機械による文献検索の実際

第2回 国内情報を利用する文献検索法[†]

山 中 啓

(筑波大学応用生物化学系)

承 前

われわれは長い間、自分の求める文献を専門分野の抄録誌の分野別記事と索引から検索するとともに、近着雑誌に直接目を通して探してきた。しかし、情報量の増加の早さは著しく、蓄積される情報量があまりにも多くなったので（化学の論文は年間40万件）、従来の方法ではもはや完全な検索が不可能になってきた。情報量が多いために「情報の洪水」と呼ばれるが、同時に情報が多いにもかかわらず「情報の不足」が叫ばれている。それは自分の求める文献をたやすく得ることが、マニュアルによる方法ではしだいに困難になってきたからである。しかし、情報の生産速度や生産量を制限できないので、われわれのとるべき解決法は処理速度の速いコンピューターを利用する以外に方法はない。それが最近になってやっと可能になったのである。それが機械によるオンライン情報検索である。

第1回では、この背景と基礎について説明をした。今回は、われわれが手許で実際に使用できる国内情報の利用の仕方について具体的に説明をしよう。

7. データベース

われわれが利用している *Chemical Abstracts* 誌、*Biological Abstracts* 誌のような抄録誌（二次資料）はそのままの形ではコンピューターで検索できない。これらとはほぼ同じ内容を磁気テープに入れ機械での読取りが可能になったものをデータベースと呼ぶ。すなわち、ある目標をもって収集され、編集された情報集団である。国立大学図書館協議会大学図書館改善協議会⁽⁶⁾によ

れば、データベースは次のように定義されている。

イ。大量のデータが組織化された形でまとめられている。

ロ。1ないし数個の電子計算機で読み出すことのできる形でファイルに入っている。

ハ。データの蓄積検索、維持管理のためのシステムおよびプログラムが完備している。

したがってデータベースは文献情報のみではなく、数値情報もデータベースである。事務処理上の多くの数値情報、ガスクロや液クロのデータ、マスペクトルのデータ等も収録するとデータベースとなる。これらの情報が一定の基準のもとに収集され、編集され、さらに新しい情報が追加または更新された情報集団がデータベースである。科学技術全般にわたる文献情報では一定の基準は幅広く設定されているが、特定の研究者に非常に有用なデータベースとして、特定の問題に関して情報を集めたデータベースも多い。たとえば、抗生物質のデータベース (Index of Antibiotics from Actinomycetes)⁽⁷⁾、たんぱく質構造データベース^(8,9)等があげられよう。

データベースの特徴は共有化である。異なったシステムからもアクセスできるように設計され、利用者の多方面の要求に対応できるように管理運営される。この点が対象範囲を限定した電算機のファイルと根本的に異なる点である。

昭和53年現在で世界で市販されているデータベースは130種以上といわれている⁽⁶⁾。そのほかに大学、企業等で独自に開発されたデータベースがあるが、これらは通常一般に公開されていない。われわれは市販されているデータベースを直接に利用するのではなく、これらのデータベースを購入し整備している機関を通して初めてアクセスができる。わが国で利用可能なデータベースと取扱機関を表2にまとめた。個々のデータベースをデータファイルと呼び、機械検索システムは各機関でそれぞ

[†] A Practical Guide for Information Retrieval System. 2. How to access database.

Kei YAMANAKA (Institute of Applied Biochemistry, The University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki-ken 305)

表 2 農芸化学分野のデータベースと取扱機関

分 野	データベース名	収録件数/年	取 扱 機 関
化 学	CAS(Chemical Abstracts Service, 米)	40 万	JICST, 東大, 筑波大, 丸善, 紀伊国屋
理 工 学	JICST 理工学文献ファイル (日)	36 万	JICST, 広島大
生命科学	BIOSIS(Biological Science Information Service, 米)	25 万	筑波大, 広島大, 丸善, 紀伊国屋
医 学	MEDLARS(National Library of Medicine, NLM, 米)	23 万	JICST, 紀伊国屋
	EXCERPTA MEDICA	28 万	筑波大, 丸善, 紀伊国屋
	TOXLINE(NLM)	12 万	JICST, 紀伊国屋
農 学	AGRICOLA(National Agricultural Library)	10 万	丸善, 紀伊国屋
	CAB(Commonwealth Agricultural Bureau)	10 万	筑波大, 丸善, 紀伊国屋
環境科学	ESI(Environmental Science Index)	1 万	筑波大

