

科学技術用語の管理の現場から

——キーワード索引支援辞書の作成について——

日本科学技術情報センター
用語管理部門

岡野 弘行

JICST では、化学技術文献のデータベースを提供するため、外部協力員による日本語抄録の作成及び専門職員による校閲索引（抄録の校閲とキーワードの付与）のように、専門家に依拠する形でデータベースを作成しています。

これはデータベース作成が知的な集約作業であり、省力化が難しいためです。

抄録作成については、自動抄録も試みられ、最近では英文からの自動抄録システムも報告されていますが、現実的には学術論文については著者抄録の徹底と標準化によって、省力化の問題を解決しようとする方向へ移行しつつあります。

索引作業についても、JICST でも抄録や標題からのキーワード（KW）自動抽出と専門職員による KW 付与の 2 本立てに頼っていますが、人手作業の限界に達しつつあります。

但し、論文記事に付与・索引する KW としては、当該記事に出現する名詞すなわちテキスト語（自然語）が良いか、シソーラス等に代表される統制語が良いかについて長い論争の歴史が続いています。

前者はコストは安く、適合率が高いが、言葉

（表記）のバラツキが多いため再現性が低い。言い換えれば、思い付いた適当な言葉で「そのものズバリ」の論文が探せますが、探し洩れが多いという欠点があります。

後者は専門職員による人手作業であるため、コストは高いのですが、一般には再現性も高いとされています。即ち、洩れが少ない（網羅的に検索できる）という利点があります。

手作業による KW 抽出（索引）においては、これまでは後者を採用する流れが主流でしたが、データベース編集に計算機が導入され、かつ分かち書きに代表される言語処理が発達するに従って、適合率と再現率の論争が解決されぬままに前者が採用されつつあるようです。

しかしながら、データベースの高度化には、上記のような両者のギャップを埋めるため、テキスト語（自然語）をシソーラス語（統制語）に変換することがどうしても必要となってきます。

その様な事情を考慮した上で、JICST では自然語から統制語への変換をめざして、計算機システムの開発を試みています。自然語と統制語の関連付けには、現在 2 種類の方法を検討していま

す：

- 1) 自然語と統制語との共出現頻度に基づく関連づけ：

相互に関連の深い用語は「共出現」する度合いが高いとの仮説に基づく。

- 2) 自然語統制語の意味関係に基づく関連づけ：

語形の類似度によって意味の類似度を設定できるとする仮説の採用とシソーラスの同義関係・階層関係の利用

1) の自然語と統制語の共出現度による関連付については、単順な共出現度順といわゆる関連強度順とを比較した結果、前者が有効であるとの結果を得たため、それを採用しています。

2 年分の JICST ファイルから無作為に抽出した自然語 2692 語についての関連関係の可否を専門家が判定した結果、共出現頻度の第一位には 72.7% (1956 語)、第二位には 36.4% (981 語)、第三位には 21.7% (585 語)・・・の比率で関連する統制語が出現しています。

(注) 関連強度 (%)

$$= (\text{共出現頻度の 2 乗}) \times 100 \\ \text{ / FT 頻度} \times \text{CT 頻度}$$

2) の語形による関連付けでは、変換解析の基本思想 (基本的な仮定) として次の 3 点を仮定しました：

- 1) 同形異義語の否定

一部の例外を除くと、基本的には語形が同一なれば、意味は同一とします。

例：「言語障害」も「電波障害」も、同一の「障害」の下位概念として見なします。

- 2) 部分は全体の上位概念

文字列 A が文字列 B に完全に含まれるならば、A は B の上位概念か関連概念とします。特に後方に含まれる場合には、A は B の上位概念となる例が多い。

例： 非晶質スピングラス
→ 非晶質 + スピングラス

- 3) 友達の輪

文字列 A の上位概念あるいは同義概念 (準同

義、類義も含む) A' を含む文字列 C が、文字列 B にふくまれるならば、C は B の上位概念か関連概念である。

例：スピングラス ・ (上位) 磁性体

非晶質スピングラス → 非晶質磁性体

捕捉 (類義) 捕獲

熱電子捕捉 → 熱電子捕獲

- 4) 中抜き処理

中間にマスク文字が存在する場合、前後の合成文字列は上位概念となる。

例：国立国会図書館 → 国立図書館

(FIG. 1 「処理概要」を参照)

なお統制語抽出の方法としては、語頭および語尾からの最長一致法を採用しました

これは、その文字列に対して一番長い候補語を選択することであり、語頭および語尾からそれぞれ抽出し、その組合せで元の概念 (自然語) を表現しようとするもので、統制語索引では一般的な索引技法です。特に案内語としてシソーラスなどに登録する場合には、この手法が採用されています。

例：応力腐食割れ→応力腐食 + 腐食割れ

この抽出法で、一般的にはほぼ妥当な解析結果が得られますが、時に不都合な用語が抽出されることもあります：

例：パイプラス中間子 → パイプ + 中間子

(注) 「プラス中間子」なる語は、一般にはシソーラスに登録される可能性は少ない。但し、中抜き処理で「パイ中間子」も抽出される。

(注) 下線部は、シソーラスの登録語です。

(FIG .2 最長一致法)

