

<国内長期派遣研修報告書>

東京大学文献情報センター・セミナーに参加して

運営課 主事
辻川輝男

はじめに

このセミナーは、東京大学（教育学部および総合図書館）が「図書館情報学に関する高度の知識を習得させる」という目的で、昭和47年から図書館情報学セミナーとして始めたものである。昭和53年、情報図書館学研究センターが発足後は同センターがその運営を所管した。その後、昭和58年に同センターが文献情報センター（以下文情セと省略する）として発展的に改組されるに対応して、単に図書館情報学の知識習得にとどまらず、学術情報システムの基盤システムの開発を目的とする文情セの設置目的に沿った内容に再編成の上で継承された研修プログラムが、今回私の参加したセミナーである。1年に前・後期の2期（昭和58年は1期）、これまでの13年間に25期、合計95名のセミナー生を送り出している。開催回数の割りに参加者数が少ないのは、通算15週間におよぶ長期の期間であり、しかも単に講義を聴いて終わりというのではなく、かなり密度の濃くて厳しい内容にするため、募集人員を毎回5名程度という小数主義で実施されているためである。また当初、国立大学の図書館職員を対象とした研修制度として位置付けられていたため、私立大学から参加したのは昭和57年からである。しかも每期1～2名であるため私で7人目となる。さて、セミナーの概歴を単々と綴ったが、主な内容は次のとおりである。

[研修内容]

期間 昭和60年 5月20日～ 7月20日、 9月 3日～10月14日（全15週）

場所 東京大学文献情報センター

内容 (1) 講義

- ア.基礎理論、基礎技法（情報科学基礎、電算機システム・プログラム言語、文献情報論、情報システム論）
- イ.図書館情報システム開発（情報検索システム論、データベース・システム、システム分析、学術情報システム論）
- ウ.図書館情報システム管理（経営管理論、図書館ネットワーク論、図書館情報システム管理論、図書館オートメーション）

(2) 演習（演習、特別講義、見学）

ア. 図書館機械化演習（システム分析演習、データ処理演習）

イ. 文献情報処理演習（オンライン目録、2次情報検索、所在情報検索）

(3) OJT（オン・ザ・ジョブ・トレーニング）

OJT方式でセンター内での各業務を実習し、図書館機械化およびネットワーク化に対応しうる能力を養う。

(4) 個別指導

研修員の履歴、専攻分野、年齢、職場での地位と業務、関心領域、セミナー終了後の計画など各人が持つ個別の条件を加味した個別指導を行なう。

(5) レポート作成

セミナー期間中に身につけたことの報告、各自の目標に対する成果を中心としたレポートを作成し、文献情報センター長に提出する。

さて、5月20日にセミナー開講式が行なわれた。その中で井上教授からまず次の3つが指示された。

- (1) 講義、演習、見学などセミナーで見聞したことを、個人的意見を入れな
いで、あるがままの事実だけをワープロで業務報告として毎週提出する
こと。前年度は5000字以上としたが、今回は2000字以上とする。
- (2) セミナー・レポートはセミナー終了の必須条件ではないので提出しなけ
ればならないことはない。ただ、これまでに提出しなかった人はない。
提出する予定の人にはできるだけ早く計画をたて、相談に来ること。
- (3) 我々としてはこのセミナーについては「捨て育ち」をモットーとしてい
る。研修プログラムの講義や演習を受けることはセミナー生としての最
低の義務である。それ以上についてはセミナー生の自由であり、我々か
ら「どうこうしなさい」と言うことはないし、また忙しくてそのような
指示ができるほど余裕もない。しかし必要とすることがあれば、可能な
限りの援助やアドバイスはするし、夜を徹してもお相手する。この15
週間をどう過ごすかは、あなたたちに任されている。

これから始まるというセミナー開講式での挨拶で、まず我々セミナー生は先生方から「捨てられて」しまったのである。「レポートの提出義務なし」の言葉でホッとしたと思いきや、それは井上先生一流の言い回しで「必ず提出しろ」という指示であった。また業務報告も「2000字位なら楽々」と思ったが、さてあるがままに1週間分を書こうとすると、とても2000字に収まる

ものではなく、結局4000から5000字になってしまうのである。「さすがに文情セ、ワープロで書かせるとは！」と感心していたら、そのワープロとは何と目録端末（文情セが特にオンラインによる目録作業のために日立に製作させた特別仕様の端末機）であった。

私自信は業務で富士通のホスト端末やエプソンのワープロ・パソコン、そして個人的には日電のパソコンでワープロなどを使っている。したがって多少の機種やソフトの違いには慣れていると思い込んでいたが、この日立の端末の感触は富士通のものとはかなり異和感があった。某先生の話では、「最初に使ったものがその人にとって一番使い易く感じるもので、大学などの教育機関にメーカーがよくホスト・コンピュータやパソコンなどを寄付したりする1つの理由は、将来の顧客確保といったこともある」とのことである。

富士通の図書館業務パッケージであるILIS(Integrated Library Information System)の目録端末としては、F9450Ⅱが採用されている。しかし、大きいな違いはILISではパソコンに目録業務のために必要な仕様変更を加えることなく、単にパソコンを採用している。これに対して日立のそれは、文情セが目録システムとして必要と認めた世界各国の言語における特殊な文字記号を、キーボードから入力できるようにソフトおよびハードの仕様変更を加えている。これは、非常に重要なポイントである。日本全国の大学図書館にすべてこの目録端末を納入したとしてもその台数は限られている。しかも数社のメーカーでその数を案分すれば、全く営業ベースには乗らないであろう。しかし文情セは、日本の学術情報システムの基盤となるその目録情報システムの必須条件として、現状ではJIS外字とされている前述の特殊文字を作成した。そしてそれらをオペレータすなわち目録担当者がキーボードから比較的容易に入力できるようにメーカーに対して設計変更を要求したのである。

2日目からは早速この目録端末を用いて、業務報告書の作成に取り掛かった。2日目からといっても9時から5時は研修プログラムがあるので、この作業は5時以降である。講義や演習内容の要約、見学レポートそれにOJTの作業内容などをまず下書きでまとめるのに1時間。午前午後とも講義であったり、内容が豊富であれば、まとめるだけに2時間はかかる。そしてワープロ。目録作業のための特別仕様をこらしたとはいえ、基本的には通常の端末機であるため、そのワープロ機能は専用機に及ぶべくもない。しかもセミナー生6人に対して、極く簡単なメーカー提供のマニュアルが2冊しかなく、いかにワープロを使い馴れるかが先決問題であった。当初は報告書の作成どころでなく、ワープロの訓練のため每晚9～10時頃まで残るらざるを得ないという状況であった。ホテルに戻ってから、「使い馴れた自分のパソコンを大阪から送ってもらおうか」と何度か真剣に考えたりもしたほどである。しかし、このときの経験は自分に

とって非常に役に立った。苦労したのは最初の2～3週間で、ワープロの所要時間は次第に短くなり、むしろ内容が豊富になり、相対的に下書きのまとめの方が長くなった。そして1か月頃には、セミナー生活にも馴れてきたことも手伝って余裕ができ、下書きなしでワープロに向かうようにまでなった。

次にこの週間業務報告書の中から、とくに興味を覚えたいくつかの講義および演習、そして見学レポートを紹介する。

I 講義および演習

<学術情報システム各論>

安達 淳 助教授（東京大学文献情報センター）

<内容>

文献情報センターの図書館システム構築にあたっては、

1. Sharaed Cataloging System を作ること
2. Union Catalog を作ること

という必ずしもマッチするものでない大きな2つの基本的課題があった。OCLC(Online Computer Library Center)のシステムは総合目録といわれるが、典拠システムがなかったために各参加図書館が各々のデータを登録し、重複データがふえて現実には総合目録とは言い難い状況に陥った。UTLAS(University of Toronto Library Automation System)はSharaed Cataloging Systemという特色をもって構築されているが、総合目録とは言っていない。これに対して文情セのシステムは「Sharaed Cataloging System」によって「Union Catalog」を構築するという非常に高い理想・理念をもとに開発が進められた。具体的には書誌と所蔵のデータをリンクづけるという手法がこの2つの課題を解決可能としたのである。

システム開発の観点からすると、

1. 画面志向の使いやすいインターフェース
2. ALA文字、外字の大幅導入

といった点が特色として挙げられる。

当初、学術情報システムの構想ではデータベース・サービス（オンライン検索）がメインであり、図書館システムは1部分でしかなかった。しかし大所高所からの判断で文情セが発足し、その主従が逆転しながらも開発が進められている状況である。

ネットワークの規約として、文情セのネットワーク・システムではN-1プロトコルを採用した。その主な理由としては、日立・富士通・日電などのメーカーが既にその技術を持っており、実際的にもDDX網を通じて全国大型計算機センターのネットワークとして運用されているという実績があったからである。しかし、DDXのパケット交換網では、時々発生する少量のデータ通信には向いているが、書誌データのように通常で1Kバイト、しかも一連作業として通信する目録作業では、必ずしも適したネットワークとはいえない。また専用線・DDXを問わず、例えば東京-大阪間で年間5万冊として、それに要する通信費用は月50~60万円（年600~700万円）となる。システムが理想に近いものであるとしても、経費の面から実用性あるものとは言えなくな

るのである。

これまで文情セでは大型計算機を持つ図書館との接続を進めてきたが、次の緊急の課題としては、オフコンないしは大きめのパソコンクラスを対象とした接続の技術開発を進め、従来の「ML」(Member Library)に類する図書館も「RC」(Resional Center)を経由しない方式で接続できるようにする。そしてさらに上述のような高額の通信費用を非常に低くする根本的な解決策を検討している。

<学術情報システム各論>

宮沢 彰 助教授(東京大学文献情報センター)

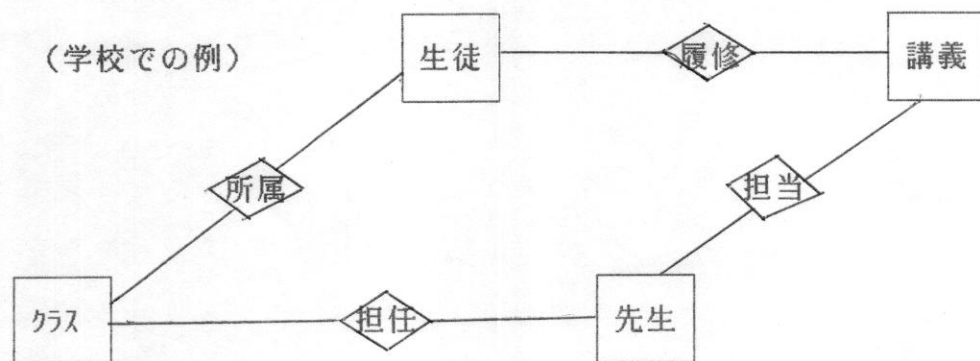
<内容>

オンライン目録演習を行なうために必要な前提知識として、文献情報センターの目録システムの基礎となっていることについて以下の通り説明があった。なお、宮沢先生は東京大学工学部出身で、文情セの前には国文学研究資料館でそのシステム作りに関与された。そして文情セの目録システム設計について中心的役割を果たされている。