れ略字名をつけている。

その他, 大型電算機センターを持つ国立大学でも一部の情報について学内利用がされている。またこれらを統合した学術情報の全国的ネットワークも計画されているので, 将来は幅広い利用が可能になるであろう。現在では, 商業ベースの丸善, 紀伊国屋が多数のデータベースをとり揃えている。東大および筑波大の学術情報検索システムではかなり幅の広い利用ができるが, 広島大ではその利用は学内者に限定されている。幅広く公開されている機関は, 特殊法人・日本科学技術情報センター (JICST) である。第 1 回に述べたように JICST は各支所が開設され, 全国的規模でのオンライン検索が可能になった。そこで JOIS 公衆回線サービスで検索できるデータベースについて説明をする^(10,11)。

(1) JICST 理工学文献ファイル

JICST で作成される理工学全般にわたる文献ファイルで「科学技術文献速報」(11 編)に対応する。対象分野は天文学, 数学を除く理工学および生命科学, 環境公害, 管理システム技術である。現在では医・農・生物学の分野については限定されているが, 将来生命科学を中心としてこれらの分野は充足される予定である。情報源は定期刊行物 8,500 種 (うち国内誌 3,000 種) のほか, 技術レポート, 会議資料等である。とくに国内情報を中心に検索する場合, これが唯一のファイルである(注1)。

(注1) 最近の JICST の調査によると, 理工学ファイルの内容は次のようである。

記事に採択された種類を国別にみると, 日本 41.2%, アメリカ 19.0%, 西ドイツ 6.9%, ソビエト 4.1% であり, 使用言語別では, 英語 58.2%, 日本語 14.5%, ロシア語 13.8%, ドイツ語 8.9% となっている。また分野別では, 化学 24.7%, 機械 14.2%, 電気 11.2%, 金属・鉱山・地学 10.6%, 管理 9.3%, 土木・建築 8.8% である。

蓄積開始は 1975 年 4 月で, 年間の情報量は平均約 36 万件である。毎月約 3 万件が追加されている。印刷を必要としないために「科学技術文献速報」よりさらに 1 カ月早くなり, カレントな検索が可能である。理工学文献ファイルは自然語では検索できず, シソーラスにもとづいて検索できるように設計されている。「JICST 科学技術用語シソーラス (1978 年版)」と「同主題カテゴリー別索引」の 2 冊が準備されている。われわれはこの 2 冊を用いて自分の検索したい項目 (自然語) を最も適切なディスクリプタに換え, これをキーワードとして検索をする。「主題カテゴリー」は表 3 に示したように 14 分野に分類されており, まず自分の分野をこれで確かめておく。と後の検索がやりやすい。この分類は科学技術文献速報の主題 (表 4) とはやや異なっている。とくに生化学の分野の分類を表 5 に示した。生物学分野もかなりカバーされているが, 医学, 土壌肥料, 飼料のディスクリプタ数は少ない。このシソーラスは医・農・生物学関係の用語を含めたものに改訂作業中ときいているので, 56 年度からはさらに充実したものになるであろう。

各収録文献にはキーワードとして平均 5 つのディスクリプタが付与されているほか, それらの上位語が 5~6 個自動付加 (アップボスティングと呼ぶ) されているので, 上位概念からの検索も可能である。

(2) CAS ファイル

Chemical Abstracts に対応するもので, 生化学, 有機化学, 高分子化学, 物理化学, 分析化学, 応用化学, 化学工学等が収録されている。情報源は 125 以上の国の雑誌 14,000 種, 27 カ国の特許, 図書, 学位論文等である。年間の情報量は約 40 万件で毎週平均 7,000 件ずつ追加されている。各文献にはキーワード, コードが付与されているが, 検索にはシソーラスを用いず, 自然語をキーワードとして用いる。したがってオンライン検索に不慣れの人でもタイプライターと全く同じ要領で簡単に検