自然語： 電気双極子モーメント

統制語： 電気双極子
双極子・モーメント

自然語： 水資源開発

統制語： 水資源
資源開発

語形が同一であれば「同義」という仮定は、逆に語形の揺れに対して全く無力であるという弱点を持っています：

例えば、

1) 漢字表記と仮名(カナ)表記

例：沸化水素 ふっ化水素 フッ化水素

2) ギリシャ文字とカナ表記

例： π 中間子 バイ中間子

3) 欧文表記とカナ表記

例：Braun 管 ブラウン管

4) 人名を冠する名詞表記の揺れ

英米人以外、とくにロシア系人名を冠した用語に多い

例：Cherenkov 計数管 Cerenkov 計数管

Bogoliubov 変換 Bogoljubov 変換

Vilsmeir-Haack 反応 Vilsmeier-Haack 反応

5) 拗音等の揺れ

例：プリューム プルूम
インタフェイス インタフェース

但し、これらは照合辞書に登録すれば解決し、それほど難しい問題ではありません。

それよりも、アクリニム及びそれを含む自然語からの統制語の抽出は難しく、現在のところほとんど打開策は見つかっていません。

なお化合物名については、部分文字列方式は、

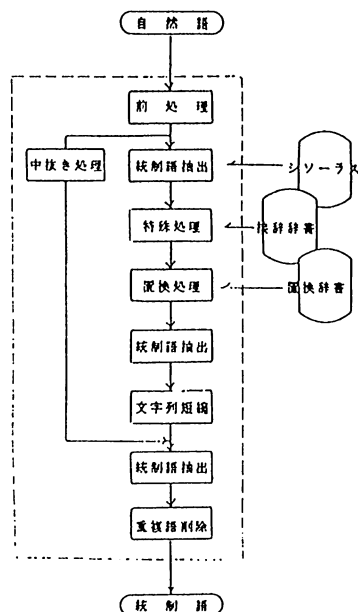


図1 処理概要

例のような JICST 表記方式とは相容れない(整合性がない)ため、別の手段が必要となります。但し、部分文字列が不要であり、かつ誤りであるか否かは別問題であり、部分文字列はそれなりに検索 KW として有効であるとの意見もあります。

例：ジェチレントリアミンペンタ酢酸

→ (抽出/変換結果)

ジェチレントリアミン
酢酸

→ (JICST 方式: 81 シソーラス)

酢酸誘導体
ジェチレントリアミン誘導体
脂肪族カルボン酸
脂肪族アミン

現在は、共出現と語形解析の2方式併用処理によって、自然語から統制語へ変換し、その変換結果を専門家が検査修正して、変換辞書を作成しています。

次の段階の構想として、ワークステーション利用の索引支援あるいは検索支援システムに変換辞書と変換ソフトを組み込み、自然語で索引あるいは検索すると、まず変換辞書(初期辞書)と照合して、既登録語は登録データで変換し、未登録の自然語についてのみ、この変換ソフトをオンライン稼働させて、適切な統制語を案内して、索引や検索に供することを計画しています。

抽出例

(重複削除前)

00000956 T I ヒト悪性メラノーマ
統制語 ヒト▲ヒト▲LS05
統制語 黒色腫▲コクショクシュ▲LS51
統制語 ヒマ▲ヒマ▲LS06 LS63
なし-BT ヒト悪性腫瘍
統制語 ヒト▲ヒト▲LS05
統制語 癌▲ガン▲LS51
なし-USE ヒト悪性黒色腫
統制語 ヒト▲ヒト▲LS05
統制語 悪性黒色腫▲アクセイコクショクシュ▲LS51

(重複削除後)

00000956 T I ヒト悪性メラノーマ
統制語 ヒト▲ヒト▲LS05
統制語 ヒマ▲ヒマ▲LS06 LS63
統制語 癌▲ガン▲LS51
統制語 悪性黒色腫▲アクセイコクショクシュ▲LS51

アメリカにおけるDTPシステムの実情調査報告

大日本印刷(株) CTS事業部システム開発第2部

斎藤 雅



1. はじめに

平成元年2月22日より11日間のスケジュールにて、(社)日本印刷産業連合会の主催による「アメリカDTPシステム実情調査団」に参加した。この調査団は日本印刷産業連合会の研究会である「電子編集システム研究会」の活動として行われた。調査団の目的はDTPの先進国であるアメリカの実情視察を通して、DTPが企業内の紙媒体としてのドキュメンテーションの作成やOHP、スライド等によるプレゼンテーションの作成また新聞・出版における活用等にどのような役割をもって利用されているのか、そしてどのようなハード、ソフトを使用しているのかを調査し、今後の日本における印刷産業としてどのような対応と準備をしなければならないかを研究することであった。

調査団の参加メンバーは29名(通訳1名)で、印刷会社、印刷機メーカー、電気メーカー等から幅広い年齢構成の調査団となった。(図1)

2月22日、御大喪前で物々しい警戒が行われていた成田からノースウェスト008便にてアメリカ西海岸へと出発した。アメリカでの訪問先は、

シアトル、サンフランシスコ、ロサンジェルス
の3都市を回り、DTPのハード・ソフト開発会社
3社、DTPのユーザ会社7社の計10社を訪問し
た。以下各訪問企業のDTP事情を報告する。

2. DTPハード・ソフト開発会社

Aldus Corporation

DTPのページレイアウトソフトであるPageMakerを開発した会社として有名であり、「デスクトップパブリッシング」という言葉を初めて提唱した会社である。1985年7月に発表されたPageMakerは世界35カ国で25万本も使用される最も普及しているDTPソフトである。通常はアップルコンピュータ社のマッキントッシュ上で使われているが、IBM-PCおよびその互換機上で動作するバージョンも提供されている。PageMakerは色々なワープロソフトで作成されたテキスト、グラフィック図形、イメージデータを取り込んでディスプレイ画面を見ながらページ上にレイアウトを行うと共にテキストの編集やグラフィックスの位置指定、サイズの変更が自由に行えるソフトである。完成した文章は、レーザービームプリンタ(LBP)または高品質の電算写植機から出力することができる。Aldus社は製品の顧客サポートが充実していた。約70名の専任スタッフが電話のヘッドフォンを付けてマッキントッシュを前にして週7000本ものユーザからの問い合わせに対応していた。

他の製品としては高度なグラフィック図形の作成がカラーで行え、ポストスクリプトに対応した図形描画ソフトFreeHand、ビデオカメラから入力されたイメージを処理するSnapShotおよびスライド等プレゼンテーションを行うために利用

されるソフト等開発していた。

私たちが訪問した時も、プレゼンテーションルームにマッキントッシュとプロセクションが接続されておりマウスを動かしながらきれいなカラー図形を見せて説明をしてもらった。

Adobe Systems Inc.

