

01-R 007

情報処理技術の応用に関する 調査研究報告書

—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 2 年 3 月



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
受けて平成元年度に実施した「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。

は じ め に

近年における情報化の進展とともに、データベースの構築が急速に進められている。

対象となる分野は拡大し、利用される局面も、従来の個別専門分野から、日常生活レベルに至るまでより多面化してきており、しかも必要に応じて遠隔地からでも瞬時に入手できるようにになっている。

しかしながら、現実利用者がデータベースにアクセスするためには、各データベースの所在、構造、収録コード体系、キーワード、検索コマンドの種類、用法等のデータベースそれ自体に関する知識と、データそのもの及び利用上の制約事項、分析ノウハウ等のデータ利用に関する知識が不可欠である。しかも、これら利用者に求められる知識は、益々複雑多様化しており、データベースの利用普及の隘路の一つとなりつつあるのが現状である。

このため本調査研究では、データベースに関する知識、データの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のデータベース化（知識ベース化）の検討を行い、利用者とデータベースの間のインターフェースの向上を図ることを目的に、昭和63年度より、3ヵ年計画で検討を進めている。

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、前年度の調査結果及び知識ベースの設計方針に基づき、貿易統計を例題として、必要となる知識を整理し、知識ベースを試作し、それら知識を利用する試作システムを開発し、知識表現の適否、知識管理方法の妥当性、動作の確認、データベース利用上の有効性等について評価、検討を行い、実用システムへの課題等を整理した。

平成2年度は、前年度までの結果に基づき、知識の追加等を行い、実際のデータベースを利用し、実用システムとしての運用可能性について評価、検討する。またこれらのシステムが他分野の統計へ拡張可能かどうかについても評価、検討を行う。

本報告書は、この3年計画の2年目にあ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、本調査研究は、株式会社CSK総合研究所に委託して実施した。

平成2年3月

「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	淵 一博	勸新世代コンピュータ技術開発機構
委員	鶴野 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	東京国際大学
委員	藤森 聿子	技術研究組合国際ファジィ工学研究所
委員	寺田 範雄	通商産業省 大臣官房 情報管理課
委員	小紫 正樹	通商産業省 大臣官房 情報管理課
委員	上原 明	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	勸CSK総合研究所

(専門委員会)

委員長	鶴野 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	野村 玲子	日本貿易振興会	
委員	篠原 昇	勸アイ・エヌ情報センター	
委員	水島 豊	勸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	勸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	太田 博親	通商産業省 産業政策局 調査課	
委員	杉原井 康男	通商産業省 貿易局 貿易保険課	
委員	上原 明	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	佐藤 安夫	通商産業省 通商産業研究所 政策情報システム部	

3.

目 次

第1章 IRSの目的

1. 1 調査研究の目的	3
1. 2 調査研究の背景	3
1.2.1 経済統計データベース利用の現状	3
1.2.2 IRSの構想と開発システムの位置づけ	5
1.2.3 本年度開発のアプローチ	5

第2章 データベース検索のためのモデル

2. 1 データベース検索支援の実体	9
2.1.1 ユーザの要求と検索支援	10
2.1.2 専門家による支援の分類	13
2. 2 概念的なモデル	20
2. 3 システム化のためのモデル	23
2.3.1 データフローモデル	23
2.3.2 データモデル	27
2.3.3 システム構成	28

第3章 IRSの知識ベース

3. 1 検索要求の表現	34
3.1.1 検索要求を表現するための属性	34
3.1.2 検索要求の属性決定に関する検討の経緯	35
3. 2 検索に必要な知識	44
3.2.1 検索要求に関する知識	44
3.2.2 類義語に関する知識	45
3.2.3 統計に関する知識及びデータベースに関する知識	47
3.2.4 画面制御及びレイアウトに関する知識	50
3.2.5 検索コマンド生成のための知識	52

3. 3 知識獲得の経緯	56
--------------	----

第4章 IRSのアルゴリズム

4. 1 データベース選択	63
4.1.1 選択の為の処理の流れ	63
4.1.2 必要な知識ベースとデータ	64
4.1.3 データベース選択のアルゴリズム	65
4.1.4 知識のメンテナンス及び拡張性	66
4. 2 導出に関するアルゴリズム	67
4.2.1 導出のための処理の流れ	67
4.2.2 必要な知識ベースとデータ	67
4.2.3 導出のアルゴリズム	68
4.2.4 知識のメンテナンス及び拡張性	69
4. 3 検索コマンド生成	70
4.3.1 検索コマンド生成部の構造	70
4.3.2 コマンド生成のアルゴリズム	70

第5章 IRSのマシニインターフェイス

5. 1 画面構成	75
5. 2 データ入力方式	79
5.2.1 操作手順	79
5.2.2 従来システムとの比較	80
5. 3 画面遷移と知識ベースとの関連	82

第6章 試作システムの評価と課題

6. 1 試作システムの概要	89
6. 2 試作システムに関する評価	92
6. 3 今後の検討課題	94
6.3.1 利用目的と必要な機能	94
6.3.2 拡張項目と技術的ポイント	96

第7章 ヨーロッパの統計DBユーザ・インタフェースの研究動向と

IRSの拡張性

7. 1	自然言語インタフェース	103
7.1.1	フィンランド中央統計局のTSEXSY	103
7.1.2	TSEXSYの評価	105
7. 2	グラフ型ユーザ・インタフェース	106
7.2.1	IASIのGRASS	106
7.2.2	GRASSインタフェースの評価	108
7. 3	対象中心アプローチのユーザ・インタフェース	109
7. 4	統計ユーザ・インタフェースと3スキーマ	112
7.4.1	ANSI/SPARCの3スキーマ	112
7.4.2	各種統計ユーザ・インタフェースの位置づけ	114
7. 5	対象中心アプローチとIRSの拡張性	116
7.5.1	LIDS86のユーザ・インタフェース	116
7.5.2	対象中心アプローチの汎用ユーザ・インタフェース	117
7.5.3	対象中心アプローチとIRSテンプレート	119
7.5.4	対象中心アプローチの外部スキーマとしてのテンプレート	120

(注. 第7章の執筆は佐藤英人委員が担当した)

付録

1. 知識ベースリスト
2. 通商白書図表タイトルに関する重要度アンケート集計結果
3. データベース検索Q&A
4. 検索要求フレームのサンプルケースへの適用
5. 貿易統計に関わりの深い経済統計用語
6. 利用目的別ユーザニーズアンケート

参考文献

第 1 章 I R S の目的

第1章 IRSの目的

本年度の調査研究及び開発システム IRS (Intelligent Retrieval System) の目標と位置づけを明確にしておく。

1. 1 調査研究の目的

データベースの普及に伴って、データベースを利用するユーザ層が多様化している状況のなかで、専門的な訓練を受けたユーザでなくても、検索が容易にできるようなデータベースの開発が望まれている。

この為には、検索方法のみではなく、収録されたデータの意味に関してもユーザに示唆を与えるシステムの開発が必要である。

本調査研究では、経済統計の分野に着目し、データベースに関する知識、データ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を行い、データベースの知的なインターフェースの研究を目的とする。

上記目的に基づいて昭和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、これらの知識ベース化の方法について設計を行った。

平成元年度は、前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った。

平成2年度は、さらに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予定である。

1. 2 調査研究の背景

1.2.1 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現状について触れ、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

経済統計データベース利用の現状を整理すると下記のようになる。

データベース利用の現状

- ・データベース及び経済統計に関して、経済統計の専門家以外の人が利用するケースが増えている。（例、貿易統計データベースについては、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、外部情報と内部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、一般化と高度化の2つの方向に向かっていくことがわかる。

次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。

データベース利用上の問題点

- ・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、検索が難しく、理解しにくい。
- ・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統合等高度な加工がしにくい。

②AI技術の適用性について

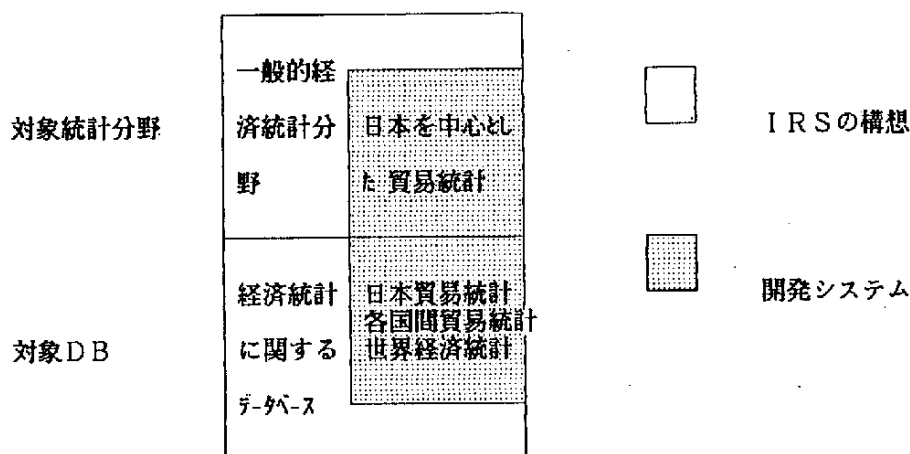
上記のようなデータベース利用上の問題点を解決する方法の一つとして、AI技術の適用可能な分野をまとめると下記ようになる。

AIの適用可能な分野

- ・ユーザの要求に対する適切なアドバイス機能について
- ・経済統計に関する専門的知識の知識ベース化による検索システムの高度化について
- ・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ースの構築について

1.2.2 IRSの構想と開発システムの位置づけ

本年度の開発は、日本を中心とした貿易統計を対象とし、データベースも日本貿易統計、各国間貿易統計、世界経済統計の3種とした。IRSの将来的な構想としては、一般的経済統計を対象とする。したがって、本年度の開発にあたっては、将来の拡張性を考慮して知識構造や、データ構造、及びアルゴリズムを検討することとした。



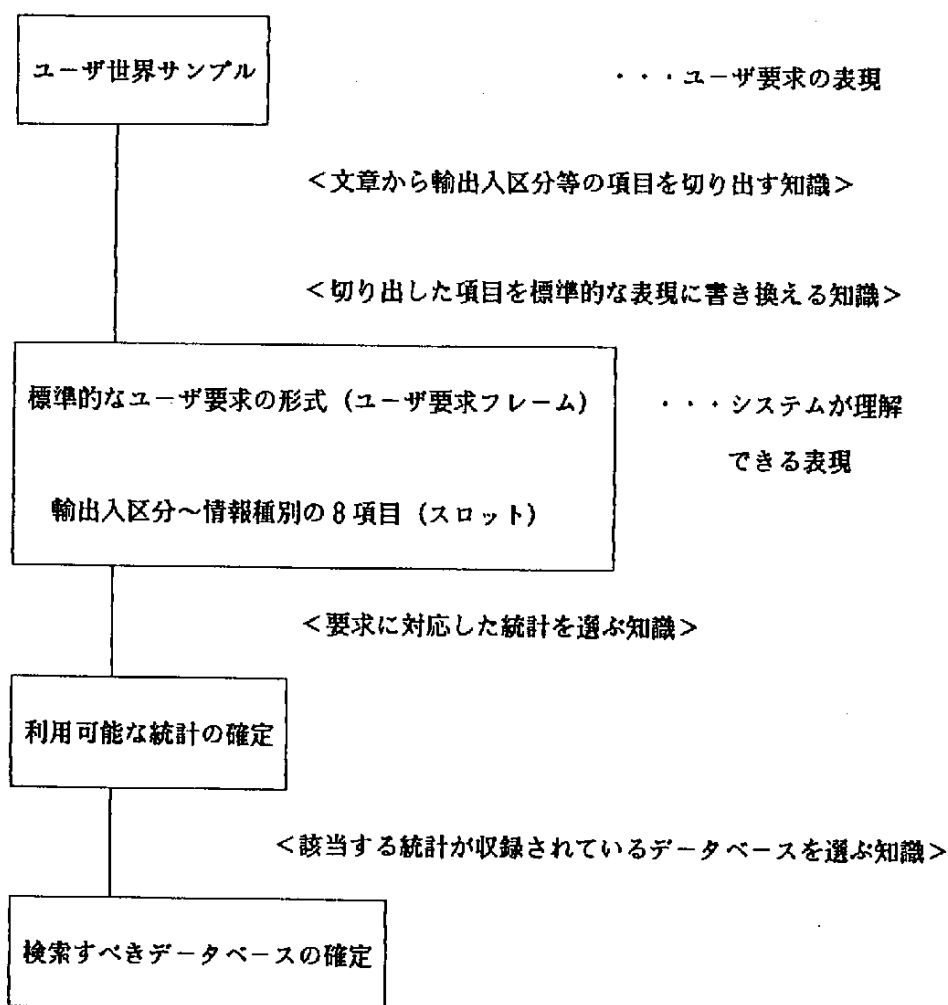
1.2.3 本年度開発のアプローチ

本年度は、検索に必要な知識ベースの構築と、検索のためのユーザインタフェースに重点をおいた開発を行った。

特に、検索に必要な知識を分析し、知識の構造を明確にし、知識獲得の方法を明らかにすることが、重要な課題であった。

対象ユーザは、現在明確に設定されていないが、昨年度収集したユーザ世界の要求サンプル100ケースのレベルを当面の対象ユーザと想定した。

ユーザ世界の要求サンプルは、日常語の文章で表現されているが、システム的に取り扱うために、これを標準的な形式に書き換え、書き換えられた検索要求をもとに統計や、データベースを確定してゆくのが基本的な検索手順になる。このように、必要な知識は、ユーザ要求を標準形式に書き換えるための知識と標準形式のユーザ要求から統計・DBの確定を行うための知識に分けられると思われる。



上図の知識に関して実際に知識獲得を行いサンプル知識ベースを作成し、実現可能性を確認したうえで、実用的な知識ベースの開発を進める。

第2章 データベース検索のためのモデル

第2章 データベース検索支援のためのモデル

昨年度は、ユーザの世界、統計の世界、データベースの世界の3世界からなるデータベース検索の概念的なモデルを作り、これをもとに、実際のシステムのコンセプトを検討してきた。本年度は、試作システムの開発の過程を通じて、引き続きデータベース検索支援の要件を抽出した。

本章では、昨年から検討してきたIRSによる検索支援を、各検討レベルでのモデルとしてまとめ、その特徴を明らかにしておく。

検討の経緯としては、以下に示す①から⑦の段階を経て各段階でのモデルを作成した。

- ① 専門家による検索支援の実体を調査し、ユーザと専門家との間でどのような情報がやりとりされているかを整理する
- ② ①で整理した内容を一般化し、図式化する …… 実体モデル
- ③ 昨年度作成した概念的なモデルについて、①、②の実体と照らして、正当性を評価する
- ④ 専門家とユーザのやりとりを媒介する「検索要求」の表現形式を決定する
…………… データモデル
- ⑤ システムとして、一件の検索が終了するまでのデータ（検索要求）の変化をモデル化する
…………… データフローモデル
- ⑥ システムとしての処理の単位、データの単位を及びこれらの階層構造を考慮して、オブジェクト指向を採用したシステム構成を決定する …… システム構成
- ⑦ ⑥で、知識ベースとして取り扱った、専門家の知識に相当するデータの構造をモデルとしてまとめる
…………… 知識構成

システムの妥当性や問題点を検討する際に、どのレベルでの問題として取り扱うべきかを明確にしておくことが重要である。

2.1 データベース検索支援の実体

始めにユーザが、実際にどのような情報を求めてデータベースを検索しているか、さらに専門家からどのような支援を受けているかを分析する。

2.1.1 ユーザの要求と検索支援

(1) ユーザの要求

昨年度の調査研究において、「通産ジャーナル」の記事より貿易統計に関する100数件の検索要求文を作成した。これらは、必ずしもデータベースの検索と直接結びついている訳ではないが、データベースを思い浮かべる前に、ユーザが欲しいと思う情報の原型と考えられる。

本調査研究の対象としては、統計やデータベースに関し深い知識のないユーザの、自然な表現（検索コマンド的な表現ではなく、概念的な表現や抽象的な表現も許した表現）を考えたので、ユーザが要求する「検索」は、この100数件によって代表されるものと想定し、これをIRSによる支援の対象とする。

(2) 専門家による支援

昨年度は、上記100数件の検索要求文に対して、経済統計に関する専門家に意味、不足情報、統計の種類を記入したワークシートを作成してもらった。

ワークシートに記入された情報は、(1)に示したユーザの要求に対する、専門家のアドバイスの過程であると考えられる。即ち、ユーザが欲しいと思う情報に対して、①統計的な意味付けを行い、②統計を確定するために不足している情報を補ったうえで、③的確な統計をユーザに示す・・・というのが、一般的な支援の形態であろう。但し、ここでいう専門家とは、「統計」の専門家である。実際にユーザが欲しい数値を入手する為には、実在するデータベースを検索する必要がある。つまり、「統計」の専門家のほかに、さらに「データベース」の専門家が存在し、（「統計」の専門家と同一である場合も多いと思われるが・・・）、コンピュータ上のデータベースシステムの利用や、検索結果に関するアドバイスを行っている。

統計に関する専門家とデータベースに関する専門家は、実際のデータベースサービスの現場では、一人の担当者が兼任していることが多いが、知識としては、両者は、別々のものであると考えられる。

すなわち、経済統計に関する専門知識は、ユーザの用途を把握し、適切な統計を示すものであるのに対して、データベースに関する専門知識は、データベースとして収録された後の「統計」に関する知識及びデータベース使用方法に関する知識と考えられる。

「データベース」の専門家の支援については、本年度さらにデータベース検索時の問題点の抽出と、専門家からのヒアリングを行い、データベース検索Q&A集（付録3）を作

成した。

専門家へのヒアリングによって得られた検索と支援の実例を図2-1、図2-2、図2-3に示す。図2-1は統計の専門家による支援、図2-2は、データベースの専門家による支援の様子を示している。また図2-3は、検索結果が得られた後の、データの加工等に関する支援を示している。

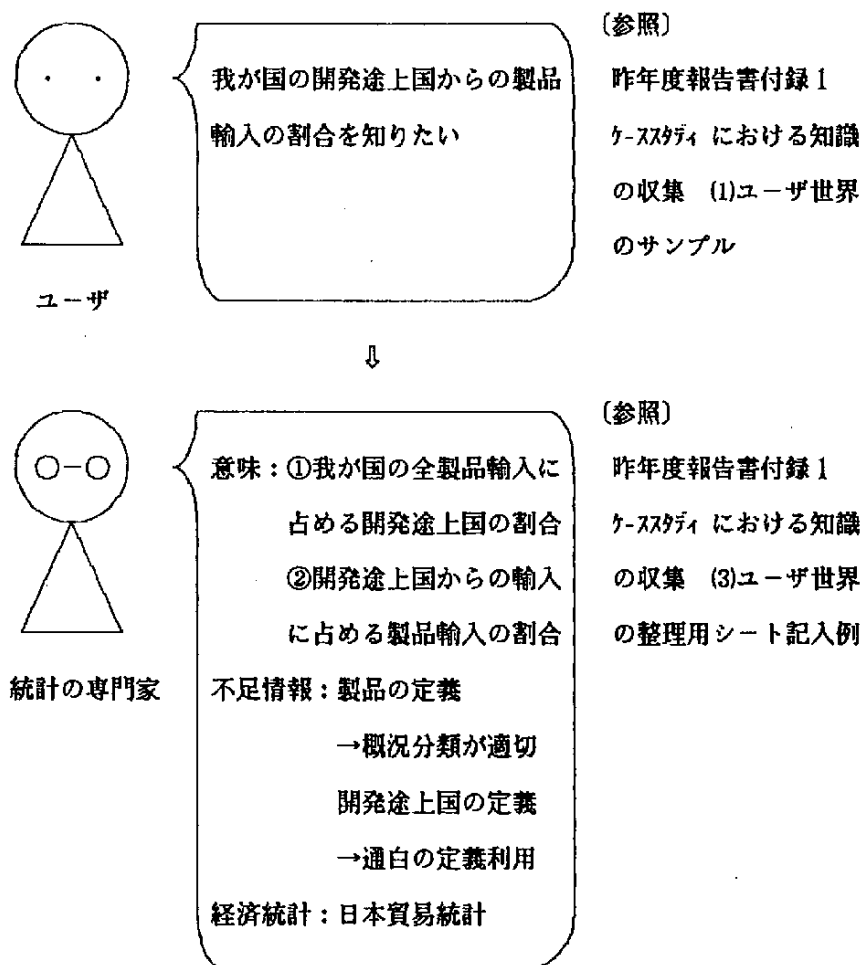


図2-1 検索支援の実例1 統計の専門家による支援

ユーザは図2-1のように、統計の専門家の支援をうけて、自分の欲しい情報を統計の用語で定義し、検索すべき統計の種類を決定し、データベースの検索を行う。

データベース検索に際しても、図2-2のように専門家の支援をうける。



ユーザ

製品の輸入額は、
 化学品（4291）、金属品（
 4292）、機械機器（4293）
 雑製品（418）の合計か

〔参照〕

昨年度報告書付録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

〔注〕（ ）内の数字

は分類コード



データベースの専門家

4312 が製品に対応しているが
 近いうちにコード変更の予定が
 あり、注意が必要。

〔参照〕

付録
 データベース検索Q&A

図2-2 検索支援の実例2 データベースの専門家による支援

このようにして、データベースの検索コマンドを作成して、データベースを検索するが
 経済統計データベースの場合は、検索に失敗するケースがかなりある。これは、

①それぞれの統計項目について、各国のデータがそろっていることが少ない。

②国によって同じ統計項目であっても集計方法が異なり、検索コードが違う。

等の理由によるものである。また、検索結果が得られた後にも、データの読み方に注意の
 必要なものがある。例えば金額や数量の単位が国によって違い、検索をしてからでない
 とどの単位で収録されているか判らない等の問題である。

このように、ユーザは検索後にその結果に関して専門家から支援を受け、必要に応じて
 検索行動を繰り返すと考えられる。また、この場合の専門家は、統計の専門家である時と、
 データベースの専門家である時の両方があり得る。すなわち、代替の統計を求めるものと、
 データベースの操作に関する内容を主とするものの、2種類である。

この場合の検索支援の実例を図2-3 に示す。

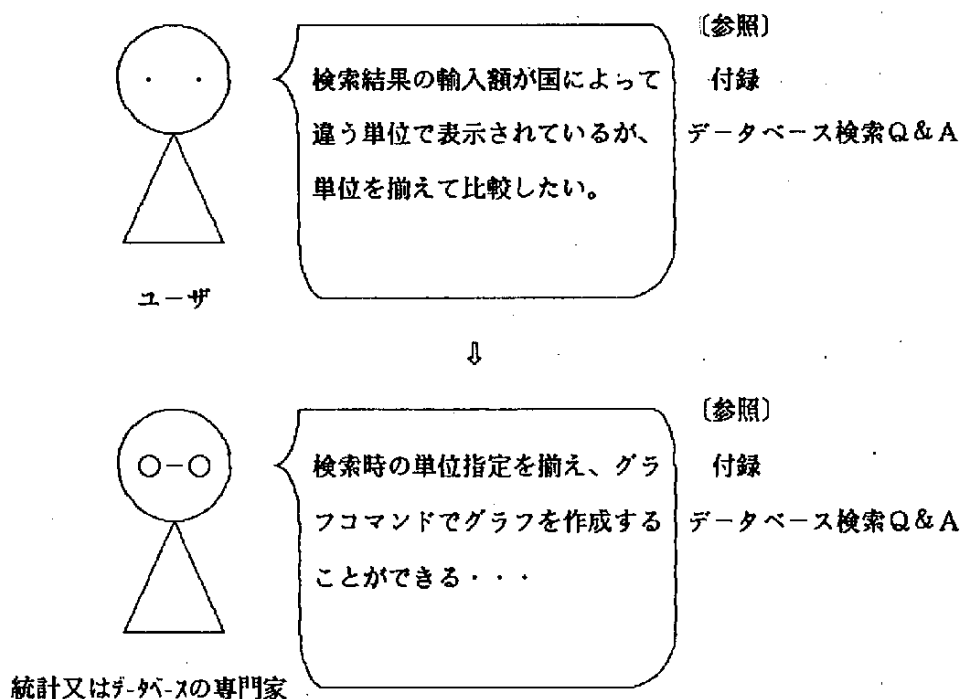


図2-3 検索支援の実例3 検索後の支援

2.1.2 専門家による支援の分類

本年度の調査研究では、主として、統計に関する専門家の知識を中心とした検討をおこなったが、専門家による支援の内容を分類し、必要な知識にどのような種類があるかを明らかにした。

(1) 統計の専門家による支援

2.1.1 で説明した統計の専門家による支援について、100数件のユーザ要求のなかにどのような種類の「検索」が含まれており、支援のしかたとして、どのようなものがあるかを整理し、下記に分類する。

支援のしかた・・・という観点からみると次の2種類に分類ができる。

① ユーザの用途を把握し、統計の概念に置き換える

② 存在する統計の中から適切なものを選択する

それぞれの内容としては、下記のようなものが含まれる。

① ユーザの用途を把握し、統計の概念に置き換えるための支援として・・・

・曖昧さを含む表現を除き統計の用語に置き換える（求める情報を得るために必要

な加工の方法を示すことも含む)。

- ・用語を統計の分類体系に合わせて置き換える(統計の分類体系として定義されていない場合は存在する分類からの類推方法を示すこともある)。
- ・不足する情報を補足する(補足できないときは、ユーザに問い合わせる)。
- ・データベース検索の範囲を超える要求に対しては、ユーザの要求に近い情報提供の方策を提案する。

②適切な統計を選択するための支援として・・・

- ・複数の統計が該当するとき収録状況を考慮し、最適のものを選択する。
- ・指定の統計がない時に次善の情報を提供する。
- ・選択した統計にデータの性質や速報性のうえて注意すべき点があれば指摘する。

具体的なユーザ要求に対して、上記の分類を適用してワークシートに記入された専門家の指摘を整理してみると、下表のようなサンプルが得られた。

表 2-1 ユーザ要求に対する専門家の支援

ユーザ要求	専門家の支援
ドル安で輸出はどの程度伸びるか。	報告国の推定(不足情報の補足) ドル安で輸出が伸びる・・・米国 曖昧な表現の置き換え ドル安→1985年秋以来の円高ドル安の現象 輸出 → 輸出額 どの程度伸びるか→変化の予測 要求全体⇒①ドル安の現象がいつまで続くか、どの程度の大きさで推移するかを予測する。 ②ドル安に伴って、輸出額がどの程度変化するかを予測する。 ユーザの要求に近い情報提供 予測は難しい(DB検索の範囲外) → エコノミストの予測が収録されていればそれを提供する → 為替レートの推移と、輸出額の推移をドル安の現象が出現する前の5～6年を含む現在までの時系列で示す。 的確な情報の選択 ・アメリカ側の貿易収支統計で月次データの収録されたものが望ましい。 ・次善の情報として、日本の通関統計もつかえる。

	・研究機関等の予測データがあれば予測値として参照可能である。
我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。	用語を統計の分類体系に合わせて置き換える ・「スーパーコンピュータ関連部品」には正式な定義はないし、慣用的な利用例もないので、範囲をより一般的に大きくとらえざるをえない。「半導体素子」等によってとらえることが可能か？ 曖昧な表現の置き換え（不足情報補足） 数量→kgでよいか、価額が必要か、 価額の場合は、円建てかドルベースか 期間→最新値でよいか？ 時系列的に見たいか？ 的確な統計の選択&注意点の指摘 ・通関統計でアメリカへの輸出を調べる→アメリカでの生産に関する統計も必要 ・スーパーコンピュータに関しては、メーカーの部品調達に関する統計を得る方法があれば類推可能
日本の石油輸入量の1974年以降の増減を知りたい。	曖昧な表現の置き換え（不足情報補足） 増減→1974年以降の推移（時系列） でよいか、1974年を基準とした指数表示がよいか？ また「増減比」の時系列も一般的に使われている。 注. 増減比 = $\frac{\text{石油の当年の輸入額}}{\text{石油の前年の輸入額}} \times 100$ 石油→通白分類の石油でよいか、 （石油製品を含むか） 石油の品目分類は1976年に変更されているので注意が必要 輸入量→数量（kt）か、価額（円建て、ドル建て）か？ 的確な統計の選択 ・通関統計 ・石油統計速報（確報ではない） ・石油資料月報（上記2統計より作成）
日本の輸出産業に占める自動車産業の割合を知りたい	曖昧な表現の置き換え（不足情報補足） 要求全体 → 捉え方がいくつかある ① 自動車の輸出額で捉えて、全輸出額との比をとる ② 自動車関連企業の生産額で捉えて、輸出産業全体の生産額との比をとる ③ 輸出産業の計上利益のランクのなかで、自動車関連産業の地位を調べる。 輸出産業 → どの定義を使うか？ 自動車産業→自動の部品メーカー等を含むか？

	期間 → 最新の情報でよいか 的確な統計の選択 ・ 上記①の場合・・・通関統計 ・ 上記②の場合・・・生産動態統計、自動車統計月報、機械統計月報 ・ 上記③の場合・・・法人企業統計、財務データファイル
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。	的確な統計の選択 ・ 輸入価格は、通関ベースでは、外国貿易概況により調べられる。物価ベースでは、輸入物価指数、卸売物価指数、消費者物価指数が使える。
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。	曖昧な表現の置き換え（不足情報補足） 耐久消費財→外国貿易概況の定義（家庭用品、家電機器、乗用車、二輪車類、玩具楽器類）でよいか 西ヨーロッパ→通関分類の定義でよいか 時点→最新の情報のみでよいか、対前年比較、国別シェア等は必要ないか 的確な統計の選択 ・ 国別品目別は、貿易月報から判るが個別品目を集計する必要がある。
日本国内で売られているカメラは、どのくらいが輸入品か知りたい。	曖昧な表現の置き換え（不足情報補足） 時点→最新の情報でよいか 国内販売→国内で販売されているカメラの数量は下式で推定可能である。 $\text{在庫増減} + \text{生産} + \text{輸入} - \text{輸出}$ カメラ→どの範囲をカメラとするか 的確な統計の選択 ・ 輸入量・・・外国貿易概況 ・ 生産量・・・生産動態統計
発展途上国の対外債務残高の推移について知りたい。	曖昧な表現の置き換え（不足情報補足） 発展途上国→世界銀行の定義でよいか 時点→最近10年程度でよいか 的確な統計の選択 ・ 世界銀行のWorld Data Table ・ IMFのInternational Financial Statistics
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。	曖昧な表現の置き換え（不足情報補足） 要求全体→発展途上国、共産国、累積債務国と3種類のグループについてと読めるが、累積債務を問題にするのなら、発展途上国、共産国諸国、各々のなかの累積債務国のみについて調べればよいか。 貿易関係→①輸出入額、②輸出入額と構成品目の傾向が考えられ

る
発展途上国、共産国、累積債務国→
それぞれの定義は？
時点→最近5～6年の推移でよいか
的確な統計の選択
・通関統計外国買概況
・Comoditty Trade Statistics

(2) データベースの専門家による支援

一つのデータベースには、一般に複数の統計が収録されている。また収録された統計は、原統計そのものではない。特定の時点を選んで、収録されていたり、数値に季節調整を加える等の操作が施されている場合が多い。

データベースになった後の情報に関して責任を持つのがデータベースの専門家の役割として、統計の専門家との区別をした。

原統計とデータベースに収録されたデータの差異を知って、ユーザにアドバイスを与えること、また実際にデータを取り出すための手続きに関してユーザの手助けを行うことがデータベースの専門家の支援であると考えられる。

以上を整理し、実際のサンプルにあたってみるとデータベースの専門家による支援は、次のように分類できた。

・各データベースに情報が収録されているか否かの判断

時点、系列、期種、国名、品目等についての、個々のデータの収録範囲、及び、データの更新時期等の管理情報が、主な知識となる。

・検索コマンドの生成

検索に必要な情報が揃った後に、検索コマンドを作成するための支援で、必要な知識は、データベースのマニュアルの内容に相当する。

例) スイスのGNPを知りたい。

→ retr tsc eq '99a..' and (cco eq 146)

