
学術情報センター紀要

Research Bulletin of
the National Center for Science Information System

第 1 号

1987年3月

学術情報センター

学術情報センター紀要 第1号

目次

創刊のことば	猪瀬 博 (学術情報センター所長)	i
研究論文		
IFLA プレコンファレンス・セミナー「多言語・多文字資料利用のための図書館利用 自動化システム」	内藤衛亮 (学術情報センター)	1
TOOL-IR 工学文献データベース (COMPENDEX, Ei ENGINEERING MEETINGS) 検索システムの開発	根岸正光 (学術情報センター)、小澤 宏 (東京大学大型計算機センター)	25
誤り訂正符号の調査票記入への応用	橋爪宏達、宮澤 彰 (学術情報センター)	61
TSS 接続による仮想画面転送 (VTSS) 方式	安達 淳、橋爪宏達、大山敬三 (学術情報センター)	73
画面指向通信用プロトコル VTSS のテスト及びプロトタイプシステム	大山敬三 (学術情報センター)	91
解説		
Technology for a Large-Scale Data Storage	安達 淳 (学術情報センター)	109
研修レポート		
駒沢大学図書館の現状と将来	阿部博則 (駒沢大学)	125
成蹊大学図書館機械化 (漢字処理) へ向けての MARC 実験	大井敏暉 (成蹊大学)	133
農学分野における情報検索の現状と問題点	為石理恵子 (京都大学)	149
文献情報センターと琉球大学附属図書館	本郷清次郎 (琉球大学)	159
Recon から総合目録データベースへ	牛崎 進 (立教大学)	167
書誌所在データベースと二次情報データベースの統合に向けて	尾城孝一 (名古屋大学)	175
中国語文献データベース化における漢字処理について	神田百合枝 (東京大学)	185
ローカルの雑誌管理システムと学術雑誌総合目録所蔵データ	辻川輝男 (関西大学)	197
東京大学文献情報センターシステムと ISSN	松野とも子 (北海道大学)	205
図書館 昨日、今日そして明日	諸富秀人 (広島大学)	211
学術情報センター日誌	219

**Research Bulletin
of
The National Center for Science Information System**

March 1987 Volume 1

Contents

<i>H. Inose</i>	i	Preface
 Contributions		
<i>E. Naito</i>	1	Issues Discussed at the "IFLA Pre-Conference Seminar on Automated Systems to Access for Multilingual and Multiscript Library Materials"
<i>M. Negishi and H. Ozawa</i>	25	Development of TOOL-IR Information Retrieval System for Bibliographic Databases in Engineering (COMPENDEX and Ei ENGINEERING MEETINGS)
<i>H. Hashizume and A. Miyazawa</i>	61	Application of Error Correcting Code to the Questionnaire Data Entry
<i>J. Adachi, H. Hashizume and K. Oyama</i>	73	Virtual Screen Transfer Method on TSS Connection (VTSS)
<i>K. Oyama</i>	91	Test and Prototype Systems for a Screen-oriented Simplified Communication Protocol VTSS
 Survey		
<i>J. Adachi</i>	109	Technology for a Large-scale Data Storage
 Seminar Theses		
<i>H. Abe</i>	125	Present and Future Condition of Komazawa University Library
<i>T. Ohi</i>	133	MARC Test toward Seikei University Library Computerization (including Kanji)
<i>R. Tameishi</i>	149	Some Recent Problems of Online Retrieval in Agriculture
<i>S. Hongo</i>	159	Center for Bibliographic Information and University of Ryukyus
<i>S. Ushizaki</i>	167	Toward a National Bibliographic Database in Japan: From the Point of Retrospective Conversion
<i>K. Ojiri</i>	175	Toward the Integration of 'Database for Bibliography and Location' and 'Secondary Information Database'
<i>J. Kanda</i>	185	How to Handle the Chinese Characters (Kanji) in Using a Computer for Research in the Document in Chinese
<i>T. Tsujikawa</i>	197	Junctioning of KULPIS to the ULSP
<i>T. Matsuno</i>	205	CBI System and ISSN
<i>H. Morotomi</i>	211	The Library: It's Yesterday, Today and Tomorrow
 Diary	 213	

創刊のことば

学術情報センター所長

猪瀬 博

近年急速に発展した情報化社会にあって、情報の重要性は学術研究の分野においても例外ではなく、的確な学術情報を迅速に収集することが研究遂行上の不可欠な手段となってきた。このような認識のもとに、昭和61年4月、学術情報センターが国立大学共同利用機関として文部省により創設された。学術情報センターは、全国の国公私立大学および共同利用機関を結ぶ、我が国の総合的な学術情報システムの中核たるべき機関として、人文、社会、自然科学の全分野にわたる学術情報の収集、整理および提供を行うとともに、学術情報や情報システムに関する総合的な研究を行うことを目的としている。

これらの事業の一部は、すでに本センターの前身である東京大学情報図書館学研究センターならびに東京大学文献情報センターにおいて行われていたのを引継ぐものであるが、学術情報センターとしてさらに発展させるべき新規事業も多岐にわたることは言うまでもない。

本センターの主体となる事業は学術情報システムの構築であり、全国の大学の参加のもとに、総合的な情報流通システムを完成させるものである。すでに複数の大学附属図書館をオンラインで接続し、その所蔵する図書および雑誌の目録システム・データベースの形成を進めており、今後さらに大規模なデータベースに発展させるべくシステム開発を行っている。最終的には、全国の460に達する大学の図書館の所蔵している約1億4千万冊の図書、および190万種の雑誌の目録所在データベースを形成することを目標としている。このような巨大データベースを構築するには、解決すべき問題点も数多く、その開発は本センターの重要な研究テーマの一つになっている。

一方、大学等の研究者を対象とした情報提供事業として、昭和62年度より、情報検索サービスを開始する。このサービスにおいては、既存の多様なデータベース中から学術研究に必要なものを選択して導入するとともに、学術情報センター独自のデータベースも積極的に構築してゆく。特に、我が国が立ち遅れているといわれる研究分野別データベースの形成も本センターの重要な事業の一つである。これらのデータベースは、高度の専門性のために従来はサービスの実現が困難であったものであり、本センターでは各分野の研究者ならびに学協会の協力を得てこれを実現することを計画している。さらに、情報検索の対象も、二次情報検索に止まらず、全文データベース検索等へと進展しつつあり、図形や特殊記号の表示手法や大量記憶方式の開発など、本センターに課せられた課題は多い。

このような学術情報システムを運営するための情報通信網として、学術情報センターでは独自の学

術情報ネットワークを構築し運用している。学術情報ネットワークは高速デジタル回線とパケット交換機を使用して全国の大学を結ぶ高度情報通信ネットワークであり、学術情報センターと各大学の大型計算機センターとを結ぶ大学間計算機ネットワーク、および学術情報センターと各図書館とを結ぶ図書館ネットワークの通信路として用いられているが、情報検索サービス、電子メール、さらには画像通信などのマルチメディア通信等への利用も計画されている。また、将来は海外に端末を設置するとともに、海外のネットワークとも接続することが考えられている。このような通信網の運用に当たっては、LAN との接続方式の決定やプロトコルの新国際標準への対応など、多くの研究開発業務を遂行していかなければならない。

本センターでは上記の他にも、学術雑誌総合目録の編集や目録端末機の開発など種々の事業を進めているが、以上に述べたいずれの業務においても、研究開発が重要な要素となっている。その内容も、センターシステムそのものの構築に関わるシステム工学的研究項目と、目録情報やデータベースの利用環境の整備あるいは利用動向の調査などの学術情報論的検討を必要とする種々の研究開発事項とがあり、これらが有機的に統合されて学術情報システムとしての成果が得られるものである。このような観点から、本センターでは文献情報センターの発足当初から研究開発に重点を置き、その成果を世に問うために『東京大学文献情報センター紀要』を発刊した。このたび創刊の運びとなった『学術情報センター紀要』はその目的において文献情報センター紀要を継承するものであるが、学術情報システムの発展に伴ってその内容もますます充実したものにして行かなくてはならないであろう。また、学術情報システムは学術情報センター職員のみで構築できるものではなく、上にも述べたように、多くの分野の研究者等の協力のもとで始めて完成する。その意味で本紀要は外部に開かれたものとしており、各方面からの投稿を歓迎するものである。今後、本紀要が大きく育つよう関係各位の御支援をお願い申し上げる次第である。

研 究 論 文

IFLAプレコンファレンス・セミナー「多言語・多文字資料利用のための図書館自動化システム」——提起された問題点

Issues Discussed at the “IFLA Pre-Conference Seminar on Automated Systems to Access for Multilingual and Multiscript Library Materials”

学術情報センター 内藤 衛亮*

はじめに

- 1 プレコンファレンス・セミナーの開催趣旨、開催団体
 1. 1 開催趣旨
 1. 2 開催団体
- 2 論点の要約
 2. 1 DURANCE 報告
 2. 2 MASSIL 報告
- 3 会議始末
 3. 1 問題点と田辺提言
 3. 2 参加者ほか
 3. 3 会議の成果と将来

付録 発表の概要

要 旨

1986年8月に多言語・多文字資料の整理と利用のためのコンピュータ処理に関する国際図書館連盟(IFLA)プレコンファレンス・セミナーが開催された。文字セット(特に漢字文字セット)の国際標準化の必要性がどのようなところにあるか、何が基本的な要求条件かを明らかにして、共通理解を深めようとする目的で開催された。中国、韓国そして日本など漢字を使用し、コンピュータ・アプリケーションの開発を進めている漢字圏の事例、東アジア各国との調整を必要とする、現に漢字処理を進めている当事国である米国(RLG)とカナダ(Utlas)などのほか、アラビア文字、タイ文字、ジャワ文字、ギリシャ文字などのアプリケーション開発とその後の問題点、そして東アジアにおける書誌調整のための体制が必要な点が指摘され、提案された。

*Naito, Eisuke : National Center for Science Information System

Abstract

IFLA Pre-Conference Seminar was held in August 1986 in Tokyo. The aim of the conference was to determine the scope and characteristics of demands for an international standard for character set; to further the common understanding on basic requirements for standardization. Developments and remaining problems are reported and discussed from China, Korea and Japan and from U.S.A(RLG) and Canada(Utlas), Arabic, Thai, Jawi, Greek, etc. Background review were made, and a propose was made to establish a standing committee activities on East Asian bibliographic control.

はじめに

UBC(Universal Bibliographic Control)は、世界的な書誌調整プログラムとして1960年代の末より国際図書館連盟(IFLA: International Federation of Library Associations and Institutions 以下IFLA)が展開してきたもので、国際標準書誌記述(ISBD: International Standard Bibliographic Descriptions)と機械可読方式で目録・書誌データベースの交換に使用するレコード規格であるUNIMARC フォーマットは、UBC プログラムの成果として世界各国の図書館界に大きな影響を与えている。

運動としての UBC は、上記 ISBD と UNIMARC の開発と普及にとどまらず、書誌レコードを国際的に交換流通(これは最近の用語でいえば TBDF:TransBoader Date Flow ということになる)するための国際標準規格として、レコードの内容、レコードの編成(arrangement)そしてレコードの物理的形態などの3つの側面での標準化が必要であることを、1974年に IFLA UBC 事務局の設立に際して初代事務局長である Dorothy ANDERSON 女史が提起している。それによれば書誌的標準化の対象と内容は次のごとくである。

1 レコードの内容

1. 1 標目(heading)として個人名と団体名の標準化ならびに法制資料(legal materials)の書誌調整の必要性と方策、統一標目のリストの作成と維持。
1. 2 記述の標準化として単行書(ISBD-M)、逐次刊行物(ISBD-S)、逐次刊行物の記事内容(ISBD-CP)、非図書資料(ISBD-NBM)などの制定と普及。
1. 3 記述対象の識別方式として単行書(ISBN)、逐次刊行物(ISSN)などの普及。
1. 4 主題(subject)アプローチの確立と普及
1. 5 翻字(transliteration)の標準化
1. 6 国際的目録規則の確立と普及

2 レコードの編成(arrangement)

- 3 書誌レコードの物理的形態として冊子体、カード、機械可読レコードなどがあり、機械可読レコードの国際的流通のためにコミュニケーション・フォーマット(ISO/2709)、内容

識別指示法(content designation)そして SUPERMARC(現在の UNIMARC)などの制定と
普及そのための整合性の必要性と維持の体制

これらの標準化の基礎となる文字セットは、アジア各国(特に中国、韓国そして日本)で次第に制定されつつあることから、データベースの互換性維持の点で注目を集めていた。

発展途上の段階では、まずテストであり試行であるものが、成り行きで規格化されていくことがままあるが、データベースの場合には、その後の流通・利用こそが情報の価値の決め手となることから、データベースに使用される文字セットの標準規格は、TBDF の基礎であり、問題の重要性は書誌情報の関係者には認識されつつあったが、IFLA を介して東アジアに関する具体的かつ公的なはたらきかけはなかったといえる。

また、文字セットの国際的標準化、特にアジア三ヵ国と北米におけるビブリオグラフィック・ユティリティ(書誌サービス機構)との間の各文字セットの互換性維持をはじめ運営上の意見・情報交換も同様に「三すくみ」の状態であった。

もちろん、実際にはシステム開発が各自の利害得失、固有の資源と固有の開発・運営方針のもとに進められていることは、当然のこととして受け入れられている一方で、みすみすデータベースの互換性が失われつつあることについては、利用者である図書館界やさらにエンド・ユーザーである研究者、学生のみならず、開発にあたっている人々の間でも危機感となっていた。

東京において1986年夏に IFLA 第52回総会が開催されたが、それに先立って多言語・多文字資料のための図書館自動化のプレコンファレンス・セミナーが開催された。この会議は、IFLA の多文化社会図書館サービス分科会(IFLA Section on Library Services to Multicultural Populations)における、特に公共図書館において地域住民向けに多様な言語資料を各々の言語で適切な目録を作成する際の問題意識に発端があった。日本側の組織委員会は、IFLA 東京大会組織委員会(永井 道雄委員長)のもとに、IFLA の第4部会である書誌部会に対応する委員会が田辺 広教授(鶴見大学)を代表として結成され、このプレコンファレンスも田辺委員会が対応することになった。1985年初夏のことであった。プレコンファレンス日本組織委員会の一員として会議開催に協力したので、以下に概要を報告する。

1 プレコンファレンス・セミナーの開催趣旨、開催団体

開催趣旨と開催団体は、会議において配布された資料によれば次の通りである。

1. 1 開催趣旨

国際図書館連盟プレコンファレンス・セミナー「多言語・多文字資料利用のための図書館自動化システムー問題と解決」IFLA Pre-Conference Seminar on Automated System for Access to Multilingual and Multiscript Library Materials: Problems and Solution は、IFLA 多文化社会図書館サービス分科会(IFLA Section on Library Services to Multicultural Populations)と IFLA 情報技術分科会(IFLA Section on Information Technology)が共同で開催するものである。

図書館の自動化と国際的情報交換が進行中の今日において、多くの言語と多様な文字の資料とに対応しなければならない図書館にとっての最大の問題は、蔵書に対するオンラインによる目録利用（アクセス）である。

1980年代の初め以来、書誌データベースへの多言語・多文字の入力／出力を可能とすべく世界各地において開発が進められている。日本のコンピュータ技術における先導性と日本語によるオンライン・データベース・アクセスに関連する問題の複雑性の両面から、今年、IFLA 総会が開催される東京において、新たに結成された多文化社会図書館サービス分科会ならびに管理・技術部会の情報技術分科会が共催で、多言語・多文字資料の処理に関するプレコンファレンス・セミナーを開催することは、とりわけ時宜を得たものと言えよう。この会合は世界各地から図書館自動化の専門家を会同することにより、現状の相互認識と今後の協力のためのチャンネルを開拓することをめざしている。会合の成果である意見と技術の両面における相互向上が、IFLA の目的であるユニバーサルな書誌調整(UBC)の推進に資することを期待されている。

1. 2 開催団体

IFLA の二つの分科会が開催母体となった。

多文化社会図書館サービス分科会は、1986年の IFLA 東京総会において正式に承認される予定であり「言語的・民族的少数派のための図書館サービス・ラウンドテーブル」が分科会に格上げされたものである。このラウンドテーブル自体が、わずか5年前に活動し始めた小さなワーキンググループを引き継いだものである。IFLA の「公共・学校図書館」部会におけるこのグループの急速な成長は、多くの国の民族言語図書館サービスの分野において現在進行中の活発な活動ならびに、この分野における国際的連携と協力の必要性を証明するものである。この新しい分科会の目的は、多言語・多文化の業務に従事する図書館員に国際的な討議の場を提供することであり、その中で図書館員は意見を交換し、経験を共有し、そして実務のための連携をとることになる。分科会は、例えば収集の国際協力の促進によってサービスの向上に役立つことが期待されているのである。

情報技術分科会は、IFLA の「管理と技術」部会に属する四つの分科会の一つである。この分科会は、IFLA のコア・プログラムとの協力計画として次のような目的をもっている。すなわち：図書館ならびに情報サービスを支援するために最新技術の利用を促進するネットワーク・システムの開発；コンピュータならびに電気通信技術の分野における作業の企画調整；書誌および所蔵データと、その情報自体が世界的規模で入手できるように結びつけること；機械可読形式による自由な情報流通に関係する国際的努力に関する組織活動への参加。分科会は、図書館への技術の適用／導入ならびに図書館・情報サービスへの光ディスク、マイクロコンピュータ技術、画像、電子出版などの新技術の開発と導入にも関与している。機械システムならびに電気通信に関係ある国際的標準化には、高い優先度を与えている。

2 論点の要約

会議の最後にカナダ国立図書館ネットワーク開発局長の Cynthia DURANCE 女史が発表論文全体を要約し、また IFLA 本会議では開催母体となった多文化社会図書館サービス分科会を代表してロンドン大学図書館の Stephen M. MASSIL 氏がプレコンファレンスの概要を報告した。IFLA 期間中に、要約と広報が進められたことは、それ自体特色とも言えるが、同時に運動課題の合意が具体的な形では得られなかった点も特徴と言える。後者については第3章で田辺提言と合わせて論ずる。本章では、海外からの参加者の論点・レビューの概要を述べる。

2.1 Durance 報告: Cynthia DURANCE は、カナダ国立図書館におけるカナダ・マーク(CAN MARC)の開発すなわち公用語として英語とフランス語というローマ文字言語における自動化システムと二ヶ国語の書誌調整の開発経験を前提にしているが、異なった言語ばかりでなく、異なった文体や構文、そしてさらには文脈的、音声的なもの、書かれた言語といった言語的複雑さの処理にかなり新鮮な違和感があったようである。

2.1.1 意識と影響力の問題: ローマ文字言語を自動的に書誌調整する技術や標準的慣行は、非ローマ文字の現下の問題を十分に処理できないという認識が、西欧社会では不足している。国際標準化機関(ISO)と IFLA は西欧社会の代表によって構成運営されてきたと言ってよい。これらの団体は自発的な組織であること、その結果、精力的に参加した国々が最終的に非常に強い影響力をもってしまったのである。アジアやアラビア等の国々がこの組織に参加し、影響力を与えるような道をさがさなくてはならない。意識と影響力の問題には、英語、文化、政治、距離、そして費用といった障害がからみあっている。今回の IFLA セミナーは正しい方向に向けての一步となるであろう。

2.1.2 標準翻字法の問題: 翻字は、ある多言語データベースや特定の機能では有効に使われている。国際逐次刊行物データシステム(ISDS)のデータベースは、その一例である。これは世界中で発行されている雑誌の識別確認を目的としており、113の言語の最新のデータが整理されている。このような計画では単一の文字でデータベースを作る以外には、目下のところ解決方法はない。

翻字が最良の解決法であるような状況下で、ローマ文字から他の文字や文字への翻字の方法が標準化されていないため困難が生じている。これらの言語の普遍的書誌調整を確立推進するための努力を複雑にするだけである。翻字は、もとの言語を読める図書館員や利用者にとっては混乱をもたらすものである。

2.1.3 東アジアにおける地域的標準化活動の必要性: 現在、アフリカ文字、キリル文字、ギリシャ文字、そしてラテン(ローマ)文字のための ISO 標準文字セットがある。加えて、アラビア字、アルメニア字、グルジア字そしてヘブライ字は、ISO 標準化の過程にある。しかしここには大きな欠陥がある。中国語、日本語、朝鮮語の文字セットを開発するために、多くの仕事が行われてきたが、単一基本文字セット〔もしくは変換メカニズム〕について緊急に国際的なとりきめをする必要がある。歴史的・文化的多様性を維持しながら新しい文字を登録できるような国際的に認められたメカニズムも必要である。その他に ISO 標準規格として日本語と朝鮮語の表音文字セットが望まれている。

る。これらは配列や分類データとして利用されているからである。これらについては国家規格または事実上の国家的標準があり、もしそうであるならば ISO が承認することは、むしろ簡単なことであろう。

2. 1. 4 現地主義: この膨大な複雑な仕事を達成させるためには、調整された努力、時間、それぞれの国ばかりではなく、共通の言語ルーツをもつ国々の言語専門家とコンピュータ専門家との協調、そして ISO の賛同が必要だろう。標準に最も依存するのはその言語が使われている国であり、とりわけその国の人々の要求に適合するものでなければならない。共通の言語ルーツをもつ国々が、文字セットについて同意しあえるのならば、彼らが ISO にうけいられる機会はずっと大きくなるだろう。データ作成やコード化のための共通の方法が見出されないかぎり、国内的、国際的な自動化書誌調整の社会的・経済的利益は、実際に危険にさらされることはなくとも大いにそこなわれるだろう。

2. 1. 5 既存規格との整合性維持: 主にローマ文字の必要性に基づいて開発された事実上の国際的標準は、非ローマ文字言語の必要を十分に考慮に入れて修正され、適合するようにしなくてはならない。例えば ISBD のデータ要素を分類するために使われている長いダッシュは、漢字の「一」とほとんど同じであるとのことである。修正を要する標準のもう一つに UNIMARC がある。表意文字の言語データのための拡張と調整が必要とされている。標準規格または事実上の標準に対しては、どんな修正や付加も責任ある国際的団体によってきちんと認められ、合意されたものとしなければならない。必要な変更や強化を行うための公的かつ永久的なメカニズムが必要である。さもなければ分裂はいよいよ大きくなり、標準は名前だけのものになってしまう。ISBD や UNIMARC のような事実上の標準への修正は、必要でも必ずしも簡単に行われるものではない。既に西欧社会の図書館システムの中で広範に使われているからであり、変更にかかる費用がかかるからである。

2. 1. 6 ハードウェア・ソフトウェア・メーカーへの働きかけ: あまり使われていない言語のコード・システムを十分なものとするためにハードウェアとソフトウェア製造者を説得することはむずかしいことである。そういった標準文字セットを完成させることは、刺激的な仕事である。市場が小さくても承認された標準をつくる要求なら、国際的な機関その他の機関からの財政的な援助はうけやすいだろう。

2. 1. 7 技術の問題: 過去20年間、ハードウェアとソフトウェアの技術は著しく進歩してきた。例えば表意文字印刷用のドットマトリックス・プリンター、パーソナル・コンピュータ、ワークステーション、そしてエキスパート・システムなどの能力や機能の向上は続き、データを処理する上で、実に柔軟性にとんだ方法を提供するであろう。問題は技術なのではなく、データ標準の欠如であり、それが技術開発のさまたげとなっている。

2. 1. 8 IFLA の関与と我々自身の貢献: IFLA の普遍的書誌調整(UBC)のためには、非ローマ文字データの自動化の開発は不可欠である。しかしながら、本当の普遍的意味における書誌調整を行う前に、やるべきことはまだまだある。IFLA はどんな支援をしようだろうか、この大きな仕事に対し、私達はどれだけ IFLA の力となることができるだろうか。

IFLA UBC 事務局は書誌的標準規格についての情報や援助を申し出るだろう。そして他の国々の書誌調整や文字セットの専門家についての情報を提供してくれるだろう。国境をこえたデータフローのための IFLA TBDF 国際事務局としてカナダ国立図書館は、図書館自動化と情報技術の新開発についての情報収集交換機関を設立する予定である。ドイツ図書館の IFLA 国際 MARC プログラム (International MARC Program) は UNIMARC フォーマットを維持しており、MARC フォーマットの情報を提供している。

しかし、IFLA を一つの組織として語る場合、私達自身が IFLA なのであり、組織は私達世界中の図書館員が支持してはじめて存在するということ、そして新しいプログラムを提起し、認識しながら共に仕事をすることによってのみ、私達の要求条件に適合する活動的な組織とすることができることを忘れてはならない。これが私達の任務である。

しめくくりとして本セミナーの第一番目の NICOLESCU 論文から引用する。「私は、一つの国家の繁栄は、地域的または地球的社会の繁栄と同時に起こるべきであるし、また起こるはずであると信じている。このことは情報システムや情報サービスにもあてはまると信じている。」

2. 2 マッシル報告: Stephen MASSIL は、資料の多様性と多文化社会的要素に富んだロンドンで、世界的にも大規模であるロンドン大学の図書館機構の再編制と自動化に1970年代から取り組んできた。貸し出しシステムのオンライン化から目録作成業務の組み込みに至る過程で、会議で語られた標準化を始めとする多くの課題に遭遇する経験を持っている。

2. 2. 1 危機的段階: プレコンファレンスは、最終的な合意事項や近い将来の行動課題を見出だしたわけではない。現時点での問題点を広範囲に抽出したこと、ならびに既に世界各地で個別に着手されている計画のゆえに選択肢がせばめられつつあり、危機的段階にあるとも言えることから、悲観的な見通しにあることが確認された。

2. 2. 2 受益者の利益: 書誌活動にコンピュータを使用するための規格は、受益者である図書館や読者が決定するというよりはコンピュータ・メーカーが決定する立場にあり、その点では受益者の意見がとり入れられるような機構が望まれる。

2. 2. 3 標準化における図書館の地位: 自動化の開発においても、上記2と同様に、標準規格がないために、ハードウェアの選択肢は狭いか、唯一のものを採らざるを得ないのであり、図書館活動には追従の立場しかないのではないか。

2. 2. 4 CJK三ヶ国の自覚: 北米におけるCJK開発では、ローマ文字と非ローマ文字の書誌データを同時に処理する要件に対応しつつあるが、日本、韓国、中国では依然として各自の国語を中心とした個別の開発を進めているのが現状であり、国家規格を進めて地域的規格とすることが緊急の課題であるにもかかわらず、まだその段階に至っていない。

2. 2. 5 重複入力の回避: 北米のCJK開発の進行にともない、例えば、JAPAN MARC レコードとの重複が進行する。各自の緊急性と投資の意義を重視するべきではあるが、ユニバーサルな書誌調整の意義も等閑視できない。UBCの基礎となる規格は、依然としてまだ無いわけで、ローカル・シ

システムとのインターフェイスに重大な影響を与えない統一的な変換機構を前提とした標準化が望まれる。この種のデータベース・インターフェイスの構築と維持は既存の規格に基づくシステムの初期投資と運営よりは高額となることも認識されている。

2. 2. 6 自動化した書誌活動において現状が複雑化している要因: ①拡張文字セット(例、アラビア文字)の ASCII あるいは EBCDIC コード体系への収容方法、文字の多さ、多様な形の文字の存在; ②印字の方向; ③混合文字列; ④文字列の配列順序; ⑤語分割(分かち書き)を必要としない文字列を使用する言語の処理原則; ⑥書法(カリグラフィ)への要求など。

2. 2. 7 緊急の課題: ①標準規格の確立であり、その際の IFLA の UBC 規格化の対応; ②各国の書誌機関は図書館の要件に適切に対応すべきこと; ③北米と日本の間では、相互認識が欠如しており、北米での開発活動、要件、成果を東アジアと関連づける機構が必要である。北米の CJK、台湾の CCCII、そして日本の開発関係者との対話が必要だったのではないか。

全体として MASSIL 自身は悲観的な立場にあるが、IFLA UBC 活動としては何らかの働きかけが必要であると述べている。

3 問題点と田辺提言

前章では、DURANCE と MASSIL の論点を要約したが、会議に参加した立場から、そして IFLA 本会議などで提起された問題点・印象について述べる。

3. 1 現段階での課題: 各システムの現状報告、相互確認は必要であり、IFLA の書誌的活動において、これ程の規模の集合は初めてであろうと IFLA 関係者は述べていた。それだけに米国議会図書館、英国図書館が MARC 開発との関連で、CJK をどのように考えているか、あるいは OCLC の最近の計画がどのようなものか、それらの参加がなかったのは残念であった。発表者全員が参加のための費用を各自に工面せざるを得なかった点で、この種の国際会議の限界を体験したことになる。

3. 2 REACC 委員会: しかし開発の進行段階としては既に後もどりの難しいところに至っており、それだけに、開発のレビュー、レビジョン、そして開発運営上の相互調整のための機構が望まれる。一つの徴候は RLG の働きかけによる REACC 委員会であり、これについては丸山昭二郎、宮澤彰両氏からいずれ報告がなされよう。

3. 3 利用者による評価: 利用者側からの発表・発言が望まれた。今回の発表者の選択過程は、すべて IFLA のもとでなされたのであるが、一方、特に北米のアジア系図書館あるいは我が国の学術情報ネットワーク参加館での各システムの利用上の問題点やネットワーク・システムを利用する際の運営課題などに関する提言が必要ではなかったか。これは各システムの現状に対するレビュー・レビジョンとしても必要な活動ではなかったか。

3. 4 端末: 多様な文字を入力し出力するための端末、ワークステーション、パーソナル・コンピュータは、国際的な情報交換の基礎である。従来型の英数字のみではない情報流通は、既に始まっているのだが、文字や言語の種類に対応してそれぞれの端末を用意することは個人では当分難しい。オ

フィスの面積からも望ましいことではない。高機能で安価そして簡便なワークステーションが望まれている。

3. 5 標準化と情報流通: CJK 書誌情報のために唯一のコード体系とするのか、多様な漢字辞書を介した、いわゆる漢字シソーラスによるコード変換とするか、漢字書誌情報の流通交換方式には大きな期待がある。ネットワーク間の相互乗り入れを円滑にするための要件、ハードウェア・ソフトウェア生産の実現可能性と市場性、「生きている」漢字の正字・異体字関係や意味を反映できる辞書体系の構築維持の費用効果、そしてこれらの事業の必然的な共同的性格など「三すくみ」の条件はそろっているのである。だからこそ次に述べる常置委員会の重さや田辺提言の苦さがいっそう鮮やかとなる。

3. 6 田辺提言: 東アジアにおける漢字書誌開発に関する標準化活動あるいは相互調整のための活動の受け皿が必要である。受け皿となるべき東アジア書誌情報常置委員会(?)に関しては、IFLA会議冒頭の書誌部会において田辺 広教授が提言している。

『東アジアの書誌情報に関する常置委員会の設立と、それに関する機関誌の発行がどうしても必要である。この委員会では目録規則の標準化、漢字文字セット、MARC フォーマット、複写を含む国際相互貸借、翻訳の推進、情報並びに人物の交流などが議せられなければならない。この委員会の困難性は議事の内容にあるのではなく、別のところにある。まず政治的な問題が存在する。委員の構成をどのように考えるのか、そして東アジアの国々をどの範囲に限定するのか、第2の問題は言語である。英語を共通語にしうるかも知れない。しかしわれわれにとって論議を深めるためにはそれでは充分ではない。日本語、中国語、韓国語が、二国語が分かる図書館員の通訳によって公用語となるべきである。第3に予算の問題がある。会議がどこで開かれても参加の費用を考えると頻りに開くことは難しいであろう。そのため IFLA、ASPAC、ICANAS (以前の CISHAAN)、IAOL のようなより大きな会議と共同して、又はその時期に併せて開くことができる。しかしながら、安定した財政的基礎のもとに常置委員会を設けることが必要である。』

田辺教授のこれまでの東アジア図書館員との交流について詳しくは知らないが、波乱がなかったはずはなく、従って提言内容も具体性がありながら、諦念の色がにじんでいる。

3. 7 参加者: 参加者122名、発表者20名と招待者19名そして日本組織委員会17名など合計178名となった。国内参加者の特色は、大部分が関東を中心とした私立大学の図書館員であり、四割近くが女性であり、一方、若干の国立大学図書館の若手がいた。プレコンファレンス日本組織委員会の委員の日ごろの個人的付き合いから意気を感じて友情参加した人々もいた。それだけにこの種の会合としては珍しく意識鮮明な人々もいたなどの点である。

1986年当初以来の「円高」で発表者以外は皆無かと絶望視されていたにもかかわらず、海外からの参加者は20%に近く、在外の日系図書館員の参加が米英国からもあり、国際会議としての色彩は維持できたと思われる。その点では発表者間の相互認識が重要であって、例えば CJK 関係では中国からの発表のないことを MASSIL をはじめ多くの人々が残念がっていた。

3. 8 古谷 夏子: UTLAS 社の古谷夏子夫人の貢献については特記すべきである。UTLAS 社の直接的利益はいざ知らず、一年以上の間、会議の企画、準備活動の運営・調整にあたられた。この間の彼女の公私にわたる時間と費用は、会の成否と直結した要因と言っても過言ではない。二日間の会議の進行調整には日本語・英語そして体力と全人的な努力を要求されていた。約一年間の準備期間があったこと、西ドイツ(ドイツ国立図書館、Cristin BOSSMEYER 夫人)、カナダ(Cynthia CURANCE 女史、Marie ZIELINSKA 夫人)などと日本組織委員会の連絡が古谷夏子夫人を介して円滑に進められた。

4 会議の成果と将来

会議の成果と将来について語ることは難しい。現状の相互確認の段階は、実は「10年も前」のことであったに違いない。今何をなすべきかについては DURANCE、MASSIL そして田辺らがそれぞれ指摘している。

地味であり地道であるだけに、当事者にさえ「一体何のためになるのか」という気持ちがあろう。しかも東アジアという政治的にも経済的にも困難が見え隠れする地域に対して、そして国境を越える以前に国内において組織の壁を越えなければならない点では、前掲田辺論文の提言での苦衷が目に見えるほどである。

付. プレコンファレンス開催支援のための活動を許諾された学術情報センターに感謝の意を表するとともに、本稿において述べた意見は、学術情報センターの公的方針ならびにプレコンファレンス日本組織委員会の意見とは関係ないことを付記する。

付録 発表の概要

第1日(8月21日)の初めに「多文化社会図書館サービス分科会代表挨拶」(Michael FOSTER)と「情報技術分科会代表挨拶並びに司会者紹介」(Christine BOSSMEYER)があり、古谷夏子氏の司会・進行によって会議は進められた。

第1の論文は、この会議の二つの基調講演の一つで米国ニューヨーク州立大学の Suzine HARNICOLESCU 教授の「アジア諸語資料のローマ字化・翻字データベース:歴史、問題、展望」(*Romanized and Transliterated Databases of Asian Language Materials: History, Problems, and Prospects*)である。もっぱら政策的、運営的側面について論じたものであった。NICOLESCU 女史の言語能力を駆使した講演に聴衆は魅了されていた。その論点は次のごとくである。

(1) 垂直的交流から水平的交流へ: アジア各国では西欧で成功したシステムの導入によって個々の書誌システムを開発しつつあるが、これはアジア諸国すなわち中国、日本ならびに韓国(CJK)と西欧との間の書誌的な垂直的交流の結果である。しかし地理的な範囲と「書きもの」において共通する文

字の使用とによって歴史的、文化的な紐帯があるという理由から、アジア諸国間の平等かつ水平的な関係が、極めて強く望まれている。

(2) **共通文字セットの必要性:** アジア資料の増加と文化的な紐帯は、アジアの三ヶ国が標準アジア文字セット(a standard“Asian”character set)ならびに地域的ネットワークの確立を緊急課題としている。

(3) **西欧における書誌的標準規格の欠陥:** ISBD に基づいた UNIMARC フォーマットならびに英米目録規則第2版(AACR 2)によって目録業務を自動化する活動は、すでにアジア諸国において受け入れられているが、これらは西欧の書誌サービスむけに開発されたものである。北米における CJK 諸語のための自動化目録作成システムの開発では、アジア各国はそれほど本格的に相談を受けたわけではない。例えば MARC フォーマットの目録カードに表れる句読記号は、アジアにおける要件としては過剰であり不適當であり、また魅力的ではない。標題紙指向の目録作業は、アジアにおける奥付の伝統と馴染まないものなのである。UNIMARC フォーマットは、アジア各国の目録作業の短所である主題アクセスには十分の手当がない。

(4) **アジアからの貢献の欠如:** 現在、北米で稼働中のアジア資料用の自動化目録作成システム—RLIN の“Sinoterm”、OCLC の“Asiagraphics”、Utlas の“Japan CATSS”—ではデータ入力の問題として、キーボード上での文字の複雑性、複雑な構造(シンタックス)と形(モルフォロジー)に起因するアジア諸文字のローマ字化の困難性などがある。システム開発の際にアジアの図書館員自体も十分に相談にあずかってはいない。

(5) **文字セット標準化における困難性:** アジア三ヶ国の漢字の適用における構文ならびに形態上の類似と差異とは、大きな挑戦である。例えば、三ヶ国を合せて、各国のローマ字化表を組み合わせれば、日本人著者の名前は少なくとも12通りにもローマ字化できる。このような問題は、書誌的な統一要件を確立する上で障壁となっている。要するに、アジア人は、ローマ字化を西欧の利用者のための手段、経済的には非現実、アジア人の利用者には概念的には不適當と考えている。

(6) **結論と勧告:** アジア諸語と西欧諸語の重大な差異を理解すべきである; アジア各国はまず各自の利用者に提供するための情報システムを開発することのメリットを認識すべきである; ローマ字化は西欧人の利用者向けオプションと考え維持すべきである; アジア各国に対して各自の利用者向けのアジア資料に適した書誌調整を開発することができるように、十分に具備された指針と支援を提供すべきである; 東洋ならびに西洋の双方から等しい数の図書館員と機械化担当者を招集して、ユニバーサルな書誌的規格の開発を促進すべきである; そして東西の意見交換の機会を増すべきである; 最後に IFLA は開発途上国からの参加に対して財政援助を提供し、またこの提案によって情報システムの世界の中での東洋各国の位置を認識すべきである。

第2の論文は米国スタンフォードにある研究図書館グループ(RLG: Research Libraries Group)の Alan M.TUCKER 氏の「非ローマ文字と多文字の書誌データベース: 設計ならびにインプリメン

テーションにおける基本的問題」(*Non-Roman and Multi-Script Bibliographic Databases: Basic Issues in Design and Implementation*)で、これは第二の基調講演として技術的側面について論じたものである。

データベース中に、二種類以上の非ローマ文字を結合させて、事柄を複雑にする理由は、第一にローマ字化の方法の多様性、既存の方法における例外、原資料の表記を翻字した書誌記述では、必要な情報がすべてあるわけではないなどの理由で混合が必要であり、第二に、二種類以上の非ローマ文字を含むデータベースが、一つの非ローマ字のための一つのデータベースより、その構築、維持あるいは利用が、本質的に複雑になるというわけではないからである。

(1) **文字セットの種類**の明示: MARC フォーマットの構成原理としてレコード、フィールドおよびデータのそれぞれのレベルにおいて、採用している文字セットの種類を指示するべきである。これによって個々のレコードは、文字セットの自己規定能力を持つことになる。この原理は、レコード作成システムでは基本的なことであり、他の書誌システムとレコードを交換する場合にはなおさら必要である。レコード自身で文字セットが特定化されるので、システムに対して処理不可能なレコードの除外や、利用者の端末機の機能に応じたデータの選択的な表示が可能となる。

(2) **文字セットの標準化**: 機械可読書誌データを交換する場合の、もう一つの原理として非ローマ文字文字セットの標準化が重要となる。書誌記述および書誌調整のための非ローマ文字文字セットの開発における重要な課題に、歴史的な正確さや、中国語のように年代や地域によって文字の形の異なる文字そのものを処理できる能力、あるいはアラビア語、デーバナーガリー(インド文字)のように正字法のある形式を処理できる能力の実現がある。文字コードの設計では、ソート(配列)の能力が重大な課題となる。

(3) **アクセスにおける利点**: 非ローマ文字による探索方法は、データの自己規定性原理が探索方法あるいは索引付けに適用されていけば、ローマ字形式での方法よりも多くの利点があり、そしてそれ程複雑ではない。

(4) **入出力装置の普及**: パーソナル・コンピュータの出現によってアルファベット文字あるいは中国語、日本語、韓国語のように連字活字印刷するよう非ローマ文字(ロググラフ)の入力は、簡単になってきた。端末機上に表示される文字は、必ずしも期待どおりとは言えないが、改善はめざましく、そのまま出版できる品質にまで到達している。

第3の論文は、国立国会図書館の丸山 昭二郎氏の「日本における書誌調整と図書館の機械化」(*Bibliographic Control and Library Automation in Japan*)であり、開催国の現状を要約するために内容豊富なものであった。

(1) **日本における書誌コントロール**: 戦前の状況、1948年の国立国会図書館(NDL)の設立、同年から開始された全国書誌の刊行、全国総合目録編纂事業、CIP が日本で実現されない理由、印刷カード・サービス、目録カード添付サービス、出版者 MARC テープやオンライン出版情報サービスなど

の概況が述べられた。

(2) **目録規則の現状:** 日本目録規則(NCR)の歴史、基本記入制の導入と現行の予備版における記述ユニット制、特に日本語標目における片かな(もしくはローマ字)と漢字の二重性、1955年以来提唱されてきた記述独立方式の経緯、そして和図書の出版慣行が洋書と相違していること、奥付の意義とそれに対応した日本独特の目録慣行など目録規則の基礎概念の特色が紹介された。目録規則としては国際的な目録法原則(標目に関するいわゆるパリ原則)や、NCRはISBDに整合性を追及しており、NCR新版予備版では複数記入制の記述ユニット記入を前提として「単一記入制目録のための標目選定表」として収録されている。

(3) **国立国会図書館の機械化:** 1970年代初頭からのNDLは、漢字のJIS規格制定前に決定されたNDL漢字コード表(NDL-70コード)を採用し、JAPAN/MARCで使用しているJIS規格「JIS C6226-1978」との関連や1986年までに追加された700字以上の外字の適用・運営上の問題が例示された。同様の事例としてUNIMARCフォーマットとJAPAN/MARCフォーマットの比較と漢字記述データと片かなやローマ字形の標目データのリンク用タグづけなど実際の解決方法を示した。JAPAN/MARCサービスについて、テープの収録範囲やレコード数、1986年現在のテープ頒布サービス予約者の内訳を示すとともに、「日本の図書館はコストは高いが漢字処理システムの導入を選ぶ傾向」について指摘した。

(4) **日本における標準化とネットワーキング:** ISBNやISSN付与率、NDC(日本十進分類法)が全館種約4万の図書館の96%で使用されており、テスト的運営段階にあるJAPAN/MARCオンライン・ネットワーク、Utlas International Canada社のプロトタイプ、学術情報センターの動きなど、日本の図書館界はオンライン総合目録、書誌情報の共有などによって新しい時代をむかえつつある。この先に、漢字圏の共通フォーマットとしてUNIMARC漢字版、共通漢字コード表、共通目録法原則の研究、特に分類記号、件名標目などの主題関連データの互換性確立のための調査の必要性がある。

第4の論文は、台湾のChing-Chun HSIEH氏ほかによる「多言語環境における汎用コーディング・システム」(*An Universal Coding System for Multi-Lingual Environment*)で、RLG/RLINやOCLCのCJK開発において基礎となった3バイト・コーディング・システムの現状報告である。

(1) **汎用コーディング・システム:** システムの構造は1980年に開発されたCCCII(Chinese Character Code for Information Interchange)を基本とし、初版は1983年に完成した。CCCIIの第三版は、ISO-646やISO-2202に基づいたすべての3バイト、2バイト、1バイトのコーディング・システムに対応でき、文字セットは台湾、日本、韓国、中国のそれぞれの文字を包含している。第3版は標準的な42,000の漢字と11,000以上の変化した形の漢字を収めており、JIS C6226、GB 2312も入っている。文字セット間での交換機構がシステムにある。

(2) **CCCIIとCCDB:** CCCIIの3バイト方式による文字属性を結合するデータベースCCDB(Chinese Character Data Base)を介して漢字属性をアクセスできる。例えばCCDBの索引機能によ

って中国の漢字から対応する日本語の漢字を呼び出せる。

CCCII は ISO-646 および ISO-2022 に基づく三つの 7 ビットバイトであり、文字のための領域としては 94×93×94 すなわち合計 821,748 の文字とデータ処理のための制御文字を収容できる。先頭(第 1)バイト(B1)はセクション中のポジション(位置)、第 2 バイト(B2)は 1 つのプレーンの中のセクション、第 3 バイト(B3)はプレーン(Plane-面)を表わしている。このように CCCII と CCDB は、7 万以上の漢字に最適であるだけでなく、中国語や準中国語ともいうべき日本語、韓国語など東洋言語に適している。

CCDB には次の三つの機能がある。①漢字の構成、読みあるいは発音記号、漢字のドット・マトリックス構成、他のコーディング・システムのコードなどの提供； ②漢字のフォントと読み方の間の交換テーブルを制御し、検索する機構； ③複数バイトの種々のコーディング・システム間の交換機能。例えば 3 バイトの CCCII から 2 バイトの JISC-6226 へ、あるいはその逆など。

このシステムには CCCII 機構があり、①核(Kernel)：CCDB は実はこの核である； ②変換ユニット：CORE(CODE RECEIVER—コード入力) と COTRA(CODE TRANSMITTER—コード変換出力) からなり、CCDB とデータ通信装置との間のインターフェイス； ③コントローラ(制御部)：CCCII 機構の中心となるコントロール・プログラムで構成する。CCCII 機構は端末やパーソナル・コンピュータやホストコンピュータに実装され、種々のアプリケーションで目的に応じた機能を提供することになる。コントローラはアプリケーションに依存する。

(3) **単一の文字セット**：CCCII の 3 バイト構造は、漢字の正規の形、簡略形、変化形の関係を維持し、日本語の仮名や韓国語のハングル、ほとんどすべてのアルファベティックな文字をそのコーディング・システムの下で処理できる。多言語を処理する場合でも複数の文字セットを使用する必要がなく、エスケープ・シーケンスも不要であり、単一のコーディング・システムで済む。

4 **製品開発の現状**：CCCII 機構は、台湾では数社が開発中なので、1986 年末には市場で供給されるものと期待されている。1986 年末あるいは 1987 年初めには 7 万字を越える文字数を収容する CCCII の第 4 版を出す予定である。

第 5 の論文は、韓国の Kyu-Seob HYEON 教授による「韓国における標準文字コード・システムの開発と韓国 MARC への適用」(*On Establishing the Standard Character Code System in Korea and Its Application on Korean MARC Database*)で、韓国 MARC(KOR MARC)開発の直接の担当者からの提言であった。ヒョン教授も流麗な日本語による発表が可能な方であった。

(1) **韓国におけるコンピュータ・文字コード・システムの概要**：韓国語で書く文章は普通ハングル(韓国文字)と漢字の組み合わせであるが、商業分野、官庁刊行物そして文学作品では漢字は使われていない。書きコトバとしてはハングルのみを使用することが一般的であるため、韓国においてコンピュータ・システムに適用できる漢字コードの開発が遅れたのである。現在の主流は、ハングルの 33 の表音文字が ASCII の各バイトに対応している 1 バイト・コード・システムである。

(2) **韓国 MARC の開発とその文字コード・システム:** KOR MARC システムは1979年に韓国国立中央図書館において創設された。共同体制の MARC 開発委員会の設置、MARC データベースを利用した印刷カードの配布などへ発展しつつある。コード・システムは、約5,000文字の漢字を適用できるように拡張されたが、依然として1バイト・コードである。KOR MARC システムは、プロジェクトそれ自体としては成功したが、文字コード・システムの拡張と標準化の課題を抱えている。

(3) **文字コードの標準化の必要性:** KAIST ならびに科学技術省は漢字コード体系の確立に尽力してきた。1982年に Korea Industrial Advancement Administration Agency は、KSC 5619 “Code of the Korean Graphic Character Set for Information Interchange”を公布した。しかし KSC コードの文字数は書誌的な用途のためには小さかったため、KOR MARC システムでは KSC コードを使用できなかった。現実的な解決策は1バイトの漢字コード・システムを5,000字から7,000字に拡張することである。

(4) **検索システムと文字コード:** 韓国における検索システムでは、実際にはハングル文字とローマ文字の2種類で探索できる。最近に至り、多くの図書館員はハングル文字、ローマ文字と同じく漢字と日本の文字の使用を強く求めるようになった。これらの要望に応えるために漢字のための辞書システムの開発が必要である。これを補うもう一つの方法は主題アクセス・システムと典拠コントロール・ファイル・システムである。

(5) **結論:** 漢字の自動化辞書システムならびに個人名の典拠コントロール・システムは多言語・多文字の図書館資料の利用のためには必須である。また多文字を使用する我々相互間の避け難い差異を橋わたしする国際的な組織の設立も必須である。

第6の論文は、マレーシアの Edward LIM Huck Tee 氏による「東南アジア諸文字のコンピュータ処理—特にジャワ語ならびにタイ語について」(*Computer Processing of Southeast Asian Scripts, with Particular Reference to Jawi and Thai*)で、我が国でもタイ語のコンピュータ処理例はあるものの、当事国からの発表として関心を呼んだものである。

(1) **序:** 東南アジア言語は方言を含めると数百にも及び、非ローマ字を使用する言語・方言があるが、キリスト教伝道会によってローマ字化がすすめられ、公用文字となってきた。東南アジア言語の文字処理では非ローマ文字の存在が自動化の障害となっている。

(2) **文字生成の原理:** 非ローマ文字をサポートするためには、特別なキーボードの他に、キャラクター・ジェネレータが必要となる。プリンタを含めたハードウェアの改良の他に、適切なソフトウェアも必要となる。非ローマ字を含む文字処理のためには、○7ないし8ビットを使用する言語のマッピングの問題。その言語が128ないし256のコードに簡単にマッピングされるかどうかは課題となる；
○言語固有の問題； ○ローマ文字と非ローマ文字のレコードを同時に処理する際の問題などを解決しなければならない。

(2. 1) **ジャワ文字:** ジャワ文字は、アラビア文字を適用し、いくつかの追加文字があり、35字の

アルファベットと lam-alif と hamzah からなる。文字の語中の位置関係によってアルファベットの形がわずかに異なり、結果的に百以上の異なる文字種がある。目録や書誌レコードでは、ローマ字とジャワ文字セットを併用する必要があるので、ASCII や EBCDIC コードを使って文字をマッピングする必要がある。しかしジャワ文字セットを符合化する基準はまだない。コンピュータ処理を複雑にする要因としては、○文字は右から左へ、数字は左から右へと書かれる。ジャワの文字セットをローマ字と同時に処理し、その際に読み方や書き方が左から右であることが処理を複雑にしている。これまで二つの解決法として、画面分割ないしはウインドーと、カーソルの移動を一方向のみにし、ローマ字からジャワ文字へ移る際に切り換えられるようにする方法が試みられ、後者が優れている。○ジャワ文字は語の中での位置関係によって形が異なるので、システムに自動的に正しい形を選択させなければならない。そのために文脈分析できる処理が必要である。○ソーティングの問題—辞書体配列の基準がない； ○特別な入出力装置の採用； ○異なる文字セットを同定するソフトウェアの問題などの要因がある。

(2. 2) **タイ文字:** タイの書法は、44の子音、32の基本母音、声調符合、特殊記号と句読点から成り、マッピング上の問題はない。事実、タイ IBM には EBCDIC による72のコードがある。タイ文字セットは、lower case にローマ字と特殊記号、upper case でローマ字とタイ文字を併用する。これは IBM を使用し、そのハードやソフトの設計をあまり変更させないようにしたための措置であった。タイ文字のコンピュータ処理を複雑にする要因として、○いくつかの母音の子音の前に書かれ、子音の後の場合でもその上や下に書かれる； ○声調符合が、すべて上付き文字である； ○語間のスペースがほとんどない； ○配列上の問題などがあり、いくつかを SEAPRINT Pilot Project で解決した。

(3) **ソフトウェアの利用:** ①ワードプロセッサ、②DBMS(MINISIS: ミナイシス)、③MALMARC(マレーシア MARC)など3つのタイプのソフトが利用ないし試行されている。

以上の六つの論文が第一日に発表され、夕刻に Pre-Conference Reception が、日本大学会館8階の会場で開催され、IFLA の会長 Hans PETER-GEH 博士夫妻、WJINSTROOM 事務局長が国立国会図書館の高橋 徳太郎副館長に案内されて来場し、IFLA 本部からの祝辞を寄せるという一幕もあった。

第2日(8月22日)には八論文が発表された。初日と同規模の出席者があったことと、質疑が詳細化しつつあったことが特色と言えよう。

第7の論文は、米国ウィスコンシン州立大学の Mohammed M. AMAN 氏の「コンピュータ化情報システムにおけるアラビア文字の使用」(*Usa of Arabic Script in Computerized Information System*)で、AMAN 博士は、非ローマ文字資料の目録作成について一方の識者として知られている。

(1) **アラビア文字:** 右から左へと書かれるアラビア文字は、約50ヶ国10億近くの人々が使用し、イラン、北インド、パキスタン(一部)の言語である Farsi でも使用される。アラビア語とペルシャ語の書法を修正したものがウルドゥ文字である。Farsi と同様、ウルドゥ語のアルファベットには、37

字(アラビア語より9字多い)と多くの上付き、下付き文字がある。アラビア文字は、ほとんど筆記体の字で連結され、最も一般的な書法が、Naskh' と Naskh'liq である。アルファベットに加え最低5つの母音がある。書法における母音の使用は任意であるが、文脈上から正確な語を類推できない場合に使用される。母音を無視することはできない。

(2) 自動化された情報処理におけるアラビア文字の三つの問題:

(2. 1) 文字表現: 1976年以来、アラビア文字の統一コードを作るための作業が進められており、ASMO(Arab Standards and Metrology Organization)の第8委員会が準備した統一コード NO. 449-1982が制定された。コード表現は ISO-646を基礎としているが、ほかに制御文字やグラフィック文字セットを規定している。グラフィック文字セットによってテキストが母音をどう取り込んでいるかの判定が可能になる。ASMOの基準には拡張性がある。アラビア文字基準は、128の文字種からなり、ほとんどは変更不可であるが、国別の事情もあり、柔軟性を持たせる規程がつくられている。

(2. 2) キーボードの問題: ローマ字とアラビア文字双方を扱うターミナルが不可欠であるが、現状ではアラビア文字用のためにローマ字(英語)の小文字(lower case)を使えなくしている。

(2. 3) アラビア語の連結文字の問題: ワードプロセッサを改良をして、キーボード上に文字をすべて配置する必要をなくした。大手のコンピュータ・メーカーは、今日までアラビア語処理に参入しておらず、若干のメーカーが、アラビア文字データを取り込めるようにオペレーション・システムを改良しようとしている。

(3) アラビア文字資料へのコンピュータ利用: 主にアラビア文字資料を所蔵している図書館での自動化計画の場合も、ローマ文字による従来と同様なアプリケーションである。しかし標準規格はまだ普及していない。例えば、MARCフォーマットはなく、ASMO基準は多くのシステムには適用されていない。アラビア人の名称の検索を容易にするような計画もない。クウェートでは、NSTICがIBMのStairsのアラビア語版を使用してアラビア語のコマンドとブール演算によって検索できるよう開発しつつある。カナダのIDRCが開発したパッケージであるMINISISのアラビア語版が、1982年にALDOCによって完成された。ほかにVTLSやDOBIS/LIBLISのようなソフトウェアのアラビア語化の研究が進行中である。

第8の論文は、オーストラリアのメルボルンにあるTechnilib(1975年にヴィクトリア州図書館評議会からの助成金を得て、公共図書館グループにより共同で組織された集中目録作成機関)のEve STOCKER夫人による「オーストラリア・ビクトリアにおけるギリシャ語の自動化目録作成システムの開発」(*Development of Greek Automated Cataloguing System, Victoria, Australia*)である。ギリシャ本国から国立図書館長NICOLOPOULOS氏がプレコンファレンスに参加していた点は、この会議の将来へのインパクトを示唆しているといえよう。

(1) 発端と意義: ヴィクトリア州メルボルンのキャリングブッシュとハイデルベルク地域公共図書館は、ギリシャ語資料をギリシャ語で目録作成する自動化システムの構築を1977年から1978年にかけて

て検討した。ギリシャ語資料の自動化目録作成システムは、英語その他のローマ字を使用する言語資料に対する図書館自動化システムの論理的展開である。1975年以来、この二つの公共図書館は、ギリシャ語その他のコミュニティ言語による蔵書を構築してきた。「コミュニティ言語」という用語は、非英語国から来た移民の母国語を指す。世界初のギリシャ語目録作成システムの開発は、ヴィクトリア州にあるデータ処理サービス会社(Librumatic System 社)が引き受けた。

(2) **システム開発:** 大文字・小文字双方のギリシャ・アルファベットの文字セットが設計され、特別に改良した Beehive B550 端末に組み込まれ、ギリシャ文字はこの端末の標準ローマ字キーボード上に刻印された。画面上には英語のアルファベット、すべての ALA 文字セット、スーパーシフトキーの打鍵によってギリシャ文字が表示される。

(3) **標準規格と製品:** 標準 MARC フォーマットを適用し、ギリシャ語のデータとそれに対応する英語のデータ(LC 件名標目)もしくは翻字形のデータ(著者、標題)という形で並列フィールドを収容している。翻字形のデータは貸出システム用に必要だったのである。Libramatic System 社は、ギリシャ語資料のマイクロフィッシュ目録を3つの配列すなわち著者/標題、英語の件名標目(LCSH)、およびギリシャ語の件名標目(翻訳 LCSH)で作成した。書架リスト、所在リスト、校正作業リスト(Validation Report) (以上マイクロフィッシュ)と差し込み票、背ラベルが作成された。最初のマイクロフィッシュ目録は1981年に作成された。

(4) **現状:** 1983年までには、ギリシャ語資料目録作成ネットワークは五つの公共図書館で構成されるまでになり、1984年にギリシャ語資料目録作成ネットワークは、その目録業務を Technilib の目録サービス部門に吸収するよう提案した。Technilibは、それまで(非ローマ字に対しては翻字を用いて)多数の言語の図書を目録作成しており、ギリシャ語資料の目録業務もその他の非英語資料目録作成サービスとの間に整合性のあるものとみなされた。Technilib の目録作成およびその他のデータ処理サービスは、当時 Libramatic System 社によって提供されていた。

1985年に Technilib は、IBM4361汎用機の上で DOBIS/LIBLIS によるインハウスのオンライン目録業務システムを導入し、Libramatic System 社のバッチ処理システムの利用を段階的に取り止めていった。1986年に Technilib では、ギリシャ語資料の目録業務を現行のシステムに吸収するという方針を決定した。それにはギリシャ文字を表示できるように Ericsson Alfaskop 端末を改良し、プログラムの修正を行うことが必要になる。

第9番目に、臨時にオーストラリア国立大学図書館アジア研究部門の Y.S.Chang 氏の発言が求められた。

彼は1983年の東京における ICTP(International Conference on Text Processing with Large Character Set—字種の多い言語の文書処理に関する国際会議)を初めとして多言語・多文字処理の多くの会議に参加した識者であり、それらの会合で語られた非ローマ文字の処理が、いずれも国内レベルまたはネットワーク内のレベルでの検討に終始している点を鋭く指摘した。

オーストラリア国立大学図書館のアジア研究部門では、CJK のほかにサンスクリット、タミール、インドネシア、タイ、ヴェトナム、ビルマ、アラビアなどの文字の資料を扱っている。このため1982年に書誌標準化、漢字書誌調整の会議を開催し、米国、英国、日本、香港などのほかに、政治的困難があったにもかかわらず二つの中国からの参加があった。この会議について成功だと皆は言い合ったが、理論的・概念的な議論に留まり、アプリケーションについて述べる場合、各論展開しかなかったのである。

1986年はじめに、同大学図書館は、RLG に三ヶ月のテスト接続をして中国語、日本語資料を750タイトル、南アジア資料500タイトル、東南アジア資料を500タイトル、3つのカテゴリー(1979、1980-82、1983以降)として入力テストした。しかし、ヒット率は極めて低く平均45%、南アジア32%、中国語42%、日本語42%であった。この場合は、ハードウェアが問題となった。次に統合システムの導入を検討したが、統合システムのゆえに維持の点で難しいことは、わかっていたが、URICA システムがオーストラリア、香港、台湾に輸出されている実績から期待していた。しかし台湾からURICA にCJK を組み込む提案があったと聞いたが、実際には技術的に可能でも経済的には不可能であり、台湾では西欧語資料のみを扱うとのことであった。そこで独自の統合システムを考えざるを得ない結論に至ったが、コンピュータ・エキスパート、図書館のエキスパートそして何よりも第一級のシステム・エンジニアが必要であるのに、「ここが一番難しいことは皆さんも良く知っていることであろう。多文字・多言語資料のためのシステムを手に入れるまでに、もう10回もこんな会議に出なければならぬのか。」氏の発言の切実さは聴衆の共感を呼んでいたが、論調は明るく会場を和ませていた。

10番目に IDRC(International Development Research Centre)のシンガポール事務所から参加した Michael Sherwood 氏の発言が求められた。彼は Ottawa の IDRC 本部の Richard Lee 氏の論文を口頭で要約した。

IDRC は政府機関(crown corporation)であり、発展途上国における開発活動の推進にあたっている事業体である。MINISIS は、カナダ IDRC がドキュメンテーション活動における書誌情報処理のために開発したソフトウェアである。目録作成よりは情報の蓄積と検索を目的としている。当初から発展途上国向けに非ローマ文字を意識して設計開発された。アラビア文字、ギリシャ文字、タイ文字、漢字/中国文字、韓国文字などが処理可能である。ただし、すべてが可能となっているかについては未詳である。非ローマ文字の処理における MINISIS の利点は、中心となるソフトウェアの修正を要しない点にあり、システム・メッセージとシステムへのコマンドは、中心ソフトウェアの外にあるメッセージ・ファイルを介して会話されるので、コンピュータと利用者との会話は、利用者自身の文字セットや言語で可能となる。

文字列の内部処理は、文字セット定義テーブルを介しているなので、配列もこのテーブルに従って可能となる。中国語の場合、ピンイン、画順、部首順のいずれをも指定できる。ギリシャ語、タイ語の

場合にも特定の文字セットを指定して、ソートキーを生成させることができる。またローマ字の次にギリシャ文字そしてタイ文字といった文字セット間の順を指摘できる。文字セット定義テーブルでは、1文字のサイズ(1バイト、2バイト、3バイトなど)を指示し、印字方向も指定できるので、アラビア語・ジャワ語の印字に対応できる。句読点の識別、大文字・小文字変換ができる。1データベースにおいて16種類の文字セットをレコード中、フィールド中に収容できる。文字セット識別コード(エスケープ・シーケンス)によってこれが可能となる。MINISISは端末や文字コードシステムとは独立であり、I/Oハンドラー(サブルーチン)が、MINISISソフトウェアと端末との間にあってコード変換を管理する。端末は何を選んでもよいことになる。MINISISには多言語シソーラス機能があり、探索キーとは別の言語の結果が表示できる。しかしこれはシソーラス機能であって件名目録作成における件名標目機能ではない。利用者規定によるキー生成ルーチンによって多言語、多文字間の変換が可能となっている。主な積み残し課題は、ワードプロセッシング機能であり、話や概念の識別能力が欠如しているが、これは例えばアラビア語、タイ語、日本語、韓国語、中国語などの言語ではスペース、区切り記号、分かち書きがないためであり、識別が不可能だからである。スペースを区切りとみなすことはできるが、利用者の入力には便利とは言えない。文脈分析の可能性を追求したが、必ずしも明確な成果とはなっていない。

Sherwood氏の要約は、MINISISの機能を簡潔に紹介し、多言語・多文字への応用可能性を示すとともに、カナダの国際協力の広がりをも示していた。

第11番目の論文は、当センターの宮澤 彰助教授による「東京大学文献情報センターにおける文字セットとその運用」(*Character Set and its Control in Online Cataloguing System: Application at the Center for Bibliographic Information*)で、ここでは1986年4月から学術情報センターとなったものの、混乱を避けるため、旧称で議論が展開された。

(1) **はじめに:** 東京大学文献情報センターは、日本の大学図書館に対してオンライン共同目録作成システム(CBIシステム)を提供している機関である。この目録作成システムには、基本的使命として大学図書館の受け入れる資料のすべての言語データを収容するデータベース・システムを構築することが求められていた。CBIシステムは、書誌ユティリティ・システムとして①データベース管理システムとコンピュータ・ネットワーク・システムに基づいた新しい設計; ②分散処理志向; ③総合目録; ④総合化された典拠コントロール・システム; ⑤インテグリティ・チェックのためのソフトウェア・システム; ⑥画面型インタフェース; ⑦漢字、ALA文字セットのサポートなど特長がある。アルファベットと漢字に加えALA文字を入出力するために目録端末が開発された。

(2) **文字セット:** 1つのデータベース・システムで使われる文字コード・テーブルは「文字セット」を形成する。多文字種を処理する際の最初の問題は、そのデータベースに対して十分なセットの作成であり、次に、その文字セットを扱うるデバイスを手入手することである。CBIシステムはN-1XNVT(拡張仮想端末)のコード・テーブルを「文字セット」としている。CBIの文字セットは、日本

語、英語、ドイツ語、フランス語、その他のローマ字を使用する言語、さらにロシア語、ギリシャ語を収容しているが、中国語、韓国・朝鮮語その他のアジア諸語はカバーしていない。

この文字セットでは、1バイト2バイトの文字の混合文字列を取り扱う必要がある。このため「混合文字」(mixed character)型の概念を導入し、この型の文字列を扱うルーチン群がシステムに用意された。

(3) **文字セット管理システム:** 漢字のように5万を超える文字種を持った文字セットでは、そのすべてを数えあげるとはほとんど不可能である。「開いた」文字セットシステムとは、固定された数の字種を持つものではなく、必要に応じて字種を増やすことのできるものである。「開いた」文字セットを管理する文字セット管理システムは、①システム外字の登録機能; ②入力外字の入力機能; ③出力外字の出力機能など三つの機能からなる。

CBI システムでは、システム外字の登録機能は集中化され、文献情報センターにおいて実行される。外字は特別なフォーマットで入力され、システムはこれを見て外字出現を報告する。文献情報センター内の一部所がこれを集め、必要に応じて文字の登録を行なう。人力の困難を避けるため、ユーザは外字を16進コード、いくつかの漢和辞典の検字番号を用いて入力することも可能である。

(4) **将来の課題:** 現在の CBI システムの文字セットは、中国語、韓国・朝鮮語やアジア諸言語をカバーしていない。この点は近い将来大きな問題となるであろう。また現在の目録末端では不可能なキリル文字の直接入力も望まれている。さらに多言語環境でのキーワード切出しなど、言語処理機能の研究開発を進めて行く必要がある。

第12番目に John W. HAEGER 氏の「RLIN システムにおける CJK (中国語/日本語/朝鮮語)の拡張」(*The CJK Enhancements to the RLLN System: a Review of Basic Issues*)の発表があり、RLG/CJK の先導性の故に、この会議におけるさまざまなコメントが RLG を意識したものとも思われた点でも聴衆の関心を呼んでいた。

(1) **RLG/CJK プロジェクト:** 米国の研究図書館グループ(RLG)の図書館ネットワークシステム(RLIN:Research Libraries Network)における東アジア言語による書誌情報の入力、蓄積、伝送、検索、画面標示、印字出力といった機能を備えるプロジェクトは、1980年4月に開発を始めた。1983年9月に最初の CJK レコードが入力された。1986年4月現在、北米の東アジア資料の約7割を所蔵する22の図書館が参加しており、約16万件の CJK レコードが蓄積され、69台の CJK 端末が接続されて稼働している。地域的には米国、カナダを中心とし、英国図書館とは専用通信回線を通して2台の CJK 端末が接続されている。プロジェクトの最終段階に CJK シソーラスの開発(1986年6月に完成予定)があり、個々の漢字の属性についてオンラインによる参照案内、RLIN システムが保有する漢字セットへの新しい漢字の追加、日本、中国、台湾、韓国の各国の標準コード間のコード変換など実現する。

(2) **基本方針:** CJK プロジェクトの経済上および設計上の基本的な方針は、新しいシステムを作る

のではなく、既存の IR システムを拡張する、そして中国語、日本語、韓国語が、これらの言語とローマ字の間で簡単に切り替えられて入力できるというものであった。このため RLG は漢字の構成要素（偏や旁など）を組み合わせた形で入力する Transtech 社の Sinoterm と呼ぶシステムを採用した。基本課題は、情報の検索であり、単にカード目録の模写ではなく、これを拡張して利用者がローマ字や漢字を自由に選択して検索できることである。

(3) **REACC コード体系:** プロジェクトで最も時間を要したのは漢字コード・システムの開発であった。問題は3種類の複雑な漢字の字体や字形であり、相互に関連するものの別のものであり、悪いことに、これらが安定していない。RLG/CJK は REACC (リアック: Research Libraries Group East Asian Character Code)を凍結して、文字の変更などはシソーラスの変換テーブルで保守していく考えである。REACC は LC によって米国の NISO(National Information Standards Organization)に、さらに ISO にも登録される計画である。REACC は書誌的な CJK データの国際的な情報交換における唯一のコード・システムであるといえる。

(4) **機器:** 現在の端末機器の構成は、1981年に採用を決定したもので、機能的にかなり制限があった。プリンタもディスプレイも16×18以上のドットでの CJK 文字表示ができない。当時はまだ端末機が幼稚で、ミニ・コンピュータも安くはなかったのである。RLG ではもっと安く機能的に秀れた、しかも RLIN システムで他の非ローマ字系言語（例えばキリル、アラビ、ヘブライなどの文字）も扱える端末機に改善する予定である。

(5) **印刷カード:** このシステムでは3×5インチのいわゆる印刷カードの作成をサポートしない方針であった。その理由は、当時印刷カードを作成できるようにプログラムを拡張することが非常に高価であったこと、RLG のスタッフや委員会のメンバーが、RLIN は東アジア図書館に利用者が直接オンラインで目録をアクセスする方式(on-line public access)を実施させるシステムであり、カード目録はなくすという信念をもっていたからである。しかし5年経過した現在、カード目録を凍結した館はわずか1館にすぎない。コロンビア大学をはじめ他の5、6館では RLIN システムからのローマ字形のカードを注文し、これに手で漢字を書き入れている。CJK カードに対する参加館の継続的な要求は無視できない。解決方法は明確ではないが、技術的解決を検討しつつある。

(6) **将来計画:** RLIN はネットワークを基盤とした中国語、日本語、韓国語の図書館資料を処理し、検索し得る現在フル稼働している世界で唯一のシステムである。今後さらに利用者の要望を考慮し、より広範囲な書誌的環境を反映すべくシステム改善を進める。

第13番にカナダの Utlas 社から Jack CAIN 氏による「Utlas における中国語と日本語の文字処理のための開発」(*Utlas Development with Japanese and Chinese Scripts*)と題する発表があった。

(1) **Utlas International Canada :** Utlas は図書館および情報産業を対象とするコンピュータ資源活用のサービス機関で、書誌情報提供機関として1973年よりオンライン・データベース・サービスおよび関連製品を英語ならびにフランス語で提供している。カナダ、アメリカそして日本で、コンソー

シア、政府機関の参加メンバーを含めた465をこえる機関、図書館において2,000以上の利用組織が Utlas の機能を通じて、利用者自身のデータベースの構築・維持を行っている。日本では丸善株式会社が Utlas の代理店である。

多文字処理にアプローチするにあたり、Utlas は、その国の伝統、言語、および文化を理解し、尊重していきたいと考えている。このようなアプローチを試みるのは、その国の要請にできる限り応えられるシステムを作り上げるためである。

(2) **オンライン多文字処理システム開発:** 環境(各国の伝統、言語、および文化)は、その国の必要とする書誌調整の形式に、基本的なレベルで影響を及ぼしている。アジアその他の地域でも、別々の文字、異なる目録の習慣が、環境の非常に大きな部分となっており、中国語と日本語では実用的な目的から二つの別々の文字、別々の目録が一般に用いられている。

(3) **多文字の機械化処理のための設計構想:** 多文字処理システム開発において検討すべき要因には、①費用; ②利用者共同体; ③文字セット; ④通信; ⑤入力方法論; ⑥ユーザーインターフェース; ⑦MARC (機械可読目録) 形式; ⑧アクセス手順などがある。

(4) **Japan CATSS(Catalogue Support Service):** 日本の現在の利用者の要求に適合すべく Utlas は Japan CATSS のひな型を開発した。従来の CATSS(ウェスタン・キャッツ)は目録作成および関連業務のために用いられているローマ字だけのオンライン共同目録作成機能のことである。プロトタイプは、①アジアの伝統に従い Japan CATSS は日本語のためだけの個別システムとする; ②文字セットは JIS C6226による; ③Japan MARC フォーマットに準拠; ④「カナ漢字変換」入力手順; ⑤オンライン会話型—日本語のメッセージおよびプロンプト; ⑥アクセス機能(ISBN、他の番号による数値検索キー; 著者名、書名、叢書名および件名による拾読通覧キー; 論理演算検索; カナ検索方—カナと漢字の表示など)の特徴がある。

(5) **中国語および韓国語の開発計画:** ①各々、個別のシステムとして計画している; ②文字セットおよび MARC フォーマットは、その国の標準に準拠する; ③韓国語の入力はハングルで、中国語は合成方式で行う; ④韓国語の拾読通覧キーはハングル順配列とし、中国語の拾読通覧キーは、発音順と字画順配列の二種類とする。

最後に第14番としてカナダ国立図書館のネットワーク開発局の局長である Cynthia DURA NCE 女史が「次に何を:会議における討議内容の要約」(*What next? Issues arising from Conference Deliberation*)を、この「多言語・多文字資料利用のための図書館自動化システム—問題と解決」と題する IFLA プレコンファランス・セミナーにおいて発表者が提示し、会議において認識された重要な問題を要約した。これについては、2. 1章で述べた。

閉会にあたりカナダ国立図書館の多言語サービス課長 Marie ZIELINSKA 夫人が本セミナーの意義を要約した。

研 究 論 文

TOOL-IR 工学文献データベース(COMPENDEX, Ei ENGINEERING MEETINGS)検索システムの開発

Development of TOOL-IR Information Retrieval System for Bibliographic
Databases in Engineering (COMPENDEX and Ei ENGINEERING
MEETINGS)

学術情報センター 根岸 正光*

東京大学大型計算機センター 小澤 宏**

要 旨

工学文献データベース COMPENDEX および Ei ENGINEERING MEETINGS に対するオンライン情報検索システムを開発した。システムの一般語索引は論文表題、会議名、CODEN の各項目、およびデータベース編集者が付与した索引用項目中のすべての語について、また特定索引は著者名、ISSN、ISBN および会議番号について作成されている。

これらの索引と照合して目的とする文献の集合を作るには、PHRASE、SEARCH、AUTHOR、ISSN、ISBN の各コマンドを使用する。PHRASE コマンドは一般語索引を対象とし、オペランドに指定された語すべてを含む文献を検索する。SEARCH コマンドは汎用の検索コマンドであり、すべての種類の索引を対象として範囲検索など複合操作を行うことができる。また LOOK コマンドを投入すると索引内容が表示され、それを吟味したのち、文献集合を作ることができる。文献集合の内容を表示するには DISPLAY コマンドを投入する。このコマンドは簡略表示から全情報の表示まで5段階の表示レベルをもつ。

本システムは東京大学大型計算機センターTOOL-IR システムの一環をなすものであり、日本の学
界において広く利用されている。

Abstract

An on-line information retrieval system for bibliographic databases in engineering, i.e. COMPENDEX and Ei ENGINEERING MEETINGS, has been

*Negishi, Masamitsu : National Center for Science Information System

**Ozawa, Hiroshi : Computer Centre, University of Tokyo

developed. A common index of the system was made on all the words in the fields of document and conference titles and CODEN, and in the indexing fields given by the database editor; specialized indices were made for author names, ISSN, ISBN and conference numbers.

