

# 専門用語研究

Journal of the Japan Terminology Association

No.15 1998-03

---

「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から

|                                |       |    |
|--------------------------------|-------|----|
| 「専門用語研究」としての「専門用語」研究 .....     | 影浦 峽  | 1  |
| データ共有と発掘の新展開 .....             | 西尾章治郎 | 8  |
| テキストからの情報・知識の獲得と自然言語処理 .....   | 辻井 潤一 | 14 |
| JICST シソーラスの改訂 .....           | 永井 賢吉 | 20 |
| 印象語の地図——マルチメディア情報の感性検索—— ..... | 加藤 俊一 | 25 |
| ISO/TC37 におけるターミノロジーの原理 .....  | 山本 昭  | 29 |
| 地図学用語辞典〔増補改訂版〕発刊の経緯——地図用語専門部会の |       |    |
| 発足から現在まで—— .....               | 坂戸 直輝 | 32 |
| 国内刊行用語辞典リスト (1997) .....       |       | 37 |
| 国内刊行用語辞典リスト 電子版 .....          |       | 45 |
| JIS用語規格制定の動き1997 .....         |       | 47 |
| 編集後記 .....                     |       | 51 |

## 第11回専門用語シンポジウム

# 知識の自己組織化と専門用語

標記シンポジウムが1997年11月14日、日本学術会議講堂にて、約60名の参加者を得て開催された。今回は、日本学術会議（学術文献情報研究連絡専門委員会）、情報知識学会、情報処理学会（情報学基礎研究会）との共催である。発表は、下記のように招待講演4件と一般講演3件の合計7件であった。

趣旨：インターネットの利用拡大により、国際的なレベルでの情報流通が年々急速に増大している。その結果として、局所的には体系化されているが大局的には未整理に近い大量の情報が増大しつつある。電子化されているが故に、印刷媒体で培われた伝統的な情報組織化手法は、量的かつ質的観点から適応困難であり、知識の自己組織化の研究が今後の情報科学において主要な研究テーマになると考えられる。

知識の自己組織化の研究において、知識を構成する中心要素である概念、およびその表現・伝達媒体である用語、特に専門用語の研究が重要かつ急務であると考えられる。また、究極の知識自己組織化システムと位置づけられる、概念の記憶・処理システムである脳の研究から上述の問題を捕らえることは極めて有用であると考えられる。

本シンポジウムでは、上述の問題に関連の深い研究者諸氏による最新の研究について報告、および討論を行う。

|      |                         |           |       |
|------|-------------------------|-----------|-------|
| 招待講演 | 脳概念学習と世界像の形成            | 大阪大学      | 中野 馨  |
| "    | データ共有と発掘の新展開*           | 大阪大学      | 西尾章治郎 |
| "    | テキストからの情報・知識の獲得と自然言語処理* | 東京大学      | 辻井 潤一 |
| "    | 印象語の地図—マルチメディア情報の感性検索—* | 中央大学      | 加藤 俊一 |
| 一般講演 | 専門用語研究としての専門用語研究*       | 学術情報センター  | 影浦 峽  |
| "    | J I C S T ソーサスの改訂*      | 科学技術振興事業団 | 永井 賢吉 |
| "    | ISO/TC37におけるターミノロジーの原理* | 関東短期大学    | 山本 昭  |

本号では、各発表者に修正・加筆をいただいた3件（西尾、影浦、永井）と予稿集からの転載3件（辻井、加藤、山本）を掲載する。

～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

## 「専門用語研究」としての「専門用語」研究 The Study of Terminology as 'Terminology'

影浦 峽\* KAGEURA Kyo

### Abstract

It is said that the terminology in modern language phenomena is gaining in importance recently, and various studies which treat terminology as their target have appeared, both 'theoretical' and application-oriented. However, the common framework within which various different viewpoints are compared or discussed is yet to be consolidated. Against this background, the paper reviews the major trends of the study of terms or terminology from the point of view of their theoretical status as 'terminology', in order to clarify the different viewpoints adopted in the different trends of the study of terms.

### 1. はじめに

様々な専門分野が細分化された現代世界において、専門分野のコミュニケーションの核を担う専門用語が言語現象の中でも重要な位置を占めるといふ認識のもと、専門用語を巡る研究が活発に行われるようになってきた。それは「理論的」と称するものから、言語処理の応用、さらにその応用としての情報検索を射程に置いたもの等、多岐にわたる。

けれども、そうした多様な研究を相互に参照し、比較したり統合したりすることを可能にするような基本的な視点についての整理はあまり進んでいないように思われる。1980年代から、いくつか「専門用語研究」に関する教科書的なものが現れてはいるが (Felber 1984; Picht & Draskau 1985; Sager 1990; Teresa Cabré 1993)、規範的なものであったり、扱う範囲が限られていたりする。実際、専門用語に関する研究や実務には、言語学を背景とする研究者、ドキュメンテーションに従事する者、計算言語学の研究者等、異なる背景のものが関わっており、統一的な視点はおろか、全体を概観して整理することすら容易ではない。

本稿では、こうした状況を念頭に、主要な研究動向を整理し、「専門用語研究」としての「専門

用語」研究の基本的枠組みを検討することとした。そのために、専門用語を巡る研究を、専門用語そのものを必要十分に限界付け、専門用語の臨界を確定しようとしているものかどうか、という観点から考察することとしたい。むしろ、専門用語を必要十分に限界付けるという目標自体は過度に抽象的で、実際には達成不能な目標であるが、けれども、異なる立場からの諸研究を整理するにあたっては、そうした原理的な問題意識との関係で整理する方がかえって都合がよいように思う。

以下では、最初に、「専門用語とは何か」という点について簡単に整理し、次いで、それを念頭に、これまでの専門用語を巡る様々な研究が、専門用語のどのような側面を扱ってきたかを、代表的な研究を取り上げて検討する。最後に、諸研究の動向を参考にしつつ、専門用語研究としての専門用語研究のためには、どのような枠組みが必要かを考えてみる。

### 2. 専門用語とは何か？

「専門用語とは何か」については、色々な議論がなされている (Condamines 1995; Kageura 1995; Desmet & Boutayeb 1994; Felber 1984; 宮島 1981; Sager 1990; Picht & Draskau 1985; Shelov 1982; Wüster 1959/60)。けれども、それらを改めて検討してみると、専門用語に対してど

\* 学術情報センター研究開発部

のような研究を具体的に進めるか、どのような研究がどのようなレベルで有効かという、研究レベルでの具体化の問題をあまり考えずに議論が進められることが比較的多いように思われる。

これに関連して、「専門用語の定義」として与えられているものも、本来的に専門用語であることの必要十分条件を設定しようと意図された定義なのか、それとも、専門用語を集めて見たときに顕著な特徴として共通に観察されるものを指摘したのか、という違いが明確にされないまま、後者をいつのまにか前者であるかのように議論している場合がしばしば見られる。このような状況が通用しているのは、専門用語の具体的な研究と、「専門用語とは何か」という議論とがきちんと結びついて論じられていないという点にも関係している。

例えば、最も典型的な専門用語の「定義」として、「概念を表すある種の言語単位が専門用語である」というものがある (Felber 1984, 中村 1997)。これは、概念の定義によって端的に誤りであるかトートロジーであるかのどちらかであり、とりたてて専門用語を規定する力を持たない (Kageura 1995)。これに対置されるものとして、宮島 (1980) の立場がある。彼は、専門語を、「それがなにをあらわすかという対象との関係が第一にある」としながら、同時に「これを専門語たらしめるには、やはりそれを主としてつかう人びと、つまり専門家がいないとてならない」とする。いわば、専門用語の内部的な特徴付けによる規定に加え、外部的な条件から専門用語の範囲を確定しようとする立場である。

後者は、曖昧性があるとはいえ、現実に専門用語の範囲を確定する具体的・操作的な条件を与え得るという意味で、より定義として適切であると思われる。Kageura (1995) は、これをもう少し言語データに引き寄せ、専門領域で流通する言語活動の特徴づける語彙的な単位として専門用語を規定している。当然、このように述べるとき、結局は「専門用語とは専門用語と見なされる用語のことである」という見解が前提されることとなる。この立場に基づき、現実的に割り切ったのが、柴田 (1989) の「専門用語とは文部省の学術用語集

の学術用語である」というものであろう。

専門用語が本質的にそのようなものであるならば、それは別に悪いことではなく、出発点としてはさしあたり専門用語は与えられた語彙であると見なさなくてはならない以上、専門用語を作業的に取り出しうる基準こそが重要であるという立場を確認するものに過ぎない。結局の所、「専門用語とは専門用語であるところのものである」という規定から逃げようがないのであれば、それを結果として隠すことになるようないたずらな神秘化をはかるよりも、それを前提としたところから始める方がよい。

こうした立場に立つならば、前節で述べた「専門用語の限界付け」との関係では、とりあえず対象とする専門用語の存在は前提としつつ、それを独立して扱っているか、専門用語以外との差異を強調しているか、といった視点が諸研究の整理の中心となる。さらに、より抽象的な「専門用語とは何か」という問題との関連では、与えられたデータの理論的な位置づけが問題となろう。すなわち、与えられたデータが専門用語そのものなのか、あるいは理論的に想定される母集団の標本となっているのか、また、後者の場合、ではどのような標本となっているのか、といった点である。いずれにせよ、これらの問題は、実際にデータとして現れる専門用語を何らかのかたちで対象とする具体的研究において・を介して立ち現れるものであるため、本稿では、具体的な研究のみを考察の対象とし、「専門用語とは何か」を巡る疑似思弁的な「研究」は、考察の対象から除外する。

### 3. 専門用語を巡る諸研究

専門用語を巡る研究は、大ざっぱに言って、処理を目的とするものと記述を目的とするものに区別することができる。処理に関する研究は、主として自然言語処理を中心とする計算機よりの背景を持つ研究者によってなされてきた。一方、記述を目的とする研究は、主に言語学系の研究者が行ってきたといえる。ドキュメンテーションを背景とするものは、具体的なソース等構築などの成果を挙げているが、その産物を記述そのものと見なすならば後者に属するものと言うことができ

よう。以下では、処理を目的とするもの、記述を目的とするものそれぞれについて、大きな研究の流れを整理することとしよう。

他に、言葉の上では「理論的な研究」という言葉を立て得るが、今のところ、いわゆる理論的な研究はないといってよい。逆に、専門用語の研究は、きわめて具体的な対象を前提とする以上、処理・記述といった具体的な作業に理論的研究が組み込まれることにならざるを得ないということもできよう。

### 3.1 処理を目的とする研究

処理を主目的とする研究は、「対象としての専門用語」に対して、どのような視点に立つか、あるいは逆に対象としての専門用語を研究においてどのように位置づけるかという構えの観点から、大きく三つに分けることができる。すなわち、

- (A1) 与えられた専門用語を出発点として、その内部的な処理を行うもの
- (A2) 専門用語を主な処理対象としているが、別に対象が専門用語であるかどうかは細かく問わないもの
- (A3) 専門用語を、それ以外の要素から区別することそのものを処理の目的とするもの

である。ただし、(A1)と(A2)の区別は実際には曖昧であり、(A1)は、例えば、専門用語データはその構造が比較的明確であるといった点を積極的に強調しているといった程度の差である。

(A1)および(A2)に相当する研究としては、例えば、頼、王&藤原(1994)の研究がある。これは、与えられた一定規模の、かつまとまりをもった専門用語の語彙を対象として、そこから、再帰的に文字列をパターンマッチさせ、それによって部分文字列が一致する異なる用語間の関係を取り出し、語彙を組織化しようとするものである。同様の系列に属する研究として、原田他(1988)がある。

森脇、河辺、辻井(1996)や影浦(1997a)、Kageura(1997a)は、やはり一定規模の、かつまとまりをもった語彙を対象とし、そこでの内部

的な情報を用いて、複合漢字列の単位切りを行っている。これらの研究では、特に専門用語を対象とすることを明確に前提としているわけではないという意味で(A2)に属する研究と言えるが、与えられたデータに対して適応的に学習する手法を適用していることの前提として、分野依存性の高い言語要素の処理という観点が存在している。Tsuji & Kageura(1997)は、一定の規模・まとまりを持つ、日英の対訳専門用語リストをもとにして、日本語部分を分割し、同時に語構成要素レベルでの日英対訳情報を取り出し、さらに将来的には語構成規則の対応関係を抽出しようというものである。

これらの研究においては、暗にあるいは陽に、一つの領域の専門用語語彙は一般語等よりも体系的であるという仮定を用いているように思われる。そのような意味では、非常に抽象的な意味で専門用語の特徴を用いている。また、特に頼、王&藤原(1994)の処理の結果やTsuji & Kageura(1997)の目指すものなどは、用語集合に見られる、何らかの内部的な規則性に関する記述を、自動的な処理によって取りだそうという方向性が強いように思われる。

これに対して、(A3)の、専門用語を、それ以外の要素から区別することそのものを処理の目的とする研究が、1990年代以降色々と行われるようになってきた。いわゆる、自動専門用語抽出に関する研究である。例えば、Bourigault(1992)、Daille, Gaussier & Langé(1994)、Enguehard & Pantera(1994)、Frantzi & Ananiadou(1995)、Justeson & Katz(1995)、Kageura, Tsuji & Takasu(1996)、Nakagawa(1997)などがある。これらについてのまとまった概観はKageura & Umino(1996)、および影浦(1997b)にある。

自動専門用語抽出の研究は、そもそも現実のテキストから、専門用語であるものと専門用語でないものとを区別することを目的としているため、そこで採用された処理のための尺度は、何らかのかたちで専門用語集合の境界を反映していると考えられる。その意味で、自動専門用語抽出処理の「手法」は専門用語と非専門用語をテキスト中の出現において区別する「理論」の現実的

反映として位置づけることが可能である。しかしながら、残念なことに、現状の自動専門用語抽出を見た場合、以下のような理由で、現実に用いられる手法と背景にあるべき理論との距離は未だかなり遠い。

- ・言語的な手法においては、単に通常の統合的なレベルで専門用語を構成しうる典型的なもの(名詞・複合名詞)を取ってくるという、間接的なものが多く、言語的のどのようなレベルで専門用語の専門用語としての特徴が抽出できるかは検討されていない。これに関しては、通常の言語的記述の枠組みでの専門用語の一般語からの差異化はそもそもできないと考えるのが適切かもしれない。
- ・数量的な手法において、処理対象となる要素の分布に対する理論的な仮定がほとんどないばあいが多い。このため、数量的な手法の結果を、背景になる言語現象における確率的なモデルとの関係で理論的に解釈し、精緻化するという手続きがとれない。
- ・さらに、確率的なモデルとの関係で統計的な議論が仮にできたとしても、そうしたモデルが、「与えられたテキスト／出現用語」のモデルなのか、テキストにおける異なるレジスタを包括する「分野の談話母集団／分野の専門用語」の関係なのか曖昧であり、ほとんどの場合、前者に属するものが多い。

### 3.1.1 記述的な研究

記述的な研究に関してもやはり、大まかに三つに区別することができよう。すなわち、

- (B1) 与えられた専門用語を出発点として、その内部構造を記述するもの
- (B2) 異なる分野の専門用語集合における差異を記述するもの
- (B3) 専門用語と一般語との相違を記述するもの

である。このうち(B1)と(B3)は、処理の研究の(A1)と(A3)とにそれぞれ対応する。

(B1)に属する典型的な例として、専門用語

の語構成パターンの記述があげられる(Pugh 1984; Kageura 1993)。これらの研究の問題は、記述の結果として現れた規則が、専門用語を特徴づけるものなのか、たまたま専門用語を対象として分析したより広範囲な現象(例えば語構成については語構成一般)を特徴づけるものなのか、その区別がどこにあるのが不明確になってしまうことである。単純に与えられた有限のデータを博物学的に並べ上げるのではなく、語構成における規則性を通常の語構成記述の形式的枠組み(語構成要素の一般化・要素間の関係の一般化・統合規則による用語全体の一般化)で一般化して記述しようとするならば、論理的必然として、この問題は起こってくる。すなわち、語構成規則は、どんなに具体的な語彙に依存させようと試みても、やはり一つの生成的な体系を構築してしまい、その体系においては、具体的な語彙を現実的に限界づけるメカニズムはそのままで持ち得ない、ということになる。Kageura (1997b)は、この問題に対する対策の可能性を論じたものであるが、上手く行っているとは言えない。

言葉を変えて言うならば、語構成規則は、いかに専門分野依存のかたちで記述しようとも、やはり許容され得るパターンの記述となってしまう、専門用語として許容されるパターンの記述とはならない。前者が言語体系的な領域に属するとするならば、後者はいわば現実的存在可能性とでも言うべき、言語の現実態に属する問題であり、なおかつ、専門用語は、前節の議論に従ってここでの言葉を適用するならば、その性質上、徹頭徹尾言語の現実態において認定される要素なのだから、現実的存在可能性を臨界づける記述の枠組みがない限り、専門用語に典型的に見られる(けれども一般語にも見られるかも知れない)一つの性質を記述することは可能でも、専門用語を一般語から区別しうる性質の記述は可能ではないということになる。

(B2)に属する研究としては、宮島(1981)や、Ishii(1987)などがある。宮島(1981)はいわば、質的な違いを重視した研究であり、Ishii(1987)は量的な構造に着目している。これらは、現実と与えられたデータに反映されている限りで

の分野間の差異を明確にしているという点で、専門用語一般という上位世界を仮定した中で分野依存の用語の性質そのものを見ようとしているものと言える。けれども、理論的かつ技術的な問題として、用いているデータが全てなのかどうか、もしそうでないとするならば、利用したデータから得られた記述結果がどのように一般化可能であるか、といった点に問題が残っている。

(B3)の、専門用語と一般語との相違を記述するものについては、著者は寡聞にしてそれ自体を主要なテーマとした研究を知らない。宮島(1981)や Sager, Dungworth & McDonald(1980)、石井(1997)に、こうした側面が見られるが、いずれも、両者の境界を定めようとするものではなく、主として専門用語の視点から、一般語の役割や一般語との距離を検討するものである。

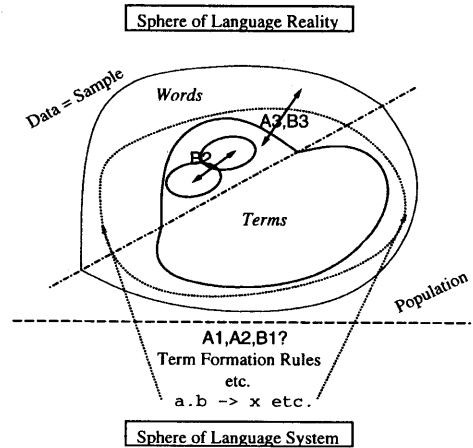


図1 専門用語諸研究の位置づけ

### 3.2 まとめ

「専門用語とは何か」という問題に対して理論的な回答を具体的に・手続き的な基準を伴ったものとして与え得るか、という観点から、これまでに見てきた研究の流れを整理してみよう。