・各データベースの分類コードに関する支援

コード表に関する知識を使った支援で、ユーザの要求する品目や、国名、系列が当該データベースのどのコードに対応するかを判断するものである。

(3) 検索後の支援

検索後に発生する問題としては、データの欠測により検索に失敗する、検索結果が希望と違う単位で収録されている等があり、支援内容としては、下記のような項目の分類ができた。

・検索に失敗した場合の対策

代替コードの指示、代替系列の指示等。

・単位が指定と異なる場合の対策

特に通貨単位に関しては、変換の際に何時の時点での為替レートによって変換を行うのが適切か、また為替レートは、どのデータベースを使用すればよいか等の支援に関して需要が大きい。

- ・ 検索結果の加工に関する支援

検索結果を見て、グラフや表等の形式に加工したい場合の方法（ダウンロード後別のソフトを利用してグラフ化する等）。

- ・ 結果の分析に関するガイド

特に複数の系列の比較可能性等、データの取り扱い方に関する支援。

2. 2 概念的なモデル

概念モデルとしては、昨年作成したモデル（図2-4）を基本として、昨年収集した検索要求に関するケーススタディを中心に議論をかさね、このモデルの各世界に相当する知識の存在を確認した。

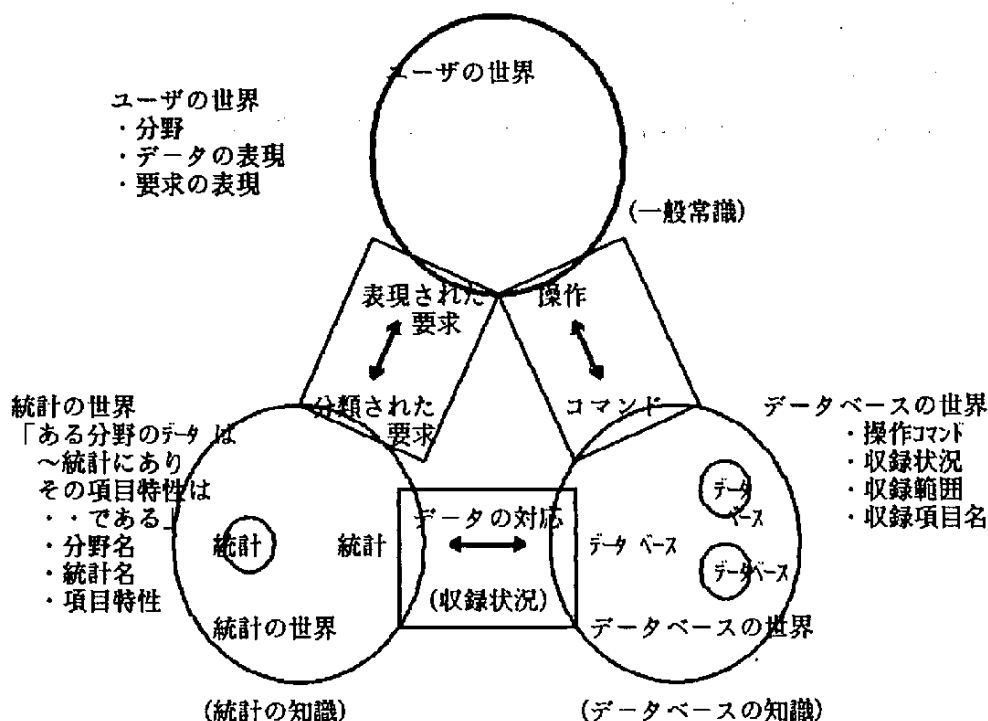


図2-4 経済統計DB検索モデル

上記概念モデルは、昨年度に提案され、本年度も継続して本調査研究の前提としているものであるが、本年度の調査により検索の実体が具体的になったので、もう一度その正当性を考察してみる。

本モデルの特徴は、必要な情報の検索をするためには、データベース検索に際して、データベースそのものに関する知識のほかに、データの使い方に係る、統計（今の場合貿易統計）に関する知識が必要である・・・という概念を示している点である。

即ち、貿易統計データベースのように、

- ① データとして蓄えられた情報以外に、データの収録条件や、補足情報等収録時に欠落した情報があり、検索したデータの使い方に注意が必要である。
- ② 同じ種類のデータが複数のデータベースに収録されている可能性があり、どのデ

ータが適切であるか判断を要する。

- ③ 欲しいデータの属性区分が収録されたデータとマッチしない等の理由により検索に失敗する場合が多く、代替の情報が求められる。

といった特徴があるデータベースの利用に際しては、単なる検索コマンドのジェネレータではなく、データの利用に関するアドバイスが必要である。

検索コマンドのジェネレータとしての知識をデータベースの世界の知識とすると、データの利用に関する知識を、統計の世界の知識と呼ぶことができる。

昨年度の調査及び、本年度の調査により、それぞれの世界に対応する知識及び、それぞれの世界間での情報のやりとりが、具体的にされてきた。

まず、ユーザの世界のなかにある情報は、ユーザの欲しいと思う情報をユーザの表現方法によって表現したものであり、通産ジャーナルからのサンプルがこれにあたる。「我が国の米国スーパーコンピュータ関連部品の輸入量を見たい」等の情報である。

注意する必要があるのは、対象とするユーザが、特定のデータベースのユーザではない点である。即ち、自分の使えるデータベースの範囲を知っていて、データベースに合わせて検索要求が出されるのではなく、データベースを意識せずに検索要求が発生していることを前提としている。

ユーザの世界から統計の世界への情報のやりとりが、2.1.1 に説明した統計の専門家による支援に相当する。

ここで、統計の専門家の役割は、ユーザの曖昧な表現を統計の用語として、具体的に確定し、統計の分類体系に合わせて、ユーザの欲しい情報を分類し、必要に応じて、統計として存在する情報へ、ユーザの表現を置き換える・・・というようなものである。

このばあい統計の専門家の働きは、経済統計の一般的な常識に基づいたものであり、データベースに関する知識とは、区別して捉えることが可能である。

またデータベースに関する知識は、データベースの収録に関する情報とデータベースの検索コマンドの生成に関する情報が主なものと考えられる。

上記概念モデルは、このように、統計データベースの検索の実体にマッチしているといえるが、さらに、この概念モデルのなかで、各世界間の情報のやりとりがどのような形でおこなわれているかを検討した。

実際の検索では、統計の専門家とデータベースの専門家（人間の場合、両者の区別がはっきりしている場合は、まれで、多くは、統計の専門家が、データベースに関して熟知し

ていて、両方の役割を兼ねている。)とデータベースの専門家のやりとりが、頻繁に行われ、ユーザの要求に対して、一旦統計用語に置き換えた後、データベースの収録状況をみて、的確なものが存在しなければ、次善の統計を示唆する。また検索結果をみたうえで、次の検索の指示をする・・・等の支援をおこなっていることが判った。

昨年度に検討した世界間の情報交換のイメージは、図2-5のような一方通行的なものであった。“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との考えに基づくものである。

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。”(昨年度報告書より引用)

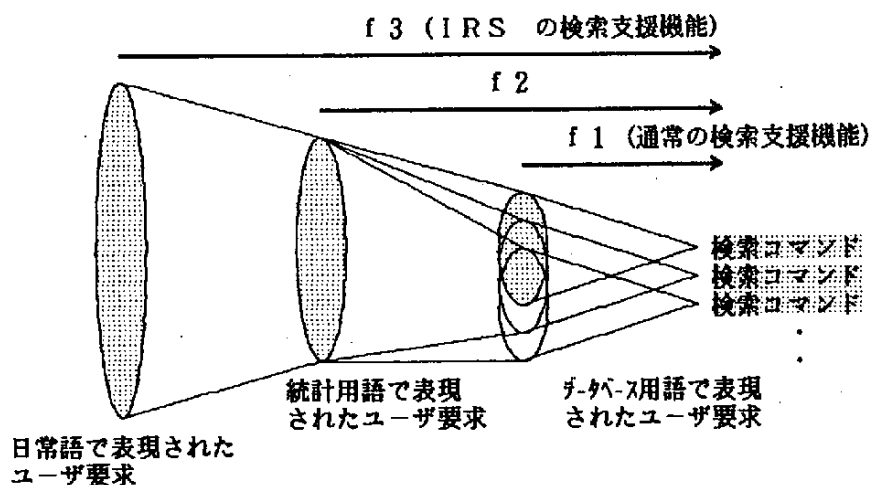


図2-5 検索支援の概念図

これに対して本年度の調査では、

ユーザ世界 —————→ 統計の世界 —————→ データベースの世界

の一方通行型ではなくユーザと統計、データベースの間でダイナミックに情報のやりとりを行うモデルがより適しているとの結論を得た。

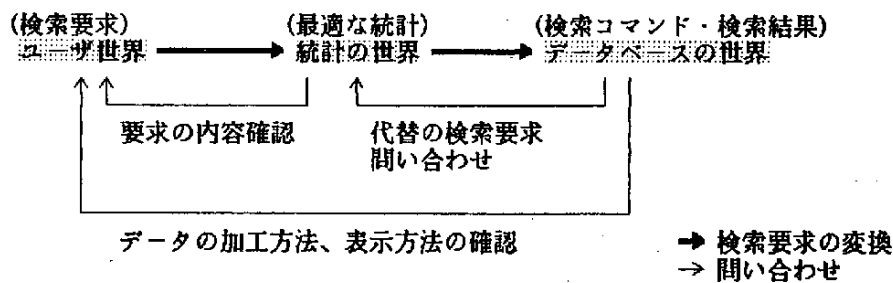


図2-6 検索支援の概念図2

2.3 システム化のためのモデル

2.2の概念的なモデルを実際にコンピュータ上のシステムにするために、システム化のためのモデルを検討した。処理とデータの流を捉えるためのデータフローモデル、検索要求を表現するためのデータモデル、及びシステム構成、知識構成である。

2.3.1 データフローモデル

ユーザの検索要求がデータベース検索コマンドに変換されてゆく過程を、これに係るデータ（知識）の作用に着目して図示すると図2-7のようになる。

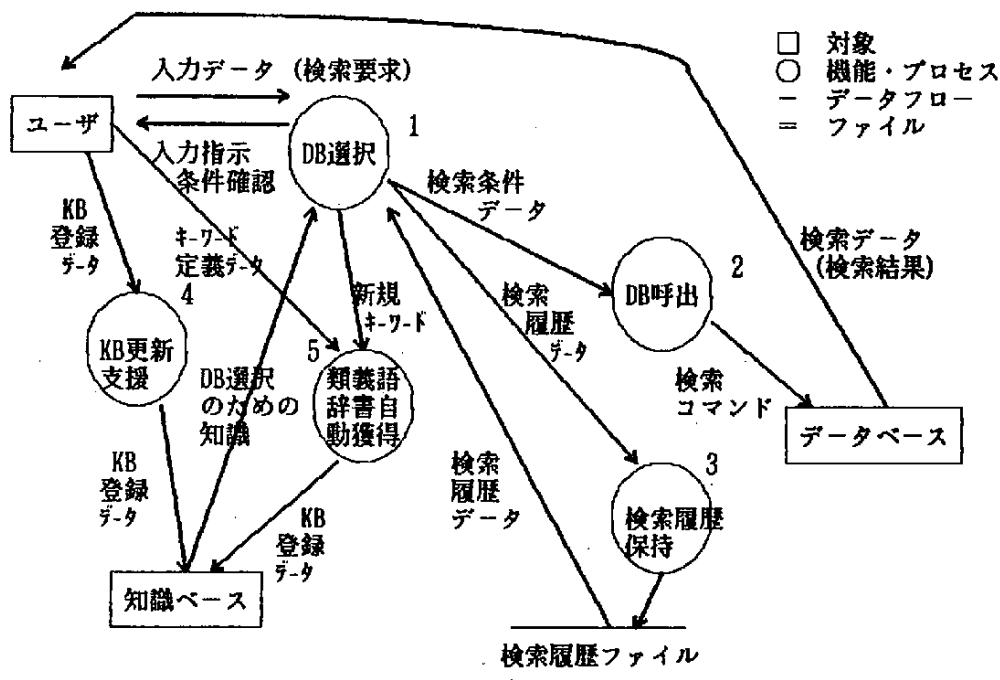


図2-7 IRSデータフロー図

図2-7で基本的なデータの流れは、

ユーザ → DB選択(1) → DB呼出し(2) → データベース → ユーザ

の流れである。

これは、ユーザの検索要求を「DB選択」によって統計の用語に置き換え、適切な統計及びデータベースを選択し、DB呼出しによってデータベース検索コマンドに変換して、データベースを呼出し、検索結果をユーザに示す処理で、本年度の試作システムで実現した機能は、この部分である。

検索履歴保持(3) は、検索の履歴（検索結果を得るためのコマンド列）を保存し、次回に同じ検索をするときに、手続きを簡略化するための機能である。

KB更新支援(4) は、知識ベースの更新作業を支援するための機能で、知識の整合性のチェック等を含む知識ベースの管理を行うものである。

類義語辞書自動獲得(5)は、ユーザが新たな（システムの知らない）用語を使用した時にそれを自動的に知識ベースに組み込むための機能である。

これらの機能の内、DB選択機能をさらに詳細なデータフローにしたものが、図2-8である。

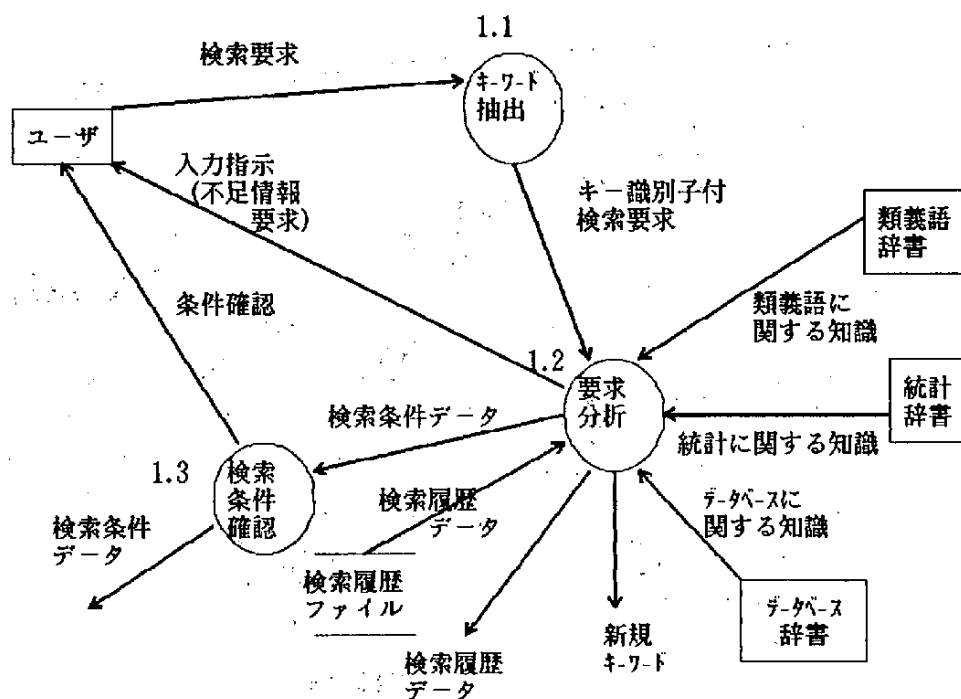


図2-8 DB選択データフロー図

検索条件確定(1.3.2) は、標準化されたキーワードによって表現された検索要求から、統計に関する知識を使って統計の概念としての検索条件を確定し(1.3.3) 、データベースに関する知識を使ってデータベース検索の際の検索条件確定(1.3.5) をおこなう機能である。

検索条件確定の際に、確定の為に不足している情報があれば、不足情報補足(1.3.6) によって補足する。またユーザの指定する統計やデータベースが存在しない場合は、次善情報選択(1.3.4) によって次善の情報を選択する。

次善情報の選択については、本年度は、具体的な支援の方策が固まらず、試作システムの機能として実現されていないが、検索支援のうえでは重要な機能であり、今後の課題とされている。

2.3.2 データモデル

システムの各モジュールでやりとりされる検索要求を図2-10のように表現した。
データモデルができるまでの経緯は、3.1で詳しく説明する。

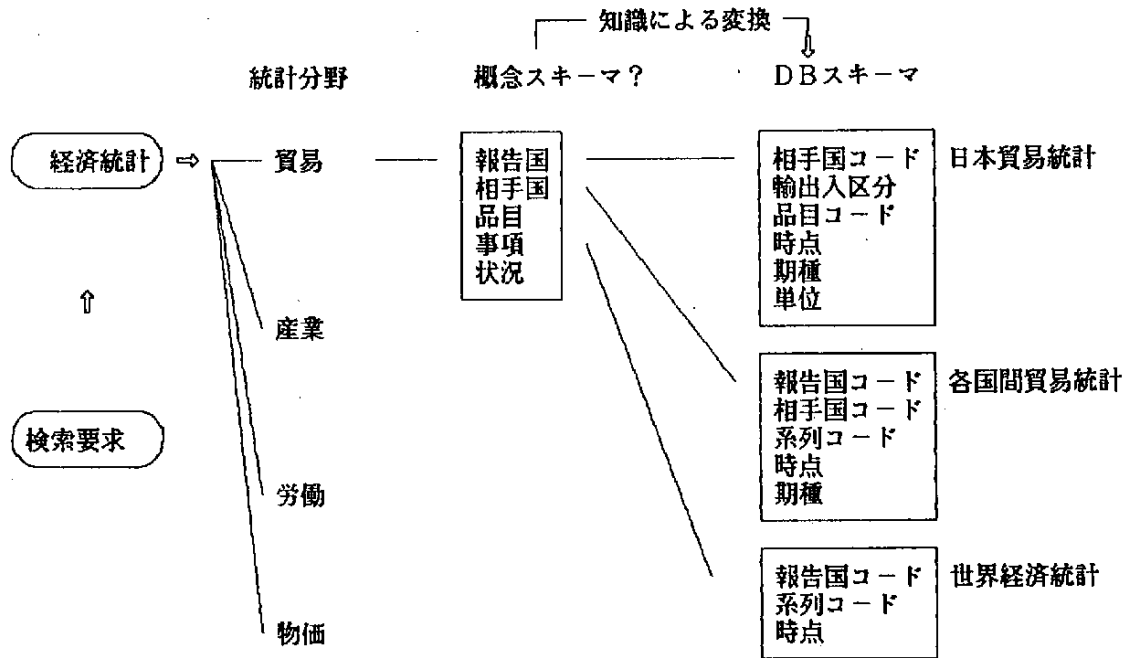


図2-10 データモデル

図2-10で概念スキーマと表現された項目（報告国、相手国、品目、事項、状態）はユーザの世界の検索要求を表現するためのキーワード群で、フレームと呼ばれる知識表現である。概念スキーマは、検索要求に対するユーザからの表現と言える。この表現を決定するためには、ユーザ検索要求という無制限の広がりを持つ対象をとらえる必要があり、本年度の調査研究でも、長期間の検討や、調査を要した。

この概念スキーマがユーザの検索要求として入力され、統計の世界、データベースの世界間の情報の媒体となり、最終的にDBスキーマと呼ばれるキーワード群に変換されて、検索コマンドが生成される。

DBスキーマは、データベースごとに定義されるキーワード群で、データベースの構造に依存してキーワードが定められる。言い方をかえれば、検索コマンドを生成するために必要なキーワードの集合として、定義されたものであるといえる。即ち、DBスキーマは概念スキーマと異なり、データベースによって一意的に決められるものである。

試作システムでは、検索要求＝貿易統計の概念スキーマと考え、概念スキーマの下にユ

ーザ要求（スキーマ）を特定化させる過程で、データベースの絞りこみを行い、概念スキーマとDBスキーマを同時に形成する仕組みを実現した。

より一般的なユーザの要求を考えると、貿易統計の他に、産業、労働等の貿易分野があり、それぞれに異なる概念スキーマが存在すると考えられる。

またユーザの要求は、必ずしも経済統計を意識して発生するものではないので、実際のユーザ要求から、各分野の概念スキーマを抽出するためには、さらに、変換の処理が必要である。

2.3.3 システム構成

試作システムの開発は、オブジェクト指向という概念をとり入れた開発ツールによっておこなわれた。

オブジェクト指向とは、処理の単位をオブジェクトと呼ばれる、階層構造を持ったモジュールによって表現し、オブジェクト間の通信（メッセージパッシングと呼ばれる）によってシステムの制御を記述する考え方である。

各オブジェクトは、そのオブジェクトに特有の変数と手続きを定義することができるようになっている。

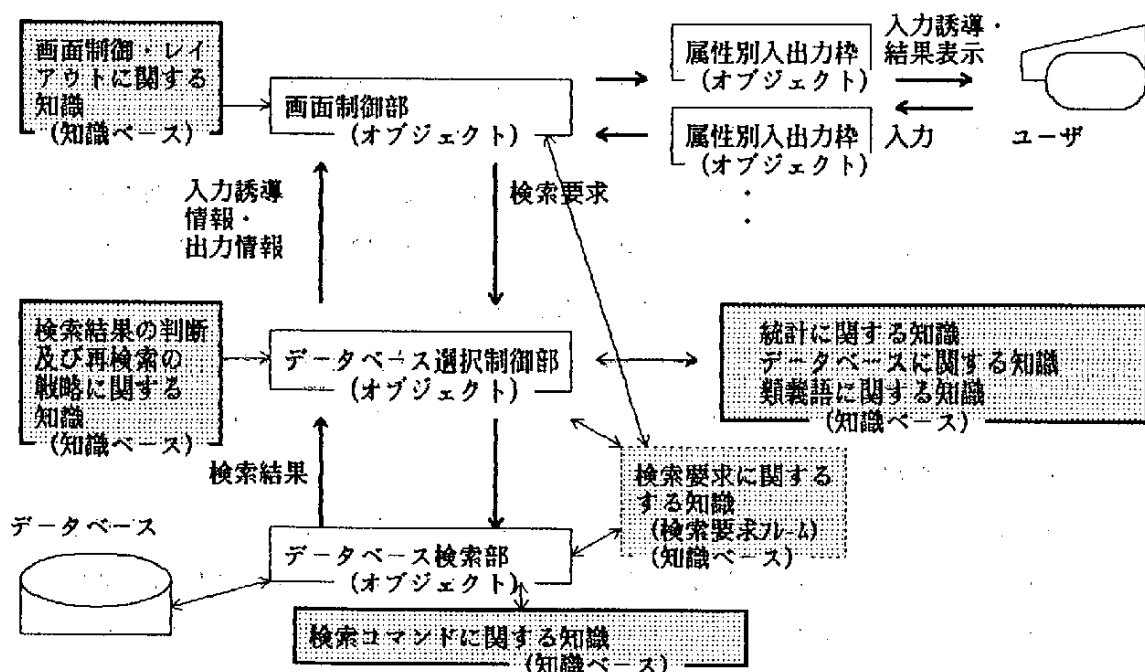


図2-11 システム構成図

図2-11は、試作システムのシステム構成図である。

試作システムは、画面制御部、データベース選択制御部、データベース検索部の各オブジェクト及び画面制御部の下位オブジェクトである属性別入力枠から構成されている。

データベース制御選択部は、試作システムの全体的な制御を行うオブジェクトである。データベース選択制御部では、概念スキーマに相当する各項目を変数として持ち（検索要求フレーム）、画面制御部を通してユーザの検索要求を取り込み、統計に関する知識、データベースに関する知識、類義語に関する知識を使って、検索条件の確定をおこない、データベース検索部にDBスキーマに変換された検索要求を引き渡す。

データベース検索部では、DBスキーマに相当する各項目を変数として持ち、各データベースの検索コマンドに関する知識を使って、検索コマンドを生成する。

画面制御部では、画面に表示する入力枠毎の下位オブジェクトを持ち、これらを制御しながらユーザ入力を誘導する。

試作システムの知識ベースは、検索要求に関する知識、画面制御及びレイアウトに関する知識、統計に関する知識、データベースに関する知識、類義語に関する知識、検索コマンドに関する知識、検索結果の判断及び再検索の戦略に関する知識の7種類である。

検索結果の判断及び再検索の戦略に関する知識は、試作システムでは、まだ実現していない。

これらの知識ベースについては、第3章で詳しく説明する。

第3章 IRSの知識ベース

第3章 IRSの知識ベース

検索支援システムとしてのIRSの特徴は、知識ベースに収録された専門家の知識を使って、支援機能を実現している点である。

知識ベースは、コンサルティングに必要な情報を、データとしてシステムに蓄え、必要に応じてこれを利用するための仕組みで、IRSの知識ベースは、その内容に応じて、①検索要求に関する知識、②類義語に関する知識、③統計に関する知識、④データベースに関する知識、⑤画面制御・レイアウトに関する知識、⑥検索コマンドに関する知識、⑦検索結果に関する判断及び再検索の戦略に関する知識の7種類に分類できる。

本章では、これらの知識ベースの構造を明らかにし、“知識を蓄える”……という観点から、IRSの知識ベースを考察する。知識ベースがどのように働いて、支援機能を実現しているか……については、次章で説明する。

各知識ベースの関連を確認する為に、下記に2章で解説したシステム構成図を再び掲載する。

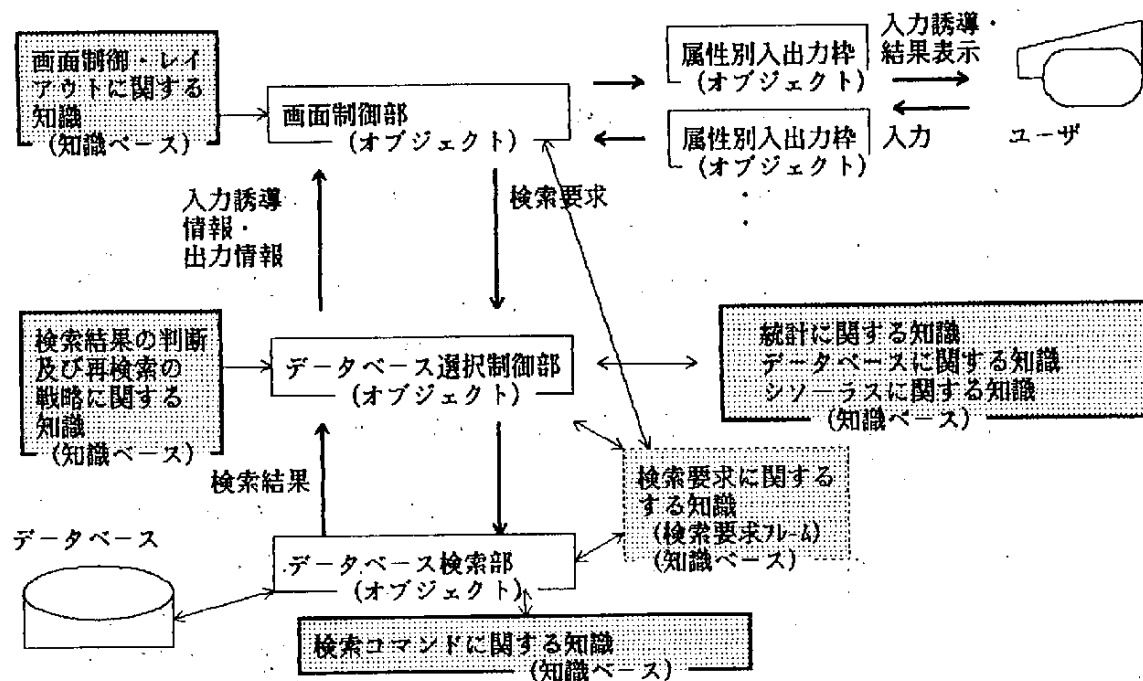


図 3-1 システム構成図

3. 1 検索要求の表現

ユーザの検索要求（このような情報が欲しい・・・という、データベースの検索コマンドを意識しない要求）もシステム内部では、一種の知識表現として蓄えられる。

但し、ユーザの検索要求は、他の知識ベースと異なり、専門的な知識として確定したものではなく、ユーザの入力に応じてシステム内部にたくわえられるもので、検索の処理の過程で参照され、処理に必要な形式に変換されるものである。

IRSでは、この検索要求をフレームと呼ばれる表現方法で、システム内部にたくわえることにした。フレームとは、表現したい実体を、有限個の属性と呼ばれる、キーワードによって表現する手法である。

属性は、属性の名称と属性の値によって定義される。

3.1.1 検索要求を表現するための属性

検索要求を表現するための属性は、最終的に下図の構成となった。これは、貿易統計に関する検索要求を表現するための検索要求であり、他の経済統計に関しては、別のフレームが存在する。

表 3-1 貿易統計に関する検索要求

属性名	意味	取りうる属性値
報告国	貿易統計の主体となる国名 (どの国からみた統計かを示す)	日本、アメリカ合衆国等の国名及び、ASEAN 諸国等の国群と、これら俗称のうちよく使われるもの（米国東独、ANIES 等）
相手国	報告国に対して、貿易の相手となる国名	報告国と同じ
事項	知りたい情報の種類	輸出額、GNP、為替レート等
状態	事項に関する表現形式	最新値、動向、推移等
品目	貿易でやりとりされる物の名称	石油、コンピュータ、電気製品等 基本的に概況分類に従って分類された品目とこれらの俗称

3.1.2 検索要求の属性決定に関する検討の経緯

前節に述べた検索要求が確定するまでに、検索要求を表現するための種々の試行錯誤をおこなった。

検討の過程で、検索要求を表現するうえでの問題点が明らかになったので、以下にまとめる。

(1) 最初に貿易統計に限定してフレームを想定し、ユーザの入力によって初期のフレームが構成された後、統計の世界、データベースの世界へ、このフレームが変換され検索が実行される・・・というシステムモデルを作成した。

また、それぞれの変換は、知識ベースにたくわえられた専門家の知識を使っておこなわれると考えた。

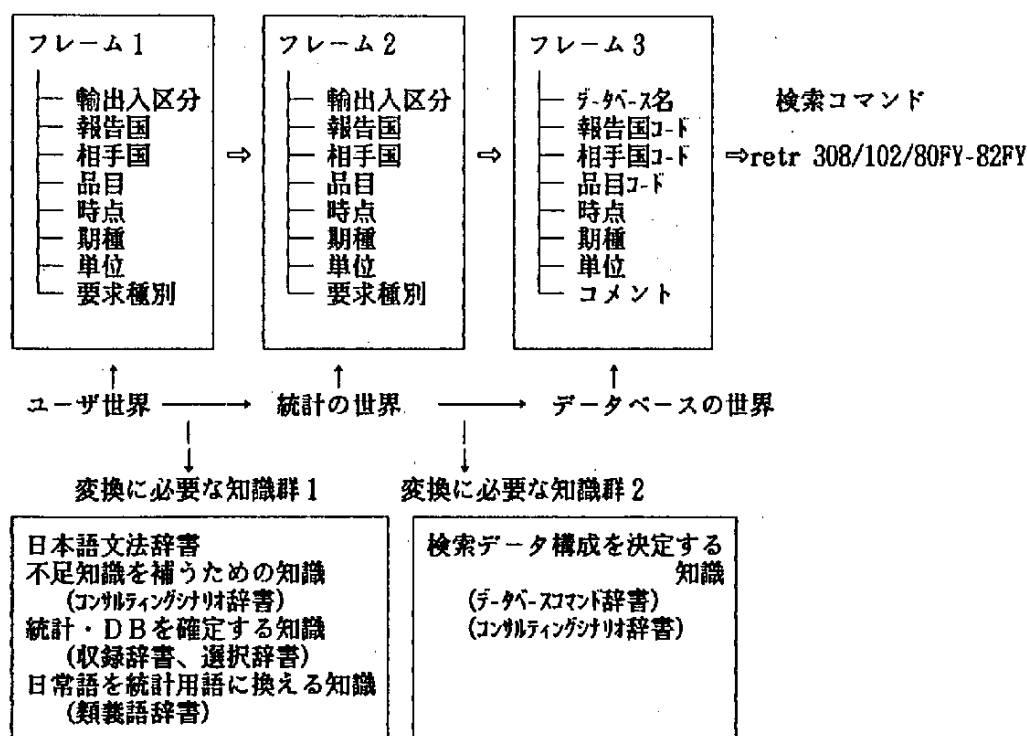


図 3-2 検索要求フレームとシステムとの関連

(2) 図 3-2 のフレーム 1 の形式に通産ジャーナルより抽出した 100 数件のユーザ要求を書き換える作業を通じて、このフレームによる表現の可能性を検討した。(付録 4)

