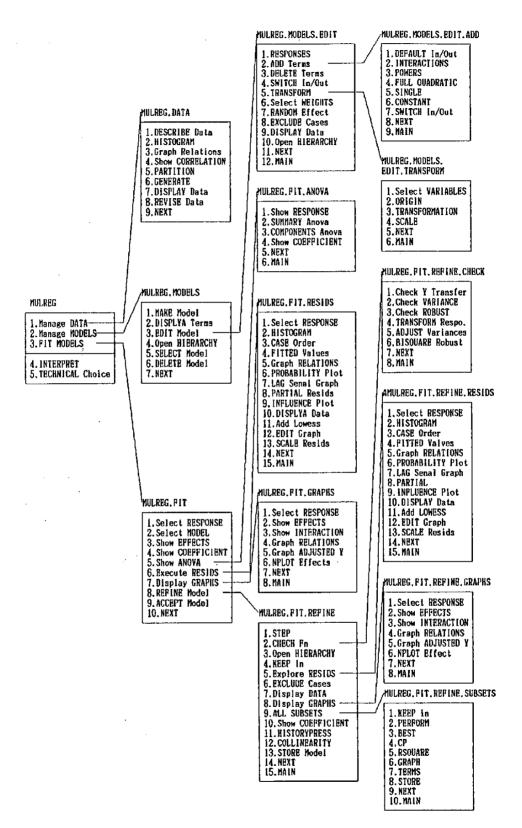


図4-15 MULREGのデータ分析及びモデル構築のメニューツリー

現在の統計エキスパート・システムは計算部分を中心に置き、ユーザ・インターアクションはほぼ分離したインターフェイスが分担している。またその多くは統計分析手法の中でユーザを導くという限定された目的に向けられている。これはエキスパートの知識を云々する以前に統計パッケージがすでに存在していたということからくる不可避の状況である。

数年来、著者は人間の側に一層注意を向けるべきであると感じるようになった。そのためにKENSシステム(Knowledge Enhancement System for Nonparametric Statistics)が開発された。知識のエンハンス、あるいは増強という用語を用い、エキスパート・システムと言わない理由は、KENSの目的がブラック・ボックスが答えを用意するのではなく、ユーザと対話することにあるからである。開発の主要動機は下記の通りである。

- ・ルールに基づくシステムよりも弾力的なシステムの方が有効である。
- ・ 人間のアドバイザーは問題に対して厳密な解決策を提示するだけでなく、関連する 問題に対する研究者の理解を深める役割を果たしている。
- 人間は手法の詳細やパッケージへのアクセス法などを容易に忘れる性癖を持つ。
- ・ これまでの文章による知識表現では不十分であること。

KENSはユーザの記述とシステムが持つテキストとの類似性の尺度を計算する。ユーザは、提示された結果を見てテキスト内容を検討するか、或いはプロセスを繰り返す。これは、ハイパーテキストの機能に相当する。

記述子としては、"is a generalization of," "is a special case of," "is the opposite of," "is a related concept of," "is an equivalent concept of," などがある。

論文はさらに、LMG (Linear Models Guide) という名称のハイパーテキスト機能を紹介している。そのセッション例を第4-16図に示す。このうち〔1〕及び〔2〕がハイパーテキスト・ボタンであり、当該用語の説明を呼び出すことが出来る。これらの用語を知らないユーザは1ないし2をタイプ・インするか、或いはメニューから"?"を選択すればよい。

Is this variable continuous (1) or categorical (2)?

- A. Continuous
- B. Categorical
- ?. Don't know

Does this variable take only a few values (such as the number of children in a family) or can it take any of a large set of values (such as heights of four year olds)?

- A. Only a few values
- B. Many values possible
- ?. I'm uncertain.

LMG has determined that in your sample this variable takes fewer than eight different values. As a consequence, it will be treated as categorical (1) rather than continuous (2).

## 図4-16 LMGの応用例

統計データベースはフラット・ファイルに収められるのがこれまでのやり方であった。 著者は統計理論の知識を組み込んだオブジェクト指向のデータベースを提案する。それにより一層強力且つ自然な統計検定を行うことが出来ると主張される。著者の提案するシステムの概要を図4-17に示す。

論文で著者が実際に行っていることは、当初に述べられる意図よりもかなり限定的である。著者は2つのシンタックスを提案する。第1のものは統計知識ベースを定義するために使用するものである。第2のシンタックスはデータベース管理者がアプリケーション・データベースをセットアップするために用いるものである。想定されているケースは、変数の分布型に関するものであり、2つのランダム変数を組み合わせたとき結果する分布に関する知識である。この知識自体は予め知られており、システムの外から専門家がシステムに教えることになる。

以上のような知識ベースの判断により、ランダム変数の比較(平均値、中位値、最頻値、 その他)を正しく行うことが出来る。

第5のテーマは不確実性の処理であるが、これについては我々の関心と直接関連する報告はなかったので省略する。

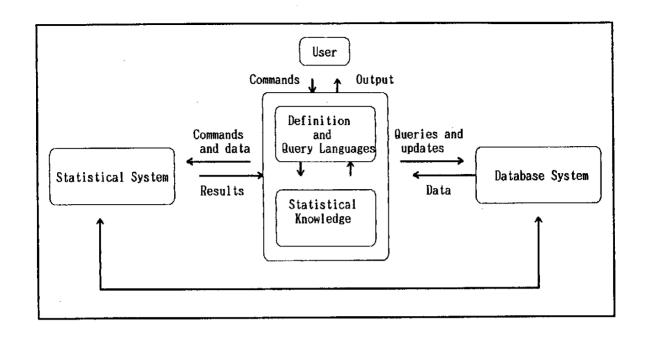


図4-17 提案システムの概要

4.3 海外に於ける動向とわが国の研究へのインプリケーション

海外に於ける最近の研究発表は以上のように展望することが出来よう。個々の論文の要約をこえて、全体の動向を見るとき幾つかの特徴を指摘することが出来る。

- ① 佐藤助教授らによる「統計データモデル」のコンセプトは、我が国で今や標準的な考え方として定着しつつあるが、欧米ではこれを論ずる者は少ない。これは、欧米では「統計」という時、まず念頭におかれるのがサーベイ・データ (プロセス・コントロールなどは勿論であるが)であることに原因を求めることが出来よう。統計利用に関して欧米ではミクロ指向が強いと言えよう。日本では欧米に比較して産業データ、地域データ、年令階層別データなどといったレベルで集計したデータが多様される傾向がある。個票を基礎にしたデータベースを構築するのであれば、関係データベースに基礎をおいて、通常のデータベース構築に進むことが出来る。これに対して、集計表を出発点として選ばざるを得ない場合は事情が異なる。個別・具体的な集計表に対して、一般的なコンセプトとしての統計のメタデータを構想する必然性が出てくるのである。
- ② 知識ペースに対する極めて実用的なアプローチが目立つ。
- ③ 技術的には知識ベースのブラック・ボックス化を避けて、ユーザに対するコンサ (注2) ルティング機能を拡充する傾向が見られる。その一環としてハイパーテキストの利用が目立った。
- ④ 今回展望した論文においては、統計分野における知識ベースの利用は日本に比べて特に先行しているとは思えない。但し、官庁の統計部局等においては統計に関する情報がマシンリーダブルな形で保存されており、一般ユーザに対するオンラインでの提供サービスが既に実用化されている。この面では、我が国は欧米の水準と比べ遅れていると言える。
- ⑤ 展望した論文には大きな研究チームによる成果は含まれていなかった。我が国における状況も社会経済データに関しては依然として研究室内部にとどまっているので、両者にそれほどの差は存在しないと言えよう。これは冒頭に述べたECの状況からみても言えることである。米国においては、統計データベースに関する限り、

さらに状況は楽観出来ない。

統計データベースに対する知識ベースの組み込みには、多様なユーザの有する知識を体系的に吸い上げるプロセスが重要である。実用を目指すほどその重要性が増すのは当然である。従って、この分野の研究の成否は、まずそうした協力関係を樹立出来るか否かにかかっていると言えよう。

## (注2)

ハイパーテキストの概念自体は、ノード(Node)と呼ばれる情報のひとかたまり(この場合、文字、ストリング列からグラフィック、イメージ、そして音声、ミュージック、動画といったものまで含めて)をネットワーク型に自由にリンクさせるメカニズムを持たせるシステムである。インタラクティブにリンクを作成していくオーサモード(author mode)、作成されたハイパードキュメントをリンクに応じて読み進んでいくブラウズ(browsing mode)の二つのモードに分かれる場合が多い。

第5章 人工知能を応用した経済統計 データベース利用支援システム の構築

## 第5章 人工知能技術を応用した経済統計データベース利用支援システムの構築

本章では、前章までの結果をもとに、構築するシステムの開発目標の設定を行う。具体 的には、人工知能利用技術を応用した経済統計データベース利用支援システムのイメージ、 システムの稼働に必要な知識、開発の際の課題等に関する検討結果を報告する。

特に本調査研究のなかでは、実際に知識を洗い出すために、経済統計の分野を限定し、 貿易に関する分野を対象として検討することとした。

### 5. 1 経済統計データベースの利用支援システムに関する提案

本調査研究を通じて得られた検討事項及び資料を基に、人工知能(AI)技術を応用した経済統計データベースの利用支援システムに関する提案を以下に示す。

#### 5.1.1 経済統計データベース利用の現状

ここでは、経済統計データベースの利用者の傾向について触れ、更にAI技術の適用範囲について説明を行う。

### ①経済統計データベース利用の現状

1章では、データベース利用の現状について、一般の調査結果を基に概観し、また 2章で特に、経済関連データベースの利用についての実例を考察し、実際の経済統計データベース検索における問題点と現行の対策についての検討を行った。

以上から経済統計データベース利用の傾向を整理すると下記のようになる。

## 

- ・データベース及び、経済統計に関して、素人が、利用するケースが増えている。
  - (例、国際化の進展により従来関心を持たなかった層の利用が増えた)
- ・従来の利用者(経済統計の専門家)のデータ利用ニーズが多様化している。 (例、データのグラフ化、外部情報と自社情報との統合加工等の要求が増えた)

このように、データベースの利用の様相が、普及と高度化の2つの方向に向かってい

#### ることがわかる。

次ページに挙げたデータベース利用の問題点にも、これらの傾向が反映されている。

#### 

- ・データベース及び経済統計に関する専門知識のガイダンスがないと、検索と内容 理解がしにくい。
- ・既存のガイダンス機能が使いにくい (メニュー構成上の問題等による)
- ・品目分類について、データベースの分類と利用者の必要とする分類にギャップが あり、検索がしにくい。
- ・データの性質(例えば、調査対象範囲の定義により同じ名前のデータでも性質が 異なる)に関する情報が、利用者に判りにくい。
- ・欠測データに関する情報が判りにくい。
- ・異なるデータベースからの情報の合成等高度な加工がしにくい。

#### ②A I 技術の適用性について

上記のようなデータベース利用上の問題点を解決する方法の一つとして、AI技術の適用可能性についての検討を行った。データベース検索へのAI利用の例が3章に報告されている。

これらをもとにAIの適用可能分野をまとめると下記のようになる。

#### · :---- A I の適用可能分野-------

- ・ユーザの要求に対する適切なアドバイス機能について
- ・経済統計の専門的知識の知識ベース化による検索システムの髙度化について
- ・統計スキーマの知識ベース化による柔軟なインターフェイスの構築について