アメリカのDTPではほぼスタンダードとして普及したページ記述言語ポストスクリプトを開発した会社として知られる。創立者の2人はゼロックス社のPARC（パロアルト研究所）でページ記述言語インタープレスの開発に携わっていた。ポストスクリプトはIBMを始め28社がOEM契約を行っている。日本では、モリサワ、日本電気、富士通、リコー、松下電器産業等が契約している。ポストスクリプトの日本語対応も行われ、日本語ポストスクリプトプリンタもアップルコンピュータ、日本電気から発売されるようになった。現在はディスプレイポストスクリプトも開発しており近い将来ポストスクリプトをベースとしたWYSIWYG（What you see is what you get）環境のDTPに利用されていくと思われる。

た、ポストスクリプトのカラー対応も行われておりAdobe社の代表的なグラフィックソフトであるIllustrator 88で作られた図形がカラープリンタへ出力できる。さらにカラーデータの4色分解が可能でLinotronicにてフィルム出力される。

ポストスクリプトがもたらしたインパクトは

- ・DTPのデータの標準化
- ・高品質出力
- ・文字とグラフィックスの統合
- ・出力装置からの独立

などがあげられる。

Apple Computer

マッキントッシュに代表されるパーソナルコンピュータメーカーでDTPのプラットフォームを提供している。特にネットワーク、グラフィック関係に強く、映像、音声を含めたマルチメディア化を進めている。

AppleのマッキントッシュがDTPの世界で広く使われる要因としてユーザインターフェイスの統一をApple主体で行っている所にある。即ち、プルダウン・メニュー・ツール・パレットメニュー

・バー、アイコン、そしてウィンドウ等のは操作性の優れた環境をユーザに提供し、さらに統一された操作環境が異なるメーカーのソフトを同じ操作により扱えるため、使うソフトごとに煩わしいオペレーションを個々に覚える作業から解放してくれた。

ハイパーカードはマルチメディア対応のツールとして今後も注目される。一例として、マセラッティ社のテクニカル・ドキュメント（パーツマニュアル）の情報出版について説明があった。

CD-ROM上にパーツデータとして、テキスト、画像、サウンド等を持ち、それらをハイパーカードによって統合し、ユーザは画面を見ながら自分の望む部品が表示されるまでブレイクダウンしてパーツの指示を行うとパーツリストを作成してくれるとのことであった。

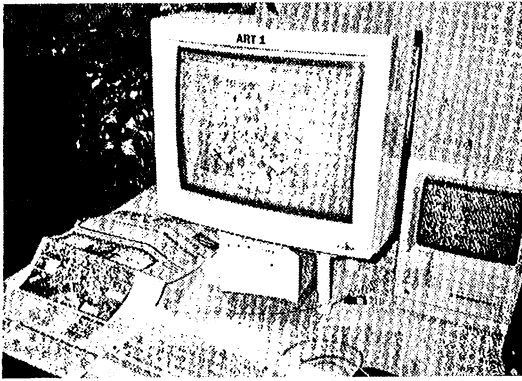
ネットワークで重要なことは「plug & play」感覚でコンピュータが接続できることである。「plug & play」とはプラグをネットワークに差し込むことですぐにコミュニケーションが行える環境ができることである。3. 3. DTPユーザ

Seattle Press Publishing

従業員6名の地方新聞社であり、隔週紙のシアトルプレス他6誌を発行している。DTPを使用する前は、一人で12ページ/月しか生産出来なかったが、今では一人で48ページ作れるようになった。ここではマッキントッシュとPageMaker、Microsoft Word、FreeHand、等を使っており文書とイラストの作成と編集を行い写真は貼り込んでいた。版下は300 dpiのLBP出力そのままを使う。また高品質な版下が必要なときは印画紙出力を外注で依頼していた。

Consolidated Press

新聞、書籍、カタログの印刷を行っている印刷会社。前のSeattle Press Publishing社で作成した版下から製版、印刷を行っている。DTPと印刷とで問題となることにカラー画像の問題がある。1つはそのデータ量が膨大になり設備、処理時間等コストの問題で、もう1つはDTPのディスプレイ画面の色と印刷時の色とのマッチングの問題であった。



Seattle Post Intelligencer

約 213,000 部の新聞を発行しているハースト財閥の新聞社である。現在印刷部門はなくライバル会社の Seattle Times が所有する印刷会社で印刷をしている。新聞の編集は専用のシステムを使用して文字データを伝送している。一部イラストはマッキントッシュ上でグラフィックソフトを使い作成していた。(図2)

IDG Communications, Inc.

コンピュータ・パソコン関係の情報出版会社で、「PC WORLD」「MAC WORLD」「PUBLISH!」などを出版している。「PUBLISH!」はその日本語版が今年より日本で出版されている。英語版 PUBLISH! では、文書、イラストは 100% DTP 化しておりカラーイメージは現在色々とテストを行っていた。DTP が組版ルールの知識のないデザイナーによって使用されることでタイポグラフィの質の低下が起こりつつあるとのことであった。また、コンピュータをいかに使いこなせるかよりデザイナーとして、また写植マンとしていかに優れているかが重要であるとのことであった。

Krishna Copy Center

全米を中心にチェーン化されたコピーサービス店であったが現在は DTP の出力サービス (スライド、印画紙、フィルム)、システムのレンタル、トレーニング等業務が拡大している。店内にはマッキントッシュが 18 台に LBP、Linotronic 300 がネットワーク化され多くの人がマッキントッシュを前に文書作成を行っていた。