検索要求を標準的な用語に置き換える際に必要な知識、問題点、知識構造、考えられる知識獲得の方法についてまとめたものが次ページの表である。

表 3-2 属性別検索要求フレーム検証結果

スロット名	必要情報、問題点、知識構造、知識獲得の方法
輸出入	☆必要情報及び問題点 輸出・輸入以外に貿易や貿易外収支などがある。 貿易は、輸出+輸入でよいのか。 貿易関連と言えないテーマの場合は、出入の区分がない場合がある。 例 我が国の外国製品の消費の状況を知りたい(14) 発展途上国の対外債務残高について知りたい(104) ⇒ 14は、国内需要額に対する輸入額の比と解釈可能。 ⇒ 104は、世界銀行'World Debt Table'等参照 対象統計の案内のみをおこなうか。 外国製品の消費を外国製品の輸入と置き換え、次善情報として提供するか。 ☆必要な知識の構造 出、入、出+入の3種類に対する類義語辞書と次善情報ルール。 ☆知識獲得 サンプル100ケースの範囲で類義語辞書を作成する。
報告国	☆必要情報及び問題点 デフォルトを日本にしてよいのか。 報告国と相手国を逆転するとデータが有る場合は、注釈つきで示す。 例 '東ドイツから日本への輸出'はないが '日本の東ドイツからの輸入'はある場合 報告国が日本以外の場合は、各国間貿易統計で総額は得られる。 報告国がグループの場合は、定義を類義語辞書に記述する必要がある (NICS、アジア等)。 複数の国の場合、国別データが必要か合計で良いか。 ☆必要な知識の構造 国名に対する同義語辞書(別称)。 例 大韓民国⇒韓国、アメリカ合衆国⇒合衆国、米国、USA、アメリカ グループ名定義リスト()の中が構成国を示す)。 例 アジアNICS=[香港、韓国、シンガポール、台湾] 国が複数の場合は、国別、総計の区分が追加される。ユーザの指定がない場合は、国別と総計の両方と解釈する。 ☆知識獲得 日本貿易統計、各国間貿易統計の収録国名から国名を収集し、それぞれに対して別名を対応させた同義語辞書を作成する(辞書、専門家へのインタビュー)。 サンプル100ケースの範囲でグループ名を抽出し、専門家が補足したものをもとにグループ名定義リストを作成する。
相手国	☆必要情報及び問題点 グループ名の場合の定義は、報告国同様に定義が必要。 デフォルトは、全世界でよいのか(ユーザに確認必要か)。 その他、報告国と同様。 ☆必要な知識の構造 報告国と同様。 ☆知識獲得 報告国と同様。
品目	☆必要情報及び問題点 製造業等業種別分類の定義が必要(通白分類コードで分類可能か)。 製品、原料等加工状態による分類の定義(通白分類、SITC)。 サービス等は、検索可能か。 分類体系は、通白分類等を基準にした階層構造でよいのか DBのコード変更に対応できるか ☆必要な知識の構造

(1) 品目類義語辞書

例 テレビ受像機→テレビ、テレビジョン

(2) 品目分類フレーム

業種別・産業別・加工別等のカテゴリに階層化する。
最下層の分類が、関税分類とリンクする。

例

製造業 原料 . . . ←抽象的カテゴリ

輸送機械 精密機械 . . . 木材 動物性油脂 . . . ←大分類

自動車 自動車部品 化学光学機器 . . . ←中分類

鉄道用及び軌道用以外の車両 . . . ←関税分類

☆知識獲得

通白分類、概況分類のいずれかをベースとする。

専門家へのインタビューをもとに必要な抽象的カテゴリを設定する。

選択した抽象カテゴリに属する分類に関して、類義語辞書を作成する
同様に選択した抽象カテゴリと連結する大分類、小分類に関して類義語辞書を作成する。

類義語辞書は、ワーキンググループ等を組織して作成する。

期種

☆必要情報及び問題点

期種に関してはユーザから指定のない場合がほとんどである。

期種は、基本的にDBの収録状況に依存し、指定のない場合は、最も細かい分類により表示することとしてよい。

年次暦年をデフォルトとするほうがよい。

時点で期種が決まる場合があるか。

☆必要な知識の構造

月次、四半期暦年、四半期年度、半期暦年、半期年度、年次暦年

年次年度の区分がある。それぞれの表現方法は複数あると思われるので
区分別類義語辞書が必要である。

☆知識の獲得

専門家へのインタビューまたは、ワーキンググループ

時点

☆必要情報及び問題点

ユーザからの具体的な指定のない場合が多い。'最近'ならば最新データ、要求種別が状況などならば、最新5年分という判断をシステム側でしてよい。

デフォルトを最新としてよい。

特に指定がなければ西暦としてよい。

☆必要な知識の構造

基本的に X~Y という形式で表現できる (X~Xなら単一年次指定)

上記指定がない場合は、最新、最近、過去 i 年等時点に変わる用語と数字に付随する形容詞の解釈によって時点を特定する。

用語	解釈
最新/最近 過去 i 年	DB収録データの最新のもの 最新から遡って i 年分

X～Yという指定がなくて、要求種別に時点の指定がある時（要求種別参照）は、要求種別からの指定を優先させる。

☆知識獲得
ワーキンググループによる協議

単位

☆必要情報及び問題点
ユーザからの指定がない場合がほとんどである。
指定がない場合や、指定単位と収録単位が合わないときは、DBに収録されている単位にあわせる。指定単位と収録単位が合わない場合は換算方法のメッセージを表示。
要求種別によって単位の指定をしたい場合があるか。

☆必要な知識の構成
金額の単位としては、円、ドル等。
数量単位としては、立法メートル、カラット、排水トン、ダース、グラム、グロス、総トン、キログラム、キロリットル、キロワットアワー、リットル、メートル、トン、数（個、本、枚、頭、匹、羽、台、両等）、平方メートル、組、千本、千枚等。
類義語辞書は不要か。
単位の換算方法。
要求種別に付随する単位の定義（要求種別参照）。

☆知識の獲得
要求種別に付随して単位が決定される場合は、要求種別ごとにワーキンググループの協議によって決定。

要求種別

☆必要情報及び問題点
影響、比較、割合等は、補足情報が必要。
（情報1、情報2のスロットを追加）
補足情報の要不要の区別を含め2段階の知識が必要か。

☆必要な知識の構造
(1) 要求種別類義語辞書
(2) 要求種別構成辞書

要求種別	情報1	情報2	単位	時点
影響	・・・の	・・・に対する	—	
金額	—	—	円	

要求種別の用語に関しては、下記の3タイプがあると思われるが、他に分類方法がないか。要検討。

タイプ1（単独で要求種別を表すもの）

数量、金額、貿易額 GNP 等

タイプ2（AのBという形式で2つの用語の組で表すもの）

額 の 変動 等

└ タイプ1に含まれる用語

タイプ3（AのA' に対するBのように2つの量の比較等を表すもの）

輸入額 と GNP の 関係

輸出 に対する 円高 の 影響

輸入額、GNP、輸出、円高は、タイプ1に含まれる

- (3) 要求種別と統計を結び付ける知識
タイプ1の範囲であれば、統計やDBを確定しやすいが、影響や関係は、ルールで蓄積しておく。

例

輸出関連の製造業の状況

⇒ 構造的な情報が求められる場合は、

- ①国民経済計算
- ②産業連関表

具体的な品目に関する情報及び速報性が求められる場合は

- ③生産動態統計調査
- ④週間統計

回答にあたっては、輸出の増加率、生産に占める輸出の比率の動向が必要

☆知識獲得

ワーキンググループによる協議

以上のようにサンプルケースを(1)のフレームを使って表現した場合の問題点が属性別にまとめられた。ここで最も大きい問題点は、属性「要求種別」の属性値の選び方である。

付録4に示すように要求種別のなかには、補足情報を必要とするものが多く、ケーススタディの過程で、当初考えていたフレームに「情報1」、「情報2」の2つの属性が追加された。

表3-2の要求種別欄で示したタイプ1からタイプ3を暫定的な要求種別の分類方法としたが、情報1、情報2を必要とするケースを見るとその内容が、一意的に表現されることは難しく、同じ検索要求が何通りもの表現に書き表される危険を含むことが判った。

(3) (2)の反省をもとに検索要求フレームを再度検討し、検索要求の内容を示す用語と検索要求の構造を示す用語を区別して捉える必要があるとの認識を持った。

以下に、新たな検索要求表現と検索要求の構造について説明する。

①貿易統計の検索要求表現（改良案）

従来検討されている属性（フレーム）を貿易統計の世界で見直すと、要求種別は次ページに示すように、検索要求の構造を示す用語が多く含まれるので、意味的には、従来検討してきた、輸出入区分、報告国・・・等検索要求の意味を示す属性と区別すべきである。 要求種別のなかで検索要求の意味を表すもの（GNP、為替レート等）は「輸出入区分」のとりうる値の一つとして扱う。

また、報告国、相手国がグループの場合国別の情報が欲しいかグループ合計の情報が欲しいかを指定する属性（表示方法）が必要である。

上記観点から検索要求に関する属性（フレーム）をまとめなおすと下記のような。< >の中は、デフォルト値を示す。（輸出入区分に限り入力情報が必要。輸出入区分が輸出で、他の属性に相当するキーワードが見つからないときは、「日本の年間輸出総額について最新の情報がドルベースで欲しい」と解釈されることになる。）

輸出入区分によって必要な属性と不要な属性が決まる。（国民総生産等は、相手国、品目、は不要）

A. 検索要求の意味を定義する属性

輸出入区分	< >
報告国／表示方法（グループの場合国別か合計か）	<日本／合計 >
相手国／表示方法（グループの場合国別か合計か）	<世界／合計 >
品目	<全品目>
期種	<年 >
時点	<最新>
単位	<ドル>

B. 検索要求の構造を定義する属性

要求種別／要求構造・時点構成・表示形式 <金額／単独・単一・表型>

②対象とする検索要求の構造

—— 検索要求から読み取る必要のある構造（要求種別の分類案） ——

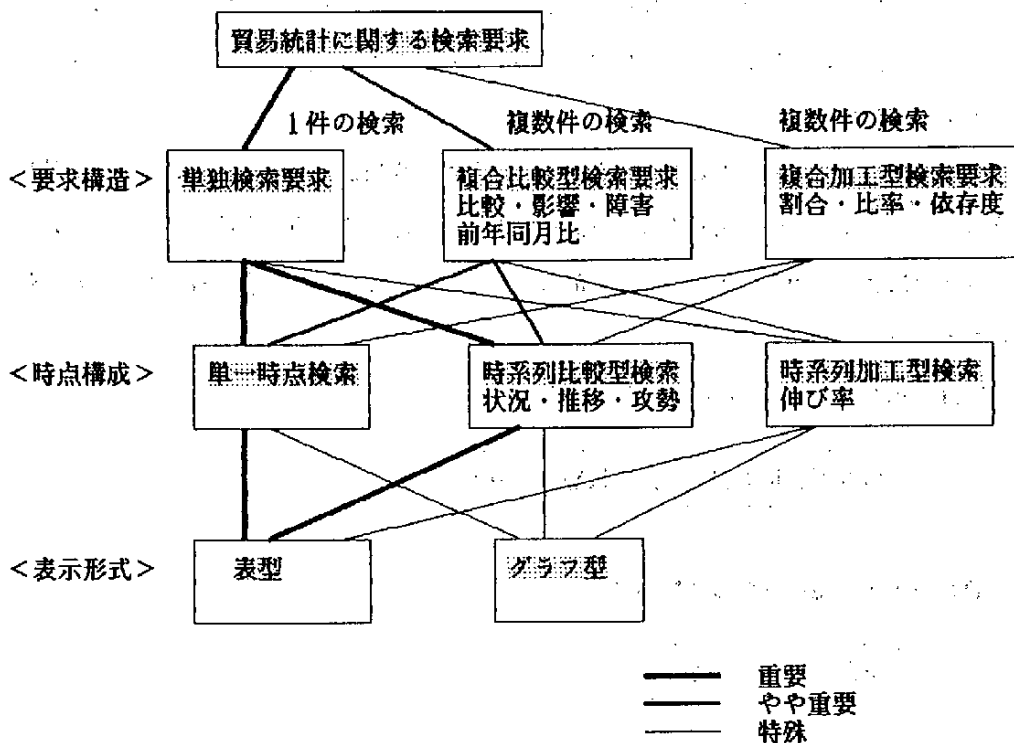


図3-3 検索要求の構造

構造を示す用語を検索要求から除くために、インターフェイスの設計の際に、これらの用語が入力されないようなテンプレートを用意することを検討した。

(4) さらに、現在のGNP、為替レート等の貿易統計の範囲を越えた内容については、(3)で検討したフレームも十分な表現能力を持つとはいえないため、貿易統計以外の分野も考えたうえで、さらに一般的なフレームを定義することができるか否かの検討をおこなった。

今回の調査対象としたユーザの検索要求のなかで、貿易統計そのものではないもののようなものがある。

GNP、生産額、為替レート、消費、取引額、対外債務残高、直接投資高
貸付資金フロー、純資金フロー、国内販売量

これらの用語は、(3)で定義したフレームを使うと相手国、品目等が不要になり、基本的に、貿易統計とは異なるフレームで表現されるべきであると考えられた。

検討の結果、GNP、生産額、為替レート等は、輸出、輸入等と同じレベルの概念として取り扱う以外に適切な属性を定義することができず、輸出入区分の代わりに、「事項」という属性を設けて、輸出入を含むこれらの用語を収容することにした。

- (5) さらにフレームを構成する属性のなかで統計レベルのもの（DBを選択する基準となる属性）とDBレベルのもの（DBが選択された後に決定されるべき属性）を区別して考えることにした。

また統計レベルの属性として、状態を付け加えた。状態は、最新値、動向、推移等検索要求のサンプルのなかによく現れる概念的な用語を収容するための属性である。状態は、属性入力によって当初考えていた概念的な入力が難しくなることを避けるため、新たに付け加えられた。

統計レベルの属性は、次の5種類である

報告国、相手国、品目、状態、事項

DBレベルの属性は次の3種類である

時点、期種、単位

このような分類をおこなったとき、相手国、品目の各属性は、事項の内容によっては、不要あるいは、無意味となる。表3. 3に属性の組合せの例を示す。

表3. 3 事項と対応するその他の属性の関係

	輸出額	輸入額	GNP	為替レート	価格指数	
報告国	○	○	○	○	○	
相手国	○	○		○		
品目	○	○				
状態	A	A	B	B	B	

Aは最新値、前年同期比、推移、傾向、対GNP比等
Bは最新値、前年同期比、推移、傾向等

- (6) 要求種別のなかに、貿易依存度のように、2種の統計の演算によって得られる統計が存在することが確認され、これらを表現するための知識ベースとして、導出辞書という名前の知識ベースを検討した。

例えば輸出依存度について、専門家は次のような知識を持っている。

輸出依存度・・・	定義式は	$(\text{輸出額} / \text{GNP}) \times 100$	である
	意味は	国民総生産のなかで輸出の占める割合 (%)	である
	収録統計	は 輸出額は通関統計、GNPはIMF統計	である

ユーザが輸出依存度を希望した場合、輸出依存度そのものは、データベースに収録されていないことが多いが、輸出額とGNPが収録されていれば、上記の定義式を使って算出することが可能である。

導出辞書を使うことにより、次善情報に関する支援の一部を実現することができたと言えよう。

3. 2 検索に必要な知識

前節で説明したように、ユーザの検索要求をシステムの内部に蓄えるための表現形式をフレームを使って決定した。

この検索要求に対して、種々の専門家の知識を使って、統計の世界の共通用語での変換、データベース世界での共通用語への変換をおこなって、データベース検索コマンドを生成し、データベースを検索し、ユーザに情報を提供するのが、IRSの目的である。

試作システムでは、この変換に必要な知識を(1)検索要求に関する知識、(2)類義語に関する知識、(3)統計に関する知識、(4)データベースに関する知識、(5)画面及びレイアウトに関する知識、(6)検索コマンドに関する知識の6種類の知識ベースに収集した。尚、本システムでの知識ベースは、prolog述語の形式で蓄えられている。

3.2.1 検索要求に関する知識

検索要求を蓄えるフレームの構造を認識するための知識として、(1) 検索要求の定義に関する知識、(2) 属性間の関連に関する知識の2種類の知識が存在する。

(1) 検索要求の定義に関する知識

検索要求を表現したフレームの定義に関する知識で、下記の書式で記述される。

辞書名：属性定義辞書 基本書式：def-atr([A1,A2...]).		
変数の意味	変数名	意味
	A-i	検索要求を表現するフレームを構成する各属性の識別子 (注)
注. 現在 r-c(報告国), d-c(相手国), kind(事項), cond(状態), goods(品目) の5種		

(2) 属性の関係に関する知識

属性と属性間に特別の関連がある場合は、下記の書式に従って関連を定義する。

辞書名：属性定義辞書 基本書式：def-atr([A1,A2...]).		
変数の意味	変数名	意味
	R	属性間の関連を示す識別子
	A-i	関連R によって関連を定義される属性 (注)
注. 属性定義辞書で登録された属性名を使う。基本的に順序はR によって固定されている		

検索要求に関する知識は、基本的にフレームの構造定義に関するもので、フレームの定義そのものを変更したい場合以外は、修正することはまれである。但しこのように、データとして、フレーム構造が定義できれば、貿易統計以外の分野を追加する場合に、現在のシステムの制御そのものはほとんど手をくわえることなく、拡張が可能になる。

3.2.2 類義語に関する知識

ユーザが入力した用語を統計の標準的な用語に置き換えるために、基本名詞に対する類義語の集合を記述する辞書を定義した。

試作システムでは、国名と品目に関する類義語辞書が下記の書式で記述されている。

辞書名：類義語辞書 基本書式：dic(Key, Name, [Name1, Name2...]).		
変数の意味	変数名	意味
	Key	国・品目の区分を示すキー r __c, goods の2種類 (注)
	Name	基本名詞
	Name-i	類義語
注. r __c : 報告国及び相手国 (共通) goods : 品目		

このような形式で知識を蓄えることにより、1件の基本名詞に類義語を何種類でも定義

することができる。但し、現在のところ同じ類義語をもった基本名詞が複数存在するときの取扱は、明確に決められておらず、最初に見つかった基本名詞が採用されるようになっている。また、基本名詞と基本名詞の間の意味的な関連も考慮していない。

類義語辞書は、最も件数の多い辞書であり、今後内容を充実してゆくうえで、高速の検索を行うしくみを検討する必要がある。また更新の際に、既存の知識と矛盾や齟齬が発生しないかどうかチェックを行うしくみも必要である。

実用的な類義語辞書を構築するには、統計の分類体系をもとにした大規模な作業が必要であるが、ユーザが使用した用語を逐次類義語辞書に蓄える機能を工夫すれば、稼働回数が増えるにつれて、自動的に、辞書を充実させることが可能である。

但し、この場合類義語のなかで、個人的な言い回しに相当する、特別な用語が共通知識ベースに蓄えられないように、監視をする必要がある。（個人辞書と共有辞書の区別の必要性も指摘された）

3.2.3 統計に関する知識及びデータベースに関する知識

統計とデータベースに関する知識は、本来区別して考えるべきであると思われるが、實際上、データベースの専門家が統計に関しても知識を持ってコンサルティングしている場合が多く、試作システムの開発では、専門家の知識のうち、検索コマンドに係るものと、それ以外の知識に分けて知識ベース化した。

本節では、検索コマンド以外の統計及びデータベースに関する知識について説明する。

統計及びデータベースに関する知識としては、(1)DB種別辞書、(2)DB属性辞書、(3)状態メニュー辞書、(4)説明辞書、(5)導出辞書、(6)導出式辞書の6種類である。

(1)DB種別辞書

取り扱うデータベースの種類を下記の書式で定義する。

辞書名：DB種別辞書 基本書式：def-db([D1,D2...])		
	変数名	意味
変数の意味	D-i	取り扱うデータベースを示す識別子
注. 現在 日本貿易統計、各国間貿易統計、世界経済統計 の3種		

データベースの種類の更新に対応可能である。

(2) DB属性辞書

データベースの属性及び収録条件に関する知識を下記の書式で定義する。

辞書名：DB属性辞書 基本書式：dbd(D, [A1,A2...], [K1,K2...], [T1,T2...], [V1,Y2...], [P1,P2...], [U1,U2...])		
	変数名	意味
変数の意味	D	データベースを示す識別子 (def-dbで定義したもの)
	A-i	D が収録するデータの属性を示す識別子
	K-i	D が収録する事項を示す識別子

T1, T2	D の収録時点 (T1~T2), T1<T2, T1とT2は西暦年号
Y-i	年 (年) 、暦年 (年度) の内、D が収録するもの
P-i	D が収録する期種 (年、半期、四半期、月の内)
U-i	D が収録する単位 (円、ドルの内)

(3) 状態メニュー辞書

状態のメニューに関する知識を、下記の書式で蓄える。

辞書名：状態メニュー辞書 基本書式：def-cond(C, [P1,P2,P3] , [K1,K2...]) .		
変数の意味	変数名	意味
	C	状態を示す識別子
	P1	Cの示す期間の開始時点に関する指示 (注1)
	P2	Cの示す期間の終了時点に関する指示 (注2)
	P3	Cの示す期間の年数に関する指示 (注3)
	K-1	Cを状態として持つ事項の識別子
注1 ユーザに指示をさせるとき=y、それ以外るとき=D 注2 ユーザに指示をさせるとき=y、最新値のとき=new、それ以外るとき=D 注3 数字で、期間に含まれる年数を指定する		

(4)説明辞書

事項及び状態に関する説明文を下記の書式で蓄える。

辞書名：説明辞書 基本書式：exp(K,'E').		
変数の意味	変数名	意味
	K	事項または、状態を示す識別子
	E	説明文
注. 説明文は、説明アイコンがクリックされた時に説明ウィンドウに表示される		

(5)導出辞書

導出の必要な事項に関して、導出の方法を、下記の書式で定義する。

辞書名：導出辞書 基本書式：rd(K, ['V1','V2'...], [N1,N2...],\$E\$, [A1,A2:...]).		
変数の意味	変数名	意味
	K	事項を示す識別子
	V-i	K を算出するために必要な変数名 (注1)
	N-i	K を算出する為の演算式番号 (注2、注3)
	E	説明文
	A-i	K の属性を示す識別子
注1. 必ず取り扱うデータベースに収録されていること 注2. 必ず導出式辞書で定義されていること 注3. 導出式は複数指定が可能である		

(6)導出式辞書

導出に必要な算式を、下記の書式で定義する。

辞書名：導出式辞書 基本書式：rd-f([NO,\$FORM\$]).		
変数	変数名	意味
	NO	導出式番号

の 意 味	FORM	導出式 (注)
注. 導出式に使うことのできる変数は、A,B,C,D,E の5変数である		

3.2.4 画面制御及びレイアウトに関する知識

画面制御に関する知識は、画面上で、入力用の枠をどのように表示するか・・・という知識で、制御情報に近い。しかし、内容として、各属性のデフォルト値等の情報を持ち、また、対象分野を追加し、入力フレームと画面を連結させる情報なので、知識として取り扱う。

(1)入力枠定義辞書

属性入力枠の位置とデフォルト（初期に表示される文字列）の定義を、下記の書式で記述する。

辞書名：入力枠定義辞書 基本書式：iw(A, [C1,C2], [\$S1\$, \$S2\$, SW])		
変 数 の 意 味	変数名	意味
	A	属性を示す識別子 (注1)
	C-i	表示位置の x y 座標
	S1	Aのデフォルト
	S2	Aの枠の後に表示する文字列 (注2)
	SW	連動するメニューウィンドウ名 (注3)
注1 現在 r-c, d-c, kind, cond, goods の5種		
注2 「の」、「に対する」等		
注3 連動するメニューがないときは、off と記入		

(2) メニュー定義辞書

メニュー画面に関する定義を、下記の書式で行う。

辞書名：メニュー定義辞書 基本書式：mw(A, [C1,C2], [N1,N2,AR,NAME]).		
変数の意味	変数名	意味
	A	タイトル
	C-i	表示位置の x y 座標
	N1,N2	X方向, Y方向の分割数
	AR	属性を示す識別子 (注1)
	NAME	ウィンドウ名 (注2)
注1 現在 kind,cond, の2 種 注2 ウィンドウ定義辞書で定義されていること		

(3) ウィンドウ定義辞書

属性入力枠以外のウィンドウの位置の定義は、下記の書式で記述する

辞書名：ウィンドウ 定義辞書 基本書式：wind(NAME,C1,C2,C3,C4,KIND).		
変数の意味	変数名	意味
	NAME	ウィンドウ名
	C-i	表示位置の x y 座標 (左上、右下)
	KIND	ウィンドウの種類 (注)
注 現在 icw(メニューウィンドウ) のみ		

3.2.5 検索コマンド生成のための知識

検索コマンドを生成するためには、次の2つの知識が必要である。

(1) コマンドに関する知識

(2) コードに関する知識

以下、試作システムで実現したこれらの知識について述べる。

(1) コマンドに関する知識

一般的に検索コマンドは、各データベースの特徴を反映するように設計されている。すなわち、同一システム内で提供するデータベース群について、検索し易いようになっているため各システムによりコマンド体系が異なる。このため、検索時には異なるコマンド体系毎に検索目的に応じたコマンドを選択したり、同一コマンド体系においても適切なパラメータ設定を行うことが必要となる。コマンドに関する知識とは、検索目的に応じたコマンドを選択し、パラメータ設定を行う知識といえる。

試作システムでは、経済・貿易分野に関する次の3つのデータベースを対象にしている

①世界経済統計：IMF加盟国の主要経済指標に関する情報等を収録したデータベース。

②日本貿易統計：日本の輸出入に関する貿易情報を収録したデータベース。

③各国間貿易統計：任意の2国間に関する貿易総額を収録したデータベース。

これらのデータベースでは、データベース検索用に汎用検索言語（以下、EUL:End User Language と略記）を備えており、検索コマンドとして表3.4に示すようなものが用意されている。今回のプロトタイプでは、検索のみを目的とし、表示、加工等の検索後の処理については対象としていない。

このため、項目コードを直接指定して検索指定を行う基本コマンド「RETR」コマンドについて、3つのデータベース毎にどのようなパラメータを設定すればよいかという知識を検索コマンド生成プログラムで表現している。

その他、導出を必要とする事項が選択された場合について、次のようなホスト側で加工を行うコマンドを試験的に生成するようにしている。

・上位nカ国：MAJORコマンド（日本貿易統計）

・上位n品目：MAJORコマンド（日本貿易統計）

・前年同期比：DISP-SRATIOコマンド（世界経済統計・日本貿易統計）

RATIOコマンド（日本貿易統計）

表3.4 EULのコマンド

- ◇RETR: データを検索する。(項目を直接指定できるもの)
- ◇FIND: データを検索する。(項目を直接指定できないもの)
- ◇DISP: データを表示する。“前年同期比”, “成長率”などの関数を使用できる。
- ◇GRAP: データをグラフ表示する。折れ線グラフと棒グラフが表示できる。
- ◇KEYL: 指定されたデータ項目について、データ数を出力する。
- ◇MOVE: 検索結果をホストの作業領域に出力する。
- ◇SAVE: 検索結果を指定されたファイルに保存する。
- ◇AND, OR, NOT: 保存されたファイル間の集合演算を行う。
- ◇LET: 演算式を定義する。
- ◇SORT: データの並べ替えを行う。
- ◇PTON LP: 検索結果をプリンターへ出力する。
- ◇PTOF: PTONコマンドを解除する。
- ◇DONE: 検索を終了する。

(2) コードに関する知識

貿易分野においては、「いつ」、「どこ(国)」から「どこ(国)」へ、「何(品物・価値)」が移動したかということが問題となる。すなわち、データベースの項目として時点、国・地域名、品目名、価額などに関する情報が収録されている。これらの収録項目について、例えば次のような対応コードが決められている。検索時には、これらのコードを指定することにより、検索が行われる。

◎時点: 暦年…CY

会計年度…FY

四半期…CQ

月…M

◎国・地域名: 日本←→158(世界経済統計・各国間貿易統計)

西ドイツ←→134(世界経済統計・各国間貿易統計), 213(日本貿易統計)

◎品目名：自動車←→2934（日本貿易統計）

テレビ←→2941（日本貿易統計）

コードに関する知識は、データベース選択・制御部から引き渡された項目情報を対応するコードに変換する知識である。プロトタイプでは、コードに関する知識を知識ベースとして表現している。コードに関する知識ベースの概要は次の通りである。

◆知識ベースの構築

- ・品目：概況分類約900品目
- ・国・地域：約230カ国、15地域
- ・IFS系列：約40系列

◆知識ベースの構造

- ・品目の知識ベース

g_code (Name, Code) .

Name：品目名（基本名詞）

Code：概況分類の品目コード

一般品目… [輸出コード, 輸入コード]

特殊品目… '先頭を除く2桁以降のコード'

例. g_code (食料品及び動物, ['310', '410']) .

g_code (鉄鋼, '2921') .

- ・国・地域の知識ベース（基本名詞）

c_code (Name, Code_list) .

Name：国名、地域名

Code_list：[世界経済統計コード、各国間貿易統計コード、日本貿易統計コード]

例. c_code (日本, [158, 158, *]) . ; '*' は、対応するコードがないことを示す。

- ・系列コードの知識ベース

k_code (Name, Code, Attribute-List) .

Name：系列コード名

Code：系列コード

Attribute-list：同一の系列名に対して複数の系列コードが存在するので、一つのコードを特定するための属性リスト。

例. k_code (GNP, [unit('NC')]) .

K_code (GNP, [unit('NC'), r_c(西ドイツ)]) .