さらに、A I 技術の適用範囲を明確化するために、利用者が必要とする知識を分離し、 図式化すると次頁のような3つの世界が想定される。利用者は、「ユーザーの世界」から 「データベースの世界」に直接アクセスして、データを取り出しているように見えるが、 「ユーザの世界」での情報表現を「統計の世界」での情報表現に替えるための知識、「統計の世界」の情報表現を「データベースの世界」での情報表現に替えるための知識がなければ正確な情報検索はできない。

本調査研究では、特に「ユーザの世界」から「統計の世界」を通って「データベースの世界」への橋渡しをスムーズに行う為に、AI技術の利用を試みることが中心的検討課題となった。

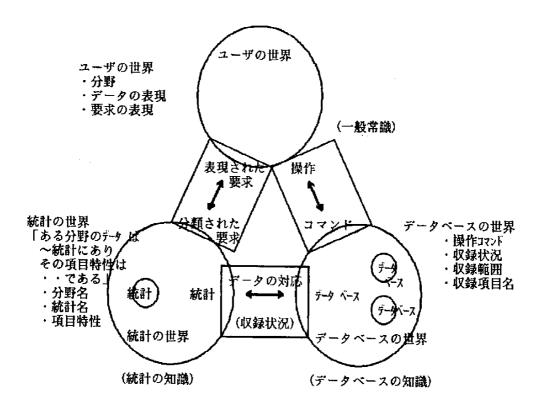


図5-1 経済統計DBとAiの関係図①

上図の「ユーザの世界」は、ある情報を必要としているユーザが (データベースや 統計に関する専門知識なしに) 表現した、欲しい情報の姿である。日常的な言葉で表現されており、データベースや経済統計の専門家から見ると曖昧な表現であったり、 複数の解釈が可能であったりするものも含まれる。

「統計の世界」は、統計の観点から情報を分類した世界で、統計の専門家が表現した情報の姿である。<u>~という統計の~という分類に計上されたデータを~</u>の期間に関

<u>して合計したもの・・・というように、どの統計にどういう形で収録されている情報であるかによって情報を表現する。</u>

「データベースの世界」は、データベースの観点から情報を分類した世界で、データベースの専門家が表現した情報の姿である。どのデータベースに収録されており、 どのようなコマンドによって検索ができるという形で表現される。

「ユーザの世界」から「統計の世界」へ情報を変換するためには、世の中に存在する統計の性質や特徴に関する知識を持ち、一般常識の範囲で表現された要求を解釈して統計に当てはめることが必要である。

「統計の世界」から「データベースの世界」へ情報を変換するためには、データベースのコマンド体系に関する知識を持ち、統計がデータベースにどのような条件で、 収録されているかを知る必要がある。

ユーザが、独力で(専門家の手助けなしに)的確なデータベース検索を行うためには、上述の知識をAI技術を使って知識ベース化し、ユーザの世界から直接データベースをアクセスするシステムが考えられる。(下図参照)

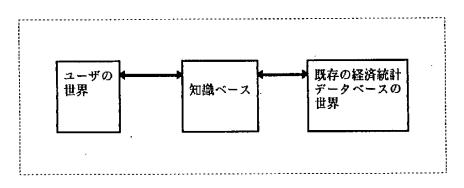


図5-2 経済統計DBとAIの関係図②

(この知識ベースの機能は、ユーザの要求(入力情報)に合う適切な 経済統計データの検索を支援し、ユーザに提供するものである)

#### 5. 1. 2 知識ベースの構築におけるケーススタディ

前節で考察された知識ベースを実際に構築するためには、どのようなアプローチが可能 で、実際の問題解決に役立つ知識が、どのような形式で表現されるものかを知るために、 本調査研究の一環として、ケーススタディを行った。本節ではその結果を報告する。

### ①知識の収集方法

経済統計データベースを利用する人すべての要求に対して、システムが対応できるような知識ベースを持つことは、膨大な作業とコンピュータ資源を要する。そこで、今回の調査では、経済統計の世界を「貿易」の分野に限定し、また、対象ユーザを通産省内の利用者に限定した形式でケーススタディを行い、実験的に知識の収集を実施した。

更に、この「ユーザの世界」と「統計の世界」に介在する専門的知識を明確にするために、ユーザの要求文を通産省の月刊誌「通産ジャーナル」中の文により作成し、それらの要求文を満たすためにはどのような「知識」が必要であり、どのような検索が行われるかについての調査を行った。

…「ユーザの世界」の把握と知識収集(獲得)の過程 …………

経済活動表現事例の 抽出(約100 事例) 「通産ジャーナル」(1年分)から「貿易情報」に関る記事(文)の抽出を行った。

Î

貿易統計に関する要 求文の作成

貿易統計に関する要求文の作成 (約100 事例)

Î

各要求文に関する知 織(コメント)の収 集

- ・経済統計の専門家による要求文の意味の解釈
- ・要求文の「不足情報」の指摘
- ・参照する経済統計についての説明
- ・使用するD/Bについての説明

#### ②知識の整理

①の方法によって収集された専門的知識に関して、整理用のシートを利用して経済統計 に関する専門的知識、データベースに関する知識等の整理を行った。

これらの知識のうち、「ユーザ世界」と「統計の世界」の間に必要とされる知識としては以下のようなものが存在することが判った。

## -検索支援に必要となる知識の種類(ユーザ〜統計)-----

- ○ユーザの意図と統計概念を対応させる知識
- ○品目分類に関する知識
- ○数量・金額の単位に関する知識
- ○国・地域分類に関する知識
- ○時点や期種に関する知識
- ○データの加工に関する知識
- ○経済行為 (活動) を理解する知識

さらに「統計の世界」と「データベースの世界」の間に必要とされる知識としては、以下 のものが挙げられる。

### ・検索支援に必要となる知識の種類(統計~データベース) …………

- ○統計が収録されているデータベースの所在に関する知識
- ○データベースに収録されているデータの範囲(系列分類・期間等)の 知識
- ○データベースのデータを加工して得られる内容に関する知識
- ○複数のデータベースを組み合わせて利用する知識

ユーザの意図と統計概念を対応させる知識を表現するために、今回の知識整理の結果から、特に検索したい情報の種類を表す概念「要求種別」を定義した。

「要求種別」とは、要求の本質を示す言葉で、参照すべき統計を決定し、検索に必要な補足情報を決定するための要因となる。ユーザの表現方法は多様であるが、統計の用語で表現すれば、限定されたカテゴリーに分類できると思われる。

要求種別に相当する言葉を整理用シートから抽出したものが付録3である。(付録参照)要求種別の表現を「輸入額」等のように一つの名詞句で表現できるものと、「輸入の推移」のように状態や、変化を示す言葉との組み合わせで表現されるものの2種類に分けて、整理した。最終的に要求種別を確定するためには、さらに詳細な分析が必要であるが、こ

れらを基に、要求種別ごとの検索行動や、統計の選ばれ方に関する検討を行うことができる。

ケーススタディの結果から、要求種別が、日常語で表現されたユーザの要求を、統計の 世界へ変換する際の「取り掛かり」となる言葉であるという仮説をたて、開発システムの 基本機能を設定する際に、この要求種別を最初のキーワードとして、検索支援を行ってゆ く仕組みを検討した。

## 5.1.3 システム開発へのアプローチ

上述のように、本調査研究の中心的課題は、<u>「専門的知識のない人にも、簡単に経済統</u> <u>計データベースが検索できることを可能にする</u>」ことである。

このような機能を、システムとして実現するためには、どのような開発の方向が考えられるのかについて、特に人工知能技術の効果的な利用という見地から、以下のような2つのアプローチが考えられる。

### i) 知識ベース付加型

このシステムは、検索機能の高度化 (知的検索システム) を中心に人工知能技術の利用を行うものであり、現在まで実験的に実施している知識収集、整理作業は、この型のシステム構築の為の知識源として利用されるものである。

検索時のユーザへのコンサルテーション機能の向上を中心的機能とすることから、このシステムは、既存のデータベースの内部構造への直接的な知識ベースの利用ではなく、サブシステム的に独立したものとなる。

このシステムの構築により、ユーザは、経済統計データベースの検索を専門家と会話しているような状況で簡単に実行することが可能となる。

このシステムの構築には、既存データベースの利用の分析、経済統計に関する知識の体 系化と知識ベース化、さらにユーザの要求パターンの分析、柔軟なユーザインタフェース の開発等が必要である。

知識ベース化のメリットは、検索時に必要となる知識の追加・変更・修正等が可能であり、システム機能を高められる点にある。但し、経済統計の利用に必要とされる知識は膨大であることから、知識の構造化において拡張性のあるシステム構成が必要となる。

最終的には、経済統計の各分野ごとに特化された知識ベースが複数構築されることから

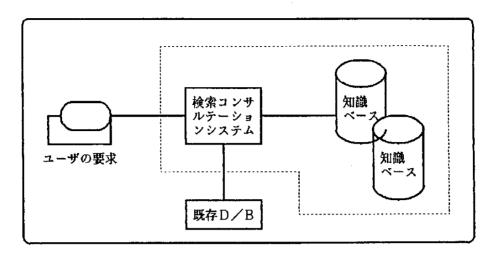


図5-3 検索コンサルテーション型の概念図

### ii) 知識データベース内蔵型

このシステムは、AI技術、特に知識ベースの特長を活かし、DD (Bata Dictionary)の知識ベース化を実現するシステムである。このシステムの開発には、既存データベースの概念スキーマの整理、データスキーマの構築、外部スキーマの構築等が必要となり、大規模な開発工数を要する。

知識ベースをデータ・ディクショナリー・ツールとして活用するシステムであり、このシステムの構築には、既存D/B において集合演算できるものを分析する必要があり、また、フィールド間の関係についても分析が必要となる。

知識ベースの利用の効果としては、知識ベースの階層表現、属性及び属性値の継承等が考えられる。

このようなシステムが実現されれば、個々の経済統計データベースの異種性が改善されメタデータベース的に統合化が図れることから、実世界 (Real World) をより反映したデータベースが実現される。

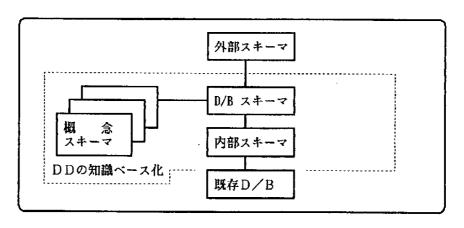


図5-4 知識データベース型の概念図

#### 5. 2 システムの概要

本節では、前節までに解説した検討結果を基にまとめられた、開発システムの概要を説明する。検討中のシステムを以降の文中でIRS (Intelligent Retrieval System) と呼ぶことにする。

システム開発へのアプローチとして、知識ベース付加型と知識ベース内蔵型の2つ が検討されたが、主に下記の理由により、知識ベース付加型のアプローチが、採用された。

- ① 知識ベース内容としては、ユーザごとに内容の異なるものも考えられるので データベースと独立した形で知識ベースを持つことが望ましい。
- ② 既存のデータベースの利用を支援することを目的としているので、データベースそのものに手を加えることは避けたい。
- ③ 頻繁に使う知識とそうでないものは、比較的明確に区別できると思われるので、すべてのデータに対して、関係定義付けを行う知識ベース内蔵型のアプローチよりは、知識ベース付加型のほうが、効率的な開発ができると思われる。

#### 5. 2. 1 システムの特徴

統計情報の為の知的検索システム IRS は、データベース利用技術調査委員会の検討結果をもとに開発される貿易統計情報利用システムである。以下に本システムの特徴を説明する。