- ・専門用語の内的特徴付けに関する(を用いた)研究、すなわち、処理に関する研究の(A1) (と(A2)), および記述に関する研究の(B1)および(B2)は、すでに専門用語が与えられていることを前提としている。そのため、第一段階において、専門用語とは何かという問題に対する答が事前に与えられていることを前提としている。
- ・従って、これらの研究が、なおかつ専門用語とは何かという点に貢献しうるのは、例えば、既に与えられている専門用語データを標本とした、直接には与えられていない理論的な専門用語の母集団に対する特徴付けを構成している、といった枠組みのもとである。ところが、ほとんど常に、このような解釈を許容しようとする専門用語の記述の体系は、専門用語母集団の範囲を越えて適用可能なものとなっている。

- ・逆に、専門用語と非専門用語とを区別しようとするような研究((A3, B3))は、専門用語を含む何らかのデータが与えられていることを前提としており、その意味では、専門用語とは何かという問題に対する回答が事前には与えられていないという前提にたっている(ただし、評価可能である場合は常に回答が別途与えられざるをえない)。その意味では、これらの研究で得られた知見は、専門用語の内的特徴付けに先行する。
- ・けれども、これらの研究が対象としてきたのは、これまでのところ、あくまで現実に与えられた言語データに現れる専門用語であり、それを標本とする理論母集団としての専門用語は扱われていない。

以上の点を、いささか強引に図式化したのが図1である。現実態と体系との区別とそれに対する専門用語の位置づけに関しては、ここで性急に導入したが、これについての詳しい議論は、影浦(1995)を参照していただきたい。

### 4. おわりに

「専門用語とは何か」という問題に対する具体的な貢献が論理的に可能でありうるような枠組み

を持った研究を、いささか勝手に「専門用語研究」であるところの「専門用語」研究とするならば、そのような研究は、以下の(いずれかの)ような要件を満たしていることが必要であろう。

- 専門用語と非専門用語との差異。これによる特徴付けなしに、専門用語の特徴付けを専門用語の特徴として言いきるかたちで行うことは、トートロジーにならない限り非常にむずかしい。残念ながら、人参をキャベツと区別できるように、専門用語を一般語とを区別することは簡単ではない。これは、後者における両概念が本来的に異なる位相で認定されているものだからである。
- 出発点として与えられた専門用語集合が標本であるような母集団の意識。これは、理論的には標本から母集団へと向かう研究の基本であるが、現実的には、具体的に母集団の性格を明確にするために役立つというよりも、むしろそのもとの標本の適用範囲の現実的な確定により大きく関わってくるであろう。

むしろ、「主に」専門用語を対象とした処理の研究や記述の研究があってもかまわないし、それは専門用語を巡る研究の活発化という点からは望ましいものであろう。むしろ戒めるべきは、専門用語を巡る具体的な研究(専門用語に必要十分なものであれ、必要なものであれ、十分なものであれ)を進展させる契機とならないような、専門用語に関する「思弁的研究」を繰り返すことである。前述したように、少なくとも現実的な出発点としては、「専門用語とは専門用語であるところのものである」という規定から逃げようがないならば、それをもとにした具体的・分析的な研究を進めることこそが「理論」的に重要なのであり、「専門用語研究」としての「専門用語」研究は、そうした具体的な研究における理論的布置の問題として議論されるべきであろう。

#### 参考文献

**Bourigault, D. 1992.** 'Surface Grammatical Analysis for the Extraction of Terminological

Noun Phrases.' *Proceedings of COLING92* 977-981.

**Condamines, A. 1995.** 'Terminology: New Needs, New Perspectives.' *Terminology* 2(2), 219-238.

**Daille, B., Gaussier, E. and Langé, JM. 1994.** 'Towards Automatic Extraction of Monolingual and Bilingual Terminology.' *Proceedings of COLING94* 515-521.

**Desmet, I. and S. Boutayeb. 1994.** 'Terms and Words: Propositions for Terminology.' *Terminology* 1(2), 303-325.

**Enguehard, C. and Pantera, L. 1994.** 'Automatic Natural Acquisition of a Terminology.' *Journal of Quantitative Linguistics* 2(1), 27-32.

**Felber, F. 1984.** *Terminology Manual*. Paris: Unesco and Infoterm.

**Frantzi, K. T. and Ananiadou, S. 1995.** 'Statistical Measures for Terminological Extraction.' *Proc. of the Third Int'l Conference on Statistical Analysis of Textual Data*. 297-308.

**原田隆史他. 1998.**「複合語の解析による語の上位下位関係の自動抽出についての基礎的研究」『三田図書館・情報学会1988年度研究大会予稿集』49-52.

**Ishii, M. 1987.** 'Economy in Japanese Scientific Terminology.' In Czap, H. and Galinski, C. (eds.), *TKE'87: Terminology and Knowledge Engineering*. Frankfurt: Indeks Verlag. 123-136.

**石井正彦. 1997.**「専門用語の語構成」『日本語学』16(2), 21-30.

**Justeson, J. S. and Katz, S. M. 1995.** 'Technical Terminology: Some Linguistic Properties and an Algorithm for Identification in Text.' *Natural Language Engineering* 1(1), 9-27.

**Kageura, K. 1993.** *A Conceptual Analysis of Japanese Complex Noun Terms with Special Reference to the Field of Documentation*. PhD Thesis, University of Manchester.

**Kageura, K. 1995.** 'Toward the Theoretical



- Study of Terms.' *Terminology* 2(2), 239-257.
- 影浦 峽. 1995. 「語」と「専門用語」—専門用語に関する理論的研究へ向けての試論—『学術情報センター紀要』7, 217-229.
- Kageura, K. 1997a. 'Type-based and Tokenbased Learning of Kanji Morphemes.' *Proceedings of the Third International Conference on Quantitative Linguistics*. 146-151. (<http://www.rd.nacsis.ac.jp/~kyo/papers/alist.html>)
- Kageura, K. 1997b. 'On Intra-Term Relations of Complex Terms in the Description of Term Formation Patterns.' *Mélanges de Linguistique offerts à Rostislav Kocourek*. Halifax: Les Presses d'ALFA ( $\alpha$ ). 105-111.
- 影浦 峽. 1997a. 「文字単位の bigram 尺度を用いた複合漢字列の単位切り手法」『言語処理学会第3回年次大会発表論文集』477-480.
- 影浦 峽. 1997b. 「自動専門用語抽出の諸問題」『整理技術研究』(<http://www.rd.nacsis.ac.jp/~kyo/papers/alist-j.html>)
- Kageura, K. and Umino, B. 1996. 'Methods of Automatic Term Recognition: A Review.' *Terminology* 3(2), 259-289.
- Kageura, K., Tsuji, K. and Takasu, A. 1996. 'Some Statistical Characterizations of Terminological and Non-Terminological Elements: Evaluation and Examination in Japanese Technical Abstracts.' In Galinski, C. and Schmitz, K-D. (eds.), *TKE'96: Terminology and Knowledge Engineering*. Frankfurt: Indeks Verlag. 130-138.
- 宮島達夫. 1980. 「専門語研究の視点」『言語』9(4), 50-56.
- 宮島達夫. 1981. 『専門語の諸問題』東京: 秀英出版.
- 森脇敏, 河辺恒, 辻井潤一. 1996. 「辞書を使わない日本語専門用語の自動分割」『言語処理学会第2回年次大会発表論文集』
- Nakagawa, H. 1997. 'Term Extraction from Small Texts.' *NLPRS'97 Satellite Workshop on Information Retrieval*. December 1, 1997. Phuket, Thailand.
- 中村幸雄. 1997. 「整理技術研究会」(1997年9月6日)における発言.
- Picht, H. and Draskau, J. 1985. *Terminology: An Introduction*. Surrey: University of Surrey Press.
- Pugh, J. M. 1984. *A Contrastive Conceptual Analysis and Classification of Complex Noun Terms in English, French and Spanish*. PhD Thesis, University of Manchester.
- 頼静娟, 王曉晶, 藤原讓. 1994. 「専門用語における階層関係及び関連関係抽出法」『情報知識学会第2回(1994年度)研究報告会講演論文集』p. 9-16.
- Sager, J. C. Dungworth, D. and P. F. McDonald. 1980. *English Special Languages*. Wiesbaden: Oscar Brandstetter.
- Sager, J. C. 1990. *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.
- 柴田武. 1989. 「日常言語と専門用語」『日本語学』8(4), 10-17.
- Shelov, S. D. 1982. 'The Linguistic Nature of the Term.' *Nauchno-Tekhnicheskaya Informatsiya, Seriya 2*, 16(9), 1-6.
- Teresa Cabré, M. 1993. *La Terminología*. Barcelona: Editorial Antartida.
- Tsuji, K. and Kageura, K. 1997. 'An HMM-based Method for Segmenting Japanese Complex Terms and Keywords using Bilingual Corpora.' *Proceedings of NLPRS'97*. 557-560. (<http://www.rd.nacsis.ac.jp/~kyo/papers/alist.html>)
- Wüster, E. 1959/60. 'Das Wort der Welt, Schaubildlich und Terminologisch.' *Sprachform* 3/4, 183-204.

～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

## データ共有と発掘の新展開

西尾 章治郎\* NISHIO Shojiro

概要：最近，WWW (World Wide Web) およびデータ発掘 (data mining) はともに応用上非常に注目を集めているテーマであるが，これらに関連したデータ共有と発掘に関する新たな展開を追いながら，データ表現に関する最近の話題を取り上げ，データベース分野において従来からの諸概念がいかに見直されているかを論じることにする。まず最初に，ネットワークを介してマルチメディア・コンテンツを空間的に広めることを可能にしているWWWをベースとした新たなデータ共有の方法について述べる。つぎに，特定のタスクに関連したデータを集中的にあるサイトに集め，そのなかから有用な情報を収集することを目指しているデータ発掘の研究を通して，従来からのデータ表現モデルが再考されつつあることを述べる。

### 1. はじめに

最近のデータベースシステム分野での研究・開発に関する新たな展開のなかで，WWW (World Wide Web) をベースとしたデータ共有の浸透と，大量に蓄積されたデータを発掘し有用な知識 (規制性やルールなど) を獲得しようという研究の急速な進展が際立っている。

WWW は，元来データベースシステムの分野以外で開発されたものであるが，近年データベースシステムのアーキテクチャに根底から変革を迫り，従来のデータベースクライアント，データベースサーバという2階層 (two-tier) のシステムから，クライアントサイドのWWWブラウザとデータベースサーバの中間にWWWサーバを持つ3階層 (three-tier) のシステムに移行しつつある (図1参照)。また，現在非常に話題になっているイントラネットにおける情報ベースシステムもほとんどがWWWを基盤技術として構築されている。このように，データベースシステムの研究・開発の分野では，最近，まさにWWWという「黒船」が来航した感がある [13]。ここで重要なのは，WWWをベースとしたシステムがデータ

共有という観点からなぜこれだけ多くの人々に受け入れられ広がっているのかを探ることである。

一方，デジタル情報化時代を迎え，コンピュータ技術の著しい発展とハードウェアの低価格化にともない，最近では，多種多様な膨大な量の「生」データが十分に解析されないまま，日々刻々とデータベースに格納されている。ある統計では，20ヶ月ごとに世界の情報量は2倍に膨れ上がり，さらに今後も，医療，商業，経済，科学技術，工業などさまざまな分野のデータベースにおいて，データの質・量の両面から着実な拡張が見込まれる。実際，宇宙関係をはじめ，ヒトゲノム関係のプロジェクト，より日常生活に近いものとしては国勢調査データの格納などのために用いられているデータベースなどは，すでに数百万件から数十億件 (つまり，テラバイト単位) のデータが格納された大規模データベースとなっている。

このように蓄積されているデータは，「知識」の根源である。しかし，データを大量に保有していることと有用な知識を多くもっていることはまったく別のことである。大量な共用データを有効利用するためには，格納されたデータをよく理解し，有用な知識を迅速かつ確に発見する必要がある。このようなデータ発掘 (datamining) により，科学分野における新たな真理の探究，工

\* 大阪大学大学院工学研究科情報システム工学専攻  
E-mail: nishio@ise.ing.osaka-u.ac.jp

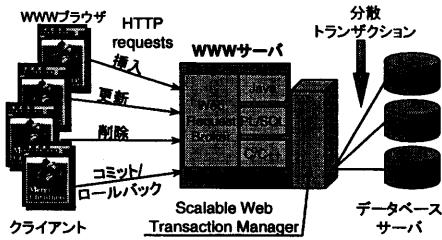


図1 WWWサーバをもつ3階層システム

場の生産性向上のキー・ファクターの発見、さらにはビジネス分野における斬新な商業戦略の立案などが可能になると考えられる。例えば、米国のスーパーマーケットでは、レジでの売上データを蓄積してデータ発掘をしたところ、「ビールと紙おむつ」が同時に売れることがしばしばあることが分かり、それらの二つの商品を近くに配置することにより、販売成績を高めることに成功した例が報告されている。そこで、このようなデータ発掘の研究と深く関連した、例えばデータウェアハウス (data warehouse) の研究[7]などが従来のデータベースの研究にどのような影響を及ぼしているかを探ることは重要である。

本稿では、以上で述べたようにWWWをベースとしたデータ共有の新たな方法の広がり、データ発掘に関連したデータ表現モデルの再考の話題を中心にして、これらの分野の新たな展開を論じることとする。これら二つの分野は、一方はネットワークを通じていかにマルチメディア・コンテンツを空間的に広めて共有していくかを指しており、他方は特定のタスクに関連したデータを集中的にあるサイトに集めて、そのなかから有用な情報を収集することを目指している。そのような意味では両極端の分野でのデータ表現に関する最近の話題を取り上げながら、この分野の従来からの諸概念がいかに見直されているかを論じることとする。

2. データ共有のための記述モデル

WWWは、そのインタフェースが一様で使い易く、ほとんどのマシン環境で使えること、さら

```

<html>
<head><title>Nishio Laboratory Home Page</title></head>
|
|body bgcolor="white"      TEXT  ="black"
|                        LINK   ="blue"
|                        VLINK  ="red"
|<basefont size=3>
|
|<center>
|<a href="/index.html"></a>
|<a href="/index-jp.html">
|</a><p>
|<table border=0 bgcolor="#ffffff">
|<tr align=center>
|<td<font size=3>西尾研究室</font></td>
|</tr>
|</table>
|</center>
|
|<table border=0 bgcolor="#ffffff">
|<tr align=center><td>
|西尾研究室ホームページへようこそ。ここは、
|<a href="http://www.osaka-u.ac.jp/">大阪大学</a><a href="http://www.eng.osaka-u.ac.jp/homepage-j.html">
|工学研究所</a>
|<a href="http://www.ise.eng.osaka-u.ac.jp/">
|情報システム工学専攻</a>
|<a href="/Nishio-Lab/location-jp.html">
|社会情報システム工学講座情報ベース工学領域
|(西尾研究室)</a>です。
|<a href="http://www.vml.org/">VML</a><a href="http://www.vml/virtual-nishio-lab.html">こちら
|</a>もどうぞ。
|</td></tr>
|</table>
|
|</body>
</html>
    
```

図2 HTML記法の例

にブラウザが無料で利用できることなどの長所から爆発的に利用者が増え続けている。さらに最近では、データベースなどへの情報検索におけるクライアント側の最も一般的なアクセスインタフェースとしてWWWブラウザを利用することが定着しつつある。このWWWでデータを提供するために用いられる言語であるHTML (HyperText Markup Language)[1]は、文章管理のための標準化規格であるSGML (Standard Generalized Markup Language)[8]をベースとした記述法であるが、SGMLおよびHTML記述法の普及が企業のデータベース構築に大きな変革をもたらしつつある。

従来、企業などにおいて共用データとしてデータベース化されてきたデータは、給与データを含む人事データや (CAD関連の企業では) CADデータなどが主であり、それらは全データ量の2割に満たないと言われている。ところが、このような状態に対して一つの変革が起きつつある。それは、

企業内の書類を SGML 形式で統一的に記述して書類そのもののデータベース化を推進し、イントラネットを通じて企業内の異なる部・課の間で文書を共有したり、冗長な書類作成などを省いて事務効率を上げようという動きである。さらに、このような先進的な動きは、異なる企業間での共同プロジェクトに代表される仮想企業 (virtual company) を構築する際の共通の情報基盤システムであるエクストラネットの構築にまで発展している。

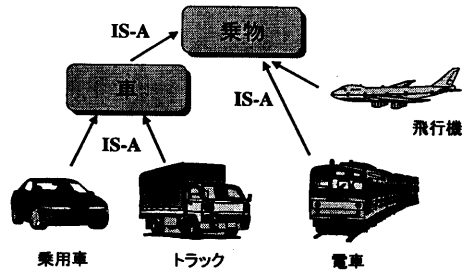


図3 IS-A関係の例

以上のような SGML, HTML などとは、いわゆる「タグ付きデータ表現記法」という点で共通性がある (HTML 記法の例として図2参照)。果たして、これらの記述法がデータ共有のための表現法として十分な能力を有しているのかという疑問がある。その問題を考える際には、データを複数の人間で共有する場合に、どこまで意味 (semantics) 表現を記述法のなかで可能にするかを考えることが重要である。つまり、データ共有のための表現法のなかでは、現実社会を表現する上で誰もが認める概念の範囲でしか意味的な表現を許すことができず、また、データベースシステムの構築技術からすれば、複雑な意味表現を許したことに起因してシステムの応答時間が実用に耐えなくなってしまうことは避けなければならない。

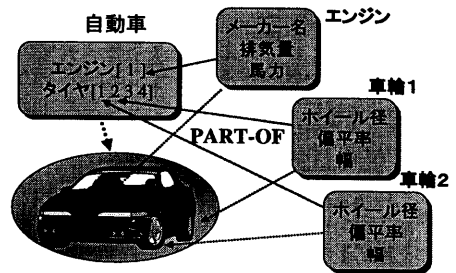


図4 PART-OF関係の例

例えば、近年オブジェクト指向表現によるデータベースシステムの構築技術が急速に進んでいる [10]。しかし、本来ならばオブジェクト間の複雑な関連をも表現してこそオブジェクト指向の概念に基づいていると言えるが、実際は誰でもが認める階層関係を表す「IS-A 関係」(IS-A 関係の例として図3参照) と対象物が何によって構成されているかを表す「PART-OF 関係」(PART-OF 関係の例として図4参照) までの意味表現でとどめることによって、システムの機能するようにしているとも言える。その意味では、SGML, HTML などの記法は、個々人で多様に異なってくる意味論的部分まで立ち入らないぎりぎりのところでデータ構造表現を実現しているという見方をすることが可能である。その観点からすれば、このようなタグ付きの表現が普及する理由が明らかになってくる。