To make a set of documents of interest matching the indices, PHRASE, SEARCH, AUTHOR, ISSN and ISBN commands are used. The PHRASE command makes use of the common index and retrieves documents which contain all the terms one specifies on the operand. The SEARCH is the most general command, where the use of any kind of the indices and complexed operations including range searching are permitted. By entering a LOOK command, one gets the contents of the indices and can make a document set after examining them. In order to get the contents of the document sets, one enters a DISPLAY command, where five levels of output, from the simplest to the full, were implemented.

The present system forms part of the TOOL-IR system of the Computer Centre, University of Tokyo, and is widely used in the Japanese academic community.

1. はじめに

東京大学オンライン情報検索システム TOOL-IR(university of TOkyo On-Line Information Retrieval system)における工学関係文献データベースとしては、昭和52年に COMPENDEX を導入し、その公開を続けてきたが、今回新たに、同じデータベースの姉妹編で工学分野の会議録文献を収録した、Ei ENGINEERING MEETINGS を加えることになった。このデータベース導入のためのオンライン化設計とプログラム開発を進めるとともに、COMPENDEX 検索用プログラムにも相当の改良を行った。本稿は両データベースの TOOL-IR への導入における設計上の諸論点に言及しつつ、そのオンライン検索全般に亘って論述するものである。

まず手始めに Ei ENGINEERIG MEETINGS データベース (EIM) の検索例を示して、データベースの内容と検索方法を概観してみる (図1)。TOOL-IR は、文献の標題、著者名、各種見出し句類等から抽出した索引語 (キーワード) を手掛りにして、会話方式で適切な文献抄録を検索・出力するシステムである。この例ではアモルファス太陽電池に関する文献を検索している。まず「太陽電池」ということで検索してみると (SEARCH コマンド)、EIM1985年1月号6197文献中の41件が SOLAR と CELL という単語を含んでいる。この41件の文献群は文献集合1番として保持される。一方「アモルファス」で検索すると92文献が検索され、これは文献集合2番になる。そこで2つの文献群に共通に含まれる文献だけを抽出すると (AND コマンド) 11件となるので、その内容を表示して (DISPLAY コマンド)、検索を終了する (END コマンド)。

この例は比較的単純な検索指示で有意な結果が得られた場合であるが、検索テーマによってはかなり手の込んだ検索指定 (検索戦略などと称される) を使う必要も生じる。以下本稿では、データベー

図1 Ei ENGINEERING MEETINGS データベースの検索例

```

>>EIM                                     ← EIMの検索開始
WELCOME TO TOOL-IR/ORION "EIM" DATABASE (Rel 850511).
Data: Ei ENGINEERING MEETINGS from Engineering Information Inc.
This Database Contains 6197 Records (JAN -- JAN 1985).

TYPE IN COMMAND
  1/ SEA SOLAR.AND.CELL                    ← 「太陽電池」の検索
    SEA SOLAR.AND.CELL

*      281          SOLAR
*      122          CELL
*      41      1/  ("SOLAR") AND ("CELL") END NOSAVE ← 該当41文献 (文献集合1)

TYPE IN COMMAND
  2/ S AMORPHOUS                            ← 「非結晶」の検索
    S AMORPHOUS

*      92      2/  AMORPHOU ← 該当92文献 (文献集合2)

TYPE IN COMMAND
  3/ AND 1,2                                ← 「太陽電池」と「非結晶」の組合せ
    AND 1,2

*      11      3/  1 AND 2 ← 該当11文献 (文献集合3)

TYPE IN COMMAND
  4/ DIS M.D                               ← 該当文献の表示 (Dモード〈詳細表示〉による)
    DIS M.D
-----
(      1)
ACC#: 85002398      EIM#: EIM8501-002398
AUTH: Shimada, Katsunori
AFFN: JPL, Pasadena, CA, USA
TITL: U.S. AND JAPANESE AMORPHOUS SILICON TECHNOLOGY PROGRAMS - A
      COMPARISON.
CITN: American Society of Mechanical Engineers (Paper) Publ by ASME, New
      York, NY, USA 84-WA/Sol-11, 4p PUBL: 1984 CODN: ASMSA4
ISSN: 0402-1215 LANG: English
CONF: Winter Annual Meeting of the American Society of Mechanical
      Engineers. SPSR: ASME, New York, NY, USA LOCN: New Orleans,
      LA, USA DATE: 1984 Dec 9-14 CNFN: 04949
TEXT: The U.S. Department of Energy (DOE)/Solar Energy Research Institute
      (SERI) Amorphous Silicon (a-Si) Solar Cell Program performs R&D on
      thin-film hydrogenated amorphous silicon for eventual development of
      stable amorphous silicon cells with 12% efficiency by 1988. The
      Amorphous Silicon Solar Cell Program in Japan is sponsored by the
      Sunshine Project to develop an alternate energy technology. While
      the objectives of both programs are to eventually develop a-Si
      photovoltaic (PV) modules and arrays that would produce electricity
      to compete with utility electricity cost, the U.S. program approach
      is research oriented and the Japanese is development oriented. 13
      refs.
MSBJ: SOLAR CELLS SSBJ: Silicon
XREF: PHOTOVOLTAIC CELLS - Research
FIDX: AMORPHOUS SILICON SOLAR CELL PROGRAMS / SOLAR-CELL EFFICIENCY
      IMPROVEMENT / PHOTOVOLTAIC PROGRAM BUDGET / JAPAN SUNSHINE
      PHOTOVOLTAIC PROJECT / CELL LIGHT-INDUCED DEGRADATION
CALC: 00-A702 / 00-A549 / 00-A913 / 00-A714 / 00-A741 / 00-A941
-----
(      2)
ACC#: 85002984      EIM#: EIM8501-002984
AUTH: Haneman, D.
AFFN: Univ of New South Wales, Sch of Physics, Kensington, Aust
TITL: SOLAR ELECTRICITY AND RECENT PROGRESS IN THIN FILM PHOTOVOLTAICS.
CITN: Australian Journal of Physics v 37 n 4 1984 p 449-462 PUBL: 1984
CODN: AUJPAS ISSN: 0004-9506 LANG: English

. . . . .
. . . . .

TYPE IN COMMAND
  4/ END                                    ← 検索終了
    END

END OF TOOL-IR/ORION "EIM" DATABASE.
>>

```

スの内容、TOOL-IR でのオンライン索引抽出方式、そして検索のためのコマンドという順で述べる。

なお、本稿とほぼ同一の内容のものは、別途、東京大学大型計算機センター・データベース・マニュアル10（1985年12月）に収録されている。

付記:両データベースの運用は、1987年4月より、学術情報センターにて行われることになった。

2. COMPENDEX データベース

2. 1 概要

COMPENDEX は、米国 Engineering Information, Inc (Ei 社) にて編集発行されている工学関係文献抄録データベースである。その源流は1884年10月の Journal of the Association of Engineering Societies に掲載された Index Notes に始まる。これは、当時セントルイスのワシントン大学土木工学教授であった John B. Johnson 博士が、工学関係文献抄録集の必要を感じ、100種の雑誌から1000件程の文献を抄録・編集したものである。この Index Notes は、1895年にニューヨークの Engineering Magazine 社によって1892-1895年累積版が出版されるに際し、現行の The Engineering Index (Ei)と改名された。1919年には American Society of Mechanical Engineers (ASME)が Ei の編集を継承し、同学会誌にこれが毎月掲載されるようになった。同時に Engineering Societies Library との提携関係が成立し、同図書館の受入雑誌の掲載論文を中心とした抄録編集が行われるようになり、この方式は今日に及んでいる。1934年には Ei の編集を専門に担当するための非営利団体 Engineering Index, Inc が設立され、同社の手により Ei の編集が続けられて現在に至る。なお同社は、業務の拡張を期して、1981年に社名を現行の Engineering Information, Inc (略称の Ei はそのまま)に変更している。

このように Engineering Index は工学文献抄録として長い歴史を持つものであるが、COMPENDEX は、その名称が COMPuterized ENgineering inDEX という造語であることからわかるとおり、Ei の磁気テープ版にあたる。Ei は1962年以降 Ei Monthly (Ei 誌) と称する独立の月刊誌として刊行されるようになり、これに伴う収録件数の拡大に対処するべく、その編集工程の電算化が進められた。この電算化の成果の一つとして、1969年に磁気テープ版の Ei である COMPENDEX が発行されるに至ったというわけである。

COMPENDEX は世界40カ国の3,500種の雑誌、報告、図書からの抄録を行っており、1969年以降の累積件数は150万件を越え、最近は年間14万件程の収録数に達している。

2. 2 内容

COMPENDEX データベースの内容は、図2のようである。すなわち各文献についての記述（文献レコード）は、下記の項目から構成されている（項目名の後に示した ACC#などは TOOL-IR における見出し文字）。また表1は後出の EIM の分も含めたデータ項目の一覧表である。

(1) アクセション番号 (ACC#) —— アクセション番号は、文献レコードをオンライン・データ

ベースに收容する際に割り振った TOOL-IR での固有番号である。この番号は次項の EIX#の数字部分そのままであり、年、月、連番の合計9桁となっている。

(2) COMPENDEX 内部番号 (EIX#) — この番号は COMPENDEX への収録順を示す原データベース中でのレコード識別番号であり、「EIX、西暦年号下2桁、月表示2桁、連番5桁」という形式である。年と月により、このレコードが COMPENDEX の何年何月号所収のものであるかが判別できる。

(3) 抄録番号 (ABS#) — この番号は冊子体の Ei Monthly (Ei 誌) における抄録番号である。Ei 誌では抄録を後述の「主見出し—副見出し」という索引語のアルファベット順に排列しており、その各号における1番からの排列番号が、この抄録番号である。従って、特定文献を COMPENDEX で検索した後、さらに Ei 誌の方でも見てみようとする場合には、この番号が有用になるわけである。しかし実際には、COMPENDEX と Ei 誌は内容的には同一で、Ei 誌の方が詳しいといったことはないので、上記のような相互参照は通常では不要であろう。

(4) 著者名 (AUTH) — 当該論文の著者名を、(共著の場合は、スラントで区切って、並べて) 記す。名前は「姓、名 中間名 …」という形式で収録されている (TOOL-IR では、これを一定の方式で「正規化」して、著者名のオンライン索引を作っている。この点は後述)。この項目には、「…… (Ed.)」の形で編者名が記されることもある。またこの項目には個人名のみを記すという方針なので、会社名、団体名による公表論文の場合、ここには「Anon」と書いておき、会社名等を次項の著者所属 (AFFN) に示すという表記方法が採られている。なお「Anon」は雑誌等の無署名記事の収録にあたって用いられる (図6)。

(5) 著者所属 (AFFN) — 第1著者の所属を「機関名 (大学名、団体名等)、学部・部門名、所在都市名、国名」といった体裁で記す。ここでは略記法が用いられており (univ など)、そのための略語表も一応規定されているが (参考資料⑤)、慣習的、特例的な記法もあり、さ程統一されていないようにみえない。従って字面 (じづら) 走査による検索の際には (後述、LIMIT コマンドによる)、注意が必要である。

(6) 標題 (TITL) — 論文、図書の標題を記す。また会議録については会議名を記す。英文のものはそのまま収録されるのであるが、英語以外の言語の場合、「原語による標題. \$left bracket\$ 英訳表題. \$right bracket\$」のような体裁で、原語と英訳を併記する (ここで \$left bracket\$ などは、文字どおり左角括弧などの記号類を記号自体を用いずに表す記法で、データベースではよくみられる)。原語標題ではアクセントの類は削除し、またウムラウトには e を補うという表記法が採られている。ロシア語は翻字される。なお日本語、中国語原文のものは英訳標題のみ記される。

(7) 引用 (CITN) — 原論文の掲載雑誌名、巻、号、年次、頁等を図2にみられるような略記法で記す。完全な誌名を知るには、PIE (参考資料③) と称する収録雑誌一覧表によらねばならないが、大抵は想像がつく。

図2 COMPENDEX データベースの内容

 (1)
 ACC#: 850105471 EIX#: EIX850105471 ABS#: 001981
 AUTH: Coats, Pamela K.
 AFFN: Florida State Univ, Coll of Business, Tallahassee, FL, USA
 TITL: ELECTRONIC FUNDS TRANSFER NETWORKS MODELING THEIR IMPACT ON BANK SERVICE GOALS.
 CITN: Comput Oper Res v 11 n 4 1984 p 415-426 CODN: CMORAP
 ISSN: 0305-0548 LANG: ENGLISH
 DTYP: JA (Journal Article) TRMT: A (Applications) / N (Numeric/Statistical) / T (Theoretical)
 TEXT: The purpose of this research is to illustrate how an EFTS decision support model can assist bankers. It demonstrates how such a model is constructed and applied to the simulation and evaluation of a proposed EFTS network design. Furthermore, it shows how the model is altered to reflect changes in the network's design and, thus, its performance. A simple case is used to do this. In addition, the paper describes: (1) the simulation technique; (2) the network components to be modeled; (3) the method of altering the network's operational characteristics; (4) the model's computer program; (5) validation of the model; and (6) the model's output statistics for measuring the network's performance. 25 refs.
 MSBJ: DATA PROCESSING SSBJ: Financial Applications
 XREF: DECISION THEORY AND ANALYSIS
 FIDX: ELECTRONIC FUNDS TRANSFER NETWORKS / ELECTRONIC BANKING
 CALC: 00-A723 / 00-A922

表1 COMPENDEX, Ei ENGINEERING MEETINGS データベースの
 データ項目と TOOL-IR オンライン索引, 検索コマンド例

区分	項目見出し, 項目内容, 実例 (〈C〉はCOMPENDEXのみ, 〈M〉はEIMのみ該当)	オンライン索引	検索コマンド例
識別番号	ACCN: アクセション番号 ACCN: 850105813 <C> ACCN: 85000317 <M>		DISPLAY A. 850105813
	EIX #: COMPENDEX 内部番号 <C> EIX #: EIX 850105813		
	ABS #: 抄録番号 <C> ABS #: 005942		
	EIM #: 出版番号 <M> EIM #: EIM 8501-00317		
著者名	AUTH: 著者名 AUTH: Gadalla, Ahmed M.	A. 正規化著者名 A. GADALLA AM	AUTHOR GADALLA AM AUTHOR CADALLA
	AFFN: 著者所属 AFFN: Texas A&M Univ, Chemical ...		
標題	TITL: 標題 TITL: THERMAL CATALYTIC POLYMERIZATION OF MORTARS ...	標題中のキーワード THERMAL CATALYTI * ATALYTIC POLYMERI * RIZATION MORTAR	PHRASE THERMAL CATALYTIC
	TTIL: 英訳表題 <M>	(TITL と同様の単語索引)	(TITL と同様)

表1 つづき

原 誌 関 係	CITN: 引用 CITN: Polym Eng Sci v 24 n 16 ...		
	PUBL: 出版年 <M> PUBL: 1984		
	CODN: CODEN CODN: PYESAZ	PYESAZ	S PYESAZ
	ISSN: 国際標準雑誌番号 ISSN: 0032-3888	S. 0032388	ISSN 0032-3888
	ISBN: 国際標準図書番号 ISBN: 0-08-031733 2	B. 008031733	ISBN 0 08 031733-2
	LANG: 言語種別 LANG: ENGLISH		
種 別	DTYP: 資料種別 <C> DTYP: JA (Journal Article)	D. 種別コード D. JA	S D. JA
	TRMT: 内容種別 <C> TRMT: T (Theoretical)	T. 種別コード T. T	S T. T
会 議 関 係	CONF: 会議名 <M> CONF: Proceedings of the 2nd European Conference ...	会議名中のキーワード PROCEEDI *OCEEDING EUROPEAN CONFEREN *NFERENCE	PHRASE EUROPEAN CONFERENCE
	CNFX: 会議関係補足情報 <M> CNFN: 04675	(CONF と同様の単語 索引)	(CONF と同様)
	SPSR: 主催者・後援者 <M> SPSR: Cent Natl d'Etudes ...		
	LOCN: 開催地 <M> LOCN: Munich, West Ger		
	DATE: 開催日付 <M> DATE: 1982 Oct 19-22		
	CNFN: 会議番号 <M> CEFN: 04675	C. 会議番号 C. 04675	S C. 04675
要 旨	TEXT: 要旨 TEXT: This conference proceedings contains...		
見 出 し 類	MSBJ: 主見出し MSBJ: POLYMERS	各キーワード POLYMER	S POLYMER
	SSBJ: 副見出し SSBJ: Mixing	各キーワード MIXING	S MIXING
	XREF: 相互参照 XREF: INDUSTRIAL PLANTS - Control	各キーワード INDUSTRT *DUSTRIAL PLANT CONTROL	S INDUSTRIAL PHRASE INDUSTRIAL PLANTS
	FIDX: 自由語索引 FIDX: LASER OUTPUTS	各キーワード LASER OUTPUT	S LASER. AND. OUTPUT
	CALC: 分類コード CALC: 00 A744	Annn A744	S A744

(8) CODEN (CODN)、(9) 国際標準雑誌番号 (ISSN)、(10) 国際標準図書番号 (ISBN) — 原論文掲載誌の CODEN、ISSN、ISBN を示す。COMPENDEX への収録雑誌は大部分 CODEN 付きのものであり、このデータベースでは、雑誌の識別記号として、ISSN よりも CODEN の方が重視されている。

(11) 言語種別 (LANG) — 原報の言語種別を示す。

(12) 資料種別 (DTYP) — 原報の形態(document type)を下記の10種に区分して示す。なお右の数字は1985年1月号での該当件数。

CP (Conference Proceedings)	175
DS (Dissertation)	1
JA (Journal Article)	8202
MC (Monograph Chapter)	0
MR (Monograph Review)	2
RC (Report Chapter)	0
RR (Report Review)	143
ST (Standard)	2
TX (Textbook)	0
UP (Unpublished Paper or Preprint)	0

(13) 内容種別 (TRMT) — 論文の内容を下記の10種に区分して示す(Treatment Designator)。複数の区分に該当するときは、それらをスラントで区切って並べる。右の数字は1985年1月号での該当件数。

A (Applications)	3155
B (Biographical)	8
E (Economic / Cost Data / Market Survey)	253
G (General Review)	504
H (Historical)	72
L (Literature Review / Bibliography)	42
M (Management Aspects)	190
N (Numerical / Statistical)	456
T (Theoretical)	3291
X (Experimental)	4245

なお、ISBN、DTYP、TRMT は、1985年1月号から新たに設けられた項目である。

(14) 要旨 (TEXT) — 論文要旨あるいは図書、会議録の書評を収録する。書評の場合は自由語索引 (FIDX、後出) に「EIREV」なるキーワードを挿入して、それらだけを抽出できるようにして

いる（図5）。要旨末尾には「10 refs.」のように参考文献数を記す。また会議録についての書評の場合、Ei ENGINEERING MEETINGS データベースにおける会議番号（CNFN）を記して、両データベース間の連絡を図っている。

上記各項目の他に、主見出し（MSBJ）、副見出し（SSBJ）、相互参照（XREF）、自由語索引（FIDX）、分類コード（CALC）があるが、これらは検索の効率化のために設けられた索引項目であるので、別項にて説明する。

3. Ei ENGINEERING MEETINGS データベース

3.1 概要

Ei ENGINEERING MEETINGS データベース（EIM）は、一口にいえば、従来 COMPENDEX に含まれていた会議録所収論文の抄録を分離独立させ、拡大したものである。会議録関連文献は、量的に増加すると同時にその重要度も高まってきているという状況に対応して、会議録関係の収録範囲をひろげ、雑誌論文を中心とする COMPENDEX とは別建てのデータベースとして編集・頒布することにしたのである。EIM は1982年7月から月刊にて発行されており、約2,000種の会議録から年間11万件程の論文が収録されている。EIM の場合、COMPENDEX に対する Ei 誌のような冊子体抄録誌は発行されていなかったが、1985年春に The Engineering Conference Index と題する抄録の第1巻（7分冊）が刊行された。これは EIM の1983年7月号から1984年6月号所収の91,000文献を印刷したものであり、今後このような年鑑形式での出版が維持されるものと思われる。

EIM には、個々の論文についての抄録の他に、会議全体について説明を行ったレコードが挿入されている。このレコードは COMPENDEX 中の会議録の書評レコードと同じものであって、これにより両データベース間の連絡が図られている。つまり、COMPENDEX の検索によっても会議の概要、目次程度の情報は得られ、また EIM 中で会議を特定するための会議番号（CNFN）もわかるから、これらによって個々の論文抄録を EIM から容易に検索することができる。

上記の会議録書評レコードは COMPENDEX と EIM のつなぎの役割を果たすために重複的に収録されているものであるが、この他にも両データベースに重複して収録される文献が若干存在する。学会の定期刊行雑誌の特定号に特集といった形で学会主催の会議での発表論文が掲載された場合、この種の論文は通常の雑誌論文でもあり、また会議発表論文でもあるから、COMPENDEX、EIM の双方のこれを収録するというのが Ei 社の方針であるとのことである。こうした重複は5～9%程度と見込まれている。

3.2 内容

EIM の各文献レコードには、下記の項目が含まれる（図3）。

- (1) アクセション番号（ACC#）—— EIM におけるアクセション番号は、次記の出版番号（EIM#）の中の年号2桁と連番6桁を連結して、「85001234」のような形式にした。
- (2) 出版番号（EIM#）—— 出版番号は COMPENDEX の EIX# と同等の抄録識別番号である

図3 Ei ENGINEERING MEETINGS データベースの内容

TYPE IN COMMAND

1/ D A.85000344,M.D
D A.85000344,M.D

```
-----
(      1)
ACC#: 85000344      EIM#: EIM8501-000344
AUTH: Saburi, Yoshikazu / Yoshimura, Kazuyuki / Kawajiri, Nobuhiro /
      Kawano, Nobuyuki / Takahashi, Fujinobu / Kawaguchi, Noriyuki
AFFN: Radio Research Lab, Tokyo, Jpn
TITL: K-3 VLBI SYSTEM DEVELOPED IN RRL FOR US-JAPAN JOINT EXPERIMENT.
CITN: Publ by Cepadues-Editions, Toulouse, Fr p 277-290      PUBL: 1983
      ISBN: 2-85428-087-3      LANG: English
CONF: Techniques d'Interferometrie a Tres Grande Base.      SPSR: Cent Natl
      d'Etudes Spatiales, Toulouse, Fr      LOCN: Toulouse, Fr      DATE: 1982
      Aug 31-Sep 2      CNFN: 05022
TEXT: In the Radio Research Laboratories (RRL), a high precision VLBI
      system (K-3) for geodesy has been developed since 1979 according to
      the five-year plan, which is compatible with Mark-III system of the
      U.S.A. Both the hardware and the software of K-3 system completed in
      1982, except the S/X-band receiving system. The joint experiments
      between NASA and RRL started at the beginning of 1984 and will be
      continued at least for five years. 7 refs.
MSBJ: INTERFEROMETRY      SSBJ: Applications
XREF: ANTENNAS - Applications / MASERS - Applications
FIDX: DATA ACQUISITION TERMINAL / DATABASE SOFTWARE / LEAST SQUARE
      ESTIMATION SOFTWARE / HARDWARE SYSTEM ERRORS
CALC: 00-A941 / 00-A655 / 00-A723 / 00-A716 / 00-A714 / 00-A657
-----
```

が、形式は異っており、「EIM、西暦年号下2桁、月表示2桁、ハイフン、連番6桁」を連結して、「EIM8501-001234」のような形になっている。これで何年何月号所収の抄録であるかが判別できる。

(3) 著者名 (AUTH) — COMPENDEX に同じ。

(4) 著者所属 (AFFN) — COMPENDEX に同じ。

(5) 標題 (TITL) — 論文標題。会議全体について解説する書評レコードの場合は、本項目に会議名そのものが記入される。英文以外のものは、COMPENDEX と同じ方式で翻字されている。なお、これらに対する英訳標題が別項目になっている点は、COMPENDEX と異なる。中国語、日本語原文のものは、TITL に直接英訳標題が記入され、原文は収録されない。

(6) 英訳標題 (TTIL) — 英語以外の欧文の標題の英訳。

(7) 引用 (CITN) — 定期刊行雑誌所収の会議論文の場合、前述のとおり COMPENDEX と併載になるが、その雑誌名は略語形でなく完全綴りで本項目に記される。単行書形式の会議録の場合、その会議録名 (書名) は会議名 (CONF) 項目の方に記されるので、本項目には出版社名、入手先等が記入されている (会議録名がこの項目にも記されているレコードもみられる。この点 Ei 社での編集作業にまだ不安定な部分があるように見える)。

(8) 出版年 (PUBL) — 原報の出版年。

(9) CODEN (CODN)、(10) 国際標準雑誌番号 (ISSN)、(11) 国際標準図書番号 (ISBN)、(12) 言語種別 (LANG) — COMPENDEX に同じ。

(13) 会議名 (CONF) — 会議名が完全綴りで記入される。正確には会議録名というべきで、

「Proceedings of …」という書き方が多い。

- (14) 会議関係補足情報 (CNFX) —— CONF 項目への補足ということであるが、ほとんどみられない。
- (15) 主催者・後援者 (SPSR) —— 会議のスポンサー名とその所在地を記す。
- (16) 開催地 (LOCN) —— 会議の開催地名。
- (17) 開催日付 (DATE) —— 会議開催日付を「1984 Aug 31-Sep 2」のような体裁で記す。
- (18) 会議番号 (CNFN) —— Ei 社にて各会議ごとに割付けた 5 桁の整理番号。会議の識別番号として、同一会議の全論文の一括検索などに有用。また、COMPENDEX 中の会議録書評レコードの要旨にもこの番号が書かれており、COMPENDEX から EIM へのつなぎ項目である。
- (19) 要旨 (TEXT) —— 論文要旨。会議全体についての書評レコードでは、会議の構成、発表件数等が記されている。

以上の他、COMPENDEX と同じく、主見出し (MSBJ)、副見出し (SSBJ)、相互参照 (XREF)、自由語索引 (FIDX)、分類コード (CALC) がある。これらについては次項にまとめて解説する。

4. 見出し語と分類

4. 1 主見出し (MSBJ)、副見出し (SSBJ)

今日のようにオンライン検索が普及する以前、文献の検索には抄録誌が主に用いられた。本の形態となっている抄録誌では、検索に便利なように文献抄録をうまく並べることが大切である。一般的には、論文の内容を表す種々のキーワード（見出し語、descriptor という）の辞書を予め作っておき、その辞書の排列に従って各キーワードへの該当文献の抄録を並べるといった方法が採られる。利用者は、自分の探索したいテーマをよく表わすキーワードをその辞書から見つけ出し、その見出し語のもとに集められている文献を点検すればよい。もっとも 1 つの論文が 2 つ以上の分野にまたがっている場合、これらをすべて重複して掲載しては抄録誌がかさばるばかりであるから、「をも見よ」という参照表示で代替するなど、編集上様々の工夫が必要である。そもそもこうした問題があまり生じないよう、見出し語の選択、体系化、つまり辞書（統制語彙）の整備と更新には、相当力を入れる必要がある。

Ei 社では、Ei 誌編集のために「主見出し (Main Subject Heading) — 副見出し (Subheading)」という 2 段階の見出し語辞書を作っており、これを SHE (Subject Headings for Engineering) と称している (参考資料①)。現行の SHE は 1983 年版であり、これには 12,000 件の見出しが登録されている。SHE には、各見出しの適用される場合、されない場合等の注意書きや、対応する分類コード (CALC) も記されている (図 4)。Ei の慣例として、主見出しは大文字のみ、副見出しは小文字まじりで表記する。「主見出し — 副見出し」の対は、各文献について 1 件だけ割当てられ、Ei 誌では、この見出しのアルファベット順に文献抄録が排列される。

図4 SHE(Subject Headings for Engineering)の内容例

NUCLEAR ENERGY	621, 932
<i>(Use for general subject and for applications not elsewhere classifiable. Otherwise use subheading -Nuclear Energy under heading for application)</i>	
Fission Reactions	
Fusion Reactions	
Thermonuclear Reactions	
NUCLEAR ENGINEERING	(613), (621), (622), 901
<i>(Beginning 01/77. For subheadings, see ENGINEERING)</i>	
NUCLEAR EXPLOSIONS	(404), 621, 932
<i>(For subheadings, see EXPLOSIONS)</i>	
NUCLEAR FUELS	621, 622
Analysis	
Cladding <i>(Beginning 01/77)</i>	
Coated Particles <i>(Beginning 01/77)</i>	
Explosions <i>(See also NUCLEAR EXPLOSIONS)</i>	
Fission <i>(Beginning 01/77)</i>	
Fuel Elements <i>See NUCLEAR REACTORS-Fuel Elements</i>	
Fusion <i>(Beginning 01/77)</i>	
Irradiation	
Management	
Measurements <i>(See also NUCLEAR INSTRUMENTATION)</i>	
	944
Metallography	
	531
Metallurgy	
	531
Mixed Oxides <i>(Beginning 01/77)</i>	
Moderators <i>See NUCLEAR REACTORS-Moderators</i>	
Pelletizing	
Processing <i>(Beginning 01/77)</i>	
Properties	
	932
Reactor Core <i>See NUCLEAR REACTORS-Cores</i>	
Reprocessing	
	802
Research	
Safe Handling	
	914
Standards	
Testing	
Uranium Carbide <i>See URANIUM CARBIDE</i>	
NUCLEAR INDUSTRY	(613), (621), (622)
<i>(Beginning 01/83)</i>	
NUCLEAR INSTRUMENTATION	944
<i>(Use for general subject or for nuclear instrumentation for which no specific heading is available. Otherwise use specific heading)</i>	
NUCLEAR MEDICINE	462, (622)
<i>(Beginning 10/75)</i>	
NUCLEAR PHYSICS	<i>See PHYSICS-Nuclear</i>

主見出しには形態上、単語、句（複数語）、逆転句という3つのスタイルがある（「WATER」、
「WATER COOLING SYSTEMS」、
「WATER TREATMENT, INDUSTRIAL」）。内容的には、物（物質、機械等）と処理（操作、現象等）が選ばれているということである。

副見出しは、主見出しの文献上での重要なアスペクトを記述するものとされ、主見出しの、部分、
応用、研究法、性質などである。そこで、異なる主見出しに従属する副見出し群が同一であることも

多い (AUTOMOBILE と MOTOR CYCLE に対する副見出し群など)。副見出しと主見出しの両方に登録されている用語もある。例えば「FUEL ECONOMY」は一般的なその場合は主見出しとなるが、自動車の燃費の場合には「AUTOMOBILE - Fuel Economy」のように副見出しになる。副見出しの方から、それがどのような主見出しとともに使われるかをリストした SUBHEADING INDEX が COMPENDEX Search Manual (参考資料⑤) に収められている。

「主見出し-副見出し」は統制語彙の性格上時代遅れになりがちであり、また文献あたり1件に限定して付与されることで、相当無理が生じる場合も多かろうが、まるで見当違いであるということはないはずである。これは案外重要な性質で、検索者が当該分野にあまり詳しくない場合とか代行検索者に検索を依頼する場合などに、一定の検索精度を保つための有効な情報といえよう。

4. 2 相互参照 (XREF)

前項の「主見出し-副見出し」は、各文献につき1件付与される。これに次いで重要とみられる「主見出し-副見出し」を、最大5件まで、この相互参照項目に記入する。上にも述べたとおり、1論文あたり見出し1件というのは無理のあることも多いと思われ、相互参照中に収録された見出し語も、その重要度において、それ程差はないと考えた方がよかろう。

図5 書評レコードの例(FIDX<自由語索引>に「EIREVを含む」)

```

-----
(      1)
ACC#: 850104847      EIX#: EIX850104847      ABS#: 002774
AUTH: Anderson, Alun M.
AFFN: Kyoto Univ, Kyoto, Jpn
TITL: SCIENCE AND TECHNOLOGY IN JAPAN.
CITN: Sci and Technol in Jpn Publ by Longman Group Ltd (Longman Guide to
World Science and Technology, v 4), Harlow, Engl, 1984. Distributed
by Gale Research Co, Detroit, MI, USA 421p      LANG: ENGLISH
DTYP: MR (Monograph Review)      TRMT: A (Applications) / G (General Review)
/ X (Experimental)
TEXT: This is a publication in the series 'Longman Guide to World Science
and Technology'. It charts the major research programs now underway
in Japan and provides a detailed directory of the research
institutes, universities and industrial companies where they are
being carried out. The book is the first to describe comprehensively
all of Japan's major government-industry cooperative research
projects and the attempts being made to stimulate basic research. In
introductory chapters the strengths and weaknesses of Japanese
technology and its historical development are discussed, and the
political systems and national bodies involved in science policy
decision making are described. The background necessary to
understand Japan's scientific and technological position is
completed with chapters on government science policy, the ministries
and major revenue programs; education and academic research; and
industrial research and development. Separate chapters are devoted
to detailing the work of the major research institutes in the field
of agriculture; biotechnology, medicine and life sciences; energy;
aerospace, aviation and railways; electronics; industrial research
and development; earthquake and disaster prevention, environmental
protection and earth sciences; marine science and technology; and
defense. Relevant societies and information sources are also
described.
MSBJ: ENGINEERING RESEARCH      SSBJ: Japan
XREF: TECHNOLOGY - Japan / EDUCATION - Japan / ENGINEERING - Public Policy
/ RESEARCH LABORATORIES - Japan / ENERGY POLICY - Japan
FIDX: AEROSPACE RESEARCH / AGRICULTURAL ENGINEERING / ELECTRONIC /
BIOTECHNOLOGY / EIREV
CALC: 00-A901 / 00-A912 / 00-A714 / 00-A715 / 00-A621 / 00-A461
-----

```

4. 3 自由語索引 (FIDX)

自由語索引は、SHE 登録の見出し語で表現しきれない重要な内容、特色などについて、Ei 抄録者が原報から抜き出し、また独自に考案して付与する非統制の索引語であり、文献あたり最大5件まで付けられている。統制語彙はどうしても時代遅れになるし、また汎用的、一般的で、個々の文献の特色を出しにくい。これらを補う意味で自由語索引は重要である。また Ei 社にとっても自由語索引は SHE 改定のための資料になっているようであり、この意味では暫定版の SHE という見方もできる。なお書評レコードについては、その識別のために「EIREV」という特別の索引語がこの項目に収容されている (図5)。

4. 4 分類コード (CALC)

COMPENDEX、EIM における文献の分類は、CAL 分類コードという3桁の数字によってなされ、1文献につき最大6個まで割付けられる。CAL コードというのは、1975年まで Ei 社で実施されていたカードによる SDI サービス (CARD-A-LERT サービス) の際に使われたコード体系で、同サービス廃止後も分類コードとして利用されているものである。CAL コードは、データベース中では先頭に「00-A」が付加され、「00-A403」のような形で表記されている。3桁の数字は、100の位が大分類、10の位が中分類、1の位が小分類を表し、各分類数は6大分類、38中分類、171小分類である (表2)。表では、TOOL-IR での索引方式 (後述) に合わせて、大分類を A4*、中分類を A40*、小分類を A401のように記した。

5. オンライン索引

TOOL-IR による COMPENDEX、EIM の検索は、CA Search などと同様、一次検索と二次検索の二様によって行われる。一次検索とは、本の巻末索引を引くように予め抽出されている索引語の中から適当なものを指定して、その語を含む文献群 (それを文献集合という) を一時的に作り出すことである。二次検索には、一次検索によって作られたいくつかの文献集合について、さらにその全体 (OR) あるいは重複部分のみ (AND) の文献集合を作り出すことと、一つの文献集合内の各文献について特定の項目の字面を調べてゆき、適当なものだけを選択して新たな文献集合とすることの、2つがある。いずれにせよ二次検索とは、現に作られている文献集合を対象として、それらを組合せたりさらに絞り込んだりする操作である。従ってオンライン検索にとって第一義的に重要なのは一次検索ということになる。

一次検索は、上述のとおり、予め用意されている一次検索用の索引、すなわちオンライン索引による検索である。この索引は原データベースをオンライン化する際に一定基準で作成されるものであるから、そこにおける索引語の抽出方法を一通り心得ておくのは、利用者にとって必要なことである。TOOL-IR における COMPENDEX と EIM データベースでは、次の8種のオンライン索引を用意している。表1では各データ項目から抽出される索引語を例示し、同時にそれらに対する検索コマンドの例も示した。

表2 分類コード(CALC)表

```

=====
A4* Civil - Environmental - Geological - Bioengineering
-----
A40* CIVIL ENGINEERING, GENERAL
A401 Bridges and Tunnels
A402 Buildings and Towers
A403 Urban and Regional Planning and Development
A404 Civil Defense Military Engineering
A405 Construction Equipment and Methods; Surveying
A406 Highway Engineering
A407 Maritime and Port Structures; Rivers and Other Waterways
A408 Structural Design
A409 Civil Engineering
A41* CONSTRUCTION MATERIALS
A411 Bituminous Materials
A412 Concrete
A413 Insulating Materials
A414 Masonry Materials
A415 Metals, Plastics, Wood and Other Structural Materials
A42* MATERIALS PROPERTIES AND TESTING
A421 Strength of Materials; Mechanical Properties
A422 Strength of Materials; Test Equipment and Methods
A423 Miscellaneous Properties and Tests of Materials
A43* TRANSPORTATION
A431 Air Transportation
A432 Highway Transportation
A433 Railroad Transportation
A434 Waterway Transportation
A44* WATER AND WATERWORKS ENGINEERING
A441 Dams and Reservoirs; Hydro Development
A442 Flood Control; Land Reclamation
A443 Meteorology
A444 Water Resources
A445 Water Treatment, General and Industrial
A446 Waterworks
A45* POLLUTION, SANITARY ENGINEERING, WASTES
A451 Air Pollution
A452 Sewage and Industrial Wastes Treatment
A453 Water Pollution
A454 Environmental Engineering
A46* BIOENGINEERING
A461 Biotechnology
A462 Medical Engineering and Equipment
A47* OCEAN AND UNDERWATER TECHNOLOGY
A471 Marine Science and Oceanography
A472 Ocean Engineering
A48* ENGINEERING GEOLOGY
A481 Geology and Geophysics
A482 Mineralogy and Petrology
A483 Soil Mechanics and Foundations
A484 Seismology
=====
A5* Mining - Metals - Petroleum - Fuel Engineering
-----
A50* MINING ENGINEERING, GENERAL
A501 Exploration and prospecting
A502 Mine and Quarry Equipment and operations
A503 Mines and Mining, Coal
A504 Mines and Mining, Metal
A505 Mines and Mining, Nonmetallic
A506 Mining Engineering
A51* PETROLEUM ENGINEERING
A511 Oil Field Equipment and Production Operations
A512 Petroleum and Related Deposits
A513 Petroleum Refining
A52* FUEL TECHNOLOGY
A521 Combustion and Fuels
A522 Gas Fuels
A523 Liquid Fuels
A524 Solid Fuels
A525 Energy Management
A53* METALLURGICAL ENGINEERING, GENERAL
A531 Metallurgy and Metallography
A532 Metallurgical Furnaces
A533 Ore Treatment and Metal Refining
A534 Foundry Practice
A535 Rolling, Forging and Forming
A536 Powder Metallurgy
A537 Heat Treatment

```

- A538 Welding and Bonding
- A539 Metals Corrosion and Protection; Metal Plating
- A54* METALLURGICAL ENGINEERING, METAL GROUPS
- A541 Aluminum and Alloys
- A542 Beryllium, Magnesium, Titanium and Other Light Metals and Alloys
- A543 Chromium, Manganese, Molybdenum, Tantalum, Tungsten, Vanadium and Alloys
- A544 Copper and Alloys
- A545 Iron and Steel
- A546 Lead, Tin, Zinc, Antimony and Alloys
- A547 Minor, Precious and Rare Earth Metals and Alloys
- A548 Nickel and Alloys
- A549 Nonferrous Metals and Alloys in General
-
- A6* Mechanical - Automotive - Nuclear - Aerospace Engineering
-
- A60* MECHANICAL ENGINEERING, GENERAL
- A601 Mechanical Design
- A602 Mechanical Drives and Transmissions
- A603 Machine Tools
- A604 Metal Cutting and Machining
- A605 Small Tools and Hardware
- A606 Abrasives
- A607 Lubricants and Lubrication
- A608 Mechanical Engineering
- A61* MECHANICAL ENGINEERING, PLANT AND POWER
- A611 Hydro and Tidal Power Plants
- A612 Combustion Engines
- A613 Nuclear Power Plants
- A614 Steam Power Plants
- A615 Thermoelectric, Magnetohydrodynamic and Other Power Generators
- A616 Heat Exchangers
- A617 Turbines and Steam Engines
- A618 Compressors and Pumps
- A619 Pipes, Tanks and Accessories; Plant Engineering Generally
- A62* NUCLEAR TECHNOLOGY
- A621 Nuclear Reactors
- A622 Radioactive Materials
- A63* FLUID FLOW; HYDRAULICS, PNEUMATICS, AND VACUUM
- A631 Fluid Flow; Hydrodynamics
- A632 Hydraulics, Pneumatics and Related Equipment
- A633 Vacuum Technology
- A64* HEAT AND THERMODYNAMICS
- A641 Heat and Mass Transfer; Thermodynamics
- A642 Industrial Furnaces and Process Heating
- A643 Space Heating and Air Conditioning
- A644 Refrigeration and Cryogenics
- A65* AEROSPACE ENGINEERING
- A651 Aerodynamics
- A652 Aircraft
- A653 Aircraft Engines
- A654 Rockets and Rocket Propulsion
- A655 Spacecraft
- A656 Space Flight
- A657 Space Physics
- A658 Aerospace Engineering
- A66* AUTOMOTIVE ENGINEERING
- A661 Automotive Engines and Related Equipment
- A662 Automobiles and Smaller Vehicles
- A663 Buses, Tractors, Trucks and Vehicle Traction
- A664 Automotive Engineering
- A67* MARINE ENGINEERING
- A671 Naval Architecture
- A672 Naval Vessels
- A673 Shipbuilding and Shipyards
- A674 Small Craft and Other Marine Craft
- A675 Marine Engineering and Naval Architecture
- A68* RAILROAD ENGINEERING
- A681 Railroad Plant and Structures
- A682 Railroad Rolling Stock
- A69* MATERIALS HANDLING
- A691 Bulk Handling and Unit Loads
- A692 Conveyors and Elevators
- A693 Cranes and Derricks
- A694 Packaging and Storing
-
- A7* Electrical - Electromics - Control Engineering
-
- A70* ELECTRICAL ENGINEERING, GENERAL
- A701 Electricity and Magnetism
- A702 Electric Batteries and Fuel Cells
- A703 Electric Circuits

- A704 Electric Components and Equipment
- A705 Electric Generators and Motors
- A706 Electric Transmission and Distribution
- A707 Illuminating Engineering
- A708 Electric and Magnetic Materials
- A709 Electrical Engineering
- A71* ELECTRONICS AND COMMUNICATION ENGINEERING
- A711 Electromagnetic Waves
- A712 Electronic and Thermionic Materials
- A713 Electronic Circuits
- A714 Electronic Components and Tubes
- A715 Electronic Equipment, General Purpose and Industrial
- A716 Electronic Equipment, Radar, Radio and Television
- A717 Electro-Optical Communication
- A718 Telephone and Other Line Communications
- A72* COMPUTERS AND DATA PROCESSING
- A721 Computer Circuits and Logic Elements
- A722 Computer Hardware
- A723 Computer Software and Data Handling
- A73* CONTROL ENGINEERING
- A731 Automatic Control Principles
- A732 Control Devices
- A74* LIGHT AND OPTICAL TECHNOLOGY
- A741 Light, Optics and Optical Devices
- A742 Cameras and Photography
- A743 Holography
- A744 Lasers
- A745 Printing and Reprography
- A75* SOUND AND ACOUSTICAL TECHNOLOGY
- A751 Acoustics, Noise, Sound
- A752 Sound Devices, Equipment and Systems
- A753 Sound Technology and Ultrasonics
- =====
- A8* Chemical - Agricultural - Food Engineering
-
- A80* CHEMICAL ENGINEERING, GENERAL
- A801 Chemical Analysis and Physical Chemistry
- A802 Chemical Apparatus and Plants; Unit Operations; Unit Processes
- A803 Chemical Agents and Basic Industrial Chemicals
- A804 Chemical Products Generally
- A805 Chemical Engineering
- A81* CHEMICAL ENGINEERING, PROCESS INDUSTRIES
- A811 Cellulose, Paper and Wood Products
- A812 Ceramics and refractories
- A813 Coating and Finishes
- A814 Leather and Tanning
- A815 Plastics, Chemistry and Materials
- A816 Plastics, Plant Equipment and Processes
- A817 Plastics, Products and Applications
- A818 Rubber and Elastomers
- A819 Synthetic and Natural Fibers; Textile Technology
- A82* AGRICULTURAL ENGINEERING AND FOOD TECHNOLOGY
- A821 Agricultural Equipment and Methods
- A822 Food Technology
- =====
- A9* Industrial Engineering - Management - Mathematics - Physics
- Instruments
-
- A90* ENGINEERING, GENERAL
- A901 Engineering Profession
- A902 Engineering Graphics; Engineering Standards; Patents
- A903 Information Science
- A91* ENGINEERING MANAGEMENT
- A911 Cost and Value Engineering; Industrial Economics
- A912 Industrial Engineering and Management
- A913 Production Planning and Control
- A914 Safety Engineering
- A92* ENGINEERING MATHEMATICS
- A921 Applied Mathematics
- A922 Statistical Methods
- A93* ENGINEERING PHYSICS
- A931 Applied Physics Generally
- A932 High Energy Physics; Nuclear Physics; Plasma Physics
- A933 Solid State Physics
- A94* INSTRUMENTS AND MEASUREMENT
- A941 Acoustical and Optical Measuring Instruments
- A942 Electric and Electronic Measuring Instruments
- A943 Mechanical and Miscellaneous Measuring Instruments
- A944 Moisture, Pressure and Temperature,
and Radiation Measuring Instruments
- =====

(1) 一般語索引—— TITL、TTIT、CONF、CNFX、MSBJ、SSBJ、XREF、FIDX (すなわち標題と見出し類) 中の各単語と CODEN および分類コード (CALC) を、索引語として、一般語索引に収録している (TTIT、CONF、CNFX は EIM のみに該当)。ここではその索引語が TITL に由来するものか MSBJ からのものかといった区別はないので、一様に検索できる (こういう区分を意識した検索は二次検索による)。

各索引語は、語尾変化による意図せざる検索漏れを防止するために、最大8文字とし、それ以降は切落している。同様の趣旨から4文字以上の語の語尾のSは複数形のSとみなし、これも切落している (SYSTEMS →SYSTEM)。一方、語尾による検索が重要な場合もあるので (接頭辞のついた単語や複合語など)、8文字以上の語については語尾から8文字分の逆索引を作っている。

分類コード (CALC) はデータベース内では「00-A401」のような記法になっているが、索引化の際に「00-」を切落し、「A401」のようにAおよびコード3桁の形で索引語としている。

なお、以下の各種索引も含めて、キャピタリゼーションの差異を吸収するため、本文中の英小文字はすべて大文字に変換して索引化している。

(2) 著者名索引 (A.) —— AUTH 項目に記されている個人名に次のような「正規化」処理を施し、著者名索引に収録する。AUTH では「姓, 名 中間名 ……」のような記法になっているが、この中の名や中間名類の表記は、場合によって頭文字化されていたりして、不安定である。これらをそのまま索引化すると同一人について多数の索引が生じ、検索する方ではそうした表記の各種変形を想定しつつ検索指定を行わなければならなくなり、不都合である。そこで TOOL-IR では、名や中間名はすべて頭文字だけを採録し連結して、著者名索引とすることになっている (下例)。

Carpenter, B. Stephen—— →CARPENTER BS

Dalla Betta, Ralph A. —— →DALLA BETTA RA

なお著者不詳の際の「Anon」は、そのまま索引化されている。

実際の著者名索引は、著者名索引であることの識別のために、上記による正規化氏名の前に「A.」を付加し、「A.CARPENTER BS」のような形式としている。従って後述の LOOK コマンドで著者名索引を通覧する際には、「A.」を付加した形での指定が必要になる。

(3) 国際標準雑誌番号 (ISSN) 索引 (S.) —— ISSN を「正規化」(後述) し「S.」を付加して収録する。

(4) 国際標準図書番号 (ISBN) 索引 (B.) —— ISBN を「正規化」(後述) し「B.」を付加して収録する。

(5) 資料種別索引 (D.) <COMPENDEX のみ> —— 資料種別 (DTYP) の各コードを、先頭に「D.」を付加して、「D. CP」のような形式で索引化している。

(6) 内容種別索引 (T.) <COMPENDEX のみ> —— 内容種別 (TRMT) の各コードを、先頭に「T.」を付加して、「T. A」のような形式で索引化している。

(7) 会議番号索引 (C.) <EIM のみ> —— 会議番号 (CNFN) を、先頭に「C.」を付加して、「C. 0

1234」のような形式で索引化している。

上記のオンライン索引に対して、次項で説明するように SEARCH、PHRASE、AUTHOR、ISSN、ISBN の各検索コマンドを用いて、一次検索を行うわけである。

6. 検索コマンド

6. 1 検索コマンドの構成

COMPENDEX、EIM データベース検索のためのコマンドは、CA Search 等のデータベース検索の場合と基本的に共通であり、これらは TOOL-IR の標準的な検索コマンド群といえる。ただしデータベース毎に収録項目に特異性があるから、一部に COMPENDEX、EIM に特有の検索指定が含まれる。

コマンドの全体をまとめて示すと表3のようになる。SEARCH コマンドは最も基本的な検索コマンドであり、前述の一般語索引等を対象にして検索を行うものである。PHRASE コマンドは多重 AND 操作を内蔵させたコマンドであって、主として論文標題、会議名等がおよそわかっている場合に、効率的に検索を行うためのものである。AUTHOR、ISSN、ISBN は、これらの項目（索引）の検索に特有な処理を内蔵させたコマンドである。以上が一次検索用のコマンドであり、これらを用いて文献集合を作成する。こうして作られた文献集合に対してさらに操作を加えるのが、二次検索用のコマンドである。これには、集合間の重ね合せを行う AND、OR、DIF の各コマンド、および特定項目の字面を詳しく調べて文献集合をさらに絞り込むための LIMIT コマンドがある。

このようなコマンドを用いて所期の文献群が検索された後、その中の各文献の内容を実際に表示させるために、DISPLAY コマンドがある。もっとも DISPLAY コマンドは、検索結果を最終的に表示させるときばかりでなく、会話型検索の過程で次々と作成される中間結果的な集合の内容を一部覗いて点検するために、随時用いられるべきものである。

検索にあたっては、どのような索引語が採録されているのか、オンライン索引自体を直接見たいことがある。このため、オンライン索引通覧用として LOOK コマンドが用意されている。これは索引の表示を主機能とするものであるが、そのついでに一次検索も行えるようになっている。

表3には各検索コマンドの書き方も示されている。そこでは、コンピュータ関係のマニュアルでよく見かける角括弧 [] と中括弧 { } を用いた記法により、コマンドの指定法の一般形を記しておいた。

角括弧 [] は、その中に記された項目が省略可能であることを示す。つまり必要に応じて指定すればよいという意味である。なお DISPLAY コマンドと LIMIT コマンドでは、対象とする文献集合番号や表示項目指定 (M.) 等に標準値 (デフォルト値) が決められており、これらの指定が省略されたときは、その標準値の指定があったものとみなされる。標準値は一般によく使われような値を想定して決められているので、このような項目は実際には指定しないで済むことが多い。中括弧 { } は、これで囲まれたものの中から一つを選択して指定するという意味に用いられる。相互に排

表3 検索コマンドの種類と書き方

種別	コマンド名称	機能	書き方
一次検索	SEARCH (Sでも可)	キーワードによる検索	SEARCH kw (---kw) [, kw(---kw)]... $\left[\left\{ \begin{array}{l} \text{AND} \\ \text{NOT} \end{array} \right\} kw(---kw) [, kw(---kw)] \dots \right]$ kwは標題, 見出し類中のキーワード, 分類コード (Annn), CODEN, 資料種別 (D. x), 内容種別 (T. x), 会議番号 (C. x)。kw*で語頭一致, *kwで語尾一致, kw---kwで範囲限定。
	PHRASE	キーワード群による検索 (キーワード間の論理積)	PHRASE kw kw kw ...
	AUTHOR	著者名による検索	AUTHOR 正規化著者名[*] ... 正規化著者名とは「姓 名の頭字」。
	ISSN	国際標準雑誌番号による検索	ISSN issn [, issn] ...
	ISBN	国際標準図書番号による検索	ISBN isbn [, isbn] ...
二次検索	AND OR DIF	文献集合間の組合せ	AND 集合番号, 集合番号 [, 集合番号] ... OR 集号番号, 集合番号 [, 集合番号] ... DIF 集合番号-集合番号
	LIMIT	データ項目の内容走査による文献の選択	LIMIT {集合番号, } $\left\{ \begin{array}{l} \text{項目見出し} \\ \text{ALL} \end{array} \right\} \cdot \left\{ \begin{array}{l} [*] \text{文字列} [*] \\ [*] \text{文字列} 1 * \text{文字列} 2 [*] \end{array} \right\}$ 集合番号を省略すると最新の集合を仮定。文字列を当該項目に含む文献を選択。文字列1, 2を指定すると両者をこの順序で当該項目に含む文献を選択。*は文字列前後の空白の存在を明示するのに用い, 単語の走査を指示。
表示	DISPLAY (Dでも可)	文献内容の表示	DISPLAY $\left\{ \begin{array}{l} \text{A. アクセション番号} \\ \text{[S. 集合番号]} \end{array} \right\} \left[\begin{array}{l} \left\{ \begin{array}{l} \text{N. n} \\ \text{I. i1, i2, ...} \\ \text{I. i1[-i2]} \end{array} \right\} \end{array} \right]$ $\left[\begin{array}{l} \text{A} \\ \text{B} \\ \text{C} \\ \text{D} \\ \text{X} \end{array} \right]$ ([, M. ...]) 集合番号を省略すると最新の集合を仮定。Nは集合内の最初のn件のファイルを表示する指定。Iは集合内のi番目の文献を表示する指定。NおよびIを省略すると全件表示。Mは表示項目指定でAからDの順に詳しい内容表示となり, Dは全項目表示, Xは標題のみ表示。省略時は M. Xを仮定。
	LOOK	オンライン索引の通覧と検索	LOOK kw[*(*)] LOOK *kw (語尾一致用逆索引を通覧) LOOK A. 著者名[*(*)] (著者名索引を通覧) kwはSEARCHの項参照。末尾の*は語頭一致索引表示, **は後続索引表示。A.のほかS, B, D, T, C.の指定も可。表示件数はSETコマンドで変更。
	REMIND	検索経過の表示	REMIND $\left\{ \begin{array}{l} \text{集合番号[, 集合番号]} \dots \\ \text{集合番号-集合番号} \\ \text{[ALL]} \end{array} \right\}$ 指定された集合を生成した検索コマンドの内容と該当の文献数を表示。
	? INFO	データ項目表の表示	? INFO
終了	END	検索の終了	END

反の指定項目を並べて示すわけである。「…」は繰り返し指定が可能であることを示す。

例えば「AND 集合番号, 集合番号 [, 集合番号] …」とあるのは、AND コマンドでは集合番号をコンマで区切りながら2個以上指定せよという意味で、「AND 1, 2」、「AND 1, 3, 4, 6」などが可能である。次に DISPLAY コマンドに

$$[, M. \left. \begin{array}{c} A \\ B \\ C \\ D \\ X \end{array} \right\}]$$

とあるのは、まず M. による表示項目指定として「M. A」、「M. B」等々が可能であることを示し、次に全体の角括弧でこの指定自体が省略可能であることを示す。なお省略の際の標準値は後述のとおり「M. X」となっている。

6. 2 各検索コマンド

(1) SEARCH コマンド

SEARCH コマンドは、一般語索引（標題・見出し中のキーワード、CODEN、分類コード）および資料種別 (D.)、内容種別 (T.)、会議番号 (C.) を検索するコマンドである。指定の一般形式は表3のとおりであるが、基本的には「同義語」をコンマで区切って並べて、一括検索するのを主眼にしている。このときコンマは OR の意味付けを持ち、

SEARCH COMPUTER, ORDINATEUR, RECHNER

S COMPUTER, MICROPROCESSOR, MICROCOMPUTER

のように指定する (SEARCH はその略記として S や SEA でもよい)。同一の対象に対して多種の表現が用いられるのが通常であるから、検索の目的に応じて「同義語」を上手に設定することが肝要である。

前方一致指定 (単語*) と複数形——「*」記号を索引語の末尾に用いて「S COMPUT*」のように指定すると、COMPUT で始まる索引語すべてを検索対象とすることができ、COMPUTER、COMPUTING、COMPUTATION などが一度に検索される。これを前方一致指定あるいは語頭一致指定、右側切捨指定などといい、語尾変化を吸収して検索漏れを防ぐためにしばしば用いられる。ただし COM* のような短かい語幹を指定すると、該当する索引語が極端に多くなり、無意味なばかりか検索処理時間も長くなるので、注意すべきである。なお TOOL-IR では、前方一致の有効性に着目して、一般語索引を生成する際に8文字以上の単語はすべて8文字に切詰めて採録するという方式を取り、長い単語についてはいわば前方一致を内蔵したような形にしている。これに合わせて SEARCH コマンドでも8文字以上の長い単語は8文字に切詰めて検索される (上例の ORDINATEUR は ORDINATE として検索される)。複数形についても同様の問題が生じ、例えば「S SYSTEM, SYSTEMS」または「S SYSTEM*」のような指定を常用するのは面倒でありま

た忘れがちでもあるから、4文字以上の単語は索引化の際に語尾のSを除去している。SEARCHコマンドでも同じ処理を施すので、「S SYSTEM」でも「S SYSTEMS」でもSYSTEMを検索することになる。なお、GASES、SOCIETIES、AUTOMATAなどを含めた本格的単複統合化は、処理プログラムの複雑化と同時に副作用の発生が予見され、さ程実用的とは思われないので、上記のような単純な変換に止めている。従ってこれらの事例については単数複数を意識した検索が必要である。

後方一致指定（*単語）と逆索引——接頭辞などの差異を吸収して一括検索を行うために後方一致指定（語尾一致指定、左側切捨指定）がある。「*」記号を語頭に用いて「S *COMPUTER」とすれば、COMPUTERの他、MICROCOMPUTERなども同時に検索される。これは、技術的には、逆索引を用いて実現されている。逆索引は、データベース容量の利用効率を考慮して、8文字以上の長い単語についてだけ作られている。従って、「S *BOAT」とすればLIFEBOATは検索されるが、BOATやTUGBOATは検索されないので注意すべきである。

範囲指定（--）——前方一致指定は語頭の数字を指定してそれ以後は任意という意味であるが、さらに細かい単語の語尾変形を指示するために、範囲指定という方法がある。すなわち「--」（マイナス2個）を用いて「S ABCDN--ABCDP」とすれば辞書順でこの範囲におさまる索引語が検索の対象になる。範囲指定の記号がマイナス1個でなく2個連結になっているのは、索引語自体の中にハイフン（マイナス）が含まれる可能性を考慮したためである。

「.AND.」と「.NOT.」——SEARCHコマンドは「同義語」のOR結合を主体とするコマンドであるが、AND結合あるいはAND NOT結合が1回に限り許される。これにはFORTRANにみられる「.AND.」と「.NOT.」という記号を用い、

SEARCH AUTOMO*, CAR.AND.ENGINE,POWER

のように指定する。「.AND.」の両側の同義語群が別々に検索された後、その両方を含む文献だけで文献集合が作られる。つまり「(AUTOMO* or CAR) and (ENGINE or POWER)」という論理展開である。「.NOT.」についても同様で、

S GENERATOR.NOT.COMPUTER,SOFTWARE

とすれば、プログラム生成プログラムの類が除外されて、発電機関係が効率よく検索されることになろう。

コード類の指定（CODEN、分類コード D、T、C）——CODENをSEARCHコマンドに指定して「S ENREAU」とすれば、Engineering News Record掲載の全論文が掲載される。EIMの場合、書評レコードの検索表示などで会議番号を得れば、会議番号索引を「S C.05727」のように指定して、当該会議の発表論文すべてを検索できる。

資料種別（D）や内容種別（T）は、単独に「S D.JA」（雑誌論文の検索）などとしても、あまり意味がない。これらは他の検索条件と組合せると有効になることが多く、「S COMPUTER.AND.D.TX」ではコンピュータ関連の教科書という意味になるし、「S ROBOTICS.NOT.