1. データベースの形成

データベースの形成の仕方にはいろいろあるが、ここではデータベースの構造をまづ作り、その後で周囲を形成していく方法を説明する。

「実体とその間の関係を捉える。」



システム化する対象を上図のように実体と関係で概念を整理してからデータベースの形成に着手する。上のような概念整理は大抵の場合は成り立つが、そのままの状態で実現しようとしてもシステム自体の冗長性あるいは非現実性のため、できあがったシステムが使えないといったことがある。

システム化したい対象を実体と関係で概念を捉える。しかし、何が実体かという問題は認識の問題であって、人あるいは視点によっても異なるものである。たとえば、「A」は英語のAなのか、それともロシア語、ギリシャ語のAなのか、これをつきつめて議論していてもシステムはできないのであって仕様がなない。したがって、実体の把握は共通して認識できる程度に捉えることが大事である。

2. 文献情報センターの目録システム

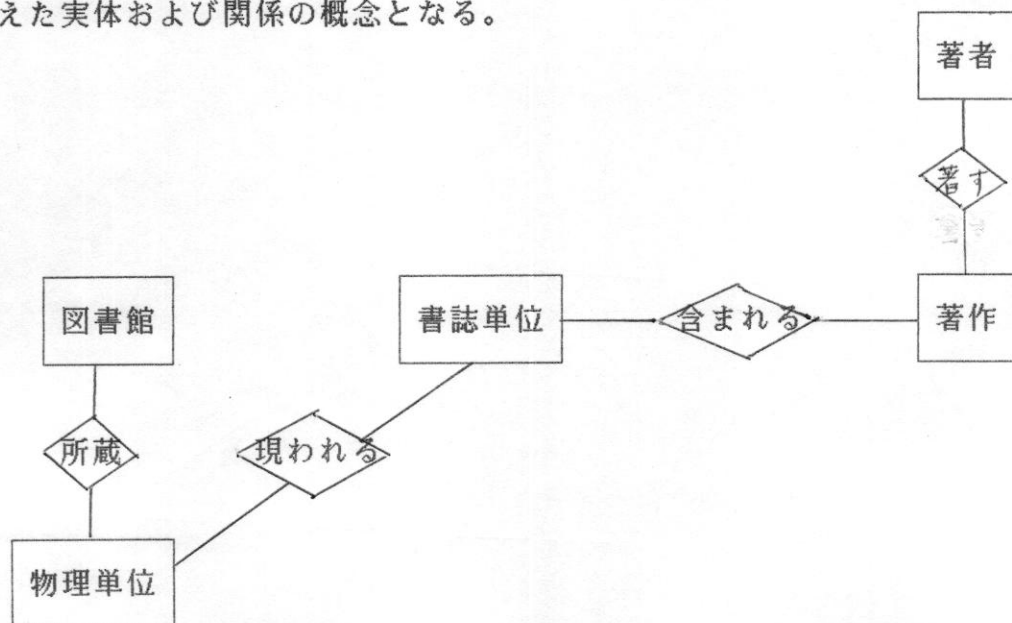
文情セでは、個々の図書館におけるハウス・キーピングを対象としていない。なぜなら計算機システムの能力の限界を超えるものであり、またそれは各図書館が行なうべきものと考えからである。

さて、1での例と同様に、目録ないしは書誌について整理をしてみると、

物理単位 …… 1冊の本それ自体。図書館での仕事の多くはこの物理単位の本を扱っている。

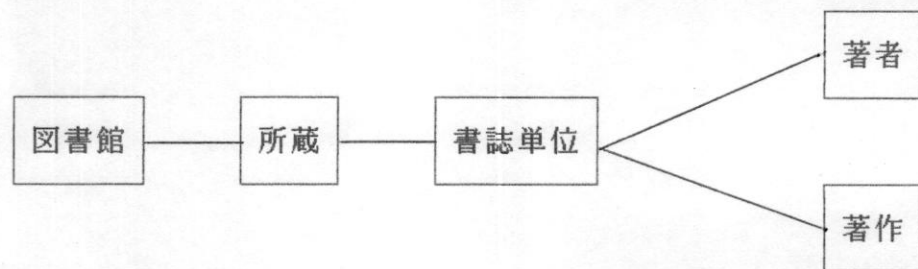
書誌単位 …… いわゆる同一本のことで、成立（出版）の過程からいって1バッチとして出来上がったもの。

図書館は物理単位としての本を所蔵している。複本は書誌単位が同じで物理単位は2つとして見る事ができる。書誌単位は本屋の都合でできることが多く、著作とは異なる。著作について細かく見ると詩集と詩の関係のように階層構造でみることもできる。しかし、システム形成の上でそれを議論しても進まない。というわけで、目録システムとして対象とするのは、次の図のように捉えた実体および関係の概念となる。

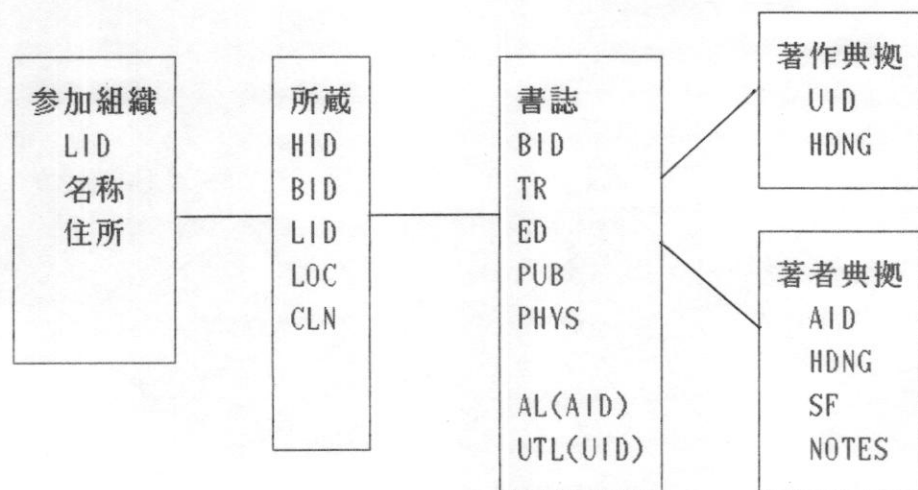


図書館のカード目録は物理単位毎に作られていると見ることができる。しかし、上の図から明らかなように書誌単位あるいは著者単位といった観点からも捉えることができるはずである。しかし、カードという物理的制約があるために、自由にその観点を換えて見ることはできない。また、現在の多くの図書館で作成されているカードの記載事項は、本の物理単位か書誌単位かがあまりはっきりしない面もある。

さて、上の概念図に現実的な解釈を加えると次のようになる。



図書館と物理単位、書誌単位と物理単位それぞれの関係は、図書館と書誌単位との関係すなわち所蔵という関係として包含して考えられる。著者と著作との関係は著者と書誌単位との関係として結び付ける方が現実的となる。これが文情セの目録システムが対象とする世界である。そしてファイルとして書き直すと次のようになる。



5つのファイルのレコードには、それぞれ固有のID（キー）が振られていて、それらによって各レコードが関係づけられている(Link)。

3. オンラインによる目録データの入力と出力

現在、文情セのシステムには参照MARCとして、LC-MARC, UK-MARC, JP-MARC いずれも全件をディスクファイルに格納してあり、毎週メンテナンスしている。ネットワーク参加館は、オンライン・システムでNCファイル(National Center File)や周囲の参照ファイルを検索し、

- (1) NCファイルに書誌データがすでにあれば、所蔵データだけを所蔵ファイルに書き込む。
- (2) NCファイルになく、参照ファイルにあればそれをNCファイルに登録した上で、所蔵データを入力する。
- (3) いずれにもない場合は、オリジナルとして書誌データを作成してNCファイルに登録の上、所蔵データを入力する

という流れになる。なお、検索時に検索対象ファイルを指示できるが、無指定の場合は、NCファイル・LCファイル・UKファイル(和図書の場合はNCファイルの次にJPファイル)を自動的に順次検索し、マッチする書誌データがあった段階で端末画面に表示される。また必要なら著者典拠ファイルや著作典拠ファイルとコントロールさせることもできる。目録作成の流れを細かく述べるとそれだけで1冊になるので、ここでは初めて利用して気付いた特徴を以下に記す。

- (1) UTLASではLINEモードでの目録作成であるが、文情セのシステムでは画面对応でできるため、格段に利用しやすい。画面設計をするとデータ項目が固定化されがちであるが、このシステムでは長い書名は尻切れにしないで改行して表示するなど、柔軟に可変的表示が可能となっている。
- (2) データにはEBCDICデータと漢字データと混在が許されている。また拡張EBCDICが採用されているので、英子文字と片仮名が自由に扱える。
- (3) 検索項目は、書名と責任表示、出版者、出版年など豊富である。検索ワードを複数入力すると自然とAND検索になる。検索結果を取り消して元に戻しても先に入力した検索ワードはそのまま残っているので、再度検索する場合も便利である。
- (4) LC-MARC だけでも200万件近いが、レスポンスは約2~3秒である。
- (5) 典拠コントロールも自動的にできないかという注文も起こり得るが、LC-MARCですら典拠ファイルとリンクされておらず、間違いが多い。また最大の理由として参照ファイルとNC典拠ファイルとで機械的にマッチングすることは極めて困難であるため、人間の判断を介在しなければで

きない。また和書の典拠について大きな問題がある。それはJP-MARCでは典拠ファイルが作成されていないことである。

- (6) 叢書やシリーズ物では、親書名あるいは個別書名いずれを採用するかということがカード目録の場合によくあるが、文情セの目録データベースでは、いずれも(両方とも)書誌ファイルに入力が可能であり、しかも親書誌と子書誌とをリンクづけることもできる。
- (7) モノグラフなどは図書・雑誌いずれの範中に入るか問題になることが多いが、図書と雑誌とでファイルが異なるので、どちらにも書誌データを入力する方が望ましいということになる。。
- (8) よく似た画面が使われているため、何の作業が行なわれていたのか分からなくなることがある。しかし、この目録システムは第1にオペレートに慣れた目録担当者を想定しており、OPAC(Online Public Access Catlog System)ということでは第2の目的である。
- (9) このNCファイルに登録したデータのMTによる書誌データの提供は、昭和60年の秋に可能となる。但し、当該図書館に対してのみである。
- (10) 文情セの目録システムでは、極端に言えば全世界の言語を扱う必要がある。従って、JIS第1、第2水準の漢字では全く不足する。とりあえず、欧文によく見られるウムラウトやアクセントなどを可能とするために、それら特殊文字の685個について外字登録されている。中国語やハングル文字などについては、これらの文字を入出力する端末装置などの開発が現在の技術では容易でないと判断されてたため、文字セットが作成されていない。
- (11) 文情セの目録システムと参加館の端末との間でのやりとり(通信)は、下図のようにして行なわれるが、基本的にはVS(Vertial Screen)として捉えることができる。VSはフィールド(単純に言えばデータ項目)の集まりである。これは参加館側の目録システムでは様々な機種の端末が考えられるからである。したがって、参加館側のUIPで扱われるデータは常に画面对応フォーマットになっている。ちなみに東工大ではそのデータを一度インターフェース・ファイルに落とし、後処理として自館ファイルをメンテナンスするようになっていっているらしい。ただし、自館システムに採り込んでしまった後にNCファイルがメンテナンスされたとしても、採り込み側では更新できないことになる。
- (12) 既に蓄積されている自館の目録データをNCファイルに入力する方法はいくつか考えられるが、著者典拠あるいは著作典拠ファイルとのリンク付けが極めて困難な作業になると思われる。