本システム開発の基本課題は、「専門的知識のない人にも、簡単に経済統計データ ベースが検索できることを可能にする」ことである。

上記課題及び、 "経済統計 "データベースを対象としている点から、本システムは、一般のデータベース利用システムと比べて、下記の特徴を持っている。

(1) ユーザは求める情報をデータベースのレベルで知っている必要はない。 また統計用語のレベルで知っている必要もない。

以下の図は、この考え方を図に示したものである。

ユーザが求める情報(ユーザ要求)が、始めは、日常語で表現されており ( 左端の楕円)、統計用語に置き換えられ、さらにデータベース用語に変換され て右端の検索コマンドの形になって検索がおこなわれる、というのが一般的な データベース検索のプロセスである。

一般的には、日常用語から統計用語への変換は、統計の専門家によるコンサルティングという形で実現され、統計用語からデータベースへの変換は、データベースの専門家または、データベースに付属するガイダンス機能によっておこなわれているのが現状である。

IRSでは、日常語の世界から検索コマンドに至る変換(図中矢印f3で 示した変換)を、システムの機能として実現しようとするものである。

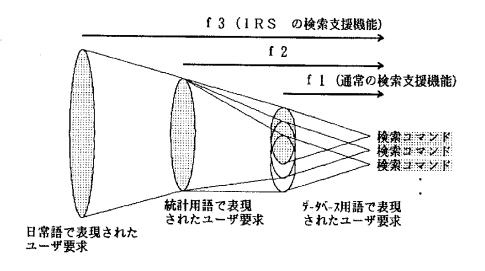


図5-5 検索支援機能概念図

(2) ユーザの求める情報が存在しない場合が多いので、次善の情報を選択する機能を持つ。

以下の図は、この機能を具体的に示したものである。

統計に詳しくないユーザは、実際にどのような統計にも収録されていない情報を求める場合が多い。同様にデータベースの収録情報を知らないユーザは、しばしばデータベースにない情報を要求する。このような場合に「データが存在しないので検索ができない」という回答を返すのでは、ユーザ要求の大部分に対応できないシステムになってしまう。

本システムでは、太線で囲んだユーザからの要求に対して、求められた通りの回答は、得られないものの、斜線で囲んだ情報が、次善の情報として、 提供される。 この機能により、ユーザは求めた情報に近い情報を得ることができ、与えられた情報で満足するか、別の情報を検索するか、という意思決定がしやすくなる。

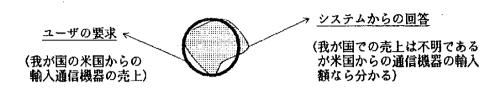


図5-6ユーザの要求に対するシステムの応答

以上、システムの特徴について概念的に述べた。このような特徴は、実際のシステム上では、下記のように現れる。

### ① 柔軟な入力機能を持つ

通常のデータベース検索システムは、メニュー方式で、限定された選択肢のなかから自分の希望に近いものを選択することにより、検索条件を決定するもが多いがIRSの場合は、日常語の世界で表現されたユーザ要求を入力させる都合上、入力が多種多様となることが予想される。

例えば、同じ意味の統計的な用語に対応する日常語は、通常複数存在し、メニュー方式では、ユーザの知っている用語が、メニューの中に存在しないケースが多くなることが、懸念される。

また、求める情報の内容も、統計的な表現を知らないユーザによって指定される ため、統計的なカテゴリーを用意するだけでは、入力が不可能となる場合が、考え られる。

このように統計的知識を知らないユーザからの入力を受け付けるためには、メニュー方式のほかに、自然言語処理機能を組み合わせる等の工夫をして、柔軟な入力機能を実現する事が必要になる。

#### ② 説明機能が充実している

ユーザが、統計やデータベースに関する専門家ではないため、検索された情報の 性質に関する説明をする必要がある。例えば、統計中の欠測データに関する説明や、 品目分類の定義に関する説明を必要に応じて行う機能が必要である。

#### ③ 次善の情報を提供する機能を持つ

求められた情報が存在しない場合に、その情報が統計として(または、データベースとして)存在しないことを示し、それに替わり得る情報を構成し、ユーザに提供する機能を持つ。

5. 2. 2 エンドユーザから見た機能概要 (外部仕様)

前節で述べたIRSの特徴を受けて、エンドユーザから見た、機能概要を整理する。

- (1) 【RSによるデータベース利用支援の内容
  - IRSが実現すべきデータベース利用支援機能を下記に列挙する。
    - ① 日本語による検索要求(表題程度)の入力を受け付ける機能
    - ② 不足情報の指摘や、検索条件の確認等のガイダンス機能
    - ③ 要求した情報が存在しないときに代替の情報を提供する機能
    - ④ 統計やデータベースの自動選定機能
    - ⑤ データベースの自動検索機能
    - ⑥ 検索履歴の保存と連続的な検索支援機能
    - ⑦ 類義語の定義に関して、ユーザ定義を受け付ける機能
- (2) ユーザインターフェイス
  - (1) で設定した機能を実現した場合IRSとエンドユーザの間の情報のやりとりは概略下記のイメージになると想定される。
  - ① 初期情報入力

初期画面では、IRSとユーザとの間の最初の接点となるキーワード (要求種別) を決定する。

(A)

検索する情報のタイトルを入力してください

→ わが国の発展途上国からの 製品輸入の割合 要求種別は、ユーザ要求を表現する為の 項目として、ユーザ世界の分析により 抽出された概念で、ユーザ要求を分類 する最初のキーワードである。

方式としては、(A) に示す表題の自然 言語入力によるものと、(B) に示すメ ニュー方式のものが考えられる。

自然言語入力の場合は、ユーザの入力の 手間を考え、出力すべき表のタイトルか らキーワードを抽出する。 欲しい情報の種類をメニューから 選んで下さい

- 1. 貿易額 101.\*\*の変動
- 2. 貿易量 102.\*\*の状況 3. 貿易収支 103.\*\*の比率

(C)は、上記の両者を選択できるようにしたものである。

(A)、(B)、(C)を組み合わせて 初期情報の入力を促進する。

(C)

入力の方法を選んで下さい

- 1. 欲しい情報を言葉で入力
- 2. 欲しい情報をメニュー選択
- → 1を選択すれば(A)画面へ 2を選択すれば(B)画面へ

### ② 補足情報入力

初期情報として、入力された情報や抽出されたキーワードに応じて必要な不足情報を問い合わせる

例1 (我が国の国民総生産に対する製品輸入の割合)

比率(A/B)のA、Bにどのような情報が対応しますか

- 1. 貿易額
- 2. 貿易量
- 3. 貿易収支
- 4. GNP

A → 1

 $B \rightarrow 4$ 

例1は、2種類の情報の間の演算や、比較を含むキーワードの場合(ここでは、 比率)

例2 (我が国の発展途上国からの製品輸入額)

貿易額に関して、下記の情報を入 力して下さい。

輸出入区分(輸出)→輸入

報告国(日本) →

相手国(世界) →発展途上国

品目 (上位10)→製品

時点 (最新1年)→

期種 (年次) →

助性 (午八) → 単位 (ドル建) → 例 2 は、単独の情報で、定義が明確に なれば、検索可能である例(貿易額)

→複数の系列を同時に指定することも可能 (例 ドル建、円建で)

## ③ 検索条件確認

システムが推論によって導いた検索条件をユーザが確認する。

下記の条件で検索を行います。

- 1. 品目は通白分類の定義
- 2. 国別内訳なし 3. 単位は円建て
- 4. 時点は最新 (1987)
- 5. データベースはTRADE 下線の番号には、他選択が可能で す他選択をする番号かりを入力し て下さい

他選択を希望する場合は、他選択を希望 する番号を入力する。 例えば2を入力すれば、 2. 国別内訳なし

- の替わりに
  - 2. 国別内訳あり
- と変更された画面が表示される。

## ④ 結果表示

IRSが呼び出したデータベースによる結果表示

ſ	貿易額 日本-発展途上国						
1	1988 (Fin) DB=TRADE						
	国名	金額					
			,				
		_					

## 5. 2. 3 システム機能構成

## (1) 機能構成図

利用者からの機能要求を実現するために、統計の専門家や、データベースの専門家の知識をAI的な手法で蓄積、利用するシステムを想定した。また主要機能とユーザ、データベースとの関係を、下図に示す。

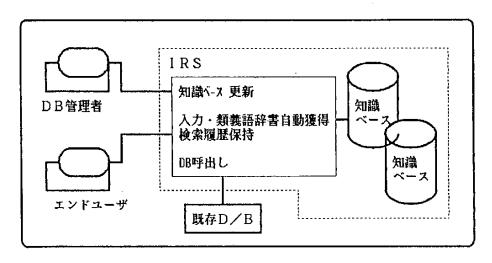


図5-7 機能とユーザ/DBとの関連図

システムとしての機能構成は下図のようになる。

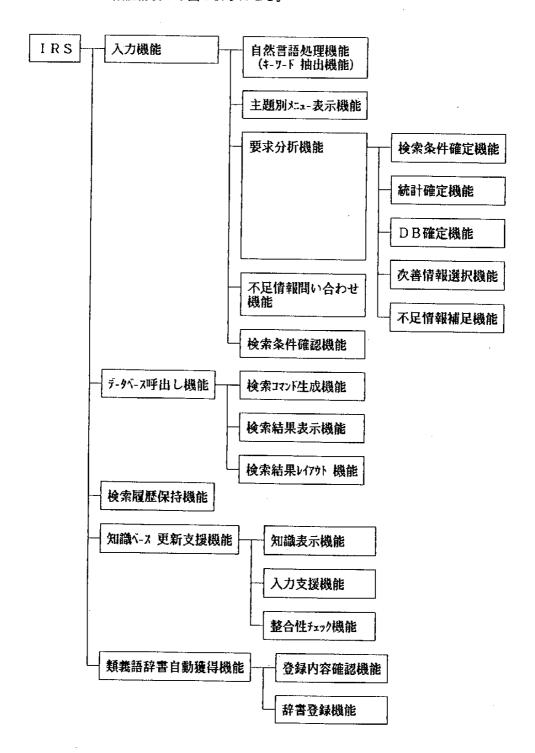


図5-8 機能構成図

それぞれの機能について、以下に説明する。

#### ①入力機能

ユーザからの検索要求を入力する機能である。自然言語処理によるインターフェイスとメニューによるインターフェイスの2種類を組み合わせてユーザからの入力を受け付け、システムの内部表現であるユーザ要求フレームに変換する。以下の部分機能を持つ。

#### ・自然言語処理機能

ユーザから入力された表題を分析し、情報種別、品目、国名等のキーワードを、特定する。意味解析は行わない。

## ・主題別にユー表示機能

メニューによるユーザ要求入力のガイダンス機能

#### · 要求分析機能

ユーザからの要求を分析し、システムの内部表現 (フレーム) に変換する機能、以下の部分機能より構成される。

## ·検索条件確定機能

ユーザからの入力から、検索に必要なキーワードを抽出し、共通用語に 変換する。

### ·統計確定機能

入力情報から参照すべき統計を確定する。

#### ·DB確定機能

入力情報から検索すべきデータベースを確定する。

#### ·次善情報選択機能

ユーザの求める情報が存在しない場合に次善の情報を構成する。

#### ・不足情報補足機能

検索条件を確定するうえで不足する情報を発見し、常識的に補足出来る 内容であるか否かの判定を行い、補足できる場合は補足する。

## ・不足情報問い合わせ機能

不足情報で、ユーザに問い合わせる必要のあるものに関して、ユーザに問い 合わせを行う。

### ・検索条件確認機能

確定した検索条件について、ユーザに確認の問い合わせを行い、必要に応じて検索条件の変更を行う。

### ②データベース呼出し機能

データベースを検索するためのデータベースとのインターフェイス機能。下 記の部分機能を持つ。

- ・検索コマンド生成機能 ユーザの要求に合わせてデータベースを検索するための、検索コマンド を生成する。
- ・検索結果表示機能 データベースの検索結果をそのまま表示させる機能。
- ・検索結果レイアウト機能 複数の検索結果を合成する場合等に、データベースからの検索結果を編 集する機能。