表 3 JICTS 科学技術用語シソーラス主題
カテゴリー別索引の分野構成

分 野	記号	分 野	記号
土 建	A	鉱 山	G B
管理システム	B	地 球 の 科 学	G C
基礎化学分野	C A	共 通 分 野	I A
生 化 学 分 野	C B	環 境 公 害	K A
工業化学分野	C C	機 械	M
電 気	E	原 子 力	N
金 属	G A	物 理	P

表 4 科学技術文献速報昭和 52 年度収録記事
分類別一覧

	収録件数
物理・応用物理編	71,005
化学・化学工業編 (外国編)	144,026
機械工学編	80,562
電気工学編	48,420
金属工学・鉱山工学・地球科学編	52,329
土木・建築工学編	41,329
原子力工学編	22,689
管理・システム技術編	25,478
環境公害編	18,847
化学・化学工業編 (国内編)	25,213
計	529,896

表 5 JICST シソーラス生化学分野 (CB) 分類

01 生 化 学	14 動 植 物
02 生化学・実験技術	15 病 気
03 生 理 学	16 薬 学
04 遺 伝 学	18 医 学
05 免 疫 学	21 土 壤 肥 料
10 生 体 成 分	22 農 薬
11 酵 素	23 飼 料
12 細胞・組織・器官	24 食 品
13 微 生 物	25 発 酵

索できる。Subject index をみて検索する従来の方法と対比すれば、その呆気ないほどの簡便さと 1 回の検索結果が数秒～1 分以内に示されるスピードを特筆したい。もちろんキーワードの使用頻度リストによって自分の検索目的に最も適合した用語を組み込んで検索することが重要である。分類コードとしてのセクション番号による検索もできる。

(3) MEDLARS 医学文献ファイル

米国国立医学図書館で作成される医学文献ファイル (Medical Literature Analysis and Retrieval System, MEDLARS) で Index Medicus に対応するものである。年間収録件数は約 25 万件でキーワード、MeSH トリー番号および MEDLARS の qualifier で検索できる。

(4) TOXLINE 毒性文献ファイル

米国国立医学図書館が CAS, BIOSIS, Index Medicus 等 11 種のファイルより毒物学を中心にして再編集したものである。年間収録件数は約 12 万件で、各種化合物の毒性、安全性を知るのに有用なファイルである。

(5) CLEARING 国内研究案内ファイル

上記 4 つのファイルと趣を異にするファイルで、JICST が国内 400 の公共試験研究機関における現在進行中 (ON GOING) の研究課題のファイルである。対象は理工学文献ファイルと同じく、現在は生物学、医学、農学の分野が含まれていない。国公立の試験研究機関のみを対象とし、大学の学部および民間会社の研究課題は含まれない。JICST の特異なファイルでアメリカでは Smithsonian Science Information Exchange (SSIE) に相当する。検索は JICST の分類コード、ディスクリプタ、研究機関コード、研究機関略名および研究者名をキーワードとして行う。

(6) 米国クリアリング情報ファイル (SSIE)

理工学全般から生物科学、社会科学、行動科学などの分野について米国内の研究プロジェクトに関する情報のファイルである。SSIE 発行のもので、現在までの蓄積量は約 10 万件である。

8. オンライン検索の準備

オンライン文献検索を始めるに先立って、われわれの準備すべき事項の 1, 2 について述べてみよう。それはぐずぐずしていると時間が空費されてむだな経費がかかるので、事前に充分準備しておくことが最も大切であるからである。