<電算機システム、プログラム言語>

長谷部 紀元 助教授（図書館情報大学）

（内容）

計算機の内部処理は2種類の表現単位で済む2進位取り記法が使用されている。この最小単位を1ビットと呼び、8ビットで256種類の情報単位を表現できる。英数文字の符号化基準はISO（国際基準機構）とIBMの2つのものが出回っている。ISO仕様のものは、ANSIのASCIIと呼ばれるもので、これに対してIBMのものはEBCDIC(Extended Binary Coded Decimal Internal Code)である。日本ではカタカナのコード体系をこのEBCDICに含めたものを採用しているが、英小文字のコードの一部をカタカナのコードに使用してしまっている。日立では、英小文字とカタカナが重ならないコード体系をEBCDICの拡張として採用したものがある。文情セのシステムはこれを採用している。

ところで、日本語（漢字）の表現は8ビットではできないため、16ビット（約64,000種）で表現する。漢字のコード体系はIBMの進出が遅かったこともあり、JIS規格で標準化されている。

計算機内部では8の情報を並列的に転送して計算処理しているが、通信の場合はコスト面からケーブル1本でデータを直列的に転送する方式が採られている。通信での最も基本的な原理に同期の方法がある。現在、次のような3種類がある。

（1）調歩同期式

通常は0の情報を送信しておき、データを送る時はその語の前にスタート・ビットをつける。また、後ろにはストップ・ビットをつけて送信する。したがって、1文字を送信するのに10ビットが必要となる。この方式では送信の速度が非常に遅くなる。(300-2400bps)

（2）同期式

通常は特別なコード（SYN）を送信しておき、受信側では8ビット毎に受信データをチェックする。SYNコードでないデータを受信すると有効なデータと認識して処理を行なう。1文字8ビットで済むために転送速度が早くなる。しかし高性能すなわち高価なMODEM（モデム）が必要となる。

（3）High Lebel Data Link

N-1 ネットなどで採用されている。文字だけでなくファックスなどの画像情報をも扱うことができるように配慮されている。専用回線使用で64Kbpsが可能である。

（4）同軸ケーブル

構内用として採用されている。高品質のケーブルが使用され極めて高速で送信される。(1Mbps)

<図書館情報システム管理論>

上田 修一 助教授(慶応義塾大学)

<内容>

この日の午前中は上田助教授の都合が悪いのが幸いして、慶応義塾大学三田情報センター(三田キャンパスにおける図書館で、昭和59年4月完成)の見学と機械化状況の説明を受けることになった。先生の講義は午後となった。

昭和60年秋の稼働を目指して閲覧貸出システムの開発を行なっている。今週中にも同システム専用として、FACOMのオフコンK-280が同センター5階のコンピュータ・ルームに導入される。

同センターでは以前から機械化のための図書館業務の分析や基本的事項について検討を積み重ねていた。昭和58年11月に閲覧貸出業務の機械化予算が明確になり、富士通(もしくはその下請け)のSE1名と図書館員4名とで開発に取り掛かった。三田の70万冊の図書にはすでにOCRラベルを貼付し終えている。利用者カードはラミネート加工のものだが、日吉キャンパスでは従来からバー・コードによるシステムが運用されているためカードの右上にはOCRコード(三田用)、左下にはバー・コード(日吉用)を印字している。三田は約120万冊の図書を所蔵しているが、書誌データが整っていないため関西大学の閲覧貸出システムと違って、トランザクション方式にせざるを得ない。

なお、メーカーのパッケージ・ソフトを採用しなかった理由として、

- (1) 伝統ある慶応の図書館では制度的にも業務的にも流れを大幅変更するには、利用者を含めて困難がある。
- (2) 貸出条件なども複雑で汎用性のあるパッケージ・ソフトには乗せにくい。
- (3) 従来から図書館員には機械化の知識、開発能力、その意欲などの蓄積があった

ことなどが挙げられていた。また、文情セとのネットワークについては、そのために必要なハードおよびソフトの経費が昭和60年度に予算化されており、昭和61年4月に接続を予定している。

[午後の講義より]

慶応では選書課に見計らい本が陳列されていて、教員がそれを見て購入希望を出したり、毎月選書委員会を開いて購入図書を選定している。また、予算上では各教員一律に年間15万円の特別図書費が割り当てられ、教員はそれぞれの研究に必要な図書資料を自由に購入できる。選書については各大学で様々な方法が試みられているが、結果的には数10名の教員が約7割程度の図書の収

集を決定し、蔵書構成に偏りが出来てしまう。特徴的なことは選定教員の得意とする分野の図書資料が充実するということである。

選定について図書館側から教員に対して必要な情報の提供がされないで、むしろ書店（その販売員）が営業戦略の一環としてその役割を担うという状況が各大学で見られるようになりつつある。特定教員の研究分野・関心領域についての情報を研究者との会話の中から蓄積し、必要と思われると図書資料の情報や新刊情報を提供している。

図書館員はかつて図書館を訪れる教員との会話の中から教員の研究動向などの情報を得ていたが、近年押し寄せてきた膨大な受け入れ図書の整理に追いつけられ、そのような余裕がなくなっている。また、その合理化を図るための機械化も若い人材を幾人も必要とし、利用者に対するサービスの充実について、より本質的な改善あるいはそのためのノウハウの蓄積があまりされていないように思える。

一方、書店の販売員は営業成績がかかっているため、販売ルートにのる学術情報の提供のための業務改善やシステム開発、そして意識改革は、図書館員よりはるかに早く進んでいくのではないかと。やがて「大学図書館は本当に学術情報の拠点か」と教員から問われる時が来るのではないかと。現在、図書館がすべきこととしては、教員に対するサービスの在り方を掘り下げて検討すること。そして、そのためにも教員との接触を蜜にし、教員の研究動向の把握を行なうべきであろう。（先生は決して図書館の機械化反対論者ではなく、最近の動向に対する良い意味での警告であった。なお、他に上田先生が関心をもって研究されている「酸性紙などの保存」について講義があったが、ここでは省略する）

<情報検索システム論>

松村 多美子 教授（図書館情報大学）

<内容>

1. 情報検索

DIALOG, JOIS などコンピュータを使っている情報検索は主題検索が中心である。従来カード目録で行なわれている分類・件名検索などよりはコンピュータの特性を生かした柔軟で多面的な検索が可能のため、よりきめ細かな検索ができる。

現在のところ、出版情報は1次情報の後に2次情報が生産され、流通する流れとなっているが、近い将来に電子出版が実用・普及すると索引・抄録は1次情報と同時、もしくはそれより前に検索手段として必要になり、現在の流通過程が大幅に再編成されるかもしれない。

情報検索システムでは、組織的、系統的に情報源を収集し、一定方針で整理したものを検索対象とする。つまり情報の蓄積と検索（INFORMATION STORAGE A

ND RETRIEVAL) という2つの側面を分析することができる。情報検索としての究極の目標は文献検索(ドキュメント・リトリバル)であるが、現在の情報検索は2次情報の検索(レファレンス・リトリバル)が中心である。

どの分野でどういう情報が必要か。図書か、雑誌か、報告か、レポートか、それぞれの分野の特性によってコミュニケーションの主要な媒体が異なる。また一過性でない情報であることが情報検索にとっての一つの条件ともなる。速報誌・予稿集などは研究情報として重要であるが、内容が変更する場合もあり、情報検索の対象となりにくい。

2. 主題分析について

1つの論文の中ではたくさんのトピックや概念が扱われている。これは我々に認識し得る文字という符号系を使って、語・句あるいは文で表現される。しかしこれら全部が主題を指し示すのではなく、それらの中の最も中心的に扱われている概念が主題を組み立てている。これらによってテーマが構成され索引語が指示される。さて問題は、多くの概念の中でどれが主題を構成しているのかという判断基準である。KWICやKWOKというような語句索引は予め不要語とされているものを抜いた語句によって作成された索引であるが、これでは文献で扱われているほとんどの語句・概念が索引語として採用されることになり、主題索引とは異なる。主題索引では主題に関する語句と、そうでない語句とを区別する。この区別は現在のところ人が判断せざるを得ない。ちなみに語句索引ならばコンピュータを用いて可能となっている。この主題決定基準は、1つのシステムにおいて一貫性をもたねばならない。でなければ検索タームに重大な影響を及ぼすからである。

3. 索引語の典拠リスト

分類体系・件名表のいずれでも可能であるが、最近の大きな情報検索システム(あるいはデータベース)の多くは、両方の機能をもったシソーラスが整備されている。これがなければ索引作成および検索に非常に不便を来す。

機械化、電算化、コンピュータ化、自動化、オートメーション化のように同一概念を意味する語が多く存在する。著者はいずれをも使用する自由をもっているので、シソーラスの作成にあたってはあらゆる語を網羅する必要がある。

情報検索システムの1文献からは、文献番号・書誌データ・索引語群・抄録がファイル化される。しかし、これらすべての語について索引づけることは不経済であり、また不必要でもある。主として索引語群、そして書誌データの書名・著者について逆ファイルが作成される。

シソーラス作成では、イメージネーションを働かせて1つの概念について多

く自然語・索引語を洗い出す作業を行なう。学術文献については当然のことながら専門用語が多くなる。したがってその分野についての主題知識も必要となる。外国の文献情報検索なら、さらに外国語での索引語のイメージネーションが必要となる。これは検索においても同様のことがいえる。

索引用語リストとしては、索引・検索語の50音順リストと体系化リストがある。前者は一覧することで、その検索システムで使われている検索語あるいは特定の検索語に類似した検索語を探すことが容易にできる。後者のリストは数多くの索引語について概念的な上下関係や関連を階層構造で表わしたものである。大きなシステム（データベース）では1-2万語にもなる。これらの索引語をすべて関係づける作業は極めて困難である。NDCでの階層化はすべてを10個の類に分けて、その中で関係づけていると見ることができる。しかし主題索引における索引語の概念の階層化はその類を超えて様々に関係づけられる。これは単純な階層関係では捉えることが不可能なためだが、50音リストにおいては、「をも見よ」参照をつけることで解決することができる。

参考例 胃かいよう、胃炎（急性・慢性）、胃アトニー、神経性胃炎
ストレス、胃酸、胃酸過多、喫煙、

システムによっては1つの索引語を入力すれば関連語が表示されるものもある。しかし、これは永い検索経験の蓄積の上にシステム化されるものである。いずれの検索システムでもサーチャーの役割が大きい、特に上述のような典拠リストが整備されていないシステムでは、サーチ戦略がすべてサーチャーの力量に委ねられることになる。