## ③検索履歴保持機能

再度検索を行う際に、質問や検索の重複を避けるために、検索の履歴を記憶 する機能。

## ④知識ベース更新支援機能

知識ベースを更新する際のユーザ支援機能。下記の部分機能を持つ。

- ・知識表示機能 既に存在する知識を分かりやすく表示する。
- ・入力支援機能

  ユーザが、必要最低限の操作で知識の入力ができるように、入力を支援し
  入力されたデータのチェックを行う機能。
- ・整合性チェック機能
  ユーザの入力した知識に関して、可能な範囲で矛盾のないことをチェックする。

#### ⑤類義語辞書自動獲得機能

IRSの辞書のなかで、品目等の類義語に関するものは、ユーザによって内容の変わる部分(ユーザ辞書)があるが、IRSを利用中にユーザ辞書の追加更新をおこなうための機能。下記の部分機能がある。

#### · 登録内容確認機能

今まで辞書にない用語が、ユーザ要求として、入力された時に、その用語を辞書として、登録するか否か、またどういう意味で登録すべきかをユーザに問い合わせる機能。

#### ·辞書登録機能

新たな用語定義を辞書に追加する機能。

## (2) システムの機能と知識との関わり

(1) に説明した機能が、どのような知識を使って実現され、ユーザとのインターフェイスとしては、どのようなイメージが得られるかを次ページに図示する次ページの例は、「わが国の発展途上国からの製品輸入の割合」という表題が、ユーザから入力された場合の例である。

はじめに入力機能によって入力文から「わが国(日本)」、「発展途上国」、「製品」、「輸入」、「割合」・・・というキーワードが切り出され、ユーザ要求のフレームが形成される。

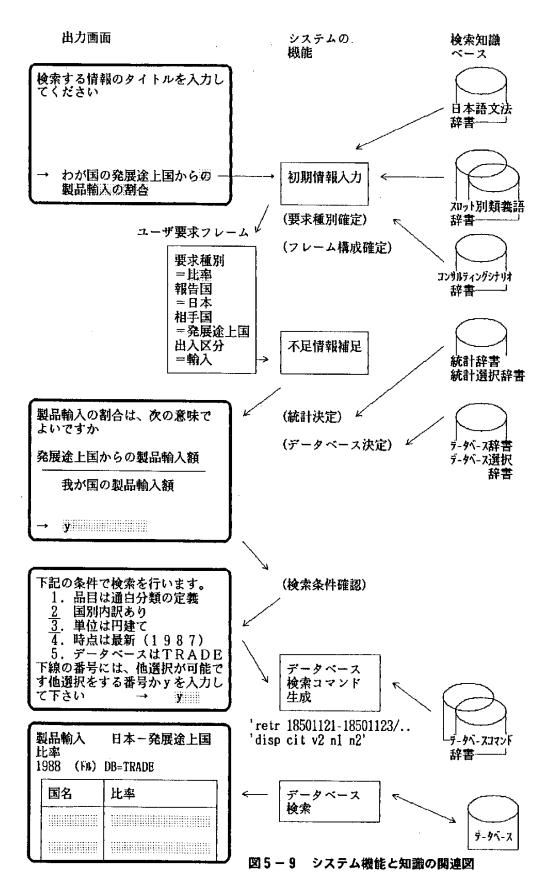
要求文から名詞句を抽出するために、日本語の文法に関する知識が必要となり、 わが国→日本 等の書換えを行うために、各種の類義語に関する知識(要求種別、 品目、国名、辞典、単位等)を使う。また情報検索に必要なコンサルティングのシ ナリオを決めるために、要求種別ごとのコンサルティング手順に関する知識が必要 である。

ここで、割合という言葉に複数の解釈があるので、どの意味であるかの問い合わせが行われ、検索条件が確定する。同時に、使用すべき統計の種類、データベースの種類が決定される。キーワードの統計的意味に関する知識、各種統計の性質に関する知識、データベースの収録状況に関する知識等が必要となる。

検索条件をユーザに確認するための問い合わせを行った後、データベース検索コマンドの生成を行う。このためには、データベースのコマンド体系に関する知識、

検索コマンドとユーザ要求フレームを結びつけるための知識が必要である。

最後にデータベースの検索を行い、データベースから出力された画面をそのまま 表示させる。



#### 5. 3 システムに必要な知識

本節では、システムの機能を実現するために必要な知識について、種類と構成を概観 する。

#### 5.3.1 検索支援に必要な知識

ユーザからの要求を理解しデータベース検索をおこなうことは、ユーザの要求を、何らかの(一般性を持った)形式で表現し、これを統計用語の世界に変換するための知識、さらにデータベース用語の世界に変換するための知識をシステム上に蓄積することと同義であると考えられる。(仮にフレーム変換型の知識処理と呼ぶことにする)

#### (1) ユーザ要求の表現方法

ユーザの要求は、日常語の世界、統計用語の世界、データベース用語の世界の3つの世界でそれぞれ表現される、本システムでは、ユーザの要求を人工知能の知識表現の一種であるフレームを使って表現する。フレームはユーザの要求を表現するための情報の枠組みである。

日常語世界と統計用語の世界ではユーザ要求は同じ形のフレームで表現され、フレーム を構成する個々の情報(スロット)の表現が、統計用語の世界では統計用語に置き換えられていると考えられる。(これを要求フレームと呼ぶ)

要求フレームを構成するスロットには、どのようなユーザ要求にも必要な必須スロットと要求の内容によって必要となる補助スロットの2種類がある。

データベース用語の世界では、ユーザ要求はデータベース検索に必要な情報に変換され、フレームの構成も先の2つの世界と異なる。 (これを検索データ構成フレームと呼ぶ)

次頁にこれらのフレーム構成と変換の状態を図示する。

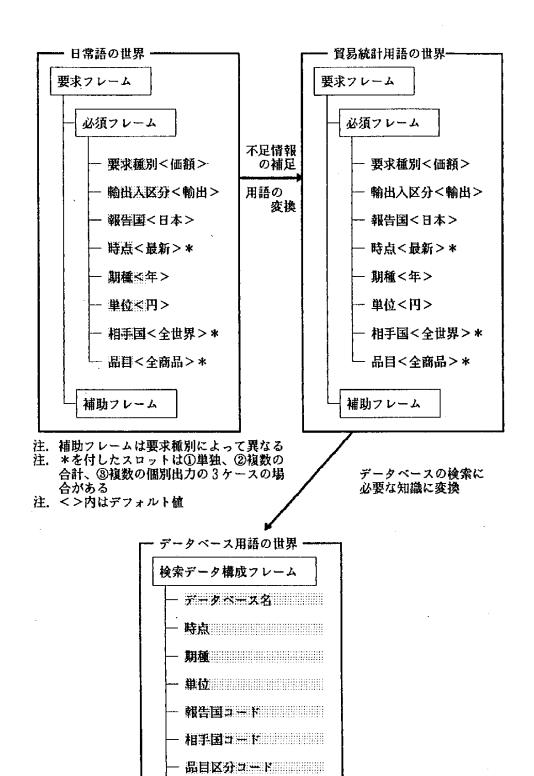


図5-10 ユーザ要求の表現と変換

データに関するコメント (要求と検索結果の違いに

関する情報)

#### (2) 検索のための知識構成

前章で述べた特徴を満たす検索支援を行うための知識を専門家のコンサルティ グをモデルにして、整理すると下図の知識構成が考えられる。 <>のなかは、次節で示す知識ベースとの対応を示す。

日常語の世界から貿易統計用語の世界へ要求を変換するための知識 日本語文法辞書 不足情報を補うための知識 (ユーザの要求を分類し、統計のレベルで必要な情報を補足する知識) 不足する情報を示し、ユーザに問い合わせを行う知識 要求種別ごとに必要な情報や、検索手順に関する知識 <コンサルティングシナリオ辞書> 検索すべき統計/データベースを特定する知識 統計/データベースの収録情報に関する知識 <データベース辞書・統計辞書> 要求を満足する統計/データベースを選択する知識 <データベース選択辞書・統計選択辞書> ユーザの指示した検索条件を正確に定義するための知識 「日常用語を統一された統計用語に置き換える知識) 品目分類を確定する知識 報告国、相手国の確定をする知識 単位を確定する知識

<スロット別類義語辞書>

#### 貿易統計用語の世界から貿易テータヘース用語の世界へ要求を変換するための知識

時点を確定する知識

検索データ構成を決定する知識 (データベース検索に必要な情報を整理する知識) データベース検索コマンドを作成する知識 < タータヘースコマント辞書> 指定された情報が存在しない時に、それに近い情報を示す知識 <コンサルティングシナリオ辞書> 5. 3. 2 知識ベースの具体例

検索に必要な知識を知識ベースとして捉えた時に、具体的には、どのような内容になるかを以下に概説する。

① スロット別類義語辞書(ユーザの指定した検索条件を正確に定義するための知識) 要求種別、国名、時点、単位、品目分類に関して、類義語の辞書を作成する。

例 貿易関係→貿易額 円高→為替レート 予測、増減、推移→変動

② 統計辞書(検索すべき統計/データベースを確定する知識)

統計に関する辞書

例 統計名 通関統計 データ形式 月次 発表時期 翌月13日 収録 外国貿易概況・日本貿易月報 品目分類 cccn コメント (計上価格) 輸出=FOB 輸入=CIF

③ 統計選択辞書(検索すべき統計/データベースを確定する知識)

例 要求種別が貿易額で商品区分が全商品→輸入報告統計

④ データベース辞書(検索すべき統計/データベースを確定する知識)

データベースを選択するために、データベースに関する収録情報の辞書を作成する

 例
 データベース名
 TRADE

 期間
 86·1~88·10

 機種
 月次

 欠測データ
 86·11

 単位
 円建て

 国
 〔米国、英国、・・・〕

 品目
 〔食品、工業製品・・・〕

⑤ データベース選択規則(検索すべき統計/データベースを確定する知識)

例 要求種別が貿易額で期間が1960年以前の場合→(国連統計) 要求種別が貿易額で期間が1960年以後の場合→TRADE

⑤ コンサルティングシナリオ辞書(不足情報を補うための知識) コンサルティングの手順を要求種別に応じて決定するための辞書を作成する。

例 要求種別 貿易額 手順 〔報告国、相手国、DB、品目、期間、期種、単位

⑦ データベースコマンド辞書データベースの検索コマンドを生成するための辞書

#### 5. 4 システム実現のための方向性

#### 5. 4. 1 知識ベース獲得の方法

前節で説明した知識ベースには、データベース管理者の知識、統計専門家の知識、データベースのエンドユーザの知識があると思われる。下記の表に各知識ベースの属性および考えられる知識獲得方法を示す。

知識ベース名	知識ベースの源泉	知識獲得方法
スロット別類義語辞書	データベース管理者及び エンドユーザ	通白分類等 ユーザによる更新
統計辞書	統計専門家	統計に関する書物
統計選択辞書	同上	同上
データベース辞書	データベース管理者	データベース仕様書
データベース選択辞書	データベース管理者及び 統計専門家	統計専門家よりヒアリ ング
コンザルティングシナリオ辞書	IRS開発者及び 統計専門家	統計専門家よりヒアリ ング
データベースコマンド 辞書	データベース管理者	データベース仕様書