(a) 目標を明確に

オンライン検索法では、検索結果は該当文献数が何件あるかをまず答えてくるので、質問内容を機械が誤りなく理解できるように十分に検討する。関連語、上位語をうまく用いて広く検索することもできるし、限定した問題について該当文献の有無を確認するような調査もできる。この点はマニュアル検索法ではほとんど不可能な調査法である。たとえば、合成洗剤の分解に関しては、50 以上の報告があるが、カテコールの分解に関しては文献がたいへん少ないといった調査が瞬時にできる。しかし、漠然としたテーマの調査は機械になじまない。たとえば、「たんばく質に関する研究」とか「抗生物質の構造と作用機作」とか「微生物のエネルギー代謝」のようなテーマでは機械には通用しない。もっと自分の検索目標を具体的に示すことが肝要である。たとえば、「たんばく質の X 線結晶解析」とか「DNA 合成×阻害×抗生物質」のように内容をより具体化したり、分解したりする

表 6 JICST 科学技術用語シソーラス・主題
カテゴリー別索引の一部

CB 04 遺 伝 学

DNA修復	* 作動遺伝子
DNA複製	USE 遺伝子
* F因子	サブレクション
USE 性因子	シストロン
* R因子	自然とうた
USE 薬物耐性因子	性因子
アロメトリー	* 制御遺伝子
アンチコドン	USE 遺伝子
鋳型[遺伝]	生物進化
遺伝	生命の起源
* 遺伝因子	染色体異常
USE 遺伝子	* 染色体外遺伝
* 遺伝現象	USE 細胞質遺伝
USE 遺伝	* 染色体外遺伝因子
遺伝コード	USE プラスミド
遺伝子	前生物学
遺伝子型	組織適合性遺伝子
☆ 遺伝子工学	* 対立遺伝子
遺伝子座	USE 遺伝子
* 遺伝子制御	致死遺伝子
USE 遺伝的調節	* 致死因子
遺伝子地図	USE 致死遺伝子
遺伝情報	* 調節遺伝子
遺伝情報因子	USE 遺伝子
遺伝的影響	* 低カロリー食品
遺伝的調節	USE ダイエット食品
インホルモゾーム	適応
ウイルス複製	適応放散
宇宙生化学	転写
栄養素要求性	突然変異
エビゾーム	* 突然変異株
* オペレータ遺伝子	USE 突然変異体
USE 遺伝子	突然変異体
オペロン	* 突然変異誘導
化学進化	USE 変異誘起
組換	* 突然変異誘導体
形質転換	USE 変異誘起物質
形質導入	* トリプレットコード
形質発現	USE 遺伝コード
* 系統関係	* ナンセンスコード
USE 系統発生	USE 遺伝コード
系統樹	☆ 倍数性
系統発生	表現型
交差	複製[遺伝]
交雑	プラスミド
* 構造遺伝子	* プロモータ遺伝子
USE 遺伝子	USE 遺伝子
コーディング活性	分子遺伝学
コード縮退	分子交雑
コード認識	分子進化
* コドン	変異
USE 遺伝コード	変異誘起
コリン因子	変異誘起物質
細胞質遺伝	翻訳[遺伝]
雑種	薬物耐性因子
* 雑種形成	レプリコン
USE 交雑	レプレッサ

CB 04 遺伝学の部分を示した。☆印の付されている遺伝子工学は 1978 年に新しくディスクリプタと

して採用されたことを意味し、*印は非ディスクリプタであることを示している。たとえば、オペレーター遺伝子も構造遺伝子も非ディスクリプタで、検索は USE 遺伝子をキーワードとして用いよという意味である。クロニングとかベクターという用語はなく、これらでは検索は不可能である。

とよくなる。

(b) データベースの選択

取扱機関(表2を参照)の所有するデータベースのうちより、自分の検索に最も適合したファイルを選定する。JICST の上述の3つのファイルについて、収録された主要雑誌 200 タイトル中の収録報文数の比較調査もされている⁽¹¹⁾ ので、ファイルの特徴を理解して選択する。化学関係の場合、CAS と理工学の2つより検索すると、重複はあっても、検索もれが少なくなる^(注2)。ただし、1回の検索は1つのデータファイルについて行われるので、2つ以上のファイルの同時混合使用はできない。