<学術情報システム論>

開原 成充 教授（東京大学医学部附属病院）

<内容>

同病院中央医療情報部ではBRS/SAUNDERSの情報検索システムであるCOLLEAGUEを使っている。講義は同情報部の1室で行なわれた。

1. 医学関係の学術情報システムとしては、次のような特徴が挙げられる。

- (1) 化学に次いで情報量が多く、また歴史も古い。
- (2) 医療面において極めて重要性をもつ。新しい医療情報や薬の副作用などを知らずに行なっていると、患者の生命に関わる場合もある。これまでも数多くの悲劇や裁判の争いがあり、今後も起こり得る危険性はある。
- (3) 文献情報以外にも重要な学術情報が多くある。医療技術では人から人に

伝わる技術情報も重要である。

- (4) 実用性と学術性の混然化。世界的に重要な医学雑誌である「胃と腸」のように商業雑誌と言えども学術情報として極めて重要で、学術雑誌との区別をつけがたいものも多い。
- (5) X線写真や統計データ、特にコンピュータを利用した医療技術・医学研究では画像情報の占める割合が大きい。
- (6) 単にキーワードを入力して検索するだけでなく、更に推論をして関連情報を自動的に検索してくれるような知識ベースの開発もエンド・ユーザー側から出てきつつある。

2. 医学における学術情報とはなにか …… 省略

3. 医学文献情報 …… 省略

4. これからの検索システムの特徴

- (1) フル・テキスト・データベースの開発・実用化がいよいよ進んできている。アメリカでは多くの有名医学雑誌がフル・テキストとしてデータベース化され、サービスを開始している。唯一の弱点は、画像・図形情報を技術的に載せられないということだけである。
- (2) オペレータとしてエンド・ユーザを志向している。端末（パソコン）の低価格化と検索方法の統一・標準化ソフト（Gate-Way System）の開発が進み、研究者自身が機器を購入して容易に検索ができるようになってきた。システム販売側は、図書志向型の図書館よりも市場が極めて豊富な研究者を販売相手としつつある。
- (3) 医学情報を扱う立場としては、他で作成されたデータベースを検索することもさることながら、自分たちのデータベースを構築していく必要性も認識されつつある。これまで研究者が個人的に作成してきたものもあるが、索引語・抄録の作成といった専門知識・技術を要する領域が、研究・開発をバック・アップする重要な役割を果たす。情報を扱う担当者は検索のみならず、作成の領域も自分たちの職務領域であるという認識が必要となろう。

<図書館ネットワーク論>

浅野 正一郎 助教授（東京大学工学部境界領域研究施設）

<内容>

1. アナログ通信とデジタル通信

通信の始まりはモールス信号であったが、これはいわゆるデジタル通信で

あった。その後、ベルが音声をそのまま通信する目的で電話を発明した。ここでは通信はアナログとなり、以降の通信はアナログが中心となって発展した。ところが、コンピュータのネットワーク開発が盛んになるにつれて、その特徴に見合ったデジタル通信が脚光を浴びるようになってきた。その理由の一つに伝送路として、従来のケーブルやマイクロ・ウェーブに加えて光ファイバーが開発されたことが挙げられる。これはディジタル伝送が基本となっており、信頼性が高く、しかも装置のサイズはより小さくできる。また伝送容量も極めて多く、その分だけ経済的でもあり、今後の通信情報量の急増に対処することができる。従来の電話やファクシミリ・TVもディジタルの世界に取り込まれつつある。最終的にディジタルの世界でこれらの通信を総合的に扱う体系を作ることが通信の世界の大きな目標となっている。

2. 電話・データ通信・INS(Integrated Network System)

電話は誰と通信するかが最も大きい関心事で、多少の雑音が入っても余り気にしなくてよい。周波数 300HZから 3,400HZが伝送範囲であるから音楽などは最初から期待していない。これに対してデータ通信では品質として極めて高い正確さが要求される。さらに性能として、大量のアクセスおよび一定速度のレスポンスの保証が必要である。INSに至っては、さらにその装置属性を選ばないという汎用性まで要求されるようになる。

3. 通信速度について

現在の通信の中で約9割が電話通信であり、1割がデータ通信、そして0.数パーセントがデータ通信としての画像通信である。通信速度としてはいろいろなスピードがある。

50b/s, 100b/s, 200b/s, 300b/s, ... 1200b/s, 2400b/s, 4800b/s, 9600b/s ...

今日では 50b/s, 100b/s はほとんど使われていない。カード・パンチ機とコンピュータの間では1200~2400b/s、ライン・プリンタとコンピュータ間では4800~9600b/s、コンピュータ間では 19.2Kb/s や48Kb/sなどが使われている。

画像情報では極めて大量の情報が送られるが、現在のところファクシミリは 2400b/s, 4800b/s, 9600b/sが使用されており、A4版のワープロ原稿で約10秒かかる。キャプテン通信では 4800b/sである。動く画像情報であるTV会議では伝送情報が極めて多く、しかも速度が要求されるため 1.5Mb/sが使われる。INSでは標準として 64Kb/s が使用され、電話やFAXも可能となっている。364Kb/sでは高品質の音声通信、解像度の高いFAX、そして1.5Mb/sではテ

レビ会議ができる。

4. データ通信のための回線

今日、最も多く使われているデータ通信回線は電話回線である。これには公衆電話回線網と構内回線がある。アナログ情報とディジタル情報の相互互換を行なう変復調装置（モデム）を使用すると、その間のデータ通信はアナログであるが、利用する立場からみるとディジタルとして考えることができる。この方式の利点としては手軽であり、近距離で利用頻度が少ない場合は経済的である。しかし、低速度に限られ品質的には余り良くない。速度は 300～1200bps で双方向同時通信ができる（全二重）。2400～4800bps も可能だが瞬時方向通信となる（半二重）。

専用回線には帯域貸し（帯域品目）と速度貸し（符号品目）とがある。前者には目的による制限はなく、データ通信でも電話でも良い。ただし伝送データの品質保証はない。後者はデータ通信用となっていて品質が保証されている。低速よりも高速の方が割安となっているので、高速のケーブル1本で多重化装置を使い、複数の端末をつなげる方が経済的となる。料金としては、公衆回線で1日当たり3時間程度の使用料金が専用回線の価格と引き合う。したがって通信量が多い場合は専用回線の方が安上がりとなる。

DDX網のパケット交換方式は伝送するデータを一定の長さに分割し、それぞれのデータに発信者や着信者の情報を荷札として付けたもの（1パケット）が小包みのようにして伝送される。特徴的なことは、従来の伝送方式は使用時間量に対して課金されるが、この方式ではデータ量について課金されることである。1パケットに付き近距離で0.4円、中距離で0.5円、遠距離で0.6円となる。1パケットの最大データ量は128文字までとなっている。この方式ではデータ量が多くなるほど料金が割高になる。そのかわりデータの品質は保証される。現在、文情セのネットワークは、この方式による大学間ネットワークを通じて実現しようとしている。しかし全国の大学が参加したとして通信料金が1年に約10億円にも及ぶ。そこで、NTTから高速ディジタル伝送回線を借りて、日本の北から南まで結ぶ計画すなわち学術情報VANが検討されている。これによると費用が約3億円程度に収まると想定される。

5. 通信制度について

日本では従来、データ通信のための専用回線の使い方について制限を設けてきた。1つは借りた人だけが使用すること。2つ目は任意のメッセージ（データ）を流してはいけない、というものである。昭和46年に若干緩和されて回線の共同利用が可能になった。つまり、A社、B社が共同して1本の回線を借

りて、お互いのデータをやりとりをすることが可能となったのである（全銀システム）。しかし、依然として他人の使用やメッセージ通信は制限されていた。昭和57年になってこの制限が漸くはずされたが、通信されるデータは情報処理を目的とするものという条件つきであった。これが昭和60年4月になって自由化されたのである。

アメリカでは、1970年頃にARPA(Advanced Research Provanced Agency: 国防省の1機関)がスポンサーとなって、大学や研究所に対してネットワーク化の研究・開発を進めた。このARPAネットワークは現在も稼働しており、数多くある機能の中で最もよく使われているのは新聞のフル・テキスト・サービスとメール・ボックスである。

6. コンピュータ間の通信

TELEXなどのような通信では1対1で符号のやりとりが行なわれる。文字符号も規準化されており、国際的にはIA5単位方式、国内的にはIA6単位方式が採用されている。この通信の欠点は伝送誤りについての回復機能がないことで、受信者が判断をしなければならない。これに対して銀行のオンライン・システムでは伝送誤りの回復機能をもっている。さらに発展した銀行間を結ぶコンピュータ・ネットワーク・システムでは、表面上は1対1の通信のようであるが、技術的には通信相手を選択している1対Nの通信である。

端末とコンピュータとの通信では端末側の人間がデータを入力し、また出力データを見て判断するが、コンピュータ間では双方の応用プログラム同士がデータのやりとりを行なう。したがって、装置とプログラムを指定した通信となる。プログラムが扱えるデータを通信するのであるから、データの通信のための多種多様な標準化・統一化が重要になってくる。

(1) 情報の表現

文字符号には例えば次のようなものがある。

ISO規準.....ASCII (日電など)

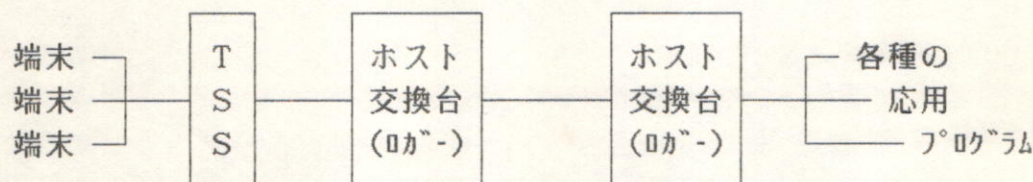
IBMEBCDIC (富士通、日立など)

この2つにもさらにバラエティがあり、それらの間で通信のための変換ユーティリティを作成するとなると、極めて多くのソフトが必要になる。そこで1種類の情報符号体系を作ると、データ通信ではそれぞれこの体系に合わせれば良いことになる。同様に標準化されたデータ構造・標準図形表現・標準画面表現というように、ローカル・マシーンでは各種の体系でデータが扱われていても、通信では特定の体系に変換してやりとりをするという技術を仮想化技術という。このような考え方はARPAネットワークにも採用されている。同ネットワークは効率の良い通信網、プログラム間通信、仮想化技術を研究・開発し

た。日本でも1974年頃から文部省が大型計算機センター間でのネットワーク開発を進めた。この時の目的はまず開発することにあつたため、その応用や需要については全くの不明であった。システムを運用する人の確保も保証されないこともあったが、それが幸いしてシステムの信頼性、拡張性に富んだネットワークが出来上がった。これが今日のN-1ネットワークである。このネットワークはDDX網を利用しているので、各センター（ユーザー）はそれに加入すれば済むことになる。但し、DDX網で結ばれているのはホスト・コンピュータだけであり、ホスト・マシンから各エンド・ユーザーの端末まではホスト・システムで内部のルート制御のためのプログラムが必要になる。