T. G」として総説 (General Review) を除外するといった用法が考えられる。

分類コードは、「S A453」のように小分類 (Water Pollution) まで指定して該当文献の検索に用いることもありうるが、「A45*」のような中分類 (Pollution, Sanitary Engineering, Wastes)、さらには「A4*」のような大分類指定は、やはり他の検索条件との組合せで文献の対象分野を限定する際に用いることになる。なお分類コードの一括指定には、上記の「*」による前方一致の他に、範囲指定が有用である。すなわち、「S A401—A404」では A401、A402、A403、A404 が対象になるが、A405以降は対象にならない。このとき「S A40*」では A401から A408までのすべてが対象になる (A409は欠番)。

(2) PHRASE コマンド

PHRASE コマンドは、そのコマンド名のとおり句 (複数単語) を指定して、これを含む文献を検索するコマンドである。すなわち

PHRASE BOX GIRDER BRIDGE

のようにすればよい。この場合単に「BOX GIRDER BRIDGE」という文字を検索するようにみえるが、実際の操作は3つの単語各々を別個に検索し、その結果の重複部分 (AND) を取り出すようになっている。つまり、「SEARCH BOX」、「S GIRDER」、「S BRIDGE」、「AND 1,2,3」という操作と同等であり、指定された索引語間で多重に AND 結合を行った結果の文献集合を作る。SEARCH が OR 論理を主体にしたコマンドであるのに対して、PHRASE は AND 論理のためのコマンドである。SEARCH の場合は、同義語と並べるといった感覚から索引語をコンマで区切るようになっているが、多重 AND 型の PHRASE コマンドでは、まさしくフレーズを指定する感じを生かし、索引語を空白で区切って並べればよい (コンマ等を用いてもよい)。

PHRASE コマンドは内部的には多重 AND 操作を行うので、指定するキーワードの語順は意味を持たない (「PHRASE BRIDGE, BOX GIRDER」などとしても同じ結果になる)。また長い単語については逆索引による語尾一致検索も同時に展開されるので、完全一致が概ね確保される。このような適度に幅がありかつ厳密な指定は、会議名による検索の場合など、不完全な記憶によっても検索が可能であるという点で有用である。例えば、

PHRASE IEEE INTERNATIONAL COMPUTER SOFTWARE CONFERENCE

のように指定すると、IEEE Computer Society's International Computer Software & Applications Conference の書評レコードが検索されてくるといった具合である (図6)。

PHRASE コマンドは一次検索用コマンドとして上記のような幅を持たせた検索を行うようになっているが、その結果、必ずしも「BOX GIRDER BRIDGE」という一連の文字列を含む文献のみが検索されるとは限らない。3語が標題の見出しの中に分散して存在しているようなものがあれば、その文献も検索される。この種の非該当文献が多数混入して困る場合には、後述の LIMIT コマンドによる二次検索の過程でそれらを除去すればよい。

なお、THE、AND 等々、英語の常用単語はオンライン索引抽出の際に除外される (これらをス

図6 PHRASE コマンドの使用例

```

TYPE IN COMMAND
1/ PHRASE IEEE INTERNATIONAL COMPUTER SOFTWARE CONFERENCE
   PHRASE IEEE INTERNATIONAL COMPUTER SOFTWARE CONFERENCE

*      85          IEEE
*     2023        INTERNAT
*     1992        #LANOITAN
*      874        COMPUTER
*     220        SOFTWARE
*     2522        CONFEREN
*     2521        #ECNEREFN
*      82      1/  "IEEE" AND "INTERNAT" AND "#LANOITAN" AND "COMPUTER" A

TYPE IN COMMAND
2/ D.M.B,N.1
   D.M.B,N.1
    
```

```

-----
(      1)
ACC#: 85002165      EIM#: EIM8501-002165
AUTH: Anon
TITL: PROCEEDINGS - IEEE COMPUTER SOCIETY'S EIGHTH INTERNATIONAL COMPUTER
      SOFTWARE & APPLICATIONS CONFERENCE, COMPSAC 84.
CITN: Proceedings - IEEE Computer Society's International Computer
      Software & Applications Conference 8th. Publ by IEEE, New York, NY,
      USA. Available from IEEE Service Cent (Cat n 84CH2096-6),
      Piscataway, NJ, USA 550p      PUBL: 1984      CODN: PSICD2
      ISBN: 0-8186-0550-2      LANG: English
-----
    
```

トップワードと称する)。従ってこれらを SEARCH すると、該当文献 0 件という結果になる。ところで PHRASE コマンドの場合は、フレーズをそのまま指定して有効な検索が行われるよう、コマンド処理中にストップワードの除去手続が内蔵されている。従って「PHRASE FLOW OF SOLIDS」としても OF は検索対象語とならず、結果として該当 0 件になることはない。

(3) AUTHOR コマンド

AUTHOR コマンドは、著者名索引を検索するコマンドである。著者名については、先に述べたとおり、「姓 名の頭字連結」という形式への正規化が施されているから、そのつもりで

AUTHOR JOHNSON JB

のように指定する。AUTHOR コマンドでは前方一致（右側切捨）指定も可能であるので、「AUTHOR ABE*」とすれば ABE、ABEL、ABELE、ABELES 等の各氏が検索の対象になる。なお姓だけわかっている場合には「AUTHOR ABE」とすればよい。この場合機械的に解釈すると、ABE という姓だけで名の部分のない索引を検索することになり、このようなものは通常存在しないので、結果は 0 件になる。従って名の部分は任意という意味で、「AUTHOR ABE *」という指定をすべきである。しかしこれは利用上面倒でかつ忘れ易いと考え、AUTHOR コマンドの内部で空白と「*」を補うようにした。そこで姓だけで検索する場合には、上記のとおり単に姓のみを指定することで、「ABE H」、「ABE T」等々の各氏が検索されることになる（図7）。

(4) ISSN、ISBN コマンド

ISSN（国際標準雑誌番号）と ISBN（国際標準図書番号）は、末尾の桁がいわゆるチェック文字になっており、また電話番号風に中間部にハイフンを置いた形が正規の形式である（ISBN では

図7 AUTHOR コマンドの使用例

```
TYPE IN COMMAND
  2/ AUTHOR ABE
    AUTHOR ABE

*      2      2/  A.ABE (  2  TERMS COMBINED)
```

```
TYPE IN COMMAND
  3/ D.M.B
    D M.B
```

```
-----
(      1)
ACC#: 85002622      EIM#: EIM8501-002622
AUTH: Tanaka, T. / Fujikawa, T. / Abe, T. / Utsono, H.
AFFN: Kobe Steel Ltd, Kobe, Jpn
TITL: METHOD FOR THE ANALYTICAL PREDICTION OF INSERTION LOSS OF A
      TWO-DIMENSIONAL MUFFLER MODEL BASED ON THE TRANSFER MATRIX DERIVED
      FROM THE BOUNDARY ELEMENT METHOD.
CITN: American Society of Mechanical Engineers (Paper) Publ by ASME, New
      York, NY, USA 84-WA/NCA-7, 6p      PUBL: 1984      CODN: ASMSA4
ISSN: 0402-1215      LANG: English
-----
```

```
-----
(      2)
ACC#: 85006090      EIM#: EIM8501-006090
AUTH: Nakamura, T. / Abe, H. / Miyano, I. / Shinozaki, J. / Okino, H.
AFFN: Tokai Univ Sch of Medicine, Dep of Physiology, Isehara, Jpn
TITL: LARGE DEFORMATION ANALYSIS OF THE EJECTING LEFT VENTRICLE.
CITN: Proceedings of the Annual Conference on Engineering in Medicine and
      Biology 37th. Publ by Alliance for Engineering in Medicine &
      Biology, Bethesda, MD, USA p 274      PUBL: 1984      CODN: CEMBAD
ISSN: 0589-1019      LANG: English
-----
```

全体の桁数やハイフンの位置と数は一定ではない)。ところが、実際にはこうした正規の記法によらない表記例がしばしばみられるので、TOOL-IR では、これらの索引化の段階で、ハイフンとチェック文字の削除という「正規化」を行っている。ISSN、ISBN コマンドはこの正規化手順を内蔵した検索コマンドである。これにより、利用者は上述のような ISSN、ISBN の本来形を気にすることなく、それら番号を引用文献、目録等で見つけたままの体裁で指定して、正しい検索結果を得ることができる。つまり、「ISBN 0-8186-0548-0」でも「ISBN 081860548-0」でも「ISBN 0818605480」でもよい。ISBN は「0141-061X」のような「4 桁-4 桁」の固定形式なので表記の変動は少ないが、それでも中間のハイフンが省略されることもある。ISSN コマンドでは「ISSN 0141061X」としても正しく検索される。指定された末尾の文字がチェック文字としてまちがっている場合には、ISSN コマンド、ISBN コマンドは、正しいチェック文字を表示した後、末尾の文字を訂正して検索を行う。

(5) AND、OR、DIF コマンド

TOOL-IR では SEARCH、PHRASE 等のコマンドの実行結果として検索条件に適合した文献の群が作成される。この群を文献集合と称し、作成された順に 1 番からの連続番号が振られる。そしてこの番号を指定して、特定の文献集合を後から呼出すことができる。AND、OR、DIF の各コマンドは、この機能を用いて既存の集合間の組合せ処理（集合演算）を行い、その結果を新たな文献集合として生成するものである。例えば

```
AND 1, 3, 4
```

は、既に作られている1番、3番、4番の各文献集合のすべてに共通に含まれる文献だけを探し出すという指示であって、こうした文献だけを集めた新たな文献集合が作られる。この集合には、最近に作られた集合の番号の次の番号が振られる。同様に、

OR 4, 10, 7, 2

は、4、10、7、2番の集合のいずれかに含まれている文献を集約して新しい集合を作る。DIFは集合の差をとるコマンドで

DIF 4-3

は、4番に含まれ3番には含まれない文献だけの集合を生成する。

(6) DISPLAY コマンド

DISPLAY コマンドは、SEARCH、PHRASE、AND等のコマンドによって作成された文献集合中の、各文献の内容を表示させるコマンドである。従って、何番の文献集合中の、何番目の文献の、どの項目を表示させるかといった指定が必要になる。例えば

DISPLAY S. 5, M. B, N. 2

というのは、5番の集合に含まれる文献の最初の2件について、その内容を「Bモード」(書誌的項目の表示)で表示せよという指示である。もっとも一々こういう指示をするのは面倒なので、各指定項目に次のような標準値を設定して、省略を可としている。

集合番号 (Set) S.s——最新の集合 (番号が最大のもの)

表示文献数 (Number) N.n——集合中の全文献

表示項目 (Mode) M.m——Xモード (論文標題のみ、検索過程点検用)

DISPLAY は DIS や D という略記が可能であるから、すべてのオペランドを省略すると、DISPLAY の最も簡略な指定は「D」1文字になる。SEARCHなどで今検索したばかりの集合の内容を一瞥するには、「D」でよい。この場合、表示文献数(N.)は全件が指定されたことになるから表示がいつまでも続くようになるが、内容を点検しつつ頃合を見計って割込キーを押せば、DISPLAYは打切られる。「D N. 10」のように表示件数を予め限定するのもよいが、表示内容を見ながら適当なところで止める方が何かと有利であることが多い。

内容を詳細にみるには、「D M.D」のように、表示項目だけ指定すればよい。表示モード(M.)にはA、B、C、D、Xの5種があり、AからDの順で詳細な表示になる(表4、図1、図7、図8)。M.Dではデータベース中の全項目が出力される。M.Xは標題のみを表示するモードで、出力料金(ヒット・チャージ)が適用されず、検索条件の適否点検用に用いられる。

ところで、M.Aなどで簡略表示させた文献の中から、特定のものだけを再度詳しく出力させたことがある。このために、何番目の文献という指定(I.)があり、

D I. 8, M. D

D I. 1, 3, 5, M. D

D I. 3-7, M. D

表4 データ項目, 索引項目, 表示項目一覧

"Ei ENGINEERING MEETINGS"			
DESCRIPTION	HEADER	INDEX	DISPLAY MODE
Accession Number	ACC#:		X A B C D
Publication Number	EIM#:		X A B C D
Author(s)	AUTH:	A.	- A B C D
Author Affiliation (First Author)	AFFN:		- A B C D
Original Title	TITL:	**	X A B C D
Translated Title	TTIT:	**	- A B C D
Citation (PIE)	CITN:		- - B C D
Publication Year	PUBL:		- - B C D
CODEN	CODN:	**	- - B C D
ISSN	ISSN:	S.	- - B C D
ISBN	ISBN:	B.	- - B C D
Language	LANG:		- - B C D
Conference Title	CONF:	**	- - - C D
Supplementary Conference Information	CNFX:	**	- - - C D
Sponsor	SPSR:		- - - C D
Conference Location	LOCN:		- - - C D
Date Conference Held	DATE:		- - - C D
Ei Conference Code Number	CNFN:	C.	- - - C D
Abstract	TEXT:		- - - - D
Main Subject Heading (SHE)	MSBJ:	**	- - - C D
Subheadings (SHE)	SSBJ:	**	- - - C D
Cross Referance Terms (SHE)	XREF:	**	- - - C D
Free Language Terms	FIDX:	**	- - - C D
CARD-A-LERT Classification Codes (SHG)	CALC:	**	- - - C D
"COMPENDEX"			
DESCRIPTION	HEADER	INDEX	DISPLAY MODE
Accession Number	ACC#:		X A B C D
Internally Assigned Accession Number	EIX#:		X A B C D
Monthly Publication Number	ABS#:		- A B C D
Author(s)	AUTH:	A.	- A B C D
Author Affiliation (First Author)	AFFN:		- A B C D
Title	TITL:	**	X A B C D
Citation (PIE)	CITN:		- - B C D
CODEN	CODN:	**	- - B C D
ISSN	ISSN:	S.	- - B C D
ISBN	ISBN:	B.	- - B C D
Language	LANG:		- - B C D
Document Type	DTYP:	D.	- - - C D
Treatment Designator	TRMT:	T.	- - - C D
Abstract	TEXT:		- - - - D
Main Subject Heading (SHE)	MSBJ:	**	- - - C D
Subheadings (SHE)	SSBJ:	**	- - - C D
Cross Referance Terms (SHE)	XREF:	**	- - - C D
Free Language Terms	FIDX:	**	- - - C D
CARD-A-LERT Classification Codes (SHG)	CALC:	**	- - - C D

とすれば、最新の集合中の、8番目の文献、1、3、5番目の文献、3番目から7番目までの文献の全項目が表示される。何番目の文献かという集合内での文献の順序番号は、出力の左肩の括弧内に示されている。

また DISPLAY コマンドでは、アクセシオン番号を直接指示して、その文献の内容を表示させることもできる。すなわち「A.a」（aはアクセシオン番号）という指定を行い、

D A. 85000344, M. D

のようにする。アクセシオン番号指定では、文献集合とは全く無関係に、特定の文献を選択・表示できる（図3）。文献集合をあらかじめ作っておく必要はない。

(7) LIMIT コマンド

LIMIT コマンドは、文献集合中の各文献について、指定された項目の字面を文字合わせ的に照合して指定の文字列の有無を判別し、条件に適合した文献だけで新たな集合を生成するコマンドである。つまり、SEARCH や PHRASE などの一次検索で一旦集合を作った上で、その中の文献1件ずつの内容を走査するためのコマンドである。

検索用のキーワード群がある程度特異なもの、多義的でないもの場合は、それによる一次検索

図8 LIMIT コマンドの使用例

```

TYPE IN COMMAND
  3/ PHRASE AIR CONDITIONING
    PHRASE AIR CONDITIONING

*      211          AIR
*      213          CONDITIO
*         55        #GNINOITI
*         32      3/ "AIR" AND "CONDITIO" AND "#GNINOITI" END NOSAVE

TYPE IN COMMAND
  4/ LIMIT MSBJ:AIR CONDITIONING
    LIMIT MSBJ:AIR CONDITIONING

    SCAN MSBJ *"AIR CONDITIONING"* END NOSAVE

*         8      4/ SCAN 3  MSBJ *"AIR CONDITIONING"*

TYPE IN COMMAND
  5/ DIS M.C.N.1
    DIS M.C,N.1
-----
(      1)
ACC#: 85000553   EIM#: EIM8501-000553
AUTH: Sheridan, Norman R.
AFN: Univ of Queensland, Brisbane, Aust
TITL: SOLAR AIR CONDITIONING.
CITN: Publ by D. Reidel Publishing Co, Dordrecht, Neth and Boston, MA, USA
      p 57-89   PUBL: 1983   ISBN: 90-277-1506-8   LANG: English
CONF: Solar Energy Applications in the Tropics, Proceedings of a Regional
      Seminar and Workshop on the Utilization of Solar Energy in Hot Humid
      Urban Development.   SPSR: Natl Univ of Singapore, Singapore,
      Singapore   LOCN: Singapore, Singapore   DATE: 1980 Oct 30-Nov 1
      CNFN: 04610
MSBJ: AIR CONDITIONING   SSBJ: Solar Energy Systems
XREF: ECONOMICS / HEATING - Solar
FIDX: SPACE HEATING SYSTEM / PASSIVE SOLAR AIR CONDITIONERS / ACTIVE SOLAR
      AIR CONDITIONERS / PREDICTED PERCENTAGE DISSATISFIED / ROOF POOL /
      SOLAR FRACTION
-----

```

だけで所期の文献に絞り込まれた結果が得られるので、LIMIT コマンドを使わなくて済む。LIMIT コマンドによる走査が必要なのは、結局、一般的な単語を数語連ねたフレーズが特有の意味をもつような場合で、これはコンピュータ関連を中心とする新しい分野に多い。

SHE には「AIR CONDITIONING」という主見出しが登録されている。そこで

PHRASE AIR CONDITIONING

として検索すれば、標題や見出しのどこかに AIR と CONDITIONING が書かれている文献すべてが検索される。こうした網羅性がオンライン検索の特徴ではあるが、上例のような場合、2つの索引語ともそれ程特異性が強くないので、大きな結果集合が作られ、そこでは不要文献ばかりが目立つということになり兼ねない。Ei 社が見出しを「AIR CONDITIONING」にした文献だけに絞り込むには、上記の PHRASE コマンドにつづけて、

LIMIT MSBJ: AIR CONDITIONING

というコマンドを入力すればよい。これは、最新の文献集合中の文献から、MSBJ 項目に「AIR CONDITIONING」という文字列が含まれている文献だけを選別せよという指示である(図8)。

LIMIT コマンドは、一次検索の際とは異って、オンライン索引を見るのではなく、データベースの本体部分を直接走査点検してゆく。従って、先にふれた8文字での語尾切捨とか複数形Sの除去などの内部処理は介入しない。飽くまでも字面に忠実な「機械的」検索が行われるから、指定上注意を要すると同時に、表記の微妙な差を生かした選別が可能になっている。また走査という処理はかなり計算機時間を消費するものであるから、対象とする文献集合が100件以上の文献を含むような場合には、コマンド投入後相当待たされる。SEARCH とか、AND、DIF 等を用いて数10件程度の大きさの集合に絞り込んでから、それに対して LIMIT を適用してゆくのがよいと思われる。

LIMIT コマンドの一般的指定形式は表3にみるとおりで、対象とする文献集合番号、走査対象の項目名、走査すべき文字列の3者を指定する(文献集合番号は後述のとおり標準値による省略可)。

走査項目の指定——どの項目を走査・照合の対象にするかを、DISPLAY の時に表示される項目別の見出し文字(「TITL:」、「TEXT:」など)を用いて指定する。これで個々の項目別の走査が可能であるが、一方、全項目に亘って特定の文字列の有無を点検したい場合もある。この目的のために「ALL:」という項目指定を設けている。ただし「ALL:」指定の際本当に全項目を当たるのは処理時間上無駄が多いので、いわゆるキーワードが含まれる項目群、すなわち標題、要旨、主見出し、副見出し、相互参照、自由語索引を対象にする。そこで

LIMIT ALL: AIR CONDITIONING

と指定すれば、最新の文献集合中の各文献について上記の項目を点検してゆき、どこかに「AIR CONDITIONING」とあれば、その文献は合格と判定される。

走査文字列の指定——走査したい文字列をそのままの形で指定する。その際、空白、コンマの類も有効であるから、

LIMIT MSBJ: CONTROL SYSTEMS, NUMERICAL

のような逆転句の指定も可能である。このような一連の文字列だけでなく、2つの文字列を同時に指定することもできる。2つの文字列を「*」をはさんで記せばよく、

LIMIT TITL:HIGH POWER*LASER

とすれば、「HIGH POWER LASER」に加えて「HIGH POWER GAS LASER」なども適合とされる。つまり「*」の部分には任意の文字列があって構わないという意味の走査である。数語からなる専門用語の場合、その中間に各種の形容詞類が割込むことが多いが、こうした用例の走査には有効であろう。

ところで、

LIMIT ALL:AIR CONDITION

という指定は、このような文字列を「含む」という意味合いであって、「AIR CONDITIONING」、「AIR CONDITIONER」なども合格になる。つまり前方一致と後方一致の指定が含意され「*AIR CONDITION*」という意味になっている。そこで単語の区切りを特に指定したい場合は、文字列の前または後に「*」と空白を記して、これを明示する必要がある。

LIMIT ALL:* AIR CONDITION *

とすれば、CONDISIONER の類は除外される。文字列を2つ指定する場合も、

LIMIT FIDX:* MARINE * ENGINE *

では SUBMARINES や ENGINEERING は対象外になる。

文献集合番号の指定——LIMIT コマンドで対象とする文献集合番号は、

LIMIT 5,MSBJ:LASERS

のように、走査項目指定の前に記す。ただしこれまでの例でわかるとおり、最新の集合番号が標準値となっているので、直前に生成された集合を対象にする場合には、集合番号を指定する必要はない。

(8) LOOK コマンド

オンライン検索では、索引語となるキーワードの類を利用者が次々と想起しながら検索を進めることになる。しかし、オンライン索引自体を通覧し、実際にどのようなキーワードが登録されているかを点検することで、当面の検索にあたっての適切なキーワードが発見されることも多い。LOOK コマンドはこのためのもので、指定された単語の前後16語程の範囲についてオンライン索引を表示する。ちょうど辞書の一部を切取ったような形で、各行に索引語とその該当文献数が示される。これにより、指定した語の前後にどのような類似のキーワードがあり、各々の該当文献数（使用頻度）がどの程度であるかが判明するから、検索上有益である。LOOK コマンドは、指定した語そのものが索引に登録されていない場合にも、辞書式配列でその語のあるべき位置の前後の索引語が表示される点で便利である。つまり LOOK コマンドでの指定は、索引語を特定するといよりも、索引上の表示すべき位置を指示すると考えればよい。

一般語索引を通覧するには、単に

LOOK COMPUT

のように、キーワードを指定すればよい。逆引の指定は、「LOOK *OXIDE」のように、先頭に「*」を付す。資料種別索引を通覧には、索引化の際に付加した「D.」を含めて、「LOOK D. JA」のように指定する（この場合「LOOK D.」、「LOOK D.*」などでもよい。「D.*」は後述の語頭一致索引表示の一例）。内容種別についても「LOOK T. A」とすればよいが、これらは元来コードであるから、この種の LOOK はコード表の確認という意味合いになる。著者名索引を通覧の際は、名前の前に「A.」を付加して、

LOOK A. ABE

のように指定する（図9）。この他、「S.」（ISSN）、「B.」（ISBN）、「C.」（会議番号）も可能である。

LOOK コマンドによる索引語の表示終了後には、「ENTER MORE TO CONTINUE TERM LIST, ENTER LETTERS TO BE COMBINED SEPARATED BY COMMAS OR DASHES」なるメッセージが出てから、次のコマンド受付状態になる。ここで索引のその先を続けて見たければ、メッセージにいうとおり「MORE」と入力する。これでさらに16件の索引語が表示される。「MORE」を繰返せば索引を16件ずつに区切って見てゆくことができる。

表示索引数の変更（SET コマンド）——LOOK による索引表示の件数（行数）は16件に設定してあるが、これは4～26件の範囲で変更可能である。すなわち

SET ADJ 6

のように SET コマンドを入力すると、それ以後、LOOK による表示は6件ずつ行われるようになる。少しずつあちこち LOOK するには、予め「SET ADJ 4」などで1回の表示件数を少なくしておくといし、上述の「MORE」の繰返しによって通覧してゆく場合には、「SET ADJ 26」としておくとも率がよい。

語頭一致索引表示と後続索引表示——「LOOK COMPUT」は COMPUT の前後にある索引語を表示するが、末尾に「*」を付して「LOOK COMPUT*」とすると COMPUT で始まる索引語が、また「**」を付して「LOOK COMPUT**」とすると COMPUT よりうしろにある索引語が表示される。これらを語頭一致索引表示、後続索引表示という。前者の場合は指定された文字列で始まる索引語がなければ LOOK コマンドは空振りになるが、後者あるいは通常の LOOK コマンドは指定された索引語よりうしろあるいはその前後の索引語を表示するので、空振りということはない。語頭一致索引表示、後続索引表示は一般語索引以外に対しても可能であり、その場合は「A.」などを付して、「LOOK A. ABE*」、「LOOK A. ABE**」のように指定する。ただし逆引索引（*kw）に対しては指定できない。標準の表示索引数は16件に設定してあるが、これを変更するには「SET TERMS n」コマンドを入力する。

LOOK 後の検索——LOOK によって適当なキーワードが見付かった場合、その該当文献を検索

図9 LOOK コマンドの使用例

```

TYPE IN COMMAND
  1/ LOOK A.OZA
    LOOK A.OZA

. ITEMS.      TERMS
A           1  A.OWEN KM
B           2  A.OWEN RG
C           1  A.OWENS A
D           1  A.OWENS EC
E           1  A.OWENS FS
F           1  A.OWENS JM
***** YOUR TERM *****
G           1  A.OZAKI K
H           1  A.OZAWA H
I           1  A.OZAWA Y
J           1  A.OZDAMAR O
K           1  A.OZEKI T
L           2  A.OZISIK MN

ENTER MORE TO CONTINUE TERM LIST
ENTER LETTERS TO BE COMBINED SEPARATED BY COMMAS OR DASHES
  1/ H
    1 ITEMS  SAVED AS SET 1

SELECT TERMS OR TYPE IN COMMAND
  2/ DIS M.D
    DIS M.D

-----
( 1)
ACC#: 85002721      EIM#: EIM8501-002721
AUTH: Negishi, Masamitsu / Ozawa, Hiroshi
AFFN: Univ of Tokyo, Research Cent for Library & Information Science,
      Tokyo, Jpn
TITL: FURTHER DEVELOPMENT ON UNION LIST OF SCIENTIFIC PERIODICALS (ULP)
      DATABASE: UPDATING OF DATABASE TO 1982 VERSION, NATIONWIDE ON-LINE
      RETRIEVAL SERVICE FOR JAPANESE ACADEMIC COMMUNITY, AND COMPILATION
      OF DATABASE FOR PERIODICALS IN JAPANESE.
CITN: Publ by North-Holland Publ Co (FID Publ 662), Amsterdam, Neth and
      New York, NY, USA p 171-179      PUBL: 1983      ISBN: 0-444-86646-9
LANG: English
CONF: Challenge of Information Technology, Proceedings of the Forty-First
      FID Congress.      SPSR: FID, The Hague, Neth      LOCN: Hong Kong
DATE: 1982 Sep 13-16      CNFN: 04929
TEXT: The project to develop the Union List of Scientific Periodicals
      (ULP) database has been carried out at Research Center for Library
      and Information Science (RCLIS), The University of Tokyo since 1977.
      Five subprojects involved are: 1. Generation of 'Title Change Map'
      based on the ULP database of humanities and social sciences, 2.
      Conversion of the structure of ULP file for periodicals on science
      and technology to the database schema adopted in the ULP database
      for humanities and social sciences, and consolidation of the two
      databases, 3. Revision of the ULP database by collecting the
      updating data of holdings which reflect new acquisitions after 1978,
      when the data for the previous version were collected, 4.
      Development of an on-line retrieval system for the ULP database, 5.
      Preparation of bibliographic check list of scientific periodicals in
      Japanese aided by computer facility of Japanese text processing,
      which is the first step for compiling the ULP database for Japanese
      periodicals. The present paper describes the results and the current
      status of these subprojects, and a future configuration of the
      scientific information system for Japanese academic community
      referring to the master plan announced by the Ministry of Education.
      7 refs.
MSBJ: INFORMATION RETRIEVAL SYSTEMS      SSBJ: Online Searching
XREF: DATABASE SYSTEMS
FIDX: SCIENTIFIC PERIODICALS / JAPANESE LITERATURE / BIBLIOGRAPHIC
      CHECKLIST
CALC: 00-A901 / 00-A723

-----

TYPE IN COMMAND
  2/

```


することになる。これには SEARCH コマンドにそのキーワードを指定して検索してもよいが、LOOK の直後ではもっと簡便な検索法が用意されている。

16件の索引表示各行の左端には、A から P までのアルファベット（行記号）が表示される。「ENTER LETTERS……」というメッセージの意味は、このアルファベットを指定すれば、その行記号に対応するレコードを実際に検索して、結果集合を作ることができるということである。そこで、「A」とか「H」とかの行記号を打鍵すればよい（図9）。このとき「F, I, M」のようにいくつか選んでもよいし、「D-F」のようにしてD行からF行までの範囲を指定してもよい。その結果、「4 ITEMS SAVED AS SET 6」、「SELECT TERMS OR TEPE IN COMMAND」といった表示があって、次の入力受付となる。この検索では行記号の間で OR 結合が行われるから、作られる結果集合中のレコード数（ITEMS）は、索引表示の際に示された各行記号別のレコード数の合計より少ない値になることがある。

もちろんこうした行記号指示でなく、新たに LOOK や PHRASE など別のコマンドを入力してもよい。ここで注意すべきは、LOOK 後の DISPLAY コマンドの入力である。DISPLAY の最も省略した指定である「D」だけのコマンド入力は、LOOK 直後の場合、行記号の D（4行目の索引語）という方に解釈されて、これに対応する集合が作られる。LOOK 後の行記号入力可能状態は上述の「ENTER LETTERS ……」または「SELECT TERMS ……」のメッセージによって判別でき、この状態は新たなコマンド入力があるまで解除されない。つまり「A」とか「C」とかを次々に入力して、1回の LOOK の後に複数の検索結果集合を作ることができる反面、DISPLAY コマンドの入力については少々注意を要するわけである。この場合、「DIS」あるいは「DISPLAY」とするか、「D M.D」のようにオペランドを付けて指示してやれば、行記号と混同されずに DISPLAY が行われ、同時に行記号入力可能状態も解除される。なお「/」を入力しても行記号入力可能状態は解除される。

(9) REMIND コマンド

検索過程が長くなってくると、何番の集合の内容がどんなものであったか、不明になりがちである。特にディスプレイ型端末では、用紙をたぐって検索過程を調べるわけにゆかない。こうした場合に備えて、集合番号、該当文献数、検索キーワードがシステムに保存・記憶されており、これを REMIND コマンドで表示させることができる。指定法は表3のとおりで、これにより検索過程を再確認できる。

(10) ?INFO コマンド

表4に掲げたデータ項目、索引項目、表示項目等の一覧表は、「?INFO」によって、オンライン検索の途中で出力可能である。

(11) END コマンド

END コマンドを入力することにより検索は終了し、システムに記憶されている集合番号、該当文献数などの情報は解放される。

7. 運用方式

COMPENDEX、Ei ENGINEERING MEETINGS 両データベースは、東京大学大型計算センターと東京大学文献情報センター（現学術情報センター）の共同運用の形で、大型計算機センターの TOOL-IR システムにおけるセンター提供データベースとして公開されている。従って、大型計算機センターの利用者であれば、特段の手續等を要することなく、COMPENDEX については「>>ENG」、EIM については「>>EIM」により検索を開始することができる。大学間コンピュータ・ネットワークを介すれば、他の計算機センターからも利用可能である。なお、この検索システムを動かすには、TSS の標準主記憶容量 1MB では不十分なので、ログオンの際「LOGON 利用者番号, S (1500)」のように、Size オペランドに 1.5MB 以上を指定する。

両データベースとも月刊であるが、その磁気テープの当方への到着日程はやや不安定である。実績をみると概ね 2 カ月程の期間を要するようで、つまり 1 月号が 3 月末から 4 月初めにかけて到着するというような状況である。大型計算機センターでは到着の都度オンライン・データベース化しているが、この作業期間を見込めば、1 月号が 4 月中頃から末にかけて公開されるといった日程になる。なお、データベースの累積量が大きくなると作業上の効率が落ち、また検索利用上も年次による選別の手間が必要になるというような点を考慮して、オンライン・データベースは、1 年あるいは半年単位で累積・完結させてゆくという方式をとっている。どの範囲を検索対象とするかは、検索開始時の「>>ENG」、「>>EIM」各コマンドのオペランドで指定する（下例）。

COMPENDEX	EI ENGINEERING MEETINGS	
>>ENG	>>EIM	……最新号
>>ENG 1986	>>EIM 1986	……1986年分
>>ENG 1985	>>EIM 1985	……1985年分
>>ENG 1984B		……1984年後半分
>>ENG 1984A		……1984年前半分
(以下同様)	>>EIM 1984	……1984年分

このような累積状況は、「>>ENG」等による検索開始時に常に表示されるので、これで現況を確認する。なお、両データベースとも文献内容の表示は英小文字まじりで行われる。特に英小文字を大文字に変換して表示したい場合には、「>>ENG 1985, CAPS」、「>>ENG , CAPS」のように、オペランドで CAPS を指定する。

付記：両データベースは 1987 年 4 月より、学術情報センターにて運用されることとなり、これに伴ない、上記運用方式も若干改定される。

8. 参考資料

COMPENDEX と Ei ENGINEERING MEETINGS に関しては、Ei 社から次のような参考資料が出版されている。

- ① SHE, Subject Headings for Engineering, 1983, 167p, ISBN 0-911820-25-6, \$35.
- ② SHG, Subject Heading Guide to Engineering Categories, 1984, 168p,
ISBN 0-911820-27-2, \$30.
- ③ PIE, Publications Indexed for Engineering, 1985, 156p, ISSN 0085-4581, \$30 (年刊).
- ④ CAL Classification Codes, 1982, 32p.
- ⑤ COMPENDEX Search Manual, 1983, 400p, ISBN 0-911820-14-0, \$75.

① SHE は、既に述べたように、主見出し、副見出しをアルファベット順排列し、適用範囲等の説明を加えた、Ei の統制語彙の基本資料である。② SHE は、「主見出し—副見出し」を分類コード (CALC) によって排列したものである。分類コード表とその説明は④にある。③ PIE は Ei の収録雑誌・会議録のリストで、年間の収録分をまとめて編集し、翌年の6月頃刊行している。⑤は COMPENDEX データベースの検索法を各社のオンライン・システム別に説明したマニュアルで、DIALOG、SDC、BRS 等8種のシステムが含まれている。なお COMPENDEX 自体の構成についても概略的な説明がなされている他、分類コード表、略語表、副見出しリストなどが付録に収められている。

研 究 論 文

誤り訂正符号の調査票記入への応用

Application of Error Correcting Code to the Questionnaire Data Entry

学術情報センター 橋爪 宏達*

学術情報センター 宮澤 彰**

要 旨

調査票のようなオフライン形式のデータ・エントリに適用可能な誤り訂正符号系について述べる。またこの誤り訂正符号を、学術雑誌総合目録の調査に適用した例を紹介する。この調査で収集した75万件のデータにつき、誤りパターンを解析する。更にこのデータから人間の誤り特性につき考察する。

Abstract

This paper introduces the error correcting code system applicable to the offline data entries such as questionnaires. The application of this error correcting code to the investigation of the Union List of Periodicals is also mentioned as a practical example. The error patterns, extracted from about 750,000 collected records of the investigation, are analyzed and yield the human's error characteristics.

1. はじめに

最近では、日常生活の中でいわゆる物品をコード番号（数字ないしアルファベットの列）を記入する機会が多くなった。たとえば通信販売の商品コード、銀行口座番号、社員コードなどである。しかしこれらコード番号は一般に人間にとっては無意味な文字の羅列であることから、絶えず誤記入や誤読の可能性が付きまとっている。またコードを光学文字読み取り機（OCR）のような装置に入力する場合は、読み取り過程でもいくつかの誤読が発生しうる。以上のような人間ないし機械の誤りを検出するために、通常はコードを構成する文字列中に、さらに確認のための文字を付加するのが普通である。これはチェック・ディジットなどと呼ばれている。

チェック・ディジットは元来、符号理論で体系的に扱われるべきものである。しかし従来のチェッ

*Hashizume, Hiromichi : National Center for Science Information System

**Miyazawa, Akira : National Center for Science Information System

ク・ディジットの構成法は、単純に各桁の数字を合計して、10の剰余をとり、その数値をチェック・ディジットする式の便宜的なものが多かった。従って誤記入・誤読があった場合の誤り検出能力などについて、必ずしも深い考慮がなされていなかったのが実情である。その上、それでは人間—機械系の誤り特性に基づき、誤りに対して抵抗力のあるコード系を作ろうとしても、コード記入の際に人間ないし機械がどのような誤りを生起するかについて、特に資料がなかったのが現状である。

一方、学術情報センター（1986年3月までは東京大学文献情報センター）では1985年10月に、学術雑誌総合目録（以下、学総目と略記する）事業の一環として、大学図書館で保有している雑誌を調査するため、全国の大学図書館に調査票を発送した。図書館側では各館が購読しているすべての雑誌について、どういうタイトルの何巻・何号を持っているかを OCR 用紙の調査票に記入し、返送する。（一部図書館では磁気テープで回答した。）この記入の際に、どの雑誌について記入しているかを明らかにするため、回答用紙に「確認コード」という10桁程度のコードを記入してもらった。全国の図書館で記入される確認コードの総数（レコード総数）は100万件内外に達すると見込まれた。ここに付加するチェック・ディジットの構成法については、総数が膨大であることから、方式について事前に考察し、可能な限り高い能力を付与しておくべきだと判断された。そこで採用したのが、本論文で述べるリード・ソロモン符号¹⁾（Reed-Solomon Code、以下 RS 符号と書く）による2重誤り訂正符号系である。RS 符号そのものは既知の技術であるが、これを調査表票記入時のチェック・ディジットとすることは誤り訂正符号の応用として新しい分野と思われる。

調査票発送から1年程が経過した現在、約74万件のレコードを回収している。そのうち9割強が OCR 用紙による回答で、残り一割が磁気テープの回答であった。本論文では RS 符号による誤り訂正符号を紹介するとともに、この回答データの解析も試み、採用した符号系が所期の目的を果たしたことを確認する。また、今後同種のコード設計をする際の基礎データとするために、人間—機械系の誤り特性の観点からの考察も試みる。

2. 符号の選択

自動誤り検出／訂正符号に関しては多数の構成理論が知られており、応用上の要求に応じて選択できる。特にハミングの多重パリティによる誤り訂正符号²⁾に端を発する線形誤り検出／訂正符号は、その後ブランジ³⁾により巡回符号として一般化され、ポーズ、チョードーリ、オッケンジェム⁴⁾らにより BCH 符号として体系化されている。この BCH 符号ないしそこから派生した符号は、データ通信や情報蓄積装置の構築に不可欠な誤り検出／訂正符号系として広く利用されている。

今回の調査票の確認コードについては、まず、誤り検出符号として構成するか誤り訂正符号として構成するかを決定する必要がある。次の2点の考察から誤り訂正符号とした。第一点は、事前の調査から、OCR 装置はかなりの頻度で読み誤りを起こすことが予想されたことである。この誤りを事後に救済するためには、誤り訂正符号である必要がある。第二点は、誤り訂正符号で訂正されたデータと元のデータとを比較することで、人間ないし OCR の犯す誤りについて資料が得られることであ

る。これは今後同種の調査をするときの参考として貴重である。

調査票のコード記入に応用されるべき誤り訂正符号は、データ通信や情報記録で求められるもの比べて、次に次のような状況に対処できることが求められる。

1) 符号長が短い

主に手書きで OCR 用紙の欄に記入するわけだから、符号長（情報点+検査点）の桁数は数桁～十数桁の応用が一般的である。通常の誤り検出/訂正符号の応用では数百～数千ビット以上の情報点を扱うのに比べて、OCR 用紙のものは著しく少ない。

2) 多元情報を扱う

現在一般に使用されている手書き OCR は数字、アルファベットなど数十種類の文字を扱える。従って OCR 文字に適用する符号系としては、通常の 2 元情報（0 と 1）のみを対象とした 2 元符号ではなく、数十種類の符号アルファベットを対象とする多元符号の枠で考えねばならない。

3) 2 重誤りが発生しうる

欧米における主にタイプライター打鍵時の誤りの解析から、人間の犯す誤りは下記の 4 種類の基本モードが複合さともものと考えられる⁵⁾。

- | | |
|---------|-----------------|
| i) 単純誤り | 例:simple error |
| ii) 転置 | 例:trasnposition |
| iii) 冗字 | 例:extora letter |
| iv) 欠字 | 例:mising letter |

特に固定長の調査票欄にコードを記入する応用に限ると、iii) と iv) のモードはなくなり、目的とする符号は i) と ii) の誤りモードに対処できることが必要となる。復号時には i) は長さ 1 の（ランダム）誤りとして、ii) は長さ 2 のバースト誤りとして出現する。

4) 符号化、復号化に時間を要してもかまわない

通常の応用では符号化ないし復号化を実時間で行なわねばならないことが多く、採用しうる符号は計算コストの制約を強く受ける。これに対して今回の応用では、事前にコードを符号化しておき、事後にある程度時間をかけて復号化してかまわないから、計算コストからの制約は小さい。

以上の条件から、今回は RS 符号を採用した。RS 符号は巡回符号ないし BCH 符号の理論体系から導出されるもので、これを特に多元符号として整備したものと位置付けられる。誤り訂正のために付加する検査点数について、同じ訂正能力を持つ線形符号中で最小である特徴がある。1) の要求から、記入時に情報点に付加する検査点をなるべく少なく留める必要があり、最少検査点数を保証する RS 符号の性質は、目的に合致するものである。また 2)、3) の要求から、多元（20～30 元程度）の 2 重誤り訂正 RS 符号とすべきことがわかる。2 重誤り訂正 RS 符号は、2 個までのランダム誤りを全て訂正する。RS 符号以外にバースト誤り訂正符号の採用も考えられるが、2 重誤り訂正程度ではあまり検査点数の節約にならず、より能力の高いランダム誤り訂正能力をもつ RS 符号を選んだ。

2重誤り訂正 RS 符号の最小符号間距離は 5 であって、3重以上の誤りは別の符号語に誤訂正されてしまうことがある。しかし次章に述べる符号の短縮により、誤訂正の可能性はある程度小さくなっている。

従来より RS 符号は半導体記憶装置、磁気記憶装置、光ディスク、PCM オーディオ記録などに実績の多い符号ではあるが、復号コストが比較的大きいのが欠点であった。しかし今回は 4) の特性から、特に実時間性の配慮の必要がなく、復号コストは問題にならない。

3. RS 符号の構成

多元 (q 元) RS 符号は q 元のガロア数体 $GF(q)$ 上に構成される。ここで q は符号語 (コード) に使用する文字の種類である。また RS 符号の性質から、 $q-1$ が構成できる符号語の最大長でもある。 q は上述の両方の条件から決定されることになる。一般に文字の種類としての q を制限すると、情報点の桁数 k 以内で識別しうる符号語の種類 q^k が少なくなってしまう。この意味で、 q はなるべく大きく取るほうが有利である。さらに、比較的大きな q を使って短い符号語を構成することは、符号の短縮操作に相当する。短縮された符号は、多重誤りの検出に有効である。なぜなら RS 符号の訂正能力を越えた多重誤りは、誤った符号語に誤訂正されるが、その際に訂正が短縮して使用している桁の外側で起これば、これが誤訂正であることを判別できる。つまり符号の短縮は実効的な符号間距離を大きくする作用がある。

以上のような理由から、 q は人間が混乱しない範囲で、また OCR の読み取れる範囲で、なるべく大きく取る方がよい。数字が 10 種類、英大文字が 26 種類であり、それらの読取りが現在の OCR の基本機能であることを考慮すると、 q については $10 \leq q \leq 40$ 程度の応用が多いと思われる。ただし q はこの範囲で自由に選べるわけではなく、ガロアの数体理論から更に次の制約を受ける。すなわち

p を任意の素数、 m を任意の自然数として、 q は

$$q = p^m$$

と書けなければならない。

従って前記の範囲では使用可能な q は次のようなものである。

$$11, 13, 16 (=2^4), 17, 19, 23, 25 (=5^2), \\ 27 (=3^3), 29, 31, 32 (=2^5), 37, 39$$

いずれの q からも数体 $GF(q)$ が構成でき、(つまり q 種類の数字とこれらを使った加減乗除の四則演算が定義され)、その上で q 元 RS 符号が作成できる。

今回の学総目の調査票では、 $q = 31$ にとり、31 元の 2 重誤り訂正 RS 符号を構成した。この場合

最大符号長は30になるが、実符号長を9に短縮して使用する。うち情報点は5、検査点は4である。実際のデータ記入に使用する文字はOCRが読み取れる文字種の中から次の31文字を選んだ。

0123456789ABCEFGHJKLMNPQRSTUVWXYZ

D, I, O, Zなどのアルファベットは、それぞれP, 1, 0, 2などと紛れることを恐れ、使用を避けた。これらの文字を、

$GF(31)$ の元

$$0, a^0, a^1, \dots, a^{29}$$

に対応させる。対応順序は任意である。

RS符号の符号化／復号化アルゴリズムは本論文の直接の主題ではないので省略する。必要があれば参考文献^{6) 7)}などを参照されたい。

4. 学術雑誌総合目録の調査票

今回の学総目録調査について、回答者側での記入作業を簡単に紹介しておく。まず大学図書館には「館別所蔵リスト」と呼ばれる、図1のような記録が並んだリストが渡される。これは一回前の学総目録調査の時点で、その図書館で所蔵していた雑誌タイトルである。特別な事情がない限り、今回もその図書館では同じタイトルを継続して受け入れ、所蔵巻号を増やしているはずである。館別所蔵リストのレコードのうち、「確認コード」なるものが、前述の2重誤り訂正RS符号として構成したものである。確認コードは「誌名番号」を符号化して作ってある。従って復号化手順で確認コードを復号するとここから誌名番号を再生できる。

図書館側の回答者は、館別所蔵リストを参照しつつ、前回の調査時点より増えた所蔵（雑誌の巻号名）を図2のOCR帳票に記入する。誌名番号はあらかじめ印刷されているので、対応する誌名番号の行の「巻次データ欄」に、増えた所蔵巻号を記入するのである。例えば図1のレコードの誌名番号はAA00218190であるが、これに対応するのは図2では15行目である。だから本タイトルについて所蔵が増えていたら、回答用紙の15行目に記入する。ただし回答者側では、この際に誤って別の誌名番号の欄に記入してしまっている可能性がある。調査票を回収した段階で、間違いなく対応する誌名番号の行に記入したことを確認できる必要がある。確認コードを館別所蔵リストからOCR帳票へ転記してもらい、調査票回収後に誌名番号と確認コードの照合を行なう。そのためOCR帳票に「確認コード」欄を設けてある。回答者は調査票記入時に、9桁の確認コードを館別所蔵リストから調査票に転記する。学術情報センターでは調査票回収後、調査票に記入された確認コードから誌名番号を再生する。そして、ここから再生した誌名番号と隣の欄の誌名番号との一致を試みる。もし不一致であれば、誤った欄に記入した誤記入であると判断される。

確認コードは2重誤り訂正RS符号として符号化してあるから、事後に学術情報センターで確認

コードを復号する際、9桁中2字までの記入誤しなく OCR の誤読があっても、正しく誌名番号を再生し、チェックできる。このような方策により、学総目調査の信頼性を高めることを意図しているわけである。

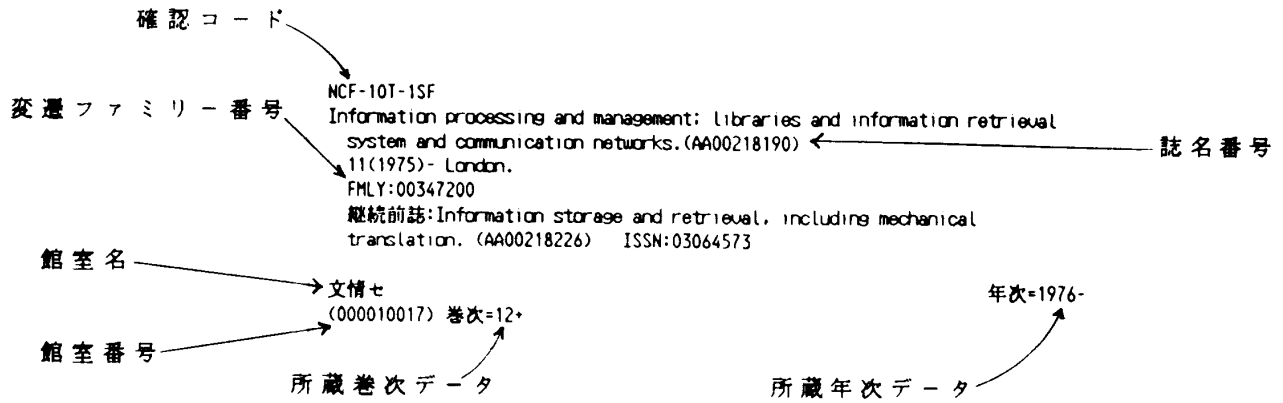


図1 館別所蔵リストのレコード

01|1|2|3|4|5|6|7|8|9 A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z +|-|.|#|(|)|

0011

学術雑誌総合目録欧文編
所蔵データシート(更新用)

館室番号	000010017	文情セ	000010017	巻次=12+	年次=1976-
01	AA00073757	S A G - 1 0 T - 1 S F			
02	AA00103349	Y G N - 1 0 A K - 7 1 8 0			
03	AA00113340	8 7 3 - 1 4 0 W - 1 A 2 1			
04	AA00144185	6 0 Y - 1 0 0 7 - 1 7 4 1			
05	AA00153729	Q Y 0 - 1 0 0 9 - 1 1 M 0			
06	AA00153360	1 3 1 - 1 1 0 0 - 1 8 R 1 0			
07	AA00153987	4 9 2 - 1 1 0 1 - 1 B 3 1 1			
08	AA00244049	2 1 5 - 1 4 0 M - 1 N 0 1 1			
09	AA00263699	Y 7 7 - 1 4 0 3 - 1 0 W 1 1			
10	AA00263645	1 1 8 - 1 1 0 1 - 1 8 T 0			
11	AA00225152	K T G - 1 0 A K - 1 P W 1 2			
12	AA00225169	M T G - 1 0 6 - 1 8 Y 0 2			
13	AA00231020	1 2 1 - 1 1 0 B - 1 A S 0			
14	AA00232994	5 6 1 - 1 1 0 N - 1 F 1 3 0			
15	AA00218190	N C F - 1 0 T - 1 S F 3			
16	AA00227045	R 7 H - 1 1 0 M - 1 P 8 0 3			
17	AA00227003	5 8 H - 1 1 0 9 - 1 9 5 8 2			
18	AA00240445	7 1 1 - 1 1 0 1 - 1 P 1 2 0			
19	AA00241578	H 8 L - 1 1 0 0 - 1 P 1 0 0			
20	AA00253653	5 0 N - 1 1 0 1 - 1 1 Q 5 0			

AA001 | AA00667302 | A|5|3|2|3|M|9|3|3| | 5|6|7|0|1|3| | 8|2|4| | 9|0|+ | | 9|7|2| | 9|8|5|

1. テータ記入上の注意) 2. テータ記入に際しては、テータ記入欄を必ずしも満たさない。 3. 各巻文刊行時に、次の場合を除き、原則として、行の所蔵テータを必ず入力し、2桁の0の場合、1. 巻次テータの行に必ず入力する。 2. 所蔵の中途に巻次テータの変更がある場合。 3. (所蔵テータが2桁以上ある場合は、所蔵テータシート(追加用)に記入を要する。)

東京大学文献情報センター

図2 学総目 OCR 帳票の記入例

なお9桁の確認コードは、記入時の便宜を考慮し、3桁毎にハイフンで区切った表記をしている。このうち5桁目(中央の桁)がほとんど「0」ばかりなのは、今回の符号アルゴリズムの癖である。

大部分の図書館では、手書きOCR用紙を使って回答した。この場合にはRS符号により、回答用紙の確認コード欄について、2字まで、回答者の転記誤り及びOCRの誤読を訂正できる。一部の図書館は所蔵レコードを磁気テープにより提出した。この際も、磁気テープのレコードは図2のOCR帳票と同一になるよう要請している。従って磁気テープのレコード中にもRS符号の確認コードが含まれる。磁気テープに記入されたデータを磁気テープ読み取り装置が読み誤ることはまず無いから、この場合RS符号での訂正対象は回答者の転記誤りのみである。

5. 磁気テープ回収データの解析

今回の調査では60,304件の雑誌所蔵レコードが磁気テープで提出された。この中で正しいフィールド位置に確認コードを記載していなかったものが若干あり、44件を確認コード自動訂正の対象から外した。自動訂正の結果は次のようなものであった。

解析レコード数	(N_R)	60,260
単一誤りレコード数	(N_{E1})	477
二重誤りレコード数	(N_{E2})	45
訂正不能レコード数*	(N_{NR})	14

N_{E1} および N_{E2} から、1文字あたりの平均誤り率 Pe を求める。確認コードが9文字からなることを考慮し、

$$Pe = (N_{E1} + 2N_{E2}) / 9 (N_R - N_{NR}) = 1.047 \times 10^{-3}$$

即ち磁気テープへのデータ・エントリー時、オペレータは約1000文字に1回の割合で記入を誤っている。

次に少し詳細に誤りの生起率について考察する。今回の確認コードで使用した31種類の文字について、各文字毎の誤り率を表1に示す。誤り率の大きな文字と小さな文字では、生起率に10倍以上の差がある。記入誤りは全ての文字に平等にランダム発生するのではなく、特に誤り易い文字があることがわかる。

具体的に、どの文字をどう誤ったかを調べてみる。特に発生率の高かった誤りパターンは表2のようなものである。になる。ここでたとえばEFという見出しは、本来FであるべきものをEと誤記入したことを示す。またEFの生起率である 1.83×10^{-3} とは次のようなものである。すなわち表1からFについての誤り生起率全体は 2.27×10^{-3} であることがわかるが、このうちの 1.83×10^{-3} すなわち約80%をEFという誤りパターンが占めていることを示す。

*訂正不能レコードとは3重以上の誤りがあり、今回のRS符号では自動訂正できなかったものを指す。

表 1. 符号文字毎の誤り率($\times 10^{-3}$)
 Table 1 Observed error rate of code character.
 (Factor 10^{-3})

文字 頻度*		文字 頻度		文字 頻度		文字 頻度	
0	1.09	8	1.72	H	0.76	R	1.03
1	1.38	9	1.13	J	1.94	S	1.06
2	0.64	A	0.36	K	0.92	T	1.26
3	0.26	B	1.73	L	0.76	V	1.41
4	0.79	C	0.37	M	3.10	W	1.57
5	0.69	E	0.82	N	1.47	X	1.49
6	0.69	F	2.27	P	1.97	Y	1.09
7	0.49	G	2.06	Q	2.47		

* 頻度1.00はその文字が1000回現れる度に平均1回誤記入されたことを示す。

表 2. 発生率の高い誤りパターン
 Table 2 Frequently observed error patterns.

パターン 頻度*		パターン 頻度		パターン 頻度	
EF	1.83	OO	0.97	HM	0.74
NM	1.81	S8	0.93	GQ	0.74
6G	1.25	KX	0.80	MW	0.74
I1	0.99	8B	0.79	MN	0.63

* 表1と同じ

以上の結果で興味深かったのは、符号文字として採用しなかった I ないし O という文字が、I1, OOとして4位、5位に出現していることだ。まさにこの誤りパターンを予期して採用を見送った符号文字であるが、それでもなお出現したのである。ところが同様の予想から不採用としたD, U, Zについては、

DP : 0.14×10^{-3}

UV : 0.15×10^{-3}

Z2 : 0.18×10^{-3}

程度の頻度でしか出現しなかった。I1, OOという誤りパターンから以下のような考察ができる。つまり I と 1、O と 0 の混同は、単なる識別誤りではなく、むしろ人間側で積極的に同一視している形跡があるということだ。また誤りパターンに方向性があるのも興味深い。例えば EF, FE という対

称なパターンは、

$$EF : 1.83 \times 10^{-3}$$

$$FE : 0.29 \times 10^{-3}$$

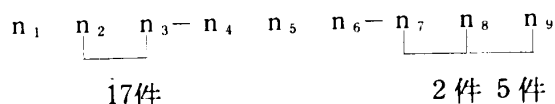
という全く異なった生起率を持っている。Fは誤ってEと記入され易いが、Eが誤ってFと記入される頻度はそれほど高くない。

最後に、今回の誤り訂正符号を設計した時に強く考慮した、転置誤りについて解析を行なう。1文字あたりの平均誤り率 Pe から2重誤りの生起率 N_{E2} を予想してみると、

$$N_{E2}' = N_R \binom{9}{2} Pe^2 (1 - Pe)^7 = 2.4$$

となり、 N_{E2} の実績である45から大きく隔たっている。これは2重誤りが単一誤りの複合以外のメカニズムで発生していることを示唆する。そのかなりの部分が転置誤りで占められる。

転置誤りは今回24件発見された。すなわち N_{E2} の約半分である。転置の起き易い場所には興味深い傾向が見られた。転置が起きた位置を確認コードの桁と照合すると、



であり、ハイフンで区切られた3文字組の末尾2桁の間で発生し易い。 $n_5 n_6$ 間で発生していないのは、符号化アルゴリズムの癖でほとんど $n_5 = 0$ となっている規則性によるものと考えられる。このような規則性は転置の発生を妨げる作用を持つようだ。

6. OCR 回収データの解析

OCR では674,788件のレコードが回収された。この内、自動訂正の対象にできたものは668,908件であった。誤り訂正の結果の内訳は下記のとおりである。

解析レコード数	(N_R)	668,908
単一誤りレコード数	(N_{E1})	20,957
二重誤りレコード	(N_{E2})	1,252
訂正不能レコード数	(N_{NR})	177

N_{E1} および N_{E2} から、前章と同様に1文字あたりの平均誤り率 Pe を求めると、

$$Pe = (N_{E1} + 2N_{E2}) / 9 (N_R - N_{NR}) = 3.898 \times 10^{-3}$$

となり、磁気テープ回収データの3倍に達する。以下の考察でも述べるとおり、OCRの誤読が頻発するためである。OCR回収データではOCRの誤りが人間の誤りをほとんど掩蔽しており、人間の誤り特性を解析する目的では使用が困難なものである。

表3に各符号文字毎の誤り率を示す。また表4に主な誤りパターンを示す。特に多いのは5Sと8Bというパターンで、この両者で5,833件、誤り全体の1/4強を占める。結果論としては、S、Bの両文字とも今回のOCR帳票では使用を避けた方が賢明であった。なお今回使用したOCR装置は、数

表 3 . 符号文字毎の誤り率($\times 10^{-3}$)
 Table 3 Observed error rate of code character.
 (Factor 10^{-3})

文字		文字		文字		文字	
頻度*	頻度*	頻度	頻度	頻度	頻度	頻度	頻度
0	0.92	8	5.166	H	1.92	R	3.95
1	0.84	9	3.48	J	3.91	S	28.97
2	0.79	A	3.07	K	4.15	T	2.42
3	1.01	B	15.44	L	1.73	V	13.39
4	2.11	C	3.57	M	4.65	W	2.39
5	3.70	E	2.03	N	2.97	X	2.75
6	1.34	F	3.91	P	3.39	Y	3.46
7	1.80	G	6.86	Q	9.99		

* 頻度1.00はその文字が1000回表れる度に平均1回誤記入されたことを示す。

表 4 . 発生率の高い誤りパターン
 Table 4 Frequently observed error patterns.

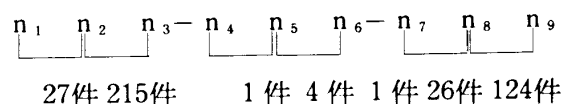
パターン	頻度*	パターン	頻度	パターン	頻度
5 S	25.90	0 V	4.28	X K	2.64
8 B	13.58	6 G	3.08	H M	2.24
9 Q	4.80	E F	2.90	B 8	2.19
9 V	4.30	S 5	2.74	Y V	1.82

* 表3と同じ

と字形の類似したアルファベットがあった場合、それを数字に読み誤る傾向があった。一方数字の認識率は良好である。数字を優先して認識するアルゴリズムになっているようだ。

表4で注目すべきなのは第3位に表れた9Qというパターンである。9とQはOCR装置にとって粉らわしい文字とは思えず、人間が「きゅう」という音に惑わされて犯した誤りであろう。明らかに人間側の誤りと判断できるパターンはこれだけである。ところで磁気テープ回収データで観測された9Qの生起率は 0.44×10^{-3} であり、OCRのものに比べて1桁低い。キーボードを操作する場合とOCR用紙にペンで記入する場合とでは誤りの発生機構が異なることがわかる。

転置誤りは397件検出された。2重誤りの約1/3である。発生した桁の位置は、



であり、磁気テープ回収データと同様な分布をしている。

どのような2文字間で転置誤りが生じ易いかも調べてみたが、これは全体としてはランダムに起きていた。その中で多少集中の傾向を見せたのは次の3種のパターンである。

2) XYYX 12例

2) GBBG 9例

3) A44A 7例

このうち1)は、人間がXYなるよく馴染んだアルファベット列に惑わされた可能性がある。

7. 結論

帳票の記入のような、オフラインのコード記入に適合した誤り訂正符号の構成法と、その学総目調査への応用例、回収データに関する若干の考察を述べた。もしこの符号系によらなければ、磁気テープ回収データでは約500件(0.8%)、OCR回収データでは約22,000件(3.3%)のレコードにつき再調査を要したことになる。この意味でRS符号による誤り訂正符号は十分目的を達したと考えてよいだろう。また今後同種の調査を行なう場合は、この解析結果に基づき更に目的に適合した符号系を作ることができると思われる。

謝辞

今回のデータ解析に協力された、学術情報センターの石井、牧村、郷端、羽田の各氏に感謝する。

参考文献

- 1) I.S.Reed & G.Solomon, "Polynomial Codes over Certain Finite Field", J.Soc.Indust., Appl. Math., 8, pp.300-304, 1960.
- 2) R.W.Hamming, "Error Detecting and Error Correcting Codes" B.S.T.J., 29, pp.147-150, 1950.
- 3) E.Prange, "Cyclic Error Correcting Codes in two Symbols", AFCRC-TN-57-103, Air Force Cambridge Research Center, Cambridge, Mass., Sept., 1957.
- 4) R.C.Bose & D.K.Ray-Chaudhuri, "On a Class of Error Correcting Binary Group Codes", Information and Control, 3, pp.68-79, 1960.
- 5) 根岸正光: 「欧文雑誌名の中の綴り誤り、類似語の検出方法(1)」東京大学文献情報センター紀要, pp9, 1, 1985.
- 6) 宮川, 岩垂, 今井: 符号理論, 昭晃堂, 東京, 1973.
- 7) 今井(編): 誤り訂正符号技術の要点, 日本工業技術センター, 東京, 1986.

研 究 論 文

TSS 接続による仮想画面転送 (VTSS) 方式

A Virtual Screen Transfer Method on TSS Connection (VTSS)

学術情報センター 安達 淳*

学術情報センター 橋爪 宏達**

学術情報センター 大山 敬三***

要 旨

本論文は、1985年に東京大学文献情報センターにおいて開発された、小規模端末システム用通信方式(VTSS)の実現方法について述べている。この方法は比較的小規模な計算機による端末システムに対して、画面型アプリケーションの仮想画面情報の転送を可能にすることを目的とし、分散処理に適した機能を有している。方式の中核は、東京大学文献情報センターおよび学術情報センターにおいてすでに運用を行っている N-1ネットワークに基づいた「画面転送プロトコル」を、N-1機能を持たない計算機に対しても適用可能にするための簡易化された下位プロトコルであり、本論文の中でその開発の背景、および実現方法について概要を述べるとともに、付録として仕様書を掲載している。

Abstract

A VTSS protocol, a new simplified communication protocol applicable to screen-oriented applications for small scale terminal systems, was designed in the Center for Bibliographic Information, University of Tokyo in 1985. Communication software for host and terminal systems was developed based on the protocol and has been in operation as a part of the online shared cataloging system of the National Center for Science Information System (NACSIS). The purpose of the VTSS protocol is to enable relatively small scale computers to be used as terminal systems for screen-oriented applications by simplifying the transfer method of virtual screen information, which is suitable for distributed processing. The National Center for Science Information System has already been operating an online shared cataloging system using the virtual screen transfer method based on N-1 Network. The VTSS is a simplified communication method for computers

*Adachi, Jun : National Center for Science Information System

**Hashizume, Hiromichi : National Center for Science Information System

***Oyama, Keizo : National Center for Science Information System

without complicated N-1 facilities and is used as a lower layer of the virtual screen transfer method as a substitute for direct use of the N-1 Network. In this paper, the background of development of the VTSS protocol is described first, the conditions of applicable network are discussed next, and then the implementation method of the communication software of the VTSS host system is presented. The detailed specification of the VTSS protocol is included in the appendix.

1. はじめに

学術情報センターの前身である東京大学文献情報センターでは1985年7月に小規模図書館向けの画面指向通信用簡易プロトコルである VTSS 方式のプロトコルの開発を行い、これに基づいてホスト及び端末システムの開発が進められ、現在数種類の小規模計算機による端末システムが図書館業務に供されている。

本論文では、VTSS 方式の開発の背景について述べ、接続形態について検討するとともに、ホストシステムにおけるネットワークソフトウェアの実現方法について述べる。付録として、VTSS プロトコルの仕様を添付する。

2. VTSS プロトコル開発の背景

学術情報センターを中心として拡充されつつある図書館ネットワークでは、異機種間コンピュータネットワークの上で画面型のアプリケーションを実現し、使い勝手の良いユーザインタフェースを提供することを、その1つの目的としている。学術情報センター、およびその前身の東京大学文献情報センターでは1984年から大学間コンピュータネットワーク(N-1ネットワーク)を基礎にして、画面型のユーザインタフェースを持つオンライン総合目録システムの開発と運用を行ってきた。N-1ネットワークのノードにはさまざまなメーカーの機種 of 計算機が用いられているが、N-1ネットワークがホスト-ホスト間通信用に開発されたため、これらはいずれも比較的大型の計算機であり、またプロトコル自体も複雑である。このため小規模な計算機システムでは N-1ネットワークを利用することは困難である。しかし、通常の大学図書館ではこのような計算機を導入または利用することが可能である場合はむしろ少ない。したがって、学術情報センターの図書館ネットワークに、より多くの図書館が参画するためには小型計算機でも利用可能な簡易型の画面転送プロトコルが必須である。

一方で、最近のデータ通信ネットワークの発達により、高性能で信頼性の高い通信が日本国内のほぼ全域で利用可能になってきた。これらのネットワークを用いると、回線上の通信エラーが実効上無視しうるため、ホストおよび端末では信頼性確保のための複雑な制御をする必要がない。

このような背景から VTSS プロトコルの設計を行い、学術情報センターホストシステム用のネットワークソフトウェアの開発を行うと同時に、主な計算機メーカーに対して端末システム用のネットワークソフトウェアの開発支援を行ってきた。この結果、1986年4月からホストシステムでの VTSS

方式によるサービスが開始され、数種類の端末システムがすでにセンターと接続されて図書館業務に供されている。

3. VTSS 方式による接続形態

VTSS 方式の位置付けについては付録の仕様書の「1. TSS 接続による仮想画面転送方式 (VTSS)」において、N-1ネットワークに基づいたシステム構成との対比により詳細な説明がなされているので、ここでは省略する。

接続形態については付録の仕様書の「2. VTSS のネットワーク形態」において、可能なものとして3つの例をあげているが、これらはオンライン総合目録システムの目録業務、すなわちデータベースの登録過程における通信上の品質に対する要求を考慮したものである。また、通信コストに関しては考慮せず、論理的な可能性についてのみに基づいている。

ネットワークとしてはデジタル専用回線（符号品目専用回線）全二重通信路と X.25パケット網 (NTT DDX-P など) が示されているが、ここで注意が必要なのは、VTSS プロトコルには N-1プロトコルのような TSS 通信リンクの多重化機能がない点である。このため、N-1ネットワークでは1本の専用回線で複数の TSS セッションを張ることができるのに対して、VTSS では1本の専用回線では1つの TSS セッションしか張ることができない。したがって、ごく僅かな特殊な場合以外はデジタル専用回線で VTSS 接続をすることはコスト的にみて現実性がない。

X.25パケット網では、X.25インタフェース自体に仮想リンクの多重化の機能があるため、VTSS であっても1本の X.25回線で複数の TSS セッションを張ることができる。このため、実際上 VTSS は X.25を前提にしていると考えて差し支えない。

その名のとおり VTSS は TSS 接続の上で画面転送をするための通信方式であり、通信の品質を無視すれば、TSS による利用が可能ないかなる手段によっても VTSS の利用も可能である。しかし、現在、VTSS 方式の唯一のアプリケーションであるオンライン総合目録システムではデータに高い品質が要求されるため、サービスを上述のような高信頼なネットワークに限定している。

今後、それほど高い信頼性の要求されないアプリケーションを VTSS 方式でサービスすることがあった場合、モデムを用いた電話網経由の利用などをどのように扱うかは、その時点における諸々の状況に鑑みて決定されることになるであろう。

4. ホストシステムにおける VTSS 方式の実現方法

学術情報センターホストシステム（サーバホスト）における N-1ベースの画面型アプリケーションは現在 TSS 環境下で稼働している。このため、VTSS 接続による画面型アプリケーションも同様に TSS 環境下で稼働することが自然な実現方法である。この場合、VTSS 方式はプロトコルレイヤの形式をなしているため、アプリケーションの一部として実現する方法と、N-1プロトコルのように OS に組み込む形で実現する方法の2つが考えられる。しかし、サーバホストには汎用大型計算機と

汎用 OS を用いているため、OS の改造にはかなりの困難がともなう。そこで、機能的に可能であればアプリケーションの一部として VTSS を実現することが現実的である。この考え方に基づいて検討を加えた結果、機能的な問題はないとの結論に達し、実現に向けての設計を行った。

ここで最も問題となるのは付録の仕様書の「6. VTSS レコード管理」に述べられたサブレイヤの機能の実現である。画面型のアプリケーションが TSS 環境下で稼働することから、アプリケーション自体以外にも端末に対して出力をしてくるものがある。例えば、OS あるいは TSS からのエラーメッセージ、コンソールからのオペレータメッセージ等がアプリケーションとは非同期に TSS 接続を通して端末側に出力されてくる。どのようにしてこれらのメッセージから仮想画面データを区別して送信するかが問題となった。幸い、用いている OS では 1 回の出力処理のデータ中に上記のような非同期のデータが混入することはないので、VTSS サブレコードを必ずひとまとまりのデータとして出力することによりこの問題を回避することができた。

もう一つの問題点は拡張 NVT(XNVT)コード (16ビット) と VTSS コード (7ビット) との間のコード変換である。使用している OS の内部文字コードは XNVT コードとは直接の対応がないため、OS に備わっているコード変換機能が使えない。この問題は、OS の TSS 端末入出力ソフトウェアを透過モードとし、必要な変換はすべて VTSS 側で行うことにより解決した。しかし、上記のような、画面型アプリケーション以外からの非同期なデータについては VTSS プロトコルの処理を受けないため、OS 独自のコード変換を受けることになる。幸い、XMVT コードは、EXC 文字等の外字を除いては JIS-C6220 および JIS-C6226 (1983年版) に準拠しており、また、これらのメッセージには外字が含まれていないことから、画面型アプリケーション以外のデータに関しては OS の変換機能を用いることによって、VTSS の処理を施したのと同じ結果が得られる。

なお、以上の議論から明らかなように、VTSS プロトコル自体の規定によると、仮想画面情報だけでなく、通常の TSS セッションの入力・応答や各種メッセージも VTSS コードデータによっているため、適切なコード変換が必要である。現在は行モードにおいて外字を扱うアプリケーションが存在しないが、将来的にはこの可能性もあるため、端末システムの実現時にはこの点も注意しておく必要がある。

5. おわりに

本論文では、東京大学文献情報センターおよび学術情報センターが小規模図書館に対して画面型のアプリケーションをサービスするために採用した簡易型の画面転送方式の実現方法について述べた。その中核となる VTSS プロトコルの仕様を付録に載せ、技術資料としても参照可能とした。

本方式は、参考文献 2) に述べたようにその妥当性と実現性の検証が行われ、これにもとづくネットワークソフトウェアは 1985 年にインプリメントが行われ、図書館業務に供されている。

これにより、学術情報センターのサービスの基盤となるべき通信手段として、N-1 ネットワークと VTSS 方式が利用可能となったことになり、既存の図書館ネットワーク、あるいは今後開始される

予定の一般ユーザに対するサービスのより一層の拡大が期待される。

謝辞

本論文で述べたソフトウェア構成手法および仕様をまとめるにあたっては、日立製作所の松原氏、水村氏に負うところが大きい。また、仕様書の作成に際して、学術情報センターの渡辺博氏、ならびに大日方一男氏の協力を得たことを記して、謝意を表す。

参考文献

- 1) 安達淳:「画面志向アプリケーション向けネットワークプロトコルの開発」東京大学文献情報センター紀要、第1号、1985年8月。
- 2) 大山敬三:「画面指向通信用簡易プロトコル VTSS のテストシステムおよびプロトタイプシステム」学術情報センター紀要、第1号、1987年。

付録 TSS 接続による仮想画面転送 (VTSS) 方式

第0版 昭和60年7月1日

第1版 昭和61年10月1日

はじめに

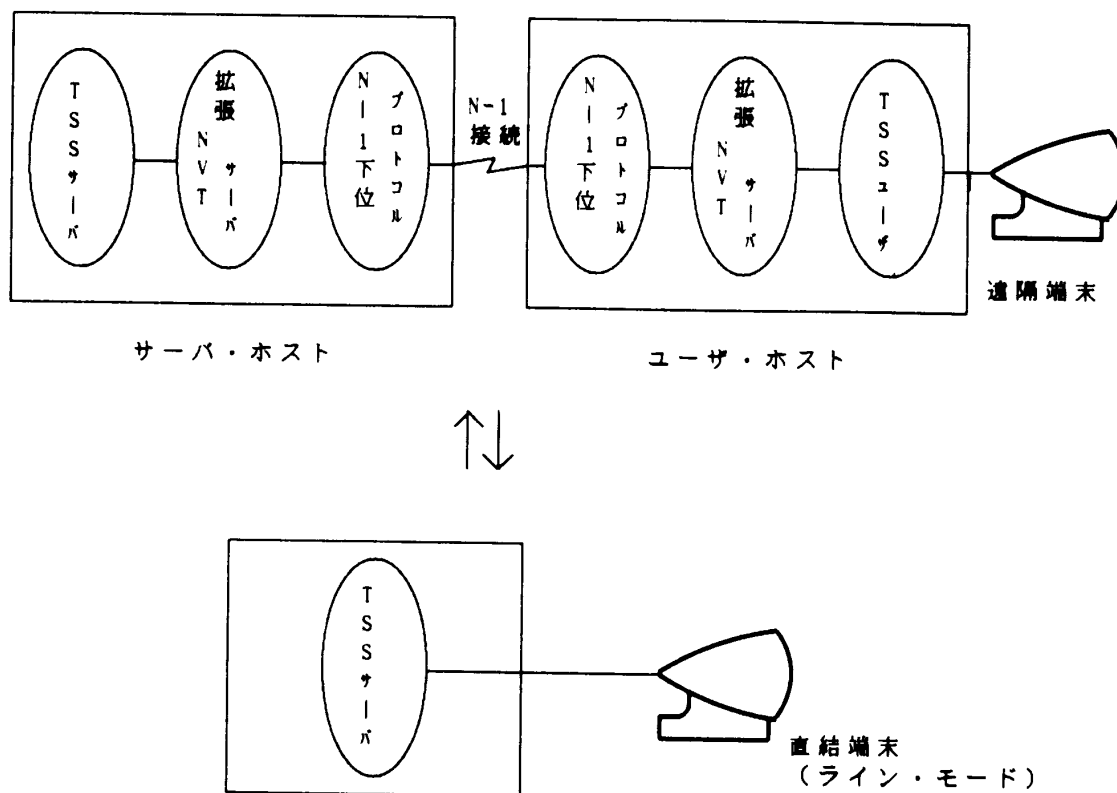
本仕様書は、TSS 接続による仮想画面転送方式について記述したものであり、これは学術情報センターが開発する TSS 端末向けの仮想画面転送に基づくアプリケーション (業務) プログラムに共通に適用することを意図したものである。そのようなアプリケーションの例としては、例えば目録作成、目録検索、ILL (相互貸借、複写) 業務、二次情報検索等のものがある。

なお本仕様書に記述する仕様のすべては学術情報センターの管理下にあり、また今後は関係各者との協議を経て最も実現性の高いものに改訂される可能性があることを言明する。

1. TSS 接続による仮想画面転送 (VTSS) 方式

TSS 接続による仮想画面転送 (VTSS) 方式とは、N-1 拡張 NVT プロトコルを使用し、昭和59年

(a) 通常の TSS の利用



(利用者から見た接続のイメージ)

図1 N-1 拡張NVTプロトコルの概念

(b) 図書館ネットワークでの形態

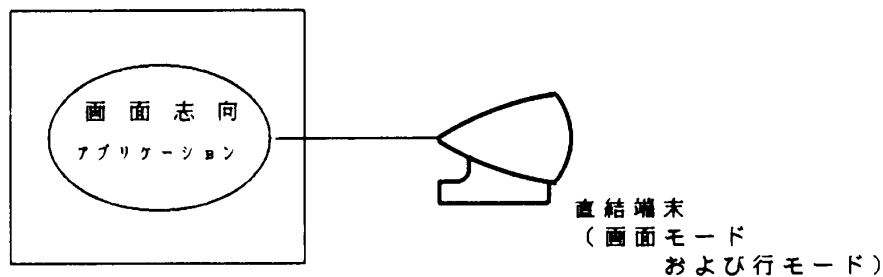
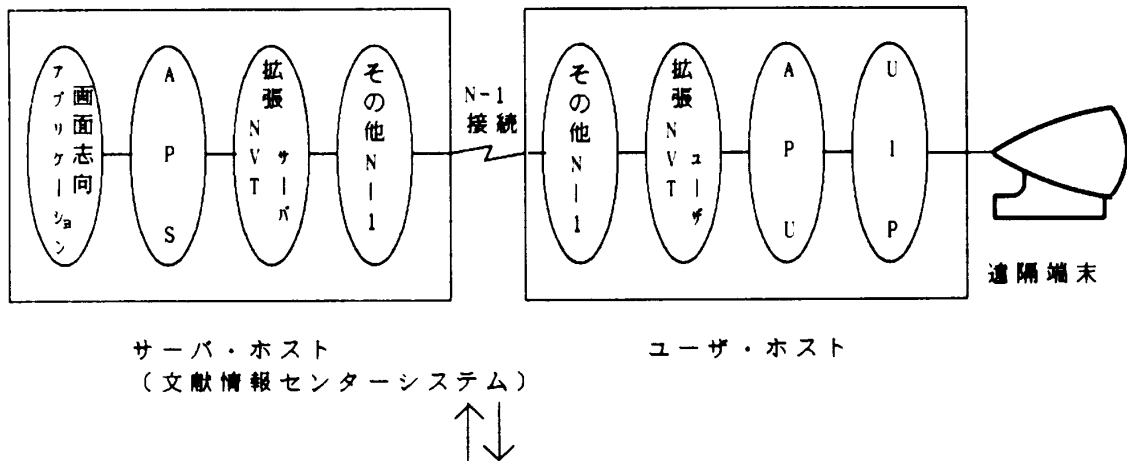


図 1 (続き)

に東大文献情報センターが開発したアプリケーション間通信データ転送方式(APP (Application Program) 間転送方式) を、通常の TSS 接続された小規模計算機端末においても適用可能とするものである。