(c) 質問式

コンピューターとの会話は、こちらが出した質問に対してのみ機械が返事をしてくるので、質問のしかたが検索結果の成否を決める。質問は、1つのキーワードを、たとえば detergent とタイプライターで打つ。すると、detergent を持つ該当文献の数が返答となって返ってくる。1つのキーワードでも検索ができるが、より正確に検索するために、第1回で述べた論理演算法により、演算子を用いてキーワードの組合せをつくる。これを質問式という。プロファイルと呼んでいる。この質問式の作製がオンライン検索において最も重要な作業である。しかし利用者はコンピューターの利用には慣れていないため、利用者が作製した検索式では完全に利用者の要望を満たさないことが多い。経験のある検索の専門家の助言を求めることである。しかし、専門家に任せてしまうと、利用者の専門知識が生かされないので、両者の協同作業によって初めて質問式が完成される。

(d) シソーラス

JICST の理工学文献はシソーラスにより管理されている。7.で述べたように、まずわれわれの自然語がシソーラスに登録されているかを必ず確認する。そのためには、主題カテゴリー別索引で目的の分野の中でアルファ

(注2) JICST の理工学ファイルと CAS との重複度についての JICST の最近の調査によれば、理工学ファイルの 8,000 種の資料のうち、2,040 種が双方で収録されていた。収録頻度の高い上位 1,000 種の資料では、その 81.3% の資料が双方で収録されていた。

表 7 JICST 科学技術用語シソーラスの例

遺伝コード(イデンコード)	
CB04	92, 92
UF	コドン
	トリプレットコード
	ナンセンスコード
BT	遺伝情報因子
見出語 →	遺伝子(イデンシ) ← ふりがな
主題カテゴリー →	CB04 979, 822 ← 索引使用 コード 頻度
UF	遺伝因子
	オペレータ遺伝子
	構造遺伝子
	作動遺伝子
	制御遺伝子
	対立遺伝子
	調節遺伝子
	プロモータ遺伝子
NT	オペロン
	シストロン
	組織適合性遺伝子
	致死遺伝子
	レプリコン
BT	遺伝情報因子 ← 上位語
遺伝子型(イデンシガク)	
CB04	79, 79
BT	遺伝
遺伝子工学(イデンシコウガク)	
CB04	0, 0
ヒストリー →	■ 78 ~
表示	遺伝子座(イデンシザ)
	CB04 165, 165
BT	遺伝子地図
*遺伝子制御(イデンシセイギョ)	
CB04	
USE	遺伝的調節
遺伝子地図(イデンシチズ)	
CB04	312, 150
NT	遺伝子座
遺伝情報(イデンジョウホウ)	
CB04	19, 19
遺伝情報因子(イデンジョウホウインシ)	
CB04	1346, 9
NT	遺伝コード
	遺伝子
	・オペロン
	・シストロン
	・組織適合性遺伝子
	・致死遺伝子
	・レプリコン
	・インホルモゾーム
	・エピゾーム
	・コリシン因子
	・性因子
	・薬物耐性因子
	・プラスミド

表 6 に対応する遺伝子(イデンシ)の部分を示す。
表 6 に示した非ディスクリプタは*印で示し、USE
の次にディスクリプタを示すとともに、別のところ

では UF 関係語としてもまとめて示されている。また上位語と下位語の関係が、遺伝情報因子→遺伝子→オペロンであることが、遺伝情報因子の項、遺伝子の項にも示されている。もちろんオペロンの項は BT として示されている。UF, Use for; NT, Narrower Term; BT, Broader Term.

ペット順に並べられているのを探そうが簡単である。

「索引」の一部とこれに関連した「シソーラス」の部分を表 6, 7 に示す。索引に用いた自然語がシソーラスにのっていると、これをディスクリプタと呼ぶ。したがってこの言語をキーワードとして検索ができるが、一度シソーラスを開いてみることをお勧めする。すると、この言葉の上位語と下位語が示され、かつ使用頻度も示されるので、検索を正確にするためには、どのディスクリプタを採用すればよいかが明白になる。