(2) プログラム間通信

プログラムとプログラムを結びつけるには、基本的にそれぞれのマシンをTSSのレベルで結合する必要がある。



このネットワークの典型的な応用として

ア.文字符号による通信（TSS）・・・標準文字符号（NVTコード）

イ.ジョブ転送標準ジョブ

ウ.ファイル転送標準ファイル

ファイル内のデータのみでなく、ファイルの属性をも転送する必要がある。

N-1ネットは1976年に初めて東京大学・京都大学間で稼働した。現在では、公立・私立を含めて約40以上が結ばれている。課題として、利用申請手続きの簡略化・課金制の整備がある。N-1ネットはメーカーの標準製品として作成されている。富士通ではOS IV/F4.F2.X8、日立ではVOS3、日電ではACOS6,ACOS4 といったOSに組み込まれて機能する。昭和54年から発足した学術情報センター協力者会議では、当面（昭和63年頃まで？）はN-1ネットを使って開発を進めること、図書館システムから開発するという事で、昭和58年に文献情報センターが発足した。

<学術情報システム各論>

橋爪 宏達 助手（東京大学文献情報センター）

<内容>

1. 通信路に何を選ぶか。

その判断基準としては、

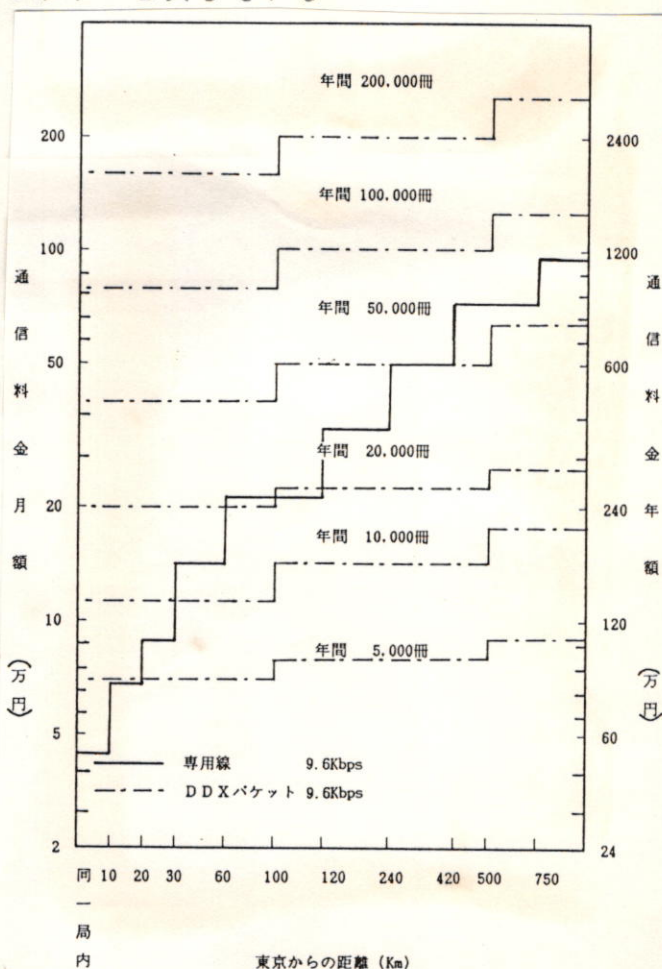
- ・ 伝送する情報の性質（音声・データ・画像＜動画、静止画＞）
- ・ 伝送する情報の速度（BIT毎秒<BPS>）
- ・ 伝送コスト

文情セのシステムではコンピュータの画面情報すなわちデータが対象になる。これを伝送する手段としてはNTTがサービスしている

- ・ 符号品目専用線
- ・ DDXパケット交換網

という2つがある。

この2つの通信手段による通信コストの比較が右のグラフである。これは昭和59年度の始めの頃、具体的なシステム開発前の段階での試算であったが、通信データ量（約1Kバイト）の誤差はあまりなく、現在もそのまま適用できる。専用線では使用量（通信量）に関係なく一定の費用が必要になる。また、料金の遠近格差が非常に高い。これに対してDDXは従量制で通信したデータ量に応じて料金が計算される。遠近格差は専用線に比べて非常に低い。東京からの距離と年間通信データ量（受入れ図書数）に応じて通信方法を選択すればよい。



2. NC(National Center) - RC(Resional Center) - ML(Member Library)問題

NCはともかくRC-MLとは何か。様々な解釈があり、整理すると次の3つが比較的明確である。

- ・ 図書館における年間の計算機予算額の違い
- ・ 情報蓄積の役割の違い
- ・ 電線の張り方の問題

ネットワークの観点からすると、配付資料の通りAからEまでの5つの接続

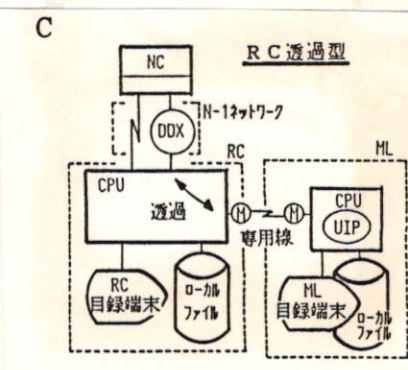
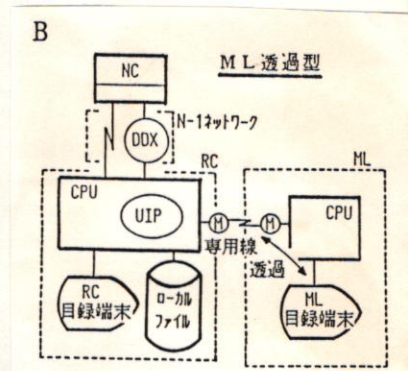
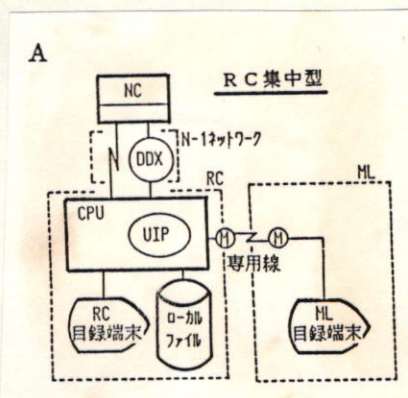
形態が考えられる。この5つでMLに相当する図書館の接続形態がどうなるかを見ると

A：RCのホスト・コンピュータに接続した目録端末を距離的に延長したもの。NCとしてはRCだけを接続対象として考えればよいことになるが、RCにとっては自己の計算機資源をMLに流用する必要がある。またMLにとっては使用機種がRCと同一機種に限定されるとともに運用においてもRCの運用制限に拘束される。しかし、新規開発は不要で容易に導入が可能となる。

B：RCのホスト・コンピュータとMLのホスト・コンピュータを接続し、UIP (User Interface Program) をローカル分についてもRC側で用意する。Aタイプと同様でRCはMLに資源を流用する必要がある。MLでは計算機は透過型となるため比較的小型で機能が果たせ、しかもローカル業務が自己のCPUで行なえる。しかしMLでのオンライン業務とローカル業務との結合が技術的に困難である。

C：Bタイプと同様の接続で、UIPをMLにもつ方法である。RCのホストは透過型になる。そのためA・Bに比べRCがMLに流用すべき資源は少ない。しかし、ML側ではUIP搭載のためある程度以上の計算機能力が必要で、機種もAタイプと同様にRCの計算機メーカーに限定される。また現在、MLに相当する図書館に導入されている規模の計算機では、能力が小さすぎるためメーカーによってはUIP搭載が困難な場合もある。

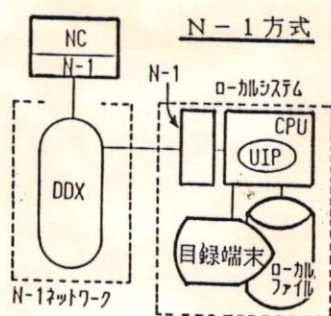
D：RCを通さずにML独自にN-1ネット機能、UIP，ローカルシステムを持つ。NCとしては、RCとMLの差異がなくなり、しかも通信機能をN-1ネットに統一



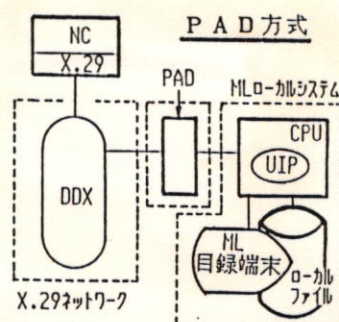
できる。RCにとってMLは全くの通信ネットワークの観点から関係がなくなる。しかし、上述のように現在MLに導入している程度の計算機でRCと同様にシステム開発が必要となり、メーカー側では当面对応が難しい。この接続ではNC-RC-MLという地域ネットワーク構想は通信ネットワークの観点ではあまり意味を持たないことになる

E: Dタイプの方式でN-1ネットの代わりにPAD接続による方法。N-1ネットとは別にPAD接続手順をNCが新規にサポートする。N-1方式に比べて通信手順が容易で、また安価で接続ができる。現在、南山大学がPAD方式でIBM5550(パソコン)での接続を開発中である。

D



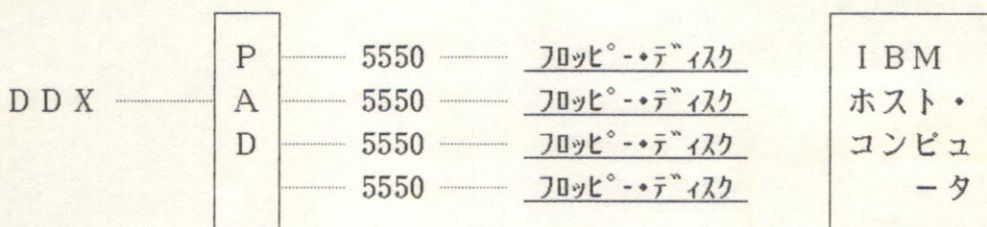
E



3. PADの紹介

現在、外国製品しかないが、その1つとして日本ダイナテック社が販売しているMULTI-PAD X25という機械(80-150万円)がある。これをNTTからの通信回線と自己の計算機に接続する。これ1台に4台までの端末が接続可能である。

<南山大学の方式>



ホスト・コンピュータを通じて接続をするN-1ネット方式の場合は、ホスト・コンピュータの主記憶をかなり占有することになる。その点、このPADを使用した方式(VTSS方式)では、通信に関してはホスト・コンピュータとは無関係であるため、大きなコンピュータを持たない小規模な図書館でも接続が可能となる。

4. 将来の図書館ネットワーク

文献情報センターのシステムに対する利用者の究極の要求はドキュメント・デリバリにある。そういう観点からすると現状の通信ネットワーク・システムは、

(1) 通信コストが高すぎる

(2) 大量の情報を伝送する原文書通信には利用できない。

という大きい課題がある。

これまで電話は電話線、テレビは電波としてそれぞれの特色に合わせて通信媒体が開発されてきたが、最近の技術によると音声・画像・データいずれもデジタル信号として伝送すれば、ネットは1つでまかなえることに注目されている。利点としては、