本仕様書を理解するために、まず APP 間通信の基本理念について述べる。N-1の拡張 NVT (Network Virtual Terminal) プロトコルとは図 1 (a)のように、サーバホスト・コンピュータのアプリケーション・プログラムを、大学間コンピュータネットワーク内のユーザ・ホストに接続された端末からアクセスするために使用されている。通常、サーバ側のアプリケーションとは、TSS サービスプログラムであり、ユーザ側端末からサーバホストの TSS を利用するために使われているプロトコルである。現状では、いわゆるラインモード (行モード) の形態で TSS を利用している。一方、学術情報センターのサービスする画面志向業務プログラムでは、ユーザ側の画面端末を対象とするアプリケーションを実行するために、サーバホスト TSS 業務プログラムを起動すると同時にユーザホストでも画面を受信するためのプログラム(APU (Application User) および UIP(User Interface Program))を起動する (図 1 (b))。APS(Application Server) および APU-UIP の機能を付加することにより、単に行モードの TSS 利用のみならず、拡張 NVT 単体では果たしえなかった画面型会話処理が実行できるようになる。

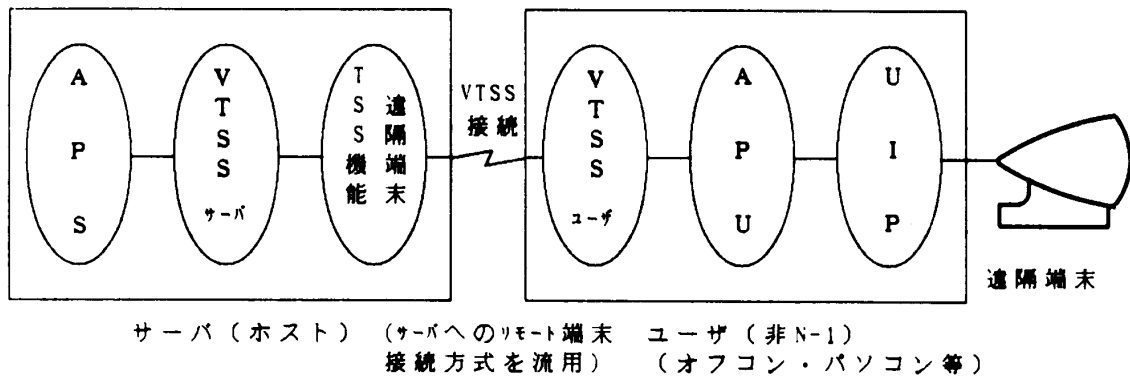


図2 VTSS接続

この方式で使用した N-1 拡張 NVT プロトコルとは、本来は単にユーザ・ホストの端末を遠隔サーバホストの TSS 端末に仕立てあげるためのものであったことに留意されたい。もし図2のようにユーザ側端末システム（オフコン、パソコン等）に画面アプリケーション用 APU、UIP を持ちこめば、これを通常の TSS 遠隔端末としてサーバ・ホストに接続することで、機能配置としては図1 (b) とほぼ等価になる。VTSS 方式とは、このような形態のネットワーク機能を可能にし、N-1 プロトコルが搭載されていないコンピュータ・システムをも学術情報センターシステムへ参入できるようにするものである。

2. VTSS のネットワーク形態

VTSS では、サーバ／ユーザ間の遠隔 TSS 接続のために各種のネットワーク接続方式が選択可能であるが、次に述べる要求条件から運用上はさしあたり図3の3種類のうちのいずれかの形態に適用を限ることとする。形態 A は専用線によるものであり、X-on、X-off 機能制御文字によりフロー制御を行なうものとする。形態 B は X.25 パケット網（NTT の DDX パケット網等）によるものであり、更に形態 B はパケット形態端末による形態 B-1 と PAD を使用する形態 B-2 に分かれる。

要求条件:

- (1) 回線上の通信エラーが実効上無視しうる回線品質をもつこと。
- (2) 連続送信の一時中断、継続要求が可能であること。

(1) の理由から、遠隔端末接続には最もよく用いられているモデム接続、ないし網間接続などアナログ電話回線経由の接続は棄却される。また(2)を可能とするため、パケット通信によるフロー制御、または X-on/X-off の制御手順を使用することになる。

BASIC 手順(BSC 手順) も上記の要求仕様を満たすものであるが、今後のネットワーク展開計画上の判断から、採用を保留している。

3. 拡張 NVT 機能との代替

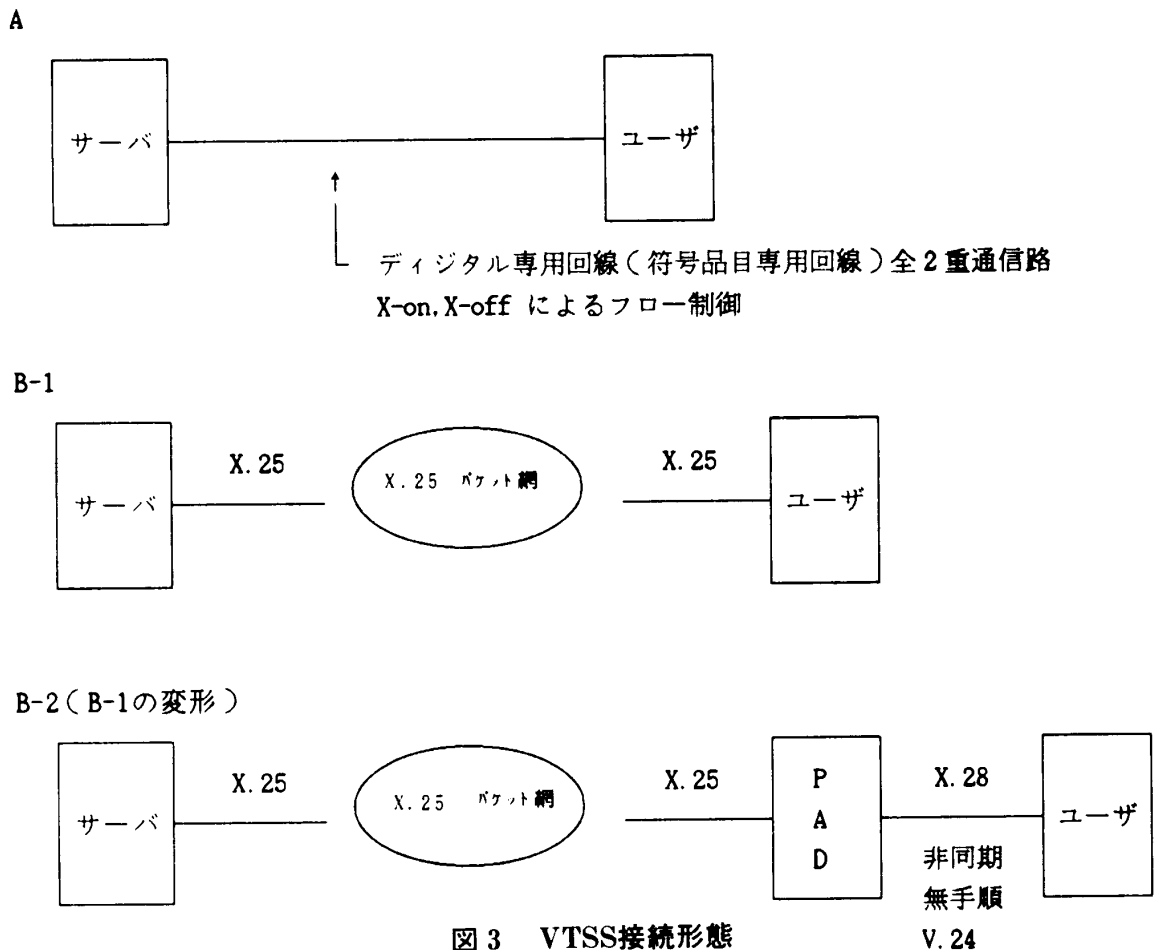


図3 VTSS接続形態

VTSS方式は、形式的にはTSS接続上でAPU-APS間通信(APP間通信)を行なうためのものであり、拡張NVTの機能と通常のTSS遠隔端末データ転送機能とにおいて等価でない部分の機能の補完を行なう必要がある。従来の拡張NVTプロトコルを使用したAPP間転送においては、拡張NVTプロトコル及びAPP転送手順は次のような機能を果たしている。

- (1) 16ビット文字コード通信機能を持ち、拡張NVTコードとして定義された文字セットによる通信をする。1バイト文字はJIS C6220、2バイト文字はJIS C6226(1983年版)を包含している。
- (2) 拡張NVTは同一文字列の圧縮機能を有す。
- (3) 通常のTSSセッションのためのライン転送モードと、画面志向アプリケーションのための仮想画面転送モードが共存している。APU-APSが実現する機能である。両者のモードのレコードを多重化して、同時に通信できる。
- (4) ユーザ側からの割込み要求を受けつける。
- (5) 拡張NVTは全2重通信をサポートする。ただしAPU-APS間通信は半2重であるため、APU-APSが送信権制御を実行している。

これに対し、学術情報センターシステムで利用可能なTSSでは、次のような仕様で通信機能をサ

ービスしている。

- (1) 7ビット情報+1偶数パリティの文字セットであって、制御文字のうちに何種類かはサーバや通信ファシリティの事情から非透過なものがある。
- (2) データの圧縮機能はない。
- (3) APU-APS 間通信に必要な2つのリンク（ラインモードと画面モード）を、1本の TSS リンクに多重化しなければならない。そのための制御キャラクターの透過性や、転送タイミングの違いなどに留意しなければならない。
- (4) TSS において、TSS 固有の「割込み要求」の機能が使用可能である。
- (5) TSS は半2重通信しかサポートしない。TSS の送信権の所在と APU-APS 間通信の送信権の関係を考慮しておかねばならない。

この違いを吸収するために、APS と TSS サーバの中間に、図4(a)のように VTSS 層を設ける。VTSS 層が実現すべき機能は次のようなものである。

- (1) 16ビット文字コードを7ビットコード系上に変換する。これによって拡張 NVT コードの通信を可能とする。
- (2) 文字列の圧縮機能をサポートする。
- (3) ラインモードおよび画面モードのいずれもの転送要求を、「VTSS レコード」概念で統合する。
- (4) TSS 固有の「割り込み要求」の機能を使用する。
- (5) APU-APS 間送信権制御と TSS 送信権制御の調整・処理を行なう。

これらの事情から、VTSS 層は、更に図4(b)のような3層のプロトコルの機能単位に分解できる。

4. 16ビット7ビット変換機能

図5のように、SI、SO 制御文字(P0/F および P0/E)を使用して文字コードの8ビット目の on-off を制御する。これにより8ビットないし16ビット長の文字コードによる通信を可能とする。ここで P はパリティ・ビットを表す。

通信で使用する文字コードは、拡張 NVT コードを基本とする。このため、コード変換は

(各システムの内部表現) ⇔ (拡張 NVT コード) ⇔ (VTSS リンク上の7ビットコード)

というように論理的には2段階が存在する。文字列変換の例を図6に示す。

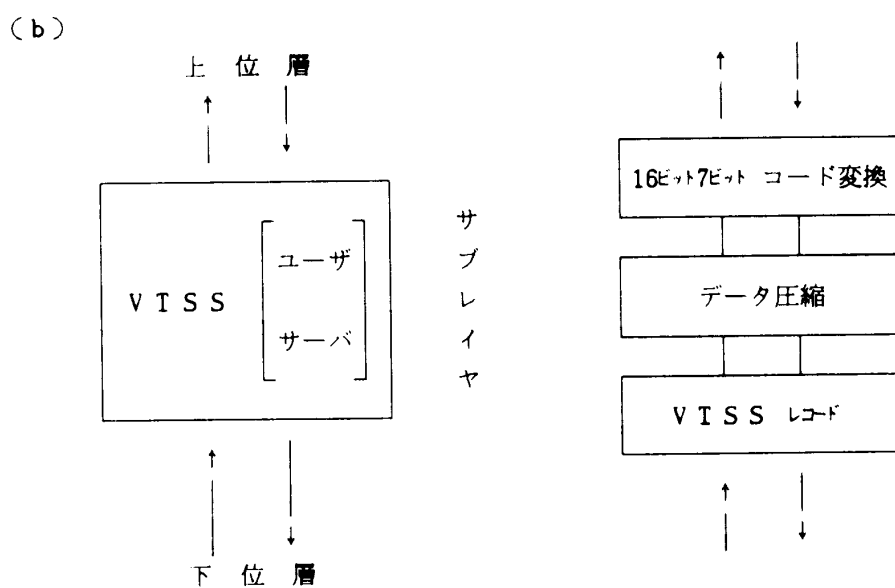
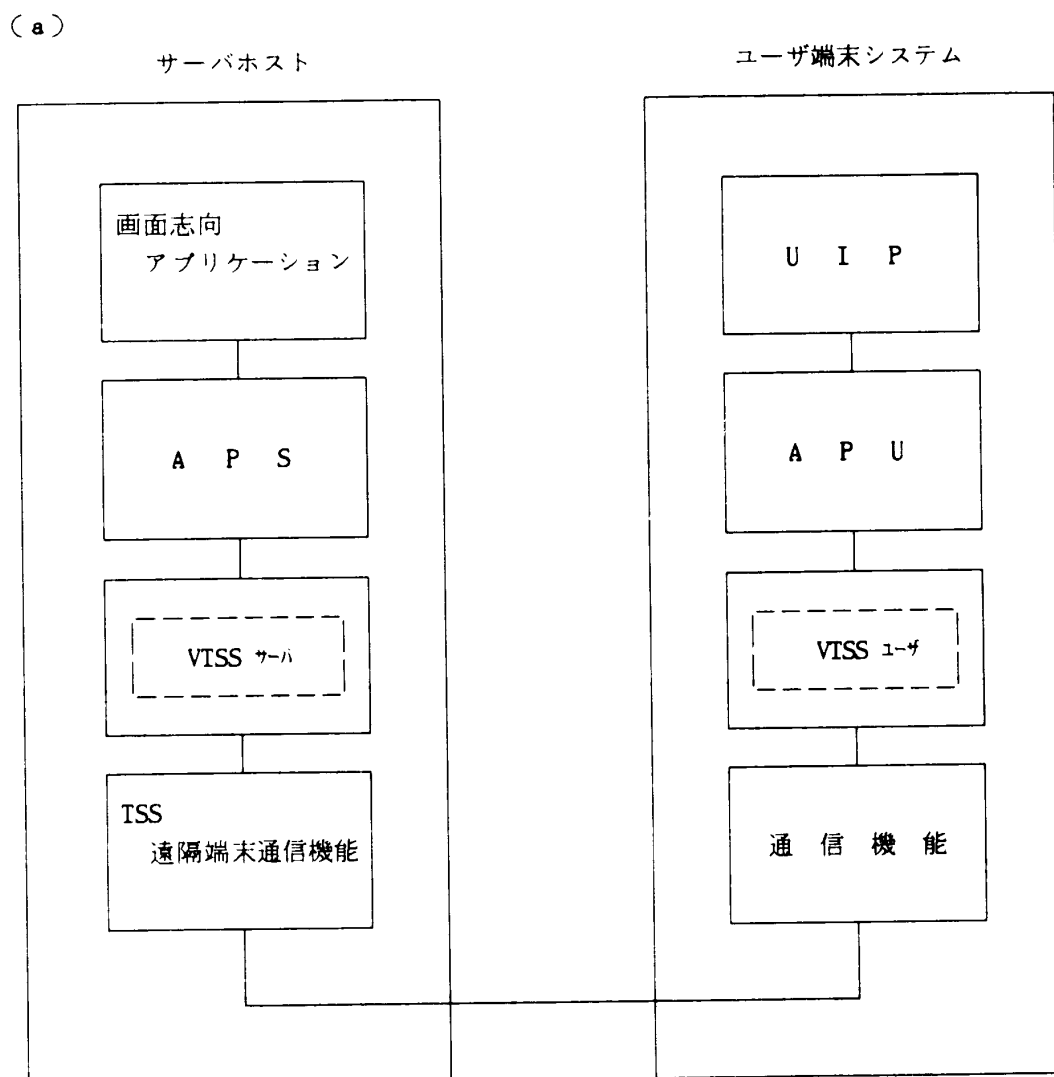


図4 VTSS層の位置づけ

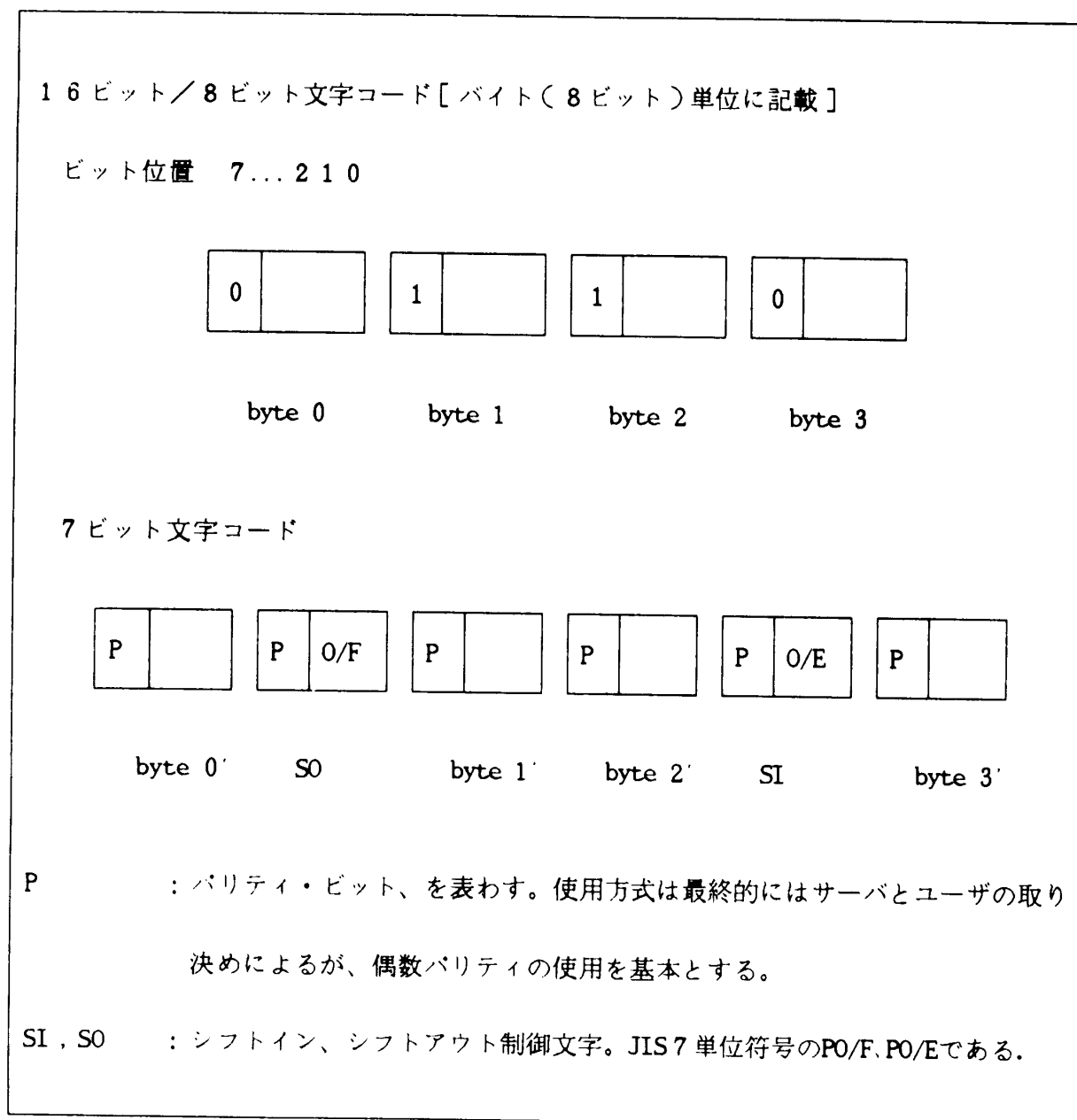


図5 16ビット-7ビット変換

5. データ圧縮機能

転送データ中に同一文字が連続する場合、VTSS 方式においては図7のようなエスケープ・シーケンスで圧縮できるものとする。

圧縮シーケンス中で表記可能な連続文字数 NN は $0 \leq NN \leq 99$ である。論理的には、 $0 \leq NN \leq 99$ が意味を持ち、受信に際しては正しく解釈されねばならないが、エスケープ・シーケンス自体が5文字の列であることから、圧縮としての実効上は $6 \leq NN \leq 99$ が意味を持つ。データの送信者は、この範囲で圧縮機能を使用するようにインプリメントしなければならない。ただし、データ送信時に圧縮機能を使用するか否かは、サーバ、ユーザ側とも任意である。受信時に圧縮を解く機能を用意してお

文字列の例：

ABアイ亜以ⓐEF

(1) 内部表現の例

C1 C2 81 82 0A 42 B0 A1 B0 CA
 ↓ ↓ ↓ ↓ └──┬──┬──┘
 A B ア イ 漢シフト 亜 以

0A 41 C3 C4 0A 42 88 D2 0A 41 C5 C6
 └──┬──┘ ↓ ↓ └──┬──┬──┘ ↓ ↓
 英シフト C D 漢シフト ä 英シフト E F

(2) 拡張NVTコード

41 42 B1 B2 1B 24 42 30 21 30 4A
 ↓ ↓ ↓ ↓ └──┬──┬──┘ └──┬──┬──┘
 A B ア イ 漢字シフト 亜 以

1B 28 4A 43 44 1B 24 42 E6 52 1B 28 4A 45 46
 └──┬──┘ ↓ ↓ └──┬──┬──┘ └──┬──┬──┘ ↓ ↓
 英シフト C D 漢シフト ä 英シフト E F

(3) VISS 7ビット・コード(パリティは無視して表記)

41 42 0E 31 32 0F 1B 24 42 30 21 30 4A
 ↓ ↓ ↓ ↓ ↓ ↓ └──┬──┬──┘ └──┬──┬──┘
 A B S0 ア イ SI 漢シフト 亜 以

1B 28 4A 43 44 1B 24 42 0E 66 0F 52 1B 28 4A 45 46
 └──┬──┘ ↓ ↓ └──┬──┬──┘ ↓ └──┬──┬──┘ ↓ ↓
 英シフト C D 漢シフト S0 SI 英シフト E F
 ↑ ↑
 ä

図6 文字変換の例

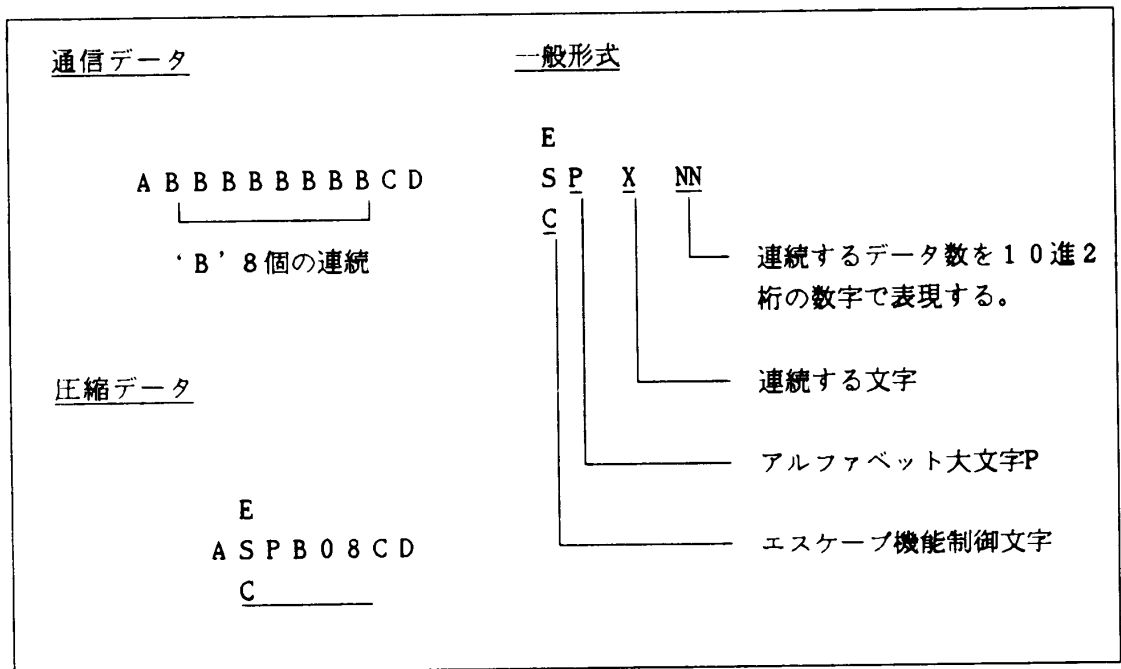


図7 データ圧縮機能

くことは両者とも必須とする。

6. VTSS レコード管理

このサブレイヤの機能は、VTSS で使用するレコードの単位を規定することにある。

ホストとユーザの間には、TSS コマンド-応答リンクと、APU-APS 間転送リンクの2本の仮想リンクが存在する。前者は LOGON コマンドにより TSS セッションを起動した時点で生成し、通常の TSS コマンドやその応答はこのリンク上を流れる。一方で、後者の APU-APS 間通信リンクは、前者のリンクを使用してホストに画面志向アプリケーション起動コマンドを投入した時点で成立する(図8)。VTSS レコード層の機能は、TSS コマンド-応答リンクと APU-APS 間通信リンクの多重化を主眼としている。VTSS レコード層により成立した多重リンクを、VTSS リンクと呼ぶことにする。VTSS リンクは、ホストの実現上の理由から現在は半2重通信のみが可能である。従ってVTSS 管理ではこのリンクの送信権も管理する。

[VTSS レコード形式]

図9にVTSS リンク上を流れるレコードの形式を示す。

受信側がリンク上を流れるデータ・ストリームからVTSS レコードを構成するためには、次の2つの原則を適用する。

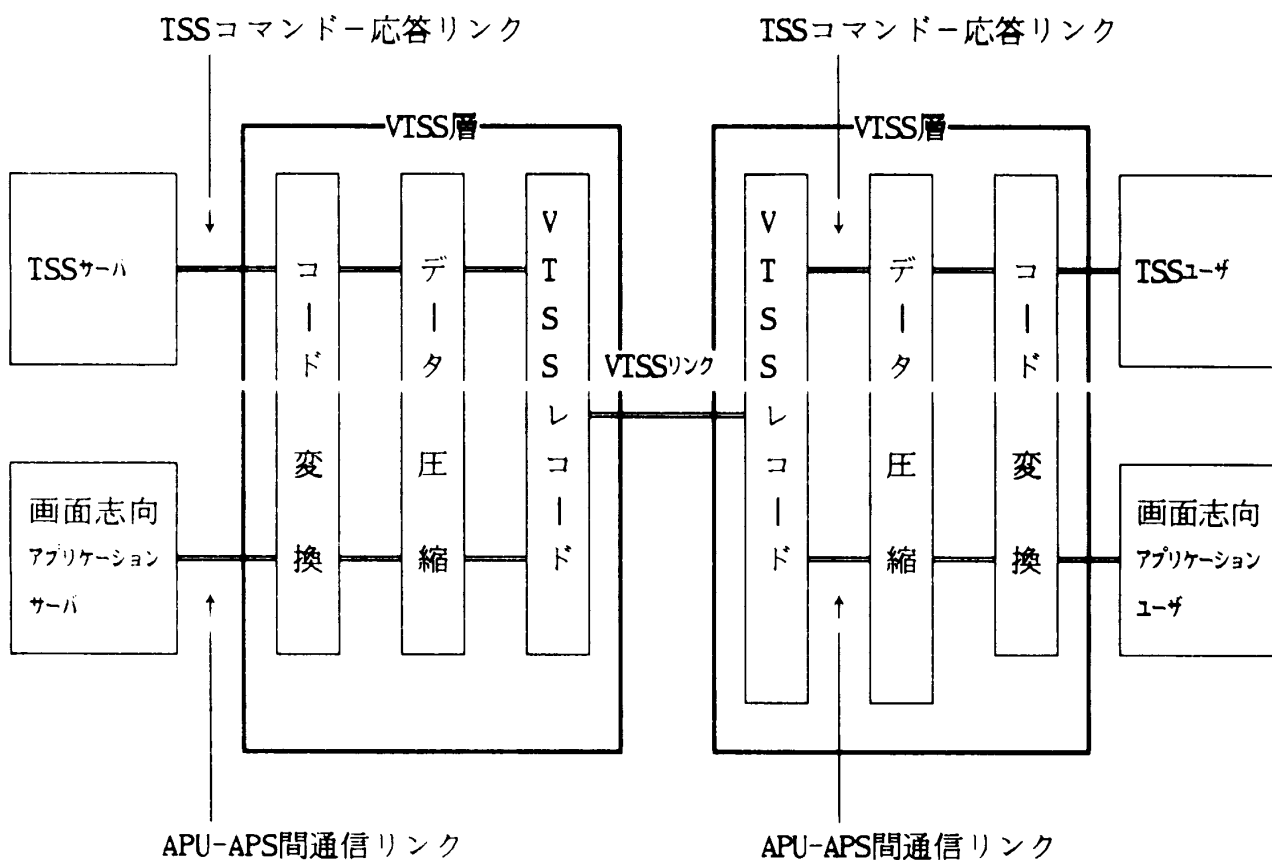


図8 画面志向アプリケーションの使用するリンク

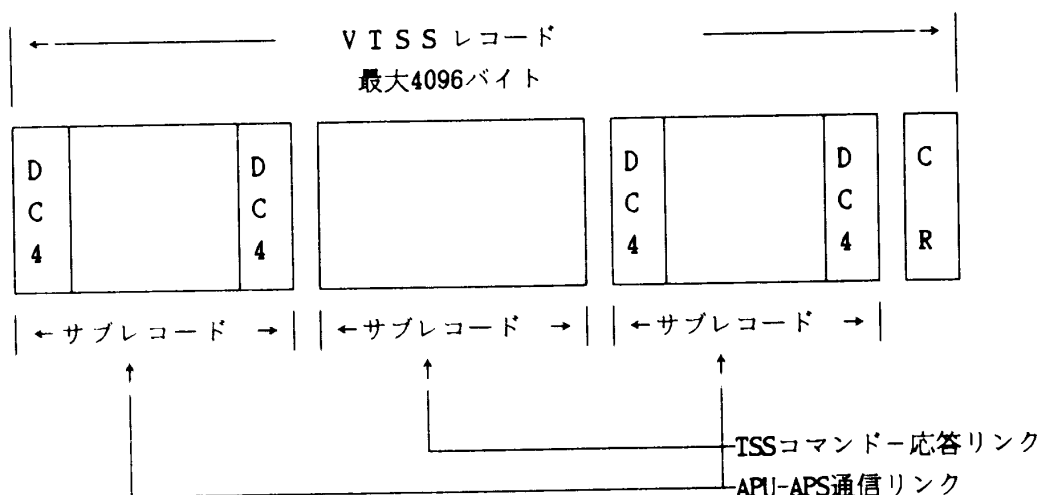


図9 VTSSレコード形式

- (1)データ・ストリームの中の LF コード(P0/A)はすべて無視する。
- (2)データ・ストリームの中の CR コード(P0/D)をレコードの区切りとする。

なお以上において、PAD を使用する接続形態ではパケット送出契機文字として CR を使用することを想定している。

このようにして得られた VTSS レコードに対して DC4制御文字を区切り文字としてサブレコードを切出す。DC4制御文字(P1/4)で囲まれたサブレコードが APU-APS 間通信リンクのレコードであり、それ以外のサブレコードが TSS コマンド-応答リンク上のレコードである。

この VTSS レコードには更に次のような制限事項が付帯する。

- (1) VTSS レコードの先頭から、終結文字 CR までは、最大4096バイト以内とする。(これは、コード変換、圧縮の操作を行った後の文字数で考える。)
- (2) ひとつの VTSS レコードが複数個のサブレコードを含んでよい。逆に、ひとつのサブレコードを複数個の VTSS レコードに分割することはできない。
- (3) TSS コマンド-応答リンク中には、DC4制御文字が出現してはならない。また CR 文字の出現は、TSS コマンド-応答リンクサブレコード内の文字とは見なされず、VTSS レコードの終結文字として扱われる。
- (4) APU-APS 間通信リンク中には、サブレコードの両端をのぞいて DC4制御文字が現われてはならない。また CR 文字が現われてはならない。

[送信権制御]

VTSS 層ではサーバ-ユーザ間の VTSS リンクの送信権も管理する。送信権とはユーザ・ノードが VTSS レコードの送信を行う権利と定義する。サーバ・ノードはあらゆるタイミングでレコードを送信しうるものとし、ユーザの送信権を無視してデータ送信を行ってよい。従ってユーザ・ノードは自ノードの送信権在否にかかわらずこれを受信しうることを期待される。

ただし、APU-APS 間通信リンク及び定常状態の TSS コマンド-応答リンク上の通信は、ユーザ

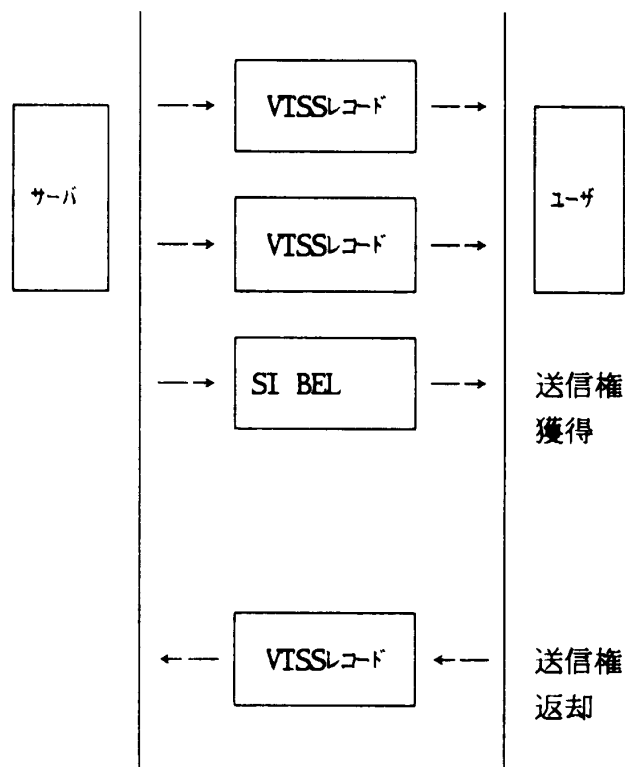


図10 送信権獲得、返却

側の送信権を尊重しているため、送信権がユーザ側にあるときにサーバが積極的にこのリンク上に送信することはないと考えてよい。例外はサーバ・ノードからの緊急メッセージである。これはユーザ・ノードの送信権を無視して流れるので、上述のような送信権制御が必要になる。インプリメント時は、緊急メッセージを除いてユーザ・ノードに送信権があるときにサーバが送信を行うことはないと考えてよい。

ユーザ・ノードの送信権獲得、返却は次のように行なわれる。(図10参照)

(1) 送信権獲得:

サーバから受信したデータ中で、TSS コマンド-応答リンク中に SI 制御文字、BEL 制御文字(P0/F、P0/7)が連続して、この順で出現していた時、これを受信したユーザは自分が送信権を獲得したものとみなす。

(2) 送信権返却:

ユーザ・ノードがひとつの VTSS レコードの送信を終了する度に、そのユーザの送信権はホストへ返却されたものと見なす。従ってユーザは VTSS レコードの連続送出はできず、必ず次の SI、BEL の到着を待たねばならない。

7. VTSS 層と上位層との関係

端末側の上位層は APU 層である。APU 層と VTSS 層の間のインターフェースでは、以下のよう

な情報がやりとりされる。

(a) 送信権の有無

(b) 割込みの通知(APU 層から VTSS 層)

(c) APU/APPS レコードの受信:

VTSS 層は、VTSS レコードから抽出された APU/APPS 通信リンク上のサブレコードを APU/APPS レコードとして、APU 層に与える。

(d) TSS レコードの受信:

VTSS 層は、VTSS レコードから抽出された TSS コマンド-応答リンク上のサブレコードを TSS レコードとして APU 層に与える。これに対して APU 層は、TSS レコードを文字列メッセージとして各レコードを一行毎に画面上へ（画面モードの表示とは区別しつつ）表示する。

(e) APU/APPS レコードの送信

および

(f) TSS レコードの送信

APU 層の与えるこれらのレコードを、VTSS 層は順次連結する。そして長さが VTSS レコードの最大長に達する以前で、かつ送信単位(VTSS サブレコード、すなわちひとつの APU/APPS レコード、ひとつの TSS レコード) が完結する毎に VTSS レコードとして送信する。

(g) 情報送出契機(APU 層から VTSS 層に対して)

APU 層は、送信契機を VTSS 層に与えることができる。これに回答して、VTSS 層は VTSS レコードを終結してリンク上に送出する。

参考文献

以上で VTSS のためのデータ転送手順の解説をおわる。VTSS 接続による業務システムを作成する場合は、本仕様書に併せて下記の文献を参考にしていきたい。(1)の N-1 拡張 NVT プロトコルは、VTSS 手順により代替されるものであるが、VTSS の文字コードは拡張 NVT コードを中間的に援用している。(2)のアプリケーション間通信方式は、VTSS の上位の通信規定である。APU-APS 間通信リンクはこのサービスを行うためにある。なお、実際の画面アプリケーションを作成する際の細かい実現仕様については、(3)が参考になろう。

- | | |
|------------------------------------|--------------|
| (1) 大学間コンピュータネットワーク用仮想端末プロトコル第 2 版 | 1985年 4 月26日 |
| (2) アプリケーション間通信用データ転送方式 | 1984年 8 月 |
| (3) ユーザ・アプリケーション・ソフトウェアの機能 | 1984年 8 月 |

—実現のための指針—

研 究 論 文

画面指向通信用簡易プロトコルVTSSのテストシステムおよび プロトタイプ端末システム

Test and Prototype Systems for a Screen-Oriented Simplified
Communication Protocol VTSS

学術情報センター 大山 敬三*

要 旨

東京大学文献情報センターにおいて画面型アプリケーションに適用する小規模端末システム用通信プロトコル (VTSS プロトコル) の開発を1985年に行い、これに基づいてホストおよび端末システム用の通信ソフトウェアの開発を行い、学術情報センターにおいてひきつづきサービスを行っている。この開発に際し、ミニコンピュータを用いてホストおよび端末のシミュレータを開発し、これらのシステムのテストおよびデバッグに用いた。本論文ではこのシミュレータについて、その構成と利用方法について述べる。また、本プロトコルの妥当性と実現性を検証し、より使いやすい端末システムのユーザインタフェースのあり方を検討するために端末シミュレータを発展させ、プロトタイプ端末システムの開発を進めている。この基本的構成と考え方についても述べる。

Abstract

A VTSS protocol, a new simplified communication protocol applicable to screen-oriented applications for small scale terminal systems, was designed at the Center for Bibliographic Information, University of Tokyo in 1985. Communication software for host and terminal systems was developed based on the protocol and has been in operation as a part of the online shared cataloging system of the National Center for Science Information System. The author developed a VTSS protocol simulator, which is able to operate both as a host tester (i.e. a terminal simulator) and as a terminal tester (i.e. a host simulator), on a mini-computer with UNIX¹ OS for the purpose of testing and debugging of these systems during development. In this paper, the configuration and uses of the simulator are described. In order to confirm and improve the VTSS protocol and to make a user interface model of a useful terminal system, a prototype terminal system is under

*Oyama, Keizo : National Center for Science Information System
¹UNIX is a trade mark of AT&T Bell Laboratories

development. The basic concept and configuration of the prototype system are also discussed.

1. はじめに

学術情報センターおよびその前身である東京大学文献情報センターを中心に形成が進みつつある図書館ネットワークでは、当初、大学間ネットワーク（N-1ネットワーク）を基礎として異機種計算機間ネットワーク上で動作する画面型のアプリケーションを開発した。N-1ネットワークの開発には日本の主な計算機メーカーが参画しており、ノードにはさまざまな機種の計算機が加わっているが、N-1ネットワークがホスト-ホスト間通信用に開発されたことから、これらはいずれも中・大型計算機であり、またプロトコル自体も複雑である。このため小規模な計算機システムではN-1ネットワーク用の通信ソフトウェアを開発および実行することは開発量および性能の面で困難である。一方、通常の大学図書館では上述のような中・大型計算機を導入することは容易ではない。したがって、学術情報センターの図書館ネットワークを拡大するためには小型計算機でも開発および実装の容易な簡易型の画面転送プロトコルが必須である。

このような背景から VTSS プロトコル（Virtual screen transfer protocol using TSS connection: TSS 接続による仮想画面転送プロトコル）の設計を行い、学術情報センターホストシステム用のネットワークソフトウェアの開発を行うと同時に、主な計算機メーカーに対して端末システム用のネットワークソフトウェアの開発支援を行ってきた。この結果、1986年4月からホストシステムでの VTSS 方式によるサービスが開始され、数種類の端末システムがすでにセンターと接続されて図書館業務に供されている。

なお、VTSS プロトコル自体は図書館ネットワーク向けに限らずある程度の汎用性をもっているため、今後新たなサービスを展開するにあたり画面型のユーザインタフェースを実現する際にも有効である。

VTSS プロトコルは新規に開発されたため、ホストおよび端末システムの開発過程では、プロトコルの実現の正しさが確認されたシステムが存在せず、開発されたシステムのテストすらできない状況にあった。そこで、まずホストシステムをテストする目的で端末シミュレータを作成し、次に、各メーカーにより開発される端末システムをテストする目的でホストシミュレータの機能を拡張した。

一方で、VTSS プロトコルそのものの検証と改良の目的でセンター独自に VTSS 端末システムのプロトタイプの開発を進め、当初の目的を達するとともに、さらに使い勝手の良い端末システムのユーザインタフェースを検討するために改良を進めている。

以下では VTSS シミュレータおよび VTSS 端末プロトタイプシステムの作成方針と構成の概要について述べることにする。

2. VTSS プロトコルシミュレータの機能と構成

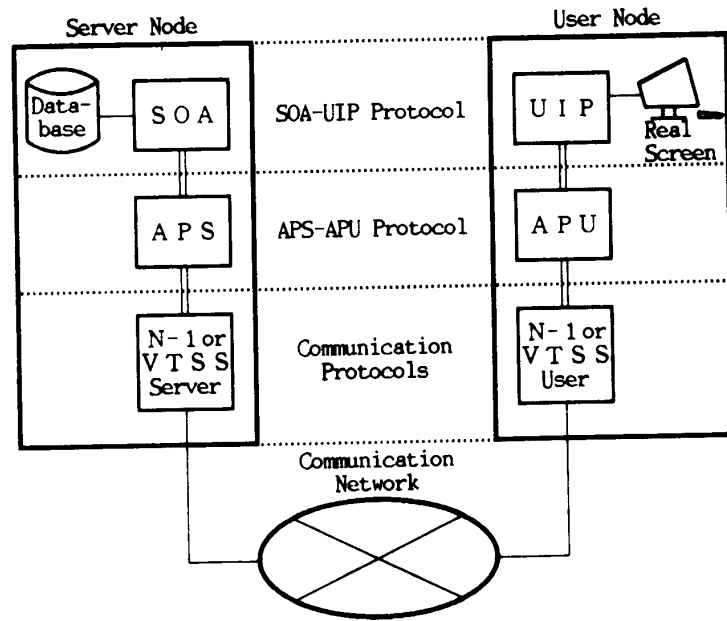


図1. 画面型アプリケーションにおけるVTSSプロトコルの位置付け

2. 1. VTSS プロトコルの概要

VTSS プロトコルの詳細に関しては参考文献1)にて述べており、ここではその概略を述べるにとどめる。

画面型アプリケーションにおけるVTSSの位置付けは図1に示すとおりである。画面型アプリケーションはサーバとユーザにまたがって存在し、それぞれSOA(Screen Oriented Application)、APS(APplication Server)とUIP(User Interface Application)、APU(APplication User)に機能分けして捉えることができる。APS-APUレイヤは仮想画面の転送を行うためのプロトコルを実現する部分であり、SOAはこの仮想画面から検索条件などを抽出してデータベース操作をおこない、UIPは送られてきた仮想画面を実端末の画面に展開し、ローカルエディット機能を提供する部分である。

VTSSはこれらの下位のレイヤであり、N-1ネットワークをベースとするシステムにおける拡張NVT(eXtended Network Virtual Terminal、以下XNVT)の機能に相当する部分を実現する。

VTSSは具体的には以下の五種類の機能に大別される。

- (1) 文字コードの16ビット-7ビット変換機能
- (2) 同一文字列の圧縮機能
- (3) 画面モードと行モードの仮想的なリンクの多重化機能
- (4) 送信権制御機能
- (5) 割り込みの発行および同期機能 (VTSS ユーザのみ)

このうち、(1)~(3)はほぼサブレイヤをなしている。(1)は上位レイヤのXNVT文字コード(16ビットコード)と下位のTSS接続用の文字コード(7ビットコード)を変換するものである。(2)は仮想画面中に多数存在すると考えられる空白文字を効率的に伝送することを目的に設けられた機能である。(3)は通常のTSSセッションと画面型セッションを同一のTSS接続の上で実現するための機能であ

る。(4)は学術情報センターホスト計算機がその実現上の制約から半二重動作となるため、正しく通信を行うために必要となる機能である。(5)は異常処理機能に属するものであり、通常の無手順端末ではブレーク信号として処理されているものを、画面型アプリケーションにおける同期の確保のために設けられた機能である。

以上のように、VTSS では XNVT コードの透過性の確保と画面転送の効率化、画面型アプリケーション同士のセッションの制御を行うものである。

なお、VTSS 端末システムでは当然、VTSS プロトコルの処理だけでなく、APU、UIP の機能も実現しなければならない。

2. 2. VTSS プロトコルシミュレータの機能

VTSS プロトコルシミュレータはホスト及び端末システムの画面型アプリケーション全体のうち、VTSS プロトコルレイヤの実現方法のテストとデバッグを行うことを目的としており、以下のような形態で利用される。

(1)学術情報センターホストシステムにおける VTSS 通信ソフトウェア (以下 VTSS サーバ) のテストおよびデバッグ (VTSS 端末シミュレータとして)

(2)図書館ユーザシステムにおける VTSS 通信ソフトウェア (以下 VTSS ユーザ) のテストおよびデバッグ (VTSS ホストシミュレータとして)

(1)については、すでに VTSS サーバの開発が完了した現時点ではほぼその目的を果たした。(2)については、初期の端末システムの中には開発を一応完了したものもあるが、今後とも新たなメーカーあるいは機種において端末システムの開発が行われることも予想され、本シミュレータの有効な利用が必要である。

VTSS プロトコルシミュレータは単一のソフトウェアではなく、実際に相手とやりとりするメインのソフトウェアと、これが用いるさまざまなデータを作成したり、あるいは採取した通信のトレースデータ等を解析したりするための付随的なツールとから構成される。

VTSS プロトコルはホストと端末 (あるいはサーバとユーザ) が必ずしも対等でない (すなわち非平衡型である) ため、シミュレータの機能もホストシミュレータと端末シミュレータとでは幾分異なるのではあるが、本シミュレータの実現にあたっては双方の機能を含んだスーパーセットとして設計を行った。VTSS プロトコルそのものの処理の機能としては以下のようなものが必要である。

- 1) 文字コードの16ビット-7ビット変換機能
- 2) 同一文字列の圧縮・伸張機能
- 3) 割り込みの発行機能

VTSS プロトコルシミュレータは単純にプロトコルを正しく実行すればよいわけではなく、意図的に誤った動作をさせることも可能でなければならない。このため、VTSS の機能のうち、仮想的なリンクの多重化機能と送信権制御機能は本シミュレータには組み込まれていない。

このほかに、シミュレータとしての機能としては以下のようなものが必要である。

- 4) コード変換および圧縮・伸張の各ステップにおける中間的データのトレース機能
- 5) コード変換および圧縮・伸張の任意のステップからのデータ送出機能（故意に誤ったデータを送出するような目的で使用する）
- 6) 文字として見えない制御文字を視覚化して表示・出力する機能
- 7) 漢字などを漢字として、あるいはそのコードに対応した文字列として表示・出力する機能
- 8) 画面モードと行モードのデータを区別しやすい形式で表示・出力する機能
- 9) APU および UIP あるいは APS および SOA の動作を人間がシミュレートできるようなユーザインタフェース機能
- 10) APU および UIP あるいは APS および SOA の動作を他のソフトウェアがシミュレートできるようなインタフェース機能

VTSS プロトコルシミュレータはあくまでも VTSS 機能の確認のために用いられるものであり、APU および UIP 自体の機能はこの中には含まれていない。

2. 3. VTSS プロトコルシミュレータの構成

2. 3. 1. 実現環境

VTSS プロトコルシミュレータの実現にあたっては次のような点を考慮して開発および実現環境となる計算機や通信機器を選択した。

- (1)通信制御については既存の装置、ソフトウェア等が利用可能で、かつ充分柔軟性を持っており、アプリケーションソフトウェアからすべての制御が行えること
- (2)全二重通信の取扱いができるような機能がサポートされていること
- (3)トレースデータなどの各種情報の加工・出力用のツールが完備していること
- (4)短時間で誤りのないソフトウェアの開発を行うために優れたソフトウェア開発環境を持つこと

VTSS プロトコルシミュレータは VTSS サーバおよびユーザシステムのソフトウェアの開発に先立ち正しい動作が確認されていなければならない、最終的にあるべき姿を念頭に置きながらも短期間で実現するための方式を強く意識して設計されている。

開発のベースとなる機器構成として以下のようなものを用いることとした。

- PAD (Packet Assembly Disassembly)装置

無手順端末用の RS232C 回線を X.25 に変換してインタフェースをとる装置。X.3、X.28、X.29の規格に従う。

- ミニコンピュータ (OS は UNIX)

VTSS プロトコルの実現、トレースデータの採取、データの加工・出力等を行う。UNIX はさまざまな規模の計算機に用いられているマルチタスク、マルチユーザのオペレーティングシステム (OS) であり、そのソフトウェア開発環境の良さには定評がある。

・漢字端末

1バイト文字コードとして JIS7ビット（カナは SI/SO による）を用いる。漢字コードとして JIS-C6226、1983年版をサポート。

PAD 装置は通信網側のインタフェースとして X.25 プロトコルを実現する一方、端末側のインタフェースとしては XON/XOFF 制御の RS232C 接続を提供する装置であり、市販の製品が手にはいる。PAD の制御用のさまざまなパラメータは PAD パラメータとして端末側から設定可能である。全二重通信が可能であり、端末側回線の伝送速度も可変である。

ミニコンピュータは通常柔軟な利用が可能で通信系も扱いやすくできている。PAD との接続を考えた場合、PAD 側回線の各種パラメータをアプリケーションから制御可能であることが必要であるが、UNIX はこの要求を満たすに十分なインタフェースを提供している。通常、ハードウェア上は全二重通信が可能であるが、UNIX（特に UNIX SYSTEM V 系）はアプリケーションに対して非同期の全二重通信機能をサポートしていない点が問題である。しかしこれは UNIX のマルチタスク機能を用いることで対応可能であり、他に適当なシステムも見当たらないため UNIX を搭載したミニコンピュータが適当であると判断した。特に UNIX 上にはさまざまなソフトウェアツールが用意されており、データの取扱いが容易であると同時に、ソフトウェア開発環境としても優れているため、今回のような緊急性のあるソフトウェア開発には最適である。

漢字端末では、VTSS プロトコルシミュレータには UIP が含まれないことから画面型の表示機能は不要であり、ライン型での JIS 第1および第2水準の漢字表示機能があれば充分である。JIS 外字（EXC 文字など）の表示機能もあることが望ましいが、ここでは VTSS プロトコル自体に主眼があるので、VTSS プロトコルシミュレータではこの機能は省略した。

開発のベースとなるソフトウェアとしては、UNIX 自体、あるいは各種ソフトウェアツール群のほかにも、C 言語処理系やシンボリックデバッガ等が非常に有効であった。

2. 3. 2. ハードウェア構成

前節に述べたような実現環境において VTSS プロトコルシミュレータを実現するにあたってのハードウェア構成を図2に示す。必要なハードウェアは以下の通りである。

・PAD 装置

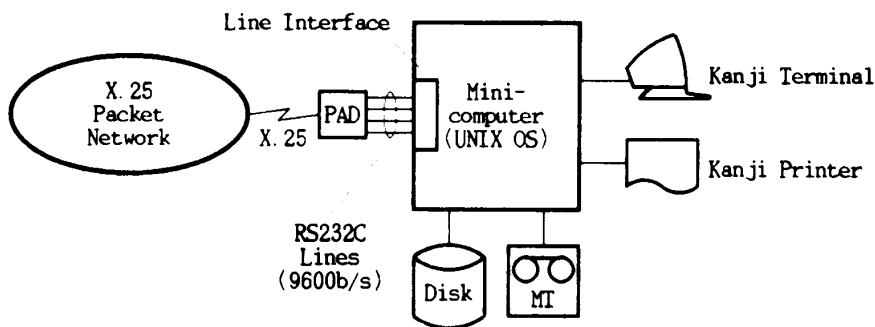


図2. VTSSプロトコルシミュレータのハードウェア構成

X.25回線速度は9600bps であり、4本のRS232C回線を論理リンクとして多重化する
 (例. 米国 Dynatech Packet Technology, Inc.社製 MULTI-PAD.25、米国 MEMOTEC DATA INC.社製 MPAC SP/830など)

- ミニコンピュータ本体
 CPUとしては最新の32ビットマイクロプロセッサを搭載し、4MB以上の主記憶と8KB程度のキャッシュメモリを持つ
 (例. 米国 DEC 社製 MicroVAX-II、住友電気社製 U-Station E/20など)
- RS232C回線インタフェース装置 (300~9600bps、可変)
 9600bpsのデータを無手順で受信可能である
- ディスク装置 (プログラム、データ格納用)
 容量約400MBのドライブを2台持つ
- 磁気テープ装置 (サーバホスト計算機とのデータ交換用)
 テープ速度125inch/s、6250bpiの性能を持つ
- 漢字端末
- 漢字プリンタ

2. 3. 3. ソフトウェア構成

VTSS プロトコルシミュレータのソフトウェア構成を図3に示す。

- UNIX カーネル
 UNIX OSの核であり、マルチタスク制御、メモリ管理、ファイル管理、デバイス (通信回線、周辺装置等) 管理等を行う。
- プロトコルシミュレータ本体

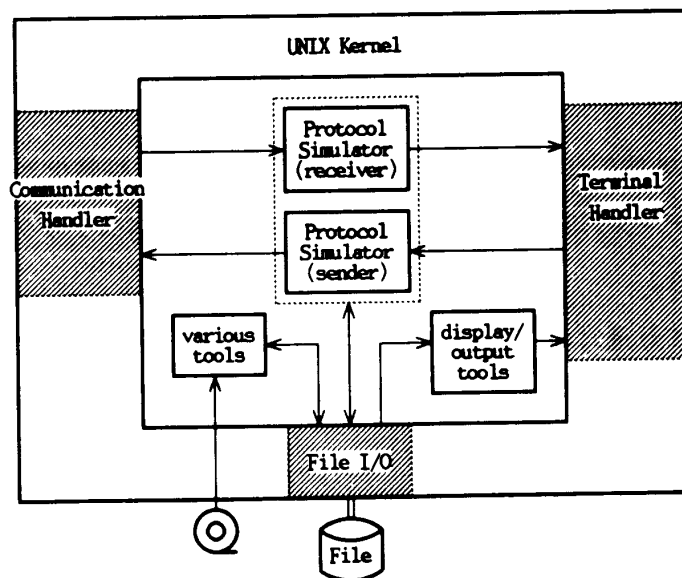


図3 . VTSSプロトコルシミュレータのソフトウェア構成

実際に回線を経由して接続相手（端末シミュレータのときはサーバホスト、ホストシミュレータのときは端末システム）と相互に通信を行う。UNIX では単体のアプリケーションでは非同期の全二重通信を扱うことができないので、送信と受信を別々のプロセスに分けて全二重動作を実現している。送信データは送信単位ごとに分割して別々のファイルに格納されており、個々のデータの送信の指示はオペレータ（通常は人間）が行う。

- コード変換・圧縮／伸張ツール

磁気テープでもらったサーバホスト上のサンプル画面等のデータや、実際のセッションのトレースデータを加工するためのツール群であり、主に異常データ等を作成するために用いる。

- VTSS データ表示・出力プログラム

実際のセッションのトレースデータを画面上に表示、あるいはプリンタに印字するためのプログラムであるが、プロトコルシミュレータ本体の各種の処理と整合性を保つために、実体はプロトコルシミュレータ本体と同一で、コマンド名の違いによりそれぞれの動作をするように実現されている。

- UNIX ソフトウェアツール

UNIX システムにもともと備え付けられている各種ソフトウェアツール群である。

2. 3. 4. プロトコルシミュレータとしての機能の実現方式

プロトコルシミュレータの機能としては、2. 2. 節にも述べたように、VTSS プロトコルの機能以外に、シミュレータであるがゆえに要求される機能がある。これらは基本的にはプロトコルシミュレータ本体に備えられるものであるが、ここに含まれていないような個別の要求事項はシミュレータ本体とは切り離された別個のソフトウェアツールとして実現されるべきである。したがって、本節ではシミュレータ本体の実現方式について述べることにする。

ソフトウェアの構成を図4に示す。送信と受信はそれぞれ別プロセスとして動作することにより全二重動作を実現している。

VTSS プロトコルの各機能はモジュールに分割され、実現上も階層化されているので、VTSS コードにおけるトレースデータが記録されていれば、プロトコルの各ステップにおけるトレースデータは容易に再現可能である。

データの送信に関しては通常の VTSS プロトコルによるデータの変換を行う経路と行わない経路をコマンドにより指定できるようにすることにより、予め用意した任意のデータをも送出可能である。

受信したデータ、あるいは必要であれば送信したデータについても、文字として見えない制御文字を視覚化して表示したり、漢字などの文字を漢字として、あるいはそのコードに対応する文字列として表示・出力する必要がある。これらの各機能もモジュールに分割され、階層化されてユーザインタ

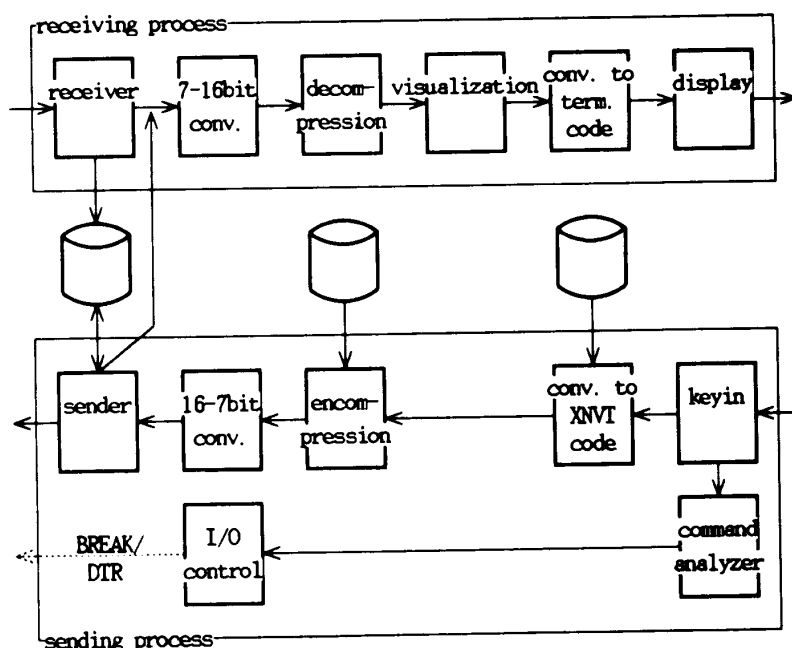


図4. プロトコルシミュレータ本体の構成

フェースに組み込まれており、それぞれの機能を有効にするか無効にするかをオプションで選択可能である。さらに、仮想画面中のデータを見やすく表示するためにフィールドの区切り文字により改行するようなモジュールも追加してある。

APU および UIP の動作を人間がシミュレートするためのユーザインタフェースに関しては、アルファベット1文字のコマンドにより、

- ブレーク信号を送る
 - X. 25の割り込みパケットを発生させる
- DTR 信号を2秒間落とす
 - X. 25の回路を切断する
- キーボード入力モードにする
 - TSS セッション (漢字を含む) を会話的に行う
- ファイル中の端末依存コードデータをプロトコルに従って変換して送出する
 - 別途キーボードから入力して作成したファイル (漢字を含む) を VTSS プロトコルに従って送信する
- ファイル中の XNVT コードデータをプロトコルに従って変換して送出する
 - サーバホストから採取した仮想画面データ等 (漢字、EXC データ等を含む) を VTSS プロトコルに従って送信する
- ファイル中のデータを変換せずにそのまま送出する
 - 特殊なデータや故意に誤ったデータを送信する

といった操作を行えるようにしてある。

APU および UIP の動作を他のソフトウェアがシミュレートするためのインタフェースに関して

表1．シミュレータの規模

	ソース	オブジェクト
シミュレータ本体	1000行 (17643KB)	19456KB
コード変換、圧縮・伸張ツール	438行 (6577KB)	51200KB

は、UNIX のプロセス間通信機能 (パイプ) を用いることにより実現する。

このような構成により、VTSS プロトコルシミュレータの作成をプログラミング言語Cを用いて行い、周辺のツール群を含めて約2人月の開発工数で一応の完成をみた。プログラムの規模は表1に示す通りである。

2. 4. VTSS プロトコルシミュレータの応用

端末シミュレータ (すなわちホストテスタ) として用いる場合は図5のような形態となる。ホストシステムのX.25インタフェースはすでに稼働しており十分に信頼できるため、最初から DDX パケット交換網に接続してテストすることができた。

VTSS プロトコルシミュレータには予め、N-1ベースの画面型アプリケーションのセッションの受信トレースデータをホスト上で採取して移し、データ加工用のツールを用いて送信単位ごとのファイルに格納しておく (①、②)。

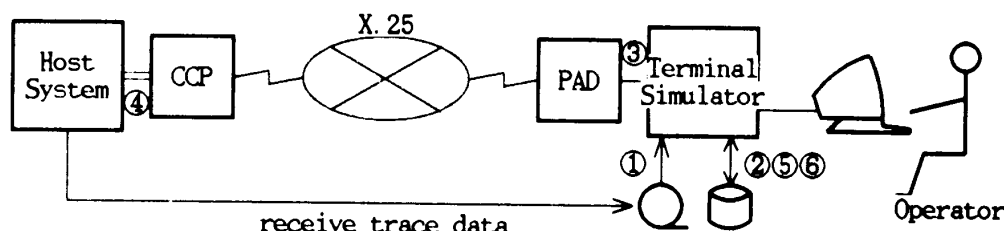


図5．端末シミュレータとしての使用

シミュレータ本体を起動し、キーボード入力モードにしておいて、まず PAD に対してコマンドを入力して回線の接続を行い (③)、つぎにホストシステムにログオンして画面型アプリケーションの起動コマンドを入力する (④)。

ホストから開始のデータが送られてくるのを画面上で確認してから、キーボード入力モードを終了し、用意してあったファイルを順次送信してゆく (⑤)。1回のセッションが終了したら、再びキーボード入力モードにしてログオフし、シミュレータ本体を終了する。

この過程の送受信データはファイルに格納されているので、あとは各種ツールを用いてジックリ確認を行えばよい (⑥)。

この一連の操作例を付録に示す。

ホストシミュレータ (すなわち端末テスタ) として用いる場合は図6のような形態となる。端末シ

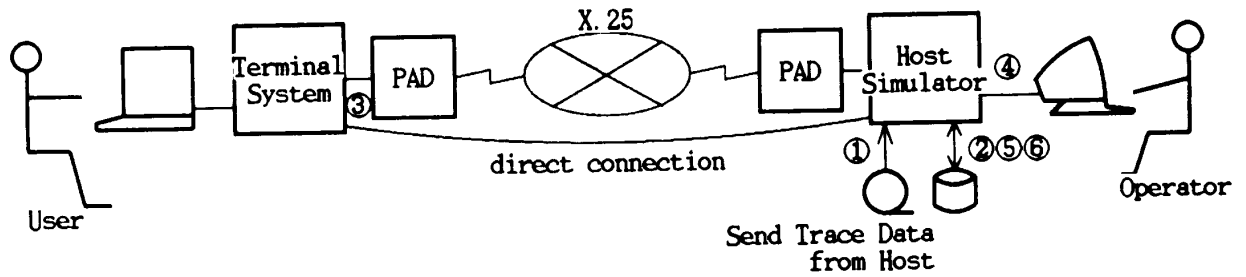


図 6 . ホストシミュレータとしての使用

システムが PAD を用いている場合はシミュレータと端末システムを直結する形で相当部分までのテストとデバッグが可能である。X.25 インタフェースの信頼性が充分でない場合は本シミュレータとは別の手段によりこの確認を行っておくことが必要である。

端末シミュレータの場合とは逆に、予め、N-1 ベースの画面型アプリケーションのセッションの送信トレースデータをホスト上で採取して送信単位ごとのファイルに格納しておく (①、②)。

シミュレータ本体を起動しておいてから、端末システムを起動し、ホストに対する接続、ログオンおよび画面型アプリケーション起動操作を行う (③)。シミュレータではこれらの受信の確認を画面上で行いながら、メッセージやプロンプトを返す (④) (キーボードから行っても良いし、ファイル上に用意しておいてこれを送信しても良い)。後は端末シミュレータと同様に用意してあったファイルを順次送信してゆく (⑤)。端末システムからログオフ操作が行われたことを確認してシミュレータ本体を終了する。

動作の確認は端末シミュレータの場合と同様に行う (⑥)。

いずれにおいても、プロトコル自体は通信上の規約を定めているだけであり、実現上はさまざまな自由度があるため、機械的にそのインプリメンテーションが正しいか否かを判定することは現状では困難であり、最終的な確認は人間の目に頼らざるをえない。

3. VTSS 端末プロトタイプシステムの基本構成

VTSS プロトコルの妥当性の確認と細部の改善を図り、小規模計算機向けの VTSS 端末システムの実現モデルを作成するために、第 2 章で述べた VTSS プロトコルシミュレータを元にして、UNIX 搭載のミニコンピュータ上に VTSS 端末プロトタイプシステムの試作を行っている。

UNIX マシンを選択した理由は UNIX の以下のような特徴による。

- 1) マルチタスク機能
- 2) アプリケーションから利用可能なメモリ管理機能
- 3) 容易で柔軟なファイル操作性
- 4) 通信ポートの細部にわたる制御性
- 5) ハードウェアおよび OS の核における全二重通信機能
- 6) 効率的なソフトウェア開発環境

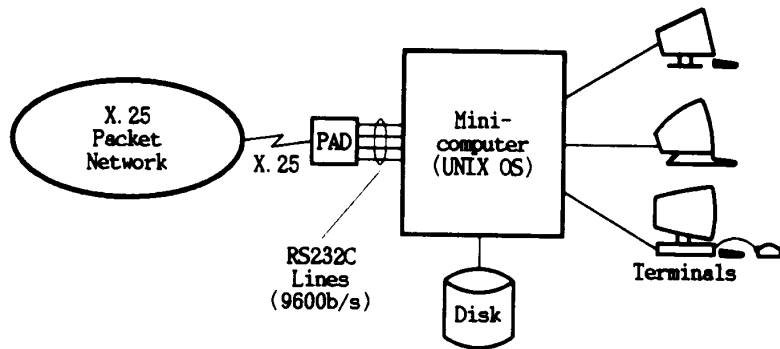


図7. VTSS 端末プロトタイプシステムのハードウェア構成

開発効率の観点からパーソナルコンピュータは採用しなかったが、実用向けには有用であり、これらのための端末システムとしてのソフトウェアの開発が行われることは十分に意味があると考えられる。

図7に VTSS 端末プロトタイプシステムのハードウェア構成を示す。X.25通信ポートを備えたUNIX システムもいくつかあるが、プロトタイプシステムであることから、前章のシミュレータと同様の理由により PAD を用いている。ユーザ端末は JIS-C6220 (1 バイトコード) および JIS-C6226 1983年版 (2 バイトコード) で定義された文字集合だけでなく、EXC 文字等の JIS 外字をも表示できることが要求される。

図8に VTSS 端末プロトタイプシステムのソフトウェア構成を示す。VTSS プロトコルレイヤはモジュール化されている。APU および UIP は仮想画面の組立、編集等の操作をする。UNIX の termcap 機能 (端末属性データベース) を利用することにより実画面ハンドラ (physical screen

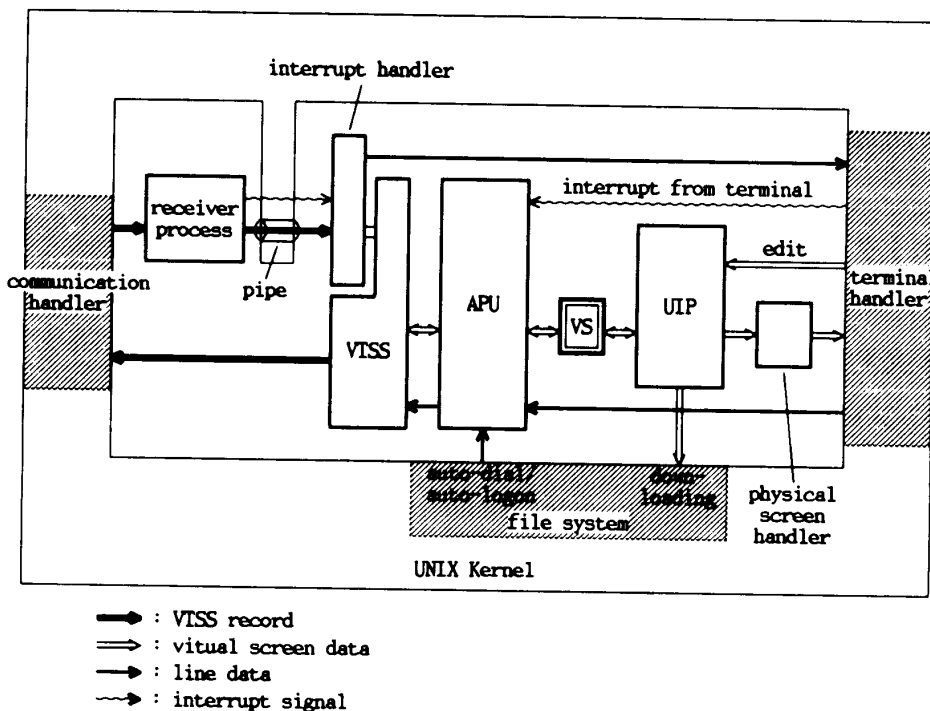


図8. VTSS 端末プロトタイプシステムのソフトウェア構成

handler)はさまざまな種類の端末を制御することが可能である。フィールド挿入、削除、拡張等のローカルな仮想画面の集編機能は UIP で実現される。

将来的な互換性を確保するため、プロトタイプシステム内部の文字コードは AT&T UNIX Pacific Co.による日本語 UNIX に関する勧告に基づいている。

仮想画面をダウンロードする機能は UIP に組み込まれる。自動接続あるいは自動ログオン機能を実現するための初期データファイル送信機能は APU に組み込まれている。

VTSS プロトコルの中ではローカルエディット中 (UIP 動作中) に到着するデータの受信および処理の実現が最も難しい。ハードウェアおよび OS の核には全二重通信機能があるにもかかわらず、UNIX (特に UNIX SYSTEM V) はアプリケーションに対してこの機能を利用するための標準的なインタフェースを提供していない。この問題には2つの解決方法がある。1つはシミュレータの時と同様に送信と受信を別プロセスにする方法であり、もう1つはキーイン待ちの時に通信ポートにデータが到着しているか繰り返し見に行く方法である。UNIX がマルチタスク、マルチユーザのシステムであることを考慮すると CPU の効率上、第2の方法はとれない。ただし、シングルユーザ、シングルタスクのパーソナルコンピュータの場合は CPU を占有しても構わないため、この方法も可能である。

第1の方法においては、2つのプロセスのうちのどちらか一方が責任を持って仮想画面を扱わなければならない。今回のプロトタイプシステムの実現方法においては受信処理は割り込み駆動としており、仮想画面の操作、表示、編集等は通常の送信プロセスで取り扱っている。この実現方法の中で用いている主な UNIX の機能はリード・ライト、I/Oコントロール、パイプ、シグナル、メモリ割り付け等である。このプロトタイプシステムは UNIX SYSTEM V でも UNIX 4.2BSD でも動くように設計されている。

UNIX SYSTEM V には共有メモリやセマフォの機能があり、仮想画面を共有メモリに置くことにより、受信プロセスが仮想画面を組立て、送信プロセスが表示および編集を行うような実現方法も採りうる。しかし、実画面は送信プロセスが管理する必要があり、なんらかの形で割り込み駆動の動作を導入せざるをえないため、構成が著しく簡潔になる可能性は少ない。

現在、基本部分については完成し、VTSS プロトコルの妥当性の確認と細部の改善の目的はほぼ達成された。今後、端末システムとしての機能の拡張・強化を行って、使い勝手の良いユーザインタフェースの研究を通して小規模計算機向けの VTSS 端末システムの実現モデルの構築を進める予定である。

4. おわりに

本論文では VTSS プロトコルの実現にあたっての支援システムを中心に、その方式と構成について述べてきた。現在は学術情報センターのオンライン総合目録システム用通信方式の一つとしてこのプロトコルが活用されている。すでにいくつかのメーカーにより、オフィスコンピュータあるいはパー

ソナルコンピュータを用いた端末システムも開発され、図書館業務に供されている。

学術情報センターでは今後、図書館およびエンドユーザ向けのさまざまなサービスを展開してゆく計画であるが、これらについてもパーソナルコンピュータ等の小規模計算機を用いた VTSS 端末システムを利用してユーザフレンドリなインタフェースを提供するための基盤ができたことになる。

一方、センター内の業務として総合目録データベースの構築を継続して行っていくわけであるが、このためには目的に特化した端末システムを実現することにより作業の効率を高めることが重要な課題となる。メーカーの製品となると改造は容易でなく、このような用途には不適切であるが、プロトタイプ端末システムの拡張・強化によりこれに対処することが可能になると期待される。

今後ともさまざまなメーカーにより、それぞれの特長を活かした端末システムが開発され、種々の利用者の要望に応えられるようになることを期待するとともに、そのための支援環境の充実を図っていくことが必要である。また、これらのシステムの実現上のモデルとなるよう、常にプロトタイプ端末システムの改善を図ることも必要である。

謝辞

本論文で述べた VTSS プロトコルシミュレータに関しては、各種のデータの準備、およびホスト、端末システムとの接続実験等において日立製作所の水村氏から多大の協力を得た。また、回線などの設備面においては学術情報センターの渡辺博氏、大日方一男氏の協力を得た。VTSS 端末プロトタイプシステムに関しては、その構成手法をまとめるにあたって学術情報センターの安達淳博士、橋爪宏達博士から多数の有益な助言を得た。ここに記して謝意を表す。

参考文献

- 1) 安達淳、橋爪宏達、大山敬三:「TSS 接続による仮想画面転送 (VTSS) 方式」学術情報センター紀要、第 1 号、1987年。
- 2) 安達淳:「画面志向アプリケーション向けネットワークプロトコルの開発」東京大学文献情報センター紀要、第 1 号、1985年 8 月。

付録. 端末シミュレータの操作例

```

% dd if=/dev/rmt0 of=mt/user.2 ibs=12b
7+0 records in
84+0 records out
% cutrec mt/user.2 send/user.2
% vtss -o trace +echo
b --- send break
d --- clear DTR
f <filename> --- send file (without conv.
x <filename> --- send file (with conv. fr
r <filename> --- send file (with conv. fr
k --- send from keyboard
q --- quit
? k
input EOF char: ^[
***** KEYBOARD *****
^M
^J^M
^Jready^M
^J*ssget_ 22:00^M
^M
^J^M
^J^M
^Jready^M
^J*c 3571150^M
^Jcom 3571150/^M
^J^J^e^eENTER TERMINAL TYPE^M
^e^e^e^e^e^J^e^e^GJIS83^M
^J^e^M
^e^e^e^e^e^eENTER LOGON^J^e^M
^e^e^e^e^e^e^Glogon z12345/password s(400
^JJGP300I Z00024 ,LAST ACCESSED DATE=86.(
JET10065I TSS Z00024 STARTED TIME=15:04:14
>> YOUR ACCOUNT (A,$94977,$5023,$100000)
>>^Gonline^M
^J
^T
S
^T
^M
^G
^[
? x S/user.2.1
***** FILE: S/user.2.1 *****

^T
S
^T
^M

^T
CO
^T

^T
TO F101X
^]OA
~00000 B業務選択
~00100 A (教育用)
~00200 B1.和図書目録登録 11.和図書目録検索 21. I L L
~00300 B2.洋図書目録登録 12.洋図書目録検索
          13.和雑誌目録検索
          ...雑誌目録検索 99.業務終了

```

MTからデータを取り込んで
ファイルに格納する

VTSSシミュレータを起動する

コマンド・メニュー

キーボードから入力を指定

PADパラメータの設定

X. 25回線接続

ホストシステムにログオン

目録システム起動

UIP起動コマンド (受信)

キーボード入力終了
ファイル送信

UIP起動レスポンス (送信)

クリアコマンド (受信)

以下、仮想画面データ

(中略)

```

^~00800 B
^~00900 B  ◎ 時差更新および教育用フー、
^~01000 B  後に行います。
^~01100 B  ◎ 8月4日より参照系データベースに、TRC/MARCを追加いたし
^~01200 B  TRC/MARCを特定して検索する場合、FILEフィールドに指定する識別名は'
^~01300 B
^~01400 B
^~01500 B
^~01600 B
^~01700 B      ジャパン・マーク共同利用実験 (JP/MARC 著作権者：国立国会
^~01800 B
^~01900 B
^~02000 B
^~02100 B
^~02200 B
^~02300 B

```

```

^T
^T
DO
^T
^M
^G

```

```

? x S/user.2.2 _____ ファイル送信
***** FILE: S/user.2.2 *****

```

```

^T
TO F101X
^]1B
^^00000 ^_11
^T _____ 送信データ
^T
NO
^T
^M

```

```

^T
CO
^T
^T
TO B401S
^]0A
^^00000 B和図書書誌検索・簡略表示
^^00100 A (教育用)
^^00200 A
^^00300 A
^^00400 A
^^00500 A
^]1B
^^00000 B>: ^_

```

(中略)

```

^~01800 B
^~01900 B
^~02000 B
^~02100 B
^~02200 B
^~02300 B

```

^T

^T

DO

^T

^M

^G

? x S/user.2.7

***** FILE: S/user.2.7 *****

^T

TO F101X

^]1B

^~00000 ^_99

^T

^T

NO

^T

^M

^T

Z

^T

^M

^J^M

>>^G

UIP終了コマンド (受信)

? k

input EOF char: ^[

***** KEYBOARD *****

^M

^J>>^Glogoff^M

^JJET11061I USER COMMAND PROFILE BEING STO

* JSN=120229 CT=00:00:03 ET=00^@03 MEMO

* FILE=(6251KB:\$1) ACCT=(A,\$73,\$95050,\$49

JET10080I Z00024 TSS SESSION ENDED TIME=15

^M

^Jclr occ^M

^J*

^[

? q

*

キーボードから入力を指定

TSSセッション終了

キーボード入力終了
VTSSシミュレータ終了

解 説

大量情報の蓄積技術

Technology for a Large-Scale Data Storage

学術情報センター 安達 淳*

要 旨

本論文は、画像データベースのような大規模データベース構築のための蓄積媒体技術の現状を調査報告するものであり、特に光学的蓄積技術に焦点をあてている。従来の磁気蓄積媒体を比較基準として参照しつつ、光ディスクの現状および将来動向を議論する。また、情報システム内における蓄積装置の機能分担の観点から、光ディスクを取り入れたシステムを設計する際、考慮すべき点についても述べる。

Abstract

As an introductory guide, this paper surveys the current state-of-the-art of storage media technology, especially focusing on optical storage technology for constructing a large database such as image databases. Based on the review of magnetic storage media, the current status and future trend of optical disks are discussed. System design considerations are also mentioned relating to the role of storage devices in information systems.