WWW の普及に関しては、上記のような要因に加えてもう少し別な側面を考える必要がある。それは、情報の「生産者」と「消費者」の接近を可能にしたという側面である。HTML による情報表現の容易さが、初心者にも自ら発信したい情報の記述と配信を可能にしたと言っても過言でない。この HTML 記述をベースとしたソフトウェアシステムにより、マルチメディア・コンテンツをネットワークを通じて世界に向けて発信することが可能になり、従来情報の消費者であった多くのユーザが、今度は情報の生産者にもなり得たという意味で非常に大きな意義をもっている。逆に言えば、上記のようなタグ付きのデータ表現方法が、その構造の簡潔さからこのような大きな変革をもたらす可能性を秘めていたとすることができる。

以上、最近のデータ共有のための記述モデルに関して、WWW の広がりの方の要因を探るところから述べてきた。しかし、やはり、HTML, SGM

L, さらには三次元データの記述法として普及している VRML (Virtual Reality Modeling Language)[15] の記述では不十分な点は多々ある。データの意味的な表現に関しては不十分であることは明白であり, また動画データなどに関して, 時空間的な広がりをデータ表現のなかにうまく採り入れることについても解決していない。これらの問題点や, データの共用のための諸要件とシステム性能を考慮しながら, マルチメディア・コンテンツのより強力な標準的な表現法の創出が大きな課題である。

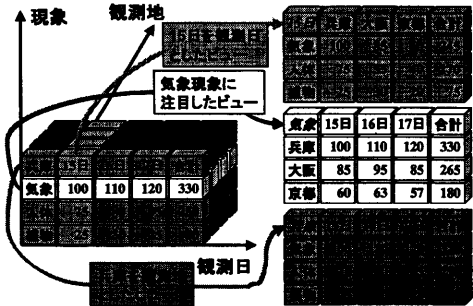


図5 データキューブモデルの例

### 3. データ発掘とデータ表現

近年, データ発掘, または, データベースからの知識発見 (KDD: Knowledge Discovery in Databases) に関する研究が非常に盛んに行われている [14,12,9,4,3,11]。これらの研究の適用領域としてデータウェアハウスの研究がある。このデータウェアハウスという用語は, 直接的には知識発見のための生データが蓄えられる倉庫という意味があるが, 最近ではそれ自身の概念は非常に広く用いられ, エンドユーザに対する意思決定支援システム (DSS: Decision Support System) の提供まで含んでいる。データモデルの観点から特に注意すべきデータウェアハウスの関連技術としては, 関係データベースにおけるデータ表現を見直した上でのデータキューブ (data cube) モデル [5] の研究がある (データキューブモデルの例として図5参照)。このデータモデルの重要性を考える上で, データウェアハウスにおいてデータ発掘を行い, 意思決定に利用可能なデータ処理を行うために要する生データの格納技術の特徴を考えると以下のことがあげられる。

まず, 従来の一般的なデータベース管理システムでは, 数百メガバイトから数ギガバイトのデータを対象として, Read/Write 操作を含む単純なトランザクションを実行する OLTP (On-Line Transaction Processing) を重視した設計がなされてきた。これに対して, データウェアハウスでは, 数百ギガバイトからテラバイト単位に及ぶ規模のデータに対して, Read 操作を中心とした複雑なトランザクションを実行し, 大量データに対

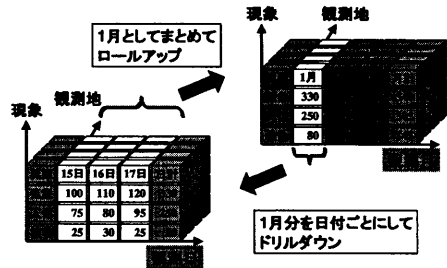


図6 ロールアップ・ドリルダウン操作の例

する解析的処理を行うための技術が要求される。従って, 問合せ出力に応じて, 視点を動的かつ対話的に切り替えることを可能とする OLAP (On-Line Analytical Processing) 技術が必要となっている。

また, データウェアハウスでは, データベース問合せ言語 SQL における GROUP-BY 記述をはじめとして集合に対する操作が頻繁に要求されるため, ドリルダウン (drill-down) と呼ばれる特殊化 (specialization) の操作と, ロールアップ (roll-up) と呼ばれるデータの一般化 (generalization) の操作を効率良く実現する技術の開発が重要な課題となる (これら二つの操作の例として図6参照)。そこで, 一般化操作の対象となるデータを単に平坦に格納するのではなく, 過去に記録されたデータの特性を的確に分類し, 要約データ (detailed and summarized data) や履歴データ

(historical data) として格納することが重要であるため [6,2], 多次元データベース (multi-dimensional database) が利用される。

さらに, 分類されて多次元データベースに格納されたデータに対して, OLAP 記事を用いるには, 頻繁に出力されるビューデータをいかに保持して効率的に再利用するかが重要な技術課題となる。そこで, この種の処理を適切に行うために重要な役割を果たす技術として, データキューブと呼ばれるデータモデルが提案された。このモデルでは従来関係データモデルで非常に重要視されてきた正規化 (normalization) の概念に逆向するデータ表現がなされ, 一つのデータモデルのなかに可能な限り多くの事実を関連づけさせて格納しておき, 必要に応じてそのモデルからデータをスプレッドシート (spread sheet) の形式で取り出すことを可能にしている。(図5において, 注目する対象を決めたビューの表示例を表している)。実際, データウェアハウスのように動的なデータ更新がなく, エンドユーザが直接利用するような環境では, 正規化技法がむしろ害になることを関係データベースの提案者である E. F. Codd 博士自身が認めている [7]。

#### 4. おわりに

本稿では, WWW をベースとした情報システム, およびデータ発掘と関連したデータウェアハウスの構築を対象としながら, データの共有と発掘という観点からデータ表現に関して起こっている新たな動きについて論じた。前者はネットワークとデータベースシステムの統合領域という観点から, 後者はデータベースシステムと人工知能の両者に関連する応用領域という観点から重要な対象領域であり, また現在最も活発に研究・開発が推進されている分野である。これらの分野の特徴としては, エンドユーザに非常に密接に関わっていることがあげられるが, その観点からユーザにとってより簡便で汎用性に富むデータ表現の研究が今後さらに進展することが期待される。

#### 参考文献

[1] Berners-Lee, T. and Connolly, D.:

Hypertext Markup Language - 2.0, RFC1866, Nov. 1995.

- [2] Brackett, M. H.: The Data Warehouse Challenge, Wiley Computer Publishing, 1996.
- [3] Chen, M.-S., Han, J. and Yu, P.S.: Data Mining: An Overview from a Database Perspective, IEEE Trans. on Knowledge and Data Engineering, Vol.8, No.6, pp.866-883, June 1996.
- [4] Fayyad U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R.: Advances in Knowledge Discovery and Data Mining, AAAI/MIT Press, 1996.
- [5] Harinarayan, V., Rajaraman, A., and Ullman, J. D.: Implementing Data Cubes Efficiently, Proc. of the 1996 ACM SIGMOD Inter'l Conf. on Management of Data, pp.205-216, June 1996.
- [6] Inmon, W. H.: The Data Warehouse and Data Mining, Comm. ACM, Vol.39, No.11, pp.49-50, Nov. 1996.
- [7] 石井義興: データ・ウェアハウス, 日本経営科学研究所, 1996.
- [8] ISO 8879: Information Processing - Text and Office Systems - Standard Generalized markup Language (SGML), 1986.
- [9] 河野浩之, 西尾章治郎, J. Han: データベースからの知識獲得技術, 人工知能学会誌, Vol.10, No.1, pp.38-44, Jan.1995.
- [10] Kim, W.: Introduction to Object-Oriented Databases, The MIT Press, 1990 (邦訳: 増永良文, 鈴木幸市 (監訳): オブジェクト指向データベース入門, 共立出版, 1996).
- [11] Mannila, H.: Methods and Problems in Data Mining, Proc. of the 6th Inter'l Conf. on Database Theory, pp.41-55, Jan. 1997.
- [12] 西尾章治郎: 大規模データベースにおける知識獲得, 情報処理, Vol.34, No.3, pp.343-350, Mar. 1993.
- [13] 西尾章治郎, 塚本昌彦: 高度ネットワーク

とデータベースシステム, Computer Today, No.79, pp.9-14, May 1997.

- [14] Piatetsky-Shapiro, G. and Frawley, W. J.: Knowledge Discovery in Databases, AAAI

/MIT Press, 1991.

- [15] 田中克己.: VRMLと3次元空間データベース, Computer Today, No.75, pp. 4 - 9, Sept. 1996.



～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

## テキストからの情報・知識の獲得と自然言語処理

辻井 潤一\* TSUJII Jun'ichi

### 1. はじめに

WWW, Inter/Intra-Net の普及に伴って、電子化され、自由にアクセスできる自然言語テキストの量は急速に増大している。また、辞書、ソーラスなど、自然言語処理に必要な言語資源の電子化も急速に進展してきている。しかしながら、情報検索・テキストからの情報抽出・知識獲得といった大量の現実テキストの集合を対象とする応用システムでは、言語の構造的な側面に焦点をあてた自然言語処理・理解研究の成果はあまり使われず、むしろ、言語の統計的な側面からの手法が主流となっており、言語の構造的な側面をとり扱う場合にも、応用分野に依存した Ad-Hoc な手法を用いるに留まっている。

我々（東京大学・辻井研究室、東京工業大学・徳永研究室、京都大学・黒橋グループ）は、学術振興会・未来開拓型プロジェクトにおいて、(1)計算言語学、言語理解研究からの手法を現実の大規模なテキスト集合に適用する際の問題点を明らかにし、その解決手法を提示すること、(2)これまで独立に研究されてきた「構造を中心とした」手法と「統計的な側面からの」手法とを統合する枠組を提示すること、および、(3)(1)(2)の成果を各種の応用システム（知的な情報検索、テキストからの情報抽出・知識獲得など）を構成することで検証すること、を目的として、5年間の研究プロジェクトを行なっている（1996年度開始）。本稿では、このプロジェクトの概要と現在までの成果を議論することにする。

大量なテキストを処理する現実的なシステムを構成するためには、必要な言語的・非言語的（実世界的）知識がすべてシステムにあらかじめ与え

られているという、従来型のシステムの前提を緩めることが不可避なことから、これらの知識をシステムが自己組織的に獲得していく技術が鍵となる。また、知的な情報検索や、テキストからの情報抽出・知識獲得においても、テキスト中に Implicit に表現された情報を顕在化し、組織化していくことが基本のプロセスとなる。したがって、本研究プロジェクトは、より広い展望からは、テキスト情報からの知識の自己組織化の方法を解明することであり、人間の言語学習・知識獲得機構の解明という、21世紀の知識科学の基本テーマに、情報科学の立場から貢献することになる、と我々は考えている。

### 2. 研究課題

言語の構造的な側面を中心にした処理といった場合、言語、あるいは、テキストに関連する構造とは何かをまず明確にする必要がある。この構造として、我々は、次の3つの構造を考える。

1. テキストの意味構造：文の持つ統語構造、意味構造を確定する文解析の研究は、計算言語学の中心課題として研究されてきた。また、米国を中心とした情報抽出（Information Extraction）の研究は、テキストの持つすべての情報（意味）を確定することは困難であっても、ある特定の観点から意味のある情報に興味を限定すると、かなりの確度で情報抽出可能なことを示している。

従来の情報検索では、テキストを特徴づけるのにキーワードやキーフレーズが使われてきたが、IEの結果得られる定型化された情報と原テキストとの相互関連を記憶することで、事実検索（Fact Retrieval）と

\* 東京大学理学系大学院情報科学専攻



文献検索 (Document Retrieval) とが融合したシステムを考えることができる。

2. 分野オントロジー, 知識の構造: シソーラスは, Recall 率を向上させるためにユーザからの Query を拡張すること, 標準的な用語による文献の特徴づけなど, 従来型の情報検索においても中心的な役割を果たしてきた。しかしながら, 用語の相互関係を, 上位語-下位語, 関連語といった表層関係でとらえる言語シソーラスは, 論理的推論関係を重視する知識処理の観点からは不十分なものである。

言語シソーラスの量的な側面を, 知識処理における知識表現が持つ論理的な側面によって補完することにより, 知的情報検索における推論の中心的な Component を構成することができる。

3. テキスト, 間テキストの構造: テキストが, 文の単位を越えた構造 (たとえば, 相互参照関係, トピックの連続性, 修辞構造, など) を持つことはよく知られている。従来型の情報検索でも, テキストの構造は, キーワード抽出の手がかりを与えるものとして研究されてきた。

一方, Inter/Intra-Net に流通する電子メールなどのテキストは, 従来の印刷テキストに比べて, テキストとしての完成度が低く, テキスト作成時の文脈への依存度が極めて高い。この種のテキストは, テキストの単位を越えて別のテキストの部分を参照するなど, 会話における発話文と共通した性質を持っている。今後, さらに増大するであろう, この種のテキストの検索においては, テキスト単位を越えた構造を適切に把握しておく必要がある。

以上3つの, 言語テキストに関連した構造を情報検索に反映するためには, そのための要素技術をまず確立し, それを目的に応じて統合してゆく必要がある。このために, 本プロジェクトでは, 次の4つの研究課題を設定している。

**研究課題 1** 高効率, 高耐性で, かつ, 言語理論的にも健全な文解析手法の確立

情報検索のように, 実世界のテキストを大量に処理しなければならない分野において, 文解析に代表される構造的処理手法が使われてこなかった理由は,

1. かなり高速な文解析アルゴリズムが開発されてきたとはいえ, 大量のテキスト処理に使えるほど, 高効率なものにはなっていなかったこと
2. 現実のテキストは, 一般的な言語を対象とした文法が想定していない構造表現や特殊な単語が含まれる。文法が精緻になればなるほど, 予期しない表現や未知語に対して Fragile であることである。この2つの欠陥に対して, 情報抽出 (IE) やコーパスからの知識獲得研究においては, 必要な部分だけを解析し, 他は無視する部分パーズング (Partial Parsing) や, 曖昧さを Implicit なまま残しておく表層パーズング (Shallow Parsing) といった手法が提唱され, 使用されてきた。

しかしながら, 言語学的な観点からの妥当性を欠くこれらの手法は, 応用分野ごとに別個の, しかも, Ad Hoc な文法を作成しなければならないこと, 意味や文脈への Systematic な写像がなく, より精緻な知識獲得・情報抽出研究へ向かう展望がないことなど, 技術・科学の観点からは, 非常に中途半端なものとなっている。

したがって, 本プロジェクトでは, 現時点の言語理論の観点からも妥当で, かつ, 効率, 耐性の観点からも実用的な文解析手法の確立を目指す。特に, 現時点での言語理論の主流である拘束中心 (Constraints-based) な文法記述が, Underspecified な記述を積極的に許容することに注目し, この点からの部分パーズング・表層パーズングの理論化を行なう。

**研究課題 2** 言語, および, 実世界知識の獲得手法の確立

実世界知識の言語的な反映としてのシソーラスを自動的に獲得する手法の確立は、分散的で多様化したネットワーク時代の情報検索にとって不可欠の技術となる。また、シソーラスだけでなく、分野やテキスト・タイプごとに变化する言語的・言語外的な制約や優先情報 (Preference) を自己組織的に獲得する技術は、将来の自然言語処理技術の鍵となる技術であり、同時に、人間の言語獲得・知識獲得のモデル化の第一歩となる。

本プロジェクトにおいては、言語の統計的な挙動と構造的な制約とを組み合わせることによって、(半)自動的な知識獲得のシステムを作成する。言語現象に関与する知識は多様であり、現在、次の3つの方向から研究を行っている。

1. 言語の統語・意味制約と統計的な挙動の統合：[研究課題1]の言語モデルは、言語表現が満足すべき制約の定式化に焦点がある。この制約的な側面を補完し、言語構造の曖昧さを優先度の形で取り込むためのモデル、および、Subcat-Frameなど語彙的情報を自動獲得するモデルの研究
2. 分野知識の自動獲得：シソーラス的な実世界知識を、特定分野のコーパスから自動獲得する技術を確立する。特に、(a)分野知識の基本単位を構成する専門用語の認定、(b)専門用語の内部構造、および、その出現の周辺環境からの専門用語間の相互関係の認定、(c)専門用語から知識単位への写像と分野の知識構造の認識、を行なうシステムの研究
3. 知識獲得のための統合システム：言語現象に関与する知識は、言語の構造に関する知識から、分野の実世界知識まで多様であり、しかも、それらが有機的に関係している。したがって、単一原理に基づくプログラムが一回の実行ですべての知識を獲得することは考えられず、初期知識・その時点までの獲得知識・コーパスを共有する一連のプ

ログラム群が漸近的 (Evolutionarily) に知識を組織化していくモデルを考える必要がある。この種の自己組織系の基盤としての、構造化されたコーパスや獲得知識を系統的に管理するデータベース系の研究

### 研究課題3 テキスト内、間テキスト構造の処理モデルの確立

間テキスト構造のもっとも大域的なものとしては、統計的な手法のテキスト・クラスタリングの研究がある。また、テキスト内構造としては、定名詞句や代名詞の相互参照の研究がある。本プロジェクトでは、これらの既存の研究を進めるとともに、テキスト・クラスタリングから、テキスト間の内容の重複、時系列関係など、より構造的な側面に焦点を当てたテキスト・スレディングに関する研究、および、相互参照関係や修辞関係を取り扱うための形式性の高い枠組 (たとえば、DRT理論) の研究を行なう。

### 研究課題4 統合化された知的検索システムのプロトタイプ・システムの開発

本プロジェクトは、[研究課題3]までの要素的・基盤的な研究を中心とするが、これらの要素技術の現実場面での有効性を示すために、知的検索システムのプロトタイプを作成する。この研究課題は、プロジェクトの後半、3年度から本格化する予定である。

## 3. 研究の成果、現状と今後の方向

研究課題ごとに現在の状況を、以下にまとめる。

**研究課題1** 文解析技術は、自然言語処理のもっとも基盤的な技術であり、この技術の確立は、本プロジェクトの今後の研究方向を決定するものであり、多大の努力を傾注し、以下のように、有望な結果をあげている。

- (a) コンパイル・データフロー解析による HPSG の高速文解析手法の確立
- (b) タイプつき素性構造抽象機械の実現とそれに基づく論理型プログラム言語 (LiLFeS) の開発
- (c) 並列計算による HPSG 文解析の高速化