(1) ネットの統合化が可能となり、効果対費用でコスト安となる。

(2) 多種のデータの同時送信が可能となり複合サービス(マルチ・メディア・通信)ができる。

国際規準としてはISDN(Integrated Service Digital Network:サービス統合デジタル網)として進められている。日本では、NTTがINSとして10数年を目指してシステム開発・実験を行なっている。将来は電話の速度でデータ通信が可能となる。この先駆けとしてNTTでは、高速デジタル伝送サービスを昭和59年11月から開始している。このサービスは従来に比べて格安の費用で通信が可能となるものである。

従来の符号品目専用線

速度 200BPS --- 48KBPS

値段 9600BPS では、東京-大阪 46万円/月

高速デジタル伝送サービス

速度 64KBPS --- 6MBPS

値段 64KBPSでは、東京-大阪 110万円/月

384KBPSでは、" 210万円/月

6MBPSでは、" 1100万円/月

1本の高速デジタル伝送サービスを多数の利用者で共同利用すれば、従来のものに比べて半分ないし4分の1程度の費用で済むことになる。また、画像通信の可能性も出て来る。静止画像の通信であるFAX(GIII)では、4~8DOT/MMでA4の1枚が約30秒から1分かかる。ISDN用のFAX(GIV)では、12~18DOT/MMで1秒から5秒程度になる見込みである。複写機とほとんど変わらなくなる。現在、文情セではこの高速デジタル伝送サービスを利用して、全国ネットの通信経費を格安に落とすことが検討されている。また、この通信では電話・ファックスなどあらゆるデジタル通信が利用可能であるため、図書館システムのネットワークの域に止らず、大学間の情報通信ネット

ワークへの展開も可能となる。

5. 次の世代の図書館システム

現在、図書館ではその規模に応じて大型コンピュータからパソコンまで、様々に導入している。これまで文情セとの接続を進めて来た図書館はいずれも大型コンピュータを持ち、その上に図書館システムを構築している。しかし、将来にわたってそのようなシステムが最良かどうか、疑わしい。処理速度だけを比べると大型コンピュータに勝るものはない。しかし費用対効果で比べると、はるかにオフコンないしは、スーパー・ミニコンといわれる規模の方が優位に立つのである。そこで、現在われわれはDEC社のスーパー・ミニコンを導入して、UNIXのもとで機能する図書館ローカルシステムの実験的モデルの模索を行なっている。

II 見学記録

<国文学研究資料館>

見学日：昭和60年5月23日

概説：安永尚志 助教授

国文学資料館の設置目的は2つある。第1は「全国に散在する国文学に関する文献資料を調査、研究し、収集、整理、保存して、これを広く研究者一般の利用に供し、国文学研究の進展に寄与すること」、第2に「研究情報の収集、整理、保存、閲覧」で、雑誌・紀要などに発表される論文や単行本などの研究情報を1年単位で「国文学年鑑」として刊行している。

同資料館の対象とする文献資料は江戸時代までの写本版本で、その多くはマイクロフィルムとして収集・整理している。現在、調査済みが約12万点、収集が約10万点、整理済みが約7万点、毎年およそ1万点位増えている。

昭和47年に共同利用施設として設立され、当時の基本構想の段階からコンピュータ利用の4つのポイントが考えられていた。

- (1) 資料（伝本）の検索
- (2) 文献（論文など）の検索
- (3) 主要語意の検索
- (4) 定本の作成（校定本分）

- (1) については、マイクロ資料目録をコンピュータ処理により毎年作成しており、現在で7冊になる。書誌データの作成は1日約18件程度。パンチ外注、そしてデータのチェックや典拠コントロールなど各種の処理過程を通してデータベースに入力される。現在、調査・整理がたいへん遅れているため、最終的に利用に供し得るのは収集後約5年位かかる見通しである。また和古書約5500冊分の目録も作成している。その特徴としては、国書総目録を典拠にして書名あるいは著者名をデータ化して統一標目、著者名典拠をコントロールしていることである。いずれもデータベース化され、日立の検索システムであるORIONによって利用できるようになっている。
- (2) については、昭和16年から37年までの約4万件について、国文学文献目録の作成がコンピュータ処理で行なわれている。最近10年分の論文データもDB化し、国文学独自の文献探索法の実現方式などの評価を加え、実用化している。
- (3) は古典テキストそのもののデータベース化で、現在万葉集など15本がDB化され、古典テキスト検索システムとして稼働している。
- (4) の課題は、将来計画として原文書そのものという性格をもつ0次情報の処

理システムが考えられている。

同資料館が抱える特徴的な問題点は、漢字処理すなわち文字セットである。JIS規格の文字約7000字の他に、約2000字の外字が作成・登録していること。それでも、毎年整理担当の掛りから外字の要求が600字、その登録の必要性などを調査・検討して3分の1にし、さらに予算の関係で削って約100字程度を登録している。外字の作成には専門のデザイナーが必要で、しかも1字当たり数万円の費用がかかる。文字の確認や追加登録、二重登録の防止などのため、文字セット管理システムも作成している。

<東京工業大学附属図書館のシステム>

見学日：昭和60年6月6日

概説：内藤英雄 専門員

1. 東工大附属図書館の概要

(同大学附属図書館)

東京工業大学は学部学生が3000人、院生が2000人、教官は600人(半数は助手)、事務は700人で、学生数の割りにスタッフの数が多く比較的ぜいたくな構成となっている。図書館本館は約7500平米、座席数600席、1日の入館者は約2000人、貸出冊数は300冊。分館は約3000平米で、座席数130席、1日の入館者は約200人、貸出冊数は20冊となっている。所蔵総冊数は約50万冊、その内の約半分は製本雑誌である。

1977年に理工学系外国雑誌のセンター館として文部省から指定された。この指定制度は雑誌に関するいわゆる拠点図書館制度で、次の通りである。

- (1) 理工学系 …… 東工大、
- (2) 医学・生物学 …… 大阪大学(中ノ島図書館)、九州大学医学部図書館、東北大学医学部図書館、
- (3) 農学系 …… 東京大学農学部図書館、鹿児島大学図書館本館、
- (4) 社会科学系 …… 神戸大学、一橋大学

文部省予算での雑誌購入予算としては、第1種(学内集中用)、第2種(地域内補充)、第3種(全国的補充)がある。第1種は行政改革や図書館の機械化予算の影響を受けて予算削減の憂き目に合っている。第2種はあまり意味がなく、立ち消えになってしまっている。第3種は上述のセンター館に予算立てされ、約2億6千万円(内1億2500万円が東工大分)となっている。つまりこの制度は、特定の大学図書館に対して雑誌購入の予算を集中し、それぞれの分野について全国的センター館として機能させるものである。

1977年、東工大での洋雑誌受け入れタイトル数は約2000であったが、この指定により現在1万タイトルを超えるに至った。

2. 電算システムについて

1973年に沖電気のミニコンが導入され、雑誌受け入れシステムを開発して運用していた。しかし拠点図書館の指定以降の受け入れ雑誌が急増し、システム能力の限界に達してしまった。このため昭和55年度から毎年、新システムの予算要求を出し続けたが、昭和58年になって漸くそれが認められた。その間、九州大学、名古屋大学、大阪大学に地域センター(RC)としての機械化予算が計上され、それぞれ独自にシステム開発を行なったが、東工大の予算化時点では、既に文情セの設置が明確になっていたため、目録システムは当初から同センターとの接続を前提とする考え方によってシステム設計が進められた。

マシンは日立のM-260D、DBMS(Data Base Management System)は、MODEL 204を採用している。このDBMSにはシステム開発のためのユーザー言語が提供されているが、広島大学が以前からこれを補強する約60程の便利なコマンド・ユーティリティ(HDM)を開発しており、これに移植して利用している。

このユーザー言語は素人にも容易に理解ができ、東工大では雑誌の受け入れシステムを除いてすべて図書館員が開発した。雑誌のシステムはレスポンスが要求されるのでメーカーに任せた。しかし、メーカー側への業務説明や連絡調整などに時間がかかり、またSEあるいはその下請けのソフト・ハウスとの意志疎通があったりして、当方の思うようにはなかなか進まなかった。しかも最初に出来上がったものは使い物にならなかった。さらに同じユーザー言語を使用しているものの、専門家のプログラムは素人の図書館員には解説が困難で、後のシステム保守が問題となっている。

図書館員でユーザー言語を用いてオペレートできる者は約10名おり、各係に1名以上配置されている。したがって、各々のシステムは各掛かりで保守・改善ができるようになっている。ちなみに我々セミナー生も見学後、簡単にMODEL 204の説明やユーザー言語の説明を受けただけで、「雑誌番号と受け入れ日を入力すれば、その受け入れ雑誌の情報を端末画面に表示する」というプログラムを、約1時間少々で作成した。

3. 文情セとの接続について

昭和59年11月に接続し、試験的に使用している。専用回線を用いており、通信費用は年間約80万円と見積もっている。

文情セのオンライン目録システムを使って目録作成を行ない、オペレータが特別のコマンド(LSAVEコマンド)を発すると、文情セから送られてきた書誌データが東工大のシステムのファイルに落ちるという仕組みである。しかし、東工大

側の目録データベース(MODEL204)に直接的に入力されるのではない。文情セから送られてきた画面構成のためV Sデータ(Vertial Screenの状態のデータ)がこのコマンドによって順編成ファイルに落とされ、後でMODEL204のファイル形式にバッチにより編集して東工大目録データベースに入力するようになっている。

また、文情セのシステムでは叢書名・シリーズ名などのデータは個別の書名と切り離された同一レベルの1件のレコードとして作成され、キーによるリンク付けにて親子関係を保っている。ところが、東工大では親と子を合わせて1件の書誌データとして持っているため、東工大の目録データベースに入力する前には、この編集も必要となっている。

現在は、文情セの目録システムで検索をして、該当書誌データがあれば採ってくる片方向の流れだけであるが、次には東工大のシステムで作成した書誌データを文情セの目録データベースに入力する逆方向の流れを開発する課題が残っている。

4. 蔵書管理システムについて

いわゆる閲覧・貸出システムである。関西大学と同様に、利用者カードと図書のOCRコードをハンド・スキャナーでなぞる方式である。図書マスタがなくトランザクション方式のシステムである点が関西大学と異なる。

<図書館情報大学の図書館システム>

見学日：昭和60年6月27日

概 説：石川徹也 助教授

1. 概要

同大学は昭和55年10月に開学し、翌56年4月に第1期生としての学生が入学、昨年漸く全学年がそろったばかりの新しい大学である。

附属図書館には、通常の大学図書館としての機能の他に、次のような施設が設置され、学内の総合的な情報中枢機関としての機能を果たしている。

(1) MEDIA機器センター

マイクロ資料をはじめA. V. 資料や写真など図書資料以外の情報媒体資料および関連機器と設備を置いている。

(2) 図書館情報システム開発センター

その名の通り図書館システムを含めた大学内の情報処理センターの機能を果たす。

(3) 実習演習室

教育・研究における実験・実習を行なう場所として提供している。

(4) 公開図書室

教育研究およびその実践の場として、また開かれた大学の事業の一環として地域住民の利用に供している。

2. 図書館システム (ULIS-NET) について

これまでの図書館システムは、トータル・システムと呼ばれているものでも実際は図書館業務の枠組みの中で構想され、開発されたものがほとんどと言っても過言でない。しかし今後は（あるいは本来的には）学内全体の情報システム系の中に位置付けて構想していく必要がある。図書館システムは同大学としての本来的な研究対象でもあるため、その特徴として