1. Introduction

Drastic progress of computer technology in the last 30 years has given us a great power of information processing. The research and development efforts in this field have been made mainly to develop high-speed processors and large capacity storage devices. As the results of the progress in semiconductor technology, laser technology and other vanguard technologies, we have already a variety of alternatives to construct an information system. On the other hand, in every aspect of social activities, the amount of

*Adachi, Jun : National Center for Science Information System

information that we have to handle is increasing rapidly in proportion to the growth of capability to manipulate a large amount of data very easily using computer systems. Nowadays, we might not get enough information necessary for research works without computer systems.

In this paper, the technical options available for constructing a large information system are surveyed, and the current status and future possibility of optical disks, which have recently become commercially available, are summarized. Section 2 summarizes the state-of-the-art of computer technology, focusing on storage devices from a technical viewpoint. In Section 3, optical storage devices are particularly examined comprehensively compared with orthodox storage technology. A document file system is also outlined as an example of practical application. In Section 4, the software technology which is used for information systems is mentioned. A variety of information sources are also overviewed relating to storage devices, and the applicability of optical disks to information sources is considered.

2. Data Storage in Computer Technology^{1),3)}

The progress in electronic instruments in general, especially storage devices, has been the major incentive to the research and development activities relating to computer technology. In this section, some considerations on storage devices are disclosed from the viewpoint of computer technology.

2.1. Overview

Storage devices are the basic components for computer systems. For the development of high performance computer system, large and fast storage devices are required. In fact, the realization of new storage device has accelerated the progress of information processing technique. We recall first the history of the progress in storage devices. The capacity of main memory of a typical mainframe* computer is increasing to approximately 10^4 times as large as that of 20 years ago, i.e., from 10^4 bits to 10^8 bits. During this period, the operation speed of memory devices has become 10^5 times faster, from the order of 10 milliseconds to the order of 10 nanoseconds. At the same time, the cost for storage becomes about 1000 times cheaper. The rate of progress is expressed in an exponential curve. Based on the innovation above mentioned, various kinds of electronic

*Mainframe means a large computer mainly used for business applications, such as IBM machines.

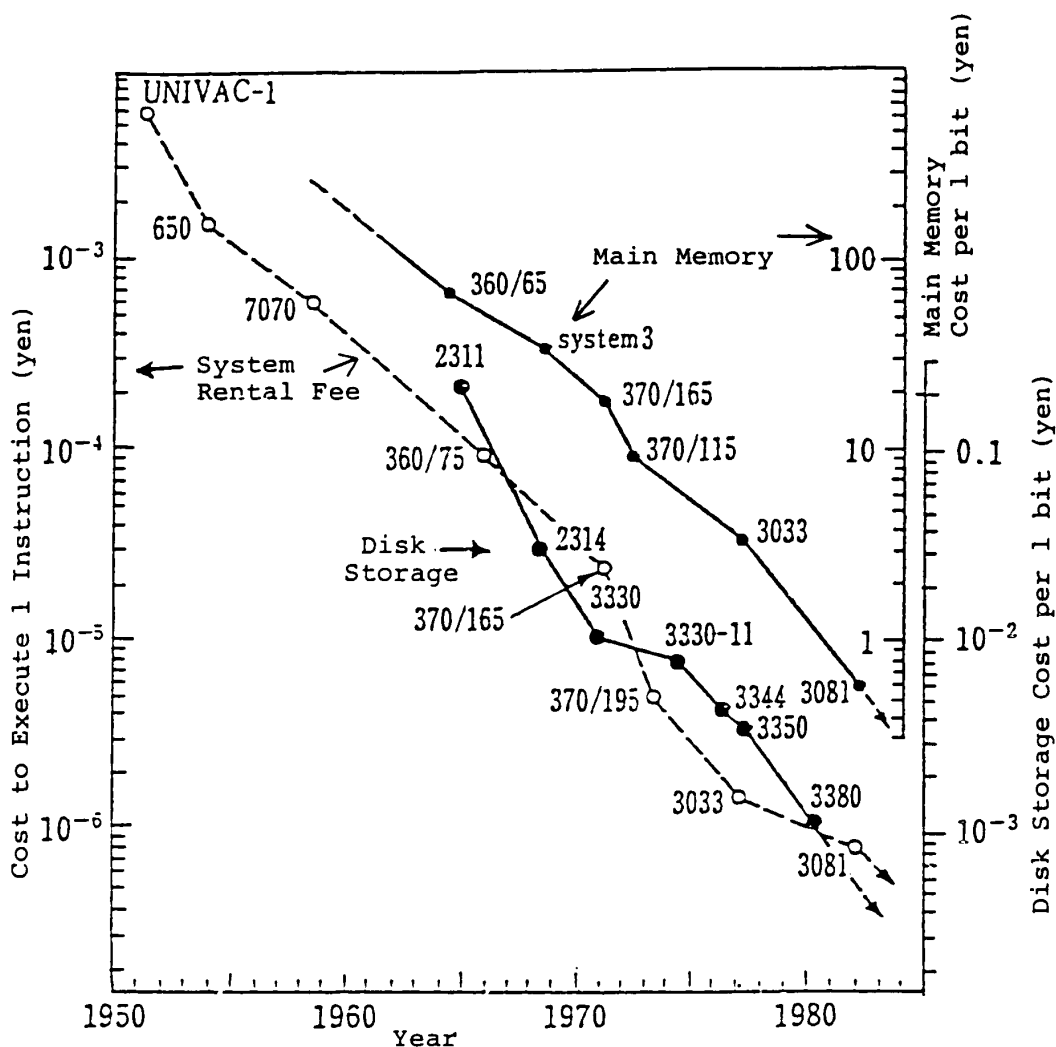


Fig.1 Cost of Computer System¹⁾

instruments have become available for a wide variety of applications using memory devices.

Fig. 1 illustrates the overall trend of computer system cost-effectiveness. The costs of disk storage, main memory and the overall cost of computer system are reduced at the rate of 1/30 in 10 years.

The leading approaches for the development of storage device are in semiconductor technology and magnetic device technology. Optical disk technology is rather new in the storage device field. Until the 1970s, the capacity of storage was the major and most strict constraint for system designers, for storage devices occupied more than half of the total system cost. However, the progress of semiconductor technology has improved the cost effectiveness of storage devices and we can now enjoy large, fast and reliable storage devices. Thus, there are now a wide variety of available options for data storage. System designers can select storage devices arbitrarily for their purpose. But a new problem

arises with the change of technical environment in designing system configurations. Such problem is discussed in Section 4.

2.2. Memory

Memory is a general concept in the field of computer technology to denote an equipment or a device to store data or programs. We can store necessary data in memories for later use, or retrieve previous data from memories. There are two types of usage in memories; one is “main memory” or “main storage,” which is used in built-in form in computer system. At present, semiconductor memories are used for main memory. Another type of memory is an “external” storage devices or a peripheral storage device such as magnetic disk, magnetic tape unit and optical disk.

The capacity of memory is counted with unit called “bit*,” which was originally introduced as a theoretical unit to measure the amount of information. Because memory handles digital information coded in binary form, information to indicate 0 or 1 is defined as a unit, i.e., “1 bit.” However another unit “byte” is much more favored to count data,

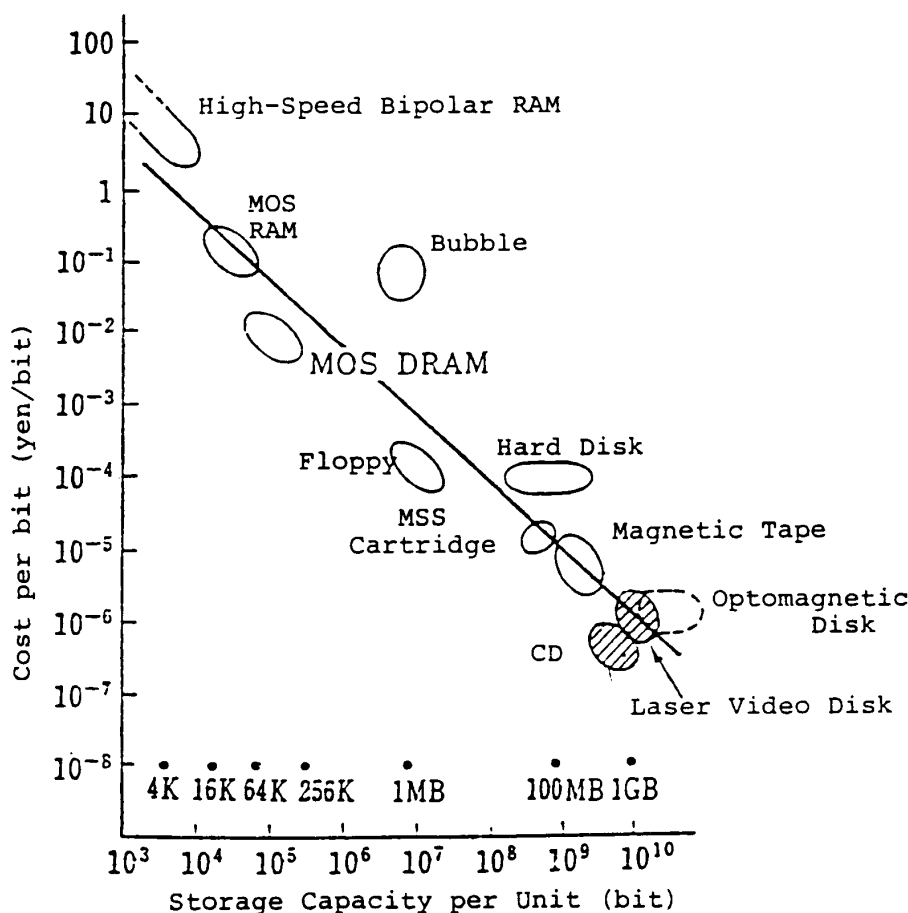


Fig.2 Storage Capacity vs. Cost-per-bit¹⁾

*“bit” is a word coming from “binary digit.”

information or memory capacity. One byte corresponds to 8 bits and one alpha-numeric character is encoded into 1 byte (8 bits) in most computer systems. Such encoding method is employed in the ASCII (American Standard Code for Information Interchange) codes used in the United States. The following notations are further used in order to measure a large capacity of storages and memories:

KB Kilobytes = 1024 bytes (i.e., 2^{10})

MB Megabytes = 1024 KB (mega means 10^6)

GB Gigabytes = 1024 MB (giga means 10^9)

The major criteria to evaluate memory characteristics are (a) the capacity per unit device, (b) the cost for unit storage (1 bit) and (c) the access time. Access time means the time necessary to get information from memory. It is very difficult to achieve both large capacity and fast access time in same one memory medium with current memory technology. Fig. 2 shows the relationship between storage capacity and cost-per-bit. Most of small capacity memories are semiconductor memories used chiefly for main storage.

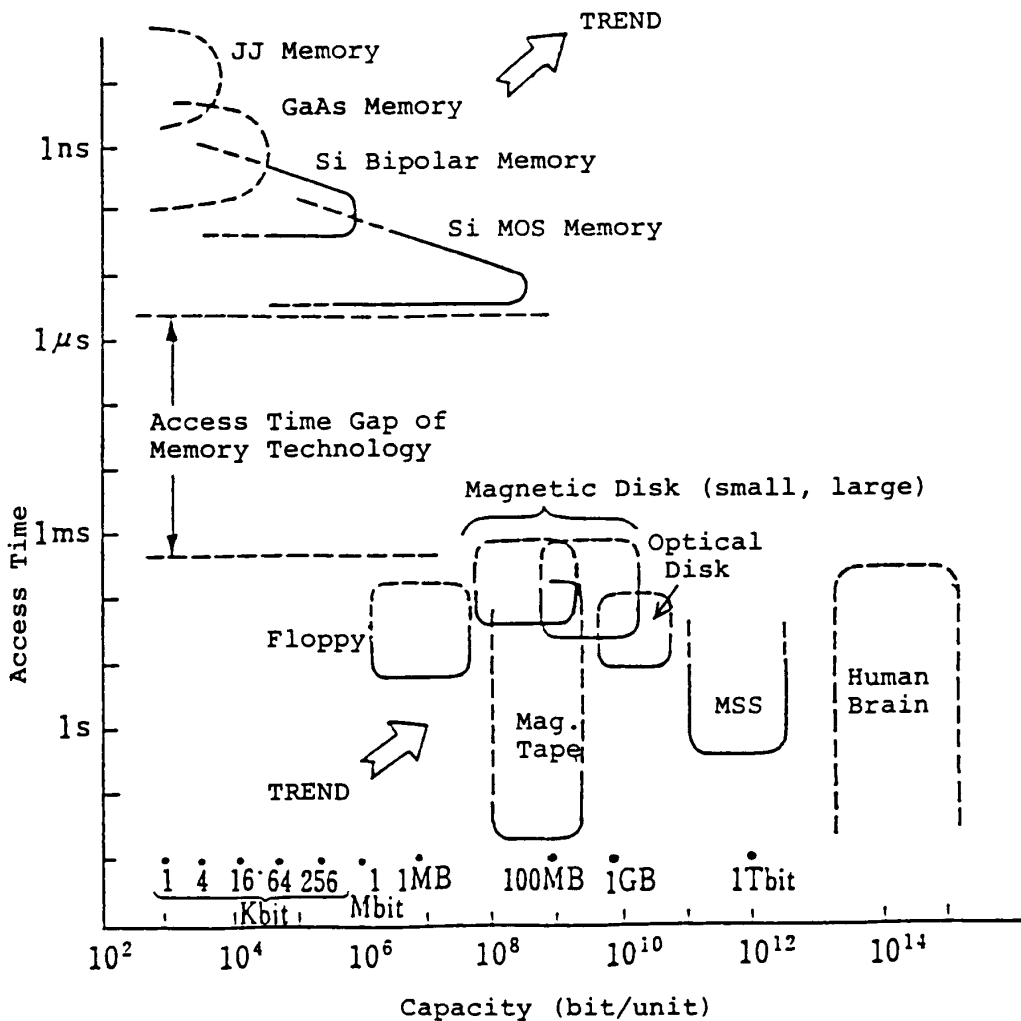


Fig.3 Storage Capacity vs. Access Time¹⁾

Larger ones are magnetic or optical devices. From logical point of view, every memory is same as a device to store binary information. However, if you use semiconductor memories for a large storage system, its price would be an astronomical figure. For example, 10 GB semiconductor memory will cost about 1 billion dollars.

At present, semiconductor memories are used primarily for main memory of computer, and even in a large system, the capacity is at most 100 MB. On the other hand, the cost-per-bit of magnetic storage devices is far cheaper than that of semiconductor memories. Such devices, which are often called “secondary store,” are used for storing a large amount of data economically. Fig. 3 illustrates this difference graphically. The gap between two groups of memories exhibits the intrinsic difference between semiconductor and magnetic technology.

The primary requirement for external storage device is the capacity. Another requirement is the “random access” capability. “Random access” means that you can get necessary information in a reasonable time whenever you want it. For example, magnetic tape units are not random access devices, because you have to rewind and read the entire tape if you get the data on the last position. Most of random access external storage devices employ rotation mechanism such as used in disks in order to realize fast access to the designated data. It takes approximately several tens of milliseconds to position the magnetic head on the designated track. It is longer than the access time of main memory which is the order of microseconds or nanoseconds.

2.3. Storage device and media

In this section, typical storage devices used in computer systems are overviewed.

Magnetic Disk

A disk is the most popular and reliable random access storage device, and is also called “direct access storage device, (DASD)” in the IBM world. A disk covers wide range of storage capacity and the maximum capacity available now is approximately 5 GB. A new type of low-cost disk, which is often called “hard disk,” covers capacity range of 10 MB to 500 MB and has become popular for small systems like personal computers. At present, we can regard magnetic disks as the best-matched storage devices to computer systems.

Magnetic Tape

A magnetic tape unit is a typical example of sequential access devices. From this reason, a tape unit is used mainly for the purpose of backup or transportation of

data. Magnetic tape of 2400 feet long stores at most 100 MB of data.

Floppy Disk

A floppy disk or diskette is the most popular storage device for microcomputers. A 5 inch or 8 inch disk in diameter is commercially available with the capacity of 1 MB. The merits of floppy disks reside in its removability and ease of handling. These points are very important in disseminating information widely. Today, hard disk becomes favored instead of floppy disks even for personal computers, because larger capacity is required in small systems.

Mass Storage System

This is often called MSS in abbreviation. This device is something like a large warehouse of cartridge tapes, each storing 50 MB of data. Cartridges are arranged in honeycomb-like containers. An automatic robot arm moves to the designated cartridge, picks it up, sets it into the reading mechanism and retrieves data from there. It takes a few seconds to access the data in MSS cartridges. The typical MSS contains 2000 cartridges, thus the capacity is about 100 GB. MSS is used mainly for the archive of data.

Table 1. Cost-effectiveness of Storage Devices
(Cost for 1 MB Storage)

High-speed bipolar RAM	20,000,000 yen
Static RAM	1,000,000 yen
Bubble memory	800,000 yen
Dynamic RAM	100,000 yen
MT and Floppy	100,000 yen
Disk	10,000 yen
MSS	500 yen

The cost-effectiveness comparison is illustrated in Table 1. All the devices above mentioned are using magnetic technology. For the last 35 years data storage devices have been dominated by magnetic technology. Magnetic technology is the most stable and reliable technology in the storage device field. Studies have been pursued to develop new devices, however, few devices were successful. As an example of exotic memory devices, CCD (Charge Coupled Device), which is used widely for image sensors of video cameras,

was not successful as memory device. Bubble memory caught researcher's interests several years ago and was expected to replace magnetic storage devices. However, in spite of a lot of research efforts, bubble memory is not yet practical due to lack of manufacturing technology of memory substrate. Research works in magnetic technology are being continued eagerly for the enhancement of recording density, so that magnetic devices are expected to be still in a dominant position in storage devices for years.

However, optical storage has another status in the field of storage devices. Several outstanding features of optical storage are much favored by computer system. Optical technology is described in detail in the next section.

3. Optical disk storage device

3.1. Technical background⁴⁾

At present, an optical disk is considered most appropriate for storage of images and documents because of its high density and large capacity. Summary of characteristics of an optical disk is as follows:

- (1) High density, large capacity and small space.
- (2) Capability to read data without contact to the disk surface.
- (3) Strength to contamination.
- (4) Removability of disks.
- (5) long life(?).
- (6) Random access capability

The characteristics described in (1) to (4) have been derived from the optical reading and recording techniques. And the resultant attractiveness resides in its cost-per-bit. In future, storage cost per bit of mass-produced optical disk will be expected to be 1/10 or 1/100 of that of magnetic disk. In the available optical disk, the capacity of one surface of 20 cm diameter is more than 10^{10} bits, which means that recording density is 100 to 1000 times higher than that of prevalent magnetic disk or floppy disk.

In optical disks, small pits (less than 10^{-6} meter) are made on a rotating disk surface to record binary information, using laser beam heating technique. Playback or reading is performed by detecting the change in strength of reflected light. Because of high track density, the major difficulty resides in the method to devise an optical head mechanism with high-quality mechanical and optical precision. Glass or aluminum is employed for disk substrate, and oxide Te is used for recording film. The optical disk above mentioned is for the *read-only* or *write-once* applications. A write-once disk is also called *DRAW*

(Direct-Read-After-Write) type disk. These types of disks are already on commercial marketplace. Current research efforts are focused on the realization of *erasable* (read/write) optical disks.

When an erasable disk is in practical use, the flexibility for applications will be greatly increased. Therefore, write-once media is now regarded as an interim step. Magneto-optical recording technique is adopted for erasable optical disk to record information on magnetic film with laser beam heating. Erasable disks require almost same technology and mechanism as those developed for write-once disks, so that know-hows compiled during the development of write-once disk will be fully applied to put erasable disks in practical use.

The rotating speed is normally 300 to 1800 rpm (revolutions-per-minute), which is very slow compared with that of magnetic disks (3600 rpm). Data transmission rate in input and output operations is determined mainly by the rotating speed. Transmission rate should be high enough according to the large capacity of optical disks. At present, transmission rate is 1 to 7 Mbit/s, which is slower than that of magnetic disk (16 to 24 Mbit/s). For example, it will take about 1 hour to read the entire data on 1.1 GB optical disk with 3 Mbit/s transmission rate. Because recording and reading are performed optically without head contact on the disk surface, it is very easy to change disks in spite of high recording density which requires high precision manufacturing.

3.2. Applications of optical disk

In early days of development, it was intended that major application areas for optical disks were commercial audio and video markets. Obviously, these applications employ analog techniques. However, it is the current trend to use digital processing technique in analog instruments supported by analog-to-digital conversion devices, which are the fruits of the progress in microelectronics.

The first optical product for data storage that is produced in volume is the Compact Disk, Read-Only Memory, or CD-ROM for audio application. CD is able to hold up to 550 MB in a disk of about 5 inches in diameter. Next product is an optical disk for video application. It should be noticed that such disks are read-only. Thus read-only applications are restricted for the category of information dissemination, taking advantage of low-cost replication. We can also anticipate that read-only media might be major means for electronic publishing and so on.

Therefore, we should examine write-once or erasable media for storage and retrieval of

information, and only write-once media is commercially available today. Major problems for optical disks are error rate and life time.

The actual error rate, which is a top secret for manufacturers, is not yet disclosed, but bit error rate of optical disk is estimated to be not a low figure because of its high recording density. A large area for error correcting bits is reserved in physical sector format and automatic error correcting technique is employed in reading data from the surface. As a result of such technique, very low error rate (10^{-8} to 10^{-12}) is achieved.

The write-once optical media on the marketplace today have only accelerated life testing to ensure their ten-year life projection. Long history of magnetic media and large amount of experiečne support much longer life projection. This problem rises from the chemical nature of complex multilayer media structure of optical disks. We can anticipate, however, that long life-time be soon achieved with the establishment of technology and advances in manufacturing.

3.3. Document file system

As a typical example of optical disk applications, a document file system (DFS), which now becomes popular in Japan, is outlined in this section. DFS is a stand-alone system to store and retrieve documents mainly for office automation (OA) purpose. Vendors say that DFS is the breakthrough to realize a “paper-less” office.

DFS is composed of 5 components, (a) optical disk unit for document storage, (b) high resolution CRT (Cathode-Ray Tube) display/keyboard unit for interaction, (c) image scanner for document input, (d) image printer for document output and (e) microcomputer system for the control of the whole system. More than 6 Japanese manufacturesrs put DFSs on market at the price of 10 to 15 million yen. In a typical system, optical disk is a DRAW type disk of 20 to 30 cm in diameter. Twenty cm disk is prevailing in space factor, while 30 cm one is good in capacity. The cost per 1 sheet (approximately letter-size paper) is estimated to be 2 or 3 yen. Some DFSs support disk selection mechanism with multiple disk library equipment. Input operation of documents is like that of Xerox photo-copy machines. Five or Ten kinds of keywords can be assigned to each document during input operation. Document retrieval is possible with such keywords. You can browse document image on the CRT display, or get a copy from the printer, which uses thermal printing or laser beam printing technique. The communication facility is available in some DFSs, according to facsimile transmission method. Thus, remote retrieval of documents is realized.

Table 2. Specifications of a Typical Document File System

optical disk media	diameter	30 cm
	substrate	acrylic material
	capacity	3.6 GB
	No. of sheets	60,000 sheets
	price	88,000 yen
input device	input method	CCD sensor
	resolution	8 line/mm or 16 line/mm
	input speed	4.5 sec
	max sheet size	A3(297mm×420mm)
	half-tone	possible
output device	output method	laser beam printing
	resolution	16 line/mm
	output speed	11 sec for 1st sheet. 5 sec for succeeding sheets
	max sheet size	A3
CRT display	size	17 inch
	pixels	1728×2300
	resolution	7 line/mm
controller	disk unit	max 8
	disk changer	25 disks
	max capacity	6 million sheets
	hard disk	30 MB(Max 130 MB)
functions	No. of keywords	5
	length in total	54 characters
	kanji input	possible
	screen editing	possible

As an example, the specification of Toshiba's TOSfile system is outlined in Table 2. Some of the features are:

- (1) Large capacity of 60 thousand sheets.
- (2) Possibility of extension up to 6 million sheets, using disk auto-changer equipment.
- (3) Maximum sheet size of A3 (297mm×420mm, 11.7"×16.5")
- (4) Network connectivity using G III FAX.
- (5) Connectivity to host computer system.

4. System design considerations

4.1. Database technology

In order to handle a large amount of information, *database* technology has been developed in computer science field. Database is defined as a collection of data formed in a

systematic manner. For example, in a bibliographic database, a record consists of such data items as title of paper, names of authors, abstract of paper, title of the journal in which the paper is published, and so on. A collection of such records is called “database.” However, the data items above mentioned are raw information, therefore, we have to cook them to make it easy to retrieve necessary information. This process is called “indexing.” Database producers extract keywords or indices from raw data. Keywords are, for example, technical terms related to the field of the paper or names of authors that can be used as clues to search necessary records. It is the case with most database producers that experts in the field extract appropriate keywords manually. In future, automatic keyword extraction will be possible supported by knowledge engineering or natural language processing technique.

In an information retrieval (IR) system, the cooked database is put into storage devices, and very sophisticated software allows the retrieval of records assisted by keywords. Such software is generally called *database management system* (DBMS). The purpose of DBMS is to realize fast and flexible access capability to each record and to keep integrity in the whole database. Therefore, storage devices for IR system should have random access capability to achieve fast access to any record in a large database. This is the reason that disks are used usually for database storage. It should be also noticed that the data items in the example are all coded data. Coded data, for example, data composed of ASCII character string, is very familiar with computer processing, because information redundancy is reduced. It takes about 4000 bytes to store the characters in this page, while it takes more than 80 KB if you regard this page as an image, which is the technique used in facsimile system. Most IR systems only deal with coded data, because of the limitation of storage space.

Today, there are the emergent necessity for database that handle information sources such as image, voice, and so on. When expressed in digital form, any information could be stored in computer storage, however the problem is the size of each data item. The followings are the approximate sizes:

Monochrome image without half-tone:

Typical case is a document. Data size will be more than 80 KB when high-performance encoding technique, which is used in G III facsimile, is employed.

Monochrome image with half-tone:

Almost same size as the case without half-tone if dither technique is used. Half-tone image looks like a picture on a newspaper with this technique. If higher quality is

required, as is the case with photographs, the amount of data will be more than 1 MB.

Color image:

Almost 3 times as large as the size of monochrome image.

Voice or audio information:

Approximately 1.2 MB or more for 1 minute recording.

Three-dimensional object:

(Future study)

It is because of the amount of data that the number of databases of noncoded information is very small. Thus a storage device of large capacity with random access capability is required. Among the information sources above mentioned, image information is rather compact. So today we can think it feasible to construct a large-scale image database with large capacity storage media. For the other information sources, further studies on storage devices and redundancy reduction technique are necessary to build practical database in a large scale. (Of course, small databases can be constructed with today's storage devices, but any database has the tendency to grow larger and larger.)

Table 3. Areas of Image Database Applications²⁾

Application Areas	Information Sources
Libraries, Archives	Books, Journals, Newspapers, Archives, etc.
Administration	Laws, Ordinances, Registrations etc.
Judicature	Judicial Precedents, Criminal Evidences, etc.
Geography	Survey maps, Ordnance maps, Weather maps, etc.
Industries	Patent Specifications, Desing diagrams. etc.
Natural Sciences	Waveforms, Photographs, Diagrams, Report etc.
Humanities and Social Sciences	Sketches, Photographs, Archives, etc.
Publication	Master pages, Master tapes, Illustrations etc.
Broadcasting and Communications	Diagrams, Photographs, Maps, Letters etc.

4.2. Considerations on storage media²⁾

In this section, we consider on several alternatives to build a large-scale image database with online or real-time retrieval feature. A large variety of application areas of image databases are shown in Table 3. Such database requires the choice of an

Table 4. Comparison of Storage Media²⁾

Storage Media	Magnetic disk	MSS	Optical disk	Microfiche
Storage Capacity	500 MB	50 MB	2000 MB	10 ⁵ pages
No. of Units for 10 ⁶ Pages	200 units	2000 cartridges	50 disks	10 units
Access Time	20 msec	15 sec	0.23 sec	5 sec
Updating	easy	easy	impossible	possible
Computer Interface	simple	simple	possible	possible
Reliability	high	high	unknown	low
Cost	very high	high	unknown	low

appropriate storage medium with high storage density and low cost. Characteristics of some of the prospective media are compared in Table 4. In Table 4, the number of units for storing 10⁶ (i.e., 100 GB) pages is shown assuming digital recording with a redundancy reduction that reduces information in one page of an original document to 100 KB, except for the case of microfiche.

Among them, microfiche is the cheapest medium. Microfilm is another storage medium of same kind, but the difference is that mechanical sorter/viewer is commercially available only for microfiche on the marketplace with modest price. The sorter/viewer stores more than 1000 sheets each containing 63 frames, and sheet picking mechanism and lens optics realize mechanical random access capability in a very primitive level. Another benefit is the fact that scientific journals are increasingly published with microfiche options. However, the experience of the feasibility experiment clearly showed such drawbacks of a microfiche sorter that its access time is quite long, that the medium is basically a read-only memory and requires manual handling for updating, and that its reliability is unsatisfactory. Another characteristic to be noticed is that microfiche sorter is mechanical instrument with lens optics. Therefore, if advanced processing of data is required using computer, high-resolution photoelectric and analog-to-digital conversions are necessary. The outstanding feature of the other media shown in Table 4 is that they deal with digital information, which enables it very easy to process, retrieve, edit and transmit. In other words, microfiche sorter is regarded as a stand-alone equipment and is not appropriate for interfacing to computer system to build online IR system.

Magnetic disks and mass storage systems (MSS) are advantageous in terms of the high reliability as well as the simplicity of interfacing with computer. However, high cost

of magnetic disks and large access time of MSS may rule out the use of these media. A large space is also necessary for the installation of equipments.

Optical disks, therefore, are chosen as the most prospective media because of their high capacity and reasonable access time, although their reliability remains dubious. One drawback to computer system is the write-once feature of DRAW type optical disk. Because of this, it becomes very troublesome to reorganize databases.

From the viewpoint of the whole information system, storage media is just a component of the system and we should also consider other things such as input devices, output devices and communication system. As is described with document file systems, a variety of devices with appropriate characteristics are necessary. To construct high-quality image database, image input device with high resolution and fast reading speed is required along with large capacity storage media. In case of DFSs, they employ the same input technique as used in facsimile. Another possible technique is the approach adopted in high-resolution TV system. In case of online retrieval system, communication system, which is the means for remote access to the database, might become the bottleneck in the system performance, because transmission line currently available is too slow and too expensive for the delivery of a large amount of data. System configuration should be designed according to the considerations on:

- (a) requirements from users,
- (b) available devices (input, output, storage and computer itself)
- (c) service menu and quality
- (d) work load for system maintenance.

The last point indicated in (d) is very important to realize a successful system.

4.3. Optical disk reconsidered

The features of an optical disk are high capacity; ease of removability; and random access capability in the order of 100 milliseconds. High capacity, however, causes another problem. In order to backup the entire data on the surface, it takes 35 to 40 hours to copy data from one disk to another in the present DFSs. Furthermore the initial input of data into disk also takes a lot of hours and human works.

Thanks to the removability, a "jukebox" of optical disks can be implemented for a large database. NASA has developed a jukebox type optical disk system, of which online capacity is 1250 GB. In this system, loaded disk can be accessed in 0.66 sec and any disk can be accessed within 6.8 sec.

The major application of optical disks is storage of images, for which a few bit errors are allowable because human eyes cannot catch such small defects on printed output. Soon, it will be possible to use optical disks for backup media of coded data, for which high reliability is required.

The pace of optical disks entering the marketplace has been rapidly increasing. However, the refinement of the following issues are required urgently:

- (a) enhancement of data transmission rate.
- (b) speed-up of access time.
- (c) reduction of error rate.

And we can expect that the day is not so far away when the erasable optical disk comes in the commercial marketplace.

5. Conclusion

A variety of storage devices are compared in this paper. Especially, the current status and future direction of optical storage devices are surveyed. One thing to remember with database design is that sometimes it becomes very difficult to change the format of database once a large amount of data has been collected. Thus, a carefully work-out design is required before you start to form actual database.

Acknowledgement : I wish to thank Dr. Hashizume of the CBI, University of Tokyo, who contributed to complete this paper. This paper was originally prepared for the Expert Meeting for the United Nations University Archive of Tradional Knowledge, Indonesia, in September, 1985. I would like to thank Dr. Walter Shearer of the United Nations University, the organizer of the Meeting.

References

- [1] A. KAWAKITA: "Introduction for special issue of memories", Vol. 67, No. 11, J. of the IECE of Japan (1984).
- [2] H. INOSE & T. SAITO: "System Configuration of an On-line Library", Vol. E66, No. 10, Trans. of IECE of Japan, (1984).
- [3] "Special issue on mass storage systems", COMPUTER magazine, IEEE, Vol. 18, No. 7, (1985).
- [4] T. YOSHIDA: "Optical disk Memory" , Vol. 67, No. 11, J. of the IECE of Japan (1984).

研修レポート

駒沢大学図書館の現状と将来

Present and Future Condition of Komazawa University Library

駒沢大学図書館 阿部 博則*

要 旨

駒沢大学図書館はこれまで大学の特色である仏教学・禅学関係図書の収集に力を注いできた。現在、本学図書館は機械化を検討中であるが、その際同関係図書の存在が大きな問題となってくる。

本学図書館ではかつて、禅学関係図書を収録した「禅籍目録」を編纂している。同目録は資料のもつ特異な性格により、独自の目録法で記述し、かつ分類作業をしている。しかし将来禅籍目録の継続編纂と、その機械化、および学術情報システムへの参加を考える時、古典籍の目録規則や、分類法の標準化が第一に望まれる。

Abstract

We have been endeavoring to collect the books chiefly concerning with Buddhism and Zen which are the basis of study and education in our university.

At present, we have a program of introducing Library's automation system. In 1964, we published "Zenseki mokuroku", a complete catalogue of Zen books in our Library. But those Zen books are of different character from other books in our Library and require specific cataloging and classification.

Now, we still keep on collecting materials for making the catalogue of "Zenseki" of better quality. In order to adopt the System for Academic Information Center in the near future, it is necessary first of all to make the special cataloging rule and standard classification for Japanese and Chinese classical texts and manuscripts.

はじめに

現代社会は、より高度なコンピュータ化への道を歩き始めている。本学図書館もその例にもれず、機械化を検討中である。

図書館の機械化の手助けとなる文献情報センターシステムは多くの期待にこたえ、実用段階となっ

*Abe, Hironori : Komazawa University Library

た。

しかし、本学図書館にとっての問題は同システムに参加するための条件整備にある。

本学図書館の機械化を考える時、蔵書構成のなかで主要な位置を占める仏教学・禅学関係の図書をいかに取扱うかは、さげられない問題である。かつて手作業で行われた禅籍目録（禅学関係図書）の刊行は本学の特色をよくあらわすものである。

したがって、このレポートでは本学の歴史と現状をふまえて、図書館と共に歩んできた禅籍目録を調査し、同目録と共通点のある国文学研究資料館「和古書目録」との比較検討をも加えて、今後における禅籍目録の取扱い及び他分野でも多くの古典籍を所蔵する本学図書館の機械化という問題にどう対処するかを考えたい。

I 駒沢大学図書館の概要

本学は明治15年10月、曹洞宗大学林専門学校を母体として開校し、百年の歴史を歩んできた。したがって設立当初の蔵書は仏典祖録漢籍（禅学関係図書）が基本となっている。

本学には、まず図書室が設けられ、やがて図書館の設置へと発展していくのであるが、歴史上からみた図書館の推移と特徴を以下に考察してみる。^[1]

明治37年当時の図書系の業務は目録の作成にあわせて、図書をよりよい環境で保存できる図書館の設置に力を注ぐことであった。また、その建設計画の基礎となる図書の収集の充実に務めることも重要であったと考えられる。特に禅学関係図書の収集は機をのがせば入手不可能となり、図書の保存ができなければ破損または散逸してしまう恐れがあったからである。

目録は、本学の所蔵図書の特徴からみて、本館独自の整理法によっている。この目録の種類は分類目録とイロハ目録（書名目録）からなっている。分類は、宗教関係の蔵書が多いという本学の特徴からまず宗乗、余乗の内典（9種類）及び一般図書からなる外典（7種類）に大別している。〔表1参照〕（これが後の駒沢大学十進分類法の基礎となる）

この方法は昭和3年の新図書館建設まで続けられたが、この時を機にこれまでの分類、目録法の変更を企図した。これが「駒沢大学十進分類法」である。

この分類法が旧分類法と大きく異なる点は、本学の蔵書構成の変化によるものと考えられる。新分類法は昭和7年4月の受入図書から開始され、次第に旧分類図書にも及んで全蔵書を新分類法で排架するにいたった。

第二次世界大戦後の昭和24年、新制大学としてスタートしたが、現在の大学は長い単科大学の歴史を経て、文科系の総合大学（大学院5研究科12専攻、5学部15学科、短期大学4科）となったのである。大学のこの拡張は必然的に図書の急激な増加現象をもたらした。

このため書庫を増設して対処したが、長期展望のもとに図書館計画をたてる必要性にせまられた。昭和48年、この計画によって30年間耐用可能な新図書館を建設した。その後もひきつづき図書の収集に力を入れてきたが、その結果、書庫の狭隘化が現在の問題ともなっている。

旧分類法 (設立時)	新分類法(昭和7年制定)	日本十進分類法
	000 総記	000 総記
図書 <ul style="list-style-type: none"> 内典 <ul style="list-style-type: none"> 宗乗部 <ul style="list-style-type: none"> 宗乗部 余乗部 <ul style="list-style-type: none"> 經疏部 華嚴部 天台部 唯識部 因明部 俱舍部 伝燈部 仏典雑部 外典 <ul style="list-style-type: none"> 宗教哲学部 文学語学部 地理歴史部 政法経済部 理化学部 叢書部 外典雑部 	100 禅学	100 哲学
	200 仏教学	200 歴史
	300 仏教各派	300 社会科学
	400 宗教・哲学・教育	400 自然科学
	500 文学・語学	500 技術
	600 歴史・地理	600 産業
	700 社会科学	700 芸術
	800 自然科学	800 言語
	900 軍事・工学・産業・芸術	900 文学

表1 駒沢大学独自分類表と日本十進分類表

戦前・戦後を通して、本学図書館の活動を列挙すれば、①二次資料の出版、②仏教図書館協会の設置、③図書館諸規程の整備、④主に機械化を前提にした業務改善の推進、⑤図書購入の充実等があげられる。

なかでも①二次資料の出版は本学の特徴をよく表わしている。本学は禅学関係図書の収集にあたっては設立当時から全力をあげている。その目録(禅籍目録)の作成(編纂)は、すでに大正2年に計画され、同8年に着手、昭和3年に刊行された。この禅籍目録はその後30余年間、斯学の内外をとわず多くの研究者の手引きとして利用されてきた。しかし、極めて短期間に、しかも小人数で編纂されたため不備な点もあった。これらの充実をはかるために改編を計画し、昭和37年に改訂増補版としての「新纂禅籍目録」、同39年に同補遺版を出版するにいたった。

本学図書館の蔵書構成の基本となる仏教学関係、とりわけ禅学関係図書の収集と禅籍目録刊行に関しては、今後もこの伝統を継承していく必要がある。

このような蔵書の特徴をもち、独自な問題を抱えた本学図書館にとって、今後の図書館業務の改善や業務電算化の検討は重要なポイントになると考える。

II 総合目録としての禅籍目録の編纂

前述の通り本学図書館の活動の重要な位置を占めるものに禅籍目録(二次資料)の編纂がある。禅籍目録は主に禅学という一分野ではあるが、本学及び他機関の書誌と所蔵情報のある程度網羅して構成されていることから、総合目録の機能を持っているといえよう。特に禅籍目録は相互協力の道具として、他に例をみない。国書総目録にも採録されていることから利用度が高いと考える。

これをふまえて、ここでは1. 禅籍目録の編纂方法、2. 国文学研究資料館「和古書目録」と禅籍目録の問題点の比較、3. 禅籍目録の継続刊行計画について考察してみる。

1. 禅籍目録の編纂方法

新纂禅籍目録の凡例及び編纂のために作成された目録（調査）カードから類推してみる。³¹

①書誌収録の対象

昭和3年に発行された禅籍目録（旧版）の収録分野が一般仏教学関係図書にあったのに対し、改訂増補版（新版）は禅学関係図書にしぼって収録している。

②書誌収集の方法

本学図書館が所蔵している禅学関係図書は整理時に作成した目録カードを別置している。

本学以外、他機関の書誌を収集する方法には二つある。一つは大学図書館、公共図書館、文庫などの所蔵目録から同関係図書を収集する。二つは所蔵の確認はとれないが、禅学関係図書の書誌を確認するための情報源（書目、目録及び和刻本、全集などの奥付前後にある出版広告）から収集する。

③書誌の所蔵先

禅籍目録収録書誌の所蔵先は下記の表2のとおりである。

所 蔵 機 関 種 別	機 関 数
大 学 図 書 館（本学図書館を含む）	15館
文 庫（内閣文庫を含む）	13館
公 共 図 書 館（国会図書館上野を含む）	4館
個 人 蔵 書	5名
外 国 所 蔵 図 書（京城大学図書館を含む）	2館
そ の 他	5館
合 計	44館

表 2 新纂禅籍目録書誌所蔵機関数

次に禅籍目録の収録書誌所蔵の調査によれば、同目録に収録されている内訳は下記表3のとおりである。

④書誌記述の方法

このように広範囲から収集された書誌は以下の記述項目により、目録（調査）カードを作成している。記述項目には1) 題号（書名）、2) 巻冊、3) 著編者名、4) 刊写の年次及び執筆者、発行者、5) 所蔵者（所蔵先名）、6) 出拠（所蔵先不明書誌の出典名）、7) 注記、8) 分類（明治以降出版のものは仏教図書館協会制定の仏教図書共通分類表）から構成されている。

書誌所蔵区分	収録書誌区分	収録書誌総数	内異版及び注釈書
第1編 慶応4年以前所出資料	① 駒沢大学所蔵分	7,987	4,126
	② 他機関所蔵分	5,697	2,475
	③ 所蔵先不明	2,743	302
	④ 小計	16,427	6,903
第2編 明治以降所出資料	① 駒沢大学所蔵分	1,567	176
	② 他機関所蔵分	144	25
	③ 所蔵先不明	230	16
	④ 小計	1,941	217
付録 禅籍洋書編	① 駒沢大学所蔵分	168	0
合計		18,536	7,120

表3 新纂禅籍目録収録書誌数

禅籍目録に収録されている書誌の内、慶応4年以前の選述、いわゆる古典籍（漢籍、和刻本、写本など）については以下の特徴があるのでこれを考慮して取扱っている。

これらの書誌は1) 書名は主として巻頭（巻首）からとっているが、略称による場合は、表題と著しく異なる場合があり、これらは別書名として注記している。2) 同一書について多数の異版や注釈書のある場合は本文と注釈書に分け、注釈書にも同一著者によって数種の著作があったり、その各々に異版や他者による副注書（注釈）がある場合は類似のものを一括して扱っている。3) 著者は明らかでも生没年代が不明な場合や4) 書名だけで著筆者も年代も不明な場合も注記している。5) 人名については號（字）と諱とをもつ場合は併記している。

書誌の各項目を調査する際、現物資料をみても上記の特徴から資料の書誌的事項の取扱いについては難しい面がある。まして本学以外の書誌の調査は確認がとりにくく、書誌同定の困難さがある。

2. 国文学研究資料館「和古書目録」と禅籍目録の問題点の比較

国文学研究資料館和古書目録の作成における問題点を「国文学研究資料館報告第10号」でとりあげているので、これを参照しながら禅籍目録との比較を検討したい。⁽⁵⁾

①収録対象

禅籍目録の収録対象は漢籍、和装本、明治以降などの図書も扱うが、両目録で共通する収録資料は慶応4年以前の和装本（版本、写本）である。しかし、禅籍目録は書誌採録の際、本学図書館所蔵以外の資料もとり入れた。この他機関の書誌は現物を確認できない場合もあるので、目録を作成するには困難な面がある。これに対して和古書目録は現物を所蔵しているので、現物そのものから詳細な書

誌事項を確認できるという違いがある。

②目録作成における資料の特徴

両目録で取扱う資料の形態的特徴は、綴り直しや改装など元の本に手を加えやすく改変しやすいことである。また合冊、合綴本も多く、原形がわかりにくい。その他、料紙、表紙の模様、筆跡の特徴の把握など、目録作成において、特殊な知識と注意の必要性が求められることは同じ問題点である。

さらに書誌的事項（書名、著者、出版事項、書写事項等）については、現代の出版物とは異なり、和古書の資料に記載されている情報のみで目録を記述すると不確実、不十分な面が多くみられることになる。

③目録規則について

現在の所、標準的な目録規則が存在していない。したがって禅籍目録は、日本目録規則を使用せず独自の目録法（書名標目）を採用している。国文学研究資料館では日本目録規則1965年版に準拠し、和古書の特徴を考慮して修正補足した「マニュアル」によっている。

④分類について

標準の分類表が存在していない。禅籍目録は明治以降の出版物は分類している。両目録で共通の明治以前所出資料は分類していない。

⑤整理の参考資料について

これらの資料を整理するには特殊な難しさがあるので参考資料に負う所が多く、その果たす役割は大きい。参考資料には整理技術、方法のための参考資料と和古書の内容や書誌的事項についての参考資料の二種類がある。国文学研究資料館での参考資料として、新編和漢古書目録法、国書総目録などがある。

禅籍目録の資料は特殊であるため、参考資料は少なく、仏書解説大辞典、読みの調査には漢和辞典を使用している。（現在は参考資料が整備され、禅学大辞典、曹洞宗全書、国書総目録などがある。）

以上の観点から、古典籍の整理には資料の特徴により標準化された目録規則、古典籍独自の標準分類表の作成、参考資料の充実をはかる必要があると考えられる。

3. 禅籍目録の継続刊行計画について

禅籍目録の増補、改訂版を刊行することは、多くの利用者や研究者にとっても待たれることであると思われる。したがって、これを実現するためには1) 作業の課題、2) 整理技術について、3) 継続計画立案上の問題等を考えてみる必要がある。

1) 作業の課題

禅籍目録のもつ機能的な側面については既に述べたが、本学図書館が今後も目録編纂を続けて行くには人的、経済的な面への配慮が必要である。本学図書館は目録を作成するための下準備と同時に方法を考えなければならない。

作業の課題としては、まず周到な編纂方針の検討とともに①関連書誌の収録、②所蔵先の確認が必

要である。

①関連書誌の収録

前回の収録書誌の再調査、増補の作業を行なわれなければならない。なぜならば、時間的なズレによる所蔵者の移動、また、新たな本学図書館の所蔵追加資料、新出資料があるからである。

②所蔵先の確認

所蔵先の確認を行なわれなければならない。前回までの目録には、所蔵先の無いものもあり、所蔵先が明らかでもその後の移動が予想されるからである。

2) 整理技術について

資料の収集、受入、整理という業務の流れは一定の方針でなされなければならない。

しかし、禅籍目録は2. 国文学研究資料館「和古書目録」と禅籍目録の問題点の比較で述べた問題点を抱えている。この内、標準的な目録規則と標準分類表の問題が重要であると考えられる。

目録規則については日本目録規則1965年版をみると、和漢書と洋書との区別をしていない。さらに古書は適用対象としているが、古典籍の標準的な目録法としては一般化していない。また日本目録規則新版予備版にはその対象から外されている。したがって、標準的な目録規則を必要としている。

一方、標準分類法としては日本十進分類法（NDC）がある。NDCは特定の専門分野、例えば本学の禅学図書をそのまま適用することは難しい。本学図書館が長く独自の分類表を使用しているのも、その理由の一つである。

仏教図書館協会もNDCの一分野のところを展開した分類表を作成しているが、細分化しすぎるきらいがある。これらの理由により、古典籍専用の分類表が必要であるという考え方もある。

このように整理技術には標準化された目録規則、分類表を必要としている。

3) 継続計画立案上の問題

以上に述べたように禅籍目録を編纂するには、多くの問題を含んでいる。ここでは①編纂計画の目標、②機械化について述べてみよう。

①編纂計画の目標

目録作成を実現するためにはまず目標を定め、これに向けてどのようにすれば実現可能なのかを考える必要がある。このような計画は短期間で達成されるものではなく、長期の展望が必要であるから全館員が目標を充分理解することが重要である。

館外においても同様に大学のトップクラスに主旨説明を行い本学の特殊性を生かすためにも、この編纂計画を採択されるよう努力する必要がある。

新纂禅籍目録の完成は個人レベルでの努力によったことは勿論であるが、学外の理解援助があったことも忘れてはならない。また、図書館員は編纂出版の際、補助的な作業に協力を惜しまなかった。このように館内、館外の協力体制により、目標が達成されたのである。

したがって、このような計画は、人的は勿論、経済的な裏付けがあつて初めて実現可能になると考えられる。

②機械化について

はじめにもふれたように現在本学図書館は機械化について検討中である。特に恒常化しつつある整理業務の滞化にまず対処することに焦点をあてている。これを実施するためには書誌データ入力の手軽化を計らなければならない。最近の出版物については各種マークを採用すれば、入力の手力は軽減される。しかし、マークを利用できない古典籍はオリジナル入力を要することに問題がある。この点については、文献情報センター目録サブシステム利用においても参加館のオリジナル入力となることが予想されるため、同様なことがいえる。

ここに、標準化された目録規則、分類表が存在しないという問題があり、このことは古典籍を所蔵する本学図書館が将来、禅籍目録の継続刊行、機械化及び学術情報システムへの参加を考える時の課題でもある。

おわりに

本学の特殊性を考える時、仏教関係がまず念頭におかれるが、なかでも本学図書館の先輩が長い年月を費やして完成させた禅籍目録が本学の特徴を顕著に表わしている。

同目録を将来も刊行するとすれば、古典籍の整理技術に必要な知識は勿論であるが目録規則、分類表の標準化がなされることがもっとも重要なことである。この点については今後の課題としたい。

引用・参考文献

- (1) 駒沢大学百周年史 駒沢大学, 1983.
- (2) 禅籍目録 駒沢大学図書館, 1938.
- (3) 新纂禅籍目録 駒沢大学図書館, 1962.
- (4) 新纂禅籍目録 追補編 駒沢大学図書館, 1964.
- (5) 国文学研究資料館報告第10号 和古書目録データベースの形成と著者名典拠ファイルー「国文学研究資料館蔵和古書目録」作成報告ー国文学研究資料館, p67-70, 1983.
- (6) 国文学研究資料館蔵和古書目録 1972-1981 国文学研究資料館, 1982.
- (7) 長澤規矩也 新編和漢古書目録法 東京, 汲古書院 1979.
- (8) 長澤規矩也 新編和漢古書分類法 東京, 汲古書院 1980.

研修レポート

成蹊大学図書館機械化（漢字処理）へ向けての MARC の実験 — リレーショナル・データベースの応用

MARC Test toward Seikei University Library Computerization (including Kanji)—An Application of Relational Database

成蹊大学図書館 大井 敏暉*

要 旨

本レポートは、PL/I による MARC 処理およびリレーショナル・データベース・マネジメント・システムの応用を述べたものである。はじめに MARC の構造分析を行い、PL/I 処理例として JP MARC レコードのコード変換、記述ブロック・データ抽出および LC MARC レコードのタグ・データ抽出を例示した。次に RDB1 の応用例を示し、最後に MARC、PL/I、RDB1 について検討を加えたものである。

Abstract

This report describes MARCs handling with PL/I programming and an application of a relational database management system RDB1.

Structures of three MARCs are analysed at first. Methods of code conversions, description block data extraction and tag data extraction of MARCs are shown with PL/I programming. Next, a simple model of relational database is constructed and information retrieval is attempted. Comment of MARCs, PL/I, and RDB1 is discussed.

はじめに

当館では、大学情報処理センターのコンピュータを使用して運用業務システム（英数字カナによる固定長データ処理で書誌登録、検索を含む）を稼働させているが、昭和60年秋の機種変更（三菱から富士通へ）に伴ない、漢字処理化および NC 接続を目指して ILIS パッケージ導入を予定している。これに伴ないデータベース・マネジメント・システムは、ネットワーク型（EDMS、処理プログラ

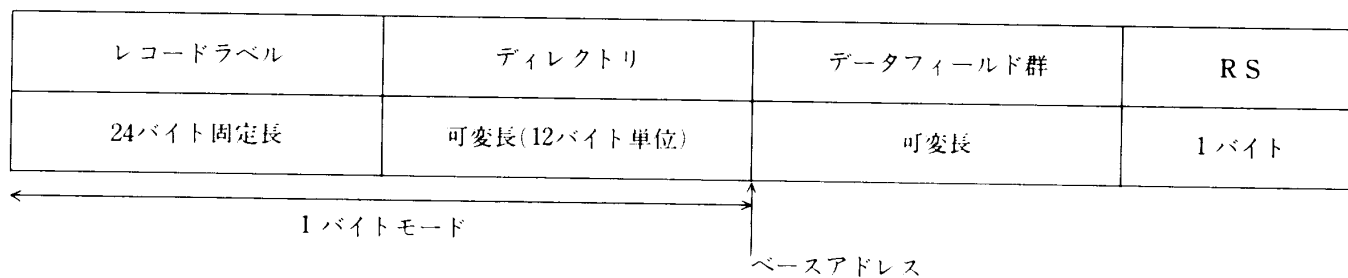
*Ohi, Toshiaki : Seikei University Library

ムは COBOL) よりリレーショナル型に変更されデータも可変長データが追加されることになる。このため研修課題として

1. JP MARC、TRC MARC、LC MARC (BOOKS)の構造理解
 2. PL/I による可変長データ、漢字処理
 3. リレーショナル・データベース・マネジメント・システム RDB1の基礎的な概念把握
- の3点を選びレポートとした。

1. MARC の構造⁽¹⁾

JP TRC LC MARC3種類の計算機処理を主眼とし、その範囲内で構造を調べる。3種類の MARCともレコードの論理構成は、レコードラベル (LC ではリーダー)、ディレクトリ、データフィールド群 (LC ではテキスト) の3部分から構成される (図1)。



RS:レコード区分文字

図1 レコードの構成

1. 1 レコードラベル

JP または TRC

		4	5	6	7	8	10	11	12	17	18	19	20		23
書誌レコード長	レコードドスタータス	書誌的状況コード		未使用	インディケータの長さ	識別子の長さ	データのベース	未使用	記述目録形式	未使用	ディレクトリ・マップ				
		レコードの別	書誌レベル								データフィールドの長さ	領域の長さ	先頭文字位置	領域の長さ	未使用
5バイト	1	1	1	2	1	1	5	1	1	1	1	1	1	2	

図2 JPまたはTRCのレコードラベル

バイト位置	内 容
1~5	レコード長
6	レコード状態区分(n-新規、c-修正、b-削除など)
7	レコード種別(a-印刷物、b-写本、h-マイクロ資料など)
8	書誌レベル(m-単行書、s-逐次刊行物、c-合綴本など)
9~10	不使用(空白)
11	インジケータ長(2)
12	サブフィールド・コード長(2)
13~17	ベース・アドレス
18	入力レベル(空白-完全、1-準完全、8-出版前など)
19~20	不使用(空白)
21	データ長記述データ長(4)
22	データ位置記述データ長(5)
23~24	不使用(00)

図3 LCのリーダー⁽²⁾

1. 2 ディレクトリ

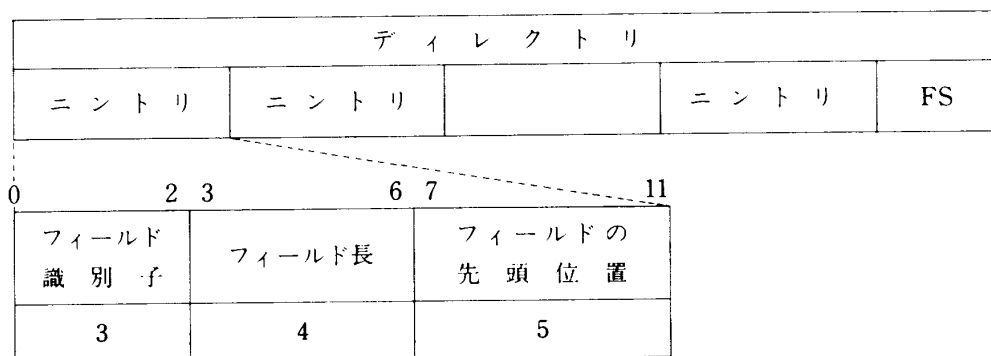


図4 ディレクトリ(JP, TRC, LC共通)

1. 3 データフィールド群

JP または TRC

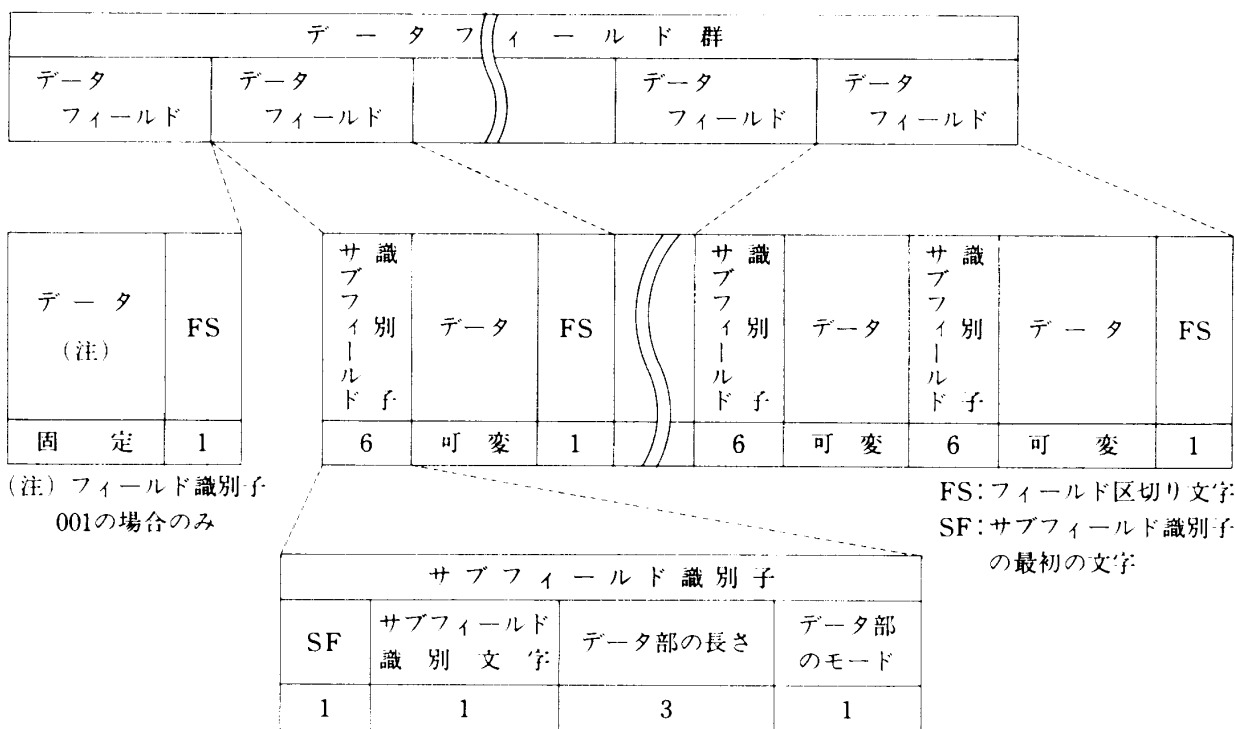


図5 JPまたはTRCのデータフィールド群

LC

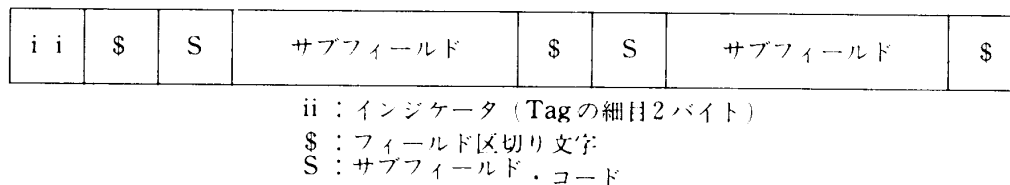


図6 LCのテキスト中のデータ⁽²⁾

以上が、JP TRC LC の MARC 構造である。これら3つの階層からなる MARC レコードから必要なタグデータを取り出すには、まずレコードラベルのレコード長とベースアドレスより12バイト単位で構成されるディレクトリ数を知り、該当タグをもつフィールド識別子をもつディレクトリを取り出す。ベースアドレスからの相対バイト数で示されるフィールドの先頭位置およびベースアドレスからタグデータが格納されているレコード位置が決まるので、その位置からフィールド長だけ抽出すればよい。

2. PL/I による MARC 処理

文献情報など可変長データを扱うには、プログラミング言語として PL/I が優れている。今回 PL/I を学ぶにあたり、根岸正光著「情報システムのための PL/I」近代科学社を用いた。

2. 1 JP と TRC

市販の MT は、両者とも JIS コードあるので、取扱う計算機に合うようにコード変換を行う必要がある。データは EBCDIC と漢字コードが混在しているので、漢字部分のみ変換する。日立の計算機と JIS では、16進で'8080'の違いがあるので、BOOL 代数で処理する。コード変換を行い、記述データを抽出するプログラムとそのブロック（タグ251~377）データ出力例を示す。TRC は、JP 仕様 MT があるので同一プログラムにより処理できる。

```

** FORMATTED PREPROCESSOR INPUT LISTING (BY FORMAT AND MACRO OPTIONS) **
:IC. NO.  -----1-----2-----3-----4-----5-----6-----7
1      JPO2:
2      PROC OPTIONS(MAIN);
3      DCL
4      INFILE FILE RECORD INPUT ENV(VB BLKSIZE(4096) RECSIZE(4092));
5      DCL
6      OUTFL1 FILE RECORD OUTPUT ENV(FB BLKSIZE(5928) RECSIZE(114));
7      DCL
8      JREC CHAR(4088) VAR;
9      DCL
10     01 JPSTRUCT BASED(P),
11         02 RL CHAR(2),
12         02 RECLN PIC'99999',
13         02 LEADERSTR,
14         03 FILLER1 CHAR(7),
15         03 BASEADDR PIC '99999',
16         03 FILLER2 CHAR(7),
17         02 JPCODE,
18         03 JC1 CHAR(3),
19         03 JC2 PIC'9999',
20         03 JC3 PIC'99999',
21         02 DIRECTORY,
22         03 DIR(200),
23         04 TAG CHAR(3),
24         04 LEN PIC '9999',
25         04 DISP PIC '99999';
26     DCL
27     OUTREC1 CHAR(114);
28     DCL
29     W1 CHAR(2000) VAR;
30     DCL
31     W5 CHAR(2000) VAR;
32     DCL
33     (W2,
34     W6) CHAR(1);
35     DCL
36     W3 PIC'9';
37     DCL
38     W4 PIC'999';
39     DCL
40     W4R CHAR(3) DEF W4;
41     DCL
42     W7 CHAR(8);
43     DCL
44     P POINTER;
45     DCL
46     KJSF CHAR(2);
47     UNSPEC(KJSF)='0A42'B4;
48     DCL
49     KNSF CHAR(2);
50     UNSPEC(KNSF)='0A41'B4;
51     DCL
52     K PIC '999';
53     P=ADDR(JREC);
54     ON ENDFILE(INFILE)
55     GOTO OWARI;
56     M=0;
57     CNT:
58     IF M>20 THEN
59     GOTO OWARI;
60     READ FILE(INFILE) INTO(JREC);
61     W4=BASEADDR+1;
62     W7=SUBSTR(JREC.W4.8);
63     PUT EDIT (RECLN.W7) (SKIP(3).X(3).A(5).X(2).A(8));
64     J = (.BASEADDR - 37 ) / 12;

```

```

65      YOMU:
66      DO I = 1 TO J BY 1;
67      W1 = SUBSTR (JREC,BASEADDR + 1 + DISP(I),LEN(I));
68      K = 0;
69      KURIKAESI:
70      IF K = LEN(I) - 1 THEN
71      GOTO KAESI;
72      W2 = SUBSTR(W1,K+2,1);
73      W3 =SUBSTR(W1,K+6,1);
74      W4R = SUBSTR(W1,K+3,3);
75      W5 = SUBSTR(W1,K+7,W4);
76      IF W3 = 2 THEN
77      DO:
78      DO I1=1 TO W4;
79      UNSPEC (W6) = BOOL(UNSPEC(SUBSTR(W5,I1,1)), '80'B4,
80      '0111'B);
81      SUBSTR (W5,I1,1) = W6;
82      END:
83      PUT EDIT (TAG(I),W2,W3,KJSF!!W5!!KNSF)
84      (SKIP,X(3),A(3),X(2),A(3),X(2),A(1),X(2),A(100));
85      IF (TAG(I))>='251' & TAG(I)<='377') THEN
86      DO:
87      OUTREC1=' ';
88      SUBSTR(OUTREC1,1,8)=W7;
89      SUBSTR(OUTREC1,9,3)=TAG(I);
90      SUBSTR(OUTREC1,12,1)=W2;
91      SUBSTR(OUTREC1,13,2)=KJSF;
92      SUBSTR(OUTREC1,15,MIN(W4,LENGTH(OUTREC1)-14))=W5;
93      WRITE FILE(OUTFL1) FROM(OUTREC1);
94      END:
95      ELSE:
96      DO:
97      PUT EDIT (TAG(I),W2,W3,W5)
98      (SKIP,X(3),A(3),X(2),A(3),X(2),A(1),X(2),A(100));
99      END:
100     K = K + W4 + 6;
101     GOTO KURIKAESI;
102
103     KAESI:
104     END:
105     M = M + 1;
106     GOTO CNT;
107
108     OWARI:
109     CLOSE FILE(INFILE),FILE(OUTFL1);
110     END JP02:

```

** NO PREPROCESSOR DIAGNOSTIC MESSAGE (BY FLAG(I) OPTION) **

図7 JP TRC MARC 記述ブロック抽出プログラム例

```

84024487251A国勢調査報告
84024487251D昭和55年版第4巻 その1 第3部
84024487251F総理府統計局 編
84024487270A東京
84024487270B総理府統計局
84024487270D1984.3
84024487275A1冊
84024487275B27cm
84024487291A抽出詳細集計結果(20%抽出集計結果) 全国編 住民の状況, 夫婦の労働力
84024488251A事業所統計調査報告
84024488251D昭和56年版第1巻 追補版
84024488251F総理府統計局 編
84024488270A東京
84024488270B総理府統計局
84024488270D1984.3
84024488275A5, 554p
84024488275B26cm
84024488291A全国編 常雇規模別集計
84024489251A全国世論調査の現況
84024489251D昭和58年版
84024489251F内閣総理大臣官房広報室 編
84024489270A(東京)
84024489270B内閣総理大臣官房広報室
84024489270D(1984)
84024489275A646p
84024489275B26cm
84024489350A昭和57年4月~昭和58年3月調査分 『世論調査年鑑』(大蔵省印刷局)
84024490251A増加図書年報
84024490251D昭和58年
84024490251F総理府統計局図書館 編
84024490270A東京
84024490270B総理府統計局図書館
84024490270D1984.3
84024490275A284p
84024490275B26cm

```

図8 JP MARC 記述ブロック抽出出力例

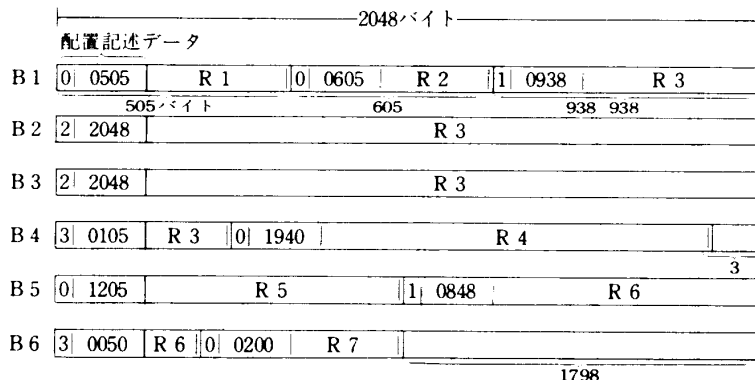
```

T2801867251AB A S I C プログラミング
T2801867251B 入門から応用まで
T2801867251FM i c h a e l   T r o m b e t t a   著
T2801867251F 石井 稔 訳
T2801867270A 東京 出版
T2801867270B 啓学 出版
T2801867270D 1 9 8 4 . 8
T2801867275A 2 9 8 P
T2801867275B 2 3 c m
T2801867360B 3 2 0 0
T2801867365AH
T2801876251AF M - 7   N e w 7   7 7 マシン語入門マニュアル
T2801876251F 中村英都 著
T2801876270A 東京
T2801876270B 秀和システムトレーディング
T2801876270D 1 9 8 4 . 7
T2801876275A 2 8 1 P
T2801876275B 2 6 c m
T2801876360B 2 2 0 0
T2801876365AJ
T2801885251A 実用 F O R T H テクニック入門
T2801885251F 西川利男 著
T2801885270A 東京
T2801885270B 誠文堂新光社
T2801885270D 1 9 8 4 . 8
T2801885275A 2 0 0 P
T2801885275B 2 1 c m
T2801885360B 2 0 0 0
T2801885365AL
    
```

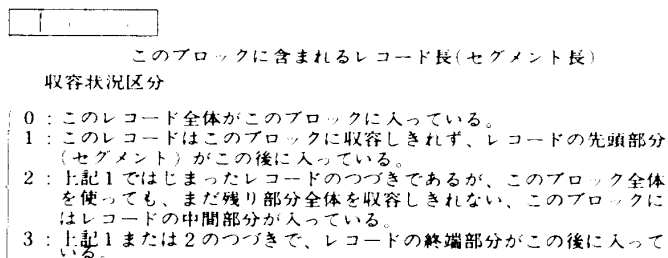
図9 TRC MARC 記述ブロック抽出出力例

2. 2 LC

LC MARC は、ASCII コードで作成されているので EBCDIC を使用している計算機では、コード変換を行う。⁽³⁾ 入力ファイルとして下記の構成をもつ LC ファイルから、タグ001と245をもつデータ抽出例を示す。このプログラムは、根岸先生の例題プログラム⁽⁴⁾を参考にして作成を行った。



「論理」レコードの配置



配置記述データ

図10 入力ファイルのレコード配置

図11 LC MARC によるタグ001, 245字一々抽出出力例

M A R C B O O K S

001: 00000294 //r31
 245:00 aGeneral digest of the decisions of the principal courts in the United States, England and Canada. bIncludes all officially reported cases and all cases not to be officially reported which were first published between Sept. 1, 1895, and <Oct. 1, 1906> ... Refers to all reports, official and unofficial. v. 1-22.