Pacific Gas And Electric

カリフォルニア州の電気ガス供給会社で従業員 26,000 人。12,000 台の IBM パソコンと 1、

000 台のマッキントッシュを導入している。DTP としては、IBM 機 100 台、マッキントッシュ 300 台が使用され社内マニュアル、カタログ、ニュースレター等に使用されていた。また社内には ITC と呼ばれる社内コンピュータ化支援部隊がハードの情報やソフトの使い方を指導していた。

Huges Aircraft Company

軍用、商用の航空機、宇宙開発などを担当している Ground System Group を訪問した。ここでは日本の衛星「JSET」を担当したそうである。2 人に 1 台のパソコン (IBM、HP、Apple) があり WP、スプレッドシート、DTP に使用されている。実際に使用するソフトはまちまちであるが、DTP のグラフィック編集に興味をもちすぎ本来の仕事をしなくなる状況も出てきたとのことであった。

4. おわりに

今回の視察を通してまず DTP が企業内文書作成、新聞、雑誌等に広く使用されつつあることを実感した。この米国における普及の特徴を見ると

- ・マッキントッシュによる DTP は統一されたユーザインターフェイスによるユーザの使い勝手の良さ、ソフトへの慣れがプラスしてる。
- ・ページ記述言語 PostScript が普通紙、印画紙、フィルムへの出力を可能にし、DTP の標準を作った。
- ・プロフェッショナルにも使えるソフトが低価格で多種提供されている。
- ・カラーの 4 色分解に対応したソフトも出て、実用化に向かっている。但し色校正の問題が残っている。
- ・DTP が単に文書、グラフィック、画像の編集、レイアウトにとどまらずプレゼンテーションツールやビジネスフォームの作成ツールへと応用分野を広げていた。

が挙げられるが、もう一つ使用者の立場を考えると、現在のソフト、ハードのレベルを認識し、その与えられた物の中で自分たちの目指すものをいかに作るかを試行し、実際に出来る所から実用化を行っていく割り切りを感じた。

日本においては、ユーザ側の環境や認識、アルファベットよりもはるかに多い文字種がある漢字の問題、ソフトのレベル、PostScript の動向等注目していく必要があると思われる。

パソコン用電子音声発信装置

最近、パソコン用の音声発信装置が注目を集めている。数社がこの音声発信装置を市販していて、中には12万円という低価格のものもある。今回、その中の一つである「音次郎」というパソコン用音声発信装置を紹介する。「音次郎」の外見は、スピーカーを両手で持ったラッコのぬいぐるみである。

「音次郎」は、漢字かな混じり文をパソコンに入力すると、アクセントをつけて入力文を読み上げてくれる。文章を入力すると「音声韻律辞書」を参照して構文解析を行ない、音読を行なう。

「音声韻律辞書」には約4万9000語が登録されている。ただし辞書の修正はできない。音声は、男声、女声それぞれ5段階、読み上げ速度は、10段階の切替えが可能である。

文章を人間が話しているように自然に音読させるには、幾つか方法がある。「音次郎」に用いられているのは「規則音声合成」という方法である。規則音声合成は、人間の音声に含まれている音素を分離して記録しておき、発生規則にしたがって音の高低や強弱を調節して音を再生させる方法である。「音次郎」では、この規則音声合成法により音をデータとして扱い、これらのデータを調節することにより音を編集することができる。

日本語文章の中に、エスケープ・シーケンス・キャラクターを埋め込むことにより、音声や発音やアクセントを変えることができる。

音読には、「滑らか読み」、「簡易読み」、詳細読み」の3つのモードがある。「滑らか読み」は、「音声韻律辞書」を参照しながら構文解析を行ない、できるだけ人間に近い発音をするモードである。「簡易読み」は、構文解析を行わず、入力された文字をただちに読み上げるモードである。

そのため、熟語の発音が不正確になる。面白いのが「詳細読み」で一字づつ説明していくモードで視覚障害者には便利である。

実際に「音次郎」が音読するのを聞いてみると、発音はちゃんと聞き取れるがアクセントは不自然で、まるで日本語を覚えてた外国人が喋っているようである。音色も男声と女声があるのだが、自動販売機の声とあまり変わらず、人間の声には聞こえない。まだ改善の余地はありそうである。

しかし、プリンターボードに接続するだけで簡単に音声を出すことができるので、アイディア次第で色々な応用が手軽にできる。例えば、視覚障害者がテキストを入力した場合、入力されたテキストを確認するために使える。通常は「滑らか読み」モードで確認し、同音異義語の部分に対しては「詳細読み」モードで確認することができる。

また、CAIにも応用ができる。視覚と聴覚を併用すれば教育の効果も上がる。「音次郎」の現在の発音とアクセントでは無理があるが、機能が向上すれば外国語教育にも応用できるであろう。

その他、機械翻訳が実現すれば、適当な入力装置と組合せることにより通訳が可能となる。本学にも、共同研究室に「音次郎」がある。毎日、学生や教官が文章を入力し、「音次郎」に喋らせている。更に今年の卒業研究では「音次郎」を使った研究もあるらしい。また、某日、工業技術院電子技術総合研究所を訪れる機会があったが、ここにも、「音次郎」があり、研究員の方が使っていた。

一度「音次郎」のお喋りを聞くことをお勧めする。

(図書館情報大学 宇陀則彦)