表 5-1 各知識ベースの源泉および知識獲得方法

#### 5. 4. 2 初期試作システムの稼働環境(案)

初期試作システムの稼働環境に関しては、未だ十分な検討は、行われなかったが基本構成として、下図の構成を想定した。これは、ホストコンピュータ上で稼働する既存のデータベースシステムに対して、端末装置(パーソナルコンピュータ)上で稼働するシステムとしてIRSを構築するものである。

この方式の利点は、IRSの稼働時にホストコンピュータに負荷がかからないので、ホストコンピュータで稼働している他のシステムのレスポンスを落とす心配がない点である。

但しIRSの知識ベースのうち、品目名に関する類義語辞書等は、膨大な量になることが予想されるので、ホストコンピュータ側の磁気ディスクに置くか、CDーROM等のパーソナルコンピュータ用大容量補助記憶装置を利用する必要がある。

またこの方式の場合は、知識のメンテナンスの際に、各端末が同時に知識を更新 するように注意する必要がある。できれば、ホストコンピュータを介して、自動的 に全ての端末の知識ベースを更新する機能を持たせることが望ましい。

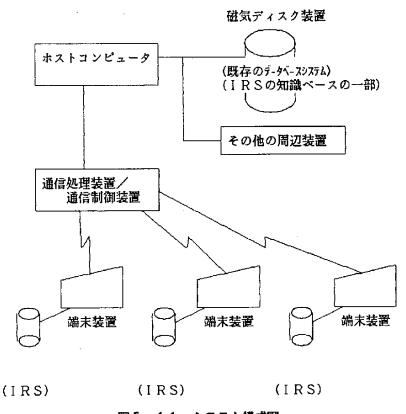


図5-11 システム構成図

#### 5. 4. 3 システム構築への課題

本章で報告したシステムイメージに関し、実用システムの開発にあたって再検討すべき課題として、下記の項目が挙げられた。

#### ①自然言語処理の適用範囲について

自然言語処理によって意味解析をおこなうのは、現状では困難であると思われる。特定のキーワードの抽出までとする等、適用範囲を設定する必要がある。

#### ②データの加工の機能について

データの加工機能は、各データベースが持っている機能の範囲としているが 本システムに特に追加すべき加工機能はないか。

#### ③判りやすいメニュー構成の工夫

対話型のシステムでは、メニューの階層が、深くなると、現在なにをしているかが判りにくくなる。インターフェイス機能を充実させて、全体のメニュー構成のなかで、現在の検索行動が、どのような位置づけであるかを判りやすくするとともに、メニューからメニューへの自在な移動を可能にすれば、メニュー方式の煩わしさが、軽減されると思われる。

#### ④検索目標の変更への対応

通常の検索行動を想定すると、検索中に、しばしば検索目標が変更する場合 があるので、このような場合の柔軟な対応を検討する必要がある。



本報告書では、3ヵ年計画の調査研究の始めの1年間の成果についての報告を行った。 委員会での検討を通じて、経済関連のデータベースについての貴重な実例が紹介され、経 済統計データベースに関する現状や、問題点が非常に具体的な形で認識されたことは、本 年度の大きな成果であると思われる。

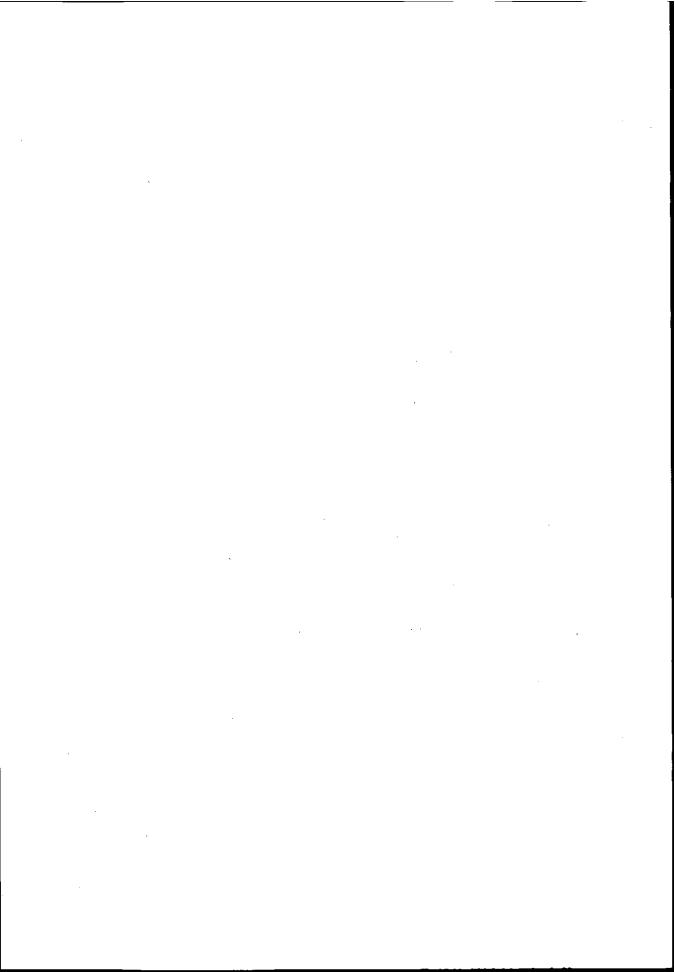
また、ユーザ世界モデルの分析作業を通じて、統計データベース (特に貿易統計) 検索 に関する専門家の問題解決手順をワークシート形式に収集することができた。本年度のシステムイメージ構築にも、この分析結果を使ったが、今後の実用システム設計開発において、重要なデータとして再度分析利用されてゆくものと思われる。

さらに海外調査によって欧米でのデータベース利用技術の動向が紹介され、本調査研究 に応用可能と思われる資料も、いくつか入手することができた。今後の調査研究活動を通 じて、これらの資料を活用し、海外の研究者との情報交換をしていくための基盤ができた という意味で有意義な海外調査となった。

次年度の調査研究のための足場として、経済統計データベースの利用現場からのAI適用可能性の指摘に基づき、システムの目標設定を行った。現在設定された目標は、システムのイメージとしては、まだ漠然としており、利用者の立場から、また開発者の立場から検討すべき多くの課題が指摘されている。

これらの課題を検討しつつ、次年度は、特にマンマシンインターフェースのあり方を中心にシステムの詳細な設計・試作を行い、最終年度には、実用システムの開発に着手する 予定であるが、実用システムが真に役に立つものとなるためには、さらに実際のデータベース利用現場を知り、的確な設計を行う必要がある。

本年度の調査研究の成果を総括するに当たり、調査研究に多大なご協力を戴いた委員各位に改めて謝意を表するとともに、今後の調査研究に対しても、引き続きご指導ご協力を 賜るよう切にお願いしたい。



付 録

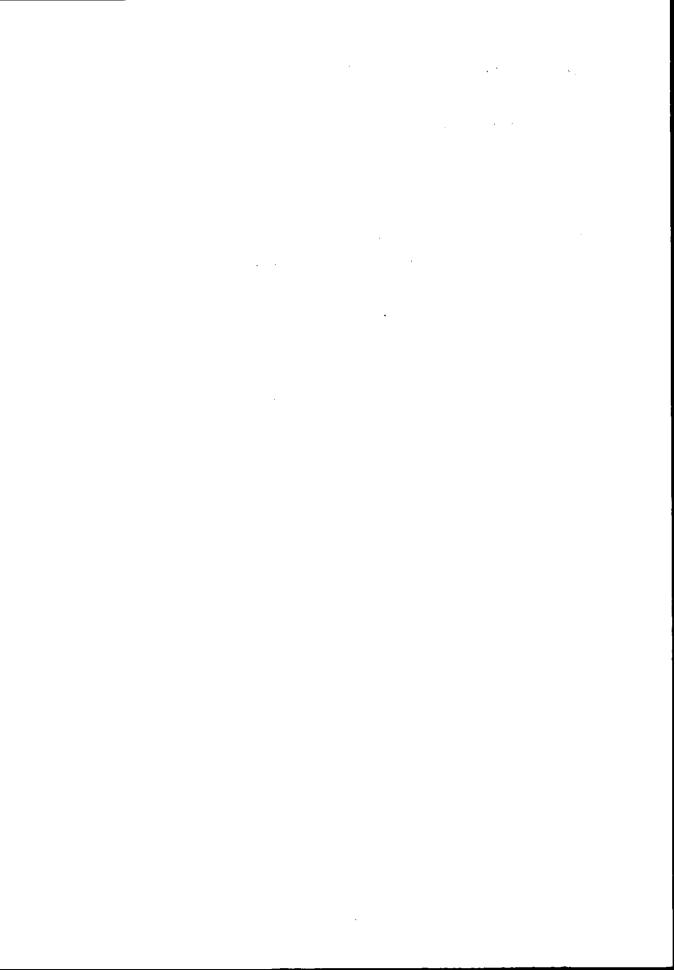
# 付録1 ケーススタディにおける知識の収集

- (1) ユーザ世界のサンプル
- (2) ユーザ世界の整理用シート
- (3) ユーザ世界の整理用シート記入例

### 付録2 ケーススタディにおける知識の整理

- (1) ユーザの世界と統計の世界を結ぶ知識の構造
- (2)検索時に必要となる知識の例
- (3)知識整理用シート記入例

付録3 ユーザ世界分析結果 (要求種別)



ケーススタディにおける知識の収集

# (1) ユーザ世界 (ユーザの要求分) のサンプル

No. 1

No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	MIL
1	ドル安で輸出はどの程度 伸びるか	ドル(単位) 輸出 伸び率		
2	最近の日本、韓国、台湾等 のアジア勢の対欧輸出攻勢 について知りたい	最近(時点)アジア 勢(地域) 対欧 輸出		
3	円髙で製造業分野の一部で 貿易面の障害が発生してい るかどうか知りたい	円高 製造業 貿易 障害		
4	61年度の中小企業の景況を 輸出の数量ベースで全体の と比較したい	61年度(時点)中小 企業 輸出 数量ベ ース 比較		
5	輸出関連の製造業の状況を 見たい。また、今後どうな るか知りたい	輸出 製造業 状況 今後		
6	1990の日本の輸出は何億ド ルになるか予測したい	1990 (時点) 輸入 ドル 予測		
7	アジアNICSからの輸入状況 を見たい	アジアNICS 輸入 状況		
8	わが国の米国からの製品並 びにサービスの輸入状況を 見たい	わが国 米国 製品 サービス 輸入 状況		
9	わが国の国民総生産(GNP) に対する製品輸入の割合を 知りたい	わが国 国民総生産 製品 輸入 割合		: : :
10	わが国の開発途上国からの 製品輸入の割合を知りたい	わが国 開発途上国 製品 輸入 割合		

資料名	通産ジャーナル
月号	1987.7
ページ	16~41

No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
11	わが国の米国からの輸入通 信機器の売上に関する統計 が見たい	わが国 米国 輸入 通信機器 売上 統計		
1 2	わが国の米国からの通信機 器の購入率を知りたい	わが国 米国 通信 機器 購入率		
1 3	米国から日本へのハイファ イのステレオスピーカーの 輸出について知りたい	米国 日本 ハイファイ ステレオスピーカー 輸出		
14	わが国の外国製品の消費の 状況を知りたい	わが国 外国製品 消費 状況		
1 5	1986年の鉱工業生産の前年 比、出荷の値が知りたい	1986年 鉱工業生産前年比 出荷 値		
1 6	1986年の輸出数量指数の前 年比の値が知りたい	1986年 輸出数量指数 前年比 値		
17	輸出産業におけるコスト削 減の状況を見たい	輸出産業 コスト削減 状況		
1 8	OEM 輸出、部品・原材料輸 出の状況を見たい	OEM 輸出 部品・原材料輸出 状況		
19	輸入報告額と前年同月比が 知りたい	輸入報告額 前年同月比		
2 0	輸出確認額と前年同月比が 知りたい	輸出確認額 前年同月比		