001: 00000547 //r28
 245:10 aPhysics and politics; bOr. Thoughts on the application of the principles of "natural selection" and "inheritance" to political society. cBy Walter Bagehot; with a special introduction by J. Laurence Laughlin.

001: 00001080
 245:00 aThe great operas; cThe romantic legends upon which the masters of song have founded their famous lyrical compositions. introduced by Giuseppe Verdi ... edited by James W. Buel, Ph.D.

001: 00001453 //r19
 245:10 aSal am an and Abs al; ban allegory ctr. from the Persian of J am i; together with a bird's-eye view of Far id-Udd in Attar's Bird-parliament. By Edward Fitzgerald; ed. by Nathan Haskell Dole.

001: 00001596 //r37
 245:00 aLife of the Rev. Henry Harbaugh, D.D. cBy Linn Harbaugh. With an introduction and eulogy ...

001: 00002117 //r782
 245:00 aTraitemnt rationnel des maladies caus es par les germes. bact eries, microbes. bMode d'emploi du glycozone et de l'hydrozone, cpar Charles Marchand ...

001: 00002163 //r832
 245:14 aThe story of the Boers, bnarrated by their own leaders, prepared under the authority of the South African republics. ch C. W. Van der Hoogt. Preceded by The policy of mediation, by Montagu White.

001: 00002566 //r80
 245:14 aThe conception of immortality, cby Josiah Royce ...

001: 00004337
 245:14 aThe Jeffersonian cyclopedia; ba comprehensive collection of the views of Thomas Jefferson classified and arranged in alphabetical order under nine thousand titles relating to government, politics, law, education, political economy, finance science, art, literature, religious freedom, morals, etc.; ced. by John P. Foley ...

001: 00005242 //r802
 245:04 aThe therapeutical applications of hydrozone and glycozone, cby Charles Marchand. Rational treatment of diseases characterized by the presence of pathogenic germs.

001: 00005742
 245:10 aPlay guide. by. I. Jan., 1900.

001: 01000833
 245:00 aMessages from the governors of Maryland and Pennsylvania, btransmitting the reports of the joint commissioners, and of Lieut. Col. Graham, U.S. topographical engineers, in relation to the intersection of the boundary lines of the states of Maryland, Pennsylvania and Delaware, being a portion of Mason and Dixon's line.

001: 01001576
 245:12 aA new geographical, historical, and commercial grammar; band present state of the several kingdoms of the world ...

001: 01002581 //r832

3. RDB1^{(5), (6), (7)}

RDB1のデータベースは、データモデルとしてリレーショナルモデルを採用した日立のリレーショナル・データベースである。

3. 1 特長

- (1)データベース内に納められるデータは、すべてテーブル（表形式）で表現する。
- (2)アクセス言語は、非手続き型であり、これを EQL (Extended Query Language) と呼び、データの定義から検索、更新および制御までデータベースのすべての操作を体系化している。
- (3)データへのアクセスパスは、データベースが最適なアクセスパスを選択する。
- (4)テーブルどうしのデータの関係付けは、実行時に行うのでユーザが事前に定義する必要がない。
- (5)データベースの定義は、実行時にできる。

3. 2 テーブル

EMPNO	NAME	DEPT	SAL

欄

行

図12 テーブルの構成

テーブルは、2次元の表で表現し、複数の欄と行で構成される（図12）。

テーブルは、集合論の関係（リレーション）を表現したものであり、下記の条件に従う。

- (1)テーブルおよび欄には、名称が付けられている。
- (2)一つのテーブルの中には、同じ行、つまりすべての欄が同じ値を持つレコードは存在しない。
- (3)一つのテーブルの中に現れる行は、どんな順序にも取り出すことができる。
- (4)一つのテーブルの中に現れる欄は、どんな順序にも取り出すことができる。

従来のファイルの概念で考えれば、テーブルはファイルに、行はレコードに、欄はフィールドにほぼ対応するが、RDB1では、物理的イメージとは切り離し論理的な見方だけで操作が行えるようになっている。

(1)-(4)の条件は、データを物理的ではなく論理的に扱うために必要な条件である。(1)によって、データベース内の任意のテーブルや欄を指定でき、(2)によって、データベース中の一つ一つの行が識別できる。また(3)および(4)によって、テーブルは単なる集合としての意味しか持たないことを示している。

この様なテーブルの集合をデータベースと呼ぶ。

3. 3 書誌データベースの作成

3. 3. 1 データベースの内容

実験用データベースとして、JP MARC の記述ブロック、書名 (仮名キーワード)、著者 (仮名) から構成される3つのテーブル集合を選んだ。これだけあれば RDB1の基本的な概念把握ができると判断したからである。なおインデックスは省略した。これはデータ量が少ないことと、インデックスは検索の応答時間を早めるだけで RDB1の本質をかえるものではないことによる。

データベースの略図を下記に示す。

書誌記述 SHOSHI 1

書誌番号		タグ	サブフィールド	記述ブロック
名称	BIBNO	TAG	SF	KIJUTU
タイプ	CHAR(8)	CHAR(3)	CHAR(1)	CHAR(102)
	83000005	251	A	日本統計年鑑
	83000005	251	D	第30回
	83000005	251	F	総理府統計局 1 編
	83000005	2	A	東京
	83000005	270	D	1980.6
	83000005	275	A	732P
	83000005	275	B	27cm
	⋮	⋮	⋮	⋮

書名 (キーワード) SHOKEY

書名キーワード		書誌番号
名称	TKEY	SNO
タイプ	CHAR(8)	CHAR(8)
	ヨウラン	84202449
	⋮	⋮
	⋮	⋮

著者名 (カナ) CHOKEY

著者キーワード		書誌番号
名称	AKEY	ANO
タイプ	CHAR(16)	CHAR(8)
	オオクラショウ	84024492
	⋮	⋮
	⋮	⋮

図13 実験用データベースの略図

3. 3. 2 テーブルの登録

テーブルを RDB1 上に定義するには、EQL コマンドで下記のように入力する。

```
CREATE TABLE [所有者名.] テーブル名
          (欄名 データタイプ [, 欄名 データタイプ] ...)
          [IN [所有者名.] データベース領域名]
```

```
EQL
SELECT CNAME, TNAME, CREATOR, COLNO, COLTYPE, LENGTH FROM SYSTEM.SYSCOLUMNS WHERE CREATOR='Z00508'
85-02-28 *** SELECT CNAME, TNAME, CREATOR, COLNO, COLTYPE, LENGTH FROM SYSTEM. ***

CNAME          TNAME          CREATOR    COLNO COLTYPE    LENGTH
-----
AKEY           CHOKEY         Z00508      1 CHAR      16
ANO            CHOKEY         Z00508      2 CHAR      8
TKEY           SHOKEY         Z00508      1 CHAR      8
SNO            SHOKEY         Z00508      2 CHAR      8
BIBNO          SHOSHI1        Z00508      1 CHAR      8
TAG            SHOSHI1        Z00508      2 CHAR      3
SF             SHOSHI1        Z00508      3 CHAR      1
KIJUTU         SHOSHI1        Z00508      4 CHAR     102
DKA7570I  8 ROW(S) SELECTED.
          SELECT COMMAND COMPLETED.
```

図14 テーブルの登録確認

3つのテーブルを上記に従って入力する。入力後、ディクショナリ・テーブルで確認したものが下図である。

3. 3. 3 データ入力

A. バッチ入力

DATALOAD コマンドを使い、書誌記述テーブルにマルチユーザモードの下に実行する。入力データは、先に PL/I で作成したものを使用。図15に DATALOAD コマンド使用例を示す。

```
/SAVE HT63B
JCT54032I /SAVE CMD ACCEPTED
L
10000 //Z00508DL JOB *****,MSGCLASS=A,TIME=(60),NOTIFY=Z00508,
20000 // REGION=3000K
30000 //DATALOAD EXEC PGM=DKADBUTL,TIME=5
40000 //STEPLIB DD DSN='SYS1.RDBPGM',DISP=SHR
50000 //RDBIPT DD *
60000 RDB SYSMODE=S,STARTUP=W,PROGNAME=DKADBUTL
70000 /*
80000 //DLDIR DD DSN=DATA,DISP=SHR
90000 //SYSPRINT DD SYSOUT=A
100000 //SYSIN DD *
110010 CONNECT Z00508 IDENTIFIED BY Z00508;
120000 DATALOAD TABLE(SHOSHI1)
130000 COLUMN(BIBNO CHAR(8),TAG CHAR(3),SF CHAR(1),KIJUTU CHAR(102))
140009 INFILE(DLOIN);
150000 /*
160000 //RDBLST DD SYSOUT=A
170000 //RDBDMP DD SYSOUT=A
180008 /*DIR DD DSN='Z09995.RDB1.DIR1',DISP=OLD
190008 /*LOG1 DD DSN='Z09995.RDB1.LOG1',DISP=SHR
200008 /*LOG2 DD DSN='Z09995.RDB1.LOG2',DISP=SHR
210008 /*DBXT1 DD DSN='Z09995.RDB1.EXT1',DISP=OLD
220008 /*DBXT2 DD DSN='Z09995.RDB1.EXT2',DISP=OLD
230008 /*DBXT3 DD DSN='Z09995.RDB1.EXT3',DISP=OLD
240008 /*DBXT4 DD DSN='Z09995.RDB1.EXT4',DISP=OLD
JTD5200I END OF DATA
E>
```

図15 DATALOAD コマンド使用例

B. 端末入力

書名および著者名テーブルのデータ入力、端末を使用した。端末入力の EQL コマンドは、行毎に入力する INSERT とデータを連続して入力する INPUT の両コマンドがある。

```
INSERT INTO テーブル名 VALUES
(欄値 [, 欄値……])
```

```
INPUT テーブル名
欄値 [, 欄値……]
END
```

両方試みたが、IPUT コマンドの方が簡単で早い。

3. 3. 4 検索例

コマンド形式の一般形を下記に示す。

```
SELECT [ ALL
DISTINCT ] { セレクトの並び
* }
FROM [所有者名.] テーブル名 [テーブル標識]
[, [所有者名.] テーブル名 [テーブル標識]] .....
[WHERE 検索条件]
[GROUP BY 欄指定 [, 欄指定] ...]
[HAVING 検索条件]
[ORDER BY { 欄指定
ソート項目指定番号 } [ASC] [, { 欄指定
ソート項目指定番号 } [ASC] ] ...]
```

検索例 1

JP 番号84024484の書誌記述の検索

```
EQL
SELECT * FROM SHOSHII
WHERE BIBNO='84024484'
```

BIBNO	TAG	SP
84024484	251	A
科学技術研究調査報告		
84024484	251	D
昭和58年		
84024484	251	F
総理府統計局 編		
84024484	270	A
東京		
84024484	270	B
総理府統計局		

検索例 2

書名の読みに「ホウコク」をもつ書誌記述（サブフィールドを除く）の検索

```
EQL
SELECT DISTINCT BIBNO, TAG, KIJUTU FROM SHOSHII, SHOKEY
WHERE TKEY='10022' AND BIBNO=SNO
```

BIBNO	TAG
84024484	251
科学技術研究調査報告	
84024484	251
昭和58年	
84024484	251
総理府統計局 編	
84024484	270
1984.3	
84024484	270
総理府統計局	

1 件分

(次頁につづく)

84024484 270 D 1 9 8 4 . 3	84024484 270 東京
84024484 275 A 2 0 7 P	84024484 275 2 0 7 P
84024484 275 B 2 6 c m	84024484 275 2 6 c m
	84024485 251 国勢調査報告
	84024485 251 昭和 5 5 年 第 4 巻 その 1 第 1 部
	84024485 251 総理府統計局 編
	84024485 270 1 9 8 4 . 3
	84024485 270 総理府統計局
	84024485 270 東京

検索例 3

今までの検索例では、毎回書誌番号が出力されて見づらいので編集を行ない、書誌記述テーブルから、書誌番号、書名、著者を出力する。

```

EQL
CREATE TABLE KAN1 (KNO1 CHAR(8), KSHO1 CHAR(102)) IN "PUBLIC".USER1
DKA7607D 'COMMIT WORK' IS ISSUED AUTOMATICALLY. IF OK, ENTER 'GO'.
        IF NOT 'CANCEL'.
GO
DKA7578I CREATE COMMAND COMPLETED.
EQL
INSERT INTO KAN1 (KNO1,KSHO1) SELECT BIBNO,KIJUTU FROM SHOSHI1 WHERE TAG='251' AND SF='A'
DKA7570I 21 ROW(S) INSERTED.
        INSERT COMMAND COMPLETED.
EQL
CREATE TABLE KAN2 (KNO2 CHAR(8), KCHO1 CHAR(102)) IN "PUBLIC".USER1
DKA7607D 'COMMIT WORK' IS ISSUED AUTOMATICALLY. IF OK, ENTER 'GO'.
        IF NOT 'CANCEL'.
GO
DKA7578I CREATE COMMAND COMPLETED.
EQL
INSERT INTO KAN2 (KNO2,KCHO1) SELECT BIBNO,KIJUTU FROM SHOSHI1 WHERE TAG='251' AND SF='F'
DKA7570I 13 ROW(S) INSERTED.
        INSERT COMMAND COMPLETED.
    
```

(次頁に続く)

```
SELECT KNO1,KSHO1,KCHO1 FROM KAN1,KAN2 WHERE KNO1=KNO2
85-03-04 *** SELECT KNO1,KSHO1,KCHO1 FROM KAN1,KAN2 WHERE KNO1=KNO2
```

KNO1

KSHO1

KCHO1

84024484

科学技術研究調査報告

総理府統計局 編

84024485

国勢調査報告

総理府統計局 編

84024486

国勢調査報告

総理府統計局 編

84024487

国勢調査報告

総理府統計局 編

84024488

事業所統計調査報告

総理府統計局 編

4. 考察

4. 1 MARC

漢字処理において JIS 第2水準外の外字が発生する。NC とメーカーが異なる場合、どこまでケアするか今後の課題となろう。JP と LC MARC とでは、JP が完全入力なのに対し LC は不完全入力（例えば書名なし）のものが存在する。

4. 2 PL/I

文字例の可変長データ処理は、COBOL でも可能であるが繁雑となる。これに対し今回 MARC 処理を通じて学んだ PL/I は、文字列処理に強力な力を発揮した。特に組込関数、擬変数、連結演算、BASED 変数等が役立つように思われる。

4. 3 RDB1

ネットワーク型では、レコードの格納属性が多様であり、かつグループ間を結ぶリンケージパス（親子関係の指定、継ぎ方の指定）がある。これらのことを知らないと扱えない。またリンクがない場合は互いに独立しており、関係づけるにはプログラムの介在が必要となる。これに対し RDB1では、格納属性は限られており、またデータ操作に関して、テーブルを集合とみなし集合演算処理を行う。このことから RDB1は、テーブルの追加等の変更が容易にできる特徴をもつ。また RDB1は、リンケージパスを持たないため、非定型業務に適しているが、定型業務にはそぐわない面がある。これは結合に関し内部を総ナメにするため時間がかかる点にある。

テーブルを作成する際の問題点としては、冗長性を除くための正規化がある。このことを考慮してテーブルを作成する必要がある。

リレーショナル・モデルは、エンドユーザに対しては、親言語を知らなくとも、非手続き言語（例 EQL）を使用してテーブル作成、データ登録、削除、検索が簡単に行えるので、利用の道を開いていくと思われる。

おわりに

PL/I、RDB1の概念把握にあたり、深夜に及ぶご指導を下さいました東京大学文献情報センターの根岸正光先生に深く感謝いたします。

参考文献

- (1) 国立国会図書館。JAPAN/MARC マニュアル。第1版, p. 9-16.
- (2) 根岸正光. 情報システムのための PL/I. 近代科学社, p. 194-196.
- (3) *ibid*(2), p. 143-145.
- (4) *ibid*(2), p. 203-207.
- (5) HITAC プログラムプロダクト VOS3 データマネジメントシステム RDB1解説. 第1版, 1984.
- (6) HITAC プログラムプロダクト VOS1/ES, VOS3データマネジメントシステム RDB1EQL 文法. 第1版, 1984.
- (7) HITAC プログラムプロダクト VOS3 データマネジメントシステム RDB1使用の手引. 第1版, 1984.

研修レポート

農学分野における情報検索の現状と問題点

Some Recent Problems of Online Retrieval in Agriculture

京都大学農学部図書室 為石 理恵子*

要 旨

京都大農学部図書室では、1982年8月からオンライン情報検索サービスを実施している。利用者がオンライン検索サービスをどのように捉えているのか、その問題点を探り、より望ましい情報検索サービスを考えるため、利用者の実態調査を行った。アンケート結果の分析から、検索効果あるいは結果に対する不満は、検索語の適否の点にあると考えられ、BIOSIS Previews をとりあげ、自然語とコードでの検索比較を試みた。さらに、農学分野におけるデータベースの問題点と検索仲介者の役割を考察した。

Abstract

Online information retrieval services are carried out from August 1982 in Library of Agriculture, Kyoto University. In order to find the needs of researchers and provide efficient searching, a survey by the questionnaire was made to the users. The analysis of investigation clarified that one of the problems for searching was choice of search terms. Searchings with free terms and codes were examined using BIOSIS Previews. Furthermore, evaluation of databases in agricultural science and a role of librarians as intermediary were discussed in this report.

1. 京都大学農学部図書室におけるオンライン情報検索サービス導入の背景と経緯

1. 1 オンライン情報検索サービス導入の背景

今日のような学術情報の氾濫の中では、研究者は自分の欲するすべての情報を入手することが困難な状況となり、そのための時間、労力が多大なものとなっている。このことは、研究者自身にとって問題となるのみならず、情報提供機関である大学図書館にも、情報をいかに迅速かつ正確に提供できるかを要求する。さらに、農学分野では、その研究領域は非常に多岐にわたっており、自然科学、社会科学、人文科学に関する多様な情報を提供しなければならない。このような状況下で、当図書室が

*Tameishi, Rieko : Library, Faculty of Agriculture, Kyoto University

研究者に十分な情報を提供するには、オンライン情報検索システムの導入が不可欠と考えられた。

1. 2 利用システム、データベース、端末機

オンライン検索サービスで利用しているシステムは、1984年12月現在、DIALOG 情報検索システム (1982年 8 月開始)、TOOL-IR/ORION (1983年 6 月開始)、UTOPIA (1983年 8 月開始) である。それぞれの利用状況は表 1 に示した通りである。

DIALOG の利用度は、研究者が35%、図書室が65%となり、1ヶ月の平均利用時間は、研究者0.536時間、図書室では0.994時間となっている。

これまでに利用したデータベースは表 2 の通りである。研究者も、図書室のレファレンス業務においても、BIOSIS Previews と CA Search をよく利用している現状を示している。

端末機としては IBM ディスプレイライターシステムを使用している。

1. 3 利用者との対応

情報検索の依頼を受けたスタッフが事前の打合わせから検索までを一貫して行う。

- (1) 検索主題を理解する。
- (2) 利用システム、データベースを選択する。
- (3) 検索式を作成する。
- (4) 端末で検索を行う。
- (5) 結果をチェックし、プリントアウトして依頼者に渡す。
- (6) 検索結果の中から依頼者が必要とする一次文献を入手する手続を行う。

上述の過程は、時間の許す限り利用者と協同で行う。殊に、端末で実際に検索を行う段階に立ち合ってもらふことは、予測しない結果が出た時の対応がスムーズにいくので効果的であった。

2. アンケート調査の分析

情報検索サービスを利用した研究者の検索結果に対する意見はその都度聞くようにしているが、検索結果に対する判断には時間がかかる場合もあり、十分な検討は行っていなかった。そこで今回、利用者がオンライン検索サービスをどのように捉えているのか、その問題点を探り、より望ましい情報検索サービスを考えるため利用者の実態調査を行った。表 3 のアンケート用紙を今までに情報検索を依頼した利用者 (22名) に配付し⁽¹⁾⁽²⁾、全員の回答を得た。

2. 1 利用の動機について

オンライン情報検索利用の動機は、「進行中の研究を進める上での参考にするため」を最優先項目としたものが45.5%で、これに副次的動機を合わせると95.5%を占め、研究遂行のために利用されていることが示された。「自分の文献が抄録されているかを知るため」、また、「論文のレフェリーとしての仕事の参考にするため」の利用は皆無であった。その他として、「ある特定の分野に関係している文献についての自分の知識水準をチェックするため」に利用したケースがあった。

2. 2 検索後の効果について

検索後、利用者の研究への影響の内容を尋ねた。最も大きな効果は、「自分の研究分野に関し、広い知識が得られた」ことであると回答した人が22.7%、次いで「二次情報を探す時間が節約できた」人が9.1%であった。副次的効果として、54.5%の利用者が「現在進行中の研究の進め方にヒントが得られた」と回答し、18.2%の人が「不必要な文献（ノイズ）が多く、時間の損であった」と悪い効果を指摘した。しかし、「研究を中止または、方向転換するきっかけとなった」人や「論文を読むのに忙しくなって研究が進まなくなった」人や「他人の論文に気をとられ、目うつりがして研究の邪魔になった」効果を挙げた人はなかった。また、最重要効果として「自分の研究分野に関しては独自のデータベースを開発する必要性を再確認することができた」点を指摘し、既製のデータベースではカバーすることのできない分野の研究者の不満が見られた。

2. 3 検索結果について

検索結果に対する満足度としては、「料金が高かった」と回答した人が54.5%で、半分以上の利用者が使用料金が高いと感じたようである。「求める文献が適切に検索できた」人は40.9%「不要な文献（ノイズ）が多かった」と回答した人は45.5%を占め、検索結果に満足した利用者とは不満をもった利用者が相半ばする状態ではあるが、満足している利用者（「求める文献が適切に検索できた」、「期待以上の文献が発見できた」、「不要な文献（ノイズ）が少なかった」、「データの質がよかった」）よりも、不満な利用者（「必要な文献が検索できなかった」、「不要な文献（ノイズ）が多かった」）の方が多少多かった。また、「その他」に、「検索された文献が本当に関連分野の文献を総て網羅しているかどうか不安を覚える」と回答したケースが見られた。

2. 4 検索語・検索式の適否

72.7%の利用者が満足し、無回答が一人であった。不満と回答した人が22.7%あり、そのいずれもが検索語に対して不満を述べている。不満の内容は

- (1) インターディシプリナリーな研究や、新しい分野での研究では適切な検索語をみつけるのが困難である。
- (2) コードによる検索はカテゴリーが適切でないように思われる。実際の研究分野の広がりや正確に反映したカテゴリーの設定が行われていると思われたい。 (BIOSIS Previews を使用した利用者) という二つの点に要約される。(1)の不満を述べた一人は、「自分の研究に関して、最適の検索語を得るためには、実は相当な試行錯誤が必要であると思われる」と答えている。

2. 5 アンケート結果から

利用者の実態は、回答を統計的手法により分析しなければ完全な把握には至らない面があるが⁹⁾、ここでは、単純集計により当図書室のオンライン情報検索サービスの改善に必要な幾つかの点を探った。

検索効果あるいは結果に概ね満足している利用者とは不満を表明している利用者がある。検索後の研究への波及効果の不満は、いずれも検索効率に対する不満である。検索効率については、22.7%の利

用者が検索語が不相当と答えていることから、検索語の選択に問題があると思われる。しかし、当図書室では、コードあるいは検索語の選択を専門家である利用者と代行者とで事前に協議して決定しているため、図書室の責任とばかりはいえないが、今後のサービスにおいては、十分に考慮しなければならない。

次に、使用したデータベース別に満足・不満の回答率をみると、CA Search(DIALOG)およびCAS (TOOL-IR/ORION) を利用した研究者は54.5%で、満足の回答率が66.7%、不満と回答した人は33.3%、BIOSIS Previews(DIALOG)およびBA・BI (UTOPIA) の利用者は36.4%で、満足の回答は25%、不満の回答は62.5%、無回答が12.5%であった。その他 (AGRICOLA、CAB、SCI) は9.1%で満足・不満ともに50%である。CA Search および CAS を利用すると満足する人が62.5%を占めるが、逆に BIOSIS Previews および BA・BI を利用すると不満が多くなる傾向を示す。しかも、BIOSIS Previews および BA・BI を使い、不満を表明した利用者は検索語が不相当であることを回答している。そこで BIOSIS Previews の検索方法の検討を行ってみることにした。

3 BIOSIS Previews を使った特定主題での検索

狭い主題と広いと思われる主題で BIOSIS Previews の検索実験をした。

3.1 検索実例その1 細菌分類学におけるピンク色をした *Methylobacterium* の検索

これは、二重結合のない C1化合物を単一の炭素源あるいはエネルギー源として資化するピンク色をした *Methylobacterium* の生理的能力について記載した文献を検索するのが目的であった。データベースとしては BIOSIS Previews が最適であり、検索にはコードも合わせて使うことにした。

検索された文献数は、表4に示されている。利用者が必要とした文献が最も多く出現したのは、PINK(W)PIGMENTED と BACTERIA を掛けた(2)と CC=30000と BC=04714を掛けた(6)であるが、前者の精度⁴⁾ (必要とした文献÷その集合に現われた文献数×100) は43%、後者のそれは8.8%で、精度の上から自然語を用いた前者の方が優位であった。また、自然語を用いた集合(2)には、他の集合に出現しなかったけれども、利用者にとって必要な文献が見られ、依頼者の要求を充足する効果があったと考えられる。

3.2 検索実例その2 遺伝子組み換え、細胞融合、組織培養技術の食糧生産への応用

バイオテクノロジーの諸領域が食品加工、品種改良といった食糧生産に直接・間接に利用された情報の有無を検索した例である。

検索語の組合わせと検索された文献数は表5に示されているが、合計32件の文献が得られ、利用者が必要としたのはその内2件であった。(1)1件、(2)1件、(3)2件、(4)では適合文献が検索されなかった。利用者が必要とした文献が(3)に出揃い、検索精度も最も高かった。この検索例で明らかなように、コードを用いた検索式では、コードの概念のカテゴリーが広いだけに精度が低くなるのは仕方がないとしても⁶⁾、自然語を用いた検索式に現れる文献しか拾えない。一方、自然語の検索式の場合にのみ出現する適合文献があり、コードを用いた検索法には一定の限界があるように思われた。

4 考察

4.1 フリータームでの検索と BC コード、CC コードでの検索比較

アンケート調査の結果に明らかなように、検索効率に対して不満をもつ利用者には、検索語の不適を指摘しているものが多い。利用者と仲介者の両者で選択したタームを使用しているにもかかわらず検索効率がよくないのは問題である。上述の検索実例が示すように、コンセプトコードに頼った検索は、コンセプトコードの概念とそのカテゴリーが、それが改訂されるまでの間に、日々発表される論文によって変化していくという欠陥を不可避免的に持つと考えられる。一方、バイオシステムティックコードは、今迄の分類体系が再検討されている分野では（例えば細菌分類学等）その概念が研究者全てに共通していない。従って、BC コードを用いて検索した場合、必要な細菌名に的中しない可能性を秘めている。

学際的研究および生物名の登場する検索では、コードを用いた検索は精度を落とし、利用者の不満を生む結果となりうる。このような主題の場合には、研究者＝利用者と検索仲介者が十分に検討した自然語を用いて検索を行った方が満足度を高めるものと考えられる。しかし、包括的な検索を行う場合には、その概念を自然語で表現するためには非常に多くの言葉を検索語として使用しなければならないので、そこには、シソーラスや統制語、あるいはコードの必要性も出てくるであろう⁹⁾。従って、検索する側は、両者の限界を見極めて検索を行わなければならないし、また、アンケート回答者の一人が指摘したように、かなりの試行錯誤が必要である。

4.2 農学分野でのデータベースの問題点

当図書室の利用者が使用するデータベースは BIOSIS Previews、CA Search、CAB、AGRICOLA に集中している。これは、Johnson⁷⁾ や Meyer⁸⁾ らの、データベースに収録される独自文献の重複率の調査で明らかにされているように、2、3種類の適切なデータベースを選択することでかなりの文献が得られることによると思われる。しかし、アンケート調査の回答に見られるように、独自のデータベースを作らねばならないことを再確認した利用者がある。これは既製のデータベースではカバーできない分野の存在を意味する。

個々の研究者が望むすべての情報を含むデータベースを既製のデータベースに求めることが困難であるなら、自らの専門分野に精通している研究者が、必要とする情報を常に収集し、整理・編集したデータベースを作成し、それらを公開することが有効と考えられる¹⁰⁾。

4.3 オンライン検索における intermediary の役割

所属研究室で利用者自身がオンライン検索を実際に行っている回答者の一人は、「自分で検索した方が適切な検索語が使えるし、試行錯誤しながらできる」と答えている。誰が検索するのがよいのかという問題には一般的な結論は出ていないようである。しかし、エンドユーザが自分で検索できる態勢が整うであろう将来には検索仲介者は廃れる可能性は高い¹¹⁾。

オンライン検索者には、

(1) 情報要求を元来持っている人 (エンドユーザ、情報要求者)

(2)エンドユーザと情報との間にあって、たとえエンドユーザがいなくても、検索仲介者として実施する図書館員ないし情報スペシャリスト

(3)検索仲介者とエンドユーザの両者協同

という三つの類型がある⁹³。英国、米国でなされたいくつかの調査・報告によれば⁹³、(3)のタイプが最も望ましいと考えられている。P.W.Williams が1974-1975年に Manchester 大学で行った研究では情報要求者自身と検索仲介者自身が検索した38の質問を比較して次のように報告している。

仲介者が1検索あたりに要した時間は、情報要求者の約43%であった。

仲介者の検索件数は、情報要求者よりも約72%多い。

仲介者が検索した適合文献数は、情報要求者よりも約43%多い。

情報要求者の検索精度（検索された総文献中に占める適合文献総数）は59.9%、仲介者の検索精度は47.5%であった。

検索精度の点に関しては情報要求者が優っている。当図書室の利用者のが指摘した点もこのことに関連していると思われる。確かに、欲する文献を明確に分かっていて、自身の研究領域を把握している研究者自身が検索を行った方が効果的であろう。それにも拘らず、検索仲介者が必要とされるのは、使用されるアクセス方法、データベース、システムの多様さに起因している。このことに関して Henry は、「複数の検索システム、データベース、アクセス方法などを互いにつきあわせることによって、最小限共通する特徴を明かにすることが出来よう。もし基本的な機能が従来より使いやすくなり、一段と複雑な機能の使用が経験豊かな検索者に残されているならば、将来は、今迄以上に情報要求者が自分自身で検索し、職業的な仲介者は残された一段と困難で特殊な検索を行うことになろう」と述べている⁹³。

それ故、現時点で検索仲介者に要求されることは、下記の点であると考えられる。

(1)検索に必要な資料の収集・整備

適切なマニュアル、シソーラス、検索補助資料を収集し、追加・訂正を行い、検索に備える。

(2)データベースの理解・選択

収録主題、収録雑誌、レコードの内容、索引の特色、検索の網羅性・包括性、経費、利用の容易さなどの観点から検索主題に合ったデータベースを選択する。

(3)データベース提供システムの理解・選択

収録データベースの範囲、検索対象としたデータベースの構造、検索フィールド、検索機能、出力形式、利用の容易さなどの観点からシステムを選択する。

表1 システムの利用件数
(1982年 8月~1984年 12月)

	研究者	図書室	合計
DIALOG	88	165	253
TOOL-IR/ORION	60	18	78
UTOPIA	18	104	122
合計	166	287	453

注) DIALOGの1ファイル検索を1件として算出した。

表2 利用データベース

	研究者	図書室	合計
BIOSIS Previews	54	115	169
CA Search	95	78	173
CAB	7	17	24
AGRICOLA	7	9	16
Excerpta Medica	0	7	7
COMPENDEX	0	17	17
INSPEC	0	4	4
METADEx	0	5	5
SCISEARCH	2	11	13
NITS	0	4	4
Food Sci. Technol.	1	0	1
LC MARC	0	8	8
BOOKS IN PRINT	0	12	12
合計	166	287	453

注) DIALOGの1ファイル検索を1として算出した。

表4 検索語の組合せと検索された文献数

(件)

1	METHYLOBACTERIUM	22
2	PINK(W)PIGMENTED AND BACTERIA	7
3	TAXONOM? AND BACTERI? AND METHYLOBACTERIUM	1
4	CC=30000 AND METHYLOBACTERIUM	5
5	TAXONOM? AND BACTERI? AND BC=04714	6
6	CC=30000 AND BC=04714	34
7	METHYLOBACTERIUM AND PINK(W)PIGMENTED AND BACTERIA	1
8	BC=04714 AND PINK(W)PIGMENTED AND BACTERIA	3

表5 検索語の組合せと検索された文献数

(件)

1	(CC=3900? OR CC=135?) AND CC=10511 AND CC=32500	15
2	FOOD? AND RECOMBINANT AND DNA	7
3	FOOD? AND (CELL(W)FUSION OR PROTOPLAST(W)FUSION)	3
4	FOOD? AND TISSUE(W)CULTURE	7

表3 アンケート用紙

オンライン情報検索利用調査

京都大学農学部図書室

該当項目に○印を付けてください。

○印は何個でも結構ですが、最優先項目には◎をお願いします。

1. どのような理由でオンライン検索を利用なさいましたか。

- A 論文に引用すべき文献を知るため
- B 研究を始めるヒントを得るため
- C 進行中の研究を進める上での参考にするため
- D 自分の文献が抄録されているか知るため
- E 他の研究者の業績を調べるため
- F 研究室や研究グループの代表として文献調べをするため
- G 総説を書くときの参考にするため
- H 論文のレフェリーとしての仕事の参考にするため
- I 情報検索システムについて知りたかったため
- J その他（具体的にお書き下さい）

3. 検索結果はいかがでしたか

- A 求める文献が適切に検索できた
- B 必要な文献が検索できなかった
- C 期待以上の文献が発見できた
- D 検索もれが多かった
- E 不要な文献（ノイズ）が少なかった
- F 不要な文献（ノイズ）が多かった
- G データの質がよかった
- H データの質が悪かった
- I 料金が思ったより安かった
- J 料金が高かった
- K その他（具体的にお書き下さい）

2. 検索後の効果はいかがでしたか。

- A 研究すべき問題のヒントが得られた
- B 現在進行中の研究の進め方にヒントが得られた
- C 研究の進め方に自信がついた
- D 研究を中止又は方向転換するきっかけとなった
- E 二次情報誌をさがす時間が節約できた
- F 原論文を読む時間が節約できた
- G 論文の引用文献とすべきものが簡単に見付かった
- H 自分の研究分野に関し広い知識が得られた
- I 競争相手の研究者について常に情報が得られて便利であった
- J 不必要な文献（ノイズ）が多く、時間の損であった
- K 論文を読むのに忙しくなって、研究が進まなくなった
- L 他人の論文に気をとられ、目うつりがして研究の邪魔になった
- M その他（具体的にお書き下さい）

4. 検索語・検索式は適切だと思いますか

- 1. 思う
- 2. 思わない（具体的にお書き下さい）

参考文献

- (1) 山本毅雄. “TOOL-IR 利用者アンケート集計 (中間報告)” 東京大学大型計算機センターニュース, 10 (9), 1978, p. 10-14.
- (2) 山本毅雄ほか. “TOOL-IR の近況 (下)” 東京大学大型計算機センターニュース, 10 (10), 1978, p. 29-36.
- (3) 根岸正光, 山本毅雄. “情報システム評価の諸問題(1): TOOL-IR 利用者アンケートの調査分析を事例として” 図書館情報大学研究報告, 1 (1), 1982, p. 41-54.
- (4) Lancaster, F.W. Information retrieval systems : characteristics, testing and evaluation. 2nd ed. New York, John Wiley, 1978, 381p.
- (5) 細野公男. “情報の蓄積・検索”. 図書館・情報学概論. 津田良成編. 東京, 勁草書房, 1983, p. 113-148.
- (6) Sievert, MaryEllen and Bert R. Boyce. “Hedge trimming and the resurrection of the controlled vocabulary in onlines searching”. Online Review, 7(6), 1983, p.489-494.
- (7) Johnson, B.K. “A Comparison of online databases in relation to agricultural research and development”. Online Review, 5(6), 1981, p.468-479.
- (8) Meyer, Daniel et al. “comparison study of overlap among 21 scientific databases in searching pesticide information”. Online Review, 7(1), 1983, p.33-43.
- (9) 山崎稔, 小澤宏. “個人用データベースシステム PDB の使い方”. 東京大学大型計算機センターニュース, 15 (5), 1983, p. 33-38.
- (10) 天野要ほか. “一般研究者によるデータベース作成の支援と管理”. 情報管理, 27 (9), 1984, p. 787-799.
- (11) 井上如. “図書館教育からユーザ教育へ: 文献情報論からの考察”. ライブラリアンズフォーラム, 1 (3), 1984, p. 21-30.
- (12) 桜井宣隆. “オンライン検索における Intermediary の教育と訓練について”. 東京大学情報図書館学研究センター紀要, 1, 1982, p. 135-144.
- (13) Henry, W.M.ほか. オンライン情報検索. 細野公男訳. 東停, 丸善 MASIS センター, 1983.

研修レポート

文献情報センターと琉球大学附属図書館

Center for Bibliographic Information and the University of the Ryukyus

琉球大学附属図書館 本郷 清次郎*

要 旨

大学図書館には学術情報センターとしての機能が益々求められてきている。琉球大学附属図書館が、情報量の蓄積が少なく、またその大規模集積地からは遠く離れている沖縄という地域環境の中で、どのようにしてその機能の強化をはかっていくかは我々の大きな課題である。このレポートでは、そのような状況を把握するために琉球大学附属図書館および沖縄の図書館環境の実情を見て、それを打開する一方策として図書館ネットワーク利用による相互協力、その具体的なものとして東京大学文献情報センターとの接続について考えていく。

Abstract

It is my belief that in this information-oriented society the Library of the University of the Ryukyus has to strengthen its function as an academic information center in Okinawa where it is not easy to get enough information that we are eagerly in need due to distant location from highly integrated nodes of academic information.

In this paper the author tried to clarify the present Library situation at the University of the Ryukyus within the insular circumstances of Okinawa, then described library network and CBI network as important means to improve the present insufficient library function.

1. 琉球大学附属図書館

琉球大学は1950年に米軍占領下、首里城址に六学部、学生562名、教職員44名で発足した。その後、布令大学、琉球政府立大学などを経て、1972年、沖縄の日本復帰にともない国に移管されて国立大学となり現在にいたっている。1977年から84年にかけて、それまで那覇市内の首里と与儀にあったキャンパスの西原町・中城村への移転がおこなわれた。1984年5月1日現在で法文・教育・理・工・医・

*Hongo, Seijiro : University of the Ryukyus Library

農・の6学部22学科3課程2研究科と教養部、それに4学科からなる夜間コースの短大部があり、学生数約五千名、教職員数約千五百名である。

戦前に高等教育機関がなかったこと、戦後は長い間日本本土と隔絶され、地域で唯一の公立の大学であったこと、更に地元で財団が結成されてその運営に関与してきたことなどもあって、県民の琉球大学に対する「郷土の大学」意識は他の国立大学の場合に比べて独特のものがあるようである。

大学の建設と同時に附属図書館も着工されている。アメリカの大学では図書館は大学の中心に位置付けられているが、米軍占領時代には米国ミシガン州立大学から顧問団の来ていた琉球大学では、その影響からか図書館の大学内での地位は高かったようである。

図書館の現況は58年度で：

総経費：35,305万円 (うち資料費17,660万円)
 受入図書冊数：36,273冊 購入雑誌種数：2,302種
 年度末蔵書冊数：419,695冊 (本館：374,434冊、分館：45,261冊)
 (うち郷土資料：19,645冊) (本館のみ)
 開館日数：293日 入館者数：280,112人(956人/日) (")
 館外貸出：73,599冊 37,365人(251冊/日 128人/日) (")
 相互貸借(貸出0,借受：54)学外文献複写(依頼：2,376,受付：257) (")である。

琉球大学附属図書館の特色として主なものをあげると：(1)中央図書館制度、(2)全館開架閲覧方式、(3)学外者への開放、(4)郷土資料の収集、などがある。(1)(2)(3)については先にも述べたような設立の経緯があり、当初からアメリカ流の大学図書館の運営がなされてきたことの伝統であろう。

その中でも特に(3)は特記されよう。58年度実績で数字をあげると：

利用者総数：1,689人

職業別内訳：高校生：84, 大学生：295, 浪人：23, 無職：639, 会社員他：648

目的別内訳：学習：839, 資料調査：715, 閲覧その他：135

となっている。学外からの利用者はカウンターにある記録簿に記帳して入館することになっているが、出入口にBDS装置が備付られていてチェックなしに誰でも入ってこれるので実数はもっと多いものと思われる。

この背景としては、当館がその設立や運営に地域住民や沖縄在住アメリカ人の協力を得ていたこと、アメリカ軍政下でインフォメーション・センターと呼ばれていた時期もあること、更に、長い間当館以外にほとんど図書館が整備されていなかったことなどがあって、沖縄の中心的図書館として機能すると同時に地域の図書館としての機能も果していた。このような経緯があり、創立以来、伝統的に学外者にもほとんど無制限で開放されてきた。利用者に浪人や高校生、目的に学習という項目があらわれているのは、そのような事情による。しかし、大学が交通の便の悪い郊外に移ったことや周辺の図書館が整備されてきたことなどにより、このような利用は減っていく傾向にあり、かわって資料調査が増えてきている。

これらの点、更に(4)の郷土資料の収集に力をいれている点などは琉球大学附属図書館が地域と深く結びついていることを示しており、大学の地域開放が奨励される中においておおいに評価されるであろう。しかし、反面、このことは沖縄の図書館事情の悪さをも反映している。

2. 沖縄の図書館環境と琉球大学附属図書館

沖縄県には国立1、私立2の3大学図書館と2私立短大図書館がある。表1にそれらの概要を示してある。その表からだけでも県内の大学数の少なさ、個々の図書館の規模の小ささ、それゆえ地域としての情報の総蓄積量の少なさがわかる。更にその中では、他の国立大学と比べると大きいとはいえない琉球大学附属図書館が中心的調査・研究図書館として機能せざるをえない事情が浮かびあがってくる。

県内には公共図書館としては県立1（分館2）、市立7（うち那覇市立4）、町立1、村立2の図書館がある。表2でそれらのうちの上位3館の概要を示すが、いずれの館も規模が小さく、単独で十分に対処できる状態ではない。それゆえここでも琉球大学附属図書館が地域の中心的公共図書館としての機能を要求されていることがわかる。

表3では「日本の図書館1984」の公立図書館都道府県別集計で人口当りの数値の最も良い2都県と最も悪い2県をとりあげて沖縄県と比較したものである。沖縄県の場合はいずれも全国平均以下、それも常に最下位グループであることが目につく。逆に東京都の場合、絶対量で全国一なのは当然なこととしても、一人当りの数値でも殆ど一位なのは驚異に値する。ちなみに沖縄県を1として東京都の数値を指数表示すると、人口は9.9、蔵書冊数は4.4、受入冊数は4.8になる。単純に絶対量を比べると、東京都は沖縄県に対して蔵書冊数で既に44倍であり、これからも48倍の早さで増加していくことになる。情報格差は益々拡大する一方である。

図1と2は那覇周辺と東京中心部、各々一遍10kmの正方形の地域の中に大学・短大・国会・公立の各図書館がどのように分布しているかを図示したものである。最も恵まれた地域を比較の対象として選んだのは、今回のレポートの主眼が図書館ネットワークによる距離のハンディの克服ということであり、それができればどこにいても最も恵まれた地域とほぼ同様の恩恵を受けたいと思うのは当然のことであるからである。

この縦横10kmの中に那覇地区だと2大学図書、2短大図書館、4公立図書館しかないのだが、東京中心部では47大学図書館、10短大図書館、19公立図書館それに国会図書館がある。同一キャンパス内の分館や国会図書館の支部図書館、それに専門図書館などは余りに数が多く煩わしくなるので描き込んでない。

3. 文献情報センターネットワーク

これまでみてきた中で琉球大学附属図書館が図書館環境の整備の遅れている沖縄県の中において、学内外から地域の中心的調査・研究図書館としての役割を果たすことを期待されていることが明らかに

3. 文献情報センターネットワーク

表1. 沖縄の大学及び短大図書館(1)(2)(3)(4)

1984年4月1日現在

図書館名	学生数	職員	蔵書冊数	受入冊数	雑誌種類	資料費
琉球大学(含分館)	5,847	32+13	419(144)	28,618(11,266)	3,110(1,629)	176,600
沖縄大学	2,807	5+2	40(8)	3,146(114)	113(1)	10,288
沖縄国際大学	6,055	9+8	76(13)	8,305(2,671)	1,101(137)	37,355
沖縄キリスト教短大	669	3	38(5)	2,922(415)	141(33)	10,296
沖縄女子短大	860	1+3	21(2)	2,680(—)	275(2)	5,584
全国国立大学平均		17	366	10.8		117,855
私立大学平均		9	129	5.9		56,428
短期大学平均		3	34	1.9		8,112
単位	人	人	千冊	冊/千冊	種	千円

表2. 沖縄の公立図書館(1)(5)(6)

1984年4月1日現在

	奉仕人口	職員	蔵書冊数	受入冊数	雑誌種類	貸出冊数	予算	図書費
沖縄県立(本館)	1,150	10+8	113	13,406	54+18	33	205,604	32,000
"(二分館計)	—	4+2	65	3,434	43+17	36	—	—
那覇市立(中央)	301	3+8	70(25)	2,922(1,933)	96	130(97)	326,333	13,000
"(その他那覇市立三館計)	—	4+6	52(23)	11,062(7,446)	137	248(171)	—	—
名護市立崎山	48	1+2	48(15)	5,346(3,752)	68	72(62)	29,933	5,683
単位	千人	人	千冊	冊	種	千冊	千円	千円

表3. 公立図書館 沖縄県と他都県との比較(1)(7)(8)

1984年4月1日現在

都県名	人口	図書館設置率		人口 専任職員	蔵書冊数 百人	受入冊数 千人	登録率	館外貸出冊数 百人
		市区	町村					
東京	11,438	100(1)	60(3)	4.0(1)	164(2)	190(1)	18.1(1)	374(1)
	9.9	2.5	8.6	4.8	4.4	4.8	7.9	8.5
富山	1,116	100(1)	100(1)	6.3(2)	199(1)	145(2)	12.7(7)	237(2)
全国	118,602	86	15	10.8	87	92	9.8	164
	102.8	2.2	2.1	1.8	2.4	2.3	4.3	3.7
長崎	1,599	63(5)	7(8)	27.1(1)	51(3)	33(2)	2.2(2)	29(1)
宮崎	1,173	89(22)	11(20)	20.9(5)	56(7)	30(1)	1.5(1)	32(2)
沖縄	1,154	40(1)	7(8)	19.2(10)	37(1)	40(6)	2.3(3)	44(4)
単位	千人	%	%	千人	冊	冊	%	冊

注：(1) 典拠は「日本の図書館1984」

(2) 職員は前半が常勤職員、後半は非常勤職員。

(3) ()内は洋書・洋雑誌。

(4) 平均値の場合は受け入れ冊数ではなく購入冊数、単位は千冊。

(5) ()内は児童書。

(6) 職員は常勤職員のみ。前半は司書あるいは司書補。後半はそれ以外の職員。

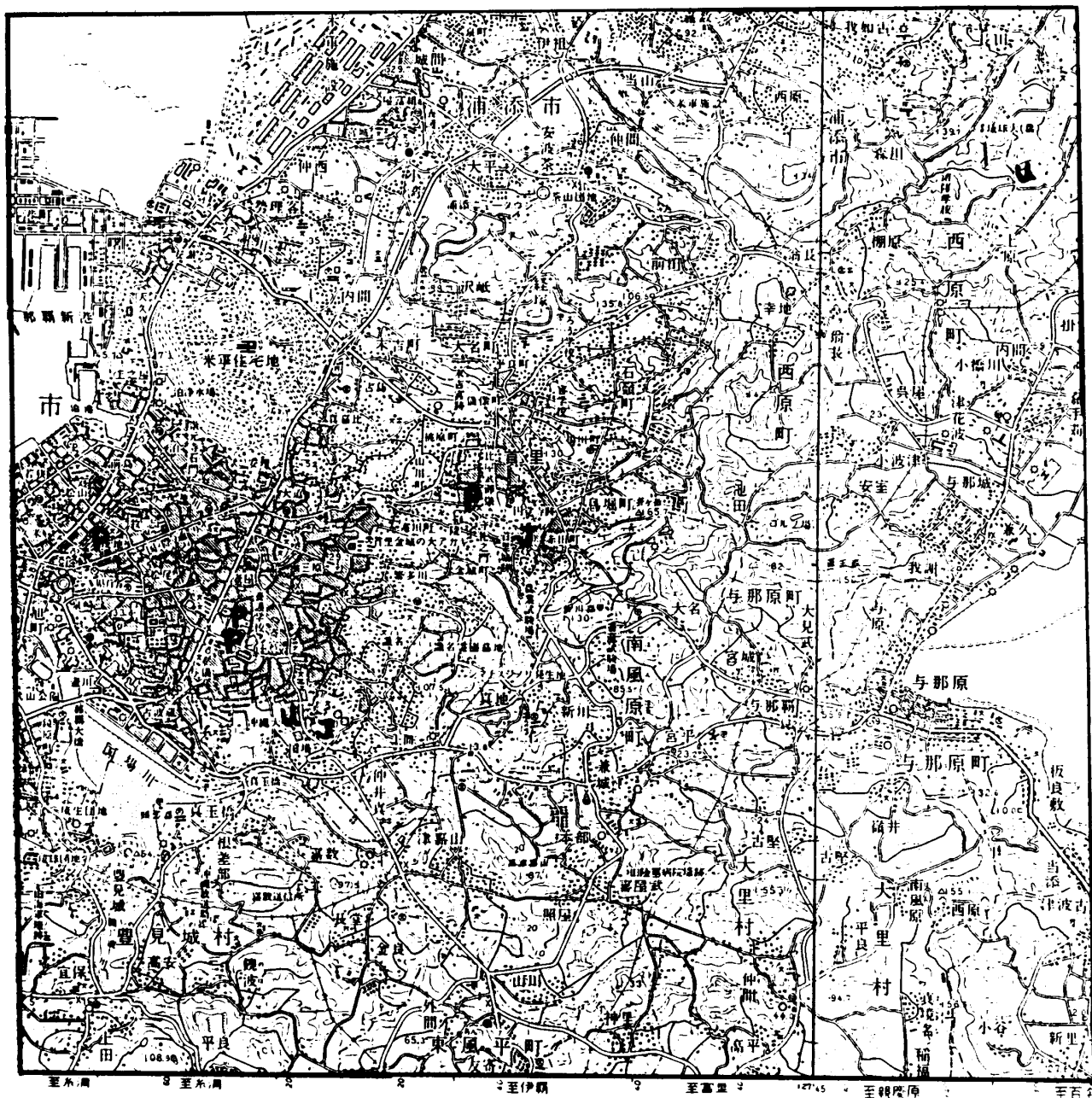
(7) ()内は47都道府県中、上位からの順。| |内は47都道府県中、下位からの順位。

(8) 東京及び全国の二行目の数値は沖縄を1としたときの指数表示。

図1 那覇周辺部の図書館

J:短大図書館 P:公立図書館

U:大学図書館

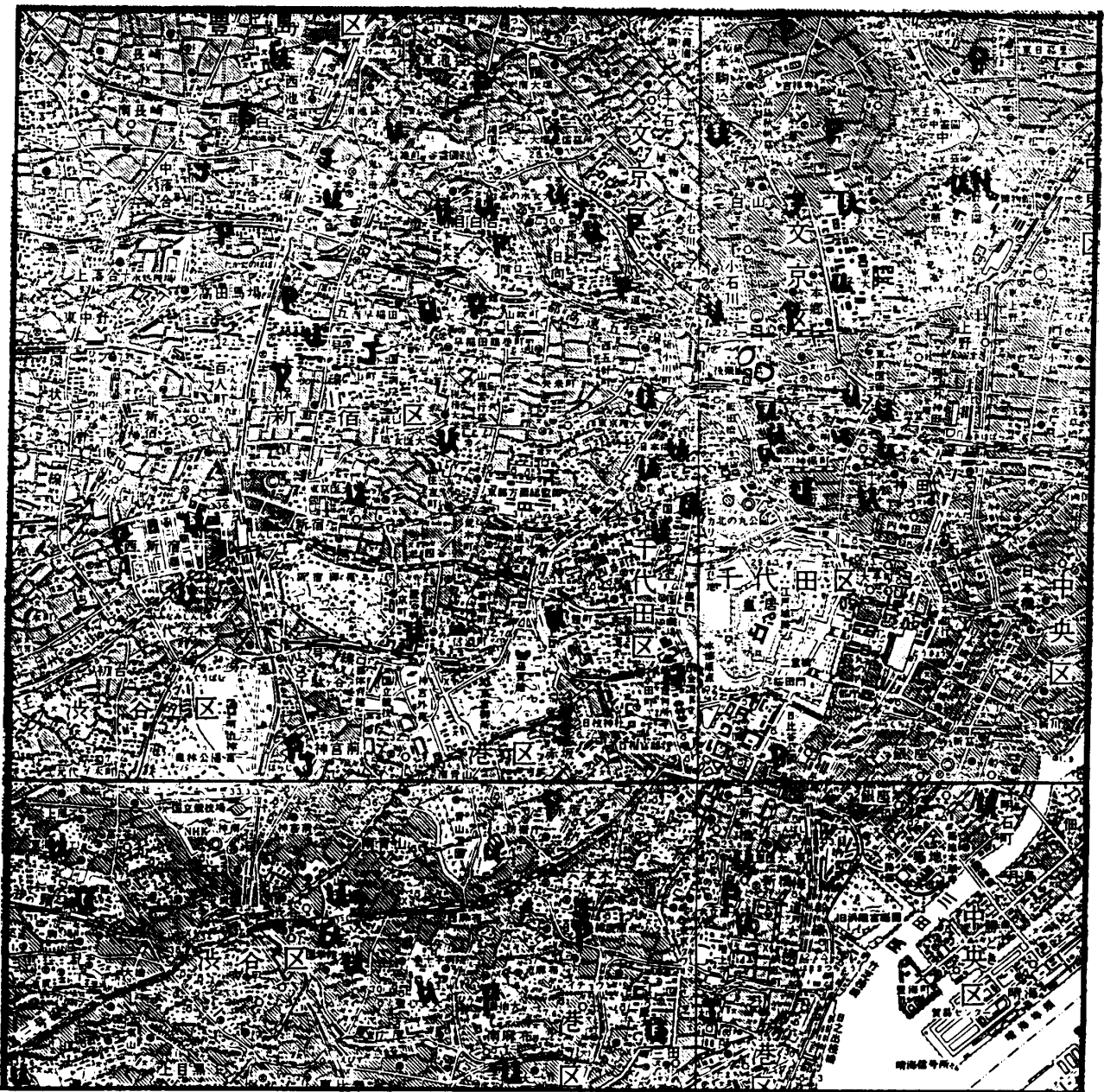


なった。

「今後における学術情報システムの在り方について」の答申の中で大学図書館に対する要望がある。ひとつには学術情報システムと情報の最終利用者である研究者とを結ぶ窓口として資料を迅速・的確に利用者に提供するため、そのサービス機能を格段に改善・強化する必要があり、そのためには図書館の事務処理システムを更に改善して利用者の情報需要に十分にこたえ得るサービス機能を備えることが必要であること。二番目には、情報検索システムが真に効果的に機能するための大学間および図書館間の相互協力の必要性が述べられている。すなわち、各種のデータベースの管理と情報提供サービ

図2 東京中心部の図書館

J:短大図書館 N:国立図書館
P:公立図書館 U:大学図書館



スを行う機関・大学図書館等が全国的観点から相互に連結されている必要があり、各大学の図書館間の相互協力を一層強力に推進していくことが重要である。個々の図書館が当該機関における情報需要のすべてにこたえることは今日においては不可能であるので、個々の図書館におけるサービス活動は図書館間の相互協力の上に成り立っていると云っても過言ではない。その意味において図書館間の資料の円滑な相互利用の体制の強化を更に進める必要がある、と述べられている。

このことはまさしく琉球大学附属図書館にも求められていることでもある。これまで見てきたように、どちらかと言えば充分でない環境の中で琉球大学附属図書館が大学図書館の使命である学内の学

術情報センターとして機能していくにはこの方向に進んで行く以外にはない。

ところで上記の相互協力という事は究極的にはネットワークということになる。

文献情報センターセミナーの講義の中でセンターの井上教授が相互協力には三つの段階があることを述べておられた。

第一の段階は相互利用の段階である。この段階の特徴としては:(1)二館の1対1の対応関係(2)トラザクションがすむと関係は解消される(3)文献の利用のみが目的である、などがある。

第二の段階は図書館組合 library consortium の段階である。この段階の特徴として:(1)多数館の相互利用関係(2)関係が恒常化してくる(3)文献の利用以外の業務を含む、などがあげられる。様々な地域・設置母体の図書館が参加し、協会の発足をみ、共同で見学・研修・目録作成なども行う段階である。

第三の段階としてネットワークがくる。第二の条件に更に(1)ハードウェア:コンピュータ+通信、(2)研究開発能力の集中という条件が整った段階である。第二と第三の段階の間には大きな段階差がある。以前から何度も計画されてきたが、殆が失敗に終わっている。しかし、現在、条件が熟してきたので、ようやく現実化可能な段階に達した。第一第二の段階では収書・予算・図書館員・利用者などに関して個々の図書館の存在を侵さないが、第三の段階ではネットワークの論理が先行し個々の図書館がそれに規制される。予算も図書館員もそれに属し、利用者はどこでも手近な図書館を利用することになる。

このネットワークの段階で行う業務の種類としては次のようなものがあげられる:

(1)総合目録の作成(2)共同目録作業(3)データベース共有(4)相互貸借(文献複写を含む)(5)職員の共同研修。これらの業務は相互に結びついている。すなわち、共同目録作業によって総合目録データベースの構築と作成が行われ、それを利用して相互貸借あるいは文献複写が行われる。共有するデータベースは総合目録に限らず市販の情報検索用のデータベースも含まれるのであるが、その一次資料の入手は総合目録を利用して相互貸借が行われる。このようなことを参加しているどの図書館でも行えるようにするには職員に一定のレベルの能力が要求されるが、そのためには共同して職員の研修を行う必要が出てくる。

文献情報センターはこのようなネットワークを志向しているようである。文献情報センター、学術情報システムについては様々なところで取上げられているのでここでは割愛する。

琉球大学附属図書館は、ところで、先程も述べたように、地域の中心的図書館として期待されている訳であるが、独自の情報蓄積量も少く、かつ情報大量蓄積地域から遠く離れているというハンディを打開する方法としては、このネットワークに参画することが最上の方策であろうと考える。

文献情報センターへの接続の形態としては、(1)RC 集中型、(2)ML 透過型、(3)RC 透過型、(4)N-1方式、(5)PAD 方式がある。琉球大学附属図書館の場合は地理的条件から(4)か(5)の直接 NC につなぐ形態を考えた方がより現実的であろう。(5)の開発がすこし先の事になるのに対して(4)の形態ならまもなく接続が可能になる。ただし、この接続形態では中型以上のコンピュータが必要となる。他にも

N-1に加入、FEP、通信用ソフトウェア、通信回線、端末を準備せねばならないなどの前提条件がある。幸にも琉球大学附属図書館の場合は学内計算センターのコンピュータ FACOM M-180ⅡAD と V-830を使って閲覧システムの機械化を行っており、またその他の部門の機械化のためにインテリジェント端末である9450-Ⅱを使う ILIS の導入をきめているので、この条件はほぼみたしている。機は熟していると言えよう。

(1986年1月13日に文献情報センターと琉球大学附属図書館の接続テストが完了し、現在、学術情報センターとの間で業務モードにより運用中である)

参考文献

1. 琉球大学. “琉球大学の概要”昭和59年度
2. 琉球大学附属図書館. “図書館年報”昭和55～58年度
3. 国立大学図書館協議会東京地区協議会東京地区国立大学図書館ネットワーク研究会“東京地区国立大学図書館ネットワーク研究会報告57年度”1982
4. 鬼頭當子. “国際基督教大学における新しい図書館へのアプローチ”1983 ICU Faculty Retreat
5. 文部省学術国際局情報図書館課. “今後における学術情報システムの在り方について(答申)”1980
6. Markuson, B. E. “Library Networks:Progress and problems” in Hammer, D. P., ed., “The Information Age:Its Developments, Its Impact”Scarecrow Press 1976
7. 文部省学術国際局情報図書館課. “昭和57年度学術情報センターシステム開発調査概要”1983. 3
8. 文部省学術国際局情報図書館課. “昭和58年度学術情報センターシステム開発調査概要”1984. 3
9. 文部省学術国際局情報図書館課. “大学図書館業務の電算化”1984. 3
10. 学術月報 vol.30,no11 (1980)
11. 東京大学文献情報センターニュース№5 (1984. 11)
12. キルガー, F. G.“OCLC の現状と将来”図書館雑誌 vol.78,no8 (1984)
13. JICST.“科学技術情報ハンドブック改訂版”1982. 9
14. 電々公社. “データ通信回線”
15. 電々公社. “新データ網サービス”

研修レポート

Recon から総合目録データベースへ

Toward a National Bibliographic Database in Japan: From the Point of Retrospective Conversion

立教大学図書館 牛崎 進*

要 旨

日本では Recon の意義や方法について言及されることが少ない [1], [2]。これはローカルに作成される書誌レコードが、外部世界と交渉を持たないできたことに一因がある。そこで、本レポートでは、先ず、Recon そのものの理解の一助として、Recon をめぐる北米と日本の現状を報告する。そして、OPAC を目指す図書館の Recon プロジェクトによって形成される書誌データベースと、図書館間を連結させて広範囲な書誌コントロールを可能にする地域的、全国的総合目録データベース（一体化もしくは仮想化されるシステムとして）との相互関係について論じている。

Abstracts

Issues reported in this paper are as follows; 1)that describe a short history of retrospective conversion in U.S., 2)that prepare general guidelines when planning retrospective conversion, 3)that discuss that the CBI(Center for Bibliographic Information, University of Tokyo)could be also a recon utility by having some systematic enhancements which enable users to convert their card catalogs to machine-readable from easily, 4)that propose that it is necessary to make a national plan for retrospective conversion in Japan.