3. 3 知識獲得の経緯

IRSの知識ベースは、6種類の内容を持っているが、今後の運用上知識の内容の正確さや、追加更新のし易さが、知識ベースの質を確保する上で重要である。

知識の獲得にあたっては、データベースのマニュアルから抽出したもの、専門家へのアンケートや、インタビュー等を行った。

現状では、試作システムとしてのIRSのための知識収集が終了した状態であるが、今後の拡張や、メンテナンスを念頭において、知識獲得の経緯をまとめておく。

(1)検索要求の表現に関する知識獲得

検索要求の表現に関しては、3. 1に詳しく経緯を説明したが、知識獲得として、まとめると、下記ようになる。

- ① 通産ジャーナルより検索要求を抽出(100数件)
- ② 専門家との協議を経て、設定したフレームに①を適用して、不具合点を検討し、再度フレームを設定しなおす。また、検索対象としてとりあげた3種類のデータベースのマニュアルより、収録項目を調べて、この範囲に関して、確かに不都合が生じるのか否か検討する。

このようにして最終的なフレームが決定するまでには、半年以上の時間がかかっている。ユーザの使用イメージや利用範囲が明確になっていれば、もう少し速く議論が収束したとおわれるが、なるべく多くのユーザに対応できるものを作りたいという目的があったため、敢えて、利用範囲を限定せずに作業をおこなった。

また、データベース側からでなく、ユーザ側からの要求表現を検討したいときに、データベースに関する知識があると、かえって純粋なユーザ要求を捉えにくい・・・という現象も見られた。素人がなにを欲しているか・・・というような知識に関しては、もう少し、方法を検討する余地があると思われる。

(2)類義語に関する知識獲得

国名に関しては、データベースのマニュアルに記載された国名を基本名詞として、基本名詞に関する知識獲得はできたが、類義語に関しては、よく使われる国名を開発者の判断で登

録した段階で、件数も少ない。

百科辞典等からの獲得や、市販のソースの利用等が今後の対策として考えられる。

品目に関しては、概況分類を基準に基本名詞を収集した。類義語は、国名と同様で、大規模な知識獲得は、行っていない。品目に関する類義語は、国名に関するものよりさらに多いことが予想され、また慣用的な呼び方が人によってまちまちである場合がある。

個人的な辞書と共有辞書の使い分けを工夫するとともに、自動的な知識獲得や、知識の整合性のチェック機構等、知識の管理を行う機能が必要である。

(3)統計及びデータベースに関する知識獲得

統計及びデータベースに関する知識に関しては、知識ベースの構造を決定するまえに、対象とするデータベースのマニュアルより、下表の知識を獲得した。

表 3-5 データベースに関する知識

注. 日本貿易統計 2 は、日本貿易統計を
品目コードに関して改訂したもの

日本貿易統計	日本貿易統計 2	各国間貿易統計	世界経済統計
・ 出典 大蔵省日本貿易月表 ・ 輸出入合わせて 12 千品目の相手国別通関金額、数量 ・ 最新 36 ヶ月分の月次データと最新 7 年分の年次データ	- 日本貿易統計と同様 ・ 輸出入合わせて 22 千品目 - 日本貿易統計と同様	・ 出典 IMF 作成 Direction Of Trade Statistics ・ 2 国間輸出入額 輸出入シェア、双極化指数 ・ 最新 40 時点分の年データと最新 72 時点分の月次データ、最新 72 時点分の四半期デ	・ 出典 IMF 作成 International Financial Statistics ・ 為替レート、国際流動性、通貨当局、預金通貨銀行、マネタリサーベイ、その他の金融、国際取引、利子、

・月末に前月のデータが収録される。 ・品目コードは、BTN、通自分類、概況分類の3種 (注 BTN=CCCN) ・最終は1987年12月以降はHSTRADEに切替え ・税関を通過する貨物に関する統計 但し10万円以下の貨物や見本品等は除く。 ・計上時期 輸出=出航月 輸入=許可、承認月 ・国 輸出=仕向け国 輸入=原産国 220国、33地域 ・数量(統計品目標による) 単位未清切捨て 重量は純重量 ・価格(千円及びドル単位) 単位未清切捨て ・12月にデータの補正行う ・グラフ、国別一覧表等作成の機能あり	・品目コードはHS ・1988以降 - 日本貿易統計と同様	・2ヵ月前のデータ(20日更新) ・品目分類無 ・国コードは日本貿易統計と違う 160国 ・100万ドル単位 輸出=F0B価格 輸入=CIF価格 ・グラフ作表あり	- 物価、生産、国際収支、政府金融、国民所得 ・1960以降の年データ ・1970:1期以降の四半期 ・1975:1以降の月データ ・2ヵ月前のデータ(10日更新) ・国コード 160国 ・グラフ作表あり
---	---	---	--

・検索コマンドは独自			・検索コマンド EUL	・EUL
------------	--	--	----------------	------

表3-5は、データベースの収録に関する情報であるが、さらにデータベースの検索属性より、つぎのようなデータベース選択手順が得られた。

日本に関する輸出入で、品目の指定があれば日本貿易統計
 日本に関する輸出入で、品目の指定がなければ日本貿易統計か各国間貿易統計
 日本に関係のない輸出入で、品目の指定があれば検索不可能
 日本に関係のない輸出入で、品目の指定がなければ。各国間貿易統計
輸出入以外で為替レートならば世界経済統計
輸出入以外でGNPならば世界経済統計

また、貿易統計以外のデータで、貿易統計と関連の深いデータについて、専門家からのアンケートにより、知識獲得を行った。（付録5）

実際に専門知識のない開発担当者が検索をおこない、不明な点について、専門家に問い合わせ、データベースQ&A集を作成した。（付録3）

状態として、蓄積すべき項目を収集するために、通商白書の図表タイトルを使って、専門家に重要度を判定してもらうアンケートを実施した（付録2）

(4)検索コマンドに関する知識獲得

データベースのマニュアルを使い知識獲得をおこなった。

マニュアルによって判断の難しい、検索コードの選定方法に関しては、専門家に問い合わせた。（付録3）

検索コードについては、マニュアルから十分な情報がえられないものがあり、専門家も、実際に検索して、得られたデータをみて、判断を行う場合もあることが判った。

1. The first part of the report is a general introduction to the subject of the study.

2. The second part of the report is a detailed description of the methods used in the study.

3. The third part of the report is a discussion of the results of the study.

4. The fourth part of the report is a conclusion and a list of references.

5. The fifth part of the report is a list of references.

6. The sixth part of the report is a list of references.

7. The seventh part of the report is a list of references.

8. The eighth part of the report is a list of references.

9. The ninth part of the report is a list of references.

10. The tenth part of the report is a list of references.

11. The eleventh part of the report is a list of references.

第4章 IRSのアルゴリズム

第4章 IRSのアルゴリズム

本章では、IRSの機能を実現するための主要なアルゴリズムに関して説明を行う。

IRSでは、“オブジェクト指向”と呼ばれる制御方式を採用しており、システムとしての処理は、各機能の主体である“オブジェクト”に定義された変数や手続きを、“メッセージパッシング”と呼ばれるオブジェクト間の通信によって起動して、実現している。

特にIRSに特有なアルゴリズムは、知識ベースを使って検索の支援をおこなう部分である。知識ベースは、3章で説明した構造をもっており、知識ベースごとに、異なるアルゴリズムが存在し、全体としてのコンサルティング機能を実現している。

支援の手順に従って、①データベース選択、②データの導出、③状態の読替え、④検索コマンドの生成・・・の各機能について、知識ベースとの関わりを中心に、アルゴリズムを説明する。

システムの拡張性を考慮すると、このアルゴリズムが、ある程度の一般性を持っていることが、重要であると思われる。

4.1 データベース選択

データベース選択は、試作システムの基本的な機能で、ユーザの入力した検索要求の属性から検索可能なデータベースを知識ベースを使って選びだし、ユーザが検索不可能な要求を入力しないよう入力画面を書き換える機能である。

4.1.1 選択のための処理の流れ

試作システムによるデータベース選択の基本的な処理は、下記のようになっている。

- ① 検索要求を5個のキーワードから構成される文章としてユーザに示し、キーワード部分をユーザに入力させる。(キーワードは検索要求の属性を示している。)
- 事項と状態については、メニュー画面及び説明画面が表示される。

画面例

(報告国)の (相手国) に対する
(事項) に関して
(状態) を
(品目) について知りたい

() がキーワード

② キーワードを1件入力する度に選択可能なデータベースの絞り込みを行い、知識ベースを使って表示されている入力画面のメニューやデフォルト値を変更する。(ユーザが入力する画面に表示されているメニューやデフォルト値は、必ず検索可能なものだけになっている)

4.1.2 必要な知識ベースとデータ

データベース選択に必要な知識ベースと処理の過程で使用するデータを下記に示す。

(1) データベースの属性、収録状況に関する知識

DB属性辞書

dbd(DB_NAME,ATR,KIND,PERIOD,C_F,STEP,UNIT)

DB_NAME : データベースの名前

ATR : データベースの属性及び定義域のリスト

下記の属性名の中から当該データベースの属性となっているものを記入する。定義域に制限があれば、属性名の後に([定義域])の形式で追加する

属性名: r _c . . . 報告国

d _c . . . 相手国

kind . . . 要求種別

cond . . . 条件

goods . . . 品目

KIND : データベースに収録された事項のリスト

PERIOD : データベースの収録期間リスト

C __F : 年、年度の種別リスト

STEP : 年、半期、四半期、月等期種のリスト

UNIT : 単位リスト

例 dbd(trade,[r __c([日本]),d __c,cond,goods,kind],
,[輸出額, 輸入額],
,[1981,1988],
,[年, 年度],
,[年, 半期, 四半期, 月],
,[円, ドル]).

(2) 検索要求に関するデータ

入力されたキーワードの属性名と入力値

例 r __c, 日本

(3) 検索要求の入力状況に関するデータ

検索可能なデータベースのリスト (入力状況に応じて内容が変更する)

例 [db1,db2,db3]

(4) 入力画面表示用データ

事項用メニューデータ

例 mwdb(事項、輸出額).

状態用メニューデータ

例 mwdb(状態、最新値).

属性別デフォルト値

例 d __c(全世界).

4.1.3 データベース選択のアルゴリズム

(1) 入力された属性名と入力値がデータベースの属性と定義域にマッチするか否か

検索可能なデータベースのリストにあるデータベースについてチェックする。

(2) マッチしないデータベースを検索可能なデータベースのリストから除く。

(3) 検索可能なデータベースのリストに含まれるデータベースの属性リストの和集

合が以降入力すべき属性となる。(定義域が固定値ならば属性値は、自動的に固定される。但し、事項が導出を伴う場合は、この処理は行わない。次節参照)

(4) 検索可能なデータベースのリストに含まれるデータベースの事項リストの和集

合が以降選択すべき事項のメニューとなる。

(5) iii、ivをもとに入力画面を変更し、すべての属性が確定するまで i から v の処理を繰り返す。

4.1.4 知識のメンテナンス及び拡張性

データベースの収録状況に関する修正は、プログラムの手直しではなく、prologのデータベースの修正で可能になっている。

貿易統計以外の分野を対象とする場合、事項、状況以外の属性は、無意味になり、新たに必要な属性が発生すると思われるが、その分野に関する検索要求を、貿易統計と同様に属性の組で表現することができるのであれば、データベースの収録状況に関しては、対応が可能である。

報告国と相手国が交換可能であるという関係等、属性間の固有の関係については、まだ対応ができていない。。このような関係の一般化は、かなり困難と思われ検討を要す。

メニュー画面は、データベース収録状況から自動的に生成されるので、メンテナンス性は良い。説明画面の修正は、prologのデータベースの修正で可能になっている。

データベースの種類が増えた時に、あるデータベースにとって意味のない属性がテンプレート上に表示される場合がある。(現在でも世界経済統計の場合相手国、品目は、無意味であるが、相手国=全世界、品目=全品目 と表示させている)

全域カテゴリ(全世界、全品目等)をデフォルトとするとテンプレートが不自然になる場合どのような対応をするべきか検討を要す。

4. 2 導出に関するアルゴリズム

検索要求の中の「事項」は直接データベースに収録されている場合と、データベースに収録されたデータを加工して得られる場合がある。後者の場合、事項として入力された用語から知識ベースに収録されているデータの加工方法を検索し、要求されたデータを導く（ここでは、これを導出と呼ぶ）

また導出の場合は、複数のデータベースを検索して得られたデータに演算を施して、1つのデータを求める場合があり、データベースの取扱上注意が必要である。

2.2.1 導出のための処理の流れ

導出は、ユーザが属性入力の際に、事項を選択し、事項のメニューの中から、導出の必要な項目を選択した時に発生する。

(1)ユーザが導出の必要な項目（輸入依存度等）を選択する

検索可能なデータベースは ()		入力確定	前画面へ	終了							
日本 輸入額 最新値 全品目	の 全世界 に関して を について知りたい	に対する									
<table border="1"><thead><tr><th>事項</th><th>説明</th></tr></thead><tbody><tr><td>輸出額</td><td rowspan="4">輸入依存度</td></tr><tr><td>輸入額</td></tr><tr><td>輸出数量</td></tr><tr><td>輸入数量</td></tr></tbody></table>		事項	説明	輸出額	輸入依存度	輸入額	輸出数量	輸入数量			
事項	説明										
輸出額	輸入依存度										
輸入額											
輸出数量											
輸入数量											

(2)選択された項目がDB属性辞書の事項として蓄えられていない場合は、導出が必要と判断する

(3)導出辞書をさがして導出に必要なデータベースを探す。

(4)導出に必要なデータベースの属性より、入力画面のメニューやデフォルト値を変更する

4.2.2 必要な知識ベースとデータ

(1) 加工に必要なデータ（直接データベースに収録されているもの）と加工方法

導出辞書

rd(KIND, [V1,V2,...], [FNO1,FNO2,...], \$COMMENT\$, A_LIST).

KIND : 事項の名前
 V1,V2... : 加工に必要なデータ
 FNO1,FNO2... : 加工式番号
 COMMENT : 説明文
 A __LIST : 属性のリスト

(注) 加工に必要なデータは5種類まで使うことができる。

FNO-i の加工式を定義した導出式辞書の加工式の中で、このデータが順にA,B,C,D,E と判断されて加工演算が行われる。

例 rd(輸出依存度, [輸入額, GNP], [1,2],

\$ 国民総生産に対する貿易の寄与 \$,

[r-c,d-c([全世界]), cond, goods([全商品]), kind]) .

(2) 加工式

導出式辞書

rd__f(FNO,\$FORMURA\$).

FNO : 加工式番号

FORMURA : 加工式

例 rd__f(1,\$A/B\$).

4.2.3 導出のアルゴリズム

- (1) ユーザが指定した事項がDB属性辞書の事項として蓄えられていない場合は、導出を伴うと判断し、一般的なデータベース選択のアルゴリズムと区別する。
- (2) 導出辞書に定義された属性リストが以降入力すべき属性となる。
- (3) 導出辞書に定義された加工に必要なデータを収録するデータベースをデータベース辞書より探索し、検索可能なデータベースのリストを変更する。
- (4) 検索可能なデータベースのリストに含まれるデータベースの事項リストの和集合が以降選択すべき事項のメニューとなる。
- (5) iii、ivをもとに入力画面を変更し、すべての属性が確定するまでiからvの処理を繰り返す。

4.2.4 知識のメンテナンス及び拡張性

導出は、基本的に、データベースから検索可能な変数と定義式によって定義される変数の変更や定義式の変更には、prologのデータベースの修正により対応ができる。

導出に用いる変数が更に導出を要する場合は、まだ対応していないが、若干のアルリズムの追加により対応可能である。

但し導出式が複雑になると、知識ベースが大変見にくくなり、入力の手がかりが多くなる懸念される。

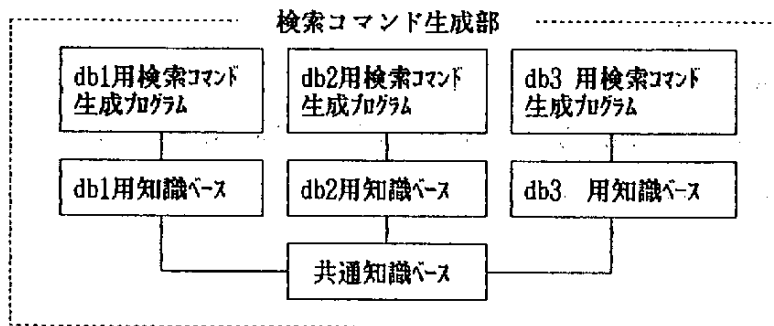
現在、導出は、2種以上のデータ間の演算を要するもの。状態は、1種のデータの表示方式に関するものと区別しているが、両方の要素をもった用語に対してどのように対応すべきか検討を要する。

4. 3 検索コマンド生成

ここでは、検索コマンド生成のアルゴリズムについて述べる。

4.3.1 検索コマンド生成部

データベース検索部の一部である検索コマンド生成部は、下図のような構成になっている。



検索コマンド生成部は、3つのプログラムと4つの知識ベースで構成されている。設計方針として各データベース毎にコマンドの生成部を設計した。これは、次のような理由による。

- (1) 検索コマンドのデータベース毎の差異が大きいこと。
- (2) データベース検索コマンド選択・制御部によって検索対象のデータベースが選択されるため、必要なコマンド生成部をその都度ロードした方がメモリーの節約になること。
- (3) コマンド生成部に変更が生じた場合、個々の生成部のみで変更が完結すること。

また、知識ベースについても共通にできるものとできないものがあり、それぞれについて設計した。

4.3.2 コマンド生成のアルゴリズム

コマンド生成は、データベース選択・制御部から引き渡された検索要求項目を知識ベースによってコードに変換し、それらをコマンドと組み合わせることによって行われる。

(1) 引き渡される検索要求項目

データベース選択制御部から引き渡される情報は基本的に以下の変数による。

- ・ jikou …GNP等の統計属性。
- ・ r_c …報告国の基本名詞。
- ・ d_c …相手国の基本名詞。
- ・ goods …品目の基本名詞
- ・ jiten …年, 年度等の時点種別、及び時点
- ・ jyoutai …上位nカ国等の状態。
- ・ unit …各国通貨建てかドル建てかの単位種別。

また、"RETR"コマンドを生成するのに必要なパラメータリストを下表に示す。

	d b 1	d b 2	d b 3
品目	○	×	×
系列	×	×	○
国・地域	○	○	○

(2)基本アルゴリズム

"RETR"コマンドを生成するアルゴリズムは次のようになる。

- i 'jikou'または'goods'の値を参照して、系列コード (db3) または品目コード (db1) を生成する。
- ii 'r_c'または'd_c'の値を参照して、国・地域コードを発生する。
- iii 'jiten'の値を参照して、時点コードを生成する。
- iv 生成したパラメータを"RETR"のコマンドフォーマットに従って並べる。

また、"RETR"コマンド以外に「状態」の値を参照にして別のコマンドを生成する場合もある。以下に、具体例を示す。

例.

「日本のソ連に対する1989年度の自動車の輸出額について知りたい。」という検索要求が入力されたとする。データベース選択・制御部は、対象データベースをd b 1 と決定し、d b 1 用のコマンド生成プログラムに次のようなパラメータを引き渡す。

<db1に引き渡されるパラメータ>

- ・ jikou …「輸出額」
- ・ r_c …TRADEでは"r_c=日本"が自明のため、このパラメータは引き渡されない。
- ・ d_c …「ソビエト連邦」（ソ連の基本名詞）
- ・ goods …「自動車」（自動車の基本名詞）
- ・ jiten …「89」及び「年度」
- ・ jyoutai …状態は指定されていないので、引き渡されない。
- ・ unit …TRADEにおける単位指定は、jyoutaiが指定されている場合に有効なのでこの場合は、引き渡されない。

<db1用検索コマンドの生成>

- i "jikou"及び"goods"の値により、自動車の輸出コード'32933'を発生する。
- ii "d_c"の値により、ソビエト連邦の国コード'224'を発生する。
- iii "jiten"の値により、'89FY'を発生する。
- ivパラメータを正しく配列して、「RETR」コマンドと結合する。
"RETR 32933/224/89FY"が生成される。

(3)ファイル出力

プロトタイプでは、検索コマンドを生成した後、そのコマンド群をフロッピーディスクにダウンロードするようになっている。この生成コマンドを用いて検索を行う場合には、一旦、IRSを終了させて通信ソフトを立ちあげてからホスト側にコマンドを送信することになる。今回は、市販の通信ソフトを用いたコマンドの送・受信テストを試み、動作を確認した。

第5章 IRSのマンマシンインターフェイス

第5章 IRSのマンマシンインターフェイス

本章では、IRSのマンマシンインターフェイスに関する説明を行う。

本年度に開発した試作システムは、パーソナルコンピュータ上で開発され、キーボードとマウスを使ってユーザが必要なデータの入力を行い、システムが、データベース検索コマンドを生成した後、ホストコンピュータ上のデータベースとの通信を行って、必要なデータ検索を行う仕組みになっている。

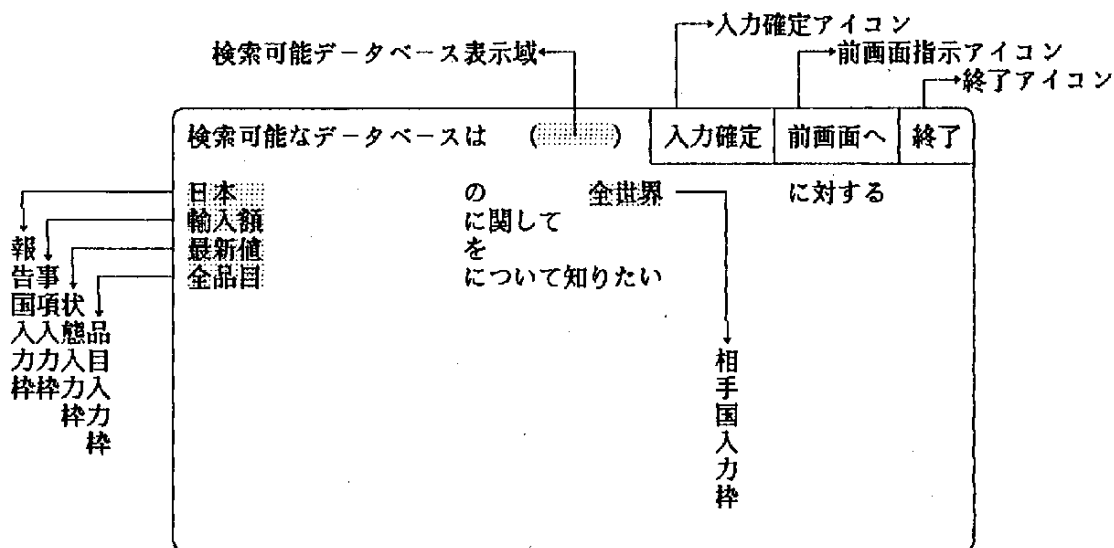
人工知能導入の一つのねらいは、ユーザが煩雑なシステム操作から開放され、自由にデータ検索をおこなうことができるようにすることである。

本章では、ユーザにとっての操作性に着目し、従来のデータベース検索と比べて、IRSが、どのような特徴を持っているかを重点的に説明する。

5.1 画面構成

(1) 検索指示画面

最初にユーザに示される画面で、検索要求として報告国、相手国、品目、事項、状態の5つの属性を入力するための画面である。



ユーザは、報告国入力枠、相手国入力枠、事項入力枠、状態入力枠、品目入力枠に検索し

たい情報の内容を入力する。これらの入力枠には、初期には、デフォルト値として、最も入力頻度の高そうなデータを表示する。

報告国入力枠、相手国入力枠、品目入力枠は、キーボードから文字列を入力する。入力の内容は、それぞれ報告国、相手国、品目である。

事項入力枠、状態入力枠は、メニューから希望する項目を選択する。事項入力枠に入力する内容は、輸出、輸入、GNP等、統計の系列名に相当するものである。

状態入力枠に入力する内容は、最新値、最近の10年間等、時点に関する概念を示すものまたは、上位10品目等、系列の選択方法に関するもの等である。

検索可能データベース表示域は、検索可能なデータベース名を表示する領域である。初期には、システムで取り扱うすべてのデータベースが表示されており、ユーザが属性を入力する度に、検索可能なデータベースのチェックを行い、その時点で検索可能なデータベース名を更新し、この領域に表示する。

入力確定アイコンは、入力枠の一部に入力をおこなった時点で、残りの入力枠は、デフォルト値をそのまま使いたい時にマウスでクリックする。

前画面指示アイコンは、入力の途中で、前の画面に戻りたい時にマウスでクリックする

終了アイコンは、システムを終了した移時にマウスでクリックする。

(2) メニュー画面 (事項と状態)

事項と状態の2属性に関して、入力値を選択するための画面である。

検索可能なデータベースは ()		入力確定	前画面へ	終了
日本 輸入額 最新値 全品目		の 全世界 に対する に関して を について知りたい		
事項	説明	状態	説明	
輸出額 輸入額 輸出数量 輸入数量	説明 ↓ 説明 アイ コン	最新値 推移 前年同期比 GNP比	説明 ↓ 説明 アイ コン	
→ 事項メニュー		→ 状態メニュー		

注. ユーザの指定に応じて事項と状態のどちらかが表示される

メニュー画面は、事項または、状態をユーザが選択したときに表示される。ユーザは、

このメニューの中から事項または、状態を選択できる。

事項の選択、状態の選択には、マウスを用いる。

またメニューに表示された用語の意味が不明のときは、説明アイコンをクリックして説明画面を表示させることができる。

(3) 説明画面

メニュー画面に表示された項目の意味を確認するための画面である。

検索可能なデータベースは ()		入力確定	前画面へ	終了
日本 輸入額 最新値 全品目	の 全世界 に対して を について知りたい			
事項		説明		
説明画面				
***** 1 輸入額/GNP 2 *****				
類義語：***** 定義式の番号を入力して下さい		注. ユーザが「説明」を指示した時に表示される		

説明画面は、ユーザが事項または、状態の入力の際に、説明アイコンをクリックした場合に、事項または状態メニューの上に上書きされる形に表示される。

説明画面に表示される内容は、説明辞書に記述された説明文を表示する。また事項の場合で、導出を伴う用語のときは、導出式も同時に表示される。

導出式が複数あるときは、ユーザがどの式を使うかをこの画面から指示する。

(4) 補助項目入力画面 及び メッセージ画面

補助入力画面は、時点、年・年度の区別、期種、単位を選択するための画面である。
メッセージ画面は、システムからのメッセージを表示するための画面である。

検索可能なデータベースは ()		入力確定	前画面へ	終了
日本 輸入額 最新値 全品目	の に関して を について知りたい	全世界	に対する	
1982 輸入額 年 (年/四半期/月)	を	1988 年 (年/年度)	の	
表示します		毎に 円 (円/ドル)	ベースで	
数字を入力してください				

→メッセージ画面

補助項目入力画面 ←

補助項目入力画面は検索すべきデータベースが決定した後、時点単位等データベース毎の補助項目を入力するために使う。

画面には予めデータベースの収録に合わせてデフォルト値が設定されている。

5. 2 データ入力方式

本節では、画面の遷移に伴って、ユーザの操作手順を説明する。

試作システムでは、ユーザは、キーボード入力とマウス入力の2通りを使い分ける必要がある。物理的な操作の手間としては、これは、煩雑であるといえる。

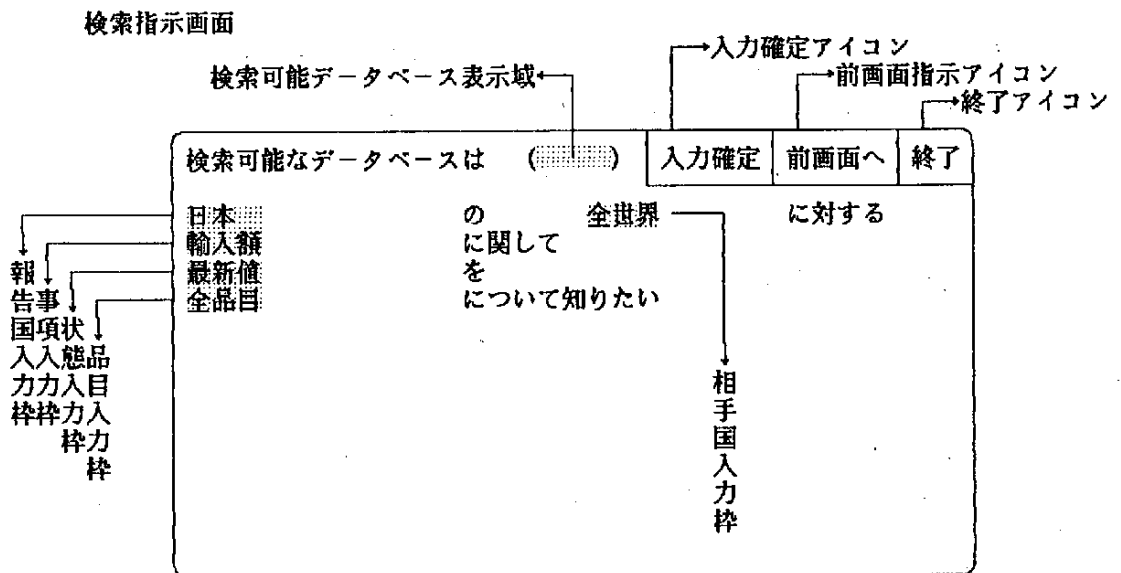
入力の操作性としては、このほかに、入力すべきデータが明確であるか、マニュアルがなくても操作が可能であるか、ユーザの望む結果が簡単に引き出せるか（検索してから、自分の思うものと違っているということがないか）等を考慮する必要がある。

以下に操作手順を示しながら、操作性について考察する。

5.2.1 操作手順

入力は、基本的にマウスを使って行う。但し報告国入力枠、相手国入力枠、品目入力枠は、キーボードから入力を行う。

マウス操作は、枠を指示したい時に左ボタンをクリックし、表示されている説明画面を消去するときや、一度指示したものをキャンセルしたいときに、右ボタンをクリックする。



基本的な作手順を下記に示す。（詳細は、付録の操作マニュアル参照）

- ①検索指示画面にユーザが検索要求を入力するときは、マウスを使って入力したい入力枠（報告国、相手国、事項、状態、品目のいずれか）をクリックする。入力枠への入力はどうのような順序でもかまわない。

②入力枠をクリックすると入力枠が入力可能な状態になる。報告国、相手国、品目は、キーボードより文字列を入力する。事項、状態は、入力枠をクリックした時点で、メニュー画面が表示されるので、メニューの中からマウスで選択する。

③入力が終了すると、入力枠の色が変わり、マウスによる指示に反応しなくなる（入力確定の状態になる）。データベースが限定され、入力可能な属性値にも制限が加えられるときは、入力枠は自動的に確定状態になる。

④全ての入力枠が入力確定の状態になると、自動的に、条件入力画面を表示する。

入力確定状態になっていない入力枠が残っていても、入力確定アイコンをクリックすると、各入力枠に表示されたデフォルト値を入力値と見做して、入力を確定し、条件入力画面を表示する。

⑤条件入力画面には、選択されたデータベースにあわせて、時点や、単位のデフォルト値を入力画面に表示する。

⑥条件入力画面の入力枠も入力順序は自由である。入力枠をクリックすると、システムメッセージ表示域が表示される。

⑦時点の入力は、西暦の数字を入力する。数字を入力して改行キーを押すと入力確定の状態になる。

⑧単位、期種は、ユーザの選択することのできる項目を（ ）内に表示する。マウスの右ボタンをクリックすると、枠内の文字が（ ）内に表示された順に変わってゆく。

ユーザは、選択したい項目が入力枠に表示された状態で、マウスの左ボタンをクリックすると入力確定状態になる。

⑨全ての条件入力終了すると検索コマンドをシステムメッセージ域に表示する。

5.2.2 従来システムとの比較

従来のデータベース検索システムと比べて試作システムのインターフェイスの特徴をまとめると、

①入力手順が柔軟である。

②検索不可能な情報を入力しないように入力の誘導をシステムが行う。

（但し、最初に検索したいと思ったものと違うものになる可能性もある）

③マウスによる入力が多いため、キーボードによる操作と比べると、操作が簡単である。

(2つのボタン押すことと、マウスの位置を変えることにより操作を行う)

- ④国名、品名等のコード表を読む必要がない。
- ⑤データベースのコマンドを意識せずに複数のデータベースを使うことができる。
- ⑥マニュアルを読む必要がない。
- ⑦データの用途に側したメニューを用意しているので、非常識な検索指示をさけることができる。

5. 3 画面遷移と知識ベースとの関連

本節では、実際の検索サンプルを通して、どのように知識ベースが使われているかを明らかにし、知識ベースの利用の効果を確認するとともに、知識ベース構造の妥当性や、拡張性を例を示し考察する。

例) '日本から米国へのテレビの輸出額をドルベースで知りたい' という検索要求の場合

① 報告国をクリックしてそのまま改行キーを押す (デフォルトが日本なので)

検索可能なデータベースは(db1/db2/db3)				
検索 輸入額 最新値 全品目	の に関して を について知りたい	全世界	に対する	

☆画面遷移 → 報告国入力枠の色が変わる (入力確定)

☆知識ベース → DB属性辞書のなかの属性の定義域に日本を含まないデータベースはないので、検索可能なデータベースは、変わらない

(注) 以下データベース名は、日本貿易統計=db1、各国間貿易統計=db2、世界経済統計=db3 と略記する。

② 相手国をクリックして「米国」と入力する

検索可能なデータベースは(db1/db2/db3)				
日本 輸入額 最新値 全品目	の に関して を について知りたい	米国	に対する	

☆画面遷移 → 相手国入力枠の色が変わる（入力確定）、米国と入力したものがアメリカに書き換えられる。

検索可能なデータベースが db1,db2 になる。

☆知識ベース→ 類義語辞書により「米国」の類義語としてアメリカを選択し、画面に表示する。

DB属性辞書のdb3 は属性として相手国を持たないので、相手国に関する指示を含む検索は不可能である。

③ 品目をクリックして「テレビ」と入力する

検索可能なデータベースは(db1/db2)			
品目	の	アメリカ	に対する
輸入額	に関して		
最新値	を		
テレビ	について知りたい		

☆画面遷移 → 品目入力枠の色が変わる（入力確定）。テレビと入力したものがテレビ受像機に書き換えられる。

検索可能なデータベースが db1のみになる。

☆知識ベース→ 類義語辞書により「テレビ」の類義語として「テレビ受像機」を選択し画面に表示する。

DB属性辞書のdb2 は属性として品目を持たないので、品目に関する指示を含む検索は不可能である。

④事項入力枠をクリックしてメニューから輸出額を選択する

検索可能なデータベースは(db1)