資料名	通産ジャーナル
月号	1987.7
ページ	42~97

				NO. 3
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	トジア
2 1	82年から85年までの米国の 貿易赤字はどのくらいか (IMF ベース)	82年から85年まで 米国 貿易赤字 IMF		
2 2	米国の貿易赤字と日本の貿 易黒字に占める為替レート の割合を知りたい	米国 貿易赤字 日本 貿易黒字 為替レート 割合		
2 3	ヨーロッパのエアバスの輸 出の状況を知りたい	ヨーロッパ エアバス 輸出 状況		
2 4	1981年から現在までの米国 日本、西独の工作機械の輸 出を比較したい	1981年 現在 米国 日本 西独 輸出 工作機械 比較		
2 5	わが国の米国スーパーコン ピュータ関連部品の輸出量 を見たい	わが国 米国 スーパーコンピュー タ 部品 輸出量		
26	韓国からカナダへの自動車 の輸出量を見たい	韓国 カナダ 自動車 輸出量		
27	日本の石油輸入量が1974年 以降の増減を知りたい	日本 石油輸入量 1974年以降 増減	-	
2 8	日本の中国からの農産物の 輸入量を見たい	日本 中国 農産物 輸入量	.W	
2 9	日本のインドからの農産物 の輸入量を見たい	日本 インド 農産物 輸入量	No.	
3 0	日本と発展途上国、共産国 累積債務国との貿易関係を 知りたい	日本 発展途上国 共産国 累積債務国 貿易関係		

資料名	通産ジャーナル
月号	1987.8
ページ	13~61

Na	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
3 1	貿易収支の不均衡について 調べたい	貿易収支 不均衡		
3 2	西独の貿易黒字について知 りたい	西独 貿易黒字		
3 3	日本と西独の貿易黒字につ いて対GNP 比の状況を知り たい	日本 西独 貿易黒字 対GNP 比 状況		ļ
3 4	日本の対米貿易黒字のドル ベースの増加率を知りたい			
3 5	日本の米国への輸出につい て数量ベースで知りたい	日本 米国 輸出 数量ベース		
3 6	米国及びアジアからの輸入 品の銘柄ごとの価格を知り たい	米国 アジア 輸入品 銘柄 価格		
3 7	中南米からのコーヒーの輸 入量について知りたい	中南米 コーヒー 輸入量		
3 8	韓国をはじめとするNICS諸 国からの鋼材の輸入動向に ついて知りたい	韓国 NICS諸国 鋼材 輸入動向		
3 9	日本の輸出産業に占める自 動車産業の割合が知りたい	日本 輸出産業 自動車産業 割合		
4 0	わが国の家電メーカーのヨ ーロッパへの輸出状況につ いて知りたい	わが国 家電メーカ ー ヨーロッパ 輸出 状況		

資料名	通産ジャーナル
月号	1987.12
ページ	9~101

				IW U
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
4 1	日本の石油輸入の湾岸諸国 への依存度を知りたい	日本 石油輸入 湾岸諸国 依存度		
42	太平洋地域の石炭輸出の状 況について知りたい	太平洋地域 石炭輸出 状況		
4 3	87年の日本の製品輸入の 増加率について知りたい	87年 日本 製品輸入 増加率		
4 4	東アジアからの製品輸入の 傾向について知りたい	東アジア 製品輸入 傾向		
4 5	87年の製品輸入の伸び率 について知りたい	87年 製品輸入 伸び率		
4 6	ビデオメーカーの輸出の動 向を知りたい	ビデオメーカー 輸出 動向		
47	鉄鋼、造船における日本と 西独との貿易関係を知りた い	鉄鋼 造船 日本 西独 貿易関係		
4 8	精密機器における日本と西 独の貿易関係について知り たい	精密機器 日本 西独 貿易関係		
49	電気機器における日本と西 独の貿易関係について知り たい	電気機器 日本 西独 貿易関係		
5 0	自動車、輸送機器における 日本と西独の貿易関係につ いて知りたい	自動車 輸送機器 日本 西独 貿易関係		

資料名	通産ジャーナル
月号	1988.1
ページ	1 2~2 5

				W. O
Na.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
5 1	85年の輸出に占める投資 関連財の比重について日、 米、西独の比較をしたい	85年 輸出 投資 関連財 比重 日 米 西独 比較		
5 2	85年の原材料の輸入の全 輸入額に占める割合を日、 米、西独で比較したい	85年 原材料 輸入 全輸入額 割合 日 米 西独		
5 3	日本カメラの西独への輸出 量を知りたい	日本西独物出量		
5 4	85年の西独の工業関係の 輸出入(実績価額)の構成 比を知りたい	85年 西独 工業関係 輸出入(実績価額) 構成比		
5 5	85年の西独の工業関係の 半製品、製品の総輸入額と 総輸出額の比を見たい	85年 西独 工業関係 半製品 製品 総輸入額 総輸出額 比		
56	わが国の83~86年度ま での製品輸入額(億ドル) の推移を見たい	わが国 83~86年度まで 製品輸入額 推移		
5 7	わが国の83~86年度ま での製品輸入比率を知りた い	わが国 83~86 年度まで 製品輸入比率		
5 8	鉄鋼の輸出入に関する今後 の展望が知りたい	鉄鋼 輸出入 今後 展望		
5 9	化学の輸出入に関する今後 の展望が知りたい	化学 輸出入 今後 展望		
6 0	一般機械の輸出入に関する 今後の展望が知りたい	一般機械 輸出入 今後 展望		

資料名	通産ジャーナル
月号	1988. 2
ページ	11~95

				<u>iNO. 1</u>
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
6 1	電気機械の輸出入に関する 今後の展望が知りたい	電気機械 輸出入 今後 展望		
6 2	自動車の輸出入に関する今 後の展望が知りたい	自動車 輸出入 今後 展望		
6 3	繊維の輸出入に関する今後 の展望が知りたい	繊維 輸出入 今後 展望		
6 4	わが国の西独からの機械機 器、化学品の製品輸入の前 年比が知りたい	わが国 西独 機械機器 化学品 製品輸入 前年比		
6 5	86年以降の輸出と成長率 の関係が知りたい	86年以降 輸出 成長率 関係		
6 6	日本の輸出依存度はGNP の何割か知りたい	日本 輸出依存度 GNP 何割		;
6 7	アセアン諸国の輸出状況に ついて知りたい	アセアン諸国 輸出状況	·	
6 8	わが国のASEANからの 輸入額について米国と比較 したい	わが国 ASEAN 輸入額 米国 比較		
6 9	わが国のASEANからの 輸入額について製品輸入に 限定して米国と比較したい	わが国 ASEAN 輸入額 製品輸入 限定 米国 比較		
70	製造業、農業、サービスに おける日米間の取引き状況 について知りたい	製造業 農業 サービス 日米間 取引き 状況		

資料名	通産ジャーナル
月号	1988.3
ページ	1 2~8 2

No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	レベル
7 1	87年の韓国の総輸出額と 世界ランキングについて知 りたい	87年 韓国 総輸出額 世界ランキング		
7 2	韓国の繊維製品、電気製品 自動車、履物等の輸出額が 知りたい	韓国 繊維製品 電気製品 自動車 履物 輸出額		
73	韓国の輸入の増加率と貿易 黒字、対外債務残高等につ いて知りたい	韓国 輸入 増加率 貿易黒字 対外債務残高		
7 4	韓国の対米貿易の黒字額に ついて86年と87年の値 が知りたい	韓国 対米貿易 黒字額 86年と87年 値		
7 5	韓国の日本や産油国との貿 易赤字の状況が知りたい	韓国 日本 産油国 貿易赤字 状況		
7 6	対米黒字額の上位4位まで が知りたい	対米黒字額 上位 4 位まで		
77	韓国の米国からの牛肉、煙 草の輸入状況が知りたい	韓国 米国 牛肉 煙草 輸入 状況		
7 8	わが国の発展途上国からの 輸入状況について知りたい	わが国 発展途上国 輸入 状況		
79	ASEAN諸国向けの日本 の直接投資の状況について 知りたい	ASEAN諸国 日本 直接投資 状況		
8 0	中南米からの米国への輸出 状況について知りたい	中南米 米国 輸出 状況		

資料名	通産ジャーナル
月号	1988.4/1988.5
ページ	45~72/12 ~22

				NU 5
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	1674
8 1	発展途上国への資金フロー (貸付) の推移が見たい	発展途上国 資金フロー(貸付) 推移		
8 2	発展途上国向けの純資金フローの推移が知りたい	発展途上国 純資金フロー 推移		
8 3	わが国の貿易収支の対GN Pの推移が見たい	わが国 貿易収支 対GNP 推移		
8 4	英国の貿易収支の対GNP の推移が見たい	英国 貿易収支 対GNP 推移		
8 5	輸入品の価格が為替レート の変動に伴って下がってい るか確かめたい	輸入品 価格 為替レート 変動		
8 6	輸入消費財の価格変動につ いて知りたい	輸入消費財 価格変動		
87	62年の非耐久消費財の輸入について西ヨーロッパ、 アメリカ、カナダ、中国、 その他東南アジアの状況を 見たい	62年 非耐久消費財 輸入 西ヨーロッパ アメリカ カナダ 中国 その他東南ア ジア 状況		
8 8	62年の耐久消費財の輸入 について西ヨーロッパ、ア メリカ、カナダ、中国、そ の他東南アジアの状況を見 たい	62年 耐久消費財 輸入 西ヨーロッパ アメリカ カナダ 中国 その他東南ア ジア 状況		
8 9	輸入ブランド品の輸入価格 と小売価格が知りたい	輸入ブランド品 輸入価格 小売価格		
9 0	輸入ブランド品の輸入全体 に占める割合を知りたい	輸入ブランド品 輸入全体 割合		

資料名	通産ジャーナル
月号	1988.5/1988.6
ページ	24~26/ 70~72

		·	<u> </u>	0.10
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	トベル
9 1	為替レートの変化と日本の 輸出価格変動を対比させて みたい	為替レート 変化 日本 輸出価格変 対比		
9 2	為替変動があってもドル建 ての輸出価格が変化しない グループが知りたい	為替変動 ドル建て 輸出価格 変化 グループ		
93	為替変動の輸出価格への転 嫁率について知りたい	為替変動 輸出価格 転嫁率		
9 4	通関収支による昨年の米国 の赤字と日本、西独の黒字 について知りたい	通関収支 昨年 米国 赤字 日本 西独 黒字		
9 5	昨年一年の日本の輸入は全 体で何%増えたか知りたい	昨年一年 日本 輸入 全体 何%増	. • .	
96	昨年一年の日本の輸入の中 で製品輸入は何%増えたか 知りたい	昨年一年 日本 輸入 製品輸入 何%増		
9 7	昨年一年の日本の輸入の中 で消費財は何%増えたか知 りたい	昨年一年 日本 輸入 消費財 何%増		
98	日本国内で売られているカ メラはどのくらいが輸入品 が知りたい	日本国内 カメラ 輸入品		
99	香港、韓国からの下着の輸 入量について知りたい	香港 韓国 下着 輸入量		
100	製品輸入の状況と輸入構造 の変化を見たい	製品輸入 状況 輸入構造 変化		