- (1) 限りなく理想に近い図書館システムを目指すこと
- (2) 研究・教育の対象でもある。
- (3) パッケージ化により汎用性を持たせること

を挙げることができる。必然的にプログラムのステップ数も多くなり、システム自体が大きくなっている。しかし汎用性があるため、他の大学図書館へも容易に移植できるメリットがある。先頃、群馬大学附属図書館に全システムを移植し、貸出閲覧システムが稼働した。

3. 開発の理念と特徴

一般的に業務システムとしては、(ア) 速く、容易に、正確に、(イ) 使い勝手が良いシステムであることが要求される。さらに図書館システムでは、その特殊性として

- (ウ) 書誌データの項目が極端に数多く、その長さ、項目数が一定していない。
- (I) データ表現が自然語である。

図書館システムの開発を困難ならしめている大きな原因は、この後の2つに集約できる。つまり、キー・ボードのタッチ数（および種類）が極めて多く、必然的に誤りが増える。入力文字の正確さに対する気遣いが多くなり、また長時間におよぶ端末画面の凝視が疲労・けんしょう炎を引き起こす。また、システムとしても項目の多さと項目長のバラつきは、あらゆるシステムの中で極めて特徴的であり、これを柔軟に扱えるDBMSは未だ開発されていないし、また非常に困難である。一般にハードウェア、ソフトウェアとよく呼ばれるが、より使いやすい業務システムを開発するためには、これら以外にデータウェアと呼んでも良いものが必要と考えられる。銀行にはそれ専用のCash Dispenserが、国鉄には緑の窓口用の端末が開発されているように、日本語を扱い、しかも上述のような特徴を持った図書館システムには、それなりの端末やDBMSがなければ、業務システムの要求にあったシステムの開発は困難である。また

実務上のオペレータに与える負担が大きくなる。そこで図書館情報大学のシステムでは、データウェアの開発にも大きな力を注いできている。

(1) データの生成・育成機能

図書の購入申し込みの時点から、端末によってLC-MARCやJP-MARCを参照し、データがあれば採り込んでしまう。なければ簡単に申し込みに必要なデータだけを入力する。それ以降、発注・受け入れ・整理ではそれぞれに必要なデータを追加・修正し、書誌・所蔵データを各掛かりで生成していく流れを採用している。そして、伝票（データ）の流れである各業務と、ものとしての図書の流れを分離することが可能となり、受け入れられた図書が非常に速く利用に供せられるようになった。

(2) データ入力の軽減化のための支援機能

日本語データ作成機能として、一般の文書作成用のワードプロセッサは書誌データ作成のためには全く不十分である。そのため、次の3つの支援機能を独自に開発した。

ア. 読みのキー入力はローマ字（NDC訓令式、訓令式、ヘボン式）、仮名いづれでも可能で、しかも漢字変換後はすべてNDL訓令表記でファイルに入力される。

イ. 漢字変換後にも漢字の読みデータを残す。

ウ. 約15万語の科学技術用語およびNDC関連索引辞書を独自に作成し、搭載した。

この3つの機能を、端末側の仮名漢字変換システムと同期をとる形でホスト型仮名漢字変換システムとして開発・実用化した。

(3) システム・モジュールを容易に選択するための支援機能

従来よく使われてきたメニュー方式とコマンド方式を融合させる形で、キー1つでどのモジュールへも移れるクイック選択方式を開発した。毎回メニュー画面に戻るのではなく、ファンクション・キー1つを押すだけでどの状態からでも必要な業務画面に移ることができる。これは、閲覧・貸出システムのようなカウンタ業務では極めて効果的である。

5. 今後の課題

システムのステップ数かなり大きいので、プログラムの構造を整備する必要がある。また、学外とのネットワーク化を進める必要がある。

<日本科学技術情報センター>

見学日：昭和60年7月3日

概 説：五十嵐隆功 課長

1. 資料管理システムについて

IBMとドルムント大学が共同開発した図書館総合機械化用プログラム・パッケージ「DOBIS」をもとに資料管理システムを開発し、1980年7月から稼働している。その後3回のシステム改善を行ない、随分と使い易くなった。しかし、開発当時は漢字処理技術が未発達であったため英数字・仮名中心のシステムとなっており、漢字システムへの移行が残された最大の課題となっている。現在、国内外から約1万タイトルの雑誌を受け入れており、年々国内雑誌の比重が高まっている。これは日本の科学技術の発達が世界のレベルに追いつき、また追い越す状況を反映していると考えられる。世界各国から日本の科学技術情報に対する注目を受けており、同センターもその情報提供機能を強化する必要性に迫られている。

2. 抄録作成の説明

文献記事の選択は、原著者であることや1ページ以上に及ぶことなどの客観的な規準により行なっている。抄録の作成者として外部の各専門分野の研究者や技術者を約4000名採用している。抄録は150字、300字と2種類で外国語の文献の場合はタイトルの和訳も行なわれる。回収された抄録は、同センター情報部にある16の各専門分野の部門に割り振られ、そこで抄録の校閲・キーワードの付与などが行なわれる。主に、記事内容と抄録が掛け離れていないか、文章表現の修正、電算機の制限による漢字の修正である。先端領域では次々と新しいタームや学問分野が発生するため、オンライン検索で支障を来さないように、主題語以外にフリー・タームの中からキー・ワードを付与することも行なっている。

3. 資料部と情報部の見学

資料部では資料管理システムの端末によるデモを見学。全体としてメニュー方式を採用している。検索項目は、タイトル・請求記号・出版者・ISSNなど7つの項目を番号で選択して検索タームを入力する。検索結果はKWIC索引の形式で画面表示される。受入情報としては第1画面に巻号・出版年月日・受入ナンバーが表示される。その他の項目は必要に応じて入力するものとして第2画面が用意されている。巻号はシステム内で次号のデータが計算されてデフォルトで表示される。受入日・受入ナンバーもシステムによる自動採番になっているため、雑誌の受入データとしての入力項目が省力化されている。受入データを入力すると、端末付きプリンタのタック・シールにそのデータが印字され、オペレータが雑誌に添付する。出版者については典拠ファイルを持って

いる。また雑誌の出版パターンにかんする情報も持っている。

4. 今後の展開

現在のところ、次の3つの課題を検討し、または準備作業を行なっている。

(1) 英文データベースの作成

外国へサービスの1つとして、現在作業中で、60年度中には提供可能となる。第1のステップとして文献のタイトルおよび抄録の英訳が考えられている。学協会誌については英文の抄録が付いていることが多いので、その採用も検討している。

(2) 日英の機械翻訳プロジェクト

4か年計画で進められており、今年度が最終年度となっている。これはソフトが京都大学、アルゴリズムが電総研、辞書はJICSTが担当して共同開発しており、(1)のシステムに組み入れることが考えられている。

(3) ファクト・データベース

書誌以外のデータとして、金材研や他の研究所の協力を得て、実験データや物性データなどをデータベース化することが検討されている。

<国立国会図書館>

見学日：昭和60年9月5日

概説：今川浩一 課長

(総務部電子計算機課)

昭和44年に業務機械化準備室を設置して始まった。現在、マシンは日立のM-260K、ディスクは合計17.7GBとなっている。

来年別館がオープンとともに電子計算機課も移転する予定だが、それに合わせてオンライン用とバッチ用の2ヘッド・システムを予算化しようと考えた。しかし、政府の財政事情が苦しいことから現在機を若干レベルアップする程度に止ってしまった。

国会図書館のシステムは当初から冊子体の編集業務の機械化を主眼としたので、漢字入力は早くから意識していた。そのため、昭和23年から34年までの蔵書目録で、書名の中に出てくる漢字の使用頻度調査を行なったりして、昭和45年にNDL-70漢字コード表を作成した。これにもとづいて漢字入力機器を独自に開発し、システム化を進めてきた。

オンライン・システムは、数多くのものを日立の情報検索システムであるORIONに搭載してサービスを行なっているが、その中でも和図書システムを館外の利用に提供している。現在のところ、実験として端末機を1か月単位で各公共図書館に貸出して、オンライン検索をサービスしている。通信費以外に

データベース利用料金を徴収すべきである（文部省）という考えや、公共図書館など公共サービス機関については無料ですべきといった考え方もある。国会図書館としては第3セクター方式で運営してはどうかという考えもあるが、未だに結論が出ていない。いずれにしても課金問題は早晚結論を出さなくてはならなくなるだろう。

1. 国会会議録システム

本会議・委員会の会議録をもとに、質問者・回答者の発言内容・事項をデータ化して入力している。以前は発言内容から件名のようにキーワードを作成していたが、適切な検索が困難なため、第101国会の分からフリー・キーワードとして入力している。その結果、キーの数が6～7倍に膨れ上がった。

2. 和雑誌システム

国会図書館所蔵の和雑誌目録の編集ならびにデータベース作成システムである。ただし、現在のところ雑誌名の読みは、分かち書きをしていないのでタイトルからの検索では問題がある。

3. 雑誌記事索引システム

人文社会が約1800、自然科学が約1500の雑誌を対象として、雑誌の記事索引誌の編集およびデータベースの作成を処理するものである。また、平和情報センターの作成した自動仮名ふり・自動単語切り出しシステムを導入し、昭和59年1月分から論文タイトルに仮名ふりをしている。現在のところ9割以上の変換率を確認している。

< N T T 三鷹電ビル >

見学日：昭和60年9月19日

概 説：河岡 司 氏

坂口敏雄 氏

1. I N S の概要

I N S を推し進める技術は次の4つである。

(1) デジタル技術

デジタル伝送処理技術はもともと、コンピュータによる信号の解析処理やシミュレーションを行なうため開発され発達してきた。それがさらに音声の分析や合成、画像信号処理などの信号解析に応用され、現在では交換技術・伝送技術を始めとして電気通信の全分野に及んでいる。

(2) 光ファイバー

従来のメタリック・ケーブルに比べて低損失・広帯域・無誘導であり、数万倍の情報を超高速で伝送することができる。また、一種類のデータのみならず文字・データ・画像・音声などの多種の情報を同時に伝送で

きるため、これまでのように単一のメディア・サービスでなく、マルチ・サービスが可能となる。

(3) 通信衛生

大容量の情報を広い地域に、しかも災害などの影響を受けることなくで伝送できる手段として大きな期待が寄せられている。

(4) 超L S I

ハードの高速化・大容量化に大きく貢献する。

これらの技術開発が相まって、一家族の可処分所得の範囲内でINSの便利な機能を楽しむまでになってきつつある。

多様な通信サービスとしては、ディジタル電話・ホームファクシミリ・テレビ会議・移動通信（新幹線・自動車・船・飛行機）・テレコントロールなどがすでに実用化され、市民生活の中でその機能を発揮している。今後、INS構想のもとで展開されるサービスとしては、