(d)ペンシルバニア大学の LTAG 文法の HPSG への移行と大規模な英語文法の開発 (a)から(c)の結果は、いずれも、言語学的に健全な文法フォーマリズムに基づく文解析方式の高速化に関するものであり、(b)の大規模文法を使った予備実験の結果、膨大なコーパスを処理するのに十分な処理速度が確保できる見通しを得た。

高耐性に関する研究では、品詞だけが確定している未知語の処理を Underspecified な辞書記述で行なう方式を考案し、これも期待した結果を得ている。ただ、より広い範囲に見られる、文法範囲を逸脱した構文の処理、意味レベルでの Underspecification の取り扱いなどは、今後の研究課題として残っている。また、優先度解釈、高効率化にとって不可欠である解析結果の Packing 技法、Disjunction の取り扱い、開発された文解析手法の情報抽出への適用なども、開始した。

(c)で開発した論理型言語は、さらにその機能を拡張することにより、[研究課題 2] や [研究課題 3] のための基幹言語として整備する予定である。

**研究課題 2** 漸近的な知識獲得を行なうための統合システムの設計・開発と、個々の知識獲得手法の研究とを同時並行的に行ない、次のような成果を得た。

(a) [研究課題 1] の文法モデルと統計的な言語モデルの統合を行なう実験

(b)専門用語認定のアルゴリズム開発と医学コーパス (眼科病理) への適用

(c) [研究課題 1] で開発された論理型言語系とコーパス・データベースとの統合に関する基本設計とその開発

(d)周辺の言語構造から専門用語の意味的クラスを推定する計算モデルの作成

[研究課題 1] の研究のいくつかは、すでに完了しつつあるのに対して、上記の研究課題は、(b)を除くといずれも現在進行中のものである。また、(b)は、言語の構造的な側面と統計的な側面を使う Hybrid 型の手法であるが、統計的な側面を重視したものになってお

り、専門用語間の相互関係など、より精緻な知識を獲得するためには、構造的側面をより重視した手法が必要になると考えている。(d)は、この方向の研究を目指すものである。

また、(c)は、日本電子化辞書 (EDR) と本プロジェクトとの共同研究によるものであり、EDR のコーパス・データベースと本プロジェクトで開発している論理型言語 (LiLFeS) との統合を行なうこと、また、LiLFeS の型定義機能を知識表現の枠組として使うことを目指している。

**研究課題 3** 間テキストに見られる相互関係の認識の第一歩として、テキスト・クラスタリングの研究、テキスト内の照応関係・結束構造の解析に関する研究を行なっている。

(a)テキスト・クラスタリングの研究では、大量のテキストを高速にクラスタリングするアルゴリズムを開発した。このアルゴリズムは確率モデルに基づいており、1 万程度のテキストを現実的な時間で処理できる。1 万を越えるテキストに関しては、テキスト集合の一部でサンプリングしてクラスタリングした後、残りのテキストをこのクラスタ上で分類することによって全テキストをクラスタリングする手法を提案している。我々の手法は、確率モデルに基づいているため、初期クラスタの生成とその上での残りのテキストの分類が同一の確率尺度に基づいて行なわれるため、大規模化による性能の劣化があまりないという利点がある。

(b)テキスト内の照応関係・結束構造の解析に関する研究では、エキスパートシステムの枠組による解析システムを開発した。これらは、表層の表現を手がかりとして解候補に点数を与えるというルールを大量に用意し、最終的に最高の得点を得た解候補を選択するというシステムである。現在、このような枠組の問題点を明確化し、より頑健なシステムとなるように種々の調査・改良を行なっている。

**研究課題 4** 要素技術の研究を有効に行なうためには、応用システムのイメージを明確に持つ必要がある。このため、将来、プロトタイプ・

システムとして開発する対象分野を調査している段階である。現時点では、次のような応用分野を想定している。

(a)ゲノム文献・データベース：ゲノム研究プロジェクトにおいては、関連する非常に広い分野の文献データベース、および、定型化されたデータベースが管理されている。定型化されたデータベースを、この分野の知識ベースと考えると、文献からの知識の自動獲得・情報抽出、あるいは、知識に基づく知的情報検索などの技術の格好の実世界応用となる。

(b)東大病院における退院サマリ：東大病院における医療カルテ電子化計画の一貫として、個々の退院患者のサマリ・レポートのデータベースが計画されている。サマリレポートは、経済情報と同様に限定の強い分野のテキストとなっており、情報抽出の格好の対象となる。また、テキスト中の文は、省略や専門用語が頻出し、文解析の耐性をはかる格好の対象である。

(c)特定の話題に関する新聞記事：(a), (b)は、いずれも独立度の高いテキストを対象としており、テキスト・スレディング的な技術を研究するには不適切である。これに対して、ある特定の話題の新聞記事、電子メールでの議論などを対象としてテキスト・スレディングのプロトタイプを開発することが考えられる。プロジェクト後半からの大規模な実験に備えて、現在、これらの応用分野のテキスト収集、システム仕様を検討中である。

#### 4. 研究活動

本プロジェクトは、東京大学・辻井研究室、東京工業大学・徳永研究室、京都大学・黒橋研究室の共同研究として進められている。また、研究を効率よく推進するために国内・国外の研究グループとの研究交流を積極的に進めており、現在、以下の海外研究機関と活発な研究交流を行なっている。

ペンシルバニア大学 (Prof. A. Joshi)：英語文法の移植とその性能評価、ニューヨーク大学 (Prof. R. Grishman)：情報抽出ソフトウェアの

開発、MMU/UMIST (Dr. S. Ananiadou)：コーパスからの知識獲得、専門用語の自動認識、ピザ大学 (Prof. N. Calzolari)：コーパスからの言語知識獲得、ローマ大学 (Prof. M. Pacienza)：情報抽出と知的情報検索システム、GMD (Dr. H. Voss)：メールによる議論の間テキスト構造の研究

また、国内のグループでは、知的情報検索、テキストからの情報抽出の具体的な応用課題を持っているグループとして、

東京大学医科学研究所 (高木教授)：ゲノム・テキストの知的検索、東京大学医学部 (大江教授)：退院サマリからの情報抽出

と緊密な研究交流を計っている。同時に、日本電子化辞書 (EDR) との研究協力のもとに、富士通・東芝・松下などの民間企業の研究グループとも技術の交流を行なっている。

プロジェクトの発足年度 (1996年12月)には、上記の国内・国外の各グループとの研究計画を策定するために、Closed な workshop を開催し (参加者 25名)、1997年4月には、その follow-up とより広い研究ソサエティへ貢献を目的に、公開のシンポジウムを日立・基礎研究所と共同で開催した (参加者 120名)。

#### 参考文献

- [1] Tsujii, J., Ananiadou, S. Information Retrieval and Natural Language Processing, in Proc. of IRAL, October, 1997 (to appear).
- [2] 辻井潤一. 視点の変換 言語の理論から設計の理論へ, 人工知能学会誌, Vol.11, No 4, 1996
- [3] Ananiadou, S., Radford, I., Tsujii, J. Sublanguage Knowledge Acquisition for Hypertext Optimisation, in proc. of NLPRS-95, Seoul, 1995.
- [4] Kentaro Torisawa and Jun'ichi Tsujii. Computing phrasal-signs in HPSG prior to parsing. In Coling 96, pages 949-955, 1996.
- [5] Takaki Makino, Kentaro Torisawa, and Jun'ichi Tsujii. LiLFeS - practical

- unification-based programming system for typed feature structures, 1997. to appear in NLPRS 97.
- [6] Takashi Ninomiya, Kentaro Torisawa, Kenjiro Taura, and Jun'ichi Tsujii. A parallel cky parsing algorithm on large-scale distributed-memory parallel machines. In *PACLING '97*, pages 223-231, September 1997.
- [7] Kenji Nishida, Takaki Makino, Kentaro Torisawa, Yuka Tateisi, and Jun'ichi Tsujii. Extension of a feature structure abstract machine for partial unification. In *PACLING '97*, pages 232-243, 1997.
- [8] Radford, L., Ananiadou, S., Tsujii, J. Adding Structural Constraints to Lexically Based Context Matching Techniques, In *International Joint Conference on Artificial Intelligence (IJCAI-97)*, August 23-29, 1997, Nagoya, Japan.
- [9] Frantzi, K. T., Ananiadou, S., & Tsujii, J. Term Identification using Contextual Cues. In *Proc. of the 2nd Workshop on Multilinguality in Software Industry : the AI Contribution (MULSAIC'97)*, at the *International Joint Conference on Artificial Intelligence (IJCAI-97)*, August 23-29, 1997, Nagoya, Japan.
- [10] 建石由佳, 鳥澤健太郎, 牧野貴樹, 西田健二, 淵上正睦, 辻井潤一. LTAG 文法からの変換による HPSG 英文法の作成. *情報処理学会第 122 回自然言語処理学会 (予定)*, 1997
- [11] Iwayama, M. and Tokunaga, T. : Hierarchical Bayesian clustering for automatic text classification ; *Proceedings of IJCAI '95*, pp. 1322-1327, 1995.
- [12] 岩山真, 徳永健伸, 桜井直之. 文書検索のための大規模文書クラスタリング. *言語処理学会第 3 回年次大会*, pp. 245-248, 1997.
- [13] Sadao Kurohashi, Makoto Nagao. Automatic detection of discourse structure by checking surface information in sentences. *Proc. of 15th COLING*, Vol.2, 1994.
- [14] Masaki Murata. Anaphora resolution in Japanese sentences using surface expressions and examples. PhD Dissertation, Kyoto University, 1996.

～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

## JICST シソーラスの改訂 Outline of JICST Thesaurus 1999

永井 賢吉\* NAGAI Kenkichi

科学技術振興事業団 (JST) が作成している JICST 科学技術文献ファイルは JICST シソーラスという知識体系に基づき索引されてきた。現在 JST では、特定有機化合物について、これまでの索引方式からもう1つの組織化された知識体系である日本語化合物辞書ファイルに基づく索引へ全面的に移行する準備を進めている。この計画に沿ったシソーラス改訂作業方針を解説する。

### 1. シソーラス

シソーラスは統制された索引言語の語彙であって、あらかじめ概念間 (例えば、「上位」、「下位」) の先験的な関係を明示するように組織化したものである。シソーラスで規定される概念間の関係には階層関係、優先関係、関連関係の3つがある。階層関係は、アプリアリに決まるものばかりではなく、シソーラスの使用目的によって定められる場合がある。例えば、ウシと家畜と動物の階層関係をどのようにするかは単純ではない。JICST シソーラスの場合も JICST ファイルの検索における利便性が優先されている。優先関係は同義語間の関係であり、索引や検索に用いる用語とそうでない用語との関係を定めたものである。前者の用語をディスクリプタ、後者を非ディスクリプタと言う。関連関係は、他の2つと比べかなりあいまいであり、階層関係や優先関係以外の関係と定義されている。シソーラスにおいてこのような関係を定めている理由は、シソーラスの利用者に対して、検索・索引に用いるべき用語を案内することであり、ひいては索引における信頼性の向上と検索効率の改善が目的である。

JICST シソーラスは1975年に初版が作られて以来、4回の改訂を経て現在の1993年版に至っている。このシソーラスにより JST では年間70万件の科学技術文献の索引を行い、JOISを通し

てサービスしている。現在使用している93シソーラスは14分野208カテゴリーの用語50,798語 (ディスクリプタ数39,046語) が収録されている。1階層内の最大語数は2,840語であり、また最も深い階層は11階層である。

JICST ファイル作成工程においては、索引者が付与したディスクリプタに対して計算機によりその上位語を付加している。これにより JOIS 検索では上位語による網羅的な文献検索が可能となっている。例えば、「放射性崩壊」という用語で検索した場合、シソーラスにおいてその下位語となっている「アルファ崩壊」、「ガンマ崩壊」、・・・等が付与されている文献すべてがヒットする。階層が最も大きい用語の場合は2840語を入力したことと同じことになる。利用者が入力した検索語について、シソーラスの階層構造に沿った組織化が実現していると言える (次ページ図1参照)。

### 2. 1999年版 JICST シソーラス

現在改訂作業を進めている第6版シソーラスの主な改訂方針は、物質用語の削除、非ディスクリプタ収録基準の見直し、の2点である。これらについて、組織化された知識体系の観点から解説する。

#### (1) 物質用語の削除

JST では1988年から JOIS-F により JICST 化合物辞書データベースをサービスしている。化合物辞書データベースは化合物の名称、構造情報等

\* 科学技術振興事業団

表1 JICST シソーラス収録語数の変遷と関連事項

| 年    | シソーラス収録語数  |
|------|--|
| 1975 | 第1版 33,998語 (ディスクリプタ 29,173語, 非ディスクリプタ 4,825語)                                 |
| 1978 | 第2版 35,572語 (ディスクリプタ 30,609語, 非ディスクリプタ 4,963語)                                 |
| 1981 | 第3版 40,902語 (ディスクリプタ 34,695語, 非ディスクリプタ 6,207語)<br>JOIS-II サービス開始 (漢字検索, 自然語索引) |
| 1987 | 第4版 48,196語 (ディスクリプタ 38,407語, 非ディスクリプタ 9,789語)                                 |
| 1988 | JOIS-F サービス開始  |
| 1990 | JOIS-III サービス開始 (抄録文からの切りだし, 近接演算, 化合物索引)                                      |
| 1993 | 第5版 50,798語 (ディスクリプタ 39,046語, 非ディスクリプタ 11,752語)                                |
| 1999 | 第6版 (予定)   |

放射性崩壊 (ホウシャセイホウカイ)

PA11

NT アルファ崩壊

ガンマ崩壊

- ・ 内部制動放射
- ・ 内部転換

自発核分裂

ベータ崩壊

- ・ Fermi 遷移
- ・ Gamow - Teller 崩壊
- ・ 二重ベータ崩壊
- ・ ベータプラス崩壊
- ・ ベータマイナス崩壊

BT 崩壊

RT 遷移確率

崩壊熱

図1 JICST シソーラスの用語例

を収録したデータベースであり、有機低分子化合物を対象とした組織化された知識体系である。この化合物辞書データベースのサービスに合わせ、1990年1月からは、JICST ファイルの作成工程

でも化合物の索引に関しては、シソーラスに基づく索引の他に日本語化合物辞書に基づく索引を特定分野の文献について行い、化合物辞書データベースの蓄積を図ってきた。

化合物に対するこのような重複した索引作業を日本語化合物辞書に基づく索引に一本化するとともに全分野を対象として実施することを、現在JSTでは計画している。この計画が実施されると現行の93シソーラスに収録されている有機化合物用語の多くは索引に用いられなくなるため、シソーラスに収録している必要がなくなる。このため、今回の改訂ではこれらの有機化合物用語をシソーラスから削除する予定である。

日本語化合物辞書とシソーラス

2つの知識体系であるJICSTシソーラスと日本語化合物辞書を比較してみよう。化合物に関しては、一般の専門用語とは対照的な特徴——数が多い、名称が多岐にわたっている(体系名、慣用名、商品名、治験番号)——を持っている。JICST93シソーラスでは収録語数の1割程度(約5,000語)が有機化合物用語であるが、これは100万を越える有機化合物総数のほんのわずかに過ぎない。また、シソーラスに収録する場合、どの名称同士が同義であるかを確定し、多岐にわたる名称からどれをディスクリプタとするかを決めることも化合物に関しては難しい課題である。

化合物辞書システムは追加型の辞書システムで

あり、日々登録された名称について系統的に(体系名から)構造を組み立てることにより、同じ化合物の名称を組織化している。日本語化合物辞書の登録数は現在 85 万化合物であり、JICST93 シソーラスに収録されている特定物質 4,000 程度とは桁が異なる。前者は原則として構造が確定している有機低分子化合物のみを収録しており、水素以外の構成原子数が 2 4 7 個以下の化合物というシステムの制限がある。また、有機塩素化合物といった総称的用語は収録できない。このような制約はあるが、用語を網羅的に収録可能かどうかということに関しては前者が優っている。シソーラスにおいても語数制限を考えなければ網羅的に用語を収録することは可能であろうが、階層関係の設定や優先関係の整理という人間の知的作業が収録の際に必要であり、現実的には数万語レベルが限界であろう。ときとして、同一概念の用語が異なった名称でシソーラスに登録されてしまうこともある。化合物辞書の場合は、収録の際に必要な知的作業は構造を正しく記述することであり、同一の物質の有無はシステムが判断してくれる。同じ化合物が異なった名称で重複登録されるといったようなことは起こらない。

用語間の階層関係については、シソーラスは人間の判断できめ細かく設定することができる。化合物辞書ではそのような組織化の器は用意されていないが、構造検索という形で階層関係と同じような組織化が可能である。しかし、構造検索にはある程度の大きさの部分構造を指定できることが条件であり、有機塩素化合物(=塩素を含む有機化合物)の集合を検索するといったことはできない。

利用の際の利便性はどうかであろうか。シソーラスという知識体系では、複数ある同義語の一つをあらかじめディスクリプタに定め、索引・検索にはその選ばれた用語を用いるという制約がある。このため、冊子体その他のツールによりどの用語がディスクリプタであるかを確認することという煩わしい作業が必要となる。一方、化合物辞書データベースでは登録されているさまざまな名称でのアクセスが可能である。しかし、知識体系の明示化という点では冊子体等を見ることを前提にして

いるシソーラスが便利であろう。シソーラスに収録されている化合物に関しては、有機化合物の専門知識を持たない者であっても、それが有機化合物全体の階層のどの部分に位置するのかわかることができるのである。上位概念語による検索の場合も、シソーラスを見ることによりどのような下位概念語が付与されている文献がヒットするのかわ確認できる。化合物辞書データベースの構造検索では検索者が部分構造を指定する必要があり、有機化合物の専門知識を前提としている。

2つの知識体系で大きく扱いが異なるものにも同形異義語の扱いがある。シソーラスでは上位語により意味限定されるわけであるが、それ以外に限定句、スコープノートによりディスクリプタが表す概念を規定する。一般に同形異義語をシソーラスに収録する際は互いに限定句を付与することで区別している。一方、化合物辞書では更新作業が随時行われるため、先に収録された化合物に限定句のない名称が割り振られ、後にシステムに入ってきた化合物にのみに限定句がつけられることになる。