1. はじめに

Recon(Retrospective conversion:遡及入力) は、図書館機械化の歴史の中で、機械化の目的に応じてさまざまに姿を変えながら取り組まれてきた。その歩みは、目録の利用形態、メディアをめぐる時代の要請とともにあったといってもよい。今日、主に OPAC(Online Public Access Catalog)用の書誌データベースを過去に遡って統合化するための手段として、Recon はいっそう注目を集めている

*Ushizaki, Susumu : Rikkyo University Library

が、この趨勢は、やがて日本の図書館においても常態となる筈である。

2. Recon とは

2. 1 Recon の目的

D.Reynolds は、次に掲げる項目の1つ以上の理由から、Recon を実施する図書館が多いと指摘している [3]。

- a)総合目録（地域、州、地方）に所蔵情報を追加するため、b)COM 目録を作成するため、c)開発あるいは購入したサーキュレーション・システム、オンライン目録や統合システムに格納するデータベースを形成するため、d)将来のアプリケーションに備えて。

総合目録を目的とする Recon のみが、単館を越えた共同 Recon を実施することになるが、通常、図書館はこのプロジェクトによってカード、COM, MT など入手することが多い。その代表的例を、イリノイ州の LCS(Library Computer System) [4] とカリフォルニア州の MELVYL [5] に見ることができる。

Carter と Bruntjen は、その著「データ交換」[6] で変換する目的を次のカテゴリーに分けて説明しているが、これは Recon にもそのまま該当するものである。

- a)統合システム、b)目録、c)サーキュレーションと蔵書管理、d)受入と IPF（インプロセスファイル）、e)逐次刊行物管理、f)リソースシェアリング、g)経営情報。

Recon 計画を短期的、局所的要請によって実施せざるをえないこともあるが、その場合でも、将来のコンピュータ・アプリケーションの可能性を大にするために、通常、次の事項が検討され、決定される。

- a)MARC フォーマットの取り込みかた、b)フル・レコードに変換するか、c)最初のアプリケーションは何か、将来の目標は、d)書誌同定のキーとなる全国書誌番号類(LCCN, ISBN, ISSN, JP-NO,国図番号など)の入力（カードにこの種の番号が記載されていない時、調査するか）。

例えば、OPAC のための Recon では、すべての所蔵をフル・レコードに変換する方法を取ることになろう。しかし、そのためには、多大な資源を費やし、長期化するプロジェクトを推進することに対して、図書館内で合意を形成しておくことが必要となろう。リソース・シェアリングを目的とする Recon では、図書館間の合意形成は別の形をとるかもしれない。例えば、イリノイ州の LCS のための Recon では、図書館の蔵書数に応じて OCLC もしくは Electronic Keyboarding, Inc.や Blackwell North-America といったベンダーを使い分けてプロジェクトを進行させた [7]。この場合の合意として、LCS に形成されるデータベースでは、書誌は簡略に、所蔵は詳細にという決定がなされた。その結果、1984年4月現在、LCS は、740万のモノグラフと38万の逐次刊行物の書誌および1300万の所蔵レコードを有する北米有数の州ネットワークとなった。

2. 2 Recon 小史

北米の Recon 史を3つの時代に分けて辿ってみる。しかし、その軌跡を詳細に分析しようとすれ

ば、目録のメディア、機械化、変換フォーマット、書誌情報ユーティリティなどの発展史と共に記述されるべきであろう。

2. 2. 1 第一期 (1950年代から60年代初め)

冊子目録作成のための Recon で、シエルフ・リストをそのまま変換した。1951年、King County Library (シアトル) が、IBM の Tabulator とパンチカードを使って冊子目録を作成したのが始まりである。1953年には、Los Angeles County Public Library が、同じ方法で青少年用図書16万タイトル (160万枚のカード) の総合目録を作成。続いて、ハーバード大やスタンフォード大へと広がり、University of California Union Catalog Supplement や New York Public Library の冊子目録で全盛期を迎えた [8]。

2. 2. 2 第二期 (1960年代半ばから1970年代後半)

この期の特徴は60年代半ばから70年代初めにかけて多くの図書館で開発されたり、ベンダーが市場に出し始めたサーキュレーションのシステムで使用される簡略データによる変換であった。変換フォーマットの標準化という意味で特記されねばならないのが、1968年に始まった LC-MARC テープの頒布である。MARC レコードが多くの学術図書館の機械化システムで使われ出し、あるいは機械化を促すようにもなり、やがて OCLC を通じて普及することによって、今日に至る変換フォーマットの標準化へとつながることになったからである。しかし MARC レコードを技術的あるいは経費的に利用できなかった殆どの図書館では、その変換フォーマットは、簡略なものか非 MARC フォーマットのままであった。この時期、その後の Recon の理念的、技術的ベースを提示することになる2つのプロジェクトが、LC で企画され、学術図書館の多くがそれに参画することになった。RECON Pilot Project [9] と COMARC [10] がそれである。

2. 2. 3 第三期 (1970年代末～)

フル・レコードでの変換が開始されたが、これを可能にするいくつかの制度的、技術的背景があった。例えば、書誌情報ユーティリティは、より低コストでの Recon サブ・システムを提供し始めている [11], [12], [13], [14]。また、CLR (Council on Library Resources) や多くの財団からのファンドによって、大規模な Recon プロジェクトが展開されたのも、多様なベンダーが特にその OPAC のシステムを売り込むために Recon 市場に参入し始めたのもこの時期からであった。

Recon 用の変換フォーマットは、書誌情報ユーティリティでのオンライン目録作業が普及するにつれ、いっそう LC-MARC に準拠するようになった。しかし、Recon においては、これは入力のコスト高を招くため、フォーマットを部分的に省略する図書館が一般的となった。このような図書館ごとのばらつきを調整するために、研究図書館協会 (Association of Research Libraries) は、1985年、モノグラフ用の Recon ガイドラインのドラフト [15] を公表し、その中で Recon レコードのミニマムを設定するに至った。

LC も Recon の動向に大きな役割を果たそうとしている。1985年に発表した新たな MARC サービス (Select MARC: Retrospective Conversion) である。これはユーザが希望するレコードを LC カー

ド番号などで抽出して提供しようとするものである [16]。

2. 3 Recon のプランニング

計画段階で、検討される事項として次の項目がある。

- a) 目的・目標の設定、確認、b) 変換対象の決定 (カード、冊子目録、ファイル)、c) 変換方法のオプションの洗い出しと比較検討、d) 費用対効果の分析、e) レコードのマッチング、マーキングのアルゴリズムの検討、f) 典拠コントロール・システム、g) レコードの質のコントロール、h) 逐次刊行物の処理、i) 再分類、再カタログニング (ISBD の採用など)、j) 所蔵情報のフォーマット。

以下では、オプションとして注目を集めている REMARC とコストの問題のみを概略的に取り上げる。

2. 3. 1 REMARC システム [17], [18]

REMAR (REtrospective MARC) システムは、Carrollton Press が LC の Cumulative title index to the classified collections of the Library of Congress (1978, 132巻) を編集する一方で、1979年からコレクション中の非 MARC タイトルを遡及入力し、1984年末現在、LC-MARC を含む約700万件を収録するデータベースとなっている。その内訳は、Carrollton の変換によるローマ字レコード500万件と1968年以降の全 LC-MARC レコード200万件である。完了は1985年末の予定である。

変換レコードは、次のように簡略化されている。

- a) 固定長のいくつかのコードと DC ナンバーを省略、b) 245の書名関連情報と責任表示をしばしば省略、c) 300の形態エリアにはないものもある、d) 490のトレースされないシリーズを完全に省略、e) ローマナイズ注記を除き、注記を省略。

調査時点で、このシステムを利用しているのは、UTLAS、カリフォルニア大学 (バークレイ)、ジョンズ・ホプキンス大学・エジンバラ大学等である。検索のみ利用であれば、DIALOG でも可能である。

2. 3. 2 Recon のコスト

コストを算出する基本的な方法について、Butler, Aveney と Scholz が、詳細な検討をしている [19]。彼らによると、コストに波及する要素は以下である。

- a) 変換されるコレクションの規模、b) レコードの言語、c) レコードのフルネス、d) コピー・カタログニング出来ないタイトルの割合 (いわゆる非ヒット率)、e) スタンダード (目録規則など) との互換性、f) 要員。

しかし、実際にコストを公表している論文は少なく、変換タイトル数やヒット率、レコードの仕様書等で伺い知るしかない。

Peter と Butler は、Recon のさまざまな手法でおおよそそのコストを算定出来る方法が明らかにされていないとして、算定モデルの一例を説いている [20]。2人によれば、最も効果的な方法の追求は、必ずしも最低コストによらず、そのプロジェクトで設定した標準セットに最も合致する方法が、最低のコストと考えられるべきだと述べている。費用対効果分析の難しいところである。

3. Recon をめぐる動向

3. 1 北米における Recon 勧告

1984年、全米から29人の関係者が会して Recon 集会が開催された〔21〕。その目的は、CLR が同年に刊行した Recon に関するレポート〔22〕について議論し、この問題について今後の見通しを得るためであった。広範に及ぶ議論の末に9つの勧告がまとめられているが、近い将来の日本における Recon の共同化の可能性を示唆していると思われるので、そのいくつかを掲げてみたい。

- a) 研究図書館は、勧告を受け入れ、プログラムの推進に精力を向けるべきである。また、特殊コレクションも対象にすべきである。
- b) ARL は、プログラムの定義や運営上の責任を持つべきである。
- c) 勧められる方法は、LC 分類表に基づく主題ごとの変換か、非常に大規模なコレクションの変換である。
- d) 共同プログラムに参加する機関は、同意するスタンダード群に準拠してレコードを作成し、共有すべきである。例えば、MARC フォーマットに従い、標目は AACR2、件名は LCSH と MeSH によるべきである。
- e) Linked Systems Project を実現する通信プロトコルを採用して、いくつかの分担目録サービス・システムで作成されるレコードを共有できるメカニズムを提供すべきである。
- f) ARL の Bibliographic Control Committee は、共同プログラムを支援するために、多様なファンドを提供していくべきである。

3. 2 日本における Recon の可能性

国立国会図書館は、Japan-MARC テープを1981年以来頒布しているが、すでに和図書の Recon にも着手している。約63万タイトルに及ぶ計画の概要は、次の通りである〔23〕。

- a) 明治期まで遡及し、最終的に全国書誌として機能させる。
- b) 入力は、旧帝国図書館蔵書目録及び国立国会図書館蔵書目録収載分の目録カード（昭和23～43年）に拠って行なう。
- c) 入力は、原則として各蔵書目録類または目録カードに記載に基づいて行い、各データの統一は行わない。
- d) 入力計画（基本的な作業は職員、データシート作成は委託）。

第一次（昭和23～43年分）：276222タイトル。

第二次（昭和2～24年分）：136251タイトル。

第三次（明45～大15、大正期乙部）：103787タイトル。

第四次（明治期刊行図書目録）：120000タイトル。

なお、第一次計画は、蔵書目録第三期末分（昭和44～51）の入力完了後の65年10月に開始予定とのことである。

リソース・シェアリング型 Recon の可能性を最も期待されているのが、文献情報センターであ

る。昭和60年度の同センターの開発項目の中に、バッチによるデータ還元というメニューがあるが、これは、接続館に対して、接続館が登録したレコードをバッチ検索し、MT で還元するサービスのことである [24]。しかし、将来の接続館の Recon へのニーズを考慮すれば、北米の書誌情報ユーティリティのように、サーチキーを図書館側で入力し、それをセンターでバッチ検索し、その結果を図書館側で編集する流れが必要と思われる。この場合でも、図書館側で典拠のリンクづけ、所蔵の登録などをオンラインで行うことになろう。

このようなシステムの要件の他に、Recon を誘引するものとして、データベースの規模がある。参照 MARC の種類を増やすことはその一端であるが、現時点では、接続館数を着実に伸ばし、システムの安定利用を計ることにウエイトがおかれているようである。いずれ、Recon のファシリティとして同センターがまさに中枢となることによって初めて総合目録データベース形成への展望がひらかれることになろう。

4. Recon から総合目録データベースへ

1960年代に北米などで本格化した図書館機械化のムーブメントは、1980年代に入って、1つの成果としてオンライン目録を提供し始めた。すでに、OPAC という用語も定着したようである。まだまだ研究、開発されねばならない問題も多いが、そのパワーも広く認められている。

オンライン目録を書誌的に可能にするのは、統合化された書誌データベースであることは論を待たない。レコードのレベルにしても、少なくとも、カードと同等の情報が表示される必要がある。この“統合”の意味するところは、すべての所蔵が、単一の検索メディアで検索でき、資料が利用可能であるということである。逆説的に言えば、オンライン目録が登場する図書館では、レコードの更新はあっても、Recon そのものはなくなっていなければならないということであろうか。

P. Wilson は、オンライン環境下における目録のありかたを確認するてだてとして、Cutter の目録法の原則に引き合いに出している [25]。周知のように、Cutter のいう原則の第一は“find a book”であり、第二は、“show what the library has”であるが、第二原則を show what libraries in the country have”と読みかえたフレーズで、本レポートを締めくくりたい。

目録の持つべき要件の第一は、単一のメディアで全所蔵を示すことにあることはすでに述べた。この要件をどうにか満たそうとしているうちに、ユーザの要求は近年ますます、単館の境界を越えたところへと向かっている。従って、次世代型のオンライン目録は、地域や国内をネットする総合目録データベースとの会話を必要とすることは当然の帰結であり、そのように予測もされている。しかし、このようなデータベースをどのようにして形成してゆくかについては、残念ながら明確な戦略があるわけではない。ただ、自明と思われるのは、書誌情報ユーティリティのデータベースを利用することによって、最も恩恵を受けるものの1つに Recon があり、他方、図書館での Recon プロジェクトを誘引する顕著なもの1つに書誌情報ユーティリティのデータベースサイズがあるということである。日本では、Recon を図書館の外部から促す条件も出揃っていないが、いずれ、CLR の Recon 集会の

議論の内容をかみしめる時が間近にせまっているように思われる。

注・引用文献

1. 松井幸子. 遡及的書誌情報データベースの作成のためのファイル統合. 図書館短期大学紀要、No.18, p. 31, 33 (1980)
2. 松戸保子. これからの大学図書館. 現代の図書館、Vol.23, No.3, p.172-179(1985)
3. Reynolds, Dennis. Library automation. New York, Bowker, 1985, p. 282.
4. Hurley, Bernard J. LCS: automated resource sharing in Illinois. Resource Sharing and Library Networks, Vol. 1, No.1, p. 67-77(Fall 1981)
5. MELVYL については、Information Technology and Libraies の Vol. 1, No.4(Dec. 1982) と Vol. 2, No1(Mar. 1983)で特集されている。
6. Carter, Ruth C. ; Bruntjen, Scott. Data conversion. White Plains, N.Y., Knowledge Industry Publications, 1983, p. 1.
7. Brown, Doris R. Retrocon for LCS in Illinois academic libraries. Information Technology and Libraries, Vol. 3, No.3, p. 274-282(Sept. 1984)
8. Malinconico, S. Michael; Fasana, Paul J. The future of the catalog White Plains, N.Y., Knowledge Industry Publications, 1979, p. 20, 48-49.
9. Avram, Henriette D. RECON Pilot Project. Washington, LC, 1972, p. 13.
10. Martin, Susan K. Library networks, 1981-82. White Plains, N.Y., Knowledge Industry Publications, 1983, p. 21.
11. Microcomputer-based retrospective conversion service now available OCLC Newsletter, No.156, p. 12(Feb. 1985)
12. 丸善 MASIS センターのアナウンスメントによる。
13. WLN introduces Micro-Recon. Information Technology and Libraries, Vol.4, No.2, p. 179-180 (Jun. 1985)
14. Initial funding completed for RLG Recon Program. Information Technology and Libraies, Vol. 4, No.2, p. 181(Jun. 1985)
15. Guidelines proposed for retrospective conversion of bibliographic records of monographs. LC Information Bulletin, Vol. 44, No.12, p. 59-60(Mar. 25, 1985)
16. Libray introduces new service for retrospective conversion. LC Information Bulletin, Vol. 44, No25, p. 129-130(Jun. 24, 1985)
17. Severtson, Susan. REMARC. Program, Vol. 17, No.4, p. 224-232(Oct. 1983)
18. Drake, Virginia;Smith, Mary Paige. Retrospective conversion with REMARC at Johns Hopkins University. Information Technology and Libraries, Vol. 3, p. 282-292 (Sept. 1984)

19. Butler, Brett; Aveney, Brian; Scholz, William. The conversion of manual catalogs to collection data bases. *Library Technology Reports*, Vol. 14, No.2, p. 177-197(Mar.-Apr. 1978)
20. Peters, Stephen H.; Butler, Douglas J. A cost model for retrospective conversion alternative. *Library Resources and Technical Services*, Vol. 28, No.2, p. 149-151, 154-162(Apr./Jun. 1984)
21. Jones, C. Lee. Issues in retrospective conversion. *College and Research Libraries News*, Vol. 45, No.10, p. 528-532 (Nov. 1984)
22. Reed-Scott, Jutta. *Issues in retrospective conversion*. Washington, CLR, 1984. 57p.
23. NDL の部内メモによる。転載の許可をいただいた。
24. 文献情報センターニュース, No.8, p. 24-25 (1985. 9)
25. P. Wilson の“Catalog as access mechanism”と題した発表。Hirschon, Arnold. *Prospects for the online catalog*. *RTSD Newsletter*, Vol. 7, No.5/6, p. 41(Sept./Dec.1982)にその要旨が報告されている。

研修レポート

書誌所在データベースと二次情報データベースの統合に向けて —— モデル・システムの作成

Toward the Integration of 'Database for Bibliography and Location' and
'Secondary Information Database'— Construction of a Model-System

名古屋大学附属図書館 尾城 孝一*

要 旨

文献情報の検索から原文献入手に至る一連のプロセスを効率的に展開するシステムは考えられないものか、というのが今回のレポート作成の動機である。「書誌所在情報 DB」と「二次情報 DB」との統合というテーマを追求し、Model204を利用してモデル・システムの作成を行った。まず、E-Rモデルを援用し5つのファイルを設計し、次に、実験データを入力し、検索するプログラムをModel204のユーザ言語によって作成した。最後に、文献情報センター・システムにおける統合システムの展開について若干の言及を行った。

Abstract

It is necessary to integrate 'database for bibliography and location' and 'secondary information database' for getting documents rapidly. The main theme of this report is the integration of these two databases. In pursuit of this theme, I tried to construct a model-system by making use of DBMS, Model204. Main part of this paper is a report of the constructing this system. In addition, I made mention of the further development of this integrated system in CBI system.

1. はじめに

今日、雑誌論文等の一次文献を入手するためには、どのようなプロセスが必要とされているだろうか。利用者は、まず、二次情報検索サービスや抄録・索引誌等に頼りながら必要な文献を見つけ、その原掲載誌名を知る。続いて、例えば「学術雑誌総合目録」や各図書館の所蔵目録といったツールによって、原掲載誌の所在情報を得る。さらに、文献複写依頼を経てようやく必要とする文献にたどり着くことができる。このように現在のところ、文献入手に至るまでには煩雑な手順を踏むことがどうしても避けられない。

*Ojira, Koichi : Nagoya University Library

こうした煩しさを解消し、文献情報の検索から原文献入手に至る一連のプロセスを効率的に展開するシステムは考えられないものか、というのが今回の研修レポート作成の主たる動機であった。図書館ネットワークによって構築される「書誌所在情報」と抄録・索引作成機関等が提供する「二次情報」との有機的結合というテーマを、自分なりに考えた結果がこのレポートとなっている。

書誌所在情報データベースと二次情報データベースの統合の可能性をさぐるために、まずデータベース・システム設計の際に利用される E-R モデルを援用して、図書館目録の世界をとらえ直してみようという作業を行った。続いて、このモデルに基づいて、書誌所在情報 DB と二次情報 DB との有機的結合を目指したプロトタイプのシステムを実際に作成した。最後に、文献情報センター・システムにおける統合システムの展開について若干の考察を加えた。

2. E-R モデルによる目録世界の把握

2. 1. E-R モデル

データベースの概念設計の段階で、データモデルというものがよく利用される。今日までにさまざまなモデルが提唱されているが、E-R モデル(Entity-Relationship Model:実体-関係モデル) は、その中でも最もよく普及しているものであろう。¹⁾

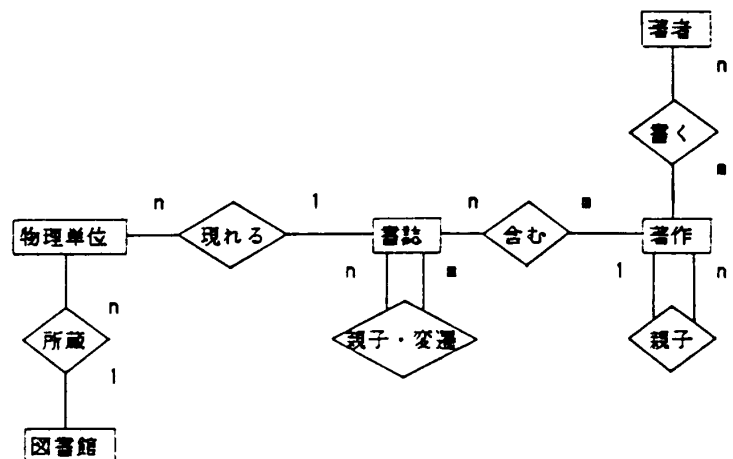
データ・モデルは、データの世界を記述するものである。E-R モデルは、実体(Entity)と実体間の関係(Relationship)を明示することにより、世界を表現しようとする。E-R モデルでは、実体は長方形の中に実体名を書いて表す。実体とは、データベース・システムで情報を管理しようとする物、人、概念、行為などのことである。関係は、菱形の中に関係名を書いて表現する。互いに関係し合う実体どうしを線で結んで、線上に二つの実体の対応比を示す添字を付する。関係のなかには同一実体間に現れるものもあり、これは内包関係と呼ぶ。

2. 2. モデルによる目録世界の把握

E-R モデルを使って、図書館目録の世界をとらえ直してみると、おおよそ図 1 のようになる。²⁾

2. 2. 1. 実体の説明

図 1 E-R図



①物理単位

一冊一冊の本、破らない限り分割することのできない一つ一つの本を物理単位と定義する。いわゆる複本であっても別の物理単位とみなす。本の登録番号というものは、個々の物理単位を区別するために付与されると考えられる。

②図書館

物理単位を収集、保存し、利用者の要求に応じて提供する機関で、名称、住所などの属性を持つ。

③書誌

書誌というのは抽象的実体であり定義しにくいのであるが、出版という行為によって生み出された、同じ物理単位の集合が書誌であると考えられる。複本というものは、書誌が同じで物理単位が異なるものと言えよう。書誌は、書名、著者名、出版者などの属性を持っている。

④著者

著者は明確な実体としてとらえることができる。著者は著作すべき内容を企画し執筆した人物あるいは団体であり、名称、生没年、所属機関などの属性を持つ。

⑤著作

著者が書いたもの。著者が書くのは「書誌」ではなく「著作」である。内容（主題）は著者の持つ属性である。

2. 2. 2. 関係の説明

①書く（著者:著作）

著作には内容という属性があり、著者はこの内容を記述した者であると定義する。この両者の間には「書く」（あるいは「書かれる」）という関係が考えられ、その対応比は n 対 m である。

②含む（書誌:著作）

多くの場合、一つの書誌は一つの著作を「含む」、あるいは、一つの著作は一つの書誌として「現れる」。書誌と著作の対応比は 1 対 1 である。

ところが、論文集、会議録、アンソロジーの類では、一つの書誌の中に複数の著作（論文等）が「含まれる」という関係が成り立つ。今回のレポートで扱う雑誌とその収録論文の関係も同様である。この場合、書誌と著作の対応比は 1 対 n となる。

一方、書誌と著作の対応比は、 n 対 1 にもなりうる。例えば、一つの小説が、単行書、全集、選集といったさまざまな書誌に現れるという現象が起る。また、雑誌に発表された論文が、後に論文集に再録されるというケースも見うけられる。さらに、古典作品を考えてみても、一つの著作がいろいろな書誌に拡散する可能性があることは明瞭である。一例を挙げると、「源氏物語」という著作は、異なる名称をもつ複数の書誌として現れてきた。図書館目録における統一書名典拠コントロールというのは、このように複数の書誌に拡散した著作を取りまとめるためのメカニズムであると考えられることができよう。³⁾

③現れる（書誌:物理単位）

一つの書誌は複数の物理単位として「現れる」。一つの書誌のオカレンスが物理単位となる。書誌と物理単位の比は1対nと定義できる。

④所蔵（図書館：物理単位）

一つの図書館は複数の物理単位を「所蔵」する。一つの書誌は複数の図書館に「所蔵」されるが、一つの物理単位は一つの図書館にしか「所蔵」されない。図書館と物理単位の対応比は1対nである。

⑤親子（著作：著作）

著作どうしが親子関係を形成することがある。例えば、古今和歌集という一つの著作の中には、個々の和歌という複数の著作が含まれる。この場合は、古今和歌集が親著作で、個々の和歌が子著作であると考えられる。対応比は1対nとなる。

⑥親子・変遷（書誌：書誌）

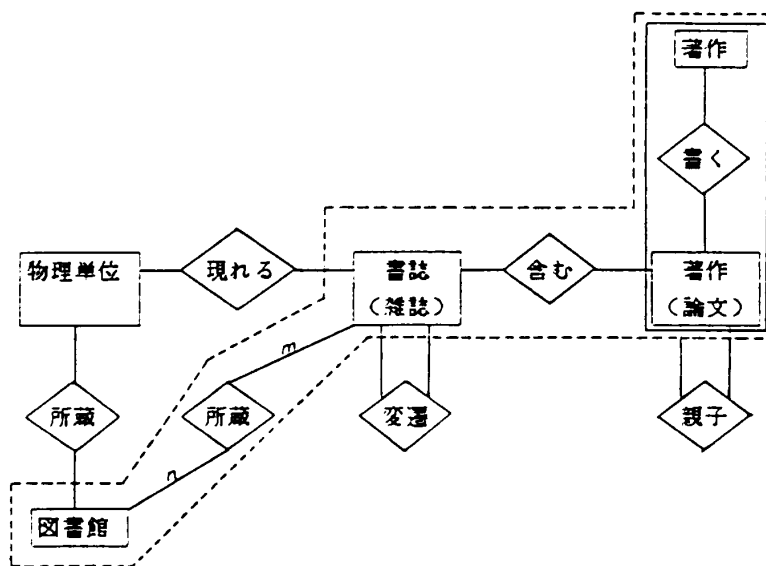
書誌どうしもまた親子関係を成す。例えば、シリーズを親書誌とすると、そのシリーズを構成する各巻は子書誌ということになり、対応比は1対nになる。また、ある本が同時に二つのシリーズに属しているというケースも稀にある。シリーズを親書誌とみなせば、一つの子書誌に二つの親書誌が結合していることになる。こうした場合対応比が逆になる。

雑誌の場合は、誌名変遷ということが起る。変遷も同一実体間に現れた関係としてとらえることができる。変遷関係には、「継続」「吸収」「派生」「分割」等のタイプがある。対応比はn対mである。

3. Model204によるモデル・システムの作成

E-R モデルに基づき、書誌所在 DB と二次情報 DB の結合のメカニズムを把握するために、Model204(DBMS)を使用して、雑誌および雑誌論文を対象とした実験システムの作成を行った。⁴⁾

図2 モデル



3. 1. E-R モデルによる概念設計

雑誌をめぐる世界を E-R モデルによって表現すると図2のようになる。

今回のモデル・システムでは、著者は著作の属性の一つと考え、著作の中に組み入れた。また、図書館は書誌ではなく物理単位を所蔵するのであるが、書誌と図書館の間にパスを設け両者を所蔵という関係でつなぐという形をとった。その結果、図中の破線で囲った範囲がモデル・システムで実現すべき部分となった。⁵⁾

3. 2. ファイル関連図

図2を実際のファイルに反映させると図3のようなファイル関連図が出来上がる。これら5つのファイルを Model204ファイルとして定義した。

今回のシステム作成の主眼は、あくまでも書誌所在情報 DB と二次情報 DB の連結のメカニズムをモデルを通して探っていくことにあるので、各ファイルのデータ項目は極力少くしてある。

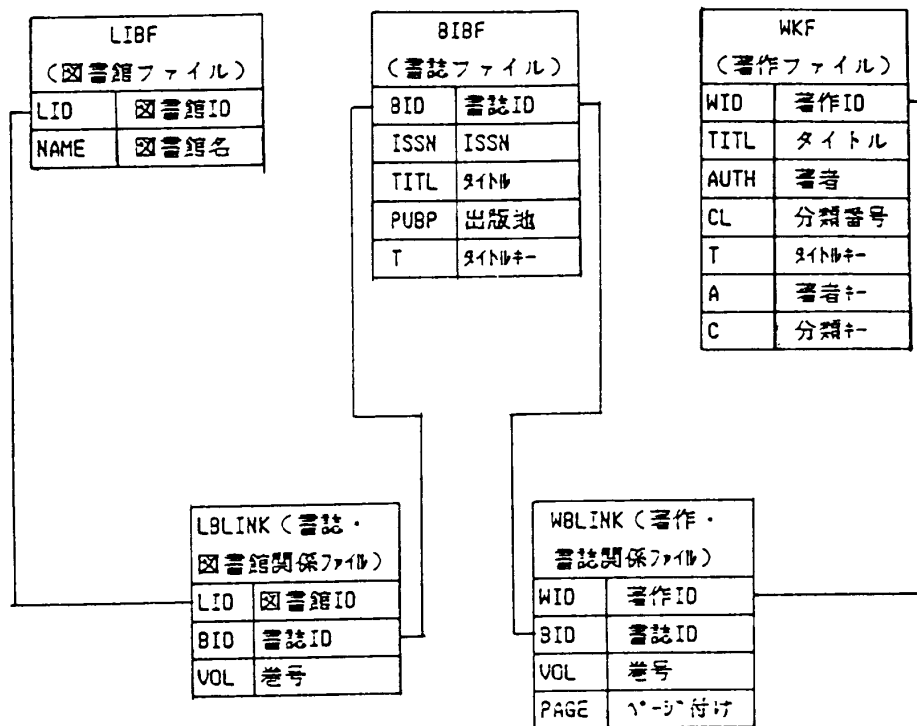


図3 ファイル関連図

3. 3. 入力

3. 3. 1. 入力データ

今回の実験では、「経済学文献季報」から適当にサンプリングしてテストデータとした。

「季報」は、全国の経済系図書館、資料室の協力のもとに経済資料協議会が編集している経済学関係の索引誌であり、111号から印刷編集がマニュアルから機械に切り替えられ DB 化された。近い将来、この DB をもとにしてオンライン検索サービスを提供することも想定されている。⁶⁾

3. 3. 2. Model204ファイル上のデータ

Model204のユーザ言語によるプログラムを通して入力されたデータが、各ファイル中に、どのように格納されているか、さらに各ファイルがどのようにリンクされているかを以下に示す。出力は PAI(PRINT ALL INFORMATION)形式である。

①WKF (著作ファイル)

```

WID = W0341
TITL = Soviet agricultural procurement pricing
AUTH = Michael L. Wyzan
CL = 0824
T = SOVIET
T = AGRICULTURAL
T = PROCUREMENT
T = PRICING
A = MICHAEL
A = L
A = WYZAN
C = 0824
C = 08
    
```

②WBLINK (著作・書誌関係ファイル)

この1レコードは、W0341の著作がOJ01100の書誌の9(1)巻の24-45ページに含まれている、という関係を表わしている。

```

WID = W0341
BID = OJ01100
VOL = 9(1)
PAGE = 24-45
    
```

③BIBF (書誌ファイル)

```

BID = OJ01100
ISSN = 01475967
TITL = Journal of comparative economics
PUBP = New York
T = JOURNAL
T = OF
T = COMPARATIVE
T = ECONOMICS
    
```

④LBLINK (図書館・書誌関係ファイル)

この1レコードは、N105の図書館がOJ01100の書誌の9(1)巻を所蔵している。という関係を表す。

```

LID = N105
BID = OJ01100
VOL = 9(1)
    
```

```

LID = N105
NAME = 名古屋大学経済学部
    
```

⑤LIBF (図書館ファイル)

3. 4. 検索

Model204のユーザ言語により、「著作検索」と「書誌検索」の二つの検索システムを作成した。著作検索では、まず著作（論文）をさがし、さらにファイル間のリンクをたどり、その著作がどの書誌（雑誌）のどの巻号に収録され、どの図書館によって所蔵されているかを知ることができる。

一方、書誌検索は、ある書誌の何巻何号にはどのような著作が収録されていて、どの図書館がその書誌を持っているか、という情報を得るためのシステムである。

図① 著作検索例

```

**著作検索開始**
ようこそ がんばってね!!

$$TITLE:                                タイトルにsovietとpricingのワードを同時に持つ探す
soviet pricing

ヒット件数 = 1

$$タイトル検索を続けますか? YES OR NO
NO                                         YESと答えるともう一度タイトルをきいてくる

$$AUTHOR:                                著者検索は必要ないので「空」を入力
_____

$$CL:                                     分類検索も必要ないので「空」を入力
_____

**著作詳細表示**                                WKFからの表示
WID: W0341
TITLE: Soviet agricultural procurement pricing
AUTHOR: Michael L. Wyzan
CL: 0824

In "Journal of comparative economics"        BIBFからの表示
9(1) pp. 24-45                               巻号とページはWBLINKからの表示

$$所蔵館情報を表示しますか? YES OR NO
YES

**所蔵館表示**
Journal of comparative economics
9(1)

<所蔵館>                                    このデータはリンクをたどってLIBFからもってくる
名古屋大学経済学部
    
```

図② 書誌検索例

****書誌検索開始****
 がんばってね!!

\$\$ISSN:
00941190

ISSNを検索キーとして入力

ヒット件数 = 1

****書誌詳細表示****

BID: OJ04300
 ISSN: 00941190
 TITLE: Journal of urban economics
 PUBP: New York

\$\$著作を表示しますか? YES OR NO
YES

\$\$巻号を指定して下さい
15(1)

15(1)に収録された著作(論文)を見たい

<コンテンツ>

TITLE: Metropolitan suburbanization and central city problems
 AUTHOR: Edwin S. Mills, Richard Price
 CL: 0530
 PAGE: 1-17

 TITLE: Changes in urban land values : 1836-1970
 AUTHOR: James S. Kau, C.F. Sirmans
 CL: 0530
 PAGE: 18-25

 TITLE: The impact of site characteristics on housing cost estimates
 AUTHOR: Steven Cobb
 CL: 0530 0870
 PAGE: 26-45

 TITLE: Measuring welfare effects of in-kind transfers

}

\$\$他の巻号の著作を表示しますか? YES OR NO
no

NOと答えると次に所蔵館を表示するかどうか
 の問い掛けがシステム側からなされる

****書誌詳細表示****

BID: OJ04300
 ISSN: 00941190
 TITLE: Journal of urban economics
 PUBP: New York

\$\$所蔵館情報を表示しますか? YES OR NO
Y

\$\$巻号を入力して下さい
15(1)

<所蔵館>
 名古屋大学経済学部

\$\$他の巻号の所蔵館情報を表示しますか? YES OR NO
NO

4. CBI システムにおける展開

文献情報センターは目録システムのサービスを既に開始している。幾つかの大学図書館とのネットワーク接続が完了し、総合目録 DB にデータが蓄積され始めた。学術情報システムの柱のひとつである「書誌所在情報システム」は、順調に軌道に乗りつつある。これと並行して、「二次情報 DB 検索システム」の方も着々と整備されつつある。

ところで、書所在情報システムは、今回の実験システム作成に際してモデルとした E-R 図（図 4 参照）の実線部分を実現したものである。一方、二次情報 DB システムは破線部分に相当すると考えられる。

従来はこの二つのシステムが分離されていたわけである。今後文献情報センター・システムが発展

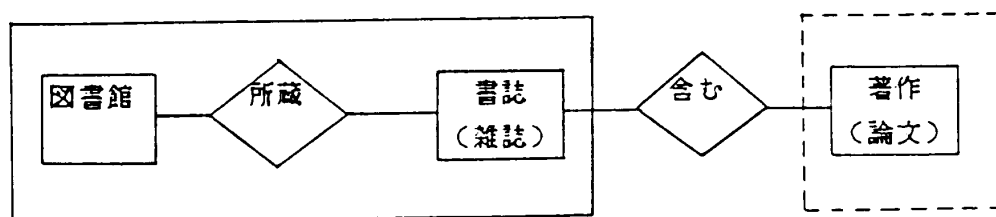


図 4

していくなかで、両システムの有機的リンクづけが実現され、両者が結合した形での検索が可能となっていくものと考えられる。こうした統合が完成すれば、二次情報 DB 検索から、原掲載誌の所在検索を経て、ILL システムによる所蔵館への文献複写依頼という一連のプロセスがオンラインで効率的に展開されることになる。こうしたシステムの一部もはやい実現が望まれるところである。

5. おわりに

今回のモデル・システムがそのまま実際に適用できるわけではないのは勿論のことである。しかしながら、実際に DBMS を使ってシステムを作ったことは貴重な経験となった。私としては、このレポートは終着点ではなくむしろ出発点であると考えたい。

註

- 1) E-R モデルについては以下の文献を参考にした。
 Chen, P. P. The entity-relationship model:toward a unified view of data. ACM transactions on database systems. Vol. 1, p. 9-36 (1976)
 Chen, P. P. The entity-relationship model:a basis for the enterprise view of data. AFIPS conference proceedings. Vol. 46, p. 77-84 (1977)
- 2) E-R モデルによる目録世界のとらえ方については、セミナーにおける宮澤先生の講義によるところが大きい。
- 3) 国文学研究資料館の古典籍総合目録 DB システムでは、「著作」という概念を「古典作品典拠ファイル」としてシステムに取り入れ、古典作品が書誌として現れる際の書名のユレに対処

している（石井啓豊、山城玲子: 古典籍を対象とした総合目録における書誌コントロール.
図書館雑誌。77(2), p. 95-97

- 4) Model204については以下のマニュアルを参考にした。

Model204ファイル管理者マニュアル

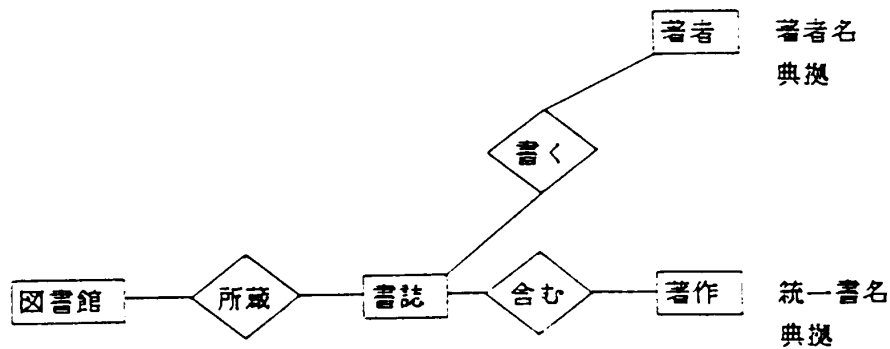
Model204コマンド参照マニュアル

Model204ユーザ言語マニュアル

- 5) 文献情報センター目録システムを E-R 図で表現すると次のようになる。

ここでは「著作」という概念は、統一書名典拠ファイルとして取り入れられている。今回のモデル・システムでは、「著作」を論文ファイルとすることにより、「書誌所在 DB」と「二次情報 DB」の統合を図ろうとするものである。

- 6) 川原和子: 「経済学文献季報」機械編集システム(KEIS プリント・システム) 大学図書館研



究. 21, p. 84-95 (1982)

研修レポート

中国語文献データベース化における漢字処理について

How to Handle the Chinese Characters (Kanji) in Using a Computer for Research in the Document in Chinese

東京大学東洋文化研究所附属東洋学文献センター 神田 百合枝*

要 旨

中国語文献を電算化する時の漢字の処理方法について調べた。

まず、歴史的観点から、米国の漢字用端末と日本の「JIS」漢字の制定について調べ、次に、日本の主要機関における対応などを調べた。

Abstract

It is a very important problem how to handle the Chinese characters(Kanji)when using a computer for the document in Chinese.

This report describes first from the historical view the Kanji terminal developed in the USA and the establishment of JIS C6226(Code of the Japanese Graphic Character Set for Information Interchange), and next the correspondence of the main facilities in Japan.

1. 漢字処理の歴史

1-1 米 国

1974年に設立された、26の研究図書館をメンバーとするグループである RLG(Research Libraries Group, Inc.)が開発した端末は、中華人民共和国、中華民国、日本、韓国の4ヶ国語を、1つの端末で処理できるものであり、これをCJK(Chinese-Japanese-Korean)端末と呼ぶ。

この端末での漢字の入力は、漢字を245種の構成要素に分解して、順次組合わせる形で行われる。たとえば、

春 = 土 + 大 + 日、 州 = 川 + 十 + 十

測 = 水 + 貝 + 刀、 李 = 木 + 子

といった形で入力され、1字を入力するのに平均3.6回、キーを押すことになる。

*Kanda, Julie : The Documentation Center for Asian Studies, Institute of Oriental Culture, University of Tokyo

図1 CJK端末のキーボード

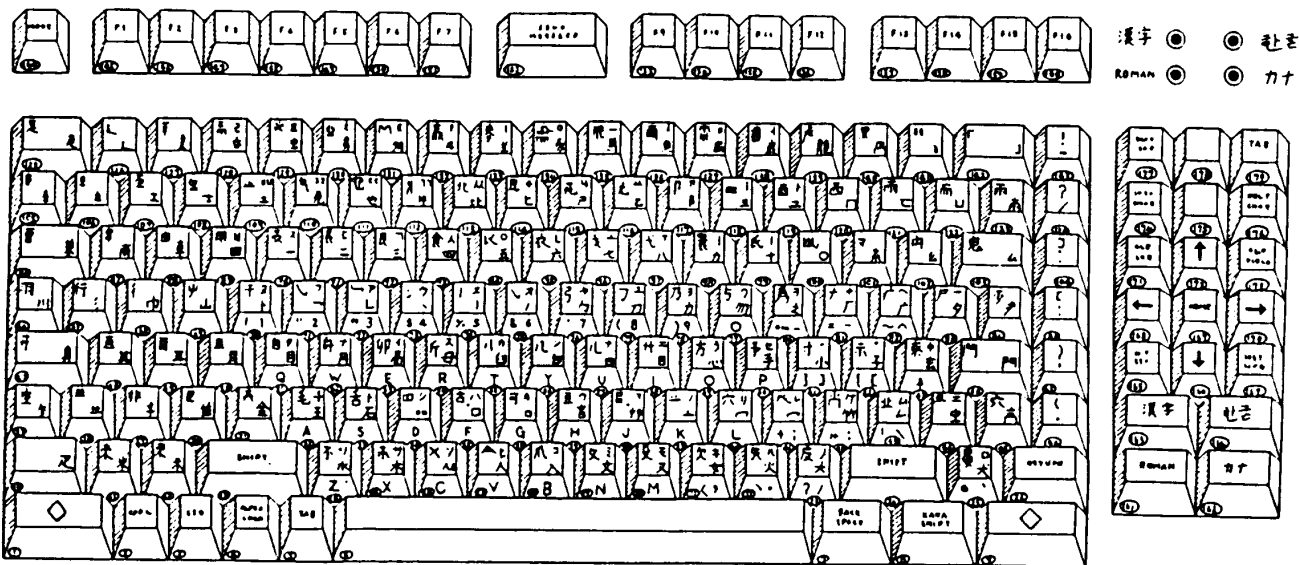


図2 中国語文献のレコード

```

PROD      Books      FUL/BIB  DCLP83-B533      Search      DCLP-IT
Cluster 1 of 1 - SAVE record
+
ID:DCLP83-B533      RTYP:c      ST:s      FRN:      NLR:      MS:  EL:  AD:10-20-83
CC:9114  BLT:am      DCF:a      CSC:      MOD:      SNR:      ATC:      UD:12-02-83
CP:cc      L:chi      INT:      GPC:      BIO:      FIC:0      CUN:b
PC:s      .PD:1982/      REP:      CPI:1      FSI:0      ILC:a      MEI:0      II:0
010      83179029/ACM
020      †cRMBY1.10
041 0      chi†beng
043      a-cc---
050 6      QC880.4.D44†bT3 1982 <Orien China>
066      †c$1
082 0      628.5/3/015515†219
245 00  Ta◊ch'i t'uan◊liu k'uo◊san chi wu◊jan ch'i◊hsiang lun◊wen◊chi /
      †cChung◊yang ch'i◊hsiang◊ch"u ch'i◊hsiang k'o◊hs"ueh yen◊chiu◊y"uan pien.
245 00  大气 湍流 扩散 及 污染 气象 论文集 /†c中央 气象局 气象 科学
      研究院 编.
250      Ti◊i◊pan.
250      第1版.
260 0      Pei-ching :†bCh'i◊hsiang ch'u◊pan◊she :†bHsin◊hua shu◊tien Pei-ching
      fa◊hsing◊so fa◊hsing,†c1982.
260 0      北京 :†b气象 出版社 :†b新華 书店 北京 发行所 发行,†c1982.
300      154 p. :†bill. ;†c27 cm.
    
```

1983年には、この端末から中国語文献の目録データが入力されている。価格が高いことと、キーの数が多くて操作が面倒であるという難点はあるが、普及しているという。¹⁾

端末のキーボードと中国語文献のレコードを図1と図2に示す。

1-2 日本

従来、計算機で扱われている漢字の字種や符号は標準化されておらず、不便であり、また非能率的であった。そこで、1978年、通産省工業技術院の委託を受け、財団法人日本情報処理開発協会漢字符号標準化調査研究委員会が JISC6226 (情報交換用漢字符号系) を制定した。²⁾

この規格に収められている図形文字は、漢字、特殊文字、数字、ローマ字、平仮名、片仮名、ギリシア文字及びロシア文字である。このうち、漢字は第1水準として2965字、第2水準として3384字、合計6349字が制定された。漢字の選定には、37種の漢字表が使用されたが、全体の選定基準が明確でなかったり、また「大漢和辞典」や「新字源」に収録されていない漢字が113字もあるなど、問題もある。³⁾

JISC6226は、1983年に改訂され、漢字は6353字にふえた。漢字表の一覧を、表1に示す。

表1 調査対象漢字表一覧

番号	漢 字 表	作成年月	新字源 漢字数	追加漢 字 数	記号 約物	漢字 数
1	IBM 2245 漢字印刷装置文字セット一覧表	47. 4	5883	909	48	6792
2	情報処理学会漢字コード委員会標準コード用漢字表試案	46.10	5778	308	0	6086
3	明治生命保険相互会社漢字コード表	46. 6	4912	443	58	5355
4	大蔵省主計局収容漢字表	48. 8 現在	4204	72	0	4276
5	富士通標準文字表	45. 3	4173	98	74	4271
6	大蔵省印刷局常用文字の調査	37. 1 ~41.12	3805	204	0	4009
7	全日本漢字配列協議会全国統一新配列による常用漢字目録	43. 7	3926	77	28	4003
8	国立国会図書館収容漢字表	48. 8 現在	3764	192	0	3956
9	林四郎、小林信子「語彙調査四種の使用度による漢字のグループ分け」	46. 9	3806	109	0	3915
10	森岡健二「漢字の層別」	49. 1 出版	3582	42	0	3624
11	国土行政区画総覧使用漢字	47年度	2971	280	0	3251
12	内閣調査室収容漢字表	48. 8 現在	3118	63	0	3181
13	日本生命収容人名漢字	48. 8 現在	2884	160	0	3044
14	読売新聞社1倍明朝(基本活字)文字数	47. 1	2899	26	17	2925
15	国立国語研究所資料集8「現代新聞の漢字調査」中間報告	41. 1 ~12	2828	48	0	2876
16	行政情報処理用基本漢字に対する符号付与に関する調査研究報告書「付表」	50. 3	2788	29	55	2817
17	日本タイプライター(株)和文タイプライター文字配列表 ^㉔	40年頃	2763	29	23	2792
18	国語シリーズ64 各種漢字表字種一覧	43. 3	2678	18	0	2696
19	日本タイプライター(株)N.T.C.漢字入力システム用文字コード表第一次分	50. 1	2642	29	38	2671
20	日本生命漢字コード表	42. 2	2620	8	36	2628
21	NHK 文研用語研究部「ラジオニュース語彙調査漢字集計表」参考 菅野謙	43. 3	2592	21	0	2613
22	(株)写研 SPICA「一寸ノ巾式」メインプレート	49. 4	2499	9	35	2508
23	サンケイ新聞社(大阪)サブトンスシステム銀孔機鍵盤文字配列及び符号表	48.10. 1	2304	15	43	2319
24	日本テレビ放送網(株)「文字表示装置用字母」	44.12	2286	16	17	2302
25	電話番号簿活字鋳造量	39.	1824	273	0	2097
26	菅野謙「当用漢字表に含まれない漢字」(放送で使う漢字の範囲の検討のために)	44. 4	1997	98	0	2095
27	NHK 文研用語研究部「国語教科書出現漢字表」	49. 6 .19	2012	10	20	2022
28	毎日新聞社さん孔機文字盤	49. 9 . 1 改	1990	2	30	1992
29	6社協定新聞社用コード表(CO-59)	34. 8	1982	3	0	1985
30	文字出度調査100万字集計結果	46. 6	1974	3	21	1977
31	日本経済新聞社KKB文字表及び符号表	48. 6 .25改	1947	5	24	1952
32	日本科学技術情報センター収容漢字表	48. 8 現在	1859	5	0	1864
33	当用漢字表	21.	1850	0	0	1850
34	電総研漢テレコード表1	40年度	1833	1	51	1834
35	電総研漢テレコード表2	44年度	1830	1	48	1831
36	雑誌記事索引自然科学編における漢字の頻度調査	42. 1	1403	8	0	1411
37	日本工業標準調査会JIS職業コード(JIS C 6266-1772)	47. 3 . 1	736	1	0	737

2. 主要機関における漢字処理

2-1 国文学研究資料館

国文学研究資料館（以下国文研と略す）は昭和47年に創設され、52年より電算化を開始した。国文研では、異体字⁴⁾の処理方法として、次の3つを考えた。

- ①まえて使用字種の範囲を決めておいて、範囲外の文字が現われた場合には、範囲内の文字に置き換える方法。
- ②使用字種の範囲は決めておかないで、新しい文字が現われた場合には、文字セットに追加して、データベース中に現われる文字はすべて、忠実に表現する方法。
- ③データベース上に、忠実に表現することを原則とするが、何らかの形で字種を管理する方法。

②の方法では、字体変化の問題があるので、ほとんど不可能である。

①の方法では、データの共有や相互利用の場合に、まえて決められた文字セットを、互いに使うことになるから、最も便宜的な方法である。しかし、実際のデータと、それを計算機上に表現したデータベースとの違いが大きくなり、データの同定に支障をきたすことも起こりうる。

そこで、国文研では③の方法が最もよいとして、独自の管理基準を設けた。それを「漢字シソーラス」と呼んでいる。

例えば、「大漢和辞典」（以下「大漢和」と略す）の中で、「国」という漢字の同義の漢字をたどっていくと、10種類の漢字が出てくる。それを図3に示す。この図の中で、同字と表示したものは「大漢和」の中で、「～に同じ」と記述されていたものである。これらの漢字群の中で、国〔04752〕の字体が当用漢字体であるので、この字体をディスクリプタと認定する。その他の漢字群は、それに対して非ディスクリプタと呼ぶ。この関係をまとめると表2のようになる。この表の中で、小括弧で示したものは「新字源」の漢字番号、中括弧〔 〕で示したものは、「大漢和」の漢字番号である。その前に書かれている16進数字は、漢字コードである。ディスクリプタと非ディスクリプタとの関係は、正字とそれに対する異体字との関係とも考えられる。

以上のように、「漢字シソーラス」とは、同じ概念を表す漢字をグルーピングし、そのグループの

図3 「国」をめぐる漢字群

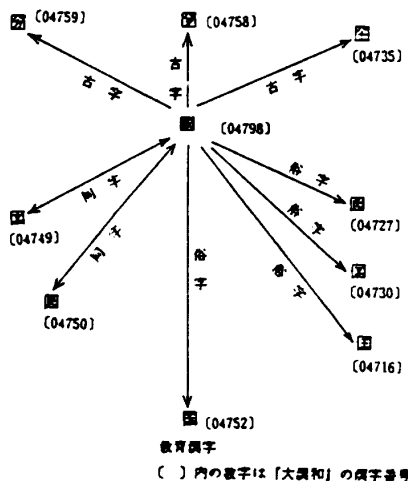


表2 漢字シソーラスの例

【国】 國 (コク、くに)

	B 9 F 1 (1 3 1 8) [0 4 7 5 2]
旧字	-(國) D 4 A 2 (1 3 1 9) [0 4 7 9 8]
古字	-(𠄎) 8 4 C A [0 4 7 3 5]
古字	-(𠄎) [0 4 7 5 8]
古字	-(𠄎) D 3 F B (1 3 2 0) [0 4 7 5 9]
同字	-(國) [0 4 7 4 9]
同字	-(𠄎) [0 4 7 5 0]
俗字	-(𠄎) (1 3 2 1) (0 4 7 1 6)
俗字	-(𠄎) (0 4 7 2 7)
俗字	-(𠄎) (0 4 7 3 0)

中で、ディスクリプタ（その漢字群を代表する字体の漢字）を決め、そして非ディスクリプタの中を、旧字、古字、俗字などのいくつかのカテゴリーに整理したものである。⁵⁾

2-2 国立国会図書館

国立国会図書館（以下国会図書館と略す）は昭和23年に創設され、46年に電算化を開始した。現在、最も流通している製品は、ジャパン・マーク（和凶書について、機械可読形式で蓄積した書誌情報データベース）である。

国会図書館では、字種の選定に当たり、当用漢字、人名用漢字、当用漢字補正資料による漢字は、すべて採用することにした。そしてその他の漢字が、国会図書館の目録、書誌類にどの位使われているかを調査するために、以下の資料をもとにした。：「国立国会図書館蔵書目録書名索引」、「外国人のための漢字辞典」、「人名漢字資料」、「漢字テレタイプ盤内文字表」、「当用漢字外地域名漢字一覧表」、「人名・地名漢字資料」。⁶⁾

現在、コード化されている漢字は、6610字である。新しく現われた漢字は、そのつど割り当てを行なっている。まず「大漢和」により検字番号を調べ、次に、すでに割り当てがすんだ漢字について検字番号順に配列されているカード・ファイルと照合して、重複しているか、していないかを確認する。重複していなければ割り当てを行なう。

国会図書館の扱う書誌は、古典籍ではなく明治以降の出版物なので、意味が同じなら、当用漢字に変えてしまう、2次資料に使う文字は、1次資料に表示された文字を厳密に表記する必要はないと考えている。⁷⁾

2-3 日本科学技術情報センター(Japan Information Center of Science and Technology, JICST)

日本科学技術情報センターは昭和32年に創設され、36年に電算化を開始した。51年から JOIS (JICST Online Information System)により、文献情報と研究課題情報のオンライン検索サービスを行なっている。

昭和54年からサービスを開始した JOIS-K の扱っている漢字は、JIS 第1水準2965字のうち、1965字である。⁸⁾

2-4 東京大学文献情報センター

東京大学文献情報センターは昭和58年に創設され、59年に電算化を開始した。前身である東京大学情報図書館学研究センターは、57年に「学術雑誌総合目録データベースオンライン検索システム TOOL-ULP」を東京大学大型計算機センターより公開している。

JIS にない漢字は、まだ登録されていない。

2-5 京都大学人文科学研究所附属東洋学文献センター

京都大学人文科学研究所附属東洋学文献センターでは、京都大学大型計算機センターと協力して、中国学に関するデータベースをこれまでに3つ開発し、現在うち2つを同大型計算機センターより公開している。データベース開発に際して、これら京大グループ（以下京大と略す）の行なった方法は、概ね以下の通りである。

(1)漢字のボタン

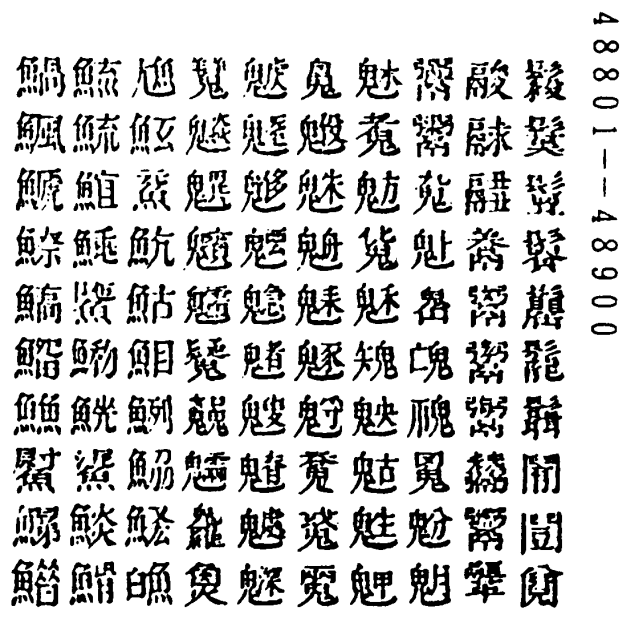
現在使用できる各メーカーの日本語処理システムでは、計算機に収容できる文字数には制限がある。JIS 第1水準と第2水準の漢字6353字と、メーカーが用意した JIS 外字のほかに、利用者が自由に文字を収容できる領域が残されているが、これだけでは多種の漢字を扱う、いわゆる中国古典を処理するのは難しい。

そこで、京大では「康熙字典」(18世紀の初めに清朝の康熙帝の命令により作られた字書で、実用的で、しかも精密であるということから、中国研究者の必携の書とされている) から文字ボタンを作り、そのボタンに一連番号を割りつけて康熙文字コードを作成した。

京大が、文字ボタンを自前で作ったのは、業者に注文して作らせると、莫大な費用がかかるからであった。しかし、当初日本の出版社が発行している「康熙字典」のリプリント本から文字ボタンの原図(字型)を作ろうとしたが、著作権などの関係で使えず、やむを得ず、台湾から発行されている文字コントラストの悪いリプリント本を使った。まず、「康熙字典」から1文字ずつを切りとり、100字1組にして10×10の罫目の台紙に貼りつけ、撮影してフィルムにした後、ドラム式画像入力装置で読み取り、磁気テープに格納する方式で行なった。文字ボタンの一部を図4に示す。

このようにして計算機に格納した字種は、全部で49188字あり、本来は利用者が文字をコードで収容する領域に、画像として出力できるようにした。中国学では、漢字の索引を作る時には「康熙字典」の画数順に配列するのが最も一般的であるが、この配列法はメーカーの日本語処理システムではできない。この文字ボタン(文字コード)を使えば、この問題は解決できるわけである。⁹⁾

図4 「康熙字典」文字ボタンの例



(2)漢字の入力方法

通常の日本語処理システムにおける漢字の入力方法を表3に示す。

表3 漢字入力方法

漢字テレタイプ方式	和文タイプ方式	タブレット方式	連想コード方式	かな漢字変換方式 (ローマ字漢字変換方式)
漢字キートップ上に複数の漢字を配列し、該当漢字をソフトで位置を示して入力する方式であり、左・右・両手で操作して行く方式。熟練すれば比較的早く入力できる。電算写植の原稿入力として採用されるケースが多い。	和文タイプライターに、文字コードの発生機構を付けて和文タイプライターを操作するのと同様の操作をして行く方式。速度は早くないが、印字ゲラリストがとれるので打鍵と同時に目検できることと、和文タイピストであれば、すぐ入力オペレーターになりうる。主力媒体が紙テープ主体であるため利用は減少している。	平面状盤面に、キートップは1漢字ずつ一定の順(音・訓・画数順等)で配列されており、オペレーターは指のかわりにペンで該当漢字をタッチして行く方式。比較的素人向きで、装置も軽量・小型になるが、ペンの上下・左右の移動に時間を要し、熟練しても入力速度はあまり上がらない。ワープロやパソコンの入力方式に採用されているが、漢字の読み方に難があり減少しつつある。	漢字を関連するカナ2桁に置きかえて入力していく方式(例、芦→トア、雄→オス等) 熟練すると、入力速度は非常に早くなる。又、既存の入力機器であればほとんどの機械がそのまま使用でき、特別な入力装置を必要としないが、漢字をカナに置き換えるので、オペレーターが連想コードを記憶するのに時間を要する。記憶できる漢字の数に限界があり、字数の多い漢字入力には不適當。	漢字の音読み、訓読みで入力し、読みに対応する漢字をディスプレイ表示し、該当漢字を選択する方法と、文節単位に文法解析して変換していく方法がある。パソコンや日本語ワードプロセッサに採用されている。同音異字、読めない、といった問題があるため、字数の少ない文書入力処理等に適用しており、入力スピードは遅いが素人にも比較的容易に入力できる。漢字の読みをローマ字で入力して行く方式もある。
プロ向け	プロ向け	素人向け	プロ向け	素人向け

ところで、中国古典に使われている多種の漢字を入力するには、視覚的に漢字をとらえるのが最善の方法ではないだろうか。

そこで、字形分解による漢字検字法として次の2つを検討してみる。

①四角号碼法

四角号碼法は、1920年代の後半に中国の王雲五氏により考案された漢字検字法で、中国の辞書や索引類に広く利用されていて、研究者の間では普及しているものである。

図5にその筆形図を示す。四角号碼法は漢字の四隅の筆形をみて、夫々を0から9までの番号(号碼)におきかえ、漢字一字を4桁の数字で表わして、0000から9999までの順に配列する方法である。同一番号の字を更に区別するには、右下隅の筆画の上方に最も近く、かつ見分けやすい一筆画をとり、「付角」として番号(付号)を与える。付号は4桁の数字のあとに小字でつける。付号も含めた10進5桁のコードは、 $10^5 - 1$ 個の文字を表現できる。

図5 四角号碼法の筆形図

号碼	筆形
0	┐
1	┌
2	└┐
3	、
4	十
5	丰
6	口
7	凵
8	八
9	小

図6 三角編號法の筆形と符号表

小	9-	小	业	少	十	米	半	尸	已	火	九
八	8-	八	金	竹	人	食	邑	舍	缶	死	爪
凵	7-	凵	凵	フ	コ	厶	凵	凵	凵	凵	凵
口	6-	口	日	目	口	田	里	邑	易	只	足
丰	5-	丰	丰	甲	戈	才	井	中	邑	夫	未
十	4-	十	土	ナ	又	彡	革	女	力	走	木
、	3-	ノ	、	彡	彡	...	馬	非	一	之	彡
┌・ノ	2-	┌	止	彡	彡	隹	牛	月	厂	欠	禾
一	1-	一	工	丁	夕	王	耳	石	乙	酉	示
┐	0-	┐	广	疒	文	亦	言	方	立	衣	
基本符	形番	0	1	2	3	4	5	6	7	8	9

四角号碼法は、10種の単純な筆形を基礎としていて、分かり易く、記憶し易いが、複数の漢字が同一の漢字検字番号を持ったり、解釈の違いにより番号が異なったりなど、漢字分解能力に限界がある。

②三角編号法

三角編号法は、台湾の胡立人、張源涓、黄克東氏により考案された漢字検字法である。図6に示すように、100種の基本的な要素が定められていて、それらに符号00～99が与えられている。100種の基本的要素は10のサブグループに分類できるが、その各グループは四角号碼法の10の筆形で、図5に対応する。3個の要素が漢字一字の四隅から取られ、夫々の符号を並べて合計10進6桁のコードは、 $10^6 - 1$ 個の文字を表現できる。これは、四角号碼法より1桁大きい。また、四角号碼法とくらべ、取り出すものを単純な筆形としないで、複数の要素としている。¹⁰⁾

京大では、「康熙字典」の文字パターンについて、2人のオペレーターによるブラインド入力を行ない、三角編号法による入力が極めて安定しているとして、独自の三角編号法コードを確立した。

3. 京都大学人文科学研究所附属東洋学文献センターの開発したデータベース

京都大学人文科学研究所附属東洋学文献センターと京都大学大型計算機センターが協力して開発したデータベースについて略述する。

3-1 CHINA1

昭和56年に、京大大型計算機センターより公開されたCHINA1には、中国明代における科挙合格者に関する情報及び明代行政地名が格納されている。そこで、例えば漢籍目録の整理をしている時に、著者や刊行者が明代の科挙合格者であることはわかっているが、字（あざな）や号しか明きらかでないような場合に、その人の本名、出身地及び生年を知ることができる。

京大大型計算機センター内にある日本語処理ターミナルを利用すると、姓名、字、号、登科録名、登科年、生年、出身地が漢字で検索できる。また、姓名、字、号は現代中国音（拼音）でも検索できる。使用されている漢字は、JIS第1水準と第2水準の漢字である。¹¹⁾

次に、大学間コンピュータネットワークを使用して、東大大型計算機センターからCHINA1を検索した例を図7に示す。ここではキーワードとして姓名を使っている。

3-2 CHINA2

昭和58年に、同じく京大大型計算機センターより公開されたCHINA2には、唐代の詩人、李商隱の作品である「樊南文集」の全文テキストが格納されている。

使用されている漢字は、JIS第1水準、第2水準の漢字のほか、メーカー提供の外字、康熙字典からの文字、新しく合成して作った文字など、全部で308060字である。¹²⁾

3-3 CHINA3

京大人文科学研究所附属東洋学文献センターでは、毎年「東洋学文献類目」という東洋学関係の文献目録書を刊行してきたが、1981年度版より、これを電算化した。これがCHINA3である。

文献の主題の分類、地域、時代などから検索ができる。また、検索者が原著に表記された字体を知らなくて検索もれが生じるのを防ぐために、漢字の字体を代表字に統一したフィールドを設けている。代表字は、JIS 第1水準から選ぶようにしている。¹³⁾

図7 姓名による検索

```
RS> SEA LA 'WANG JI ZU'
1 FOUND
RS> OUT EL(* PY LA)

CHINA1

#1      BANGO      00002100
        SEIMEI     : : : : :
        AZANA      : : : :
        SYUTEN     : : : : :
        SEINEN     : : : : :
        NPLACE     : : : : :
        GYOSEI2    08NAA000
        PY         2WANG4JI3ZU
        LA        4KE4SHAO
                WANG JI ZU
                KE SHAO
```

4. 今後の中国語文献データベース化について

4-1 漢字のパタン

中国語文献、特に中国古典をデータベース化する時の漢字のパタンについては、現時点ではやはり京大が開発したように「康熙字典」の文字パタンを作るのがよいと思われる。それは「康熙字典」が中国古典研究の基礎となる字書であり、また約5万字もの漢字を収めているからである。

4-2 漢字の入力

漢字の入力には、次のような条件が要求される。

- 正確に早く入力できること
- 複数かつ難解な漢字でも同一条件で入力できること
- 漢字の読み方がわからなくても入力ができること
- 記憶に頼らず、論理性のある入力ができること
- オペレーターの拒否反応のない方法であること

これらの諸条件を満たす入力方法として、漢字の字形分析による三角編号法が最も適していると思われる。

ところで、この三角編号法を使ってシステムを開発した機関として次の3つがある。

まず、三角編号法の考案者である黄克東氏らは、約3万字の漢字をコード化した。

また、日本で三角編号法による漢字入力方式を開発した伊藤忠データシステム株式会社の担当者の話によると、同社のシステムでは約2万字の漢字を扱うことができ、オペレーターの養成には約3ヶ月を必要とするという。

一方、京大では約5万字の漢字をコード化したのが、オペレーターの養成には約6ヶ月を要したという。

このうち、中国語文献、特に大量の漢字を扱う中国古典のデータ入力には、京大の作成した「康熙字典」文字三角編号法コードによるのが、コード化された字数も多く、よい方法だと思われる。

なお、三角編号法は、データの入力には適しているが、検索のための入力には四角号碼法が適している。それは、四角号碼法はすでによく知られている検索字法であり、また、新しく覚えるにしても規則が簡単だからである。

引用、参照文献と注

- 1) 池田秀人“アメリカ合衆国における図書館自動化システム”紀伊国屋 p.60、75～82
- 2) “JIS C6226-1978 情報交換用漢字符号系解説”日本規格協会
- 3) 田嶋一夫“JIS 漢字表の利用上の問題”情報管理 21 (10) 1979. 1 p.756
- 4) 異体字とは意味が同じで、正体とは字体が異なる文字をいう。“新字源”には本字、古字、別体字、俗字、誤字などの説明が記載されている。
- 5) 田嶋一夫“日本語情報処理における文字セットコントロールシステム”情報管理 26 (7) 1983. 10 p.556、560～561
- 6) 安田健“漢字の処理”図書館研究シリーズ 21 1980. 3 p.36～37、97
- 7) 鶴田真也“日本全国書誌における漢字字種調査”第18回情報科学技術研究集会発表論文集 p.179
- 8) 長谷川昇、細山美樹“JICST 漢字オンライン情報システム JOIS-K”情報管理 21 (10) 1979. 1 p.778
- 9) 勝村哲也“中国研究における情報処理の試み”アジア・クォーターリー 13 (2・3) 1981 p.15～18
- 10) 島崎眞昭、河野典、桶谷猪久夫、村尾義和“漢字の字形分析と漢字入力システムの応用”情報処理学会第26回(昭和58年前期)全国大会 p.1153
- 11) 勝村哲也、星野聰“データベース CHINA1について”京都大学大型計算機センター広報 14 (3) 1981 p.128～132
- 12) 星野聰、三重まゆみ、勝村哲也“データベース CHINA2について”同上 16 (3) 1983 p.122～127
- 13) 星野聰、勝村哲也“東洋学文献類目データベースの研究と開発”情報処理学会論文誌 25 (2) 1984. 3 p.187～188

図・表の出典

図1：“アメリカ合衆国における図書館自動化システム”

図2： 同上

- 図 3： “日本語情報処理における文字セットコントロールシステム”
- 図 4： 星野聰、勝村哲也編“康熙字典文字集覧” 1981
- 図 5： 島崎真昭、桶谷猪久夫“三角編号法による漢字の字形分析と漢字の入力”漢字の計算機による
処理に関する総合的研究 S59. 3
- 図 6： 同上
- 表 1： “JIS C6226-1978情報交換用漢字符号系解説”
- 表 2： 図 3 に同じ
- 表 3： “日本語情報処理の手引き”伊藤忠データシステム株式会社

研修レポート

ローカルの雑誌管理システムと学術雑誌総合目録所蔵データ — 受入れから製本・所蔵データへの流れと学総目所蔵データ形 式への交換

Junctioning of KULPIS to the ULSP—An Effort of Converting Local
Acquisition/Holding Data to the ULSP Data Format

関西大学図書館 辻川 輝男*

要 旨

関西大学図書館において稼働中の雑誌管理システム(KULPIS)は受入れシステム、製本管理システム、そして所蔵管理システムで構成されている。この受入れから製本・所蔵に至るデータの流れの処理方法は、電算化を成功させる重要なキー・ポイントである。ここでは3つの間のデータの流線を整理し、機械的処理をする場合の問題点を探る。また当館システムの所蔵データ(開始・終了方式)から学総目所蔵データ(あるもの表示)への機械的処理を試みた。

Abstract

KULPIS has been developed and operated in the full range of periodical processing for more than 7 years since 1978. The main structure is composed of 3 units. Namely, acquisition system, record holding system and binding(shelf)management system. The full range usage of coherent data in the related systems must be a key to make success on the library automation. Aiming here the reconsideration of data flow in the KULPIS and junctioning to the ULSP data automatically.