日本 の 経済力 に対する
 最新値 に関して
 システム受像機 を について知りたい

事項	説明
輸出額 輸入額	

メニュー画面からマウスで選択して下さい

☆画面遷移 → メッセージ表示域に入力指示が表示される。

事項入力枠の色が変わり輸出と表示される (入力確定)。

☆知識ベース→ 事項メニュー辞書のなかから db1 が収録するもののみをDB収録辞書より選択して画面に表示する。

⑤状態入力枠をクリックしてメニューから最新値を選択する

検索可能なデータベースは(db1)

日本 の 経済力 に対する
 輸出額 に関して
 システム受像機 を について知りたい

状態	説明
最新値 推移 前年同期比 GNP比	

メニュー画面からマウスで選択して下さい

☆画面遷移 → メッセージ表示域に入力指示が表示される。

状態入力枠の色が変わり最新値と表示される (入力確定)。

☆知識ベース→ 状態メニュー辞書のなかから db1 が収録するもののみをDB収録辞書より選択して画面に表示する。

⑥条件入力画面で単位をドルに変換

検索可能なデータベースは(db1)				
日本 輸出額 最新値 資料も集録済	の に関して を について知りたい	アメリカ	に対する	
輸出額を (円/ドル) ベースで	(年/年度) の (年/四半期/月) 毎に 表示します。			
修正の必要な枠をクリックして下さい。修正後入力確定を・・・				

☆画面遷移 → 条件入力画面が表示される。

メッセージ表示域に入力指示が表示される。

マウスの左クリックにより円とドルが交互に表示される。

マウスの右クリックにより入力確定。

☆知識ベース→ DB属性辞書より db1 の最新時点、期種、単位をデフォルトとして
条件入力画面にセットする。

⑦検索コマンドがメッセージ表示域に表示される

検索可能なデータベースは(db1)				
日本 輸出額 最新値 資料も集録済	の に関して を について知りたい	アメリカ	に対する	
輸出額を (円/ドル) ベースで	(年/年度) の (年/四半期/月) 毎に 表示します。			
retr 308/102/90CY				

☆画面遷移 → メッセージ表示域に検索コマンドが表示される。

☆知識ベース→ db1 用検索コマンド生成部が検索コマンドを生成。

以上のように、本試作システムでは、ユーザ要求を5つの属性によって表現し、類義語辞書や、DB属性辞書も、この属性をもとにして定義されている。このような知識構造であるため、任意の順序での属性入力が可能になっているわけで、今後分野の拡張を考慮する場合に、属性による要求の表現がどこまで可能であるかを検証することが重要である。

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the
the same time, the fact that the
the same time, the fact that the

the same time, the fact that the

第6章 試作システムの評価と課題

第6章 試作システムの評価と課題

試作システムの開発に伴い、いくつかの観点からの評価をおこなった。

本章では、試作システムで実現された機能を説明し、この評価及び、今後の課題を提案する。

システムの対象ユーザを明確にしない状態で開発をおこなったために、評価の焦点が絞りにくい面があるが、“統計や、データベースを詳しく知らないユーザでも検索が可能である”ことをシステムの要件として評価を行った。

評価に際しては、①問題解決能力、②操作性、③メンテナンス性（保守性）の3点を、評価のポイントとした。

さらに、①研究業務、②企画業務、③調査業務（定例業務）、④調査業務（臨時業務）の利用目的を想定した場合の、ユーザニーズに関するアンケート結果をもとに、必要な支援機能を抽出し、IRSの拡張項目と技術的ポイントをまとめた。

6.1 試作システムの概要

(1) 環境

IRS試作システムは下記の環境にインプリメントした

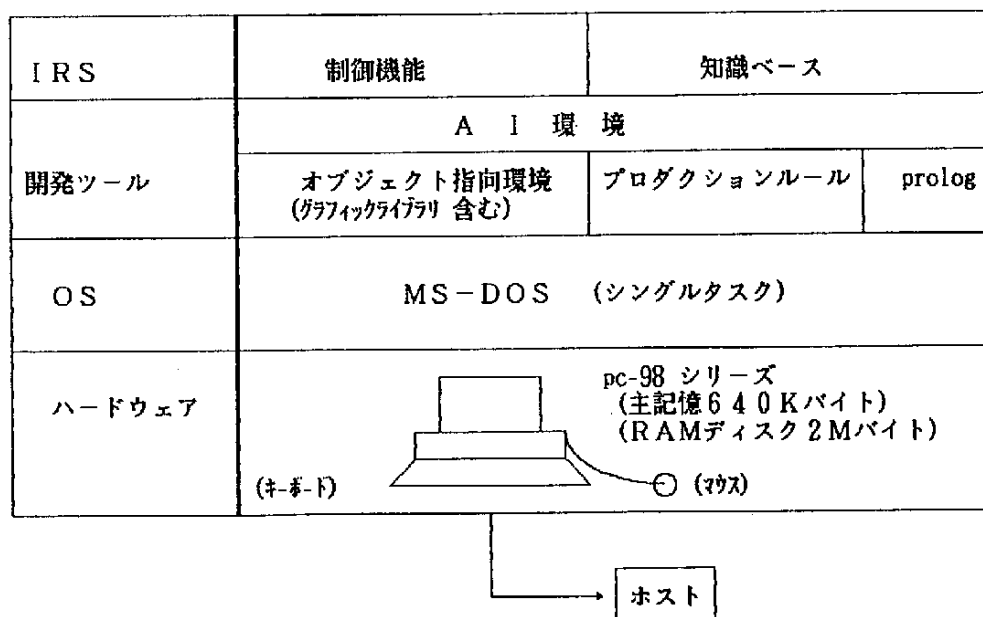
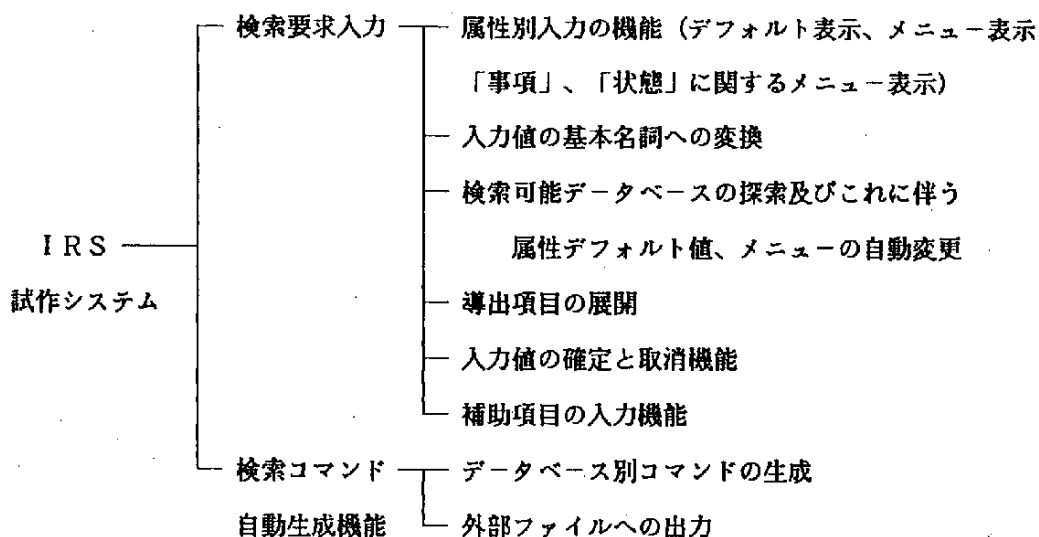


図6-1 試作システムの稼働環境

(2) 機能概要

試作システムで実現した機能は次のようなものである。



(3) 主要知識ベースのメモリ専有量 (概算)

試作システムのコンピュータ上の記憶領域は、表 6-1 のようになった。

表 6-1 主要知識ベースのメモリ専有量

知識ベース名	使用メモリ (KB)
DB 辞書	—
類義語辞書	—
国名	8 (180 か国)
品名	28 (900 語)
導出辞書	—
導出式	—
状態辞書	—
検索コード辞書	—
国名	8
品名	28
系列コード	2

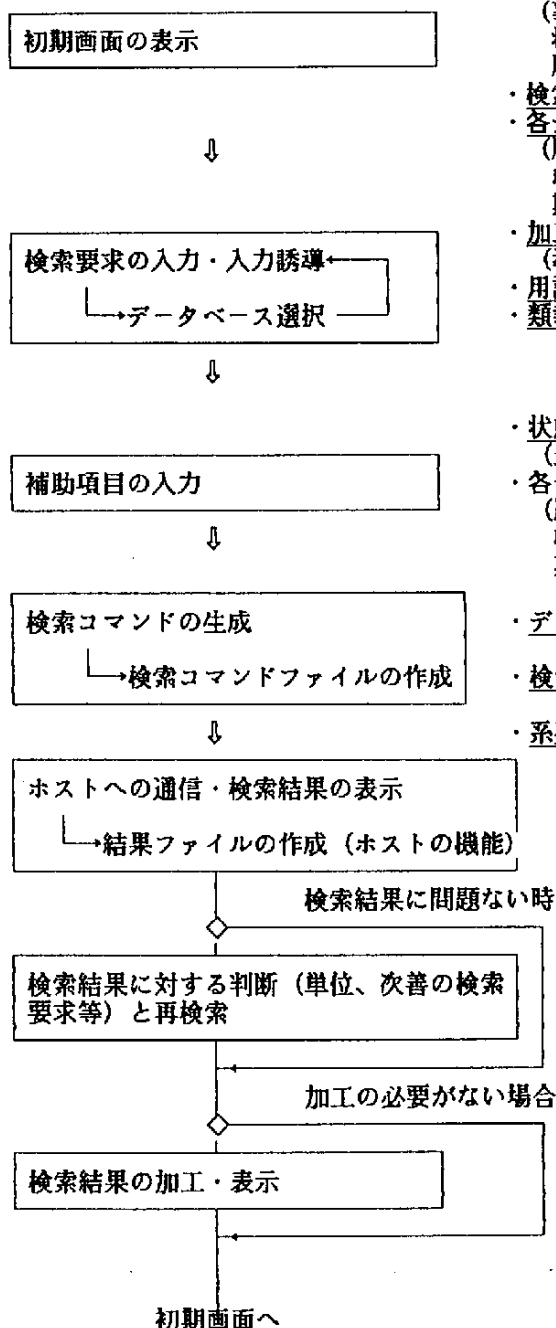
(4) 試作システムの処理と制約

試作システムに関して、処理の流れと、知識との関連、及び現在のシステムの制約をまとめたものが、図 6-2 である。

<処理の流れ>

<必要な知識>

<制約>



- ・ 入力画面表示用データ
(事項のメニューデータ
状態のメニューデータ
属性別のデフォルトデータ)
- ・ 検索可能なデータベースリスト
- ・ 各データベースの収録状況
(属性別の定義域・収録された事項
収録期間・年／年度の区別
期種の収録状況・単位)
- ・ 加工に必要なデータ
(導出辞書・導出式辞書)
- ・ 用語の解説文 (依存度等の説明)
- ・ 類義語辞書
- ・ 状態を統計的に定義するための知識
(最新値・・・最新データ1時点)
- ・ 各データベースの収録状況
(属性別の定義域・収録された事項
収録期間・年／年度の区別
期種の収録状況・単位)
- ・ データベース別検索コマンド生成手順
(検索要求から検索コマンドを生成)
- ・ 検索コードに関する知識
(品目・国名をコードに変換)
- ・ 系列コードの選択に関する知識
(例外コードを使用する国名等)

★属性は貿易統計のフレームをもとに定義してある (他分野でも汎用的なフレームが定義できれば対応は可能)

★細かい注意点等は、含まれない (何年に品目の組替えがあった等の情報)

★概況分類ベースほとんどが基本名詞のみ

★DB別にプログラムで定義している
★新規追加時の矛盾検出等は、未だ考慮していない

・ 検索結果に関する判断基準
↳ まだ作成していない

・ 加工に必要なデータ
(導出辞書・導出式辞書)

図6-2 試作システムの処理の流れと知識の関係

6. 2 試作システムに関する評価

試作システムに関して、①問題解決能力、②操作性、③メンテナンス性（保守性）の3点から評価をおこなった。

(1) 問題解決能力

試作システムでは、貿易統計に関して、日本貿易統計、各国間貿易統計より、検索できる情報及び、世界経済統計より検索できる情報の一部を対象として、ユーザからの入力を受け付ける。

ユーザは、国名や品目名を検索コードではなく、日常使用する言葉で入力することができる。また3種類のデータベースの選択は、システムがおこない、データベース毎の検索コマンド作成もシステムがおこなう。即ち、試作システムは、3種類のデータベースの範囲で、コードの検出と検索コマンドの生成に関して、人間の代行を行う。

また、ユーザは、「最近10年の推移」、「最新値」、「上位10品目」・・・等のメニューのなかから、検索する系列の指定方法を選択することができる。これは、検索コマンドによって情報を選択する場合と比べると、生のデータではなく、意味のある情報の集合としてのデータを取り出すことを可能にしているといえる。即ち試作システムは、'用途を考慮したデータ検索'の機能を一部実現している。

また、データベースに直接収録されていない用語についても導出辞書により、データベースに収録された用語を導出して、検索をおこない、演算をおこなって、回答をだすことができる。即ち有益な指標等の算出に関して、人間の代行をおこなうことができる。

但し、情報に関する判断に関しては、まだ十分な機能を持つとは言えない。「複数のデータを比較して、この関連や、影響を判断する」また、「同じデータを定期的に検索して、その変化を判断し、必要に応じて警告を出す」等は、要望の高い支援項目であり、今後の課題であるといえる。

(2) 操作性

試作システムは、ユーザが5種類の基本属性の入力と、3から5項目（データベースによって項目数が異なる）の条件を入力することによってシステムが自動的に検索コマンドを生成し、検索を実行するようになっている。

ユーザは、システムの表示するメッセージに従って、入力を行っていけば、どのデータベースになにか収録されているかを気にせずに情報を検索できる。

操作は、殆どマウスによる選択で行うが、国名と品目名に関しては、キーボード入力で

あり、操作に慣れないユーザは、マウスの操作、キーボードの操作、かな漢字変換の方法等を習得するまで、若干の煩雑さを感じると思われる。

また、「一旦入力したデータを取り消したい場合は、始めから入力をやり直さなければならない」、「検索に成功した後、検索に要した手続きを保存する方法が示されていない」等の不都合な点がある。

問題解決能力にも係わることとして、「入力誘導に関する問題」をどのように解決するかが、比較的大きな課題であると思われる。

即ち、試作システムの処理方式上、ユーザが、一件データを入力するごとに、システムが、検索できるデータベースを絞りこみ、メニュー画面や、デフォルト表示の変更をおこなう。これは、検索の失敗をなくすために有益な方式ではあるが、ユーザが最初に検索したいと思った検索要求が、入力の過程で、表現不可能になる場合が発生する。

この結果検索できたデータが、当初ユーザが欲していたデータに最も近いものである補償はなく、ユーザが不満を感じることもある。

この問題を解決するためには、改善情報のありかたを再度検討し、入力方式を含めて、ユーザにどのように提供するかを決定する必要がある。

(3) 知識の拡充及び保守

試作システムでは、知識をプログラムと切り離して、知識ベースとして、蓄えることには、注力したが、知識ベースの拡充が、保守に関しては、検討が充分におこなわれているとはいえない。

現在の知識ベースのなかで、品目に関する知識ベースは、もっとも保守の困難な知識である。新しい商品への対応、新しい呼び名が普及した時の登録、使う人によって、同じ呼び名が指し示すものが異なる場合への対応、また膨大な類義語、基本語の集合のなかから、ユーザの求めるものを高速に、的確に検出する機構等、今後の課題が多い。

また、試作システムでは、知識ベースは、開発者が収集して入力したものを使っているが、特に類義語辞書や状態辞書を実用的なものにしてゆくために、知識の自動獲得の仕組みを作っておくことが重要である。これは、人工知能技術のなかで、「学習」と呼ばれる領域に属するものであるが、検索行動を極力妨げないように、知識獲得を行うような機能が必要である。

6. 3 今後の検討課題

次年度の調査研究のための検討課題を抽出するために、検索支援システムの利用目的によって、どのような支援機能が求められているかを調べ、そこから得られた項目について技術的ポイントをまとめた。

6.3.1 利用目的と必要な機能

通産ジャーナルより抽出した100数件の事例については、前述のように分析をおこない概念モデルの作成と試作システムに反映させたが、IRS拡張の方向性を考慮するとユーザの検索目的によって、必要な支援の方法が異なると考えられる。

そこで、本調査研究では、IRSの今後の検討課題を抽出するために、実際にユーザへのサービスを行っている部門に対してユーザが業務に応じて、どのようなデータベースの使い方をしているか、専門家に支援して欲しいのは、どのような場合かをアンケート方式で調査した。(付録6)

本節では、この調査に関する結果を説明する。

(1) ユーザの利用目的について

統計データベースの利用目的として下記の4種類が挙げられた

①研究業務担当者

本格的な検索を行う前に、データの性格に関するチェックや、加工結果に関するパイロットスタディを行い、満足のいく結果が得られてから本格的な検索に入る。加工は、SAS等の汎用パッケージを使っている。

②企画業務担当者

主張したいこと、意見等の裏付けのための資料や、背景を説明するための資料作成を目的として、検索を行う場合が多い。

③調査業務担当者(定例業務)

定期的な報告書等の作成のために、決まった手順で、検索を行う。

前回と比べて、データがどのように、変化したか・・・に興味がある。

④調査業務担当者(臨時業務)

与えられたテーマに関する調査なので、データベースの分類体系とマッチしない場合が多い。データの出典や性質等に関する情報が必要である。試行錯誤的に検索を行い、テーマにあったものを選択してゆく。

(2) ユーザの利用目的と必要な支援機能

利用目的別に必要な支援機能をまとめると下記の各項目があげられた

①研究業務担当者

データ特性に関する情報提供（メタデータ）

検索データの収録情報

検索データのダウンロード

多種（多分野に渡る）データの結合

条件文の保存・再利用

②企画業務担当者

加工結果の評価

③調査業務担当者（定例業務）

処理手順の単純化（登録等）

データ更新等の監視

加工結果の変化の監視

結果出力の編集

④調査業務担当者（臨時業務）

データの収録案内

検索案内

複数データベースからの結果の結合における注意

データ比較の妥当性指示

またすべての利用目的に共通の支援課題として、試行錯誤的な使用に耐えるシステム化、多種（多分野）のデータの結合に対する支援があげられた。

このアンケートにより、最終的に利用目的を絞るという結論はでなかったが、IRSの拡張項目をその背景と共に捉えることができた。次節で、それらの項目について説明する。

6.3.2 拡張項目と技術的ポイント

試作システムの評価を踏まえて表6-2のような拡張の方向性がまとめられた。

拡張の項目としては、①知識ベースの拡充、②次善情報の提供機能の追加、③分野の拡張可能性の検討、④要求指定方法の拡張、⑤検索可能なデータベースの追加、⑥データの加工、⑦検索履歴の保持の7項目が挙げられた。

(1) 知識ベースの拡充

試作システムが十分な問題解決能力を持つためには、現在ある知識ベースをさらに拡充し、専門家の知識、マニュアルの収録内容等に近づける必要がある。

長期間の使用に耐える質の高い知識ベースを作成するためには、知識獲得をおこなって知識量を増やすとともに、運用時に知識を追加更新するための機能や、整合性をチェックするための機能を工夫することが重要である。

これらの機能を検討する際に、各知識ベースについて、内容に関する管理責任者が明確になっていなければならない。管理責任者としては、データベース管理者、統計専門家、ユーザの3者が考えられ、類義語辞書はユーザ、導出辞書は統計専門家、検索コマンドに関する知識はデータベース管理者というような、役割分担が必要であろう。これによって、ユーザが責任を持つ知識に関しては、利用時に知識獲得をおこなう。データベース管理者が責任を持つ知識に関しては、データベースの利用方法に変更のあった時にメンテナンスをおこなう等の管理方法を設定し、そこで必要なインターフェイスを設計しなければならない。

(2) 次善情報の提供機能の追加

試作システムでは、ユーザの入力に対して、システム側が、その時点での検索可能なデータベースをチェックし、常にユーザが検索可能なものを選択するよう、メニューやデフォルトを変更していた。

ユーザが当初イメージした要求に最も近いものを選択するためには、一旦ユーザの要求をすべて入力させてから、この完結したユーザ要求に対して、次善の情報を選択すべきであろう。

次善情報の定義は、明確に定義されていないが、該当する品目が収録されていない場合

は、その品目を含む上位分類名で代替する。GNPが収録されていないときは、GDPを代わりに使う等。何通りかの「次善情報」の定義を、定式化することが可能であると思われる。

また、次善情報が必要な場合としては、検索を行うまえに、データがないことが判る場合と、検索実行後に欠測等で、検索に失敗する場合の2つのケースが考えられる。それぞれに対して、どのようなシステムの対応が望ましいかを検討する必要がある。

(3) 分野の拡張可能性の検討

試作システムでは、貿易統計の範囲での、検索支援機能を実現したが、貿易統計以外の分野に適応可能であるか否かは、補償されていない。

実際に新に分野を追加することは、先の話であるにしても、同じ手法が、別の分野に関しても適用可能であることを、検証しておくべきであろう。即ち、別の分野の検索要求は、貿易統計と同様のフレーム形式で表現することができるか。検索に必要な各辞書に相当する知識が存在するか。また複数の分野が共存する場合に共有する知識と固有の知識は、どのように、区別しておけばよいか。また複数の分野を同時に検索領域としたい場合にどのような仕組みにしておくのが好都合か等が検討課題になる。

(4) 要求指定方法の拡張

試作システムでは、検索したい要求を属性別に入力するが、文献等の分析事例等から検索要求が発生する場合、実績のある検索要求そのものをメニュー選択できるような仕組みが有効である。

また、複数の検索結果を同時に画面上に表示させたいという要望は、利用形態によらず多くのユーザから出されており、これに対する対応も、要求指定方法の拡張に含まれる。

(5) 検索可能なデータベースの追加

試作システムは、3種類のデータベースを検索対象としたが、データベースの種類を増やし、貿易に関する情報提供を充実させる・・・という拡張の方向がある。同じデータが複数のデータベースに収録されている場合の、選択のしかた等が問題となってくると思われる。また数値データベース以外に、文献データベースを追加すると、特に、入出力方式等に関して数値データと文献データの混在させかたを十分に検討する必要がある。

(6) データの加工

報告書作成を最終目的とするユーザにとっては、グラフ作成や表作成の機能と、データの分析に関する機能が重要である。

表やグラフの作成は、当面、データをダウンロードして、市販の作表、グラフィックソフトウェアで、加工する方法が考えられる。検索後のデータの分析は、今のところ、人間にとっても判断の難しいものが多いと思われ、本年度の調査研究では、まだ十分な調査はできていない。

(7) 検索履歴の保持

同じ内容の検索を定期的に行いたいユーザにとって、一度入力した検索要求が、次回から簡単に指定できるような機能が有効である。作成された検索コマンドファイルを保存するのが最も簡単な方法であるが、さらに高次の支援としては、そのユーザの検索傾向や使用する用語等を蓄積して、頻度の高いものがデフォルトになるような誘導をおこなう等の拡張が考えられる。

以上の項目のなかで、次年度の検討課題は、知識ベースの拡充と分野の拡張可能性の検討が優先課題である。

表6-2 拡張の方向性

項目	技術的ポイント	備考
知識ベースの充実 類義語辞書の充実 状態辞書の充実 導出辞書の充実 説明文の充実（データの属性等） 知識の追加更新機能の充実 知識の整合性チェック	☆類義語の収集 ☆基本名詞への変換速度の維持 ☆品目分類の体系化（既存の分類体系ではなくユーザオリエンテッドな分類体系を考慮する必要がある） ☆経済白書等を使って、状態として登録すべきテクニカルタームを抽出する ☆テクニカルタームを検索条件に展開して、知識ベースを作成する ☆経済白書等を使って、導出辞書に登録すべき（因果関係によって導かれる）テクニカルタームを抽出する ☆データの属性として、ユーザに提示すべき情報の収集、整理 ☆知識ベース追加更新用インターフェイスの仕様検討 ☆知識の矛盾検出のしくみ（知識間の関連性を考慮する必要がある）	◇類義語辞書は、大規模になればホストに移行も検討 ◇状態辞書と導出辞書の切り分け方は、再検討の余地がある
次善情報の提供機能の追加	☆品目分類、情報の因果関係等から次善情報を選択するための情報尺度を定義する ☆上記尺度を使って再検索を行う仕組みを作る	
分野の拡張可能性（P P I S 関連分野） 国際経済（財政・金融） 国内経済（マクロ経済・産業・企業活動・物価・労働・産業連関表等） 資源情報、地域別経済情報、公害情報 企業財務、商品市況、法令情報等	☆分野別フレームの検討 ☆分野間のメタ知識の抽出と知識ベース化（品目分類と産業分類の関連に関する知識等） ☆複合分野検索のためのインターフェイス検討 ☆知識ベースの高速逆引き機能（例えば貿易→自動車→自動車産業→産業）	
要求指定方法の拡張 過去の事例により指定 フラットメニューによる指定 複数の検索要求入力	☆事例の収集と知識ベース化 ☆検索手順の知識ベース化 ☆入力インターフェイスの検討	
検索可能なデータベースの追加（貿易統計の） 数値情報データベースの追加 文献情報データベースの追加	☆検索コマンドの追加 ☆キーワード検索・キーワード合成等 ☆数値情報と文書情報の併用システムとしての要件抽出（入力フレームの共有等）	◇文献検索は、本質的には別システム
データの加工 統計パッケージとのリンク 統計分析手法のアドバイス	☆データのダウンロードと統計ソフトウェアファイルへの変換 ☆統計分析のノウハウ収集と知識ベース化	
検索履歴の保持 前回入力値を次回デフォルト値として利用 メニュー・辞書の並べ換え（前回使用したものを先頭に置く） 前回の検索要求をメニューに追加	☆履歴ファイルの作成（ユーザ別カルテ） ☆知識ベースのメンテナンス機構 同上	

第7章 ヨーロッパの統計DBユーザ・

インタフェースの研究動向と

IRSの拡張性

第7章 ヨーロッパの統計DBユーザ・インタフェースの研究動向とIRSの拡張性

本章は、当調査の一環として行われた、海外調査の報告をまとめたものである。この海外調査は、統計DBシステムへの人工知能・オブジェクト指向DBの応用の現状を把握し、当調査のシステムの今後の発展に資することを目的としたもので、1990年1月8日から19日にかけて、スウェーデン、フィンランド、イタリアの3カ国の統計局、研究機関を訪問し、先端的な研究開発動向を取材すると共に、研究者らと研究の方向性について討議を行った。

取材・討議した内容は、運用中の統計DBシステムとその機能拡張の方向、統計データモデル、統計データ操作概念、統計DBユーザ・インタフェースなど多岐にわたるが、本章では、これらの内、当調査のシステムIRSと密接に関係する統計DBのユーザ・インタフェースの研究・開発状況に的を絞って報告する。

社会経済統計DBは、種々の集計レベルをもち、互いに関連をもつ多種のデータで構成されており、個々のユーザがその全体像を把握し、必要なデータを正しく検索することは難しい。そこで、ユーザの統計データ検索を支援するためのユーザ・インタフェースが、統計DB研究の主要なテーマの1つとして研究・開発されてきた。これまでに提案された新しいタイプのユーザ・インタフェースは、おおむね以下の3種に分けられる。

- ① 自然言語インタフェース
- ② グラフ型の統計データモデルに基づくインタフェース
- ③ 対象中心の統計データモデルに基づくインタフェース

以下では、上述の海外調査の結果を踏まえつつ、これら3種のユーザ・インタフェースの研究開発状況を説明し、最後にIRSの拡張方向との関係に言及する。

7.1 自然言語インタフェース

7.1.1 フィンランド中央統計局のTSEXSY

自然言語インタフェースは、DBの内容についてユーザの特別な予備知識を前提とせず済む。このため、人工知能研究の応用として自然言語解析が容易になるにつれ、一般のDBの世界では、自然言語インタフェースを提供するDBMSが数多く登場してきている。しかしながら、これを統計DBのように、データの種類が多く、かつ、多様な加工方法が

存在するデータに適用しようとする、1つのジレンマが発生する。

それは、ユーザが予備知識を持たずに検索要求を出せるためには、ユーザが使いそうな構文の全てを予め登録しておかなければならないということである。これは、データの対象世界が広いとき、言うは易く、行うのは難しい。このため実際には、使える構文のある程度限定せざるを得なくなり、構文に関するユーザの予備知識を前提とせざるを得なくなる。

このジレンマに対し、フィンランド中央統計局の Janhunen らが開発したシステムは、1つの回避策を示している。それは、厳密な構文解析は行わず、単なるキーワード・マッチングだけで、自然言語検索文を解釈しようとするものである。このシステムは、中央統計局がもつ13、000も経済時系列を対象に自然言語のデータ要求文の入力を受け付け、対応するデータを提供するものである。これは、TSEXSY(Time Series Expert System)と呼ばれており、現在ベータ・テスト版が提供されているところである。

このシステムでは、例えば、

「1978年～1988年のサービス業の時間当たり平均賃金の実質変化率は？」

という問に対して、文を構成する単語を以下のように種類分けし、対応するコマンドを生成する。

- ① 主たるデータは何か、... この例では、「賃金」
- ② それはいかに特定化されるか、... 「サービス業」「時間当たり」「平均」
- ③ 期間はいつか、... 「1978年～1988年」
- ④ どう加工するか、... 「実質」「変化率」

このシステムは全て、IBM PC上で作動するAPLで書かれており、生成されるコマンドもAPLの合成関数の形をとる。例えば、

「全消費支出に占める個人消費支出の割合」

という入力文は、以下のAPL関数に変換される。

ただし、50010、50008はそれぞれ、個人消費支出、全消費支出の時系列コードである。

7.1.2 TSEXSYの評価

フィンランド中央統計局で実際にテスト使用されているシステムは、フィンランド語を用いるものである。著者が見せてもらったものは、英語のデモ版で、こちらにはごくわずかしかな単語が登録されておらず、その能力を十分に知ることはできなかった。このデモと担当者から聞きだしたことをもとに評価すると、その長所は以下のようにいえそうである。

① どのような入力文に対しても、一応答えることができる。

これは厳密な構文解析を行わず、また分からない単語は無視するという方法をとっているためである。実際、語順を一部入れ替えて、文法的には正しくない文を入力しても、同じ答えが返ってきた。

② システム構築は比較的容易である。

知識の拡充は、(a)単語の定義とそのAPL関数への対応付け、(b)同義語の定義、の2種のみであるようで、問い合わせ画面の中から知識メンテナンス画面を呼びだせるようになっている。

以上の長所と裏腹の関係になるが、その短所は以下のようにいえる。

① 複雑な問に正しく答えられない。

これは、構文解析を行わないことの当然の帰結である。例えば、先の例で、「平均」は「特定化」の単語に分類されており、「加工」の単語とみなされていない。従って、「業種別賃金の平均」を求めることは難しいと予測される。

② 対話型インタフェースにならない。

TSEXSYは構文依存でないといっているが、前節で述べたように使用単語を類別して解釈している。このことは、単一の構文のみを認めていることと実質的に同値であり、

データ問い合わせ以外の入力文を認めるようにシステムを拡張することは困難である。つまり、システムから疑問点をただす質問をユーザに出して答えさせるといった、自然言語インタフェースに特徴的な対話を実現することは難しい。

③ 意図せざる結果が得られることがある。

これは、構文解析しないことに加え、未知の単語は無視するというアルゴリズムの結果である。

これらの欠点は、開発者の承知しており、その対策として、advisorなるものを用意している。これは、いわゆる助言システムではなく、推論プロセスを記録したログ情報、及び、データに関する注意事項などのメタデータの表示機能であり、これを見ることにより、導出されたデータが意図したものと同じか否か確認することができる。しかし、著者の印象では、ログ情報よりも、システムが理解した内容をもとに問い合わせ文を言い直す文章生成の方が、利用者には分かりやすいと思われる。