#### SN-KW 自動付与

シソーラスによる索引から日本語化合物辞書による索引への有機化合物の索引方式の変更は、特定物質そのものの検索においては利点の多いものである。しかし前述したように上位概念語により特定化合物を網羅的に検索しようとする場合は、化合物の専門知識を持たない検索者にとっては不便になる。このためシソーラス改訂に合わせ、日本語化合物辞書により索引された用語(SW)から日化辞番号(SN)を通して、上位語に相当するシソーラス用語(KW)を自動的に索引する仕組みを開発する予定である。あらかじめ、SNとその上位語のKWとの対応テーブルを用意しておき、JICSTファイル作成工程の中で系統的にシソーラス用語を付加するものである。もちろん、80万を越える化合物辞書ファイルの用語すべてにこのようなことを行うわけではない。現行シソーラスから削除される用語と出現頻度の高い化合物用語を合わせ 12,000 語程度(物質数では 10,000 程度)を対象に対応関係を作成する予定



| 物質用語<br>(Substance Word) | 日化辞番号<br>(Substance Number) | シソーラス用語<br>(KW) |
|--------------------------|-----------------------------|-----------------|
| ペニシリンG                   | J2.342F                     | 窒素複素環化合物        |
| ベンジルペニシリン                |                             | 硫黄複素環化合物        |
| ベンジルペニシリン酸               |                             | カルボン酸           |
| スペシリンG                   |                             | 芳香族化合物          |
| リカシリン                    |                             | ラクタム            |

図2 SW-KW 対応データの例

である。関係付ける KW としては、シソーラスに残す予定の総称的用語約 600 語の中から化合物上位語にあたる用語を選定することとしている。

(2) 非ディスクリプタの収録基準の見直し

シソーラスの優先関係は同義語間の関係であると述べたが、これは必ずしも正しくない。収録語数(ディスクリプタ数)を押さえるため、本来は階層関係を定めるべき用語間に優先関係を設定する場合がある。索引規則を案内する USE A AND B 型の非ディスクリプタと合わせ、JICST シソーラスにおいては非ディスクリプタのかなりの用語が同義関係以外となっている。また、非ディスクリプタは「索引に用いないこと」というのがシソーラス利用の原則であるが、1993 年以降は非ディスクリプタを自然語として索引することをルール化している。(もちろん非ディスクリプタが自然語として索引された場合には、シソーラスを通して自動的にディスクリプタが付与されるようにしている。) 一方、第 5 版の JICST シソーラスと同時に、内部の索引ツールとして自然語-統制語変換辞書を JST では整備し利用している。これは、自然語として索引される用語とシソーラス用語との関係を規定したものである。1981 年に以降行っている自然語索引では、シソーラスにある上位概念/近似概念の用語の索引を併せて行うという規則がある。変換辞書はこの自然語-統制語の関係を組織化することを目指したものであり、現在使用している変換辞書は約 7,600 語を収録している。索引者が自然語を付与した場合にはこのテーブルを通して自動的に統制語(ディスク

リプタ)が付与されるシステムとなっており、索引者の労力低減と統制語索引のばらつきをなくす目的も有している。

自然語-統制語変換辞書の役割を見てみよう。これは索引者にディスクリプタを案内し、それらのディスクリプタの索引を自動的に行うものであり、シソーラスという知識体系の中の優先関係にある用語の知識体系と同じ体系であるといえる。シソーラスの優先関係もディスクリプタを案内する役目を有しているのである。

今回のシソーラス改訂に当たっては、シソーラスに非ディスクリプタとして登録されている用語と自然語-統制語変換辞書に登録されている用語の違いを明確にすることを旨とし、改訂方針の議論を重ねた。その結果、99 シソーラスでは、「(自然語としても)索引に用いない用語を非ディスクリプタとする」ことになった。具体的には非ディスクリプタはディスクリプタの同義語のみに制限することとし、優先関係が階層関係となっている非ディスクリプタについては、ディスクリプタに変更するか、あるいはシソーラスから削除し自然語-統制語変換辞書に登録するという作業を行っている。

3. 終わりに

シソーラスは索引・検索に用いる用語を案内し、索引における信頼性の向上と検索効率の改善のために、用語に統制を加えてきた。しかし、化合物に関してはシソーラスにより索引を保証できる範囲が極めて制限されてきた。今回のシソーラス改訂における物質用語の削除は、JICST ファイルの

索引の立場から見ると、一部の化合物についてのみ保証されていた索引を、全化合物を同じレベルで索引することへの変更に伴う措置といえる。どのような名称でも検索できる、構造からも検索できるという化合物辞書の利点を十分活用していただきたい。

非ディスクリプタの収録基準の見直しについては、「ディスクリプタと非ディスクリプタを定める」ことが実は「索引や検索に用いない用語を定める」ことであるというシソーラスの基本への回帰と言える。シソーラスの語数の問題、自然語索引の実績の中で、JST 内部でもあいまいとなってしまった非ディスクリプタの位置づけを明確にしたものといえる。

パソコンやネットワークといった情報インフラの普及に伴い、情報検索を行う者がサーチャーのみではなく、一般のエンドユーザへと拡大しつつある今日、利用者に親しみやすい形態にシソーラスも変わっていく必要がある。エンドユーザにディスクリプタで検索することを強いるのは好ま

しくないことであり、無理があろう。今後は化合物辞書のように「同義語はどれを用いてもよい」という知識体系を目指して、シソーラスも変わっていくべきである。優先関係を同義関係に限定したことはそのための第一歩である。

#### 参考文献

- 1) 岡野弘行. “JICST 科学技術用語シソーラス”. 情報の科学と技術, Vol.39, No.12, p558-566
- 2) 五味淵巨. “JICST 科学技術用語シソーラス 1993 年版の作成”. 情報管理. Vol.35, No.11. p953 - 965
- 3) 木村美実子ほか. “JICST 化合物辞書データベースの拡充 混合物の登録と文献への物質索引”. 情報管理. Vol.34, No.2. p141-162
- 4) 荒木啓介. “JICST ファクトデータベース (2) JICST 化合物辞書データベース”. 情報管理. Vol.31, No.1. p21-40

～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

### 印象語の地図

— マルチメディア情報の感性検索 —

加藤 俊一\* KATO Shunichi

#### 1. はじめに

法則以心伝心、皆令自悟自解（法ハ則チ心ヲ以テ心ニ伝ヘ、皆自ラ悟リ自ラ解カ令ム。六祖壇教）。

人間と人間の間コミュニケーションは、単にメッセージの字面のやりとりだけではなく、相手の知識や考え方、嗜好や意図、独特な表現方法など、様々な観点で相手の人となりや自然に理解しつつ行われている。それだけに、相互に相手を深く理解する必要がある一方で、数少ない言葉のやりとりからでも深いレベルでのコミュニケーションが可能となるのである。

しかし、現状の情報機器・情報システムと人間とのコミュニケーションは高度に抽象化、客観化された情報に偏っており、一人一人の人間によって異なるバックグラウンドの知識や考え方、嗜好などの個性（主観的評価尺度）は考慮されていない。これは上述した人間にとって自然なコミュニケーションの形態とは大きく異なり、円滑なコミュニケーションを行う上での阻害要因となっている。多様な人々に利用される情報機器・情報サービスを提供するためには、このような人間の側面の情報処理に視点をおいた情報技術の実現が必須である。これは情報処理の根本的な問題といえる。

#### 2. 印象語による対話と視覚感性

マルチメディア情報システムでは、種々のメディアの物理的特性から、特定個人による主観的な解釈まで反映した形で、人間と情報機器の間にマルチメディア情報を媒介とするコミュニケーション（マルチメディア対話）を成り立たせる必要があ

る。そのためには、個々の利用者がメディアを解釈する際の特徴（個性）を客観的に計測し、これをモデル化する枠組みを確立することが、マルチメディア情報システムのための最も基本的な技術と言えよう。

我々は人間の知的過程（特に視覚）での感性の工学的モデルとして、次の5つの階層を考えている（図1）。

- (i) 物理的レベル：対象をその物理的な信号レベルで内容を判断すること。
- (ii) 生理的レベル：神経回路の特徴抽出機構を経てパラメータ化された上で、内容を判断すること。個人差はあまりない。

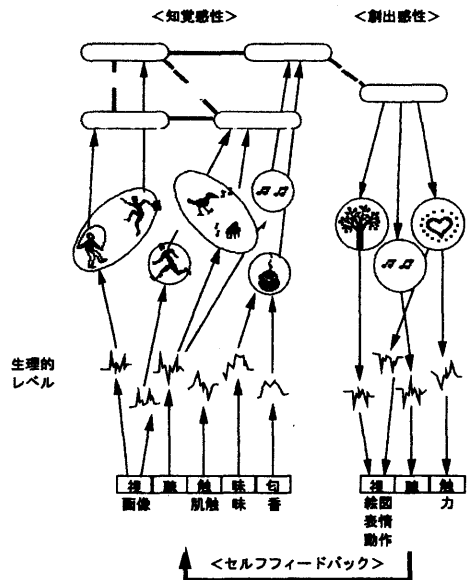


図1 感性の工学的モデル

\* 中央大学工学部 経営システム工学科  
 \* 通商産業省工業技術院 電子技術総合研究所 ヒューマンメディアラボ

- (iii) 心理的レベル：言語を媒介せずに対象間の類似性を判断する。個人差が見られる。
- (iv) 認知的レベル：あるカテゴリの対象に印象語（概念のラベル）を割り当てる。個人差だけでなく、年齢による差，文化による差も観察できる。
- (v) 論理的レベル：知識として用語間の関係（オントロジー）を理解する。

### 3. 感性分析と可視化

対象や考えを印象語で表現（いわゆる感性的表現）することが自然な場合が多い。

芸術の世界では、人間の感じる印象は絵画の色彩（組合わせと配色等）と構図に由来することが知られている。一方、好みや文化的背景などの違いにより、個人（あるいはグループ）毎に、印象語による表現と絵画の色彩との対応関係に偏りがあることも知られている。我々は、主観的評価基準を工学的にモデル化し、マルチメディア対話を

行う手法を開発した（図2）。

絵画や風景写真など、対象を限定すれば、主成分分析や因子分析等の多変量解析を用いることにより、一人一人の感性において印象語の使われ方や、印象語間の相対的關係を分析し、可視化できる。

例えば、利用者は、システムが提示する学習用の絵画それぞれに対して、各絵画を見た印象を、限定された種類の印象語に対して5段階評価で重みを与えて記述する。利用者の回答（あるいは利用者のグループの平均値）に主成分分析を適用して、主要な2軸でつくる平面に各「印象語」を射影する。これを「印象語の地図」と呼ぶことにする。これは特定の利用者の（あるいは特定の利用者グループに共通の）印象語間の関係を空間的に可視化したものである。

図3に、特定の利用者グループ（左：女子学生グループ、右：男子学生グループ）の平均値から得た印象語の地図を示す。この実験では、同じ学

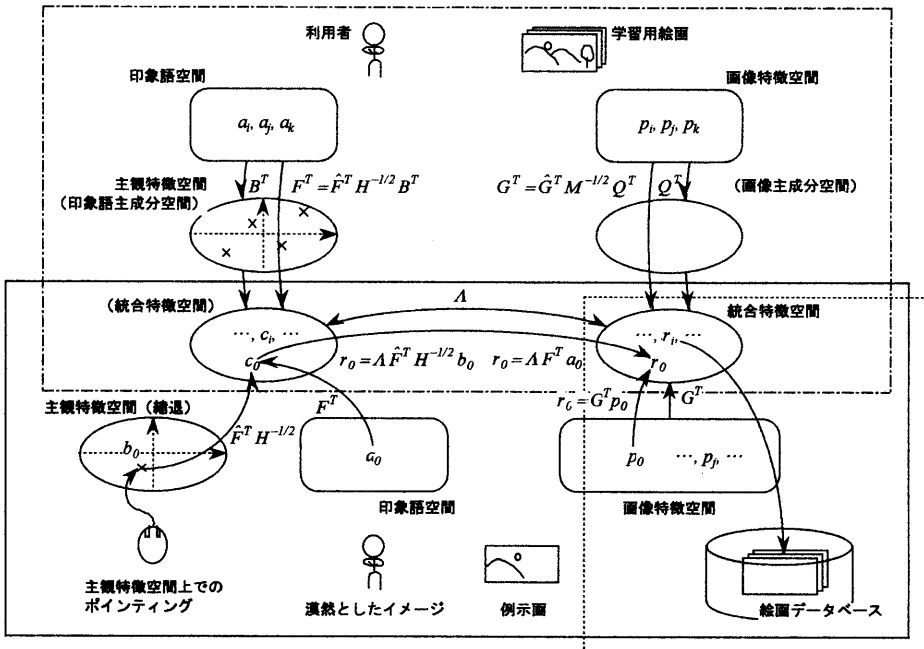


図2 感性分析による印象語の地図, 感性検索の実現法

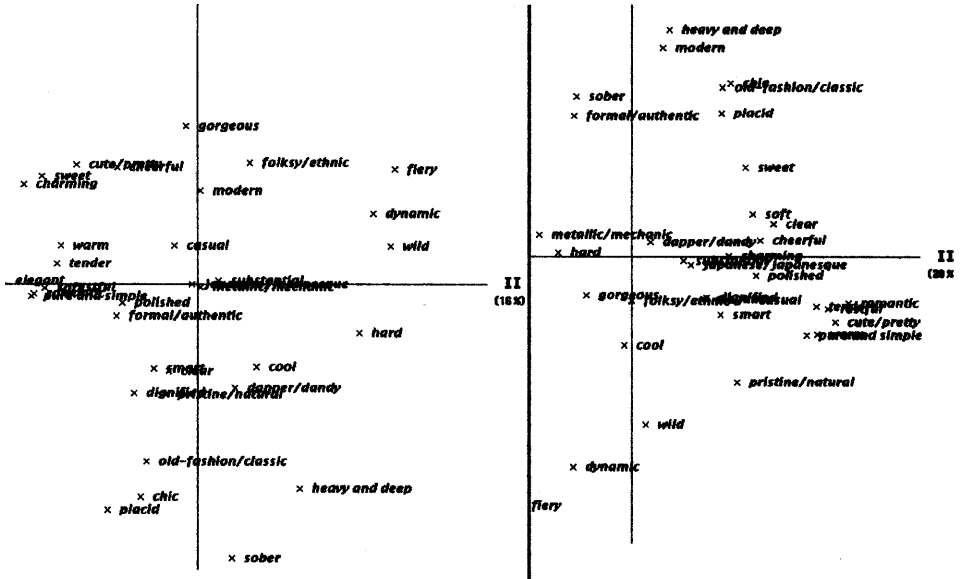


図3 印象語の地図 (左: 女子学生グループ, 右: 男子学生グループ)

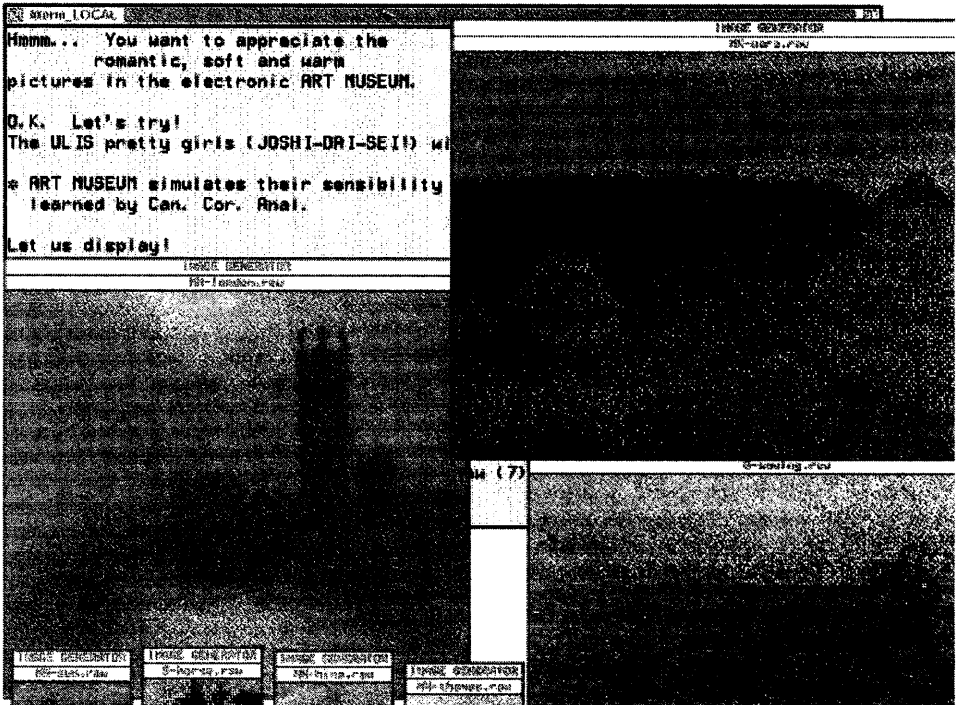


図4 印象語表現による感性検索

習用絵画 20 枚が提示されたが、このような異なる結果が2つのグループから得られた。例えば、女子学生グループの結果では、「地味な、落ち着いた」と「ゴージャス、激しい」、「可憐な、エレガントな、プリティな」と「ハード、ダイナミック、重厚な」が対極する概念であるとわかる。

この結果だけで「男女の感性の差」を論ずることはできないが、それぞれのグループに対して、異なったモデルを採用しなければならないことは明らかである。

#### 4. 印象語と感性検索

印象語による表現から、その言葉にマッチした絵画を自動的に探し出す場合を考えよう。システムは、その利用者の感性（主観的評価基準）に基づいて印象語を解釈し、また、印象語に相応しい画像特徴をもつ絵画を検索する必要がある。このような検索方式を、感性検索 QBD (query by subjective description) と呼ぶ。感性検索を実現するためには、印象語と絵画という、複数のメディアにまたがったマルチメディア処理が必要となる。

そのためには、絵画の色彩の特徴という客観的に定義し計測できる情報（画像特徴 GF）と、印象語表現という主観的評価基準を反映した情報（主観特徴 SF）の相関関係を比較的少ない数の実例から例示学習（統計的学習）により分析して、利用者の主観的評価基準のモデルとして参照すればよい。

システムは、GF 空間と SF 空間との相関が最大となるような、統合特徴空間（UF 空間）空間への写像を、正準相関分析等の多変量解析で構成する（図2）。

この方法によれば、未知の画像に対してもその画像特徴を分析すれば、利用者の主観的評価基準とマッチした印象語を推定することが可能となる。

印象が類似した画像は、UF 空間上で近い点に

写像されていると考えて良い。同様に、類似したイメージを表わす印象語  $aO$  も、回帰  $\Lambda$  により  $gO = \Lambda tF aO$  で、UF 空間上の近い点に写像されていると考えられる。

感性検索は、利用者から示された印象語による記述を UF 空間上の一点に写像し、その点の近傍に写像された絵画を距離順に求めることで実現できる（図2）。

女子学生グループの回答とサンプル絵画の色彩特徴との相関関係を、正準相関分析により統計学習させた。この視覚的印象のモデルは、女子学生グループの平均的な判断基準となっていると期待できる。

図4に、印象語「ロマンチック、ソフト、暖かい」を満たすような色彩の絵画を検索した結果を示す。視覚的印象のモデルは、非常に単純な色彩特徴と印象語との相関分析で作成したが、上位の候補の多くは、例示された印象語にかなり良くマッチしているようである。我々の予備的な実験では、被験者らの主観的な事後評価ともほぼ一致する結果が得られている。