- ・データベースサービス・・・新聞・辞書などのデータを音声入力し、電話などの問い合わせでデータベースをアクセスして回答を送る。
- ・ファームバンキング・・・・銀行と家庭・スーパー・デパートなどをネットワークで結び、料金支払いや決済などを自動的に行なう。
- ・分散オフィス
- ・通信講座
- ・書斎オフィス
- ・自宅学習
- ・ホームドクター
- ・生産管理システム
- ・ホームセキュリティ
- ・災害情報伝達システム
- ・ホームショッピング
- ・気象／市況情報サービス
- ・ホームバンキング・

雑誌や新聞などの記事情報（文字情報）を解読し、音声情報に変換してスピーカーから出力する。また、必要ならファックスに文字および図形（画像情報）を混在させて出力する。音声変換をするためには、

1. 1600ほどの音素を作成・記録し、これにアクセントなどを加えて合成する。
2. 強力な辞書
3. 文章の前後関係を解析する高度な処理が必要となってくる。

モデル・システムではブッシュ・ホーンで新聞記事のデータベースを検索し、その結果がコンピュータの音声で回答されるという実演。そしてこの基本システムにパソコンを接続し、パソコン上に作成したデータベースをユーザーがブ

ツシュ・ホーンで検索、あるいはデータ更新をするというシステムの実演を見学した。

I N Sモデルの見学

- ・ 3人電話 …… 3つの電話を利用して3人が同時に会話ができる。
- ・ キャプテンシステム …… ショッピング、新幹線予約、印象派美術のディスプレイ表示、メールボックス
- ・ 描画通信、メッセージ通信
- ・ 超高速ファクシミリ通信（A4で約2秒で、しかも超鮮明）
- ・ テレビ会議、テレビ電話

Ⅲ 研修レポートとOJT

1. 研修レポート

このセミナーで最も特徴的で、また苦勞したのがこのレポート作成であった。冒頭にも記したように、義務ではないというものの極めて強い強制的なものであり、単に作文や感想文では済まないからである。これまでのセミナー生のレポートをみても、内容的に最低でも学部の卒業論文の優秀なものであり、レベルが高いものでは修士論文を超えるものも提出されている。当然に引用・参考文献を短期間で揃える必要があるので、その調査や収集活動を計画的に進めなければならない。

またワープロによる作成が必須条件になっている。原稿用紙にして約100枚ほどをワープロするだけでも1週間はかかる。セミナーの最後の1ないし2週間はレポート作成のために用意されているが、ワープロおよびその修正で費やしてしまう（もちろん何日かは徹夜もしてのことである）。したがって普段から1日の研修の後、業務報告をワープロしてからでも、さらにレポートのための文献調査をしなければならない。私自身、セミナー後半のほとんどの日曜日は、東京大学総合図書館の参考図書室に座っていた。夏休みの間に必要と思われる文献資料をできるだけ東京に送ったものの、あとでどうしても必要になった文献や資料・データは関大の職場の方を煩わせて送って頂くなどした。この時ほど図書館、さらには資料が豊富に揃っている図書館の有り難さを感じたことはなかった。思いもかけず利用者の立場の貴重な体験が得られたのである。他のセミナー生も国会図書館や慶応の三田情報センターなどに日参していたようである。他の図書館関係の研修に比べ、このセミナーが相当に厳しいものであることが御理解していただけると思う。

さて、このようにして私が設定したテーマは、「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（受け入れから製本・所蔵データへの流れと学総目所蔵データ形式への変換）」というものである。これは、雑誌管理システムを構成する3つの主要なデータ、すなわち受け入れデータ・製本データ・所蔵データの間での相互関連を整理し、機械的処理の可能性を探ってみた。現在運用されているいくつかの雑誌管理システムについてレポートしているが、いずれも3つのデータ間での機械処理に成功しているところはなく、むしろそれは困難であるとしている。唯一関西大学の雑誌管理システムであるKULPISでは、手作業によるデータ・メンテナンスが多少とも必要だが、受入データから機械処理で製本データを作成している。さらに製本データから所蔵データを自動的に作成・更新するにはどこに問題があるのかを分析した。

また、文情セが編集している学総目への提出すべき所蔵データを、このロー

カル・システムから出来るだけ機械処理で作成できないものかという観点から、所蔵データの形式変換プログラムを2～3本作成し、そのロジックを紹介した。KULPISの将来の発展に少しでも役立てばと願うのだが。（この報告書に、研修レポート全文およびその概要を記したものを添付資料とします。）

2. OJT (On the Job Training)

実はセミナー参加が決定して、最も恐れていたのがこのOJTであった。というのは、「全国から集まった学総目の所蔵データの日チェックなどをいやというほどさせられた」との話を、以前にこのセミナーに参加した人から聞いていたからである。しかし、幸いなことに学総目の和文編は前年度で一応終了してデータベースが完成している。また欧文編の作業は昭和60年度の秋からということで、わたしのセミナーはこの端境期に当たったのである。私たちセミナー生が担当した作業は、この欧文編の全国調査をするための事前調査や調査資料作成の手伝いが主であった。具体的には、

- (1) 1980年移行に創刊した雑誌について、オンライン・システムを使って、NCファイル（学総目ファイル）での重複率を調査する。（学総目欧文編予備版作成のための準備作業）
- (2) 書誌データの調査対象機関を抽出するための条件をどのように設定したらよいかを検討するためのシミュレーションを行なう。書誌データの調査が特定の機関に偏らないようにするための条件を探る。
- (3) オンライン・システムを使って、雑誌の異誌名データのチェック調査を行なう。
- (4) 所蔵データ記入用のOCRデータシートにデモ・テスト用のデータを作成する。OCRのデータ読み取りミスについてのチェックに使用するため。
- (5) 学総目各編からサンプルとして20ページについて、雑誌名・団体名・出版地・変遷注記・異誌名・一般注記の各データが何文字で構成されているか、その分布状態・最大文字数・合計文字数を調査する。
- (6) 学総目欧文編全国調査のためのデータ作成マニュアルを校正する。
- (7) 全国調査のために各大学図書館などに発送する資料をまとめる。
- (8) ネットワーク参加館がオンライン目録システムで作成登録した書誌データについて、インデックス付けなどのバッチによる時差更新が正常に行なわれているかどうかをチェックした。
- (9) オンライン目録システムのマニュアル作成
概ね以上のような作業を行なった。

IV セミナーで得たこと：まとめ

1. 文情セの目録システムについて

規模としては、OCLCやUTLASなど欧米の書誌ユーティリティにおよぶべくもない。しかしシステム自体の内容としては文情セ自身が誇るように、正しく第3世代の書誌ユーティリティであると言えよう。つまり

(1) Shared CatalogingによるUnion Catlogの構築という理想像の具現化。

(これまでのシステムは決して総合目録とは言えない。)

(2) 異機種のコピュータ間におけるネットワーク・システムの展開

(これまではコンピュータと端末の通信であった。)

(3) エンド・ユーザにとってより使い易いインターフェイスの設計

(画面对応のシステムであり、目録端末としての特別仕様の設計)

(4) システム的発想を基本とした目録データベースの構造

(カード目録のイメージにとらわれないデータ構造)

(5) ALA文字などの大幅な拡張文字の採用

(これまでは当館のように音標符号をとった形で入力する考え方しかなかった。)

というように、新たな発想や本来あるべき主張を貫き、技術的にもむづかしい壁を次々と克服している。しかし1ユーザ(まだ実際のユーザになっていないが)の立場から敢て言えば、ローカルにおける目録システムあるいはその開発担当者(メーカーやSEも含めて)、さらには図書館員自身の思考は、まだ第2世代を抜け切れずにいるのではないだろうか。現に接続している国立の大規模な図書館の中にも、ローカル・システムとの機能的な接続というところに大きな問題点を含んでいるように思える。

検索あるいは接続が目的であれば、パソコンと若干の設備を購入して、NTTのDDX網に電話回線をつなげば済むことである。MARCデータのダウン・ローディングが目的ならば、恐らく当館が採用している出版MARCの導入方式の方が経済的であり、かつ手間が掛からないであろう。目録業務として導入するには、けっしてそんな単純ではないということを認識しておく必要があると思う。

2. 図書館システムの機械化担当の立場として

雑誌管理システム、閲覧貸出システムそして目録システムと関西大学図書館の機械化は進んできた。最近、書誌データを扱うようになって特に感じることは図書館システムの特異性である。それはまさに石川先生(図書館情報大学)がおっしゃっていた次の2点である。

(1) データ項目の多いこと。さらに可変的な長さでデータ項目数

(2) 自然語に近いデータを扱う。

この特質を十分認識した上でシステム構築に取り掛かる必要がある。でなければ、たとえ無事稼働してもオペレータに極めて大きい負担を課することになるのは明白である。機械化担当の自分としては、今後一層データウェアとしての機能充実に力を注ぎたいと考える。

3. 情報を処理するものとして

情報処理という意味の情報ではなく一般的な意味で、日常生活に必要な情報の収集は誰にでも必要であろう。しかし、図書館でその電算処理システムの開発にあたる者として、電算化に関連する情報については普段から積極的な収集活動が必要である。その意味で今回のセミナー参加により、図書館界の中心的な機能を果たしている国立国会図書館や日本科学技術情報センターそのほかいくつかの機関を見学できたことは、非常に有り難いことであった。セミナーで初めて知ったこととして、次の2つを挙げておく。

(1) 文情セ、国立国会図書館、日本科学技術情報センター、国文学研究資料館、東工大附属図書館、図書館情報大学、そして東京大学大型計算機センターおよび総合図書館、いずれも日立のマシンを導入してシステムを組んでいる。偶然なのかもしれないが、セミナーでお世話になったこれらの機関が、図書館界での主要な機関が揃って日立であった。おもしろい現象だと思った。

(2) DBMSとしてMODEL 204 (米国CCA社)を採用している機関が多い。文情セのシステム自体がそうであり、他に広島大学、名古屋大学、東京工業大学の各附属図書館が採用し、その上に図書館システムを構築している。図書館システムの構築では、かなりの実績を積み重ねたDBMSであると言えよう。このDBMS提供の簡易言語が素人にも津かい易いことも大きい要因であろう。

また、セミナー・レポート作成という体験を通じて、普段からの文献収集活動の重要性を痛感した。今後、図書館(職場)内で組織的にその活動ができるように進めて行きたいと思う。

図書館の電算化業務に携わってたった5年目で、このようなセミナー生活を体験させていただけたことについて、心から感謝したいと思います。

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変換)

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
Tsujikawa, Teruo: Kansai University Library

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れを整理し、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 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, holding record 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れとそれに付帯する各種のデータ処理関係をどう整理し、システム化するかということが電算化の最も重要なポイントであり、各大学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。

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	15,386

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

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期間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の通りである。

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(・)

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・欠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20万件を越す雑誌セ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セ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セルフデータ、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セ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。

2-2.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-

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ビータブルで持つ。

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の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.

1	2	3	...	...	24	25	26	27	28	29	30	31	32	33	34	35	36	...	...	9999

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になる。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らる。

(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 NO.131 1985.8
P.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 1984.11 P.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回トクメンテーション・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 1982.12 P.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 1984.5 P.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—記事索引とのからみで—」井上 如 図書館雑誌 76(4) P.217-219

その他