7. 2 グラフ型ユーザ・インタフェース

7.2.1 I A S I の G R A S S

統計データの構造、及び、データ相互間の関係を有向グラフで表現するモデルである S U B J E C T は、カリフォルニア大学の Shoshani 教授らにより 1981 年に提案されたものである。その後、イタリアの国立研究機関 I A S I (情報分析システム研究所) の R a f a n e l l i 教授、ローマ大学の B a t t i s t a 教授らにより、その精密化が行われている。

ここでは、I A S I の R a f a n e l l i、R i c c i らが開発したグラフ型ユーザ・インタフェースを紹介する。これは、上記 S U B J E C T の拡張であるグラフ型統計データモデル G R A S S の有向グラフ表現を、そのままユーザ・インタフェースとしたもので、Macintosh 上で作動する P a s c a l O b j e c t を用いてインプリメントされている。

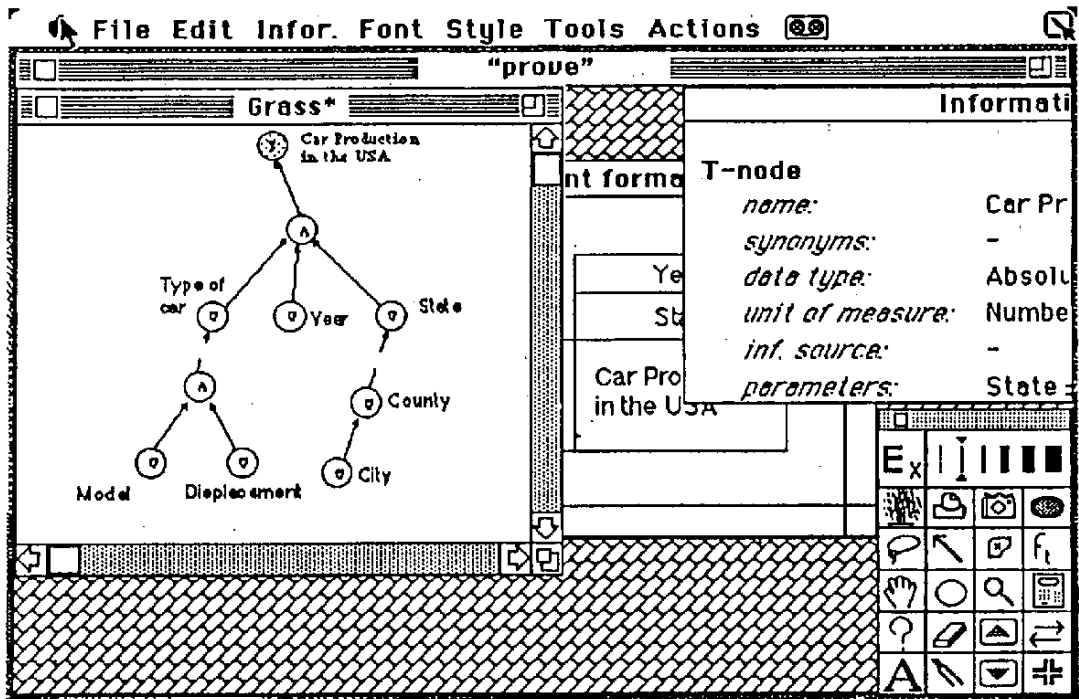


図7-1 GRASSのユーザ・インタフェース

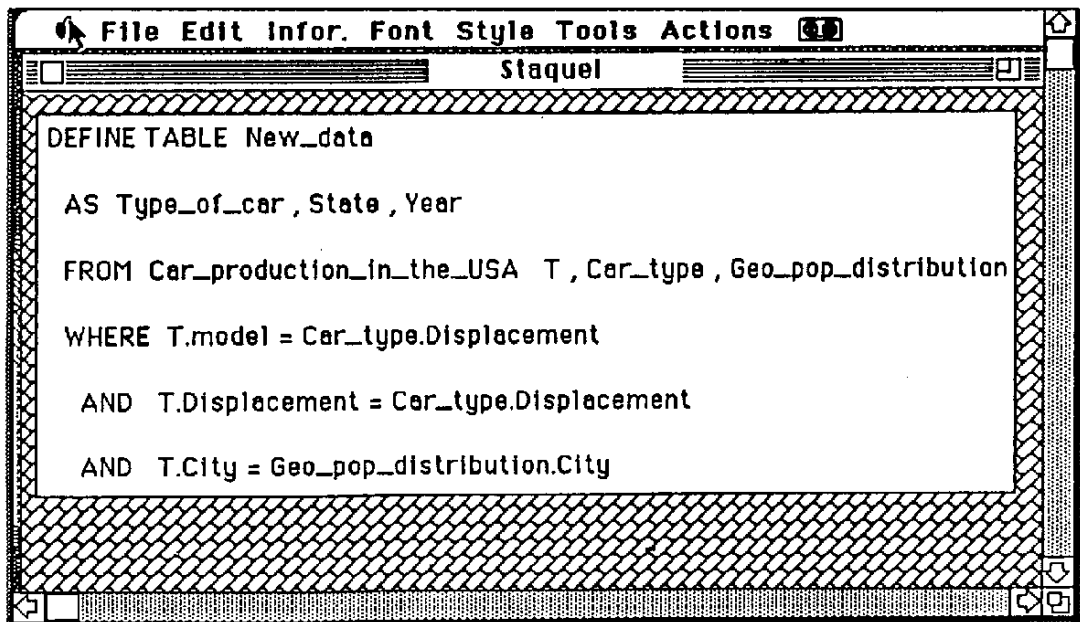


図7-2 GRASSで生成されたDB検索コマンド

図7-1はその検索画面のハードコピーである。左の副画面は、GRASSモデルに従った統計データの記述であり、「自動車の型別、年別、州別のアメリカの自動車生産台数」(Tノード)を表している。また、「自動車の型は、モデルと排気量(displacement)の組合せで与えられること」(Aノード)、及び「州(State)は、County 更に Cityに細分化できること」(Cノード)を表している。

右上の副画面は現在カーソルがおかれているノード(今のケースでは、Tノード:アメリカの自動車生産)に関する詳しい情報を表示している。

この図7-1は、自動車生産台数というただ1つの統計属性をもつ統計表を表現しているが、GRASSでは同様にして、多種の統計表とそれらの間の関係を表示することができる。また、いくつかの統計属性ノードを集めたノードをトピック(主題)と呼び、トピックを選ぶことにより、スコープを限定することができる。利用者がいくつかの統計属性ノードを指定して、独自のトピックを定義することも可能である。

なお、これらのノードの種類の違いは、画面では色で区別されており、実際は図7-1より分かりやすい。

画面には、図7-1のようにノードを名称つきで表示する詳細画面(ズームイン)とノード表示を省略したアウトライン画面(ズームアウト)が存在する。

このユーザ・インタフェースを用いた統計データ検索は概略以下のようになる。

- ① アウトライン画面で必要なデータがありそうな部分に見当をつけてズームインする。
- ② 詳細画面上でデータ検索に関する細かい指定をする。詳細画面上では、ノードをマウスで選び、そのノードに適用可能なオペレーション(選択、集計、マージなど)を選択することにより、問い合わせの指示を行う。なお、適用可能なオペレーションはノードの選択によって現れるポップアップ・メニューとして表示される。
- ③ ②の指示に基づいて、図7-2に示されるようなDBMSへの検索コマンドが生成される。

7.2.2 GRASSインタフェースの評価

GRASSの長所は、ビジュアルで分かりやすいことである。共通の分類を持つ統計属性は、同じ分類ノードに矢印でリンクしており、統計属性間の関係を見ながら検索を行う

ことができる。このリンクを利用して、例えば、同じ商品分類で分類されている統計属性だけを表示せよ、といった問い合わせは、商品分類のノードを選択することにより、容易に実行できる。

また、特定のカテゴリの選択（例えば、商品カテゴリ「自動車」の選択）や中分類への集計の指示も、ノードを選択したときに表示されるポップアップ・メニュー上で、容易に指示できる。

他方、欠点としては、多種データの記述には不向きということができよう。実際に見せてもらったデモ・システムでは、統計属性、分類属性のノードが30～50程度の小規模のものであったが、詳細画面では1度に表示できる属性ノード部分は10～15ノード程度である。また、アウトライン画面は、この程度でも蜘蛛の巣のようになり、よほど慣れていないと何か表されているのか分からない。

なお、現在のGRASSモデルは、より原理的な問題点をかかえている。それは、統計間の関係といっても、正確に分類が一致する場合のリンク関係しか表現できないことである。国際標準商品分類と日本標準商品分類のように、粗く対応関係はつけられるが、1対1にカテゴリが対応するわけではない分類が存在するとき、それぞれの分類を持つ統計間の関係を表現することはできない。同じ理由から、途中で分類変更のあった時系列統計の表現も困難である。

この欠点は、Rafanelliらも気がついており、彼等の新しい理論モデル（STORM）と提案しているところである。そこでは、一般的に「商品分類」と呼ばれるもののように、個々のカテゴリを特定化できない抽象的な分類の概念と、「〇年に制定された国際標準商品分類」のように、個々のカテゴリを列挙できる分類の概念を区別するようになった。前者を内包的な(intensional)な分類、後者を外延的な(extensional)な分類という。

彼等は、この新しいモデルに基づくユーザ・インタフェースもGRASSと同様にビジュアル化することを構想している。しかし、内包と外延の区別により、ノードの数は更に増大し、リンク関係も複雑になるので、実用的なものが作れるかどうか疑問である。

7. 3 対象中心アプローチのユーザ・インタフェース

スウェーデンは中央集権的な統計収集体制をとる国で、最も統計DBシステムの体系化

が進んでいる国である。その中央統計局は、先端的な統計DBシステムの開発・研究で知られており、そのシステムは、ヨーロッパ経済委員会、中国、ザンビアをはじめ多数の国や国際機関で採用されている。そこで特徴的な点は、彼等はDBシステムのソフトウェアを単に持ち込むだけでなく、同時に統計DBを設計するための考え方（統計データの概念モデル）の普及に努めていることである。そのモデルを、OPRモデル(Object-Property-Relationship model)という。このモデルは、彼等に固有なものではなく、一般のDBの世界で意味データモデルと呼ばれている対象中心のデータモデルと同じものであり、一般のデータモデルの概念を用いて統計データを整理する事が重要であることを主張しているのである。（注）

（注）オブジェクト指向と対象中心アプローチ

本節で対象中心と呼んでいるアプローチは、オブジェクト指向と混同されることが多い。オブジェクト指向では、共通のオペレーション（データ操作）が適用できる対象の集まりを1つのクラス（object class, object type）とみるのに対し、対象中心では、共通の性質を持つ対象の集まりを1つのクラスとみる。物理的な対象に関するデータでは、オペレーションの違いはデータ対象の性質の違いに起因することが多いため、オブジェクト指向は、対象中心アプローチを内を含むアプローチと考えることができる。他方、統計データの場合、統計データ操作の適用可能性と対象の性質との関係が、一部の例外を除き、明かでなく、オペレーションの共通性はあまり強く意識されていない。そこで、ここでは、オブジェクト指向と呼ぶずに、対象中心という用語を用いた。

なお、統計データ操作と対象の性質との関係が明白な例としては、時系列データがある。これは、時点という共通の性質をもつデータであり、この性質故に伸び率等の時系列演算が適用可能になる。

OPRモデルの基本的な考え方は、著者らが国土庁の統計DB検索プロトタイプ・システムLIDS86の開発に際し提案した統計データモデルSDS4Sと極めて類似しており、簡単に言えば以下のように要約できる。

- ① 共通の性質を持つ対象を抽象化したものを対象型（object type）という。
- ② データは、対象型毎にその性質を記述したものとして、まとめることが望ましい。

例えば、通関統計も各国間貿易統計も報告国、相手国、輸出額、輸入額という共通の性質を持っており、同じ「輸出入」という取引を表す対象型に関するデータであると考えることができる。これらが区別されるのは、先に述べた外延の違いであり、データの収集方法、対象範囲が異なるからである。また、世界経済統計、各国間貿易統計、通関統計が貿易統計として一括できるのは、これらが報告国別という共通の性質をもつからである（図7-3）。

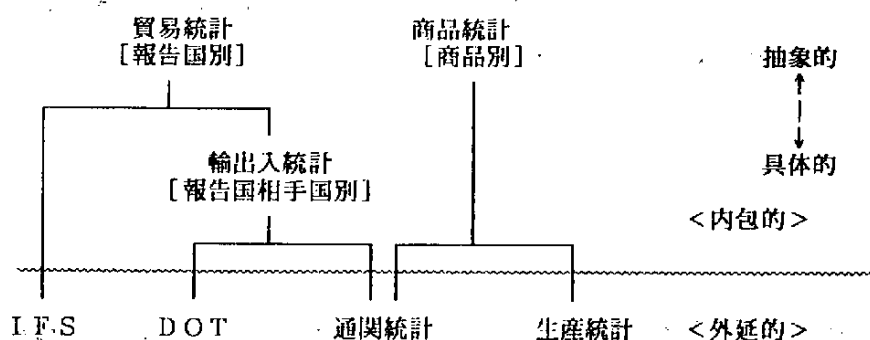


図 7-3 対象中心アプローチにおける統計データの汎化階層

先の Rafanelli の GRASS モデルでは、統計属性をまとめる抽象概念は、アドホックなトピックという概念でしかなかったが、この対象中心アプローチの下では、上述のように共通する性質に着目することで、多種多様な統計データを抽象度の高いものから具体的なものと汎化階層に整理することができる。

現在のところスウェーデンの統計 DB は、この OPR モデルに立脚したものではなく、OPR モデルは、単に統計 DB 設計の指針としての役割しか果たしていない。1988 年の国際統計 DB 学会で、スウェーデンの Malmberg は、このモデルに基づくユーザ・インタフェースの構築の計画を報告した。それは、データの抽象度に応じた階層関係を利用して DB 全体をブラウズでき、対象型別に整理された性質を利用して検索要求を特定化していくものであった。これは、国土庁の LIDS 86 のブラウザと良く似ている。LIDS 86 のブラウザは、単にブラウズの機能があるだけで、データ検索機能は別になっている。Malmberg の計画は、ブラウザを Smalltalk のマルチウィンドウを用いて分かりやすくするとともに、検索機能を付加するものであった。

今回のスウェーデン訪問の主たる目的は、この計画の進捗状況を知ることであったが、このプロジェクトは未だスタートしていないということで、この面の収穫はなかった。かわりに、OPR モデルの提唱者である Sundgren 及び 上記の Malmberg と直接会い、彼等のモデル概念が LIDS 86 のデータモデルである SDS 4 S と基本的に同じであることを確認できたことは収穫であった。また、この会合で、対象中心アプローチのユーザ・

インタフェースとして、IRSのテンプレートが有望であるということで、意見の一致をみた。この点については、7.5節で述べる。

7.4 統計ユーザ・インタフェースと3スキーマ

以上、ヨーロッパの3機関における統計DBの新しいユーザ・インタフェースの研究・開発の動向をみてきた。それらは、いずれも統計DBのインタフェースとはいうものの、それぞれ、対象とするデータの範囲、想定している利用目的に差があり、同列に評価することはできない。そこで、これらの相互関係を理解する視点として、一般DBの世界では普遍的に受け入れられている3スキーマの概念を説明し、次いで、その概念を用いて、先の3種のユーザ・インタフェースのそれぞれの位置づけを明らかにしてみよう。

7.4.1 ANSI/SPARCの3スキーマ

3スキーマの概念は、ANSI/SPARCによって1975年に提案されたDBにおけるデータ記述の構成法である。ここで3スキーマとは、以下の3種のスキーマのことを指している。

① 外部スキーマ

特定の利用目的に応じたデータの記述であり、DB内のデータから導出される特定データの記述である。

② 概念スキーマ

特定の利用目的から中立的なデータの記述であり、DB内の全データの記述である。

③ 内部スキーマ

DB内の物理的配置に依存するデータの記述である。

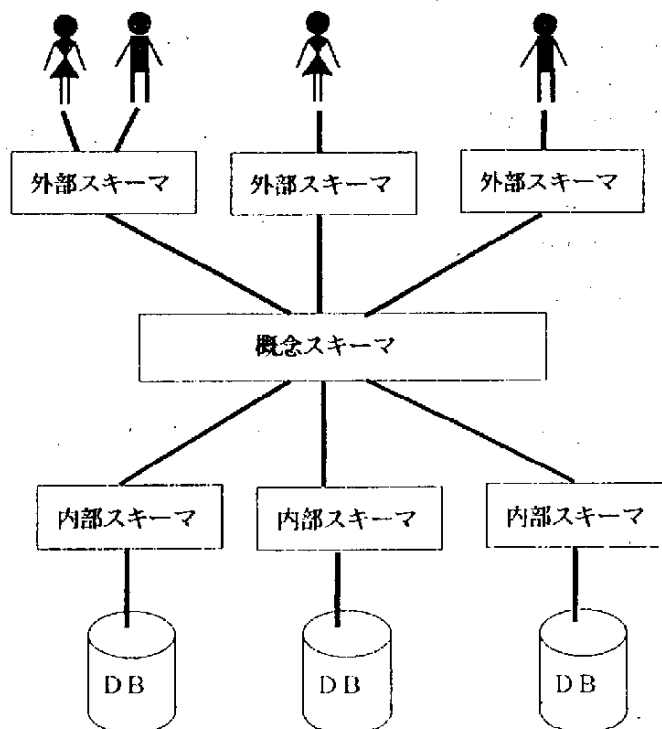


図7-4 3スキーマ

この3スキーマの関係は、図7-4のように書くことができる。3スキーマで構成されたDB上では、以下のようにデータの検索が行われる。利用者のデータ要求は、外部スキーマに対して出される。このデータ要求の情報は、外部スキーマの定義情報をもとに概念スキーマ上の用語に変換される。次いで、概念スキーマ用語のデータ要求は、内部スキーマ定義情報を参照して、物理的なデータ検索命令に変換され、物理DBからデータが検索される。このようにデータ記述を3種のスキーマに分けるのは、以下のように、変更に対する自由度を確保するためである。

外部スキーマを概念スキーマから切り離すことにより、他の利用者のことを考えずに個々の外部スキーマ（データの特定の利用形態）を変更することが可能になる。つまり、個々の利用目的に応じたデータのカスタマイズが可能になる。これを論理的データ独立性という。また、内部スキーマを概念スキーマから切り離すことにより、利用者に（外部スキーマに）影響を与えることなく、データの物理的配置を変更することが可能になる。これを物理的データ独立性という。

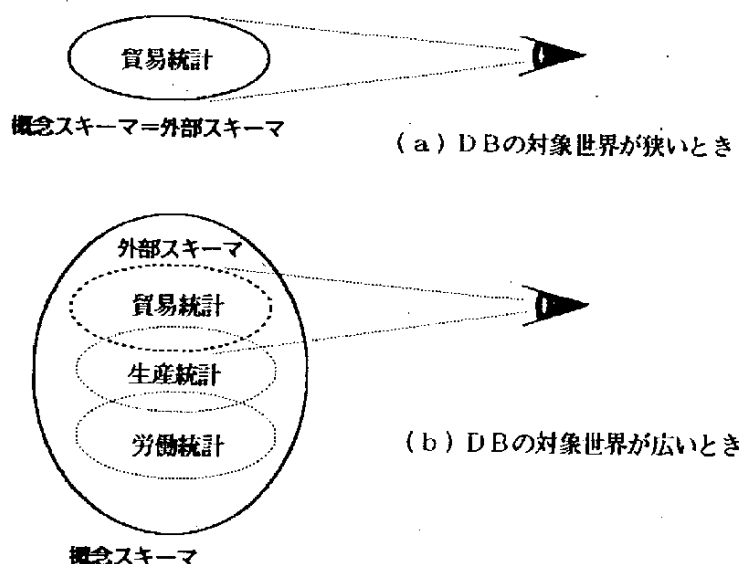


図7-5 概念スキーマと外部スキーマ

3スキーマ概念をここで持ち出したのは、IRSが現在かかえている最大の懸案事項の一つが、他の対象分野への拡張性の確保という点にあるからである。現在のIRSは、貿易統計という特定の利用形態を想定し得る狭い領域だけを対象としている。この範囲内では、概念スキーマと外部スキーマを区別する必要は生じない(図7-5(a))。しかし、他の分野のデータを、従って、別のタイプのデータ利用形態を取り込むように拡張しようとする、利用目的に応じた外部スキーマとは別に、概念スキーマを用意する必要がでてくるのである(図7-5(b))。

7.4.2 各種統計ユーザ・インタフェースの位置づけ

ここでは、上で説明した3スキーマの概念を用いて、7.1～7.3節で紹介した各種統計ユーザ・インタフェースの位置づけを考えてみる。

(1) 自然言語検索

フィンランドのTSEXSYは、入力された単語を、①データの種類、②特定化の条件、③期間、④加工方法、の4種に類別することにより、一意に検索コマンドを生成するものである。つまり、データの利用形態は、ただ1種類に限られており、現在のIRSと同じ

く、概念スキーマと外部スキーマが未分化なシステムであるといえる。

むろん、自然言語検索一般がそうだということではない。7. 1, 2節で述べたように、他の種類の入力文も許すように拡張しようとする、一般的なデータ表現と個々の構文で表現できるデータの記述とを分ける必要がでてくる。

(2) グラフ型ユーザ・インタフェース

グラフ型統計データモデルの祖であるSUBJECTは、もともと、現実の統計書に収録されているような統計表を記述する目的で考案されたものである。つまり、表頭、表側があり、それぞれが多重にネストしている統計表の構造のモデルである。このため、意味的には同じデータでも、表頭、表側が入れ替わると、異なる形で記述されることになる。つまり、物理配置を表現する内部スキーマ、あるいは、出力表の形式を定義する外部スキーマの記述モデルとしては有効であり得るが、DB内のデータを目的から独立に記述するという概念スキーマ記述には不向きである。

SUBJECTを改良した前述のGRASS、あるいは、STORMになると、統計表の物理モデルというより、統計データの論理構造を記述するものになってきているが、依然として、意味的に同じデータは同じ形式に記述されるべきであるという、概念スキーマの視点は希薄である。

(3) 対象中心のユーザ・インタフェース

DBの世界で対象中心のアプローチがとられるようになってきたのは、先の3スキーマが受け入れられ、その概念スキーマを記述するためのモデルとしてであった。目的に依存しないデータ記述のためには、目的から離れたデータ自身の意味を捉えること、すなわち、どの対象のどのような性質を記述したものであるかということを明らかにすることが、有効であったからである。

統計DBの場合も同様で、スウェーデンのOPRモデルも、国土庁のLIDS86のデータモデルSDS4Sも、共に、特定の目的に応じて集められた各種の統計データを、個々の収集目的から離れて、統合的に記述するための手段として、対象中心のモデルを採用しているのである。つまり、これらのモデルは、個々の収集目的に応じた統計ファイルの記述を内部スキーマとみたと、それらを統合する概念スキーマを記述するために提案されたものである。

それでは、このアプローチでは、統計DBの外部スキーマはどのように考えられてきたのであろうか。この問題は、次節で詳しく述べることにする。

7. 5 対象中心アプローチとIRSの拡張性

前節で述べたように現在のIRSの最大の課題の1つは、外部スキーマから独立した概念スキーマの確立であると思われる。他方、概念スキーマの記述モデルとしては、対象中心のモデルが今のところ唯一のものである。そこで、本節では、IRSのユーザ・インタフェースと対象中心アプローチとの関係について考察し、IRSの今後の拡張の方向性について私見を述べてみたい。

7.5.1 LIDS86のユーザ・インタフェース

統計DBに対する対象中心アプローチには、スウェーデンのOPRモデルと国土庁のLIDS86の2つのものがある。しかし、先にも述べたように、OPRモデルには、それに基づく実働システムは存在せず、そのユーザ・インタフェースの考え方は詳しくは明かでない。そこで、ここでは、まずLIDS86の考え方を紹介しておこう。

LIDS86では、統計データは、図7-3にあったように対象型の汎化階層に整理されている。従って、汎化階層上の内包的な対象型（概念ファイル）—例えば、貿易統計—を1つ指定すると、その下にある具体的な統計ファイル（DBファイル）の集合—この場合、世界経済統計、各国間貿易統計と通関統計—が定まる。

この概念ファイルとDBファイルの関係を分かりやすく書くと、図7-6のように書ける。LIDS86では、この図のようなDBファイルの集合の概念スキーマ上の記述を、そのまま画面に表示することにより、そのブラウザ画面が作られている。つまり、LIDS86のユーザ・インタフェースは、概念スキーマの一部を切り出したものであり、特に外部スキーマといえるものは持っていない。

なお、LIDS86では、輸入依存度のような加工データも、この図にあるように、そのDBファイル記述を用意することになっている。これは、実データと加工データの違いは単に内部スキーマ記述の違い、すなわち、物理DBから直接取り出すか、あるいは、

加工して作るかの違いに過ぎないからである。



図7-6 対象中心アプローチの概念ファイルとDBファイル

7.5.2 対象中心アプローチの汎用ユーザ・インタフェース

上で述べたLIDS86のブラウザは、単に概念スキーマ情報をブラウズできるだけで、データの検索を行うことはできない。そこで現実のLIDS86から離れて、LIDS86の概念スキーマ情報だけを利用して、特定目的でない汎用的な検索画面を作るとしたら、どのようなものができるか考えてみよう。図7-7は、IRSの検索画面に似せて作ったそのような画面の例である。なお、参考のために、IRSの検索画面を同様に示すと、図7-8のようになる。

図7-7の事項のサブ・ウィンドウは、GNP比を除き、図7-6の統計属性情報（事項の情報）より生成され得る。また、LIDS86の概念スキーマには、各々の分類属性の定義域の情報が害かれているので、それを利用することにより、報告国、相手国、品目の各サブ・ウィンドウを表示することができる。時点のサブ・ウィンドウは、月次、あるいは、四半期データを含めば、図のように、最新値、推移、前年同月比を表示するように、予め決めておけば良い。

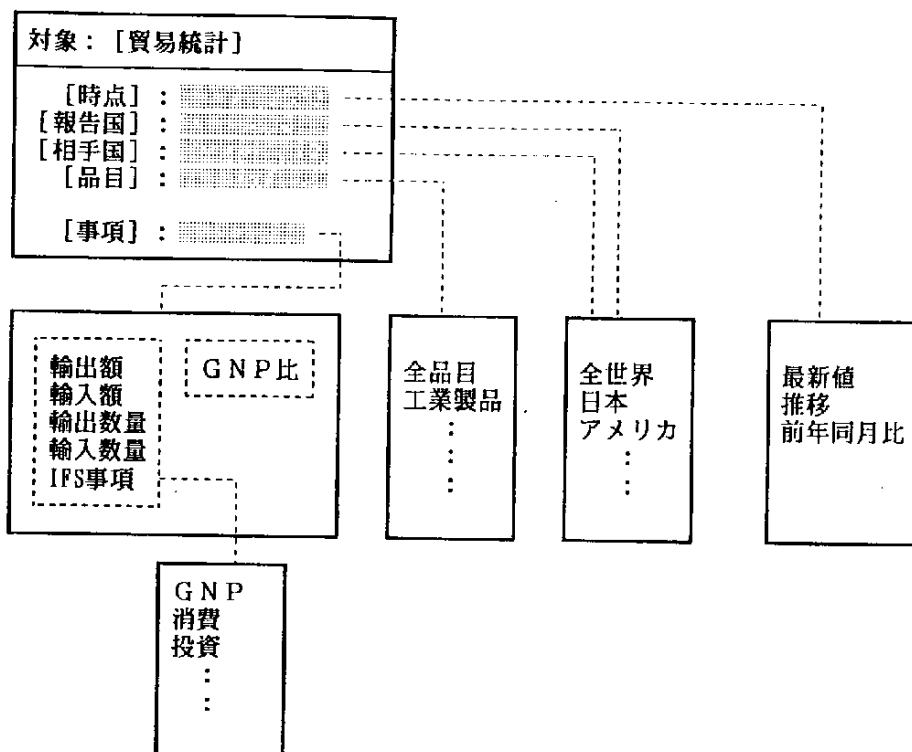


図7-7 対象中心アプローチの汎用ユーザ・インタフェースの例

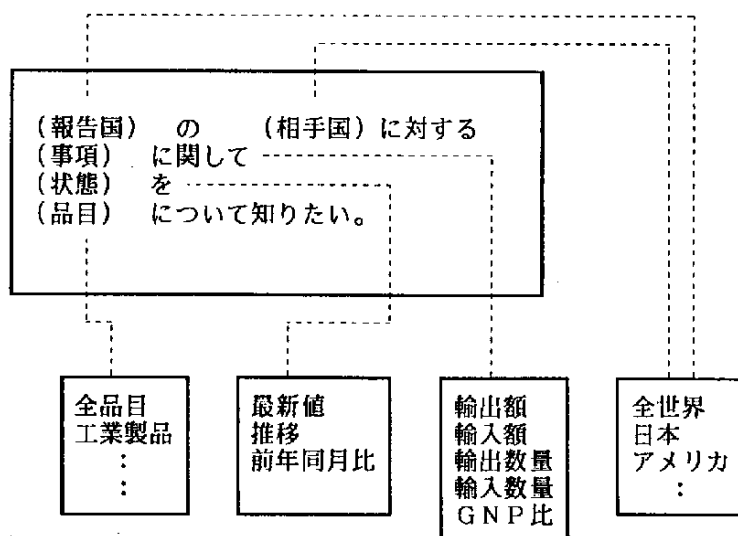


図7-8 現行のIRSのユーザ・インタフェース

GNP比は、やや扱いが面倒である。これについては以下のようないくつかの取り扱い方が考えられる。

- ① 輸入依存度と同様に1つ1つの事項のGNP比を、導出データとして、概念スキーマ上に定義する。
- ② 事項間の演算を検索画面上で定義できるようにすることで対応する。
- ③ IFS(International Financial Statistics)やSNA(System of National Accounts)のように、事項としてGNPをもつ統計ファイルがスコープ(DBファイルの集合)に入るときは、GNP比という事項オプションを表示するようにする。

①は、DB管理者に負担がかかり、②は、一般性はあるが、ユーザならびにシステム開発者に負担がかかる。③は、システム開発者に若干の負担がかかる。しかし、いずれにせよ、GNP比だけであれば、このような汎用的なユーザ・インタフェースの下でも、扱えないことはない。

図7-7と図7-8とを比較すると判るように、機能的には、汎用ユーザ・インタフェースでもIRSとほとんど同じことができそうである。

7.5.3 対象中心アプローチとIRSテンプレート

対象中心アプローチの汎用ユーザ・インタフェースとIRSのテンプレートを比較すると、以下のようにいうことができる。

まず、対象中心アプローチの汎用ユーザ・インタフェースが優れている点は、次の諸点である。

- ① ウィンドウならびに検索方法を一般化できるので、対象分野に依存しない統計DBの汎用的なユーザ・インタフェースとなり得る。
- ② 図7-3のような汎化階層は、DB内の統計ファイルの対象中心型の記述(図7-6のDBファイル記述)が与えられれば、かなり機械的に作ることができる。
- ③ ユーザが任意に自己のスコープ(DBファイルの集合)を決めた場合でも、それに対応する検索画面を自動生成できる。

すなわち、この方法では、検索システムはただ1つ作れば良く、対象分野の拡大は、概念スキーマ情報の整備のみで済むということになる。これに対し、現状のIRSテンプレートでは、テンプレートごとにそこで使用する推論知識、データ知識を個別に定義する必要がある。

他方、IRSテンプレートが優れている点は、次の諸点である。

① 「～に対する」「～に関する」等、項目に適した接続語がついているので分かりやすい。

この分かりやすさという点は、軽視すべきものではない。国土庁のLIDS86が、そのままでは、ユーザに受け入れられなかったのは、一般性を重視するあまり、この種の分かりやすさを欠いていたからである。

② 各入力項目にその初期化として適切なデフォルト値を与えられる。

それぞれの画面に適したデフォルト値を一般的に生成する方法を考えることは難しい。

③ アド・ホックな加工データ、あるいは、GNP比のような事項オプションを容易に追加できる。

将来、GNP比のようなものが多数登場し、しかも、7-5-2節で述べたようには、うまく一般化できないかも知れない。IRSテンプレートは、個別に定義するものであるから、そのようなものも取り込むことができる。