σκέτο

経済大国化と知的好奇心との関係

一橋大学経済研究所・教授 松田 芳郎

日本という国は、いつのまにか経済大国というものになったらしい。何を指して経済大国というのかは、人によってさまざまであるけれども、少なくとも身の廻りのさまざまな生活用品をみていると、欧米諸国と差がないとはいえる。場合によっては、日本の方が一見快適であるかのように見える。敗戦直前にはタブロイド判一枚にまで小さくなった新聞や、ザラ紙全紙に刷りっ放しの教科書を自分で糸でかがっていた敗戦後にみた、大判のアメリカのグラフ雑誌『ライフ』アメリカの豊かさへの目をみはったのは、ついこの間の様な気がする。その『ライフ』で一番、目を見張ったのは、様々の色刷りの食料・衣料など日用品の広告であった。この驚きは、1ドル360円の時代に欧米に留学した者にとって共通のものなのではなかろうか。この頃に国立大学で助手をしていた人が、留学先で給与をきかれ、月給のドル換算を答えて、それは、「日給か」と質問されて答えに窮した話なども、この頃の日本の状況を物語る話題である。この時代の記憶が、頭に刻みこまれている者にとっては、二十代・三十代の人々が、こともなげに欧米に出かけるのは、確かに、日本も豊かになり、経済大国になったことの明かしとも思われる。ただ気になるのは、研究会などなどででかけた先の印象をきくと、「たいして新しいこともなかった」とつまらなげに言う者の増えてきたことである。そこには、教室で欧米の文献にのみ言及を留めておくと、「日本語ではないのですか」という質問に示される欧米の文献に対する冷淡さがある。

疑問は、冒頭の、本当に日本は豊かになったのだろうか、という疑問に戻る。経済大国というのは、知的な豊かさも保障しているのであろうかということである。昨年秋、ケルン大学500年記念事業の一環として、同大学で、ケルン計算機会

議 (Cologne Computer Conference) が開かれていた。日本からの出席者も何人か居り、報告やら座長を勤められていたので、知っている方もあろうが、若干かいつまん紹介しておく、人文諸学と社会科学をめぐっての計算機の利用とデータベースの学会である。どちらかというと歴史諸学を中心とした人文諸学に著しく傾斜していた。全体は、五つの主題と、28種類のデモンストレーションとに分かれていた。主題は、

A: 新しいデータベース (46 報告)

B: ソフトウェアの発展 (38 報告)

C: データ解析 (25 報告)

D: CAI (23 報告)

E: インフラストラクチャ (データベースセンター・データ保護・各種書誌目録と図書館網など) (50 報告) である。

会議そのものは、ケルン大学の実証社会調査中央アーカイヴ所属歴史社会調査センター (Center for Historical Social Research of Central Archive for Empirical Social Research) とゲッティンゲンのマックス・プランク歴史研究所 (Max Planck Institute for History) が組織して、人文諸学・社会科学データベースの1988年度国際会議 (The Internathinal Conference on Data Bases in the Humanities and Sosial Sciences of 1988 ICDBHSS) と、歴史学と計算機の第3回年次会議 (The 3rd Annual Conference of the Association for History and Computing) と IFDO (International Federation of Data Organizations for the Sosial Sciences) の第10回国際会議とを兼ねた比較的大がかりのものである。ただそのわりには、日本人の参加者が少なかった。確かに近年海外だけでなく、日本で開かれる国際会議も増えてきたし、中国を始めとするアジア諸国で開

催されるものもあり、その一つ一つに接するゆとりは、個々人にはないであろう。それにしても、全分科会を通じて2本の報告というのは、扱われている主題の包括性と比べて少ないというのが、実感である。それでは、逆に、誰が報告しようだろうかと考えてみて、報告を勧誘出来そうなあの、この人と思い浮かべてみると、それ自体まことに少ないことに思いたった。どうも、これは私自身の付き合いのせまさだけによるのではなさそうである。

量的に確認できないままにいうのは、無責任かもしれないが、この頃の日本の人文諸学・社会科学の研究者の間で、具体的な事象に対する関心が低くなっているのではなからうか。

これに対してこのケルンの会議の報告の主要なものは、やはり人文諸学の対象としているもの、あるいは社会科学の対象としているもの、そのものに関するデータベースが中心である。例えば、絵画における主題の画像の識別とデータ化であるとか、帝政ロシアの最初にして最後の人口センサスの集計表を利用して言語別人口から民俗別人口を推定するデータベースの構築の試みといったものである。こういった事象そのものに対する関心が日本では薄れていき、それに替わって抽象的理論への関心の移行があるのではなからうか。そのあらわれが、どこの国に旅行しても天が下に新奇なるものなしという感想につながっていくのではなからうか。《百聞は一見にしかず》ということわざがあるけれども、現在の状況はむしろ《見えども見えず》という心境に近いのではなからうか。とするならば、たかだか一つの国際会議への参加者の多いか少ないかでなく、研究者の研究のあり様の問題でもある。物に飢えることがなくなると一緒に、物に対する好奇心も減少するのかもしれない。このような悲観的な思いになって来たのは、実は、経済統計の分野で、実証的研究の日本における衰退に対する対策について、この所何人かの人とラウンドテーブルで話し合う機会があったときの皆の共通のなげきに影響されたのかもしれない。

◆第14回 情報学基礎研究会

(主査: 藤原譲、幹

事: 有川節夫、岩野和生、吉田郁三)

日時 平成元年7月27日(木) 10:00~16:40

会場 国立教育会館 402 研修室

(東京都千代田区霞ヶ関3-2-3、

地下鉄銀座線、虎ノ門下車、徒歩2分、地下鉄丸の内線: 霞ヶ関下車、徒歩5分、

Tel. 03 (580) 1251

議題 特集: フルテキスト・データベースのご利益はどこにあるか?