資料名	通産ジャーナル
月号	1988.6/1988.8
さって	70~72/ 61~72

# <u>No. 1 1</u>

	· · · · · · · · · · · · · · · · · · ·			
No.	経済(貿易)活動表現事例	基本構成要素の抽出	経済統計コメント	M
101	資本財輸出の伸びについて 見たい	資本財輸出 伸び	·	
102	発展途上国の累積債務合計 に占める日本の負担の大雑 把な推計が知りたい	発展途上国 累積債務合計 日本 負担 大雑把 推計		
103	日本の直接投資についてア ジアの発展途上国、中南米 アフリカについて比較した い	日本 直接投資 アジア 発展途上 中南米 アフリカ 比較	·	
104	発展途上国の対外債務残高 の推移について知りたい	発展途上国 対外債務残高 推移		

資料名	通産ジャーナル
月号	1988.8/1988.9
ページ	61~72/ 13~61

# (2) ユーザ世界(ユーザの要求分)の整理用シート

経済統計に関するコメント No.	レベル
	*分類区分
<b>味</b>	
<b>香</b>	·······
<b>홲</b>	
経 済 計	,
<b>新</b> 計	
使 用	
D / / B	

経済統計に関するコメント No.	MIL	
<b>*</b>	*分类	類区分
<b>集</b> 不		·
<b>E</b>		
<b>律</b>		
報		
経 済 統 計		
世		
使 用 D B		
В		

# (3) ユーザ世界(ユーザの要求分)の整理用シート記入例

	経済統計に関するコメント	No.	2	ሁላĵ <b>!</b>	
意味	①アジア各国の対ヨーロッパ輸出総額を調べる ②輸出の内容を品目別に調べる			*分	類区分
不足情報	①アジア勢をどの国にするか、欧州を西欧という地域 最近をいつからとするか ②輸出の内容を品目別に調べるには、国により状況か 地域、時点の明確化が必要				るか、
経済統計	①DOT (Direction of Trade) によりアジア各国の調べられる ②日本については、我が国の通関統計で調べられる。 輸出については国連貿易統計(アジア経済研究所)で れの年データなので最近時点と言う点は問題がある ③輸入国側からの統計では、先進5カ国貿易統計によ てのみ1年遅れで年データが調べられる	、E :調へ	本以外( (られる)	のアジア( が、2 –	の国の 3 年遅
使 用 D / B	①DOT (Direction of Trade) によりアジア各国の 調べられる ②日本貿易情報システム ③先進5ヵ国貿易情報システム	対ヨ	<b>リーロッ</b> /	べの輸出	念額が

	経済統計に関するコメント	No.	7	レベル	
意味	アジアNICSからの <u>輸入状況</u> について	<del>-</del>		*分	類区分
不足情報	何時から調べるか、品目別状況は品目コードの何桁で がある	調へ	くるかを	明確にする	る必要
经游析計	我が国の通関統計を使うことにより調べられる				
使 用 D B	日本貿易情報システム				

#### 事例「アジアNICSからの輸入状況を見たい。

	<b>密発材料に関する</b> コメンド   Na. 7	L/VIL	
英	①わか国へのアジアNICSからの輸入額の推移を調べる。 ②わか国の輸入額に占める、アジアNICSからの輸入の割合の推移を調べる。 ③わか国へのアジアNICSからの輸入について品目別に推移を調べる。 etc.	* <i>分</i>	<b>区分</b>
不足情報	(1)駅食期間、期種(月次、四半期、年次、 c t c . ) (2)品日別に調べる場合には、対象品目の定義		
<b>組资基品</b>	大蔵省「通典解制」 (但し、1988/1月のデータから品目コードが設定された為、対象品目の定義に、 連続性が保たれないケースもある)	よりデー	-タの
使用ロベロ			

#### 事例「わか田の開発途上国からの製品輸入の割合を知りたい

経済機能に関するコメント No. 1 0 1/11/ \*分類区分 ①わか国の金製品輸入に占める開発金上国からの割合を調べる。 ②開発途上国からの輸入に占める製品輸入の割合を調べる。 诛 (1)關查期間,期種(月次,四半期,年次、 e t c.) 不足情報 (2)「製品」の定義 (8)「開発金上国」の定義(対象とする国を明確化) 大蔵省「通界統計」 粧 海新針 (参考) 開発途上国の輸出に占めるわか国のシェアを知りたい場合には「Commodity Trade Statistics」 (UN): 年次データで2~3年遅れ 慎用ロン

	経済統計に関するコメント No.	1 3	レベル	
意味	米国から日本への <u>ハイファイステレオスピーカー</u> の輸出に べる	ついて調	*分	領区分
不 足 情 報	ハイファイスピーカーについての品目定義を確認する必要	<b>きがある</b>		
経済統計	①先進5カ国貿易統計により、2-3カ月遅れでアメリカられる ②日本の輸入については我が国に通関統計で調べられる	の輸出月テ	<b>ニ</b> ータ ?	を調べ
使 用 D B	①先進5ヵ国貿易情報システム ②日本貿易情報システム			

	経済統計に関するコメント Na 1.1	I/\II	<u> </u>
意味	我が国の米国からの <u>輸入通信機器</u> の売り上げを調べる	*分	類区分
不足情報	通信機器の定義を明確化する必要がある		
経済統計	該当する統計はないが、輸入額なら我が国の通関統計を使うことに る	より調べ	·6n
使 用 D B	日本貿易情報システム		

		Nα	1 9	l/VI	
*				*5)\$	創区分
(1) 大型 <b>建</b>	(1)対象商品(総額、商品別) (2)対象地域・国			.,	
<b>建资资</b>	・通産省「輸入報告統計」 商品別は全地域計、地域・国別は全商品合計。 但し、アメリカ、EC、東南アジアの3地域に関しては主要商品の重大成省「通関統計」の連報より5日程度、確報より1ケ月程度発表が約価格でFOBも含まれる為、CIFの通関統計より金額は低い。	加力	がわかる。 'a 又。:	<b>集計仙</b>	名は製
世界の入込					

# 事例「輸出確認額と前年同月止が知りたい。

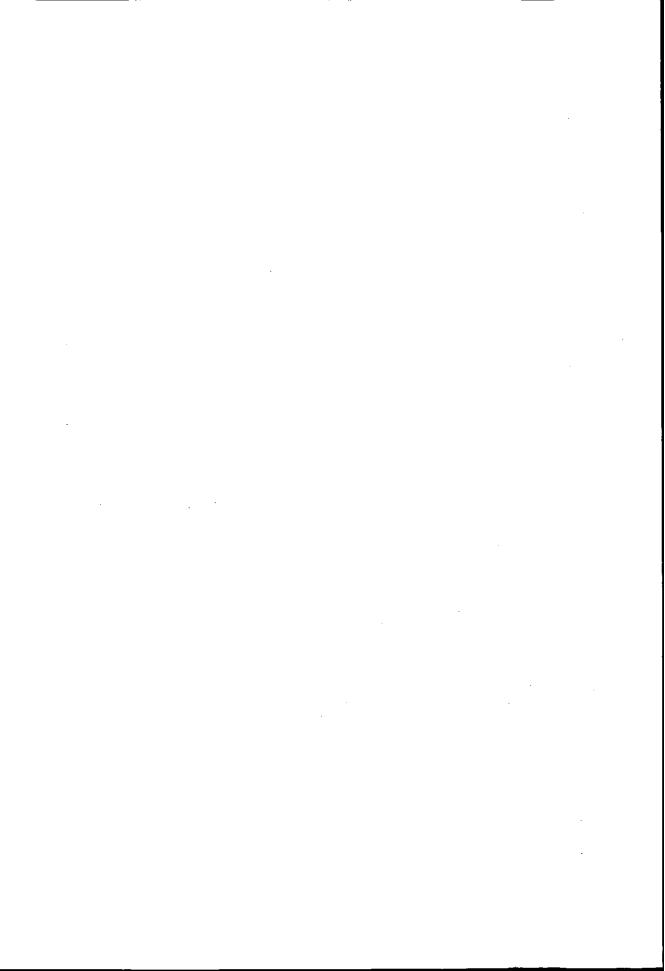
	No. 2(	ו (	M	
¥ X	①輸出確認統計の各月の値と、前年当月比を調べる。		* <del>/)</del> *	10分
	(1)对象商品(総額、商品別) (2)対象地域・国			·
<b>建肉类</b>	・通産省「輸出確認が計」 商品別は全地域計、地域・国別は全商品合計。 但し、アメリカ、EC、東南アジアの3地域に関しては主要商品の動向がわかる。 大蔵省「通界統計」の連報より5日程度、確報より1ケ月程度発表が早い。			
を持つくの				

# 事例「日本の石油輸入量の1974年以降の増減を知りたい。

	経済がはは異するコメント No	2 7	Nor	
*	①日本の石油輸入量の1974年以降の推移を調べる。 ②日本の石油輸入量を1974年を基準として指数化して見る。		*5	取分
<b>基基</b> 加州	(1)隅査期間,期種 (2)品日の定義(原油のみか、石油製品も含めるかetc.)			
推然版品	(a)大蔵省「通換統計」 (b)通産省「石油制御・連報」、確報値は「エネルギー生産・隔給統計月報」 (c)石油連盟「石油資料月報」((a)。(b)を含めた各種統計より作成)			

# 事例「日本と発展途上国、共産国、緊積債務国との貿易関係を知りたい。

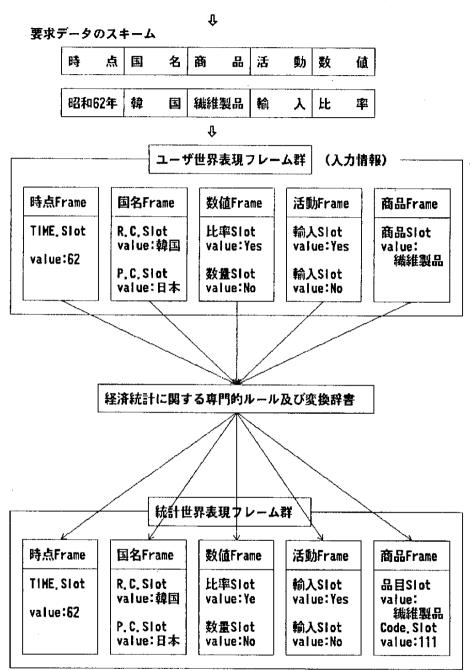
	Ma 30 Vol
新華	①日本と発展途上国、共産国、累積債務国との輸出入額を調べる。 ②日本と発展途上国、共産国、累積債務国との輸出入の構成品目の傾向を調べる。 ③日本の貿易に占める発展途上国、共産国、累積債務国との貿易の割合を調べる。等
不进實施	(1)調査期間 (2)発展途上田、共産国、累積債務国の対象国 (8)品目構成等を調べる場合、調査対象品目の定義
経済統計	(a)大蔵省「通関統計」 「外国貿易概況」に発展途上地域、共産国との商品別輸出入額データが掲載 (b)類記3地域の輸出入に占める日本の割合を調べたい場合には 「Commoditty Trade Statistics」(UN),年次データ
ながらくる	