すなわち、個々の目的に応じたカスタマイズという点で、IRSテンプレートは優れているのである。

7.5.4 対象中心アプローチの外部スキーマとしてのIRSテンプレート

上で述べたIRSテンプレートの特徴は、まさしく、7.4節で述べた外部スキーマの特徴に他ならない。つまり、IRSテンプレートは、外部スキーマとして優れているのである。これまで当調査では、このIRSテンプレートを概念スキーマとみなしてきたために、対象分野の拡張に際して、困難が生じてきたのである。

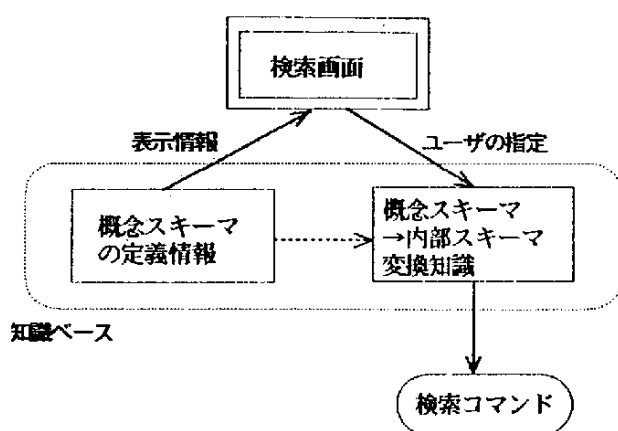
概念スキーマと外部スキーマを分けるとき、システム構成は、図7-9の(a)から(

b) へと変わることになる。ここで、「概念スキーマの定義情報」が(a)、(b)の双方にあるが、これらは異なるものであることに注意する必要がある。(a)のそれは、目的に依存したデータの記述であり、(b)のそれは、目的から中立的なデータの記述である。なお、(b)の中にある「外部スキーマのメタ知識」とは、検索画面のレイアウトなど、個々の外部スキーマに依存しない外部スキーマに関する知識である。

個々の外部スキーマに固有の知識をできるだけ少なくすることが、システムの拡張を容易なものにする上で重要である。概念スキーマでデータを記述するモデルと外部スキーマのそれとが大幅に異なると、「外部スキーマ→概念スキーマ変換知識」が複雑になり、外部スキーマの定義が難しくなる。しかしながら、図7-7と図7-8を比較すると判るように、IRSテンプレートと対象中心アプローチの項目は、極めて良く一致しており、この変換はごく簡単なものになると予想される。

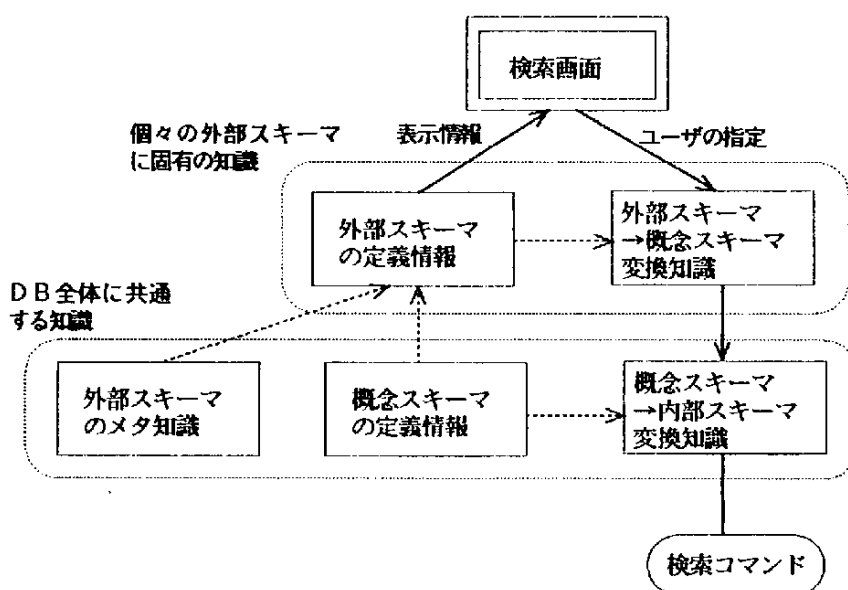
従って、対象分野の拡張という課題に対しては、概念スキーマを対象中心アプローチに従って作成し、外部スキーマをIRSテンプレートのように作ることが望ましい解であると思われる。これが7.3節で述べたスウェーデンのBo Sundgrenとの対話を通じて得た著者の結論である。

(東京国際大学 佐藤英人)



(a) 外部スキーマが未分化のとき

————→ 情報の流れ
 - - - - -> 内部的参照情報の流れ



(b) 概念スキーマと外部スキーマの区別がある時

図7-9 検索システムの知識の構成

む す び

本報告書では、3ヵ年計画の調査研究の2年目の成果についての報告を行った。

本年度は、委員会の活動を通じて、貿易統計に関するデータ検索を行う際の、検索支援に必要な知識を収集し、知識ベースを試作し、パーソナルコンピュータ上で稼動する試作システムを作成した。

試作システムは、まだ適用範囲が限定されているが、データベースや統計に関する専門知識を持たないユーザが、知識ベースの働きによって、3種のデータベースのなかから、的確なものを選択し、データベース検索コマンドを生成することができるようになっている。

また、試作システムの開発に係わる検討や、知識獲得の作業を通して、有意義な議論が行われ、今後の調査研究に大きな示唆を与える成果を残すことができた。

とくに、ユーザの検索要求の表現に関しては、時間をかけた議論を行うことができ、今後データベースの対象分野を拡張する際の基礎となる考え方がまとまったといえる。

海外調査に関しては、本年度は、フィンランド、スウェーデン、イタリアの経済統計データベース研究機関を訪問した。本調査研究に近い研究テーマを持つ機関を中心とした調査となり、有意義な意見交換を行うことができた。

試作システムに関する評価の過程で、現在の試作システムに関する次年度への拡張の方針と、問題点が明確になり、実用化へ向けての足掛かりができた。

次年度の調査研究では、本年度の成果をもとに、実用的な検索支援システム開発のための調査研究を継続してゆく予定である。

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、次年度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。

2. The second part of the document addresses the challenges of data management in a rapidly changing environment. It highlights the need for flexible and scalable solutions that can adapt to new technologies and data sources. The author argues that organizations must invest in training and development to ensure their staff are equipped to handle complex data sets and analyze them effectively.

3. The third part of the document focuses on the role of leadership in driving organizational success. It stresses that leaders must be able to inspire and motivate their teams, set clear goals, and provide ongoing support and feedback. The text also discusses the importance of communication and collaboration, suggesting that leaders should foster a culture of open dialogue and teamwork.

4. The fourth part of the document explores the impact of external factors on organizational performance. It discusses how changes in the market, regulatory environment, and global events can affect an organization's ability to achieve its goals. The author suggests that organizations should develop contingency plans and maintain a high level of agility to respond to these external challenges.

5. The fifth part of the document concludes by summarizing the key points discussed and offering final thoughts on the future of the organization. It reiterates the importance of continuous improvement and innovation, encouraging the organization to stay ahead of the curve by embracing change and seeking out new opportunities for growth.

付 録

付録1 知識ベースリスト

付録2 通商白書図表タイトルに関する重要度アンケート集計結果

付録3 データベース検索Q&A

付録4 検索要求フレームのサンプルケースへの適用

付録5 貿易統計に関わりの深い経済統計用語

付録6 利用目的別ユーザニーズアンケート

付録 1

知識ベースリスト


```

/**** window definition ****/
/* input window*/
iw(r_c,[16,16],[$日本$, $の$, off]).
iw(d_c,[304,16],[$全世界$, $に対する$, off]).
iw(kind,[16,48],[$輸入額$, $に関して$, msw3]).
iw(cond,[16,80],[$最新値$, $を$, msw3]).
iw(goods,[16,112],[$全品目$, $について知りたい$, off]).
/* menu window*/
mw(事項,[16,160],[2,8,kind,icw4]).
mw(状態,[420,160],[1,6,cond,icw5]).
/* window size and collar */
def_wind(iw,22,1,3).
def_wind(mw,22,1,0).
def_wind(emw,80,1,2).

/**** attribute definition ****/
/* kind of attribute*/
def_atr([r_c,d_c,kind,cond,goods]).
/* relation of attributes*/
def_atr_relation(exchange,[r_c,d_c]).

/**** data base definition ****/
/* kind of db*/
def_db([trade,dot,ifs]).
/* attribute of db */
dbd(trade,[r_c([日本]),d_c,cond,goods,kind],
    [輸出額,輸入額],[1960,1989],[年,年度],[年,四半期,月],[円,ドル]).
dbd(dot,[r_c,d_c,cond,goods([全品目]),kind],
    [輸出額,輸入額],[1960,1989],[年,年度],[年,四半期,月],[円,ドル]).
dbd(dot,[r_c,d_c,cond,goods([全品目]),kind],
    [輸出数量,輸入数量],[1970,1989],[年,年度],[年,四半期,月],[ ]).
dbd(dot,[r_c,d_c,cond,goods([全品目]),kind],
    [双極化指数],[1970,1989],[年,年度],[年,月],[ ]).
dbd(ifs,[r_c,d_c([全世界]),cond,goods([全品目]),kind],
    [GNP,為替レート],[1955,1985],[年,年度],[年,四半期,月],[NC]).
def_cat([全世界,全品目]).

```

```

/* condition for each kind */
def_cond(最新値,[D,new,1],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(指定時点の値,[D,y,1],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(指定期間の値,[y,y,D],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(最近10年の推移,[D,new,10],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(上位10品目,[D,new,1],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(上位10カ国,[D,new,1],
[輸出額,輸入額,輸出数量,輸入数量,双極化指数,GNP,為替レート,依存度]).
def_cond(対前年比,[D,new,1],
[輸出額,輸入額,GNP,為替レート]).
def_cond(対GNP比,[D,new,1],
[輸出額,輸入額,為替レート]).

exp(双極化指数,
'二国間の貿易の結合の強さを示す').
exp(.,'sorry,I have no message for this item.').

/**** define reduction ****/
rd(依存度,['輸入額','GNP'],[1,2],
$国民生産に対する貿易の寄与$,
[r_c,d_c([全世界]),cond,goods([全品目]),kind]).
rd_f(1,$A/B$).
rd_f(2,$abs(A/B)$).

/**** window definition ****/
wind(icw1,462,532,0,22,icw).
wind(icw2,533,602,0,22,icw).
wind(icw3,603,637,0,22,icw).
wind(icw4,350,390,160,182,icw_2).

```

```
wind(icw5,590,630,160,182,icw_2).  
wind_exw(1,21,39,229,247,exw).  
wind_exw(2,21,39,251,269,exw).  
wind_exw(3,21,39,273,291,exw).  
wind_exw(4,21,39,295,313,exw).  
wind_exw(5,21,39,317,335,exw).
```

/*****/
/* 類義語辞書です。*/
/* 1/9/26 S.T */
/*****/

/* 概況分類・特殊品目 */

dic(goods,食料品及びその他の直接消費財,[食料品,直接消費財]).

dic(goods,工業用原料,[]).

dic(goods,粗原料,[]).

dic(goods,鉱物性燃料,[]).

dic(goods,化学工業生産品,[]).

dic(goods,金属,[]).

dic(goods,繊維品,[]).

dic(goods,資本財,[]).

dic(goods,一般機械,[]).

dic(goods,電気機器,[]).

dic(goods,輸送機器,[]).

dic(goods,非耐久消費財,[]).

dic(goods,繊維製品,[]).

dic(goods,耐久消費財,[]).

dic(goods,家庭用品,[]).

dic(goods,家庭用電気機器,[]).

dic(goods,乗用車,[]).

dic(goods,二輪自動車類・自転車類,[オートバイ,自転車]).

dic(goods,がん具・楽器類,[おもちゃ,楽器]).

dic(goods,食料品,[]).

dic(goods,魚介類,[魚]).

dic(goods,原燃料,[]).

dic(goods,軽工業品,[]).

dic(goods,繊維及び同製品,[繊維,繊維製品]).

dic(goods,綿織物,[]).

dic(goods,合成繊維織物,[]).

dic(goods,非金属鉱物製品,[]).

dic(goods,重化学工業品,[]).

dic(goods,化学製品,[]).

dic(goods,化学肥料,[]).

dic(goods,金属及び同製品,[金属,金属製品]).

dic(goods,鉄鋼,[]).

dic(goods,機械機器,[]).

dic(goods,テレビ受像機,[テレビ]).

dic(goods,ラジオ受像機,[ラジオ]).

dic(goods,自動車,[車]).

dic(goods,船舶,[船]).

/* 概況分類・一般品目（輸出） */

dic(goods,食料品及び動物,[食料品,動物]).

dic(goods,生きた動物,[動物]).

dic(goods,肉類及び同調製品,[肉類,肉]).

dic(goods,酪農品及び鳥卵,[酪農品,卵]).

dic(goods,ミルク及びクリーム,[ミルク,クリーム]).

dic(goods,魚介類及び同調製品,[魚,魚介類]).

dic(goods,魚介類（生鮮）,[魚,魚介類]).

```

dic(goods,時計用貴石及び半貴石,[時計用貴石,半貴石]).
dic(goods,写真用フィルム類,[フィルム]).
dic(goods,未撮影映画用フィルム,[ ]).
dic(goods,撮影した映画用フィルム,[ ]).
dic(goods,録音用・録画用テープ類,[テープ]).
dic(goods,がん具及び遊戯用具,[がん具,遊戯用具,おもちゃ]).
dic(goods,ゴルフ用具,[ゴルフ用品]).
dic(goods,美術品・収集品及びこつとう,[美術品,収集品,骨董]).
dic(goods,成型品及び彫刻品,[彫刻]).
dic(goods,再輸入品,[ ]).

```

```

/*****
/* 国/地域の類義語辞書 */
*****/

```

```

dic(r_c,日本,['JAPAN', ]).
dic(r_c,大韓民国,['KOREA',韓国]).
dic(r_c,北朝鮮,[ ]).
dic(r_c,中華人民共和国,[中国,'CHINA' ]).
dic(r_c,台湾,[ ]).
dic(r_c,モンゴル,[ ]).
dic(r_c,香港,[ ]).
dic(r_c,ベトナム,[ ]).
dic(r_c,タイ,[ ]).
dic(r_c,シンガポール,[ ]).

```

```

dic(r_c,マレーシア,[ ]).
dic(r_c,ブルネイ,[ ]).
dic(r_c,フィリピン,[ ]).
dic(r_c,インドネシア,[ ]).
dic(r_c,カンボジア,[ ]).
dic(r_c,ラオス,[ ]).
dic(r_c,ビルマ,[ ]).
dic(r_c,インド,[印度]).
dic(r_c,パキスタン,[ ]).
dic(r_c,スリランカ,[ ]).

```

```

dic(r_c,モルディヴ,[ ]).
dic(r_c,バングラディシュ,[ ]).
dic(r_c,東チモール,[ ]).
dic(r_c,マカオ,[ ]).
dic(r_c,アフガニスタン,[アフガン]).
dic(r_c,ネパール,[ ]).
dic(r_c,ブータン,[ ]).
dic(r_c,イラン,[ ]).
dic(r_c,イラク,[ ]).
dic(r_c,バーレーン,[ ]).

```

```

dic(r_c,イエメン民主人民共和国,[ ]).
dic(r_c,サウジアラビア,[サウジ]).
dic(r_c,クウェート,[ ]).
dic(r_c,カタール,[ ]).
dic(r_c,オーマン,[ ]).
dic(r_c,イエメン,[ ]).
dic(r_c,イスラエル,[ ]).
dic(r_c,ヨルダン,[ ]).
dic(r_c,シリア,[ ]).

```

```

/*****/
/* 概況分類の品目データです。*/
/*      1/9/7  S.T  */
/*****/
/* 1. 収録データは概況分類約900品目です。      */
/*      */
/* 2. 構造は次の通りです。      */
/*      */
/*      概況分類・特殊品目の場合      */
/*      */
/*      g_code (Name, Code).      */
/*      Name : 品目名      */
/*      Code : 先頭を除いた2桁目以降のコード */
/*      */
/*      概況分類・一般品目の場合      */
/*      */
/*      g_code (Name, Code).      */
/*      Name : 品目名      */
/*      Code : [輸出コード, 輸入コード]      */
/*****/

```

```

/* 概況分類・特殊品目 */

```

```

g_code(食料品及びその他の直接消費財,'21').
g_code(工業用原料,'22').
g_code(粗原料,'221').
g_code(鉱物性燃料,'222').
g_code(化学工業生産品,'223').
g_code(金属,'224').
g_code(繊維品,'225').
g_code(資本財,'23').
g_code(一般機械,'231').
g_code(電気機器,'232').
g_code(輸送機器,'233').
g_code(非耐久消費財,'24').
g_code(繊維製品,'241').
g_code(耐久消費財,'25').
g_code(家庭用品,'251').
g_code(家庭用電気機器,'252').
g_code(乗用車,'253').
g_code(二輪自動車類・自転車類,'254').
g_code(がらん具・楽器類,'255').
g_code(食料品,'26').
g_code(魚介類,'261').
g_code(原燃料,'27').
g_code(軽工業品,'28').
g_code(繊維及び同製品,'281').
g_code(綿織物,'2811').
g_code(合成繊維織物,'2812').
g_code(非金属鉱物製品,'282').
g_code(重化学工業品,'29').
g_code(化学製品,'291').
g_code(化学肥料,'2911').
g_code(金属及び同製品,'292').
g_code(鉄鋼,'2921').
g_code(機械機器,'293').
g_code(テレビ受像機,'2931').
g_code(ラジオ受像機,'2932').

```

g_code(自動車,'2933').
g_code(船舶,'2934').

/* 概況分類・一般品目 */

g_code(食料品及び動物,['310','410']).
g_code(生きた動物,['31001','41001']).
g_code(肉類及び同調製品,['31003','41003']).
g_code(酪農品及び鳥卵,['31005','41005']).
g_code(ミルク及びクリーム,['3100501','4100501']).
g_code(魚介類及び同調製品,['31007','41007']).
g_code(魚介類(生鮮),['3100701','*']).
g_code(鮮魚及び冷凍魚,['310070101','*']).
g_code(かつお,['3100701011','*']).
g_code(まぐろ,['3100701012','*']).
g_code(かじき,['3100701013','*']).
g_code(甲殻類及び軟体動物(生鮮、冷凍、塩蔵、乾燥),['310070103','*']).
g_code(えび,['3100701031','*']).
g_code(魚介類の調整品,['3100705','4100711']).
g_code(魚介類のかん詰,['310070501','*']).
g_code(まぐろ及びかつお,['3100705011','*']).
g_code(さば,['3100705012','*']).
g_code(かき,['3100705013','*']).
g_code(穀物及び同調製品,['31009','41009']).
g_code(小麦粉,['3100901','*']).
g_code(果実及び野菜,['31011','41011']).
g_code(果実,['3101101','4101101']).
g_code(うんしゅうみかん(生鮮),['310110101','*']).
g_code(りんご(生鮮),['310110103','*']).
g_code(果実かん詰,['310110105','*']).
g_code(うんしゅうみかんかん詰,['3101101051','*']).
g_code(野菜,['3101103','4101103']).
g_code(乾しいたけ,['310110301','*']).
g_code(糖類及び同調製品・はちみつ,['31013','41013']).
g_code(コーヒー・茶・ココア・香辛料類,['31015','41015']).
g_code(茶,['3101501','*']).
g_code(飼料,['31017','41017']).
g_code(魚の粉及びミール,['3101701','4101705']).
g_code(配合飼料(ペットフードを含む),['3101703','*']).
g_code(その他の調整食料品,['31019','41019']).
g_code(飲料及びたばこ,['311','411']).
g_code(飲料,['31101','41101']).
g_code(たばこ,['31103','41103']).
g_code(葉たばこ,['3110301','4110301']).
g_code(食料に適しない原材料,['312','412']).
g_code(原皮及び毛皮(未仕上),['31201','41201']).
g_code(採油用の種・ナット及び核,['31203','41203']).
g_code(生ゴム,['31205','41205']).
g_code(合成ゴム,['3120501','4120505']).
g_code(木材及びコルク,['31207','41207']).
g_code(木材,['3120701','4120701']).
g_code(製材,['312070101','412070103']).
g_code(なら製材,['3120701011','*']).
g_code(パルプ及び紙,['31209','41209']).
g_code(織物用繊維及びくず,['31211','41211']).
g_code(人造繊維,['3121105','*']).
g_code(合成繊維短繊維,['312110501','*']).

```

/*****/
/* I F Sの系列コード辞書です。*/
/*      1/9/8  S.T  */
/*****/
/* 構造は次の通りです。 */
/*****/
/*  k_code (Name, Code, List) . */
/*      Name : 系列名 */
/*      Code : 系列コード */
/*      List : 単位 (各国通貨建てまたはドル建て) */
/*      'NC' は各国通貨建てを表す。 */
/*      '*' は各国通貨でもドルでもないことを示す。 */
/*****/

```

```
k_code(GNP,'99A..',[unit('NC')]).
```

```
k_code(GNP,'99A.C',[unit('NC')]).
```

```
k_code(GDP,'99B..',[unit('NC')]).
```

```
k_code(GDP,'99B.C',[unit('NC')]).
```

```
k_code(交換比率,'..RF.',[unit('*')]). /* NC／ドル */
```

```
k_code(輸出額,'70...',[unit('NC')]).
```

```
k_code(輸出額,'70..D',[unit('$')]).
```

```
k_code(輸出額,'70.XD',[unit('$')]).
```

```
k_code(輸出額,'70.V.',[unit('NC')]).
```

```
k_code(輸入額,'71...',[unit('NC')]). /* CIF */
```

```
k_code(輸入額,'71..D',[unit('$')]). /* CIF */
```

```
k_code(輸入額,'71.V.',[unit('NC')]). /* FOB */
```

```
k_code(輸入額,'71.VD',[unit('$')]). /* FOB */
```

```
k_code(輸入額,'71.X.',[unit('$')]).
```

```
k_code(輸入額,'71.XD',[unit('$')]).
```

```
k_code(輸出数量指数,'72...',[unit('*')]).
```

```
k_code(輸出数量指数,'72.X.',[unit('*')]).
```

```
k_code(輸入数量指数,'73...',[unit('*')]).
```

```
k_code(輸出価格指数,'74...',[unit('NC')]).
```

```
k_code(輸出価格指数,'74..D',[unit('$')]).
```

```
k_code(輸入価格指数,'75...',[unit('NC')]).
```

```
k_code(輸入価格指数,'75..D',[unit('$')]).
```

```
k_code(原油輸入額,'71AA.',[unit('NC')]).
```

```
k_code(原油輸入額,'71AAD',[unit('$')]).
```

```
k_code(MERM,'..AMX',[unit('*')]).
```

```
k_code(人口,'99Z..',[unit('*')]).
```

```
k_code(人口,'99ZX.',[unit('*')]).
```

```
k_code(卸売物価指数,'63...',[unit('*')]).
```

```
k_code(商品輸出額,'77AAD',[unit('$')]).
```

```
k_code(商品輸入額,'77ABD',[unit('$')]).
```

```

/*****/
/* 国・地域コードの辞書です。*/
/*      1/9/7  S.T  */
/*****/
/* 構造は次の通りです。 */
/* c_code(国名,[ I F Sコード,D O Tコード, T R A D Eコード]). */
/*****/

/* 国コード */

c_code(日本,[158,158,*]).
c_code(大韓民国,[542,542,103]).
c_code(北朝鮮,[*,954,104]).
c_code(中華人民共和国,[924,924,105]).
c_code(台湾,[*,*,106]).
c_code(モンゴル,[*,948,107]).
c_code(香港,[532,532,108]).
c_code(ベトナム,[582,582,110]).
c_code(タイ,[578,578,111]).
c_code(シンガポール,[576,576,112]).

c_code(マレーシア,[548,548,113]).
c_code(ブルネイ,[516,516,116]).
c_code(フィリピン,[566,566,117]).
c_code(インドネシア,[536,536,118]).
c_code(カンボジア,[522,522,120]).
c_code(ラオス,[544,544,121]).
c_code(ビルマ,[518,518,122]).
c_code(インド,[534,534,123]).
c_code(パキスタン,[564,564,124]).
c_code(スリランカ,[524,524,125]).

c_code(モルディヴ,[556,556,126]).
c_code(バングラディシュ,[513,513,127]).
c_code(東チモール,[*,*,128]).
c_code(マカオ,[546,546,129]).
c_code(アフガニスタン,[512,512,130]).
c_code(ネパール,[558,558,131]).
c_code(ブータン,[514,*,132]).
c_code(イラン,[429,429,133]).
c_code(イラク,[433,433,134]).
c_code(バーレーン,[419,419,135]).

c_code(イエメン民主人民共和国,[459,459,136]).
c_code(サウジアラビア,[456,456,137]).
c_code(クウェート,[443,443,138]).
c_code(カタール,[453,453,140]).
c_code(オーマン,[449,449,141]).
c_code(イエメン,[473,473,142]).
c_code(イスラエル,[436,436,143]).
c_code(ヨルダン,[439,439,144]).
c_code(シリア,[463,463,145]).
c_code(レバノン,[446,446,146]).

c_code(アラブ首長国連邦,[466,466,147]).
c_code(ガザ,[*,*,148]).
c_code(アイスランド,[176,176,201]).

```


付録2

通商白書図表タイトルに関する重要度アンケート集計結果

付録2 通商白書図表タイトルに関する重要度アンケート集計結果

通商白書の図表タイトル索引に掲載された図表タイトルに、貿易統計としての重要度の順に1から3までの得点を付けてもらった。

回答は、通産省、JETRO、アジア経済研究所よりよせられた回答を集計し、得点の高い順に並べたものが、以下の表である。

点	タイトル	番号
9	世界貿易の動向	1-1-7
	主要4地域間の貿易動向	1-1-11
	(アメリカの)地域別貿易動向	1-1-12
	(アメリカの)貿易収支の推移	1-1-19
	(アメリカの)商品別貿易収支動向	1-1-20
	(EC各国の)貿易収支の推移	1-1-28
	(中国の)貿易動向	1-1-31
	(ソ連の)地域別貿易動向	1-1-34
	(我が国の)地域別輸入動向	2-2-24
	地域別輸出動向	2-2-7
	財別輸出動向	2-2-5
	財別輸入動向	2-2-24
	88年の財別地域別輸出動向	2-2-7
	輸入額、輸入価格、輸入数量の推移	2-2-22
8	(ソ連の)西側先進国向け輸出構造	1-1-35
	世界の農業貿易	1-3-23
	(我が国の)輸出動向(前年比(同期比)伸び率)	2-2-1
	部品輸出の推移	2-2-8
	品目別輸出数量の推移(85年基準)	2-2-12
	電気機器輸出の推移	2-2-13
	有線電信機器の輸出数量の推移	2-2-14
	排気量別乗用車輸出台数増減率(87~88年)	2-2-15
	鉄鋼輸出の推移	2-2-16
	表面処理鋼板の輸出の推移	2-2-17
	資本財輸出の推移	2-2-18
	(アメリカの)資本財輸入動向	2-2-20
	品目別輸出入	2-2-40
	我が国の機械部品輸入の動向	2-2-44
	ECの世界貿易(輸出)に占めるシェアの推移	3-2-1
	世界の工業製品輸出額に占めるECのシェア	3-2-2
	(EC)及び世界の工業製品輸出額伸び率	3-2-5
	(ECと我が国の)工業製品輸出構造の変化	3-2-7
	工業製品輸出に占めるハイテク製品輸出のウェイトの推移	3-2-8
	ECの相手国・地域別輸出額伸び率	3-2-23
	(アジアNIES、ASEAN諸国の)輸出に占める工業製品の割合の推移	3-3-25
	ASEAN諸国の地域別輸出動向	3-3-34
	(アジアNIES・ASEAN諸国)間貿易の品目構成の推移	3-3-35
	(マレーシアの)対(ASEAN諸国)電子機器輸出動向	3-3-38
	中国の工業製品輸出の推移	3-3-42
	中国の地域別貿易動向	3-3-43

	中国の香港経由対韓国、対台湾間接貿易の推移 オーストラリアのアジアNIES向け工業製品、鉱物性燃料輸出の推移	3-3-44 3-3-54
7	《西ドイツの》域内相手国別機械輸出額（前年同期差） （EC各国の）財別貿易収支 韓国からのニットの製品及び付属品の輸入額 農業貿易の商品貿易に占めるシェアと貿易額の推移 世界の工業製品需要と我が国輸出数量の推移 通関貿易収支の推移 （アメリカの）部品輸入の推移 （台湾の）家電製品貿易 （マレーシア）を巡る半導体貿易動向 （オーストラリアの）地域別輸出動向 （ニュージーランド）地域別輸出動向 （韓国、台湾）における部分品の輸入依存	1-1-29 1-1-30 1-3-11 1-3-22 2-2-3 2-2-36 3-1-7 3-3-23 3-3-37 3-3-52 3-3-53 3-3-56
6	（我が国の）輸入弾性値と資本財輸出数量の変動要因 相手地域別製品輸入額と製品輸入比率 加工度の高い食料品等、石油製品を製品に加えた場合の製品等輸入比率（試算値） 主要先進国の製品輸入指標	2-2-19 2-2-27 2-2-28 2-2-29
5	世界経済・貿易の実質成長率の推移（対前年比伸び率） 先進国の実質輸入伸び率と主要国の寄与率 （日本・アメリカ・西ドイツの）国際収支の動向 （日本・アメリカ・西ドイツの）貿易収支の変動要因 （アメリカの）輸入数量の変動要因 （ラテンアメリカ諸国の）実質為替レートの推移 世界の対先進国輸入と我が国の輸入の品目別増加寄与度 輸出数量の変動要因 財別輸入の相手地域別増加寄与率（88年） 製品輸入比率の推移と財別寄与 （我が国の）輸入構成（円ベース）の変化 輸入数量の財別弾性値（7.9～88年） 輸出入構造変化指数 我が国の貿易特化指数の変化（85～88年） 我が国製造業の逆輸入の増加 （アメリカの）完成工業品収支の推移 ECの工業製品貿易における域内依存度の推移 （EC）の製品別競争力係数の推移 （EC）のハイテク製品競争力係数の推移 （アジアNIES）の製品用途別国際競争力の推移 （韓国・台湾の）輸出地域集中度指数の推移 （ASEAN諸国の）製品用途別国際競争力の推移 （ASEAN諸国の）貯蓄・投資バランス （中国の）製品別貿易特化係数（84年→88年） （中国の）（対日、対アジアNIES）機械輸入の動向 （我が国の）対（アジアNIES、ASEAN諸国）品目別貿易特化係数の推移	1-1-1 1-1-6 1-1-9 1-1-10 1-1-21 1-2-28 1-3-6 2-2-2 2-2-25 2-2-26 2-2-31 2-2-32 2-2-41 2-2-45 2-3-12 3-1-6 3-2-4 3-2-10 3-2-11 3-3-10 3-3-23 3-3-26 3-3-27 3-3-45 3-3-46 3-3-57
4	GATT加盟国の数の推移と世界貿易（輸出）に占める加盟国のシェア 輸出の地域集中度指数 輸出の所得弾性値の推移 財別輸入数量の前回円高時との比較 輸出額上位企業（製造業）の我が国輸出総額に占める割合 （アジアNIES、ASEAN諸国）からの耐久消費財輸入の例 （アメリカの）品目別競争力係数の変化 （アメリカの）機械・輸送機器の品目別競争力係数の動向	1-3-1 1-3-10 2-2-4 2-2-34 2-2-39 2-3-8 3-1-11 3-1-12