(1) フルテキスト・データベースの実用化における諸問題 学術情報センターの事例を踏まえて

根岸正光 (学術情報センター)

〔概要〕学術情報センターでのフルテキスト・データベースの形成、提供の経験を踏まえ、この種のデータベースの有用性について検討する。

(2) SGML とフルテキスト・データベース

芝野耕司 (東京国際大)

〔概要〕SGML の概要を述べ、データベース研究の文脈で SGML をもとにしたフルテキスト・データベースについて展望する。

(3) 特許情報のフルテキスト・データベース

大山勝弘 (日本特許情報機構)

〔概要〕日本特許庁のペーパレス計画と、その成果物のひとつであるフルテキスト・データベースを基にした JAPIO のサービスシステム開発計画について現状と今後の予定を報告する。

- 12:45~16:40 -

(4) STN のフルテキスト・データベースについて

廣田勇二 (化学情報協会)

〔概要〕STN で利用できるフルテキスト・データベースの特徴、検索時の留意点および現状について報告する。

(5) フルテキストサーチのハードウェア技術について

高橋恒介、山田八郎、本村真人 (日電)

〔概要〕フルテキストサーチのハードウェア技術の進展の経過と技術内容および今後の課題を述べ、特に、著者らの開発した文字例サーチエンジンの開発結果と応用開発結果を報告する。

(6) 大規模文書情報システム用テキストサーチマシンの研究

加藤寛次、藤沢浩道、

大山光男、川口久光、畠山 敦 (日立)

〔概要〕インデキシングが不要な全文検索方式と、これを用いたテキストサーチマシンのプロトタイプングについて報告する。

(7) テキストデータベース管理システム SIGMA とその利用 有川節夫 (九大)

〔概要〕標記システムの概要と基本的エンジンについて説明し、フルテキストデータベースの構築とその利用について概説する。

(8) General Discussion

「文字記録史料と電算機応用に関する課題と解決」研究集会について

歴史研究の素材として、さいきん絵画や器物あるいは口述資料への関心が高まっているものの、やはり主要には事件発生当時に関係当事者の作成した文字記録原史料にあるとよいであろう。考古学的発掘による木簡や漆紙つまり文字記録原史料の出土は後世の編纂叙述文献における記述にたいして貴重な証拠を提供するものである。

そうした文字記録原史料の残存量は、古代・中世・近世・近代と大雑把にみても、指数関数的に増加している。その理由は経過時間による消失もあるであろうが、むしろ文字で記録を作成しなければならないさまざまな組織体の増加によるものであり、またそうしたさまざまな組織体における執行業務の複雑化と先例参照の必要性によるものであろう。

また明治以降の近代歴史学の導入以来、さらに太平洋戦争以後における地方史の盛行によって日本史研究論文・著書も膨大な量になっている。研究遂行に当たっての先行文献参照のための検索が次第に困難になったきている。

文献検索の問題はひとまず措くとしても、膨大な総量となる文字記録原史料を利用するための検索手段の作成に、電算機の漢字入出力能力の活用を考慮しなければなるまい。

1989年4月1日2日の二日間をかけて、標記の研究集会が開催された。なお福武学術文化振興財団の研究助成を得たことを記して謝意を表しておきたい。研究集会の詳細は省略せざるを得ないので、参加者氏名と研究発表主題（略記）を記して、歴史学界での電算機応用の方向を示唆するに止める。なお敬称略、順序不同をお断りしておきたい。

研究集会代表 森安彦（国文研・史料館）

☆安澤秀一（国文研・史料館）「趣旨説明」

☆中野美智子（岡大付属図）「岡山藩人物情報データベース『諸職交代』」

☆川口洋（筑大学術情報処理センター）「『宗門改帳』データベース」☆永田治樹（金大付属図）／

増田元・竹内比呂也（東大付属図）「文書目録情報のデータベース化の問題」

☆阿部昭（国士館大）「文書、記録検索システムの試み」

☆山田哲好（国文研・史料館）「近世・近代史料所在情報のデータベース化」

☆石上英一（東大史料編）「史料学とデータベース」

☆永村真（東大史料編）「古代・中世の古文書・聖教類データベースの構築」

☆水口政次／水野保（東京都公文書館）「近代行政文書目録作成時の電算機応用」

☆周防節雄（関西大）「SASによる統計書誌情報データベースの構築」

☆松田芳郎（一橋大）「明治期の統計データの復元」

☆勝村哲也（京大人文研附属東洋学文献センター）

☆星野聰（京大大型計算機センター）

☆今岡典和（京大文）

☆根岸正光（学術情報センター）

☆澤村正信（神商大）

☆長瀬真理（東女大）

☆保坂一房（東京都公文書館）

☆大藤修（国文研・史料館）

☆安藤正人（国文研・史料館）

☆渡辺尚志（国文研・史料館）

☆広瀬睦（国文研・史料館）

☆倉地克雄（岡山大）

☆山室恭子（東大史料編）

☆田中康雄（群馬県立文書館）

（安澤秀一記）

情報知識学会第3回理事会 議事メモ

日時：平成元年5月9日18:00～20:00

場所：学士会館（神田本館）

出席者：

会長：米田幸夫、副会長：月見里禮次郎（凸版印刷）、藤原 譲（筑波大）

理事：石塚英弘（図書館情報大）、江成保徳（大日本印刷）、
近藤秀文（日立製作所）、菅原秀明（理化学）、高橋靖明（凸版印刷）、
中山 堯（神奈川大）、田隅三生（東大）、名和小太郎（旭リサーチ）
根岸正光（学術情報セ）、三浦勲（紀伊国屋書店）、
安澤秀一（国文学研究資料館）

配布資料

- 1) 前回議事録（案）（第2回、平成元年2月6日）
- 2) 理事会資料（昭和63年度収支明細、会員名簿）
- 3) 事業企画案（企画委員会）
- 4) 総会開催案内
- 5) 医学生物科学関連情報交換・研究支援システム Bio-Net 参加の手引

議事：

1. 前回議事録確認

根岸より、資料1)に基づき、前回理事会での議事内容を説明し、これを承認した。

2. 昭和63年度事業報告

藤原副会長より、昭和63年度の本会事業について、ニューズレターの発刊、論文誌発刊準備、情報学シンポジウム開催、2部会の発足などが総括的に報告され、本会創設初年度としては一応の成果があったと評価するべきものとした。