#### 5. むすびにかえて

印象語を含む用語の体系化では、用語間の論理的客観的な関係のみならず、感性的主観的な評価尺度も現実的には重要である。我々は、マルチメディアという観点から、用語に閉じた世界ではなく、用語とメディアの開かれた世界で統合的にアプローチを試みている。

謝辞

本研究は通商産業省工業技術院のヒューマンメディアプロジェクトの一環として進めている。本プロジェクトにご助力戴く電子技術総合研究所の皆様、企業・大学からの皆様に深謝致します。

～「第11回 専門用語シンポジウム：知識の自己組織化と専門用語」から～

## ISO/TC37 におけるターミノロジーの原理

山本 昭\* YAMAMOTO Akira

### 1. ISO/TC37 の活動

ISO (International Organization for Standardization) TC (Technical Committee) 37 はターミノロジーに関する規格を扱う技術委員会で、3つの Subcommittee (SC) からなる。SC/1 はターミノロジーの原則、SC/2 はターミノロジーの組織法、SC/3 はターミノロジーデータベースに関する規格をそれぞれ扱う。SCI はさらに3つの Working Group (WG) からなり、それぞれ以下の規格を扱う。

- ISO / TC37 / WG1 ISO860 Terminology Work - International Harmonization of Concepts and Terms - Multilingual
- ISO / TC37 / WG2 ISO1087 Terminology Work - Vocabulary
- ISO / TC37 / WG3 ISO704 Terminology Work - Principles and Methods

このうち、WG1 は、現在活動を休止している。WG2 の 1087 は 704 と整合性を取る必要があるため、704 がある程度固まるのを待つかたちで、改訂スケジュールを遅らせている。この夏、コペンハーゲンで行われた会議では、改訂作業中の 704 についてかなり集中的な討議が行われた。この WG3 における議論の内容を中心に、現在、ターミノロジーの原則について、どのような議論が行われているかを紹介する。

### 2. ISO CD704 に関する議論

前述のように、ISO704 は、“Terminology Work - Principles and Methods” と題された、ターミノロジーの基本原則を含んだ規格である。現在改訂に向けて Committee Draft (CD) が作成さ

れ、議論が行われている。今回の改訂作業では、基本原則の根本に関わる変更は提案されていない。しかし、議論の中では、文言上の問題とともに、現在のターミノロジーの枠組みの根幹に関わる議論に飛び火しそうな問題点も浮かび上がった。

#### 【“terminology” という語について】

英語の “terminology” という語が、多義であり、解決策が議論された。「terminology of botany (植物学用語集)」のように語彙自体を表すもの (“terminology1”), 「ターミノロジー学」 (“terminology2”), 「ターミノロジーの作業」 (“terminology3”) の 3 つの概念がある。terminology3 は「terminology work」という語を用いて区別するが、前二者は、加算名詞の “a terminology” と無冠詞の “terminology” で区別する。terminology2 に “terminology science” という語を当てる案もあったが、英語のネイティブスピーカーには “-logy” (=science) と “science” が並ぶのがすわりが悪いということで採用されなかった。これが、“terminology management” などという複合語になると混乱を招く。“terminology management” = “management of (a) terminology1” なら “terminology management” = “terminology work” になるし、“terminology management” = “management of terminology3” なら、また別の概念である。これは、実は英語特有の問題で、各国の委員 (アイルランド、フィンランド、北欧各国、仏語圏等) から、「自国語では区別できる」とのコメントがあった。

#### 【“dimension” という新概念の扱い】

下位概念へ分割する際の、分割基準を示すもの

\* 関東短期大学

として“dimension”という新概念が導入された。「自動車」という上位概念を「燃料」という“type of character”で分割すると、「ガソリン」、「軽油」という“dimension”があり、「ガソリン車」、「ディーゼル車」という下位概念に分割される。“type of character”と“dimension”はファセット分類における“facet”、“isolate”に対応する概念である。“dimension”という語の使用にあたっては、図書館学で用いられる“category”“isolate”も考慮された。

### 【定義の表記の問題】

現行の規格では、「定義は・・・“statement”である」としている。ところが、機械部品等で図による定義が使われる。文章の定義が主体で参考として図を使うのではなく、図のみでの定義を「定義」に含めるかが問題になった。結論は出ていないが、この問題は後述する、ターミノロジーの非言語表現への対応の問題ともつながってくる。

### 3. ターミノロジーの対象範囲

現在、ISO CD704 では、Designation の項に“Terms”、“Proper Names”と並列するかたちで“Symbols”を含めている。概念の視覚的表記法としてのアイコンや、非言語的に使われた文字(「Uターン」の「U」)、抽象的記号を含めている。別の観点からは、以下のように分類できる。

- ・文字を使ったもの(頭辞語やアクロニムを除く)例:化学式,紙の規格「A4」
- ・抽象的記号 例:& ; , \$
- ・視覚的に表現したもの 例:アイコン,選択タグの記号

これらの(文字から離れたという意味で)非言語的なシンボルは、専門的なコミュニケーションの中で重要な役割を占めるようになってきたが、“term”という扱いをするには至っていない。これらをどこまでターミノロジーの対象にするかという問題が次回の改訂の際に重要な問題となる。また、仮に、これらすべてをターミノロジーの対象に含めるとすると、規格上での取り扱いや、区別を如何にするかなどの難問に直面する。ターミノロジーの実践の場では、既に非言語的表現への

対応は必要になっている。換言すれば、従来、ターミノロジーは、文字の世界の「狭義の言語」のみを使用してきたが、ノンバーバルコミュニケーションを含む「広義の言語」への脱皮が要求されているのである。

そのような中での議論の一例を紹介すると、「言語的表現」と「非言語的表現」の区別を、「アルファニューメリックで表記し得るもの」という限定で行おうという意見もあった。これは、キャラクターセットが有限であることを前提としたヨーロッパ言語の発想であり、日本語の文脈では機能しない。日本語使用者は「文字」と「記号」の間にそれほど厳密な区別を設けていない。「〒」が「文字」か「記号」かは文脈依存である。

「非言語的表現」と「言語的表現」との結び付けの必要性も検討を要する。CO<sub>2</sub>に対して「二酸化炭素」に対して「no smoking」など、等価の言語表現と結び付けるか否かが重要な論点になることと思われる。結び付けないメリットは、アイコン等の非言語的表現の利点である、(英語、日本語等の)特定の言語からの独立である。しかし、そのことは、「読み上げることのできないterm」の出現を招くことになる。また、非言語的表現の中には、「言語(文字)による表現に置き換えられないもの」が存在する可能性もある。それらを無条件でターミノロジーの対象に取り込むことは、ターミノロジーの枠組みを混沌とさせる危険がある。

さらに、「非言語的表現」には「非視覚的表現」も含まれる。beep音等の音声も、特定の概念を表記するものとして用いられている。これらの非視覚的表現は、再現の困難さ、視覚的文書での表現が不可能なことなど、主として技術的な面で、現在のターミノロジーでは取り扱いが困難となっている。これらの技術の開発とともに、これらをターミノロジーの枠組みの中で取り扱えるような原則の確立も必要であろう。

### 4. 結び

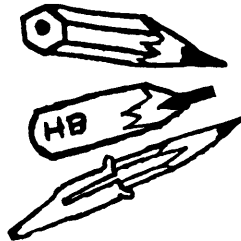
現在のところ、ターミノロジーの基本原理が根本的に問われる状態にはなっていない。しかしながら、学際的局面でのターミノロジーの実践経験



が蓄積されるに従って、現状のターミノロジーの枠組みの限界も明確になりつつあるようだ。

現在のターミノロジーの原則は、ヴュスターを起源とし、ヨーロッパ文化の中で発達してきた。ISO TC/37 も、欧米語圏のの委員が中心になっている。しかし、国際的、学際的な場におけるターミノロジーを考えると、欧米語主体の発想がど

こまで有効かは未知である。欧米語、アルファベット言語では当然のこととして受け入れられている言語現象が成立しない、非欧米語、非アルファベット圏の発想が局面を乗り越えるカギになることもあるだろう。特にターミノロジーの原則に関しては、日本の貢献は大いに求められている。



## 地図学用語辞典〔増補改訂版〕発刊の経緯

— 地図用語専門部会の発足から現在まで —

坂戸 直輝\* SAKATO Naoteru

### 1. はじめに

すべての情報媒体である地図は、伝統的な紙地図 (paper map) つまり「用紙に描かれた、また印刷された地図 (コピー、出力図等を含む)」から、電子地図 (electronic map) つまり「数値化した地図情報を、表示、検索のためのソフトウェアとともにフロッピーディスク、CD-ROM 等に収録したもの」に変わり始めている。これは時代のすう勢として当然のことといえるが、従来の紙地図は決して消滅するわけではない。これら地図に関する正しい知識を得るためには、その内容についての適確な定義を持つ専門用語を会得することが大切である。このようなことを踏まえて日本国際地図学会で今回再び世に問うたのが、標記の「地図学用語辞典〔増補改訂版〕」である。

これらについて、初版から今次の増補改訂版発刊にいたるまでの経過と学術用語集、また国際関連学会での専門用語の標準化などを含めて、説明していきたい。

### 2. 日本国際地図学会と地図用語専門部会

日本国際地図学会 (Japan Cartographers Association) は、地図に関する研究活動や情報収集、地図作成・地図情報利用技術等の発達に寄与することを目的として、1962 (昭和37) 年11月設立された。現在の会員数約1,300名。機関誌「地図」(毎月添付地図を含む) を季刊。大会・例会・専門部会活動を幅広く実施している。また、わが国を代表する地図学研究団体として国際地図学協会 (International Cartographic Association : ICA) (以下“ICA”という) と連携して国際交流の場を広げている。

地図界を支える3本の大きな柱は、地図を作る人、地図を使う人、地図の正しい使い方を教える人であるが、これら3グループの人々を傘下に持ち、お互いの間の情報交換と意志の疎通をはかりつつ今日まで発展してきた学会である。

学会創立の2年目に専門部会<sup>1)</sup>の一つとして地図用語専門部会が設置された。その主旨は、「地図用語の混乱や誤用を正し、また、できる限り用語の統一を図り、国内のみならず国際的観点に立って各国の地図用語委員会との連携のもとに、地図用語の標準化に貢献する」ことにあった。

1964 (昭和39) 年8月に第1回の専門部会を開催してから、数多くの部会、臨時小委員会を開催し、機関誌等を通してその成果を発表してきた。

### 3. 学術用語集と地図用語

部会発足当初の大きな目標は、文部省編の学術用語集と同列のものを刊行することを念頭においた。当時文部省では学術用語集の一環として「地理学用語集」<sup>2)</sup>についての編集が行なわれており、この専門委員会の委員に当時の本学会会長故野村正七氏が参画しており、しかも同氏は地図用語の専門部会員でもあった。したがって当時の文部省の青戸邦夫学術調査官からは絶えず関連資料の提供と協力が得られた。このことは部会の運営に非常にプラスとなっていた。

1965 (昭和40) 年7月に「地理学用語集」の編集に伴い文部省学術奨励審議会学術用語分科会地理学用語専門部会から「地理学用語選定原案 (最終案)」[GE-4] について、本学会あて検討の要請があった。当部会では、この用語のうち地図学に関連の用語について、数回の部会で検討審議を行ない、結論を1966 (昭和41) 年3月に報告した<sup>3)</sup>。1967 (昭和42) 年4月に作成された『地

\* 日本国際地図学会 地図用語専門部会主査

理学用語選定原案(最終案)』[GE-4]に対する修正案』[GE-6]には本学会から報告した資料が全面的に採用されていた。

#### 4. 地図学用語多国語辞典 (Multilingual Dictionary of Technical Terms in Cartography : MDC) の刊行 (以下「多国語辞典」という)

本学会では前に述べたように、国際地図学協会(ICA)の事業には積極的に協力してきており、1964年ロンドンにおけるICA第2回総会において、ICAの第II委員会として「地図学用語における定義・分類および標準化に関する委員会」(ICA-Commission II : Definition, Classification and Standardization of Technical Terms in Cartography)が設置され、加盟各国の国内委員会との協同作業による「多国語辞典」の編集が始められた。国内の専門部会員は約25名。国土地理院・海上保安庁水路部・関係官庁・大学・関係学協会・高校、民間の地図調製機関・出版会社からの参加で、第II委員会委員長のProf. Dr. E. Meynen(国内の連絡委員は筆者)から送付されてくるドイツ語の原語約1,200語に対する日本語の同義語について逐一審議し、漢字とローマ字読みについて、フィルムポジを作成し、すべての資料を1969(昭和44)年9月に委員長宛送付した。なおこれらの用語については、さきに説明の「学術用語集」との整合は、すべてとってある。この「多国語辞典」は1973(昭和48)年に初版が完成した。

28×21cm, 前文83ページ, 本文573ページで基本用語は独語で、独・英・仏・スペイン・ロシアの5ヶ国語は定義付き、チェコ・イタリア・日本(ローマ字表記と漢字)・ハンガリー・オランダ・ポルトガル・ポーランド・スウェーデン・スロヴァキアの9ヶ国語は同義語のみの計14ヶ国の用語から構成されている。見出し項目は1用語ごとに番号が付与され、主題ごとに次の10部門に分類されている。

1. 地図学とその諸部門, 2. 地図と地図シリーズ, 3. 数理地図学, 4. 地図表現方法, 5. 地図編集, 6. 地図製図・製版, 7. 地図複製, 8. 地図 [成果としての], 9. 地図の配布と利用,

0. 地図専門機関, さらに付録A : 英語・ドイツ語・フランス語・ロシア語・スペイン語で表わした地図投影の名称, 付録B : 同上5ヶ国語で表わした地図投影の分類の名称が付き, 後付として各国の標準名称および同義語のアルファベット順索引がある。前文の83ページには各国のこれに係わった学会と国内の各メンバーの氏名と所属が掲載されている。この辞典は前記のProf. Dr. E. Meynenの多大の努力によるものといえる。

この「多国語辞典」は1974(昭和49)年に改版(第2版)の編集計画が始まり初版の見直しと「地図のオートメーションに関する用語」などの追加項目が委員長から送付された。これに対応して当専門部会では再検討と審議を行ない1979(昭和54)年9月, 初版のときと同様, 改訂・追加の資料一切を委員長宛に送付した。第2版は初版のときよりも同義語を掲載する国が約10ヶ国増加し, このためか編集に日数を要していたが, 約20年を経た昨年, 別の形でドイツからProf. Dr. J. Neumann編集の名で刊行された。この間, 委員長はDr. MeynenからDr. Neumannに変わっている。なお書名は, 「地図学用語25ヶ国語辞典(改訂増補第2版)」(Encyclopedic Dictionary of Cartography in 25 languages (2nd enlarged edition))で, 内容の構成はほとんど同じである。28×15cm, 全586ページ, クロス装丁。

#### 5. 地図学用語辞典〔初版〕の刊行

本学会は1982(昭和57)年に創立20周年を迎え, その記念刊行物の一つとして, それまでの成果を集大成した「地図学用語辞典」を刊行することになった。わが国の地図界には, 今までに地図学に関する信頼できる用語辞典がなく, この刊行は長い間の要望であった。幸い当専門部会には当初から地図学に関係の専門家が参加されており, そのほか, 2. でふれた3本の大きな柱の各界の専門家が加わってその刊行計画から編集計画にと作業を進めることになった。幸いこの時点では, 「地図学用語集(案)」の各分類項目の用語については, すでに学会機関誌「地図」誌上に発表済みなことでもあり, これらの用語をきめる段階で項目ごとの定義の資料も保有していたことも幸いし,

時宜にかなっていた。それに正しい定義を伴う内容ともなれば、当初の計画である、和英・英和の「地図学用語集（案）」の一步前進ともいえた。一方ICAの地図学会議は2年ごとに開催され、当専門部会からも何回となく Technical Report 等を提出してきていたし、その点でも絶好の機会であった。今次計画の「地図学用語辞典」は、わが国で用いられている地図学用語とその定義を整理し、統一していくことを第1の目的とし、それに前記のICAの多国語辞典に所載の用語と定義をできるだけ尊重していく方針で編集を行った。

多国語辞典の中でも未だわが国では熟していない用語もあり、それについては割愛せざるを得なかった。この用語辞典は部会員・執筆会員50名の献身的協力により1985（昭和60）年7月ようやく発刊の運びとなった。企画・構成・編集・図版と索引の作成まで一切無報酬の正に手作りの辞典であった。学術用語集と同じ大きさ（B6判）、全470ページ、収録項目は同義語を含め約2,500語、本文中に関連の付図123図。ICAの多国語辞典に所載の用語については、その分類番号を見出し語の末尾に示してある。わが国の固有の用語に使用のローマ字は学術用語集にならい、また英語のつづり方は英国式つづり字を使用した。

## 6. 地図学用語辞典〔増補改訂版〕の編集計画とその経緯

地図学用語辞典の初版が刊行されて約13年、この前後の頃は、1980（昭和55）年にICAの第10回国際地図学会議が日本（東京）ではじめて開催され、また1985（昭和60）年には日本学術会議の改組に伴い地図学研究連絡委員会が独立するなど、日本の地図学が世界に伍して躍進した時期でもあった。

不統一であった地図学の専門用語がとにかく標準化され、この辞典の初版の刊行は各種の論文や著作また地図作成機関、そして地図教育などに使用することで方向づけられたことにより大きな成果をあげられた。また辞典の内容が、さきに説明したICAの多国語辞典における用語との関係が明確で、原典の翻訳や外国語による論文作成などに極めて有用であるため、好評であった。

しかしこの10年余りの間には、他の科学技術の発展と同様に地図学が成し遂げた進歩発展は極めて著しい。地図作成のための資料の収集、処理、地図表現の過程に宇宙技術やコンピュータ技術が導入されたことは大きな変革といえよう。また純粋な学問的分野においても、技術の発展と併行して、その基礎の理論の確立が地図学の重要な側面となってきた。

一方ICAでは地図の概念、理論地図学、空間データ、地図の記号論などの多くの新しい委員会や作業部会が活動して、これらに関連する新しい用語が生まれ、これからの地図学の分野で欠かせないものとなっている。

今回の「増補改訂版」はこの現状に就いての内容となっていることが、大きな特長といえる。

なお、当専門部会ではこの間、国際地理学連合(IGU)の計画による「国際地理学用語辞典」のうち「地図学に関する用語」の答申、および本学会が創立30周年記念出版の「日本主要地図集成—明治から現代まで—」の中の「地図にかかわる主要語句」の掲載に、それぞれ協力、ほか本学会の別の専門部会である、地図教育専門部会・地図学教育専門部会との連携が行われてきている。