3	(84年からのアメリカの)貿易収支変動要因 (我が国の)生産比率、輸出比率、輸入比率の推移 (環太平洋地域内)及び(E C)との間の輸出結合度(87年 域内競争指数の推移 (韓国、台湾の)実質GDP成長率及び輸出の寄与度 (日米-アジア諸国)の実質為替レートの推移 半導体各製品の地域別シェア(金額ベース)	1-2-20 2-2-42 3-3-4 3-3-5 3-3-6 3-3-7 3-3-15
2	(日本、アメリカ、日ドイツの)経常収支の対GNP比率 (発展途上国の)実質経済成長率 (中国の)主要経済指標の推移 (ソ連の)主要経済指標の推移 主要各国通貨の対ドル直物為替レート指数の推移 主要各国通貨の実行為替レートの推移 (アメリカ)卸売物価の変動要因 (アメリカ)輸入物価の推移 (ラテンアメリカ)の主要経済指標 (ラテンアメリカ諸国の)工業化率と輸出に占める工業製品の シェア 地域別経常収支動向 (1930年前後の)世界貿易と世界経済の推移 世界の貿易の伸びと鉱工業生産 (60年代の)工業製品輸出構造変化指数 東京ラウンド前後における鉱工業品平均関税率の変化 主要先進国の貿易不均衡指数と輸入額に占める非関税障壁関連 品の割合の推移 鉱工業生産・出荷動向 国内向け鉱工業出荷の推移 (カラーテレビの)対アメリカ輸出と現地生産の関係 (自動車の)対アメリカ輸出と現地生産の関係 海外生産比率の推移及び見通し 我が国の輸入数量の従来の関数による推計値からの乖離 国際収支総括表 貿易外収支と海外旅行者数の推移 対外収支の対GNP比率 工業用原料における粗原料から中間財への輸入代替化 主要耐久消費財の実質購入価格の推移 主要耐久消費財の購入達成率の推移 中間財輸入数量の変動要因 製造行の完成品輸入(逆輸入、開発輸入、OEM輸入を除く) とその要因 小売企業の輸入品仕入れ額及び輸入品仕入れ割合の推移 (ブラザ合意移行の)(アメリカの)輸出入数量と生産の変化 (アメリカの)品目別輸出入比率と実行為替レートの推移 (アメリカの)(最近の)機械・輸送機器の輸出入相対価格指 数の推移 (アメリカの)国内需要と輸入の相関関係(80~87年平均 先進7か国のGNPに占める輸出比率の比較 (アメリカの)輸出の在庫弾性値 実質経済成長率と失業率の推移 (環太平洋地域諸国の)経済成長の推移 (環太平洋地域諸国の)実質経済成長率の推移 (環太平洋地域の)経済構造変化と一人当たりのGNPの推移 (台湾)におけるNC工作機械生産・貿易動向 (韓国・台湾の)需要項目別、産業別実質成長率の推移 (台湾)と(アジアNIES、韓国)との産業内分業度指数の 推移(工業製品) (中国の)香港経由自国向け輸出の推移 (日米、アジアNIES、ASEAN諸国)における需要波及 マトリックス(88年)	1-1-8 1-1-13 1-1-32 1-1-36 1-2-1 1-2-2 1-2-16 1-2-17 1-2-24 1-2-26 1-2-31 1-3-2 1-3-4 1-3-5 1-3-7 1-3-8 2-1-9 2-1-10 2-2-9 2-2-10 2-2-11 2-2-30 2-2-35 2-2-37 2-2-38 2-2-43 2-3-5 2-3-9 2-3-10 2-3-11 2-3-14 3-1-8 3-1-10 3-1-13 3-1-16 3-1-20 3-1-21 3-2-3 3-3-1 3-3-2 3-3-3 3-3-12 3-3-16 3-3-21 3-3-50 3-3-60

(アジアNIES、ASEAN諸国) の対 (日) 需要波及係数 の変化	3-3-61
(日米の) 対 (アジアNIES) 需要波及係数の変化 (86年→88年)	3-3-62

- アンケートの集計結果と白書を参照し、①タイトルと実際の表との関係調べる
- ②事例として重要なものを抽出する...という観点からまとめ、IRSの拡張として、実現可能性をコメントする。

(1) タイトルと実際の表との関連

重要なタイトルは、下記の構成をしたものが多い

地域別	+	貿易の	+	動向
商品別		貿易収支の		推移
財別		輸出（数量、額、価格）の		
地域間		輸入（数量、額、価格）の		

① 常識的な用語の使い方

地域別→世界、先進国（アメリカ、日本、西ドイツ、その他）、発展途上国
共産圏

商品別→全商品、農産物、鉱産物、工業製品

財別 →食料品等直接消費財、工業用原料、（化学品、金属）、資本財（一般機械、電気機器、輸送機器）、非耐久消費財、耐久消費財（家電機器、乗用車）（国によって分類に違いがある）

貿易 →輸出額、輸入額の原数値及び前期比

貿易収支→輸出額と輸入額の差の意味だが、輸出額、輸入額と共に、表示するのが一般的

動向、推移 →同義に使われる、3 から 12 時点程度の時系列

国際競争力 →貿易特化係数

用途別→ 非耐久消費財、労働集約的中間財、耐久消費財、資本財、資本集約的中間財

輸入弾性値と資本財輸出数量の変動要因→ 相対価格要因と国内需要要因

貯蓄・投資バランス→ 総資本形成/名目GNPと貯蓄/名目GNPの比較

産業別 → 製造業、建設業、農林水産業、サービス業

需要項目別 → 輸出等、輸入等、民間固定資本形成、民間消費支出

② データに関する注意

4 半期毎のデータ → 季節調整をしてあるか否かに注意が必要

(2) 事例として重要な用語

① 品目として

部品、電気機器、有線電信機器、乗用車（排気量別）、鉄鋼、表面処理鋼板
資本財、機械部品、工業製品、ハイテク製品、電子機器、鉱物性燃料、ニット製
品、ニット付属品、家電製品、半導体、石油製品、完成工業品、用途別製品
NC工作機械、非関税障壁関連品、カラーテレビ、粗原料

② 事項として

シェア、輸入弾性値、製品輸入指標、品目別寄与度、地域別寄与度、財別寄与
輸出（入）構造変化指数、貿易特化係数、逆輸入、域内依存度、競争力係数
地域集中度指数、産業分業度指数、実質（名目）GNP、GDP

③ 地域分類として

GATT加盟国、EC、アジアNIES、ASEAN諸国

環太平洋地域、先進7か国

④ 時点として

前回円高時、東京ラウンド、プラザ合意

⑤ 表として

国際収支総括表

拡張の可能性

- ・一般的と認められる用語を辞書として登録することは可能だが、複数の意味を持つものは、問い合わせが必要
- ・～別・・・の指示機能（試作システムには、ない）の需要が大きい・・・追加可能
- ・国際収支総括表のように、表のイメージで出したいものがあるか。データベース側で表イメージを保持していて、コマンドで表が表示できれば可能。

IRSで表を合成するのは、負荷が大きい

- ・重要な用語のなかに、説明の煩雑なものがあるが、どのような形式で、説明すればよいか検討が必要（別紙1参照）

- ・国ごとにカテゴリーが違う場合の処理（財別分類等）は、検討が必要
- ・データに関する説明は、データベースを選択した後に、システムからのメッセージとして、表示すればよいか

（季節調整の有無、欠測時点等）

- ・事例としてメニュー化する場合、一般的で、長期の使用に耐えるものと、話題性が大きく、一時的であると思われるもの（ハイテク製品等）がある

付録 3

データベース検索Q&A

Q1. RCA指数（揭示的比較優位指数）について

定義式は

$$\frac{i \text{ 国 } k \text{ 財輸出}}{\text{世界の } k \text{ 財輸出}}$$

$$\frac{i \text{ 国の総輸出}}{\text{世界の総輸出}}$$

となっているが、k財に相当するものとして、何が考えられるか？

概況分類の特殊分類として、資本財、耐久消費財、直接消費財があるが、これらのものと考えてよいか？

A 基本的には、個別の商品について計算できればよい、一般的には、資本財、消費財等に関して使うことが多い。

Q2. 製品輸入額について

製品輸入額＝化学品＋金属品＋機械機器＋雑製品で定義されているが、概況分類の具体的なコードとしては、次でよいか？

化学品：化学製品（コード・4291）

金属品：金属及び同製品（コード・4292）

機械機器：機械機器（コード・4293）

雑製品：雑製品（コード・418）

A 通白分類の4312が製品に対応している。但し通白の特殊コードについては、

検索時にデータの積み上げを行うため検索時間がかかる。

来年は、概況分類に関して、特殊分類に関するデータ登録を行う予定なので、その時点で、概況分類の製品コードに切替えて使うのが有効。

Q 3. 市場別輸出入比率について

DOTの輸出・入シェアからどうすれば求められるか？

A 市場別構成比 (market share) (%)

$$\frac{i \text{ 地域 (国) への輸出 (入) 額 (量)}}{\text{対世界輸出 (入) 額 (量)}} \times 100$$

DOTの場合には商品別のデータがないので、総輸出 (入) 額レベルになり、TRADEでは、商品別 (特定の商品) について計算できる。

例：日本のハイテク製品輸入における米国製品の比率

Q 4. 商品別輸出入比率について

定義式は？

A 商品別構成比 (product share) (%)

$$\frac{i \text{ 地域 (国) への特定製品の輸出 (入) 額}}{i \text{ 地域 (国) への総輸出 (入) 額}} \times 100$$

例：日本の対米輸出に占める米国からのハイテク製品輸入

Q 5. 前年同期 (月) 増減比について

次の定義でよいか？

本年値－前年同期値

前年同期値 × 100

A この定義でよい

但し、為替相場の変動比較の場合は、次のIMF方式が一般的に用いられる。

(普通の増減率の計算式は、為替の場合には、便宜的に欧州方式と呼ぶ)

IMF方式： 前年同期比

(本年値 - 1) × 100

Q6. 原油輸入比率について

IFSで検索する場合・・・原油輸入額は、コード71aa、または71aad
輸入総額はCIF価格。

TRADEで検索する場合・・・原油輸入額は、原油及び粗油・コード4130301
輸入総額はコード1。

でよいか？

A これでよい。71aaがない場合は、71a で代替えすることができる。

71aa・・・原油の輸入

71aad・・・原油の輸入（米ドル建て） 末尾のdは米ドル建てを示す

71a・・・石油の輸入

71ad・・・石油の輸入（米ドル建て）

OECDの貿易統計を使用する場合は、SITC-R2分類=333を使用する

Q7. 一人あたりGNPについて

GNPを人口（コード99z.）で割ったものでよいか？

また1か国のみ別記（コード99zx.）されているが例外処理をするべきか？

A これでよい

中国の人口系列が99zx.

人口の単位は、国毎に違うので注意が必要

Q8. GNP、GDPについて

IFSのマニュアルを見ると、GNP、GDPは各々検索コードが複数記載されている。そのうちどのコードを使うかについて、国数の最も多いものを用いることにすればよいか？

具体的には、GNPは99a. . .

GDPは99b. . .

を用いてよいか？

またGNP（GDP）がないときにGDP（GNP）で代用して差し支えないか

A GNP、GDP（IFSの場合）は、おおくの場合各国の通貨単位になっている。

したがって、自国通貨建ての表示の場合と米ドル建ての表示の場合によって処理を変える必要がある。

自国通貨建てであるかどうかは、通貨コードを見て判断する。

また、1985年コンスタント・プライスのものが別系列（99a. r、99b. p）にある。

① IFSの系列コード：

GNP・・・99a (原則)

99a, c (西独)

GDP・・・99b

② 米ドル建て換算には、IFSの系列コードRFを使用する。

米ドル建てGNP = (GNP系列) ÷ RF

RF: 米ドル換算レート (自国通貨/米ドル、期中平均)

Q8' . ①コンスタント・プライスの意味と用途は?

②通貨コードとは、単位コードUNOのことか?

③IFSのマニュアルでは、全て自国通貨建てで記載されている。米ドル換算式を使うのは、次の場合と解釈してよいか?

- ・ユーザがドル建てのGNPを知りたいとき
- ・例えば、輸出額/GNPを導出するときに輸出額が米ドル建てである場合

④GNP (GDP) がないときのGDP (GNP) の代用については、可能と判断してよいか?

A ①各年次の価額評価のGDP/GNPが、99aないし99bで、1985年時点の評価価格に換算したものが99a, rないし99b, p

基準年次 (この場合は、85年) の価額評価にするのは、インフレ等により物価が異なるため、異なる年次のデータを単純に比較することができないため。

一般的には、99a (99b) を使用する。

②通貨単位は、IFSのレコード中の「単位コード」欄に表示されている。

例:

コード

- 1 100万US\$
- 2 100万自国通貨単位

③これでよい

- a. ドル建てのGNPが知りたい時
- b. 複数の系列での計算をする場合に、単位を揃える場合には、米ドル建てになるように為替レートで換算する。

一人当たりのGNP（GDP）のようなケースは、多くの場合米ドル建

④これでよい

参考までに、GNPの公表国は100か国以下だが、GDPの公表国は、120か国ある。

付録 4

検索要求フレームのサンプルケースへの適用

% no.	輸出入	報告	相手	品目	期間	時点	単位	種別	情報1	情報2
ret(1	,出	,n	,n	,n	,n	,n	,n	,影響	,ドル安	,伸び率
ret(2	,出	,アジア勢	,欧州	,n	,n	,最近	,n	,攻勢	,n	,n
ret(3	,貿易	,n	,n	,製造業	,n	,n	,n	,影響	,円高	,n
ret(4	,出	,n	,n	,n	,n	,61	,n	,数量・比較	,中小企業	,全体
ret(5	,出	,n	,n	,製造業	,n	,n	,n	,状況/予測	,n	,n
ret(6	,出	,日本	,n	,n	,n	,1990	,n	,予測	,n	,n
ret(7	,入	,日本	,nics	,n	,n	,n	,n	,状況	,n	,n
ret(8	,入	,わが国	,米国	,製品/サービス	,n	,n	,n	,状況	,n	,n
ret(9	,入	,わが国	,n	,製品	,n	,n	,n	,割合	,gnp	,n
ret(10	,入	,わが国	,発展途上国	,製品	,n	,n	,n	,割合	,n	,n
ret(11	,入	,わが国	,米国	,通信機器	,n	,n	,n	,売上	,n	,n
ret(12	,入	,わが国	,米国	,通信機器	,n	,n	,n	,購入率	,n	,n
ret(13	,出	,米国	,わが国	,ハイテク製品	,n	,n	,n	,n	,n	,n
ret(14	,入(消費)	,わが国	,外国	,製品	,n	,n	,n	,状況	,n	,n
ret(151	,n	,n	,n	,網工業	,n	,n	,n	,生産・前年比	,n	,n
ret(152	,出(出荷)	,n	,n	,網工業	,n	,n	,n	,値	,n	,n
ret(16	,出	,n	,n	,n	,n	,1986	,n	,輸出数量指数・前年比	,n	,n
ret(17	,出	,n	,n	,輸出産業	,n	,n	,n	,状況・コスト削減	,n	,n
ret(181	,出(OEM)	,n	,n	,n	,n	,n	,n	,状況	,n	,n
ret(182	,出	,n	,n	,原材料/部品	,n	,n	,n	,状況	,n	,n
ret(191	,入	,n	,n	,n	,n	,n	,n	,輸入報告額	,n	,n
ret(192	,入	,n	,n	,n	,月	,n	,n	,輸入報告額の前年同月比	,n	,n
ret(201	,出	,n	,n	,n	,n	,n	,n	,輸出確認額	,n	,n
ret(202	,出	,n	,n	,n	,月	,n	,n	,輸出確認額の前年同月比	,n	,n
ret(21	,貿易	,米国	,n	,n	,n	,82_85	,n	,貿易赤字	,n	,n
ret(221	,貿易	,米国	,n	,n	,n	,n	,n	,割合・為替レート	,貿易赤字	,n
ret(222	,貿易	,日本	,n	,n	,n	,n	,n	,割合・為替レート	,貿易黒字	,n
ret(23	,出	,ヨーロッパ	,n	,エアバス	,n	,n	,n	,状況	,n	,n
ret(24	,出	,米国/日本/西独	,n	,工作機械	,n	,1981_現在	,n	,比較	,n	,n
ret(25	,出	,わが国	,n	,米国スパコン関連部品	,n	,n	,n	,輸出货量	,n	,n
ret(26	,出	,韓国	,カナダ	,自動車	,n	,n	,n	,輸出货量	,n	,n
ret(27	,入	,日本	,n	,石油	,n	,1974以降	,n	,輸入量・増減	,n	,n
ret(28	,入	,日本	,中国	,農産物	,n	,n	,n	,輸入量	,n	,n
ret(29	,入	,日本	,インド	,農産物	,n	,n	,n	,輸入量	,n	,n
ret(30	,貿易	,日本	,発展途上国/共産国	,累積債務国	,製品	,n,n	,n	,関係	,n	,n
ret(31	,貿易	,n	,n	,n	,n	,n	,n	,貿易収支	,不均衡	,n
ret(32	,貿易	,西独	,n	,n	,n	,n	,n	,貿易黒字	,n	,n
ret(331	,貿易	,日本	,n	,n	,n	,n	,n	,貿易黒字・比較	,gnp	,n
ret(332	,貿易	,西独	,n	,n	,n	,n	,n	,貿易黒字・比較	,gnp	,n
ret(34	,貿易	,日本	,n	,n	,n	,n	,ドル	,貿易黒字・増加率	,n	,n
ret(35	,出	,日本	,米国	,n	,n	,n	,n	,数量	,n	,n
ret(36	,入	,n	,米国/アジア	,輸入品	,n	,n	,n	,価格	,銘柄毎	,n
ret(37	,入	,n	,中南米	,コーヒー	,n	,n	,n	,輸入量	,n	,n
ret(38	,入	,n	,nics諸国	,鋼材	,n	,n	,n	,動向	,n	,n
ret(39	,入	,日本	,n	,自動車産業	,n	,n	,n	,割合	,輸出産業	,n
ret(40	,出	,わが国	,ヨーロッパ	,家電メーカー	,n	,n	,n	,状況	,n	,n
ret(41	,入	,日本	,湾岸諸国	,石油	,n	,n	,n	,依存度	,n	,n
ret(42	,出	,太平洋地域	,n	,石油	,n	,n	,n	,状況	,n	,n
ret(43	,入	,日本	,n	,製品	,n	,87	,n	,増加率	,n	,n
ret(44	,入	,n	,東南アジア	,製品	,n	,n	,n	,傾向	,n	,n
ret(45	,入	,n	,n	,製品	,n	,87	,n	,伸び率	,n	,n
ret(46	,入	,n	,n	,ビデオメーカー	,n	,n	,n	,動向	,n	,n
ret(47	,貿易	,日本	,西独	,鉄鋼/造船	,n	,n	,n	,関係	,n	,n
ret(48	,貿易	,日本	,西独	,精密機器	,n	,n	,n	,関係	,n	,n
ret(49	,貿易	,日本	,西独	,電器機器	,n	,n	,n	,関係	,n	,n
ret(50	,貿易	,日本	,西独	,輸送機器	,n	,n	,n	,関係	,n	,n
ret(51	,出	,日/米/西独	,n	,n	,n	,85	,n	,比重/国間比較	,投資関連財	,n
ret(52	,入	,日/米/西独	,n	,原材料	,n	,85	,n	,割合/国間比較	,全輸入額	,n

2. The second part of the document focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides guidelines for effective communication, such as using appropriate language, listening actively, and providing feedback. It also discusses the benefits of open communication, including improved collaboration and decision-making.

付録 5

貿易統計に関わりの深い経済統計用語

属性の取り得る値 (基本名詞)	基本名詞に対する類義語
貿易依存度	類) 輸出依存度と輸入依存度
輸出依存度	式) $\text{輸出} / \text{GDP又はGNP} \times 100 (\%)$
輸入依存度	式) $\text{輸入} / \text{GDP又はGNP} \times 100 (\%)$
貿易収支	式) 輸出 - 輸入
製品輸入 (出) 比率	式) $\text{製品輸入 (出)} / \text{総輸入 (出)} \times 100 (\%)$ 類) 工業品輸入 (出) 比率 輸出の工業比率
水平分業度指数	式) $\frac{ \text{輸出} - \text{輸入} }{\text{輸出} + \text{輸入}}$ 類) 水平分業度 輸出入依存係数 貿易特化係数
外国貿易比率	式) $(\text{輸出} + \text{輸入}) / \text{GNP or GDP}$
輸入の所得弾力性	式) $2 \text{ 時点間の輸入増加率} / 2 \text{ 時点間のGNP増加率}$ GDP
輸入の価格弾力性	式) $2 \text{ 時点間の輸入増加率} / 2 \text{ 時点間のGNP増加率}$ GDP
交易条件	式) 輸出価格指数 / 輸入価格指数 類) 交易条件指数
産業内分業指数 (加-藤・ロイの定義)	式) $\frac{(\text{輸出} + \text{輸入}) - \text{輸出} - \text{輸入} }{\text{輸出} + \text{輸入}} \times 100$ 類) 水平分業度指数
貿易特化係数	式) $\frac{\text{輸出} - \text{輸入}}{\text{輸出} + \text{輸入}}$ 類) = 競争力係数
i 国 k 財の RCA 指数 (類系的比較優位指数)	式) $\frac{i \text{ 国 } k \text{ 財輸出}}{\text{世界の } k \text{ 財輸出}}$ $\frac{i \text{ 国の総輸出}}{\text{世界の総輸出}}$
i 国と j 国の輸出結合度	式) $\frac{i \rightarrow j \text{ 国への輸出}}{i \text{ 国の総輸出}}$ $\frac{j \text{ 国の総輸入}}{\text{世界の総輸入} - i \text{ 国の総輸入}}$
貿易不均衡指数	式) 一定の国グループについて 貿易収支の標準偏差 / 総貿易額
製品輸入	SITC分類の5類から8類までの合計 または

	概況分類の 化学品+金属品+機械機器+雑製品の合計
市場別輸出 (入) 比率 (%) マーケット・シェア	
商品別輸出 (入) 構成比 (%) プロダクト・シェア	
前年同期 (月) 増減率 (%)	
原油輸入比率	$(\text{原油輸入} \div \text{輸入総額}) \times 100 (\%)$
輸出入ギャップ率	$(\text{輸入総額} - \text{輸出総額}) / \text{輸入総額}$
GNP (米ドル建て) GDP (米ドル建て)	
一人あたりのGNP (\$建て) 一人あたりのGDP (\$建て)	
実質実効為替レート	$(\text{MERM} / \text{卸売り物価指数}) \times 100 (\%)$
貿易収支戻 (国際収支ベース)	(国際収支項目) より計算
貿易外収支戻	(国際収支項目) より計算
経常収支戻	(国際収支項目) より計算
長期資本収支戻	(国際収支項目) より計算
総合収支戻	(国際収支項目) より計算
外貨準備 (金を除く)	(国際収支項目) より計算

注 類) は類義語
式) は定義式

付録 6

利用目的別ユーザニーズアンケート

(1) 研究業務

特徴

パイロットスタディを繰り返して実際の検索の方針を決定する。
データを眺めながら繰り返し分析作業を行う。
1系列のみで検索を終了する場合はまれ。
レスポンスよりも、何ができるか・・・が重要。



必要な機能

データの特性に関する情報の提供。
検索データのAvailabilityに関する支援。
検索データを加工し易い形式で、SAS またはLOTUS 用のファイルとして提供する。



プロトタイプの前張項目

検索データ画面が表示されている状態で、一定のコマンドまたは、マウスのクリックで、そのデータ特性を示す画面が表示されるようにする。

例えば貿易データであれば

- (1) FOB 価格か CIF 価格かの区別
- (2) Total Export、Domestic Export、Re-export の区別
- (3) データの出所 (出版物名及び通関統計かその他の区別)
- (4) データがavailable な期間

などが表示されればよい。

検索データの存在状況に関する情報の提供。

例えば貿易データの検索の場合メニュー画面で、①国、②データの種類の、③期間を指定するマスがある場合で、メニュー画面の①に例えば?と入力すると②の条件で示したデータがある国のリストを別画面に表示する。

①を入力して②を?としたらDBの特定国データのリストを別画面に表示する。

(データの存在を本で確認する手間が省略できる)

また、加工計算をする場合その演算が可能か否かは、DB上のデータの制約による(世界計を足し上げて作りたいとき、共産圏のデータがない・・・等)、どのような制約条件のもとで、結果が得られるかを、コメントDBとして付加することが望ましい。

多種のデータの加工後、LOG OFF するまでは、そのデータを指定する変数名で、テンポラリディスクにコマンド一つで保存できるようにする。さらにべつの計算メニューで、それらを出し、計算し、結果の表示をする。

例えば計算メニュー画面を選択すると次のような画面が現れるようにする。

使用変数名	A :	ユーザが定義した変数が追加されてゆく
	B :	
	.	
	.	
	.	
計算式: $Z = (A \times B) - C - D$		

さらに、加工演算により必要とされるデータの種類、特性、Availabilityを表示し、さらにこれを上記コメントDBにリンクして、計算可能性に関する情報を利用者に与える。

条件文の保有

次回以降の検索・加工のために条件文を保有しておくことは、大切であるが、IRSのようなメタデータも提供するようなシステムでは、どのような形で、キープしておくのがよいかは、システムの使われ方による。(入力されたコマンド等をすべて保存したのでは、実用的でない)。メタデータの参照等は、印をつけて保存し

、キープされた条件を再実行する時に印のついたものは、スキップするなどの工夫が必要。なおSASでは、ログを再度コンピュータで、実行できるようになっている大量データの使用に耐える性能の確保。

大量データの処理を伴い、時間のかかるものについては、バッチジョブに流すようにする等の工夫があるとユーザにとってありがたい。

他システムへのデータの変換

例 AIDX から SAS へ

① YEAR EXP/IMP R/CNTR P.CNTRX VALUE
62 1 1 1 5634

一般的だがSAS
では、使い難い

② JPN62EXP USA62EXP JPN62IMP USA62IMP
3400 6281

変数に関する指示が
必要

(2) 調査業務 (定例資料作成)

特徴

毎回同一処理を行う。

前回と同じ処理をおこなえばよいかの判断が必要。

前回とくらべ、結果がどのように変化したかの判断が必要。

報告書としての体裁を整える必要がある。



必要な機能 (支援)

処理手順の登録による操作の簡便化。

データの属性等の変化に関する管理。

加工結果の前回結果との変化管理。

美しい結果出力。



プロトタイプの拡張項目

処理手順の登録機能。

操作手順を指定によりファイルに登録し、次回はこのファイルを指定するだけで同じ操作ができる。また、一部のパラメータ (時点等) を指定時に変更できる。さらにこのファイルを編集して新たなファイルとして登録できる。

前回と同じ処理をおこなえばよいかの判断。

データの存在、収録期間の変更、出所の変更、期別の変更、性格 (名目・実質、期別集約の方法、季節調整の有無) を前回と比較して変化があった部分については、警告をだし、処理を変更するかどうかユーザに問い合わせる。

前回との結果の評価の判断。

前回の結果と比較して著しい違いがあるものについて警告を出す。比較内容は変化率とする。

結果の編集。

グラフ表示についても処理手順を登録しておき自動的に実行する。ただし、表示内容について最適化が利用者の判断でできるような余地はのこしておく。

(2) 調査業務（通常業務）

特徴

定期的報告書の作成。

毎回略同一のフォーマットで最新値を使用した作表。

担当範囲の監視。

自分の頼りとする指標について最新値を整理してファイリングする。

使用する統計の範囲はそう頻繁には変わらない。



必要な機能（支援）

自動化。

担当者自身が起動をかけなくても帳表類を自動的に出力する。

体裁の編集支援。

上記の定期出力帳表類の体裁編集支援。（統計データを見ながら体裁を試行錯誤的にチェックしつつ編集する。ただし、頻度は極めて少ない。帳表は、そのまま報告書に使用できる体裁であること）

警告／強調等、傾向の変化に関する自動検出。（ただし、重要性は乏しい）

統計の属性変更に対する監視／チェック／及び対応の支援。（ただし、担当者のほうが詳しい場合もあるので、発生の警告のみでも充分）

担当範囲の管理機能。（自分が使用している統計に関する管理機能）

自分流の分類体系、インデックス等の追加的付加等。

定義済帳表の管理（ネーミング／配布先管理／出力実績管理等）



プロトタイプ of 拡張項目

処理手順の登録と自動化及び登録内容の管理機能。

出力形式の登録と自動出力。

定義・登録：①出力形態の定義・登録、②実際のデータを使って表示してみながらの編集、③出力タイミングの定義等。

自動出力：更新タイミングチェックして、丁度よいところで、自動的に出力されること。（ユーザが忘れていても出力される）

登録内容の管理機能。（管理帳表乃至はブラウジング機能）

個人ディレクトリの作成／管理機能。

使用統計系列の定義：定義された帳表とは別に、使用している系列の登録と定義また、帳表の登録時にも自動的に派生されることが望ましい。

ここで、一括、表示単位・ラベル等が定義できるとよい、

アライアスの使用を可能とすること：自分流の分類体系に従ったコードの定義登録及びそれらの検索を可能とすること。

☆追加すべき事項としては、担当者の担当範囲に関する知識とそれによるガイドがあると便利である。

(3) 調査臨時業務

特徴

求める情報の所在の所在調査からおこなう。
データの抽出・検索方法の決定を行う。
複数のデータベースからとった情報を統合する。
データの比較を行う。
試行錯誤的な検索。
与えられたテーマについての調査。
テーマは、データベースの分類体系と一致しないことがある。
書誌的な情報が必要になることがある。



必要な機能（支援）

所在情報の提供。
データ検索のガイド。
情報統合上の注意等支援。
比較の正当性支援。
複数分野間の移動。
テーマと統計やDBの分類体系との対応付け。



プロトタイプの拡張項目

所在情報に関する知識。
利用者の指定する用語について、世に存在する情報と現在データベースに存在する情報を区分けしてガイドする。データベースに存在しない情報についてはどこへいけば入手可能かを表示する。
複数のデータベースからとった情報の統合に関する支援。
収録期間、期別、性格（名目、実質、季節調整の有無）から比較可能かどうかを判断し、警告を出す。
複数要求種別・フレーム間の移動。
検索・加工の過程が密接・連続的に推移し、且つこれば何度も簡単にトライできる（特に銀行・証券の産業調査担当者を想定）
別分野へのスムーズな移動。
検索履歴の保持。
適切なDBの選択。

(4) 企画業務

特徴

企画立案の背景を捉える。
企画立案の結果を検証する。
特殊なデータベースを利用する場合がある。
意見、考え方等を裏付けるために検索をおこなうことがある。
既存の雛形に改良を加えたい場合がある。
国際比較等を念頭において、情報収集する。



必要な機能（支援）

目的とする調査結果は、どのように評価したらよいか。
評価基準としては、なにを使えばよいか。
広範な領域に関するアラカルトメニューの用意。
過去の事例のガイド。
データの定義等が変わっていないか確認。
比較に関する支援。
標準的な分類基準、慣例の参照。



プロトタイプの拡張項目

結果の評価。

結果の評価について一般的な基準、及び過去に公表された事例等の情報を利用者の要求に基づき表示する。
最近の輸出入動向等を状態として用意し、グラフ作成までを行う。
過去の事例に関する知識を持つ。（使用上の注意も含む）
データの特性に関する情報の提供（精度、出所等）特に時系列として扱う場合。
比較の正当性、可能性に関する知識。
複数の検索要求の入力。
データの定義、注記の表示。
検索要求の保存。

参 考 文 献

- (1) 穂鷹良介、佐藤英人 「統計データベースの設計と開発」
ーデータモデルと知識ベースの応用ー オーム社 1988年
- (2) (財)日本情報処理開発協会 「情報処理技術の応用に関する調査研究報告書」
ー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ー 1989年
- (3) 通商産業省 「通商白書」 1989年
- (4) 日本関税協会 「外国貿易概況」 1988年7月
- (5) 日本関税協会 「実行関税率表」 1988年
- (6) 佐藤英人 「統計情報に関するメタデータの在り方について」
DATABASE FORUM Vol.2 No.1 pp.3-15 1988年4月

1. The first step in the process of the development of a new product is the identification of a market need. This is often done through market research, which can be conducted in a number of ways, including surveys, focus groups, and interviews with potential customers.

2. Once a market need has been identified, the next step is to develop a concept for the new product. This involves creating a detailed description of the product, including its features, benefits, and target market.

3. The third step is to conduct a feasibility study. This involves assessing the technical, financial, and market viability of the product concept.

4.

5. The final step in the process is to launch the product into the market. This involves creating a marketing plan, developing a distribution strategy, and implementing the product.

—— 禁 無 断 転 載 ——

平成 2 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