1. ローカル・システムでのデータ処理

雑誌管理システムに特徴的なこととして、受入れデータから製本シェルフデータ、そして所蔵データというようにデータが時間的、あるいは階層的に、しかも内容的には濃縮される形で編集され流れていくことが挙げられる。そして、この3つのレベルのデータの流線とそれに付帯する各種のデータ処理関係をどう整理し、システム化するかが電算化の最も重要なポイントであり、各大学

*Tsujiikawa, Teruo : Kansai University Library

図書館あるいは機関が担う使命や特徴、雑誌業務担当者あるいは開発担当者の考え方など諸々の個別的条件の違いによって異なってくる。これまで多くの大学図書館あるいは雑誌資料収集機関で雑誌管理システムが開発され稼働している。ここでは当館のシステムを中心にデータの流れを紹介する。

1-1. 関西大学図書館のシステム(KULPIS)

KULPIS(Kansai University Library Periodicals Information Systems)は昭和53年3月よりオン

表1 雑誌所蔵タイトル件数

年 度	53	54	55	56	57	58	59	60
和 雑 誌	6,391	6,761	7,224	7,821	8,651	8,886	9,384	9,975
洋 雑 誌	3,106	3,838	4,099	4,512	4,754	5,017	5,200	5,401
合 計	9,497	10,600	11,323	12,333	13,405	13,903	14,584	15,376

表2 継続受入れ雑誌タイトル件数

年 度	53	54	55	56	57	58	59	60
和 雑 誌	2,601	3,036	3,256	3,722	3,998	4,055	4,439	4,898
洋 雑 誌	1,446	1,794	2,036	2,278	2,412	2,652	2,818	2,965
合 計	4,047	4,830	5,292	6,000	6,410	6,707	7,257	7,863

ラインによる到着雑誌の受入れ業務を稼働し、昭和56年4月までに受入れから製本シェルフ管理・整理業務、発注・支払い・予算管理など雑誌に関わる各種業務を電算化し、初期開発を終えた。マシンは現在、学内事務共用機である FACOM の M-160F を使用している。稼働当初から現在に至るまで、このシステムが処理し管理するデータ件数の変遷は次の通りである。

受入れデータはオンライン・システムの端末画面から巻・号・通号・発行年月日・価格などを入力する。製本時期が来るとこの受入れデータを製本コード・製本単位コードなどを基にして製本データがプログラムにより抽出・編集され、更に製本作業リストを出力する。このリストにより雑誌現物を製本単位毎にまとめる。この時に紛失などによる製本保留や製本単位の例外扱いなどの事例が発生すれば、製本データをメンテナンスする。そうして整えられたデータから製本発注書・同納品書・同検収書を作成し、雑誌現物と共に製本業者に引き渡す。製本された雑誌は検収を受けた後、現物は書庫に配架され、会計上は編入手続きを経て固定資産勘定に計上され、製本データは再編集されてシェルフマスタに累加される。そして年度末には図書台帳が出力される。製本を要しない「1冊もの」などは、完冊編入と称して同様に自動抽出され、請求記号ラベルもプリンタ出力され、編入手続きが行われ

表 3 編入雑誌件数

年 度	54	55	56	57	58	59
和 雑 誌	1,328	2,457	3,147	3,099	2,566	4,361
洋 雑 誌	1,508	2,405	2,676	2,099	3,358	4,136
合 計	2,836	4,862	5,823	5,198	5,924	8,497

図 1 所蔵マスタのデータ項目（簡略図）

呼 出 符 号	雑 誌 名	請 求 記 号	所 蔵 デ ー タ				各 種 コ ー ド	欠 号 デ ー タ	そ の 他
			開 始		終 了				
			巻(号)	年	巻(号)	年			
X(10)	X(135)	X(22)	X(4)	X(4)	X(4)	X(4)	X(◆)	X(68)	X(◆)

る。これも抽出されたデータがシェルフマスタまで送られる。また全く消耗品扱いの雑誌も同様に抽出され払い出し手続きが行われる。処理が進み実務上不要となった受入れデータは一定機関が過ぎると受入れマスタから削除される。なお、このシステムで処理された編入雑誌件数は表3の通りです。

ところで所蔵データは図1のような項目を持っている。これらの項目は当該雑誌の受入れ決定が行われた時に作成される。そして、雑誌到着時あるいは製本作業時に現物と異なるデータがあればメンテナンスするだけとなっている。ここで問題となるのは、

- ①受入れデータあるいは製本シェルフデータには巻・号・通号・出版年月日が揃っているのに対して、所蔵データとしては開始終了とともに巻(号)・年しかない。
- ②継続受入れ中の雑誌は、終了巻年をオープンと称して空白にしている。このようなフォーマットになった理由としては、
 - ①システム構築時、書誌所蔵データを従来のカード目録からパンチ外注して起こしたため、開始・終了・次号という表示形式が採用しやすかった。
 - ②ファイル容量の節約が必要があった。
 - ③所蔵終了データが空白であれば継続受入れ中としておけば毎年メンテナンスする必要がない。
 - ④オンラインでデータ・メンテナンスをするという考え方があったため、製本シェルフデータからの所蔵データ自動更新は余り考慮されなかった。
 - ⑤製本シェルフデータはシステム化する時点で、カード目録に約15万件あり、このデータを起こすことはその費用・労力・ファイル容量・利用価値の点から当時としては不可能であった。

ところが、昭和60年4月稼働の貸出閲覧システムが在庫管理方式を採用したため、所蔵する全ての

図書及び雑誌のシェルフデータをデータ化することになった。雑誌については昭和59年9月から約1年半の作業を行い、現在20万件を越す雑誌シェルフデータがマスタに管理されている。ここに至って漸く製本シェルフデータから所蔵データの自動更新の可能性が出てきたのであり、現在所蔵データのフォーマットを変更すべく検討がなされている。

1-2. 他大学図書館のシステム……省略

1-3. ローカル・システムの現状

セミナー受講期間に資料の入手が出来たいくつかの雑誌システムについて、そのデータ処理の流れをみたが、受入れデータに関わるサブシステムだけであるとか、書誌および所蔵データだけのシステム、製本の部分を残している場合などバラエティに富んでいる。しかし雑誌データの流れをトータルにシステム化し、さらに学総目のような全国的規模のシステム（データベース）との関連性を持たせたシステムは未だ実現していないのが現状のようである。その大きな理由としては、

- ①雑誌管理システムに内在する問題として、3つのレベルで発生する雑誌データの自動編集処理を解決して得ていない。
- ②外的要因として、学総目データベース自体がここ数年前から漸く明確になってきた段階であり、これとの有機的な繋がりをもつシステムの開発は、まさに今日的課題となりつつある。

ということが挙げられる。

2. 受入れデータから製本シェルフデータ、そして所蔵データへ

この章では、当館の雑誌管理システムである KULPIS でのデータ処理の現状をやや詳しく紹介し、受入れから所蔵データに至る流れの自動処理に焦点を合わせながら、問題点あるいは改善点を明かにした。

2-1. 受入れデータから製本シェルフデータへの編集

この処理のポイントは

- ①いかにして受入れデータを製本の合冊単位毎に抽出し編集するか。
- ②製本処理の中でのデータ・メンテナンスの量をいかにして少なくするか。

という2点である。

当館システムの製本処理は本製本処理を基本として、その他に完冊編入処理・仮製本処理・払い出し処理という4つで構成されている。いずれも受入れデータ及び各種コードに基づいてプログラムによる自動抽出を行っている。ちなみに本製本処理における抽出条件は

- ①製本コードが「1」:本製本である雑誌で製本単位（これもコード化されている）を充足する受入れ雑誌が到着している。欠号などがある場合は抽出しない。但し、別にリスト出力し次回の製本抽出までに間に合うように欠号補充の処置をする。
- ②オペレータが指定する製本対象年度の出版物で、支払い済みのもの。
- ③前回製本抽出されたが、紛失などで製本保留したもので、今回抽出まで補充され製本単位を充足

したもの。

等々である。合冊単位毎に1件目の受入れデータの巻・号・通号・発行年月日と最後のそれらをそれぞれ製本単位の開始終了データに編集し、いわゆる製本シェルフデータの元データが出来上がる。このファイルから製本作業リストが出力される。

2-2. 製本データのメンテナンス

製本作業リストをもとに担当者が雑誌現物を製本単位毎にまとめると必ず修正を要するデータが発生する。すなわち、

- ①雑誌の厚さが前回と異なる場合……抽出データを合併、もしくは分割製本単位コードの変更
- ②製本データとして不備な項目があれば補充・修正する。
- ③紛失などの事由による製本保留の場合は製本データを削除し、次回製本時に再抽出されるように受入れデータにコード設定する。
- ④本誌と別扱いすべき累積索引版などが本誌とともに受入れられていた場合は、累積索引版としての書誌データを作成する。

等々である。メンテナンス後はプログラム処理で一連の登録番号が振られ、誌名や請求記号などと共に製本発注書が出力される。製本業者においてそれらのデータが製本雑誌のカバーに表示されるので、図書館側では検収書に基づいて検収するのみとなっている。検収終了後、製本シェルフデータはシェルフマスタに累加され、編入手続きや統計リストの出力が行われる。更に製本雑誌は書庫に配架されて貸出閲覧の対象となるので、貸出閲覧システムのマスタにも累加される。昭和60年夏の製本作業は、日常の受入れ業務を行いながら通算16日間、製本単位数は和雑誌が2,660冊、洋雑誌が2,658冊、合計5,318冊、データ・メンテナンスは約3日が通算4日をかけて行われた。

2-3. 製本シェルフデータと所蔵データ

(1) シェルフデータの順序逆転の問題

欠号や紛失などによって製本保留とするとシェルフマスタでのレコードの順序が逆転する場合が生じる。これを回避するため累積分と新規分をソートマージする方法が考えられる。但し、前提として巻・号などが前詰めなら後詰めにしたたり、受入れデータでの特殊巻号表示があれば数値化あるいはコード化してソートのためのキーを作成するというテクニックが必要となる。その上で、出版年・巻・号・通号の順にソーティングすれば一応の順序付けが出来る。巻号表示の変更がある場合でも出版年によるソートが効いているので、極く特異な場合を除いて順序の逆転を防ぐことが可能である。当館システムでは、この考え方によるソート・キー作成プログラムを運用している。

(2) 所蔵データ自動更新での問題

A. 巻号表示の変更があった場合

所蔵・製本いずれのデータでも通常は開始と終了の一对の項目しか持っていない。多くは誌名変遷に伴って発生するが、同一誌名の下でも起こり得る。製本単位の切れ目の場合と合冊中に埋もれる場合があり、後者では製本シェルフデータにおいても問題が生じる。これを回避する方法としては、

①変更があったものを合冊中に含む製本自体を行わない。

②開始終了の項目をリピータブルで持つ。

ということしか方法がないと考えられる。

B. 明示されない欠号データの抜け落ち

開始終了方式での欠号データは、開始終了の範囲について欠号表示を行う。したがって、ある雑誌の欠号データはその製本シェルフデータすべての欠号データを古い順に集めて編集すれば出来上がるはずである。しかし、途中で製本保留があった場合はレコード自体がないため、その欠号状況をプログラム処理で明示することは非常に困難である。これについては製本シェルフデータのメンテナンス時に既累加分とチェックを注意深くして欠号データとしてのレコードを1件おこすという方法しかないのではないかと考える。

3. 学総目所蔵データ形式への変換

ローカル・システムの所蔵（受入れ）データから、学総目の「あるもの表示」の所蔵データにプログラム編集し、磁気テープで所蔵データを提出した事例はいくつか報告されている。ここではそれらを踏まえながら3つのレベルそれぞれからの変換を検討した。その結果として当館の現行システムでは所蔵データ（開始終了方式）からの形式変換が最も適していると考えられる。そのための変換プログラムを試験的に作成したので、その基本的ロジックを紹介する。

3-1. 所蔵データ（開始終了方式）からの変換

開始終了のデータと欠号データとの関係をプログラムで解析しながら変換することになるので、若干のテクニックが必要である。

(1) 当館システムの所蔵データ

- 例 A誌 25 (1965) ……35 (1975) 欠号:27, 31 (4-5)
- B誌 1 (1970) …… ……継続受入れ中
- C誌 3 (1973) …… 3 (1974) …… 3巻のみ所蔵

(2) 変換処理の結果としては

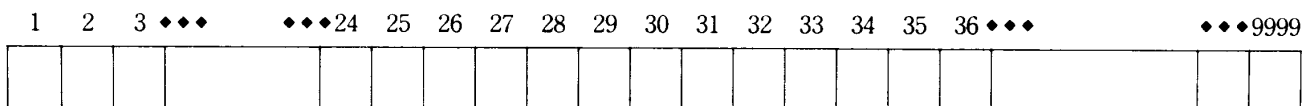
- 例 A誌 ① 25-26, 28-30, 31 (1-3, 6), 32-35 ……号レベルを表示
- ② 25-26, 28-30, 31 (), 32-35 ……号レベルを省略

今回は②への変換を行っている。

(3) 巻次データ編集ロジック

A. 巻号ワークテーブルとして、1バイトの繰り返し回数が9999個のテーブルを用意し、1バイトが各巻あるいは各号に対応していると考える。

図 2



B.開始終了の範囲内の該当テーブルに所蔵のマーキングをする。

図 3

1	2	3	◆◆◆	◆◆◆24	25	26	27	28	29	30	31	32	33	34	35	36	◆◆◆	◆◆◆9999
					H	H	H	H	H	H	H	H	H	H	H			

C.欠号データを解析し、丸カッコ付きの欠号（巻）データの場合は不完全巻、数字のみの場合は完全欠号（巻）とみなし該当テーブルにマーキングする。

図 4

1	2	3	◆◆◆	◆◆◆24	25	26	27	28	29	30	31	32	33	34	35	36	◆◆◆	◆◆◆9999
					H	H	L	H	H	H	?	H	H	H	H			

D.巻号ワークテーブルを最初から順に読み取り、学総目所蔵データの形式に編集していく。

「25巻から26巻まで所蔵し27巻は欠、28巻から30巻まで所蔵し31巻は不完全所蔵、そして32巻から35巻まで所蔵している。」

4. まとめ

学総目データベースは和文編が全国調査依頼約1年半の日数を要して漸く完成している。欧文編は目下、全国調査の真最中で完成は早くて2年先の予定である。多額の費用と大量の労力、そして可能な限りの技術と知恵とを投入して精力的にデータベース構築の作業が進められている。そうして出来上がったデータベースはあらゆる分野の学術研究・技術開発に寄与しその効果を存分に発揮するであろう。しかし、更に厳しい目でみるとそのデータベースの内容は完成時点で既に数年前のものとなっている。このタイム・ラグを埋めることは、現時点ではそれほど問題視されていないであろうし、またすべき時期でもない。だが、基本的にデータベースの内容は常にメンテナンスする必要があるし、それと共に再度所蔵データの調査が必要となる。この2つの問題、即ち

- ①データの最新性を保つこと。
- ②ローカル、センター双方にとってデータ調査の作業量を最小限にすることを達成するためには、ローカルの雑誌管理システムにおいて、本論で述べてきたようないくつかの問題点を克服して受入れから所蔵データに至るスムーズなデータの処理方法を早急に開発すること、そしてそのローカル・システムとオンライン目録作成システムとを有機的に結びつけるシステム開発が必要であると考え。

参考文献

1. 「総合図書館の電算システムについて」丸善ライブラリ・ニュースNo131 1985. 8P. 8-9 関西大学特集
2. 名古屋大学附属図書館業務電算処理システム報告書 1983
3. 「東京工業大学図書館情報システムについて—文献情報センターとの接続にむけて—（第4次報告）」昭和59年8月 東京工業大学図書館

4. 「兵庫教育大学図書館業務電算化システム」山田常雄・笹川郁雄・石井道悦 大学図書館研究 XXV 1984. 11P. 14-24
5. 「図書館トータルシステム LIAISON—逐次刊行物管理サブシステム: ACCAS—」図書館情報大学図書館業務機械化検討ワーキング・グループ (第15回ドクメンテーション・シンポジウム 1985年6月予稿集より)
6. 「学内逐次刊行物の目録作成システム—同志社大学の概況報告—」吉田和夫 大学図書館研究 XXI 1982. 12P. 12-18
7. 「ILIS/雑誌受入使用者の手引」、「同目録検索使用者の手引」、「同目録作成使用者の手引」富士通マニュアル
8. 「自館所蔵データの機械変換による学術雑誌総合目録和文編へのデータ提出の試み」高島津雪・星野雅英 大学図書館研究 XXIX 1984. 5P. 20-24
9. 学術雑誌総合目録和文編データ記入要領 (1983) 東京大学文献情報センター
10. 図書館における雑誌のあり方—記事検索とのからみで—」井上 如 図書館雑誌76 (4) P. 217-219

その他

研修レポート

東京大学文献情報センターシステムと ISSN

CBI System and ISSN

北海道大学附属図書館 松野 とも子*

要 旨

図書館のコンピュータによるシステム化が進むにつれ逐次刊行物の書誌調整がより大きな問題となっており、文献情報センターシステムと各館のシステムとが接続するためにはこれが重要な課題となる。このための一手段として ISSN を考え、その現状及び文献情報センターシステムとの関わりについて調査を試みた。

Abstract

Bibliographic control of serials is an area of computerization in library. Especially ISSN is important aid in identifying serials at exchange of data between CBI & participating libraries. Provides a survey on the history & present state of ISSN & some theoretical problems.

第一章 はじめに

図書館業務のコンピュータによるシステム化は新しい段階を迎えようとしている。現在の図書館は学術情報システム構想とネットワーク技術の進歩の中で単館からネットワークへと推移する過渡期におかれ、文献情報センター（以下、文情センター）との接続を開始しネットワークとして第一段階に入ったところもある。またネットワークの動きに伴って標準化の研究も進められている。図書館資料の中で大きな位置を占めている逐次刊行物（以下、逐刊）の標準化のための ISDS(Int. Serials Data System)が設立されてから10年が過ぎ ISSN(Int. Standard Serial Number)は目録類に採用され、検索語としても利用され始めた。各館の逐刊システムと文情センターシステムとの接続を考えると、二つのシステム間の溝を埋めるための書誌調整は最も重要なことの一つといえるが、その書誌調整の基本は書誌の同定である。そして ISSN は書誌同定の手段の一つとして機能することができる。そ

*Matsuno, Tomoko : Hokkaido University Library

ここで ISSN の現状を調べそれが期待に応えることができるものなのかどうか、また文情センターシステムとの関係について調査していくことにする。

第二章 文献情報センターシステム

文情センターの逐刊に関係したシステムは冊子体の学術雑誌総合目録（以下、学総目）、個別版、及びオンライン検索・登録の3つのサービス形態となっているが、これらは全て学総目データベースを基にしている。学総目 DB はオンラインの雑誌目録システムと学総目全国調査の融合により形成され、維持方式の開発が行われている(1)。IFLA の雑誌総合目録ガイドラインによれば雑誌総合目録の機能は収集のガイド、所在案内、保存判断の指針、書誌データの情報源、書誌、ILL システムのノードである。(2)また DB となったことにより従来の冊子体が一時的ではあるが作業の完了した完成品であったのに対し、DB となった目録では最新成果であると同時に常に中間加工品として存在し続け、また物としてではなく機能としての総合目録といえることができる(3)。

学総目と各館をつなぐ時、その書誌情報は互換性のとれたものになっていなければならない。しかし、システム側からみると書誌情報の形態は他のコンピュータシステムに比べて明確ではなく、ケタ数や形式が決まっているコード類以外は、何が入ってもある程度許されるシステムになってしまう。そのうえ文情センターは既にコンピュータ化されている各館のシステムに悪影響を与えないよう自由度の高いシステムをとっているため(4)データの取り込み方を各館で選べる反面、互換性の問題は残されたままになっている。一方で逐刊の特徴として上げられる階層的な関係（別冊、特集、臨増等）、変遷による書誌の不安定、誌名の確定困難等による書誌情報の互換性の難しさがあり、これから書誌の調整が重要な問題となってきている。このため国際的な書誌情報の標準化をめざして UBC (Universal Bibliographic Control: IFLA)、NATIS(National Information System: UNESCO)、ISDS(UNISIST/UNESCO)等の活動が進められ、ISDS により ISSN が付与されている。そこで次に書誌情報の同定手段の一つとして ISSN の現状について調べていく。

第三章 ISSN

・成立と組織

ISSN は逐刊の各誌名に与えられるチェック用数字付きの8桁の識別番号で、ISDS により管理されている。この組織は逐刊の識別に必要な情報のデータベース形成と維持に対し共同責任を持ち活動する諸センターの国際的なネットワークであり、全世界の逐刊の信頼しうる登録を行うことを目的としている。ISDS は UNESCO と ICSU が共同で検討した UNISIST 計画の一貫として作られバリエーションに国際センターを設置している。ISSN の開発には ISO が責任を負っている。1968—69年頃米国規格協会より検討されていた ISSN が米国規格となり(ANSI 39.9: ANS Identification Number for Serial Publication, 1971)これが ISDS における逐刊の国際登録の手段として採用され現在では国際規格 ISO 3297: Documentation-International Standard Serial Numbering, 1975)となっている。(5)。

当初 ISO は CODEN を国際規格として検討していたがコンピュータ処理上の問題やアルファベットを用いない国があるという理由から CODEN をやめ ISSN を採用したということである(6)。

ISDS の組織は総会、理事会、国際センター、国別センター、及び技術諮問委員会から成り(7)、その中心となるのは国際センターと国別センター（現在48ヶ国）であり、ISDS の主旨を広め、ISSN を普及させることに努めている。国際センターは主として、・ISDS ファイルの編集と維持管理、・国別センターへの ISSN のブロック割り当て、・国別センターのないところ及び国際組織により刊行された逐刊について ISSN 割り当てと登録等の業務を行う。また国別センターは国内で刊行される逐刊への ISSN 割り当てと書誌データの国際センターへの登録を主な業務としている。国別センターのほとんどが国立図書館であり我が国でも法定納本図書館として網羅的に国内刊行の逐刊を収集している国立国会図書館がセンターとなっていて1974年から正式に実務を開始している。

・付与と登録

付与の方針としては、抄録、索引、引用等に使用される国際性のある逐刊を範囲としている。最初には科学技術系のものに限定していたが学際的な研究が増すにつれ科学技術の分野の境界を定めるのが困難なため人文・社会化学系にも広がってきていて現在ではほとんど分野に関係なくなっている。対象とする逐刊は定期刊行物の他に紀要、会報、モノグラフシリーズ、会議録を含む。登録の基準は国別センターに任せているが国際性、学術性の高いものという区分は難しく、国内では学総目に載せるようなものを目安に登録している。また出版社からの要求のあったもの、国会図書館に創刊号が納本されたもの、国際センターから要求のあったものを中心に登録を行うが、逐刊であれば基本的には拒まないという方針をとっている。

書誌データの登録は識別に必要なデータエレメント（キータイトル、略誌名、異誌名、出版事項、分類、創刊年等）によって行われる(8)。日本センターでは毎月ワークシートで国際センターに送り、オーセンティケーションを受けその後データは ISSN Bulletin となり送られてくる。また ISDS では関心の限られた地域性の強い逐刊について簡易登録という方法を設けている。簡易登録は国別センターに決定権があり、登録する際必要なデータエレメントは国、センターコード、ISSN、キータイトルだけである。

・付与の現状

年	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	合計
付与数	40	700	785	1075	855	1545	1450	1750	1500	1370	11070

(表 1) 日本の ISSN の付与数

日本センターではわずか2名の担当で新規、更新、問い合わせ等の ISSN の維持管理を行っている。年間約10%の誌名変更、休廃刊等によるデータ更新がある。現在までの日本センターの付与数は(表1)であり年間平均1400誌となっている。

このうち納本が確認されているのは8858誌（1985. 7月現在）、また日本センター設立の前に国立セン

ファイル名	タイトル数	ISSN有り	北大所蔵数	北大ISSN有り
学総目日文編	40237	5688	14277	3539
学総目欧文編	90136	12993	14666	3183
LC MARC(S)	194564	102609		
北大図購入（和）	177	73		
北大図購入（欧）	383	287		

(表 2) ISSN 表示数

ター等で付与されたものもいれると、現在の登録数は約16000誌である。世界全体での登録数は20万誌（簡易登録を含めると25万）になっている(9)。次に学総目、LC 及び北大附属図書館の場合について ISSN 付与数を調べてみた。

これをみると LC MARC(S)では約53%の付与数であるが学総目では和文、欧文編共に約14%、北海道大学全体の所蔵分については和文21%、欧文25%という低い結果であった。学総目での ISSN は所蔵館の記入に任せているため、ISSN が表示されていても記入しなかったものもあると推測される。またこの数字は既に廃刊になっているものも含んでいるため、次に現在刊行されているものという事で北大附属図書館で購入しているものの調査を試みると、和雑誌は41%、洋雑誌は75%という結果になった。和雑誌については国内のみの関心の限られたものがあるためどうしても少し低くなるようである。また北大附属図書館で寄贈を受けている各大学紀要類では約90%に、北大で刊行している紀要類では約80%に ISSN が表示されていた。これから現在刊行されているもの、特に国際的、学術的なものは、かなり付与率が高くなっているといえる。ISSN の表示形については表紙、裏表紙、奥付等のわかりやすい位置、特に表紙の右上が基準とされている。(10) が約8割が右上につけられていた。表記の仕方については Colorado State Univ.での調査により誤り例(CN ISSN 0002-6638, 0428-8286-JU JSSN 等) (11) があるが今回調べた限りではこのような誤りはなかった。今後の ISSN の普及については出版界の協力と各種目録への記載がポイントになると思われる。後者については大学、研究機関等の目録に多く見られるようになってきているが、前者については、出版社にとって ISSN 表示による即効性のある利点があるわけではなく、また雑誌コードの問題もあり容易ではない。また米国における 著者権関係や郵政省の登録コードとしての例のように (12)、図書館からの利用でより普及させていくという方法もある。

第四章 ISSN と学総目

学総目では書誌識別のコードとして ISSN と CODEN(CODEN は今回の欧文編調査から加わる)を持っている。CODEN は科学技術系のものに限られているため ISSN の役割は大きいといえる。

そのため、学総目では ISSN は必須項目として扱われ (13)、ISDS の MT も発注中である。また学総目欧文編の書誌記述については完全さを求めず、出版した国の責任において完全な書誌記述を作成しそれとのインターフェイスを保持すれば良いということから ISSN を採用するという考え方もある (14)。ISSN を学総目で使用する時には ISDS と学総目との書誌情報のとらえかたのくい違いが問題となる。これは特に変遷関係で顕著にあらわれる。多発する変遷は逐刊の特徴であり書誌情報を扱う上で大きな問題である。LC MARC(S)の30万余りのレコードのうち、過半数は訂正 (更新) レコードであり、純誌名数は10数万であるという。(15) また学総目では変遷マップによれば DB に含まれる誌名数44179のうち、3割以上を占める13967誌が変遷に関係している (16)。誌名変遷のとらえかたは文情センターと ISDS とで異なっている。例えば学総目では年代順のシリーズ表示 (2nd Series 等) が変遷の対象とされるが (17)、ISDS では変遷とみなさない (18)。このとき学総目を検索すると一つの ISSN で二つの誌名が出てくることになる。また ISDS では発行機関の名称変更は全て変遷とみなしているが、学総目では誌名が識別困難な一般的名称 (Bulletin, Report 等) の時に限られている。この場合一誌に二つの ISSN がつくことになるので、旧い方を XISSN を使用して対応していくと思われる。

誌名については ISDS のキータイトルは識別のために唯一の誌名とされるため一般的名称のものは学総目のタイトルと異なる形になることがある。更にキータイトルは漢字などのアルファベットを使用しない言語については ISO 規格にそいローマナイズするため、読み方の問題もあり学総目との間にすっきりしないものが出てくる可能性がある。また日本センター設立前に BOWKER 社や国際センターが日本の逐刊に ISSN を付与していたものは、外国での付与ということもあってタイトルの取りかたの誤りが多く、約4000誌のうち約300誌の重複割り当てが発見され (19) 現在日本センターで調整を行っているということもある。

学総目は ISSN を活用できるように対応づけられていると思われるが、オンラインの雑誌目録システムが稼働し、実際に運用されていく時、上記のような書誌情報の問題からスムーズに行かないことも起こると思われる。しかしそれにより問題が明確にされることで書誌調整の必要性の意識が高まり、学総目と各館のシステムを互換性のとれたものにしていく作業も進んでいくと思われる。

第五章 おわりに

逐刊の標準化は逐刊特有の書誌的問題点である変遷、階層性、誌名の不確定性のため国際的にはもちろん国内においても難しい。ISSN は約10年間で着実に付与数を増やしつつあり、現在刊行されているもの、特に国際的、学術的なものについてはかなりの付与率になってきている。一方、図書館のコンピュータ化の中で検索キーとして、また目録の中にもよくみかけるようになってきた。今、文情センターシステムを中心とするネットワークが動き出したことは ISDS、ISSN にとって一つの転機であるといえる。今までわずか二人という人員で情報量の多い日本の ISDS 業務を行ってきた日本センターは超人的とも言えるが、その普及活動に無理があるのは当然であろう。ネットワークに参

加する図書館では今までより確実に ISSN を意識し、使用することが多くなると予測される。利用が増えるにつれて ISSN に対する関心が高まれば需要の増加、登録数の増加にもつながる。それは書誌情報の標準化を推進することにもなるだろう。ISSN が、そして文情センターシステムが実用化の時代に入りつつあることを期待する。

最後に、セミナーにおいては井上先生はじめ諸先生、職員の皆様、国立国会図書館の川添氏、東大附属図書館の石倉氏にたいへんお世話になりました。この場をかりてお礼申し上げます。

引用文献

1. 「学術雑誌総合目録欧文編新版データベースの開発」文献情報センターニュース 6 1985 p1
2. 「IFLA Guideline Cataloging of Serials」内藤先生の講義より
3. 「参加組織の条件整備その1」文献情報センターニュース6 1985 p 9
4. 安達淳「画面アプリケーション向けネットワークの基本構成」文献情報センターニュース6 1985 p17
5. 国立国会図書館「国内逐刊行物リスト」1981 piii
6. Emery Koltay「Int. Standard Serial Number」New Serial Title 1950-1970 Bowker 1973 pxv-xvi
7. 「ISSN について」国立国会図書館案内 12
8. ISDS int. Centre「ISDS Mannual 1983」p30
9. 付与の数も含め、川添氏とのインタビューより
10. 国立国会図書館「国内逐刊行物リスト」1981 piii
11. David Walker Lupton「Tracking ISSN」The Sirial Librarian 4 (2) 1979 p190-195
12. David Walker Lupton「Tracking ISSN」The Sirial Librarian 4 (2) 1979 p196-197
13. 「学術雑誌総合目録欧文編データ記入要項1985」p44
14. 柴田正美「学術雑誌総合目録人文・社会科学欧文編データベースの編集を終えて大学図書館研究 17 1980 p37
15. 永田治樹「つなぎとしての逐刊刊行目録」メディアの多様化と図書館 論集・図書館学研究の歩み 第4集 p48
16. 根岸正光 井上如 柴田正美 藤原鎮男「学術雑誌総合目録人文・社会科学欧文編データベースにおける誌名変遷状況の計量分析」ドクメンテーション研究 31(4) 1981 p152
17. 「学術雑誌総合目録欧文編データ記入要項1985」P41
18. ISDS Int. Centre「ISDS Mannual 1983」p58-60
19. 川添登美夫「ISDS と ISSN について」ドクメンテーション研究 30(8) 1980 p369

研修レポート

図書館 昨日、今日そして明日 — ニューメディア時代へ向けて

The Library : It's Yesterday, Today and Tomorrow
— Towards the New-Media Age

広島大学附属図書館 諸富 秀人*

要 旨

今日におけるコンピュータ技術・通信技術等の飛躍的な進歩が、いわゆるニューメディア時代を現出させようとしている。図書館をめぐる今日的情況は、「紙なし情報化社会」へ向かおうとしているかに見える。「ゲートウェイ」というシステムは、学術情報の流通が図書館の介在なしに行なわれる時代の到来を予想させる。この論文では、図書館の本質・役割とは何かということ踏まえ、ニューメディア時代へ向かう図書館の将来を展望しようと試みた。

Abstract

Rapid technological progress in computers and communication is bringing about a “new media age”. The library is moving fast towards what F. W. Lancaster called a “paperless information system”. If we look at information retrieval systems such as “Gateway”, we can imagine that science information needs can be served without the present day library. In this report, I tried to grasp what the library's essence is, what its duties are and what its prospects for the future are.

1. はじめに

今更、マクルーハンやトフラーの言を持ち出すまでもなく、コンピュータ技術・通信技術・エレクトロニクス技術等の技術革新が急速に進んでいる。それにともない、知識量・情報量とも増大し続け、その社会的ニーズも極めて多様化ししかも緊要なものとなっている。こうした状況のなかで、将来の図書館は社会的ニーズに対応できなくなり昔ながらの書庫に後戻りするのかわ、それとも情報センターとして生き残れるのかの岐路に立たされていると言える。この機会に図書館とは何か、その本質・役割とは何かを捉えなおし、図書館の将来を考えてみたい。

*Morotomi, Hideto : Hiroshima University Library

2. 図書館とは

古代において人類は、知的コミュニケーションの手段として「ことば」を得、更に「文字」を発明した。その文字を使って知識・情報を記録することによって、時間・空間を越えたコミュニケーションを可能とした。その記録されたものが「図書」であり、その容器「館」が図書館である。ところで、図書館という言葉は、フランス語では Bibliothèque、ドイツ語では Bibliothek、英語では Library というが、その語源はラテン語の本を意味する“biblia”と箱を意味する“theca”という語が結びついた“bibliotheca”という合成語、あるいは“librarium”からきている。最古の図書館といわれるアッシュール・バニパルの王宮図書館に始まる図書館のあゆみを眺めてみると、図書館とはまさに人類の知的発達の歴史そのものであるといえる。その初期の段階では、頂点にある権力者や一握りの知識階級のものでしかなかったが、その二千年を越える歴史の流れの中で少しずつ民衆に開かれていき、今日すべての人がそれを享有できるようになった。図書館は今後さらに、二百万年からなる人類の知的活動の成果を集積し、保存し、提供してゆく責務を負っている。

3. 図書館の機械化

計算機の歴史は、1623年 W. Schickhald の発明に始まる。そして今世紀中頃の H. Aiken の MARK や、J. P. Eckert のいまや伝説的な ENIAC を経て第五世代の時代を迎えつつある。情報の洪水に対処すべく開始された図書館業務の機械化は、1936年テキサス大学図書館が IBM の創始者 H. Hollerith が発明したパンチカードマシンを貸出業務に使用したことに始まる。それ以後、スタンフォード大学の BALLOTS をはじめ多くの試行錯誤の上に到達したのが、OCLC, RLG, UTLAS などの書誌情報共同利用機構(Bibliographic Utilities)にみられる図書館ネットワークであった。日本においても、1960年代中頃以後各図書館が個々にシステムを開発してきたが、学術情報構想の一環として1984年4月図書館のネットワークを志向する東京大学文献情報センターが発足した。その目的は、文献の目録、所在情報データベース構築とその利用を通じて大学図書館での資源共用を促進することにある。センターシステムの特徴としては、ユーザーインタフェースとしての画面制御システムや専用目録端末の開発、またネットワーク技術として異機種電算機間の接続を可能にする NVT プロトコルの機能拡張や小規模図書館の接続を可能にする PAD 方式の開発があげられる。これらのシステムの上に各種のマークからなるデータベースをのせて、図書館業務のトータルシステムを全国的な図書館オンラインネットワークで一挙に実現しようとしている。

4. ニューメディア時代の情報システム

4.1 ニューメディアの出現

人類の知的情報の伝達媒体は、西暦105年以來ずっと紙が主役であった。19世紀後半に発明された電気通信系メディア（電信・電話）は、エレクトロニクス技術やコンピュータ技術と結合し、いわゆるニューメディアとして紙の印刷系メディアを脅かす存在になりつつある。現在話題となっている光

通信は、電気信号と光源（レーザーダイオード）で光に変え、光ケーブルを伝送路とし、受光器（フォトダイオード）で電気信号に再変換する通信方法である。学術情報には文字だけでなく図表・画像があり、高速で正確な伝送が必須であることから光通信が最適である。NTT が実験的に行なっている INS は、いわゆるサービス総合デジタル網(ISDN: Integrated Services Digital Network)で、音声・データ・画像等個別の通信網をデジタル化し、光ケーブルを通信線として一つの通信網に総合し、より高速で低料金の回線を提供するものである。今世紀末までに、全国ネットによるマルチメディア通信サービスを目指している。

4. 2 学術情報流通への応用

将来の図書館を考えるうえで、現在行われている、あるいは極く近い将来において行なわれるであろうニューメディアの学術情報への応用例をみることは重要であろう。

4. 2. 1 電子出版(Electronic Publishing)

出版へのコンピュータ利用の最も早い例は、米国医学図書館(NLM)が1964年に索引誌 Index Medicus の編集を電算化し、電子写植機 GRACE を用いて出版したものであ⁽¹⁾。この時の印刷過程の副産物(By-products)が、後に情報検索の対象としてのデータベースになるのである。技術的な背景として、大容量の記憶装置・遠隔アクセス・タイムシェアリング・データベース管理システム等のコンピュータ技術の開発と、通信コスト及び端末機の価格低下などがあげられる。印刷技術の面では、ワードプロセッサとのリンク・マルチターミナルシステムの採用・レーザによる直接製版等が機械可読データの蓄積コストの低下とあいまって印刷業界を変えつつある。電子出版の一例として、1980年に Afte 社の刊行した“Academic American Encyclopedia 22vols”の機械可読版に対するオンラインアクセスの実験が行われた。これは OCLC の開発した“Channel 2000”と名付けられた3ヶ月にわたる試験的な「コンピュータ家庭情報サービス」の一環として実施されたということである⁽²⁾。電子出版の動きは、抄録誌・索引誌の分野で特に顕著で、ついで辞典類・統計書の分野に及び、早晩、雑誌・図書の分野でもそうした動きがでてくるものと思われる。その最終的な姿は、もはや印刷物としては発行されず、電子化されたままディスク等に納まり、これをオンラインで検索し必要な部分のみ On-Demand すればよい。85年10月に三修社から「最新科学技術用語辞典」の CD-ROM 版が発売され、86年4月には日外アソシエーツ社が人名辞典等の CD-ROM を発売する予定という。

4. 2. 2 フルテキスト・データベース

電子出版の一形態として、一次情報そのものを電子的な方法でユーザー端末に直接伝送するフルテキスト・データベース・サービスは、大記憶容量をもつ光ディスク等の出現で膨大なデータ入力と比較的安価にできることと、一次情報作成工程の電算化により、入力したものをそのまま利用できることから、BRS の HBRO をはじめ各種のフルテキスト・データベースが提供されている。その中のあるものは、ダウンロードさえ商品サービスとして提供している。また国際的な学術雑誌出版社の数社で新しい事業体を作り、世界の主要な学術雑誌の掲載論文を光ディスクに記録し、リクエストによりそのハードコピーを提供することを目的とする ADNIS 計画や ARTEMIS 計画⁽³⁾、通信衛

星によるドキュメントデリバリーを目指す APOLLO 計画⁽⁴⁾も開始された。

4. 2. 3 電子雑誌(Electronic Journal)

電子雑誌は、学術雑誌の編集・審査・配布を電子的な手法でのみ行なうものである。ここにニュージャージー工科大学での実験報告が紹介されている⁽⁵⁾。これは、1976-1980年にかけて行なわれた EIES(Electronic Information Exchange System)と呼ばれ、電算機端末を利用した研究者同志のメッセージ交換・研究者グループでの会議・共同著作・ワークショップなどの機能をもつ。その手順は次のとおりである。①著者は電子形態で論文を書き、②完成したら“Submit”コマンドで投稿し、③編集者はこれを“Read”し著者名を除いて査読者に送り、④査読者・著者・編集者で修正のための討議を経て、⑤論文が修正され発行が決まったら、抄録が加えられ数千の読者に提供される。イギリスにおいても、BLEND-LINC という同様の実験が行なわれた。いずれにしても、電子雑誌が社会的に受け入れられるかは、ひとえに論文の質と雑誌の権威付けにある。

4. 2. 4 ゲートウェイ・システム

データベースによる情報検索は、そのコマンド・プロトコル・ファイル構造・メッセージ類に習熟したいわば情報専門家しか検索できなかつたし、そこに図書館員が学術情報流通に介在できた。ゲートウェイ・システムは、検索方法が非常に簡単かつ多機能(Automatic Logon, Automatic database selection, Automatic term selection, Assisted very negotiation etc.)で、情報検索技術の訓練をうけていないエンドユーザーが簡単に検索できる⁽⁶⁾。米国では、1982年にパソコンの売り上げが200万台に達すると同時に、パソコン通信用のソフトと学術情報以外の分野(ビジネス情報・ニュース・消費者案内等の一般向け)の各種データベースが、ダウジョーンズ社などから提供され始めた。数万円のパソコンと通信用ソフトがあれば、これら各種の画像を含む情報をだれもが容易に入手できる。パソコンが、通信回線を通じ大型コンピュータの端末として使用されることが一般的になってきた。こうした状況は学術情報流通の面にも反映され、1980年代初めには Scimate, Micro-Disclosure, Seacher-Helper というシステムが商品化された。私はゲートウェイの一つである BRS/Sounders の Colleague というシステムで情報検索を試みたが、メニュー方式に上りシステムと会話しながら簡単に検索できることを実体験した。ゲートウェイ・システムが一般的になると、学術情報流通に図書館員が介在する必要がなくなる事態も予想される。いずれにしても、学術情報へのニューメディアの適用は、著作権の問題が解決すれば一挙に展開するだろう。

5. 図書館の将来—ニューメディア時代への対応

図書館は今、情報の洪水の中で翻弄され押し流されまいと懸命に踏み止まっている。いや、むしろ流れに乗り遅れまいとしていると言うべきであろうか。図書館にその歴史が始まってから今日まで、その本質である知識(情報)の収集・蓄積・保存・提供という与えられた役割に対して、その時代時代の社会的ニーズに答え、かつ技術的發展を吸収しつつ今日まで歩んできた。今世紀以前の2000年の歴史の中で、図書館におけるパラダイム変革は紙の発明と印刷術の発明だけであった。ところが今世

紀の中ば以後のわずか50年間に、電算機の導入とニューメディアによる情報流通というパラダイムの変革に遭遇した。

5. 1 未来予測

イリノイ大学の F. W. Lancaster は、その著書“Toward Paperless Information Systems”で、「西暦2000年の情報システムでは、学術情報はもちろん一般的な本や雑誌も電子的な媒体に記録され、研究者は簡単にしかも気楽にデータベースにアクセスし必要な情報を得る。したがって、利用者は図書館に足を運ぶ必要はなく、図書館そのものの必要性もなくなる」と予言している。また「図書館員の残された道は、目録とかレファレンスとかの個別業務の専門家ではなく、情報専門家として熟達するしかない」ともいっている。雑誌「現代の図書館」は、その21巻3号(1983. 3)と22巻2号(1984. 6)で『図書館の未来像』を特集しているが、いずれからも明確なイメージを抱くことができない。ここに未来の図書館システムの子測がなされている。それによると、①原文書電子ファイリング②目録カード自動遡及変換システム③自動目録システム④画像・音声・文字による自動図書館案内システム⑤利用者コンサルテーション⑥図書館経営意思決定支援システムなどが、西暦2010年位までには実現するという⁽⁷⁾。人工知能を備えた第五世代コンピュータならこんなシステムも夢ではなくなるのだろうか。また科学技術庁においても、1998年までの情報技術の進歩子測がなされている⁽⁸⁾。

5. 2 ニューメディアのインパクト

図書館をめぐる今日の状況(出版物の製造工程が電算化され、それに伴い原稿が磁気テープなどの機械可読の媒体に記録され、データベース化を可能にした。文字処理と画像処理の統一したシステムが開発されれば、より完全なデータベースサービスが可能となる)は、Lancaster の言うように紙なし情報化社会へ向かおうとしているように見える。図書館の収集する資料のうち、紙以外の記録メディアのものが少しずつ増えてきた。第4章でみたように、電子出版からゲートウェイに至るまでの動きが出版そのものの変容をもたらし、同時に図書館もまた変容を迫られている。紙の出版物の生産と流通コストは毎年上昇しているが、電子形態の出版物(?)のそれは逆に低下している。特に二次資料(抄録誌・索引紙)が、完全に印刷体から電子形態にとってかわるのは時間の問題である。つづいて百科事典等の参考図書(ファクトデータベース)、雑誌・図書(フルテキスト・データベース)の順に電子形態の出版物の比重が高まっていくだろう。

5. 3 レファレンスの将来

図書館の将来を論じたいくつかの論文をみると、いずれも将来の図書館員は印刷されたものと、電子的メディアの両方をサービスするのに十分な訓練を積まねばならないと言っている。図書館は、商業ベースビデオテックス等の情報業界と競合する。図書館がこれに対抗するには、図書館員が高度の情報処理能力をもたなくてはならないとも論じている⁽⁹⁾。また将来のレファレンサーは、情報ニーズを先取りし、不特定多数の将来の依頼者を想定して、専門データベースを作成するとともに、研究者間の情報流通ネットワークの形成にも責任を負うことになるかも知れない。⁽¹⁰⁾ 自然言語による知的会話・推論・問題解決・意思決定支援(DDS)の能力をもつ第五世代コンピュータの時代になり、

図書館知識ベースによるエキスパートシステムが開発されると、レファレンスサービスも質的変換をとげるだろう。

5. 4 図書館の対応

ニューメディア時代の電子技術や通信技術は、今後の図書館サービスの発展に無限の可能性をもたらしてくれることは確かである。研究者の切実な要望は、ドキュメントデリバリーである。この情報の提供という面においては、学術情報システムの電子メール機能を利用した ILL で答えることができるだろう。この ILL では、INS の高度デジタル伝送サービスを利用したフルテキストの提供も早晩可能となるはずである。NTT 三鷹電々ビル内の INS モデムルームで見学した、超高速デジタルファクシミリ(BHF)の機能は、12-16dot/mm と高精細で、A4版文書一枚の伝送速度は白と黒の2階調の場合約2秒で、中階調を含む場合は6秒である。写真など濃淡のある原稿や精細な図・画・小さな文字などもきれいに伝送できる。さらに書籍等厚みのあるものも一旦コピーしないで直接伝送でき、カラーファクシミリも用意されていて充分学術情報の伝送にも耐え得るものである。このような技術を手に入れた図書館は、まさにその時こそ、図書館の究極の目標である、だれでもが、いつでも、どこからでも、いかなる情報も、容易に、速く、安く手に入れることができる時代へと踏み出す。これまで図書館は、図書館を取り巻く環境の変化に対処すべく図書館業務の機械化、そして学術情報システムにみられる学術情報のネットワーク作りに取り組んできた。それも完成間近である。一方、これまでの図書館における情報サービスの主体は、研究者の情報入手に対する支援・補助であった。しかし、それだけでは研究者より専門知識の劣る図書館員は、やがては「無人図書館論」にみられるように無用になる時がくることもありうる。発想の転換をはかり、能動的な情報サービスを志向すべきではなからうか。そしてニューメディア時代へ向かう図書館としては、社会的・技術的・文化的状況の変化に的確かつ柔軟に対応できる態勢を常にとっておき、図書館に対する多様なニーズに答えなければならない。

6. あとがき

ニューメディア時代への対応ということで、ややもするとその技術的な側面にばかり目が奪われ、それを追うことに精一杯で、図書館の本質を忘れがちである。将来の図書館がニューテクノロジーで武装した「電子図書館」であろうとも、図書館の本質が、知識の集積・保存・提供の場であると同時に、学び、考え、安らぐ「場」でもあることは、ニューメディアの技術が進んだ将来の図書館においても変わらないだろう。むしろ技術社会においてこそ、図書館の社会的“Re-creation”の「場」としての要素を大切にしたい。

このレポートは、セミナー期間中に見・聞き・体験したことをまとめたセミナー卒論の抜粋である。

引用文献

- (1) 日埜 衆司. “電子出版”. ドクメン研究. Vol.35, No 1, p. 18 (1985)
- (2) 村上 正志訳. “オンライン百科事典—その可能性”. びぶろす. Vol.34, No 8, p. 8 (1983. 8)
- (3) 武内 寿. “学術情報のフルテキスト検索と伝送”. 科学技術文献サービス. No64, p. 1 (1983)
- (4) 武内 寿. “EC のドキュメント伝送計画”. 科学技術文献サービス. No69, p. 12 (1984)
- (5) 仲本 秀四郎. “電子版雑誌の誕生”. ドクメン研究. Vol.33, No8, p. 378 (1983)
- (6) Williams, M. E.(Univ.of Illinois). “Highlight of the online datebase field-Gateways, Front ends and Intermediary Systems”. National Online Meeting, 1985.
- (7) 池田 秀人. “学術情報システムの現状と展望”. 医学図書館. Vol.31, No 1, p. 7 (1985)
- (8) 黒沢 慎治. “科学技術分野の情報サービスの将来”. 現代の図書館. Vol.21, No 3, p. 180 (1983)
- (9) Ross, Richard W.. “Technology and the modern library. “Library Journal. Vol.109,No.11, p.1183 (1984.6)
- (10) 田屋 裕之.“ニューメディア時代の図書館とレファレンスサービスのゆくえ”.ライブラリアンズ・フォーラム.Vol.2, No 2, p. 22 (1985)

東京大学文献情報センター日誌

昭和60年

3. 30～ 4. 7 センター長, 海外出張 (オーストラリア)
4. 1 井上如助教授: 教授に昇任
梶史郎事務長 (東京大学経理部主計課課長補佐), 佐藤隆雄共同利用掛長 (東京大学新聞研究所), 郷端清人事務官 (広島大学附属図書館) 着任, 大山敬三氏助手に採用
タスク・フォース着任: 村田康彦 (徳島大学), 木村祥子 (京都大学), 庄司三千子 (千葉大学)
永田治樹目録情報掛長: 東京大学附属図書館整理課目録主任に配置換
4. 16 タスク・フォース着任: 伊藤彰 (大阪大学), 斎藤豊 (弘前大学)
5. 1 タスク・フォース着任: 米澤誠 (東北大学)
5. 13 第7回運営委員会
5. 16 小山修美 (東京大学庶務部併任) 着任
5. 20 昭和60年度文献情報センター・前期セミナー開講式 [研修員: 牛崎進 (立教大学), 尾城孝一 (名古屋大学), 神田百合枝 (東京大学), 辻川輝男 (関西大学), 松野とも子 (北海道大学), 諸富秀人 (広島大学)]
5. 31 第13回文献情報センターシステム全体会議
6. 15 第4回文献情報センターシンポジウム (於: 東京工業大学)
6. 23～ 7. 6 センター長, 海外出張 (フランス)
6. 29 画像入力装置搬入
7. 1 タスク・フォース着任: 本郷清次郎 (琉球大学)
7. 11 臨時運営委員会
8. 1 安達淳講師: 助教授に昇任
8. 15～20 センター長, 海外出張 (アメリカ合衆国)
8. 20 『東京大学文献情報センター紀要』第1号刊行
9. 20 接続図書館説明会
" 『目録所在情報サービス利用の手引』 (第1版) 刊行
9. 24 図書館ネットワーク専門委員会
9. 25～27 第1回目録システム講習会
9. 30 タスク・フォース帰任: 庄司三千子 (千葉大学), 伊藤彰 (大阪大学), 斎藤豊 (弘前大学)
10. 9 目録情報専門委員会
10. 14 昭和60年度文献情報センター・前期セミナー閉講式 (修了者6名)
10. 16～18 第2回目録システム講習会
10. 20～26 センター長, 海外出張 (フランス)
10. 21 学術雑誌総合目録欧文編データ記入説明会「東京大学」(於: 東京大学)
10. 24 学術雑誌総合目録欧文編データ記入説明会「九州・沖縄地区」(於: 九州大学)
10. 25 学術雑誌総合目録欧文編データ記入説明会「東日本地区」(於: 東京大学)

- | | |
|---|--|
| <p>10. 30 学術雑誌総合目録欧文編データ記入説明会「北海道地区」(於：北海道大学), 「中部地区」(於：名古屋大学)</p> <p>10. 31 学術雑誌総合目録欧文編データ記入説明会「関西地区」(於：大阪大学)</p> <p>11. 1 タスク・フォース着任：鹿島玲子(千葉大学)</p> <p>11. 8 運営委員会・特別ワーキンググループ</p> <p>11. 11 昭和60年度文献情報センター・後期セミナー開講式〔研修員：市川美知代(清泉女子大学), 伊藤茂夫(福井大学), 柿沼澄男(図書館情報大学), 金子進(創価大学), 末木慎一(東京工業大学), 渡部哲夫(山口大学)</p> <p>11. 13 目録情報専門委員会ワーキンググループ</p> <p>11. 16 タスク・フォース着任：星和夫(新潟大学), 土師守(群馬大学), 大久保佳典(京都工芸繊維大学), 瓜生照久(九州大学)</p> <p>11. 17~12. 1 文部省科学研究費補助金・海外学術調査／(第1班) 安達助教授, 大山助手</p> <p>11. 20~22 第3回目録システム講習会</p> <p>11. 29 第5回文献情報センターシンポジウム(於：大阪大学・中之島講堂)</p> <p>11. 30~12. 16 文部省科学研究費補助金・海外学術調査／(第2班) 根岸助教授, 橋爪助手</p> <p>11. 30 HITAC L-450システム搬入</p> <p>12. 2 昭和60年度データベース担当者養成研修開講式〔研修員：石村恵子(埼玉大学), 井関泰夫(大阪大学), 京藤貫(富山大学), 笹山達成(南山大学),</p> | <p>塩谷玲子(弘前大学), 伊達恭子(京都大学), 中井えり子(名古屋大学), 仲清志(兵庫教育大学), 丹羽弘(名古屋工業大学), 林茂榮(京都大学), 松澤志津代(茨城大学)〕</p> <p>12. 10 システム運用専門委員会</p> <p>12. 16 目録情報専門委員会ワーキンググループ</p> <p>12. 31 タスク・フォース帰任：米澤誠(東北大学), 本郷清次郎(琉球大学)</p> <p>昭和61年</p> <p>1. 5~20 文部省科学研究費補助金・海外学術調査／(第3班) 井上教授, 宮澤助教授</p> <p>1. 18 通信制御処理装置(CCP)搬入・増設</p> <p>1. 22~24 第4回目録システム講習会</p> <p>1. 24 目録情報専門委員会ワーキンググループ</p> <p>1. 27 第8回運営委員会</p> <p>1. 29 図書館ネットワーク専門委員会</p> <p>2. 2~10 センター長, 海外出張(フランス, ベルギー)</p> <p>2. 10 目録情報専門委員会ワーキンググループ</p> <p>2. 12~14 第5回目録システム講習会</p> <p>2. 24 目録情報専門委員会ワーキンググループ</p> <p>2. 27 接続図書館説明会</p> <p>3. 1 昭和60年度データベース担当者養成研修閉講式(修了者11名)</p> <p>3. 7 システム運用専門委員会</p> <p>3. 9~23 文部省科学研究費補助金・海外学</p> |
|---|--|

- | | |
|--|--|
| <p>術調査／(第4班) 内藤助教授, 橋爪
助手</p> <p>3. 10 目録情報専門委員会</p> <p>3. 12~26 宮澤助教授, 海外出張(連合王
国)</p> <p>3. 15 仕様検討委員会</p> <p>3. 17 第9回運営委員会</p> <p>3. 19 昭和60年度後期セミナー閉講式(修了
者6名)</p> <p>3. 29 内藤助教授海外研修旅行(アメリカ合
衆国/合衆国政府招請) 出発(～5. 3)</p> | <p>3. 30 安達助教授文部省(短期) 在外研究員
(アメリカ合衆国) 出発(～9. 30)</p> <p>3. 31 タスク・フォース帰任:
60. 4. 1～; 木村祥子(京都大学),
60. 11. 1～; 鹿島玲子(千葉大学),
60. 11. 16～; 星和夫(新潟大学),
土師守(群馬大学),
大久保佳典(京都工芸織
維大学),
瓜生照久(九州大学),
小山修美(東京大学庶務部) 併任解除</p> |
|--|--|

学術情報センター日誌

昭和61年

- | | |
|--|--|
| <p>4. 5 学術情報センター創設</p> <p>” 学術情報センター職員着任</p> <p>所長(事務取扱) 猪瀬 博</p> <p>[管理部]</p> <p>管理部長 伊藤 公紘</p> <p>総務課長 梶 史郎</p> <p>システム・データベース課長
雨森 弘行</p> <p>[総務課]</p> <p>総務係長 渡辺 道夫</p> <p>会計係長 飯塚 正人</p> <p>共同利用係長 佐藤 隆雄</p> <p>研修・資料係長 渡邊 正昭</p> <p>会計係(併任) 中島 茂</p> <p>会計係(併任) 早川 敦夫</p> <p>共同利用係(併任) 飯田 信之</p> | <p>[システム・データベース課]</p> <p>システム管理係長 清水 二郎</p> <p>システム業務係長 中村 茂樹</p> <p>ネットワーク係長 渡邊 博</p> <p>企画係長 石井 啓豊</p> <p>データベース形成係長 牧村 正史</p> <p>目録情報係長 郷端 清人</p> <p>ネットワーク係 大日方一男</p> <p>データベース形成係(併任) 飯島 重美</p> <p>目録情報係 羽田 和久</p> <p>目録情報係 室橋 眞</p> <p>[研究開発部]</p> <p>研究開発部長 市川 惇信</p> <p>[学術情報研究系]</p> <p>研究主幹(教授) 井上 如</p> <p>教授 根岸 正光</p> <p>助教授 宮澤 彰</p> <p>助教授 安達 淳</p> <p>助手 大山 敬三</p> |
|--|--|

- [システム研究系]
- 教授 濱田 喬
 助教授 浅野正一郎
 助教授 橋爪 宏達
4. 7 海部文部大臣，センター来訪：開所セレモニー
 “ タスク・フォース着任：松本昌子（京都大学），大石博昭（国文学研究資料館），巖本康治（金沢大学），東海安興（長崎大学），有馬秀子（愛媛大学），青山功（大阪外国語大学），真中進（静岡大学），岡本正貴（愛知教育大学）
4. 23～25 第1回目録システム講習会
 4. 25～ 5. 24 宮澤助教授，海外出張（中国）
 5. 1 タスク・フォース着任：貝田辰雄（北海道大学），早野裕士（東京大学），堤守政（名古屋大学），坂賀力（大阪大学），河津秀行（九州大学）
 5. 3 内藤助教授，海外研修旅行帰国
 5. 9 OCLC 社長 Mr.Rowland C.W.Brown 氏来訪東京大学附属図書館において講演
 5. 11～16 浅野助教授，海外出張（米国）
 5. 16 タスク・フォース着任：藤岡節夫（京都大学）
 5. 19 昭和61年度前期セミナー開講式〔研修員：井上正則（同志社大学），牛島裕（近畿大学），河野建二（愛媛大学），二村英文（名古屋商科大学），松井好次（東北大学），吉田英明（鹿児島大学）〕
 5. 23 行政管理官視察
 5. 27 昭和61年度第1回総合目録データベース実務研修開講式〔研修員：阿部佳市（東北大学），工藤一郎（東京大学），久保田良子（東京大学），栗原道夫（新潟大学），斎藤泰則（慶應義塾大学），櫻井美由喜（兵庫教育大学），高橋宏美（東京大学），高橋實智子（京都大学），田中耕二，（京都大学），松藤典生（熊本大学），山下進（大阪大学）〕
5. 28～30 第2回目録システム講習会
 6. 10～20 所長，海外出張（スウェーデン，フランス）
 6. 14～20 研究開発部長，海外出張（スウェーデン，フランス）
 6. 24 国連大学シンドラ局長，センター来訪
 6. 27 第1回評議員会議
 6. 30 第1回運営協議員会議
 7. 2 衆議院文教委員会，センター視察
 7. 11 センター創立記念パーティ（於：如水会館）
 7. 21 昭和61年度第1回総合目録データベース実務研修開講式（修了者11名）
 7. 22 オックスフォード大学ティトラー氏，センター見学
 7. 23 アメリカ議会図書館吉村ヨシコ氏，センター見学
 8. 5 学術情報センター第1回シンポジウム（於：日本学術会議講堂）
 8. 18 ケンブリッジ大学図書館小山騰氏，センター見学
 8. 20 FIZ KARLSRUHE イェンシュケ氏，センター来訪
 西ドイツ各図書館（アウグスブルグ大

- 学図書館他) 職員, センター見学
8. 23 RLG マッコイ氏, センター見学
8. 26 韓国公州師範大学玄圭燮氏, センター
見学
8. 28 IFLA 参加者, センター見学
8. 29 ドイツェ・ビブリオテーク ケルム
・バルバラ氏センター見学
9. 2 昭和61年度第2回総合目録データベー
ス実務研修開講式 [研修員: 小幡砂智
子 (新潟大学), 恩知陽子 (山口大
学), 菊地房雄 (東北大学), 金城照子
(琉球大学), 熊谷伸彦 (東京工業大
学), 河野富行 (大阪大学), 齋藤穂
(弘前大学), 堤美智子 (京都大学),
友光健二 (東京学芸大学), 中野由紀
夫 (九州大学), 平尾行藏 (慶應義塾
大学), 藤巻美恵子 (東京大学), 藤原
誠 (名古屋大学)]
9. 3~5 第3回目録システム講習会
9. 9 ILL システム協力者会議
9. 30 安達助教授, 文部省 (短期) 在外研究
員帰国
10. 1 タスク・フォース着任: 小西和信 (北
海道大学), 酒井清彦 (横浜国立大
学), 谷岡修司 (岡山大学), 辻井喜美
代 (神戸大学), 平田明夫 (東京医科
歯科大学)
10. 1 ILL システム協力者会議
- ” 『学術情報センター紀要』編集委員会
10. 9~19 文部省科学研究費補助金・海外学
術調査 (フランス, 西ドイツ) 浅野助
教授
10. 13 昭和61年度前期セミナー閉講式 (修了
者6名)
10. 24~11. 8 文部省科学研究費補助金・海
外学術調査 (カナダ, アメリカ合衆
国) 根岸教授, 大山助手
10. 27 昭和61年度第2回総合目録データベー
ス実務研修閉講式 (修了者13名)
11. 1 タスク・フォース着任: 熊谷伸彦 (東
京工業大学)
11. 10 昭和61年度後期セミナー開講式 [研修
員: 奥田正義 (大阪大学), 栗田やよ
い (京都大学), 鳥場世浩 (立正大
学), 南里竜生 (関東学院大学), 村山
英治 (東洋大学), 渡邊俊彦 (大分大
学)]
11. 14~16 第4回目録システム講習会
11. 22 電気関係プレプリント情報データベー
ス懇談会
11. 27 総合目録委員会
11. 29 データベース委員会

東京大学文献情報センター日誌

昭和60年

3. 30～ 4. 7 センター長, 海外出張 (オーストラリア)
4. 1 井上如助教授: 教授に昇任
梶史郎事務長 (東京大学経理部主計課課長補佐), 佐藤隆雄共同利用掛長 (東京大学新聞研究所), 郷端清人事務官 (広島大学附属図書館) 着任, 大山敬三氏助手に採用
タスク・フォース着任: 村田康彦 (徳島大学), 木村祥子 (京都大学), 庄司三千子 (千葉大学)
永田治樹目録情報掛長: 東京大学附属図書館整理課目録主任に配置換
4. 16 タスク・フォース着任: 伊藤彰 (大阪大学), 斎藤豊 (弘前大学)
5. 1 タスク・フォース着任: 米澤誠 (東北大学)
5. 13 第7回運営委員会
5. 16 小山修美 (東京大学庶務部併任) 着任
5. 20 昭和60年度文献情報センター・前期セミナー開講式 [研修員: 牛崎進 (立教大学), 尾城孝一 (名古屋大学), 神田百合枝 (東京大学), 辻川輝男 (関西大学), 松野とも子 (北海道大学), 諸富秀人 (広島大学)]
5. 31 第13回文献情報センターシステム全体会議
6. 15 第4回文献情報センターシンポジウム (於: 東京工業大学)
6. 23～ 7. 6 センター長, 海外出張 (フランス)
6. 29 画像入力装置搬入
7. 1 タスク・フォース着任: 本郷清次郎 (琉球大学)
7. 11 臨時運営委員会
8. 1 安達淳講師: 助教授に昇任
8. 15～20 センター長, 海外出張 (アメリカ合衆国)
8. 20 『東京大学文献情報センター紀要』第1号刊行
9. 20 接続図書館説明会
" 『目録所在情報サービス利用の手引』 (第1版) 刊行
9. 24 図書館ネットワーク専門委員会
9. 25～27 第1回目録システム講習会
9. 30 タスク・フォース帰任: 庄司三千子 (千葉大学), 伊藤彰 (大阪大学), 斎藤豊 (弘前大学)
10. 9 目録情報専門委員会
10. 14 昭和60年度文献情報センター・前期セミナー閉講式 (修了者6名)
10. 16～18 第2回目録システム講習会
10. 20～26 センター長, 海外出張 (フランス)
10. 21 学術雑誌総合目録欧文編データ記入説明会「東京大学」(於: 東京大学)
10. 24 学術雑誌総合目録欧文編データ記入説明会「九州・沖縄地区」(於: 九州大学)
10. 25 学術雑誌総合目録欧文編データ記入説明会「東日本地区」(於: 東京大学)

- | | |
|---|--|
| <p>10. 30 学術雑誌総合目録欧文編データ記入説明会「北海道地区」(於：北海道大学), 「中部地区」(於：名古屋大学)</p> <p>10. 31 学術雑誌総合目録欧文編データ記入説明会「関西地区」(於：大阪大学)</p> <p>11. 1 タスク・フォース着任：鹿島玲子(千葉大学)</p> <p>11. 8 運営委員会・特別ワーキンググループ</p> <p>11. 11 昭和60年度文献情報センター・後期セミナー開講式〔研修員：市川美知代(清泉女子大学), 伊藤茂夫(福井大学), 柿沼澄男(図書館情報大学), 金子進(創価大学), 末木慎一(東京工業大学), 渡部哲夫(山口大学)</p> <p>11. 13 目録情報専門委員会ワーキンググループ</p> <p>11. 16 タスク・フォース着任：星和夫(新潟大学), 土師守(群馬大学), 大久保佳典(京都工芸繊維大学), 瓜生照久(九州大学)</p> <p>11. 17~12. 1 文部省科学研究費補助金・海外学術調査／(第1班) 安達助教授, 大山助手</p> <p>11. 20~22 第3回目録システム講習会</p> <p>11. 29 第5回文献情報センターシンポジウム(於：大阪大学・中之島講堂)</p> <p>11. 30~12. 16 文部省科学研究費補助金・海外学術調査／(第2班) 根岸助教授, 橋爪助手</p> <p>11. 30 HITAC L-450システム搬入</p> <p>12. 2 昭和60年度データベース担当者養成研修開講式〔研修員：石村恵子(埼玉大学), 井関泰夫(大阪大学), 京藤貫(富山大学), 笹山達成(南山大学),</p> | <p>塩谷玲子(弘前大学), 伊達恭子(京都大学), 中井えり子(名古屋大学), 仲清志(兵庫教育大学), 丹羽弘(名古屋工業大学), 林茂榮(京都大学), 松澤志津代(茨城大学)〕</p> <p>12. 10 システム運用専門委員会</p> <p>12. 16 目録情報専門委員会ワーキンググループ</p> <p>12. 31 タスク・フォース帰任：米澤誠(東北大学), 本郷清次郎(琉球大学)</p> <p>昭和61年</p> <p>1. 5~20 文部省科学研究費補助金・海外学術調査／(第3班) 井上教授, 宮澤助教授</p> <p>1. 18 通信制御処理装置(CCP)搬入・増設</p> <p>1. 22~24 第4回目録システム講習会</p> <p>1. 24 目録情報専門委員会ワーキンググループ</p> <p>1. 27 第8回運営委員会</p> <p>1. 29 図書館ネットワーク専門委員会</p> <p>2. 2~10 センター長, 海外出張(フランス, ベルギー)</p> <p>2. 10 目録情報専門委員会ワーキンググループ</p> <p>2. 12~14 第5回目録システム講習会</p> <p>2. 24 目録情報専門委員会ワーキンググループ</p> <p>2. 27 接続図書館説明会</p> <p>3. 1 昭和60年度データベース担当者養成研修閉講式(修了者11名)</p> <p>3. 7 システム運用専門委員会</p> <p>3. 9~23 文部省科学研究費補助金・海外学</p> |
|---|--|

- | | |
|--|--|
| <p>術調査／(第4班) 内藤助教授, 橋爪
助手</p> <p>3. 10 目録情報専門委員会</p> <p>3. 12~26 宮澤助教授, 海外出張(連合王
国)</p> <p>3. 15 仕様検討委員会</p> <p>3. 17 第9回運営委員会</p> <p>3. 19 昭和60年度後期セミナー閉講式(修了
者6名)</p> <p>3. 29 内藤助教授海外研修旅行(アメリカ合
衆国/合衆国政府招請) 出発(～5. 3)</p> | <p>3. 30 安達助教授文部省(短期) 在外研究員
(アメリカ合衆国) 出発(～9. 30)</p> <p>3. 31 タスク・フォース帰任:
60. 4. 1～; 木村祥子(京都大学),
60. 11. 1～; 鹿島玲子(千葉大学),
60. 11. 16～; 星和夫(新潟大学),
土師守(群馬大学),
大久保佳典(京都工芸織
維大学),
瓜生照久(九州大学),
小山修美(東京大学庶務部) 併任解除</p> |
|--|--|

学術情報センター日誌

昭和61年

- | | |
|--|--|
| <p>4. 5 学術情報センター創設</p> <p>” 学術情報センター職員着任</p> <p>所長(事務取扱) 猪瀬 博</p> <p>[管理部]</p> <p>管理部長 伊藤 公紘</p> <p>総務課長 梶 史郎</p> <p>システム・データベース課長
雨森 弘行</p> <p>[総務課]</p> <p>総務係長 渡辺 道夫</p> <p>会計係長 飯塚 正人</p> <p>共同利用係長 佐藤 隆雄</p> <p>研修・資料係長 渡邊 正昭</p> <p>会計係(併任) 中島 茂</p> <p>会計係(併任) 早川 敦夫</p> <p>共同利用係(併任) 飯田 信之</p> | <p>[システム・データベース課]</p> <p>システム管理係長 清水 二郎</p> <p>システム業務係長 中村 茂樹</p> <p>ネットワーク係長 渡邊 博</p> <p>企画係長 石井 啓豊</p> <p>データベース形成係長 牧村 正史</p> <p>目録情報係長 郷端 清人</p> <p>ネットワーク係 大日方一男</p> <p>データベース形成係(併任) 飯島 重美</p> <p>目録情報係 羽田 和久</p> <p>目録情報係 室橋 眞</p> <p>[研究開発部]</p> <p>研究開発部長 市川 惇信</p> <p>[学術情報研究系]</p> <p>研究主幹(教授) 井上 如</p> <p>教授 根岸 正光</p> <p>助教授 宮澤 彰</p> <p>助教授 安達 淳</p> <p>助手 大山 敬三</p> |
|--|--|

- [システム研究系]
- 教授 濱田 喬
 助教授 浅野正一郎
 助教授 橋爪 宏達
4. 7 海部文部大臣，センター来訪：開所セレモニー
 “ タスク・フォース着任：松本昌子（京都大学），大石博昭（国文学研究資料館），巖本康治（金沢大学），東海安興（長崎大学），有馬秀子（愛媛大学），青山功（大阪外国語大学），真中進（静岡大学），岡本正貴（愛知教育大学）
4. 23～25 第1回目録システム講習会
 4. 25～ 5. 24 宮澤助教授，海外出張（中国）
 5. 1 タスク・フォース着任：貝田辰雄（北海道大学），早野裕士（東京大学），堤守政（名古屋大学），坂賀力（大阪大学），河津秀行（九州大学）
 5. 3 内藤助教授，海外研修旅行帰国
 5. 9 OCLC 社長 Mr.Rowland C.W.Brown 氏来訪東京大学附属図書館において講演
 5. 11～16 浅野助教授，海外出張（米国）
 5. 16 タスク・フォース着任：藤岡節夫（京都大学）
 5. 19 昭和61年度前期セミナー開講式〔研修員：井上正則（同志社大学），牛島裕（近畿大学），河野建二（愛媛大学），二村英文（名古屋商科大学），松井好次（東北大学），吉田英明（鹿児島大学）〕
 5. 23 行政管理官視察
 5. 27 昭和61年度第1回総合目録データベース実務研修開講式〔研修員：阿部佳市（東北大学），工藤一郎（東京大学），久保田良子（東京大学），栗原道夫（新潟大学），斎藤泰則（慶應義塾大学），櫻井美由喜（兵庫教育大学），高橋宏美（東京大学），高橋實智子（京都大学），田中耕二，（京都大学），松藤典生（熊本大学），山下進（大阪大学）〕
5. 28～30 第2回目録システム講習会
 6. 10～20 所長，海外出張（スウェーデン，フランス）
 6. 14～20 研究開発部長，海外出張（スウェーデン，フランス）
 6. 24 国連大学シンドラール局長，センター来訪
 6. 27 第1回評議員会議
 6. 30 第1回運営協議員会議
 7. 2 衆議院文教委員会，センター視察
 7. 11 センター創立記念パーティ（於：如水会館）
 7. 21 昭和61年度第1回総合目録データベース実務研修開講式（修了者11名）
 7. 22 オックスフォード大学ティトラー氏，センター見学
 7. 23 アメリカ議会図書館吉村ヨシコ氏，センター見学
 8. 5 学術情報センター第1回シンポジウム（於：日本学術会議講堂）
 8. 18 ケンブリッジ大学図書館小山騰氏，センター見学
 8. 20 FIZ KARLSRUHE イェンシュケ氏，センター来訪
 西ドイツ各図書館（アウグスブルグ大

- 学図書館他) 職員, センター見学
8. 23 RLG マッコイ氏, センター見学
8. 26 韓国公州師範大学玄圭燮氏, センター
見学
8. 28 IFLA 参加者, センター見学
8. 29 ドイツェ・ビブリオテーク ケルム
・バルバラ氏センター見学
9. 2 昭和61年度第2回総合目録データベー
ス実務研修開講式 [研修員: 小幡砂智
子 (新潟大学), 恩知陽子 (山口大
学), 菊地房雄 (東北大学), 金城照子
(琉球大学), 熊谷伸彦 (東京工業大
学), 河野富行 (大阪大学), 齋藤穂
(弘前大学), 堤美智子 (京都大学),
友光健二 (東京学芸大学), 中野由紀
夫 (九州大学), 平尾行藏 (慶應義塾
大学), 藤巻美恵子 (東京大学), 藤原
誠 (名古屋大学)]
9. 3~5 第3回目録システム講習会
9. 9 ILL システム協力者会議
9. 30 安達助教授, 文部省 (短期) 在外研究
員帰国
10. 1 タスク・フォース着任: 小西和信 (北
海道大学), 酒井清彦 (横浜国立大
学), 谷岡修司 (岡山大学), 辻井喜美
代 (神戸大学), 平田明夫 (東京医科
歯科大学)
10. 1 ILL システム協力者会議
- ” 『学術情報センター紀要』編集委員会
10. 9~19 文部省科学研究費補助金・海外学
術調査 (フランス, 西ドイツ) 浅野助
教授
10. 13 昭和61年度前期セミナー閉講式 (修了
者6名)
10. 24~11. 8 文部省科学研究費補助金・海
外学術調査 (カナダ, アメリカ合衆
国) 根岸教授, 大山助手
10. 27 昭和61年度第2回総合目録データベー
ス実務研修閉講式 (修了者13名)
11. 1 タスク・フォース着任: 熊谷伸彦 (東
京工業大学)
11. 10 昭和61年度後期セミナー開講式 [研修
員: 奥田正義 (大阪大学), 栗田やよ
い (京都大学), 鳥場世浩 (立正大
学), 南里竜生 (関東学院大学), 村山
英治 (東洋大学), 渡邊俊彦 (大分大
学)]
11. 14~16 第4回目録システム講習会
11. 22 電気関係プレプリント情報データベー
ス懇談会
11. 27 総合目録委員会
11. 29 データベース委員会