3. 昭和63年度会計報告

藤原副会長より、資料2に基づき昭和63年度の収支決算について報告された。支出超過分については、法人会員から特別会費として援助を仰ぐこととし、この点、既に目算がある旨発言があった。

4. ニューズレターの発行状況・論文誌編集の進行状況

石塚理事より、ニューズレター第4号を準備中であり、この号から、欄別に編集を分担して、編集の定常化を図っており、年間4号の発刊を維持する旨、報告された。また、論文誌については、SGML方式での編集を試みることとし、関連システムの見学を含め、技術的検討を行っていること、執筆依頼は発送済みであることが報告された。次回編集委員会は6月9日に開催予定。

5. CODATA 部会の活動状況

田隅副部長より、既提出の事業計画に基づき活動を準備中であること、文部省データベース作成科学研究費の継続申請を行ったこと、また「日本のデータソース」についても科学研究費の申請を検討したい旨、報告された。

6. 人文・社会系部会

安澤部会長より、史料のデータベース化に関する研究会を4月1～2日に開催したので、ニューズレターで紹介する予定であること、5月26日にテキスト・データベース連絡委員会を開催することが報告された。

7. 生物科学関連の活動状況

菅原理事より、生物系研究会の発足準備を進めていること、関係分野研究者によりBio-Net（資料5）運用を始めたことが報告された。

8. 事業企画案

江成理事より、企画委員会での事業企画案として資料3が示され、討議の結果、当面、（財）データベース振興センター、データベース協会主催の展示会「データベース'89」（9月5～7日、池袋サンシャイン）に合わせて、当学会のフォーラムを計画することにした。

9. 会員の拡充

標記に関しては、発足当初に集めた各種名簿を整理して入会案内を送付する、各種学会の大会での、当学会案内冊子の配布に協力を求める、同じく各種学会誌への案内記事の掲載に協力を求める、などを進めて行くことにした。

10. 次回理事会、1989年度総会の日程

総会：1989年5月23日 14：00～17：00 東大山上会館

（名和講師による講演会をあわせて開催）

次回理事会：1989年9月のフォーラムに合わせて開催

1989年度総会の報告は、次号に掲載します。

（学術情報セ・根岸記）

記念講演の概略紹介

「情報知識と知的所有権」

1. はじめに

5月23日、東京大学内の山上会議所で総会に先立って記念講演が行われた。講師は旭リサーチセンター取締役で本学会理事でもある名和小太郎氏で、演題は最近注目を集めている話題『情報知識と知的所有権』であった。著作権審議会の委員を歴任されている名和氏は、この複雑な分野をエンジニアの視点から明解にまとめられた。また、会場は満席で質疑応答も活発であった。もっとも、当日は都合により出席できなかった会員も少なくないと思われるので、当日の録音テープを元に後日詳しい内容を会誌に載せることになった。但し、それには未だ時間が掛かるので、先ずはその前ぶれも兼ねた簡単な紹介を書くことにした。

2. 講演の概略

著作権は権利の束と言われるほど、いろいろな権利があり、これまでの経緯もそれぞれにある。そのため、現在の制度はモザイク状で論理的統一がない感じになっている。この講演では、その制度を分析して述べるとともに、それが持つ問題点を指摘した。

知的所有権は著作物に関する権利、特許権、その他多数の権利（たとえば、商標権）とから構成されるが、これ以外に最近アメリカが重視している財産的情報（その実体は企業秘密）がある。知的所有権と財産的情報は共に有償であるが、前者はその内容が公開されているのに対し、後者は公開されていない点異なる。また、特許権、著作権、企業秘密は性格が異なっている。特許権は登録が必要であり、その結果独占使用が許され、権利の保護もある。著作権は登録不要で、独占使用は認められていないが、権利は保護される。企業秘密は登録はしないため、独占も権利の保護もなかったが、最近特にアメリカで権利の保護を認めるようになってきた。

特許権の場合は、先人の業績と本人の業績とを明確に区別し、本人の業績のみを特許請求の範囲とする。一方、著作権の場合は、先人の業績を引用という形式で含みうようになっているので、区別が不明確になることがある。著作権の場合は登録がなされないから、紛争が起きた時点で互いに証拠を挙げて争うことになる。アメリカでは著作権が侵されたか否かを同一性を判定することによって判断する。たとえば、プログラムの著作権の紛争では、何行似ているかによって判定することになる。

著作権には、著者に付いており売買できない著作人格権と、著作を作った人の権利である著作権と、使う人の権利である著作隣接権とがある。

何が著作物で、何が著作物でないかが問題となる。データは著作物ではないが、データベースは著作物とされた。なお、書誌データベースの場合の抄録はデータに著作物性があることになる。要旨は原著物と別という考え方と、コピーであって、それを集めることによって二次著作物になるという考え方とがある。

著作権は本来、人の権利であるが、この頃は法人にも認められるようになってきた。ただ、アルバイトに入力させた時はどうかとか、ボランティアの時はどうか、POS端末から入ってくる時は等々いろいろな問題がある。また、ダウンロードは複製に当たる。ただし、どれだけの量を取ったら複製に当たるかの判定は難しい。

ソフトウェアの権利保護も最近の話題である。これについてはも実はいろいろな場合がある。

特許は各国に登録し、登録済の国でのみ権利が生じる仕組みであるが、著作権は登録なしに発生し、一国で認められたものは他国でも認められる。たとえば、アメリカで著作権が認められたものは日本でも認められることになる。最近、日米間で知的所有権に関する紛争が起きているが、その原因はこの点にある。