ケーススタディにおける知識の整理

# (1) ユーザの世界と統計の世界を結ぶ知識ベースの構造・

「昭和62年の韓国からの繊維製品の輸入比率のデータが欲しい」



\*入力情報は必ずしも満たされているわけではなく、実際には不足した情報の場合のほうが多い。そのような時、ユーザの要求を確認し、統計世界の構成要素を聞き出す機能として、Ask、User機能があり、不明の場合には、入力情報から不足情報を推定する推論機能が必要となる。このような機能を「経済統計に関する専門的ルール及び変換辞書」によって実現する。

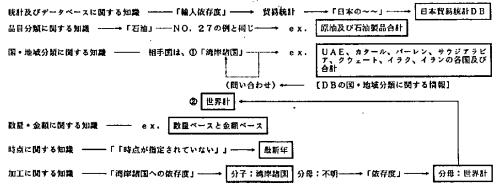
#### (2) 検索時に必要となる知識の例

#### NO. 41

日本の石油輸入の清岸諸国への依存度を知りたい。

① 使用するデータベースを探す。
 輸入 → (貿易) → (貿易・ ) → (収録計 → (収録計データベース)
 日本の石油輸入 ○ 日本貿易統計データベースの検索・表示に必要な条件。
 輸出人区分 → 輸入 → 「石油」の概念があい

#### ③ 必要となる知識例。



#### (参考) エネルギー用語辞典

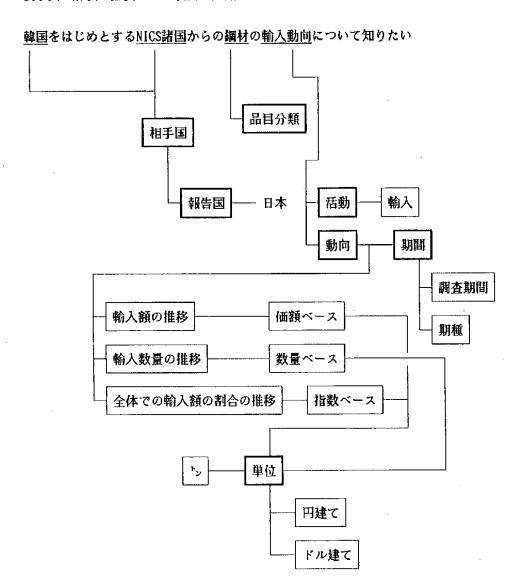
#### 海岸指国:

「石油産業の場合、一般にベルシャ湾(ニアラビア湾)岸諸国のことをいうが、具体的には<u>アラビア首長国連邦(UAE</u>)、 <u>カタール、バーレン、サウジアラビア、クウェート、イラク、イラン</u>をさす。 J

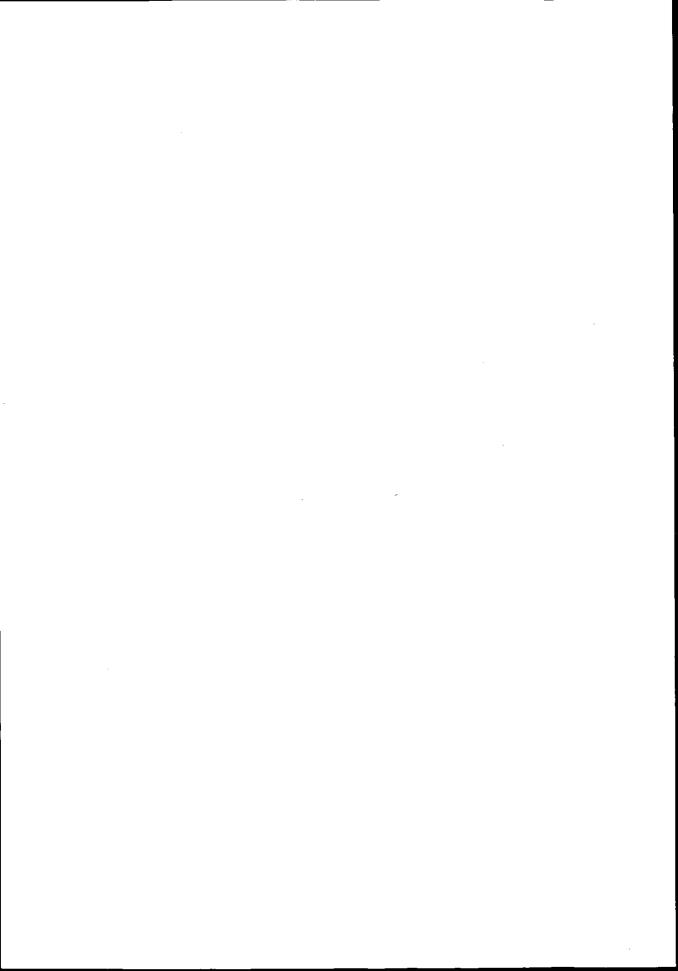
## (3) 知識整理(比較) 用シート記入例

No.38	韓国をはじめとするNICS諸国からの鋼材の輸入動向について知りたい				
	意味	不足情報	経済 統計		
(1)	①輸入額の推移 ②輸入数量の推移 ③全鋼材の内でNICS諸 国からの輸入額が占 める割合	調査期間 期種 鋼材及び各品目の定義	大蔵省 通関統計		
(2)	わが国にNICS諸国から の鋼材の輸入動向を求 める	鋼材の分類の定義 国別かNICS全体か 動向⇒数量ベース、指数 ベース、価額ベース 指数の定義	日本貿易統計 通商白書		
(3)	•				
	· · ·				
(4)					
(5)					
(6)					
	: :				

### <要求文の解釈に必要となる項目と知識>



ユーザ世界分析結果 (要求種別)



ユーザ世界分析シートを使って、ユーザ要求フレームを構成する要求種別スロットの

内容を整理した。

要求種別の形式として 下記の3通りが存在すると思われる。

- ① A (情報の内容)
- ② A (①と同じ) の B (表示や加工の方法)
- ③ ②の特殊の場合で Bが比率、影響等の場合

A (①と同じ) <u>の</u> A' (①と同じ) に対する B

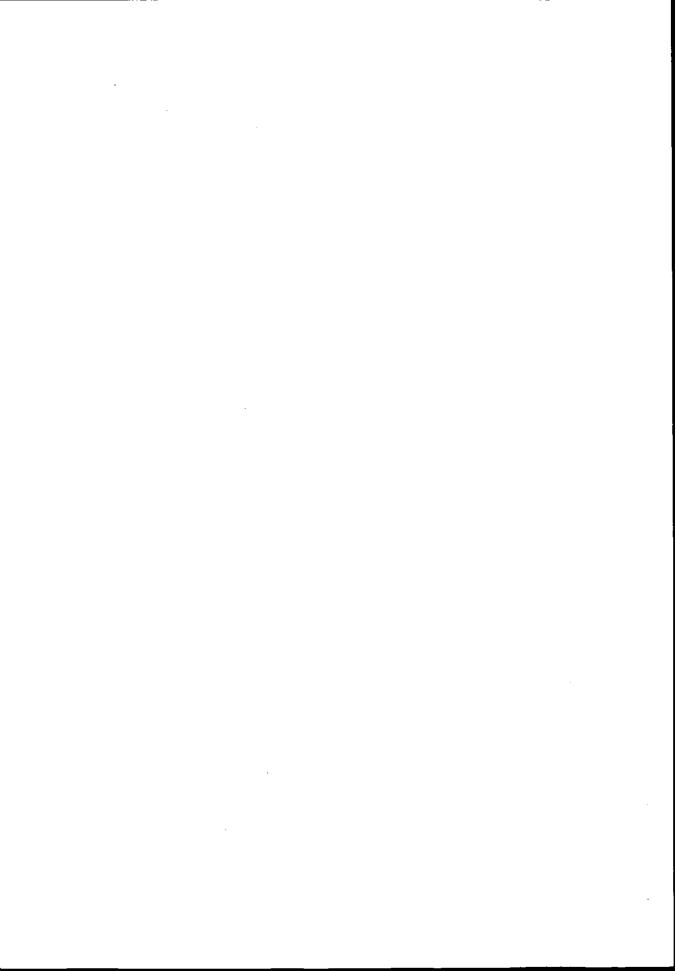
次ページにユーザ世界分析シートから抽出したAとBの内容を示す。

A	他のスロットへの作用
輸出入の数量	単位
輸出入の金額	単位
貿易量	輸出入区分→出入・・・輸出量、輸入量をそれぞれ出力
貿易髙	輸出入区分→出入 単位
消費高(量)	
貿易関係→貿易髙	輸出入区分 単位
貿易赤字	輸出入区分 単位
貿易黒字	輸出入区分 単位
貿易収支	輸出入区分
取引高 (量)	輸出入区分
対外債務残髙	輸出入区分 単位
直接投資髙	輸出入区分 単位

A (つづき)			,
貸付資金フロー	輸出入区分	単位	÷ .
純資金フロー	輸出入区分	単位	:
為替レート	輸出入区分		
国内販売量	輸出入区分		
国内販売高	輸出入区分		
(輸出入)構造			. •
国民総生産(GNP)	輸出入区分	単位	
円髙→為替レート	輸出入区分		9 (1)

В	他のスロットへの作用
予測→変動	時点
影響	時点
変動	
比較	
状况	
比率	
割合→比率	
推移→変動	
増減→変動	
動向→変動	時点
依存度→比率	
伸び率	
比重→比率	

B (つづき)		
世界ランキング		
上位X位		
		-



### 参考文献

#### 第1章

「データベース・サービスに関するユーザ意識調査」 DPC 1988年3月 「国土情報の知識処理」 国土庁計画・調整局編 1986年 「人工知能の技術と需要」 (財) 日本情報処理開発協会 1988年7月 「AIビジョン」

「Expert Systems」summer Vol.4 No.2 (株)CSK総合研究所 1988年

### 第2章

「アジ研ニュース (NO. 50)」 アジア経済研究所 1984年 「AID・XT I, II」 アジア経済研究所 1983年

### 第3章

「Expert Systems」spring Vol.3 No.1 (株) CSK総合研究所 1987年 「Expert Systems」summer Vol.4 No.2 (株) CSK総合研究所 1988年 穂鷹良介、佐藤英人 「統計データベースの設計と開発」 オーム社 1988年

#### 第4章

A.Claydon and M.Croft (U.K) 「統計コンサルテーション・誰がエキスパートか?」
P.L.Darius (Belgium) 「既存の統計ソフトウェアに知識ベース・モジュールを付加する手段」

P.de Greef (The Netherland) 「統計的問題に関する協力的解法」

J.C.Dumer Ⅲ,et al.(USA) 「データ分析のためのコンサルテーション・システムNONPARE」

P.Hietala (Finland) 「時系列分析のためのエキスパート・システム EST-ES」

B.de Vill and Ed Suen (Canada) 「データ分析のためのインタアクティブ・ツール KnowledgeSeeker」

- E.Cholawsky (USA) 「知識獲得の方法としての計量心理学的方法」
- V.L.Sauter and L.A.Madeo (USA) 「エキスパート・システムを利用者向にするための統計利用」
- E.Dawbroise and P.Massotte (FRANCE) 「エキスパート・システムMUSEにおける統計手法表示・・・統計手法の選択」
- W.DuMouchel (USA) 「MULREGの構造、設計方針、および戦略」
- D.J.Hand (U.K) 「統計知識ハンエンス・システム」
- D.J.Lubinsky (USA) 「統計データベースに統計を教える」

### ─ 禁 無 断 転 載 ───

平成元年3月 発行

発行所 財団法人 日本情報処理開発協会

東京都港区芝公園3丁目5番8号

機械振興会館内

TEL 03(432)9372

印刷所 株式会社 方 泉 社

TEL 03(253)2771

	·		