このような状況の中で「増補改訂版」の編集計画が1995（平成7）年3月から進められた。その方針は次の通りであった。

- (1) 初版の各項目については、これを基本とし、項目自体を含め、内容をすべて再調査し、最新維持をはかる。全面的の組替え改版は経済的でないで、本文のページ数は変更しない。従って項目の内容を書き改める場合は、同じ行数内で調整する。
- (2) 五十音順の並べ方の各項の最終のところの空白部を活用して新項目をそう入する。
- (3) 初版に掲載漏れの重要項目を含め、今回新しく採択の項目は、原則的に「増補の部」として本文の用語解説の後部に書き起こして追加する。
- (4) 「増補の部」との関連を明快にするため、本文中の二つの見出し項目の間に→で、五十音順で、そこに相当する新しい項目が「増補の部」にあることを示し、一方その新しい項目名は同

じページの上部欄外に記す方式をとる。

(5) 初版刊行後に収集の編集資料は十分に活用する。

(6) 新項目：約200語、現行本(本文)の全面改訂：約65語。

項目選定がすみ、執筆者約30名、編集専門委員11名により、初版と同じ方針で編集作業を進めた。幸いこの中には初版以来部会に協力され、しかも、現在も国内・国際の地図界で活躍の会員が約半数居られ、初版の経験その他をふまえて十分な協力が得られた。

1997(平成9年)10月末、終校を出版社に返送し、これに付属物、索引等を添付してようやく完成となった。なお、今回も、青戸邦夫元文部省学術調査官からは初版以来の協力が得られた。紙上をかりてお礼を申し述べる。

## 7. 地図学用語辞典〔増補改訂版〕の内容

1998(平成10)年2月16日発刊の増補改訂版の内容は次の通りである(本ページの組方見本参照)。B6判, 545ページ, 定価6,600円(税別), 日本国際地図学会, 地図用語専門部会編, 技報堂出版

でんしアトラス 389

比べ、単位がすべて平面となるため、等高線の設定は容易になるが、他の解析、表示等の処理手順は煩雑となる。

**でんしアトラス 電子アトラス electronic atlas**  
⇒ でんしちず 電子地図<sup>(1)</sup>(増補)

**でんしかいず 電子海図 electronic chart**  
紙海図\*に対し、デジタル化された海図情報と航海援助用のその他の情報をディスプレイ上に表示し使用する航海情報システムの総称。フロッピーディスクやCD-ROM等に記録された海図情報やその他の情報と表示装置を含めたソフトウェアとハードウェアにより構成される。狭義には、デジタル化された海図情報のみを指す場合もある。(⇔電子地図(増補))

**でんしかいずひょうじシステム 電子海図表示システム electronic chart display and information system (ECDIS)**  
ディスプレイ上の航海用電子海図\*に測位情報、レーダ映像等の重量表示ができ、自船の位置や進路を海図上で把握できるとともに、浅瀬などの危険海域に近づくと警告と警報を発する機能をもつ航海情報システムをいう。国際海事機関\*が規定した性能基準を満たすものをECDIS(エクティス)の略称でよび、基準を満たさない簡易なシステムを電子海図システム(electronic chart system: ECS)と呼び区別している。

**でんしきじゅんでん 電子基準点**  
GPS\*衛星を利用した基本測量\*により水平位置を求めた基準点\*。平成8年(1996)度末で885点あり、平成9年(1997)4月刊行の地形図から電子基準点の記号も設けられた。地形図にはすべての電子基準点が表示されるが、標高数値\*は表示されない。(図-125参照)

|           |                |   |     |
|-----------|----------------|---|-----|
| △52.6 三角点 | -124.7 標石のあるもの | } | 標高点 |
| △ 電子基準点   | -12.5 標石のないもの  |   |     |
| □21.7 水準点 |                |   |     |

図-125 電子基準点

**でんしちず 電子地図 electronic map<sup>(1)</sup>; electronic atlas<sup>(2)</sup>**  
(1) 数値化した地図情報\*を、表示、検索のためのソフトウェアとともにフロッピーディスク、CD-ROM等に収録したもの。伝統的な紙地図\*に対

株式会社 発行。

収録項目：約2,700語（同義語を含む）、末尾に付表・中項目索引・欧文索引を収録、本文の説明に付随する図版（128図）

収録用語：(1)総括（関係法令を含む）、(2)地図学の理論に関する用語、(3)地図投影に関する用語、(4)地図作成の基本要素に関する用語、(5)地図表現方法（地図編集を含む）に関する用語、(6)測量および調査に関する用語、(7)地図製図に関する用語、(8)地図の製版・印刷に関する用語、(9)地図の種類に関する用語、(10)コンピュータ地図学・リモートセンシングに関する用語、(11)地図の利用・地図教育に関する用語、(12)その他

収録付表：I. 地球に使われる準拠楕円体、II. 各種準拠楕円体の使用地域、III. 地図投影分類表（4表）、IV. 主要地図投影に関する公式、V. 平面直角座標系・UTM座標系、VI. 子午線・平行圏の弧長（ベッセル楕円体）、VII. 国土数値情報の項目、VIII. 国土地理院の数値地図の刊行状況、IX. 紙加工仕上寸法・紙の原紙寸法

今回の増補の40ページ（約190語）は、コンピュータ地図学に関する用語と併せて、これからの地図学の中心となる理論地図学ならびに新技術に関する用語を主として収録してある。例えば、地理情報システム（GIS）、数値地図、空間データ、電子海図、地図の記号論、地図コミュニケーション、地図学的モデル理論、地図言語、GPS、衛星画像などについて収録してある。また、初版の際に当然採用していなければならなかった用語をもこれに加えてある。なお初版に収録の用語についても付表を含め、全ページにわたり見直しを行うとともに、13年間に概念の変ってきた約100語については、その表現を現在の内容にふさわしいものに書き改めてある。

末尾の中項目を利用し分類すれば、地図学の標準的マニュアルとしても使える。さらに欧文（主として英文）索引は全ての見出し語を掲出してあるほか、見出し項目がなく解説の中だけに出てく

る欧文用語をも、→でその欧文用語が出てくる見出し項目を示す方法をとってある。従って欧文和訳の対語表としても十分活用できる。

## 8. むすび

専門用語の部会に携って34年、当初は何から手を付けてよいか迷ったが、幸いよき部会員の参加を得て今まで続けてこられている。まず最初に手がけたのは「地図投影の名称に関する用語」であったことを思い出す。時を経ていつしか地図に関する用語のことが四六時中気になる、そのとりこになってしまった。いかに統一がむずかしいものか、どんどん新しい用語が使われはじめ、逆に今までのもので死語に近いものになっていくものもあり、中には誤った使用の用語もある。用語辞典中では〔同〕、〔古〕、〔止〕、〔誤〕などの記号で、これらを示してあるのが、これを反映させたものといえる。これが国際関連の用語ともなれば、なおさらである。他の学会関係の用語関連の委員会でも同じ悩みを持っておられるであろうと思うと、少しは気持ちが楽になる。この意味で、すべての専門用語（自然・人文）についての研究を進めて居られる「専門用語研究会」の発展は益々必要となることと信ずる。

本文は、地図学用語辞典の初版・増補改訂版の序文、編集の経緯を基本としてまとめたものである。

専門用語に関心のある諸兄に少しでもこの拙文がお役に立てば幸いである。

## 参考事項

- 1) 専門部会は現在13設置
- 2) 学術用語集 地理学編, 昭和56年3月初版発行 文部省
- 3) 文部省「地理学用語選定原案」に対する答申（昭和41年3月）8ページ、「地図」Vol.4 No.2 (1966) に掲載

## 国内刊行用語辞典リスト (1997)

### 凡 例

#### 1. 収録範囲

1997年に国内で発行された専門用語辞典(集)を収録した。新語辞典, 国語辞典, 古語辞典, 方言辞典などは収録対象としていない。

#### 2. 情報源

Kinokuniya 和書データベース Book Web ほか

#### 3. 書誌記述

記述項目は, 書名, 著者(编者)名, 出版社名, 出版年月, ページ数, 価格, 国際標準図書番号 (ISBN) とした。

#### 4. 配列

日本十進分類法 (NDC) にしたがって, 分類順に配列した。

#### 5. 例示

|       |      |                        |       |         |                 |      |
|-------|------|------------------------|-------|---------|-----------------|------|
| 分類見出し | 件名   |                        |       |         |                 |      |
| ↑     | ↑    | 書名                     |       |         |                 |      |
| 007   | 情報処理 | ↓                      |       |         |                 |      |
|       |      | コンピュータ技術基礎用語事典         |       |         |                 |      |
|       |      | 新電気編集部. オーム社. 1997-12. | 248p. | 2,800円. | ISBN:4274945952 |      |
|       |      | ↓                      | ↓     | ↓       | ↓               | ↓    |
|       |      | 著者(编者)名                | 出版社名  | 出版年月    | ページ数            | 価格   |
|       |      |                        |       |         |                 | ISBN |

#### 6. その他

このリストについてのご意見やお問い合わせは, 編集委員会までお寄せください。

---

#### 007 情報処理

##### コンピュータ技術基礎用語事典

新電気編集部. オーム社. 1997-12. 248p. 2,800円. ISBN:4274945952

##### ラジカルなコンピュータ用語辞典 第2版

岩谷宏. ソフトバンク. 1997-10. 513p. 2,200円. ISBN:4797303832

##### 困ったときのサルにもわかるパソコン用語集

森田慶子. ジャパン・ミックス. 1997-02. 197p. 1,600円. ISBN:4883213366

##### デジタル伝送用語集

電気通信協会. オーム社. 1997-01. 403p. 2,700円. ISBN:4885495075

##### パソコン早わかり用語辞典-3行でわかる&くわしくもわかる

- 学習研究社. 1997-10. 390p. 1,200円. ISBN:4054007678  
とりあえずわかるパソコン用語辞典—マニュアルやパソコン雑誌がすらすら読めるようになる  
小島邦男. オーエス出版. 1997-08. 306P. 1,500円. ISBN:4871907783  
コンピュータ用語辞典 第2版  
1997-05. 9,800円  
最新パソコン用語辞典—時代のキーワードがわかる!  
邑ネットワーク. 成美堂出版. 1997-06. 278p. 1,400円. ISBN:4415084990  
最新パソコン用語事典 '97-98年版  
大島邦夫ほか. 技術評論社. 1997-02. 733p. 1,550円. ISBN:4774104078  
新版コンピュータ用語辞典  
日本経済新聞社. 1997-02. 187p. 900円. ISBN:4532013194  
Macintosh用語事典—軽快 トラブルマスター法  
山本英司. アスキー. 1997-04. 567p. 1,850円. ISBN:4756110940  
システムアドミニストレータ用語辞典—パーフェクト2,000語  
中家裕之. プレンティスホール出版. 1997-07. 400p. 2,600円. ISBN:488735052X