索引語が「主題カテゴリー別索引」に*印をつけて登載されている場合は、この言葉は非ディスクリプタであって検索には使用できない(表 6 参照)。しかし、すぐ下に使用すべきディスクリプタが示されているので、これをキーワードとして用いるか、または上述のようにさらにシソーラスで調べてみる。問題は索引に用いたい自然語が、ディスクリプタ、非ディスクリプタのいずれにも指定されていない場合には、「主題カテゴリー別索引」の分野のなかで最も近い言葉を探さなければならないので厄介である。第 1 回で例にあげた「生分解」がその例で、この言葉がシソーラスにないために、「生物」と「分解」の 2 つの言葉に分解するか、関連用語として「微生物分解」を採用する。このようなケースは多々あると予想される。

またシソーラスには、われわれの日常用語をディスクリプタとして採用しているものが多いが、日常用語とディスクリプタとが一致しないものもまた多く、感覚的にはかなりの抵抗感を覚えるものもある。やや意地悪かもしれないが、極端とも思えるような例を少し引き出してみた。

非ディスクリプタ	ディスクリプタ
金属たんぱく質	色素たんぱく質
サイトカイニン	シトキニン
アルファ酸	苦味酸
公害	環境汚染
塩化ビフェニル	ポリクロロビフェニル
石油発酵	炭化水素資化発酵

左側の用語はすべて非ディスクリプタなので、これでは検索ができない。右側のディスクリプタを使用する。石油発酵を炭化水素資化発酵とおきかえることは、なるほどとなぜけるが、カナ文字タイプで「タンカスイソ

シカハッコウ」と打つのは苦勞しそうだ。サイトカインをシトキニンと置き換えるのは、カナ文字が2字減ったが、どうもピンとこない。動植物名、とくに微生物名は必ずシソーラスを調べないと、ディスクリプタがわからない。たとえば、*Aerobacter* 属は好気菌属、*Alcaligenes* 属はアルカリ菌属とするような例が非常に多い。

これらはやむをえないのかもしれないが、利用者の立場からは、もう少しなんとかならないものだろうかと常常思ってしまう。なお、シソーラスについての詳細な解説記事^(12,13)があるので参照されたい。

9. オンライン検索の開始

それでは JOIS を例にとってオンライン検索の手順を説明しよう。

表8に示した JICST の支所へ直接行く場合も同様であるが、JICST に登録してパスワードの交付を受けた機関にある端末装置を利用する。登録機関名は JICST に照会されたい。大学関係では付属図書館または分館であるが、表9にまとめたようにまだ普及しておらず、医学関係に圧倒的に多く、次いで工学部である。オンライン検索を手近なものにするためには、大学図書館のレファレンス・サービスとして参考業務のなかに積極的に取り組んでゆくべきであろう。

端末装置はスクリーンのついたディスプレイ型と簡便なタイプライター型がある。後者はケースに入れて電話機の近くへ持ち運びが可能である(図5)。タイプライターと同じキーボードに英文字のほか、記号、カナ文字が英文字とともに配列してある。1つのキーに英数、カナの双方が入っているの、いずれかの選択のキーを押してタイプライターと同様に操作すればよい。発信および受信はすべてスクリーンにディスプレイされるか、印字される。端末機のもう1つの機能は音響カブラーであ

る。電話機の受話器をこのカブラーにセットすると、中央の計算機と端末装置がオンラインで接続されたことになる。音声と電気信号の変換機能を果た部分である。

表8に示した近くの JICST の支所へ電話をかける。つながるとピーという音がする。これがコンピューター

表8 文献検索取扱機関

取 扱 機 関	電 話 番 号
JOIS 呼び出しステーション	
札 幌	011-222-1341
仙 台	0222-75-2071
東 京	03-580-2624
筑 波	0298-51-8412
名 古 屋	052-203-0326
大 阪	06-445-6191
広 島	0822-28-7379
福 岡	092-473-5521
丸 善	03-271-6068
同 上 (大阪)	06-251-2621
紀 伊 国 屋	03-364-0751

注) JICST の支所のうち、富山、高松には特定回線の端末装置が置かれたが、呼び出しステーション化は富山6月、高松10月の予定である。沖縄は福岡を利用する。

表9 JICST 端末装置を設置している大学*
(昭和55年4月現在)