（図書館情報大学 石塚英弘記）

「1990 年情報学シンポジウム」論文募集

目的：科学における情報の円滑な流通と高度利用を促進するため、データ・知識に関する基本的問題とその整備・利用に関する討議を行い、研究交流をはかる。

本シンポジウムは 1984 年以来毎年開催されている。

内容：データ・知識に関する課題を具体化し、その理論化と体系化をめざした下記のような論文を募集する。

1. データ・知識の整備：記述、表現、評価、識別、蓄積など
2. データ・知識の流通：標準化、媒体変換、分類など
3. 基礎理論：アルゴリズム（並列・分散コンピューティング、データ構造、分類、暗号理論）、推論、モデリングなど
4. 応用：情報工学の枠にとらわれない個別的な専門研究、データベース、AI、言語処理、ソフトウェア工学での応用（分散データベース、オブジェクトデータベース、ハイパーメディア）、統合システム構築法

共同主催（予定）：

日本学術会議 情報学研究連絡委員会
学術文献情報研究連絡委員会
学術データ情報研究連絡委員会
情報工学研究連絡委員会
情報処理学会、人工知能学会、日本医学会、日本化学会、日本数学会、
日本地理学会、日本物理学会、情報知識学会

後援（予定）：学術情報センター、計測自動制御学会、国際電信電話（株）、
情報科学技術協会、情報通信学会、電子情報通信学会、
日本医療情報学会、日本科学技術情報センター、日本機械学会、
日本金属学会、日本原子力学会、日本材料科学会、
日本材料学会、日本生化学会、日本電信電話（株）、
日本動物学会、日本農学会、日本分子生物学会、
日本分析化学会、日本薬学会、化学情報協会

日時：1990 年 1 月 17 日（水）～18 日（木） 9：30～17：00

場所：日本学術会議講堂（地下鉄千代田線、乃木坂駅下車）

参加申込み：氏名、連絡先、職名、資料必要の有無を葉書に記入し 12 月 20 日までに下記に申し込む（当日受付もあるが資料不足の際は事前登録者を優先する）

参加費（資料代として）： ①共催学協会員 5,000 円 学生 1,500 円
一般 7,000 円

講演申込方法：ワープロ使用で A4 用紙 4～10 枚の論文と題目、氏名連絡先、職名を記入した別紙を添えて下記宛に申し込む。

講演申込締切：1989 年 9 月 16 日（土）必着

採否通知：プログラム委員会で審査し、採否は 1989 年 10 月 1 日までに通知する。

最終原稿締切：1989 年 11 月 30 日（木）必着

宛先：情報処理学会 シンポジウム担当 木村保明
◀ 106 東京都港区麻布台 2-4-2 保科ビル 3F
Tel. 03 (505) 0505

「原子力におけるネットワーク活用技術研究会」

主催 東京大学工学部附属原子力工学研究施設

共催 日本原子力学会「原子力分野における計算機ネットワーク利用」研究専門委員会

日時 1989 年 8 月 24 日（木）

場所 東京大学工学部 11 号館

プログラム

開会挨拶

ネットワーク技術の現状と展望

原研におけるコンピュータネットワークの現状

米国における科学技術計算用ネットワークの現状

電力会社のネットワーク利用計画

原子力学会のネットワーク関連委員会の活動

学術情報センターの情報ネットワーク

筑波ネットワークの現状について

パネル討論「原子力研究情報ネットワークを組するためには」

名 大 仁科浩二郎

東 大 石田 晴久

原 研 原田 裕夫

東工大 高橋 亮一

電中研 墨沢 欣二

東 大 岩田 修一

学術情報センタ 飯田 紀子

研究交流センタ 高木 謹一

参加無料

連絡先 318-11 茨城県那珂郡東海村 東京大学工学部附属原子力工学研究施設

古田 一雄

第15回 ベリーラージ・データベース国際会議開催予定 (注、原題 15th Int. Conf. on Very Large Databases)

開催場所 オランダ、アムステルダム RAI 会議センター

開催期間 8月22～25日

スポンサー オランダ コンベンション ビューロー)

Netherlands Conv. Buredu (NCB)

(開催窓口) 連絡先 OBA, Ms L Groot, Europaplein X,
1078 GZ, Amsterdam, 020-440807, Netherlands

(出典) "WORLD MEETINGS" Vol. 22 NO. 2

Published by Macmillan Publishing Company. 以上

情報知識学会シンポジウム開催予定

テーマ：データベースと情報知識学を考える

(内容：各界の有識者による講演、パネルディスカッション)

講演者予定：藤原鎮男（神奈川大学教授・知識情報研究所所長）

：藤川正信（図書館情報大学・学長）

：渡辺龍雄（DPX）

：宮川隆泰（三菱総合研究所）DINA

パネルディスカッション：司 会 藤原 譲（筑波大学教授）

：パネラー 上記講演者及び

三浦 勲（紀伊国屋書店）

名和小太郎（旭リサーチセンター）

日 時： データーベース '89 TOKYO

'89年9月7日（木） PM 1:00～5:00

会 場：サンシャイン文化会館 7階

編集後記

どこかの社会学者の分析では、人間は多くの時間を情報量としては少ない無駄なおしゃべりに費やしているという。そうしたおしゃべりの目的は、単なる情報交換ではなく、それぞれの存在の主張・確認であり、その無駄によってひとつのコミュニティ、明日も維持できそうだという心の安らぎを得るのだという。学問分野の細分化が進むにつれてそれぞれの特化された分野におけるしきたり、制度、価値観を生んでいるが、努力を怠り、そこに安住してしまうと、"Gentlemen's Club" と化してしまう。小さく生まれた情報知識学会が活

性を維持していくためには不断の活動による増殖が必要で、新たに活動を開始した部会に期待するところ大である。その意味で会員各位の積極的な参加を待ち望んでいます。"ニュース" が途切れないためにも。

(岩田)

☆編集委員の兼子有史氏（東北工業(株)）が4月28日死去されました。謹んでお悔み申し上げます。