010 図書館情報学

学術用語集 図書館情報学編

日本図書館学会, 文部省. 丸善. 1997-03. 165p. 2,262円. ISBN:4621043056

日・中・英対訳 図書館用語辞典

吉村善太郎. 雄松堂出版. 1997-07. 332p. 3,800円. ISBN:4841902325

図書館情報学用語辞典

日本図書館学会. 丸善. 1997-09. 244p. 3,800円. ISBN:4621043625

文書館用語集

文書館用語集研究会. 大阪大学出版会. 1997-11. 172p. 1,500円. ISBN:4872590333

021 印刷

DTP最新用語事典 編集者, デザイナー, DTPオペレータのための

帆風. アイ・ディ・ジー・コミュニケーションズ. 1997-12. 256p. 2,600円. ISBN:4872803221

210 日本史

荘園史用語辞典

阿部猛. 東京堂出版. 1997-05. 230p. 3,800円. ISBN:4490104510

319 外交・国際

国際情勢ニュースワイド (1997-1998) —最新情報&用語事典

時事総合研究所. 時事通信社. 1997-05. 356p. 1,600円. ISBN:478879716X

320 法律

中日・日中法律用語辞典

畑中和夫ほか. 晃洋書房. 1997-10. 246p. 2,700円. ISBN:4771008701



**英和契約・法律基本用語辞典**

菊地義明. 洋販出版. 1997-06. 217p. 2,200 円. ISBN:4896846516

**330 経済**

**経済学用語辞典**

佐藤武男・館野敏編. 学文社. 1997. 本体 2,000 円

**新版 英和経済用語辞典**

長谷川啓之. 富士書房. 1997-07. 874p. 3,700 円. ISBN:4938298031

**最新英和経済ビジネス用語辞典**

長谷川啓之. 春秋社. 1997-06. 1,762p. 7,600 円. ISBN:4393600061

**英和和英 経済用語辞典**

長谷川啓之. 富士書房. 1997-04. 1098,208,46p. 4,200 円. ISBN:493829804X

**新経済学用語辞典**

J.K.シムほか. 新世社. 1997-04. 327p. 2,380 円. ISBN:4915787680

**最新時事用語事典 就職試験・ビジネス常識に必須**

池田書店. 1997-11. 351p. 1,400 円. ISBN:4262151662

**335/336 経営**

**経営学用語辞典**

藤芳誠一監修. 学文社. 1997. 本体 2,427 円

**経営学用語辞典**

田島壮幸. 税務経理協会. 1997-08. 416p. 3,600 円. ISBN:4419027304

**ビジネス用語事典-決定版 仕事にすぐ活かせる**

大勝文仁. PHP エディターズ・グループ. 1997-10. 249p. 1,200 円. ISBN:4569558097

**原価計算用語辞典**

角谷光一. 同文館出版. 1997-04. 384p. 3,000 円. ISBN:449516161X

**最新会計学用語辞典**

井上良二ほか. 税務経理協会. 1997-03. 316p. 2,600 円. ISBN:441902707X

**基本簿記用語辞典**

中沢弘光ほか. 同文館出版. 1997-03. 254p. 1,800 円. ISBN:4495155121

**338 金融**

**金融ビッグバン用語辞典 大競争時代を読みこなすキーワード 170**

金融財政事情研究会. きんざい. 1997-11. 329p. 2,600 円. ISBN:4322150411

**金融・経済用語辞典 新訂**

吉野昌甫. 経済法令研究会. 1997-04. 511p. 3,495 円. ISBN:4766813634

**339 保険**

**保険用語辞典**

広海孝一. 日本経済出版社. 1997-07. 224p. 900 円. ISBN:4532013461

**340 財政**

**税務用語事典**

新井隆一監修. 日本税理士連合会編. ぎょうせい. 1997-06. 1,096p. 8,000円. ISBN:4324050066

360 社会学

**社会学用語辞典**

鈴木・森岡・秋元・安藤監修. 学文社. 1997. 本体 2,427円

**社会福祉士・介護福祉士のための用語集**

古川孝順. 誠信書房. 1997-06. 576p. 3,300円. ISBN:4414605210

370 教育

**現代保育用語辞典**

岡田正章他編. フレーベル館. 1997-02. 587p. 7,767円. ISBN:4577811308

**すぐに役立つ教職教養用語事典 (99年版)**

下村哲夫. 一ツ橋書店. 1997-10. 251p. 940円. ISBN:456599173X

**はじめての日本語教育「基本用語事典」**

高見沢孟. アスク講談社. 1997-11. 274p. 2,500円. ISBN: 487217061X

400 自然科学

**8ヶ国科学用語事典**

久保田博南. 講談社. 1997-02. 292p. 820円. ISBN:4062571579

**英和・和英工業技術用語集**

藤岡啓介. 日本アソシエーツ. 1997-10. 1,080p. 8,900円

430 化学

**分析化学用語集**

日本分析化学会, 文部省. 丸善. 1997-03. 378p. 2,800円. ISBN:4621043021

469 人類学

**人類学用語事典**

渡辺直径. 雄山閣出版. 1997-11. 305p. 5,800円. ISBN:4639014791

490 医学

**消化器集団検診用語集**

久光茂. 医学書院. 1997-06. 2,200円

**手の外科学用語集**

日本手の外科学会. 南江堂. 1997-05. 201p. 3,000円. ISBN:4524216456

**英和・和英栄養・食糧学用語集**

日本栄養・食糧学会. 建帛社. 1997-05. 612p. 3,800円. ISBN:4767960738

**消化器内視鏡用語集**

日本消化器内視鏡学会用語委員会. 医学書院. 1997-05. 3,500円

**日常会話医療用語集**

ミクス. 1997-04. 313p. 5,000円. ISBN:489587107X

**外科手術用語集**

日本外科学会. 金原出版. 1997-04. 89,38p. 5,000 円. ISBN:4307201299

**医療・病院管理用語事典**

日本病院管理学会. ミクス. 1997-07. 189p. 2,800 円. ISBN:4895871088

**衛生・公衆衛生学用語事典**

楯博. 医学出版社. 1997-03. 420p. 4,700 円. ISBN:4870550636

**わかりやすい膠原病・リウマチ用語事典 患者さんと医療関係者のために**

竹原和彦. 診療新社. 1997-01. 200p. 3,398 円. ISBN:4915917495

500 工学・技術

**学術用語集 計測工学編**

計測自動制御学会, 文部省. コロナ社. 1997-07. 622p. 3,900 円. ISBN:4339031

**特許技術用語集**

特許技術用語委員会. 日刊工業新聞社. 1997-04. 258p. 2,400 円. ISBN:4526039993

514 土木

**道路用語辞典**

日本道路協会. 丸善. 1997-04. 1,116p. 25,000 円. ISBN:4621043145

519 環境

**環境用語事典**

横山長之・市川惇信共編. オーム社. 1997-05. 452p. 4,500 円. ISBN:427402346X

520 建築

**木造建築用語辞典**

小林一元他. 井上書院. 1997-03. 484p. 3,200 円. ISBN:4753000842

530 機械工学

**機械用語大辞典**

実践教育訓練研究協会編. 日刊工業新聞社. 1997-11. 1,086p. 28,000 円. ISBN:4526040886

538 航空学

**航空実用事典—航空技術・営業用語辞典兼用**

日本航空広報部. 朝日ソノラマ. 1997-05. 473p. 3,800 円. ISBN:4257035021

544 電気

**電気工事士用語辞典**

榭井和明. 日刊工業新聞社. 1997-04. 277p. 2,000 円. ISBN:4526040037

547/549 情報通信

**日英西 情報通信技術用語辞典**

日本電気. 朝倉書店. 1997-03. 1,637p. 24,000 円. ISBN:4254222122

これで納得インターネット用語事典ーカンタン覚え方つき

藤田英時. ナツメ社. 1997-05. 175p. 1,200円. ISBN:4816322205

最新 インターネット用語事典

堤大介. 技術評論社. 1997-07. 463p. 1,680円. ISBN:477410468X

わかりやすいインターネット用語辞典ーいろいろ読んだけれどわからなかった人のための

佐藤信正. ナツメ社. 1997-08. 303p. 1,500円. ISBN:4816322035

英日中電子・情報用語辞典

工業調査会. 1997-10. 1,605p. 18,000円. ISBN:4769370547

半導体製造装置用語辞典 第4版

日本半導体製造装置協会. 日刊工業新聞社. 1997-11. 547p. 4,600円. ISBN:4526040908

566 鑄造

鑄物用語辞典 英独和 独英和

村井香一. いちい書房. 1997-04. 663p. 4,800円. ISBN:4759101004

ハンダづけ用語辞典

川口寅之輔. 日本アルミット. 1997-04. 229p. 3,800円. ISBN:4931031064

574 セラミックス

セラミックス用語辞典

日刊工業新聞社. 1997-09. 496p. 7,600円. ISBN:4526040622

589 ファッション

ファッション産業ビジネス用語辞典

ファッション総研. ダイヤモンド社. 1997-04. 374p. 3,200円. ISBN:4478501394

590 家事

コスモ「実践」家政・生活系教育用語辞典

飯田朝子他. ブレーン出版. 1997-01. 173p. 1,553円. ISBN:4892425613

596 料理

中国料理用語辞典 決定版 2版

井上敬勝. 日本経済新聞社. 1997-11. 342p. 1,400円. ISBN:4532162378

610 農学

英中日畜産学用語集

農林水産技術会議. 農林統計協会. 1997-04. 11,000円

619 農産製造

独・英・日 製粉・製パン・製菓用語辞典

長尾精一監修. 清水弘. 製粉振興会. 1997-07. 569p. 12,000円. ISBN:4384000782629

629 造園

**芝草管理用語辞典**

真木芳助. 一季出版. 1997-05. 393p. 3,800 円. ISBN:4872650344

673 商業経営

**営業・マーケティングキャリア用語辞典**

名東孝二他. 税務経理協会. 1997-03. 178p. 1,500 円. ISBN:4419022620

**宅建試験用語事典 (1997 年)**

坂本一洋他. 学太陽書房. 1997-03. 396p. 3,107 円. ISBN:4313526129

674 広告

**新版 広告用語辞典 (日経文庫 267)**

日経広告研究所. 日本経済新聞社. 1997-04. 166p. 900 円. ISBN:4532012678

675 商品・配給

**基本ロジスティクス用語辞典**

日本ロジスティクスシステム協会. 白桃書房. 1997-04. 318p. 2,500 円. ISBN:4561741003

676 相場

**先物語典－誰でも分かる最新相場用語事典 相場に勝つためのキーワード**

ゼネックス. 1997-04. 204p. 2,000 円. ISBN:4795261180

676.9 不動産取引

**不動産用語辞典**

日本不動産研究所. 日本経済新聞社. 1997-07. 182p. 900 円. ISBN:4532012465

**不動産取引用語辞典**

不動産適正取引推進機構. 住宅新報社. 1997-06. 364p. 2,190 円. ISBN:478921091X

740 写真

**写真用語事典**

上野千鶴子他. 日本カメラ社. 1997-06. 573,41p. 4,400 円. ISBN:4817900091

760 音楽

**かんたん理論 ロック音楽用語辞典**

渡辺博海. シンコー・ミュージック. 1997-07. 96p. 970 円. ISBN:4401615697

778 映画

**エヴァンゲリオン用語事典**

エヴァ用語事典編集局. 八幡書店. 1997-03. 186p. 1,200 円. ISBN:4893503227

786 登山

**仏和山岳用語集**

村西博次. 京都山の会出版会. 1997-05. 182p. 2,200 円. ISBN:4888483744

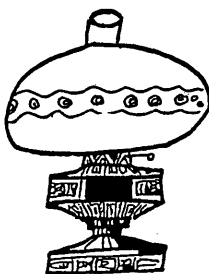
791 娯楽

ポケットモンスター用語辞典

ティーツ出版. 1997-11. 158p. 850 円. ISBN:4900700568

RPG 用語辞典-ファンタジー・ファイル (富士見文庫)

富士見書房. 1997-09. 237p. 520 円. ISBN:482914341X



## 国内刊行用語辞典リスト 電子版

### 五カ国語対訳特許用語辞典

パテトン社. 108,000 円

### 現代用語の基礎知識 '95 電子ブック版

自由国民社. 3,950 円

### 現代用語の基礎知識 '95 EPWING 版

自由国民社. 8,000 円

### 英・独・日 最新科学技術用語辞典

三修社. 19,800 円

### 絵と音が出る音楽用語辞典

リットーミュージック. 16,800 円

### 電子ブック版 朝日現代用語・知恵蔵 1996

朝日新聞社. 5,200 円

### 情報技術用語集 CD-ROM

阿部出版. 20,000 円

### マイクロソフト コンピュータ用語辞典第 2 版 CD-ROM 版

アスキー. 7,800 円

### ブリタニカ国際大百科事典 電子ブック小項目版 (現代用語収録)

TBS ブリタニカ. 130,000 円

### 入門コンピュータ用語解説 What's a Computer?

富士通ラーニングメディア. 18,000 円

### ビジネス法律用語の基礎知識 電子ブック版<増補版>

自由国民社. 4,500 円

### ビジネス法律用語の基礎知識 EPWING 版

自由国民社. 4,500 円

### 電子ブック 25 万語医学用語大辞典 英和/和英対訳

日外アソシエーツ. 28,840 円

### 電子ブック 科学技術用語大辞典 英和/和英対訳

日外アソシエーツ. 9,700 円

### 電子ブック データベース 情報検索用語辞典

日外アソシエーツ. 9,800 円

### 電子ブック コンピュータ用語辞典 第 2 版

日外アソシエーツ. 1997-05. 9,800 円. ISBN:4816970371

### 経済用語事典

三修社. 32,200 円

### CD-ROM 版 経済新語辞典

日本経済新聞社. 3,800 円

### 電辞海 for Windows

富士通. 電辞海: 24,000 円, 専門用語辞書 (6 種類): 各 50,000 円

### 最新ビジネス・技術実用英語辞典英和/和英 電子ブック版

日外アソシエーツ. 9,800 円

**最新医学大辞典スタンダード版 CD-ROM**

医歯薬出版. 28,840 円

**電子ブック 最新医学大辞典**

医歯薬出版. 15,450 円

**三省堂ワードハンターマルチ ROM 辞典 (Mac 版)**

クォリタスジャパン. 69,000 円

**ビジネス法律用語の基礎知識 電子ブック版<増補版>**

自由国民社. 4,500 円

**ビジネス法律用語の基礎知識 EPWING 版**

自由国民社. 4,500 円

**電子ブック 25万語医学用語大辞典 英和/和英対訳**

日外アソシエーツ. 28,840 円

**電子ブック 科学技術用語大辞典 英和/和英対訳**

日外アソシエーツ. 9,700 円

**McGraw-Hill Science & Technical Reference Set**

McGraw-Hill Inc. 約 40,000 円 (代理店により異なる)

**McGraw-Hill Multimedia Encyclopedia of Science & Technology**

McGraw-Hill Inc. 約 150,000 円 (代理店により異なる)

**Officer's Bookcase**

Quanta Press Inc. \$ 79.95

**TERMDOK**

Walters Lexikon. 年間 \$ 985 (6,400 クローネ)

**経済用語事典**

三修社. 32,000 円

**CD コンピュータ用語辞典 第2版**

日外アソシエーツ. 1997-05. 12,000 円. ISBN:4816980644



## J I S用語規格制定の動き

情報源：“標準化ジャーナル”，1997年1月号-1997年12月号

### 制定・改正される J I S

- 土工機械-グレーダー第1部：用語及び仕様項目（制定）  
Earth-moving machinery - Graders - Part 1: Terminology and commercial specifications
- 土工機械-スクレーパー第1部：用語及び仕様項目（制定）  
Earth-moving machinery - Scrapers - Part 1: Terminology and commercial specifications
- 土工機械-ローダー第1部：用語及び仕様項目（制定）  
Earth-moving machinery - Loaders - Part 1: Terminology and commercial specifications
- A8308 土工機械-基本機種の用語（改正）  
Earth-moving machinery - Basic types - Vocabulary
- B0660 表面粗さ-用語-第1部：表面及び表面粗さパラメータ（制定）  
Surface roughness - Terminology - Part 1: Surface and its parameters
- B3000 F A用語(改正)  
Glossary of terms used in factory automation
- Q1405 環境マネジメント-用語と定義（制定）  
Environmental management - Vocabulary
- R1600 ファインセラミックス関連用語（改正）  
Glossary of terms relating to fine ceramics
- W0125-1 航空宇宙-流体システムに関する用語-第1部：圧力に関する一般用語及び定義（制定）  
Aerospace - Fluid systems - Vocabulary - Part 1: General terms and definitions relating to pressure
- W0125-2 航空宇宙-流体システムに関する用語-第2部：流れに関する一般用語及び定義（制定）  
Aerospace - Fluid systems - Vocabulary - Part 2: General terms and definitions relating to flow
- W0125-3 航空宇宙-流体システムに関する用語-第3部：温度に関する一般用語及び定義（制定）  
Aerospace - Fluid systems - Vocabulary - Part 3: General terms and definitions relating to temperature
- X0013 情報処理用語（図形処理）（改正）  
Glossary of terms used in information processing (computergraphics)
- X0024 情報処理用語（計算機統合生産）（制定）  
Glossary of terms used in information processing (computer-integrated manufacturing)
- Z0106 パレット用語（改正）  
Glossary of terms relating pallets

官報告示・発行された J I S

- A1526:97 ドア用語(制定)  
Doors terminology
- B0111:97 プレス機械—用語(改正)  
Presswork machinery - Vocabulary
- B0114:97 木材加工機械—用語(改正)  
Wood working machinery - Vocabulary
- B0144:97 電子部品実装ロボット—用語(制定)  
PCB (printed circuit board) assembly robots - Vocabulary
- B0152:97 クラッチ及びブレーキ用語(改正)  
Clutches and brakes - Vocabulary
- B0155:97 工業プロセス計測制御用語及び定義(改正)  
Glossary of terms used in industrial process measurement and control
- B0171:97 ドリル用語(改正)  
Twist drills - Vocabulary
- B0175:96 ブローチ用語(改正)  
Broaches - Vocabulary
- B3000:97 F A—用語(改正)  
Factory automation - Vocabulary
- B8650:97 プラスチック加工機械(改正)  
Plastics moulding machinery - Vocabulary
- C0161:97 EMCに関するIEV用語(制定)  
International electrotechnical vocabulary: Electromagnetic compatibility
- C8960:97 太陽光発電用語(制定)  
Glossary of terms for photovoltaic power generation
- F0010:97 造船用語—一般(改正)  
Shipbuilding - Vocabulary - General
- F0011:97 造船用語—船体—基本計画(改正)  
Shipbuilding - Vocabulary - Basic design of hull part
- F0012:97 造船用語—船こく構造(改正)  
Shipbuilding - Vocabulary - Construction of hull part
- F0026:97 造船用語—機関ぎ装(改正)  
Shipbuilding - Vocabulary - Machinery fittings
- Q0030:97 標準物質に関連して用いられる用語及び定義(制定)  
Terms and definitions used in connection with reference materials
- T0101:97 福祉関連機器用語(義肢・装具部門)(改正)  
Glossary of terms used in prosthetics and orthotics
- T9232:97 ストーマ用品に関する用語(制定)  
Glossary of terms used in ostomy aids
- X0009:97 情報処理用語(データ通信)(改正)  
Glossary of terms used in information processing (data communication)

- X0016:97 情報処理用語(情報理論) (改正)  
Glossary of terms used in information processing (information theory)
- X0017:97 情報処理用語(データベース) (制定)  
Glossary of terms used in information processing (database)
- Z6015:97 エレクトロニックイメージング用語(制定)  
Electronic imaging - Vocabulary



## 「専門用語研究」投稿規定

1. 「専門用語研究」(以下会誌という)には、下記の内容に関する論文・記事を掲載する。

- ・ターミノロジーの理論と応用
- ・専門用語集の作成技術
- ・その他、専門用語に関するもの

2. 会員は、会誌に自由に投稿することができる。編集委員会からの依頼により執筆することもできる。

### 3. 原稿の書き方

#### 3.1 原稿用紙

原稿は、通常のA4サイズ横書き原稿用紙(20字×20字)か、ワードプロセッサを使用する。ワードプロセッサを使用する場合は、A4用紙に1行20字、20行で作成し、印刷する。また、可能であれば、MS-DOSテキスト形式でフロッピーに保存し、印刷物と一緒に送付する。

#### 3.2 原稿の長さ

全体で図表ほかを含めて、原稿用紙16枚から32枚とする。原稿用紙4枚で刷り上がり1ページとなる。執筆依頼時に別途指定ある場合はそれに従う。

#### 3.3 原稿の仕様

原稿には、以下の内容を記入する。

- ・和文と英文の、表題、著者名、所属
- ・和文の、抄録(250字前後)とキーワード(5から10語)  
(可能ならば、英文の抄録(150語前後)とキーワード(5から10語)も)
- ・本文(ページをつける)
- ・図表など(番号と表題をつけ、朱筆で文中に挿入位置を指定する)
- ・引用文献(本文中に肩付き数字<sup>1)2)</sup>・・・を記入する)
- ・参考文献、参考図書(本文を読む上で参考になるものがあれば)

### 4. 原稿の受理、査読

投稿原稿は、当研究会事務局が受け付けた日を受付日とし、会誌編集委員会で査読を行なう。査読結果をもとに、会誌編集委員会で掲載の可否を決定する。委員会で内容・表現などについて修正が必要と認めた場合、執筆者に修正依頼する。

### 5. 校正依頼

執筆者に初校を依頼する。この際、大幅な修正・加筆は行なわないこと。なお、論旨に差し支えない範囲で、編集委員会が内容の変更を求めることがある。

### 6. 掲載原稿の扱い

会誌に掲載された原稿、フロッピーは返却しない。

### 7. 謝礼

執筆者には、掲載された会誌10冊を無料贈呈する。これ以上および抜刷を希望する場合は、有料となる。校正時に申し込むこと。

### 8. 著作権

本誌に掲載された論文、記事の著作権は、当研究会に帰属する。

### 9. 原稿提出先

専門用語研究会会誌編集委員会

---

 編集後記
 

---

◆残念ながらシンポジウムには参加することができなかった。「知識の自己組織化」とむずかしい題名がついたが、人間であれば誰でも頭の中で日々行っていることを指す。さまざまなことを見て、聞いて、感じて、単なる「事象」はやがて複雑な知識になる。個々の事象は関連付けられ、足し算の集合体から、掛け算あるいは無限の集合体へ変わっていく。人間はこの知識を駆使して、あれこれ考えをめぐらせる。知識の自己組織化は、頭の中の活動のため簡単には説明されない。頭で頭のことを考える。大変なのは当然だ。

では、この動的だが、具体的にはどうなっているのか分からないことを専門用語にどう関連、応用させるのか。このへんの糸口がシンポジウムで少しでも見いだされればよいと考えている。専門用語は用語集という図書やデータベースとしての存在が目立つように思う。標準化、統一化という側面ではある程度実績をあげている。これは言葉の単なる表、足し算程度の使い勝手だと思う。意味的な概念的な側面、知識的な考えも必要だと感じている。今までとは異なる「専門用語知識集」ができあがることを期待している。(四ノ宮)

◆このところ多忙なこともあり、用語関連の研究とは縁遠い生活を送っています。しかしながら、研究者ディレクトリをはじめとするいくつかの学術研究活動に関するデータ分析に携わっていることもあり、日常的に幅広い分野の専門用語と付き合い合っています。こうしたデータを見ておきますと、先端の用語の正確な概念を把握しているわけではないのですが、学術研究のフロンティアの姿がおぼろげながら見えてくるような気がします。データの中に含まれている最先端の用語の分析を通じて各分野の研究活動の実態を把握する方法はないものかと思案しているのですが、最先端の用語というものは、当然ながらまったく統制されておらず、表記もばらばらで、手をつけられずにいます。何かいい方法はないのでしょうか。

(山下)

◆編集という仕事の中には、素材集めからはじまっていろいろある。当研究会誌では、原稿を削除することはまずないけれども、映画のような場合には、いかに切り捨てるかが、作品の出来不出来を左右するとさえいわれている。ところが、球技の中継などで野球のような単純な画作りでなく、サッカーやラグビーでは、カメラマンがその競技を知らないでボールから離れて走る選手の動きこそ勝敗を決めるのに大事な部分を撮影していないことがある。そのような素材でVTR編集する羽目になると、編集者は泣きである。同じことがテープ編集についてもいえる。未熟な編集者は、より良い作品をいかに画と音で表現し、生み出していくかが、わからない。ディレクター、演出家の腕の見せ所でもある。

文字表現についても、同様なことがいえそうである。ある文章を書いて、どうも意あって力足らず、何かが抜けている、不足であると思うことしばしばで、それでいていりもしない部分が入っていることに気がつかない。作家の綱淵謙錠は高等学校の同期生だが、彼は長い間、編集者として著名な作家の文を読まされたことが栄養になり、基礎となって作家になれたのだと、直木賞をとった時、同窓会で話をしたことは、印象深く私の心に残っている。

今、聴覚障害を持つ方々のために、日本のテレビジョン番組に日本語で字幕を入れる仕事もしているが、話す言葉を読む言葉に直すのは、意外にむずかしい。ドラマでなく、情報、教養物では、話す人が素人であり台本もないのだから、ご本人は説明をしているつもりでも、耳を聞く限りなんとなく通りすぎるにもかかわらず、文字にすると意味をなさないような場合が、時々ある。その際、問題は、適切な文が書けず、説明過剰もしくは要約しすぎの危険がある。

東京外国語大学教授 井上史雄著『日本語ウォッチング』(岩波新書)には、ラ抜き言葉や省略化「しています」が「している」、「してる」の話が面白く書かれている。これも字幕ではやはり「ラ」は付け、「します」にしている。「して」とは書いていない。少なくとも私に関してではある。

とにかく勉強を怠りなくしては、何事もつとまらないわけだが、それだけに編集という仕事はたい大変おもしろい。(中山)

**編集委員会委員**

太田 泰弘 文教大学  
四ノ宮明夫 大正製薬  
戸塚 隆哉 KMK デジテックス (委員長)  
中山 亮一 リョウ・プロダクション  
牧野 正久 東京理科大学  
山下 泰弘 学術情報センター  
山本 昭 関東短期大学

**専門用語研究 第15号**

(1998年3月31日発行)

発行所 専門用語研究会

〒102-0082 東京都千代田区一番町4-6

一番町中央ビル 2 F

日本総合技術研究所(JIST)内

Te1. 03-3262-8956

Fax. 03-3262-8960