学 部 別	設 置 数			
	国立	公立	私立	小計
医・歯・薬(医科大学、図書館医学分館を含む)	27	6	28	61
工	19	0	4	23
附属図書館その他 (研究所、研究室を含む)	47	1	14	62
計	93	7	46	146

* 高専を含む。

FACOM-1513 A (富士通)

NP 3000-20 (日本電気)

図5 公衆回線用端末機(タイプライター型)

手前がキーボード、その後ろがプリンター。電話機の受話器を接続する音響カブラーは左または右側の部分。カブラーは後ろのタイプ、本体と別のタイプもある。

の声である。この接続音を確かめて、受話器を音響カプラーにはめ込む。接続を確認すれば会話を開始できる。これから後の手順は計算機のシステムにより異なるのでそれぞれのマニュアルについて検討をしてほしい。今回は JOIS について例示するが、具体的な使用法は「JOIS の利用のてびき」を参考にいただきたい。ユーザー ID とパスワードを入力してユーザーの確認が済むとただちに利用可能の状態になる。パスワードは端末を使う人が登録された機関からであることを確認し、情報の管理を行うためのもので、キャッシュカードの暗番と同じである。東大の大型計算機センターの場合には、さらに乱数を用いて利用者の確認がされている。ここまでの操作は簡単でおそらく1分も要しないであろう。

ここでシステムからはウェルカム・メッセージが出されて会話が始まる。JOIS では、利用できるデータベース名とその蓄積範囲および収録件数が示されて、どのデータベースを利用するかを尋ねてくる。ただちに番号で答える。たとえば CAS の場合、2 のキーを打つ。するとさかのぼって5年間の CA の番号とそれぞれの収録件数が示され、利用希望の範囲を再び尋ねてくる。たとえば、0~4 のように範囲について返答をする。システムからは、「サービス オ カイシシマス」といってくるとともに時間が記録される。ここまでの手順の一例を図6

に示す。この瞬間から使用時間の計算が始まるので、ぐずぐずするのは禁物である。前もって調べておいた検索式をタイプライターと全く同じ要領で打ち込めばよい。理工学ファイル、クリアリングファイルの場合はカナ文字で、その他のファイルの場合は英文字で打つ。第1の質問のキーワードを打つと、ただちに該当文献の件数がシステムからの返事となって返ってくる。すぐ次のキーワードを打つと、また該当文献数が何百とか何千とかの件数で回答が示される、このように会話はタイプライターを媒介とした1問1答式である。2つのキーワードの論理積を求めるには、先ほど検索したキーワード（検索タグと呼んでいる）をもう1度使わないで、会話のステートメント番号を用いる。それは1, 2の質問で検索された該当文献は別のファイルへ移され、そのなかで論理積の計算をさせるためである。入力の方法としては、1*2とする。すると途端に件数が1/10から1/100位にまで減ってしまうことが多い。この内容を知るためには、表示の命令（コンピューターへの命令をコマンドという）を与えると、ビデオ型端末機ではスクリーンに表示されるし、タイプライター型端末機ではプリントされる。これをオンライン出力という。オンラインで検索し、その結果をオンラインで表示（印字）させるのがオンライン文献検索である。

```

YIN JOI
JICST オンライン ケンサウ サービス
H

① U: YIN JOIS,08D-0410,YAMANAKA
S: パスワード ハ?
② U: 0000000000000000
JICST オンライン ケンサウ サービス
③ NO データベース 検索済 収録件数 蓄積範囲 354
1 --- JICST (1975.04 - 1980.04) 1,823,770 ケン 17:00
2 --- CAS (V.084N.04-V.092N.08) 1,734,219 ケン 17:00
4 --- CLEARING (ON GOING - HISTORY) 19,208 ケン 17:00
5 --- SSIE (1978.01 - 1978.12) 104,511 ケン 17:00
S: データベース / ハンゴウ ハ? 1,2,4,5,7,8=N
④ U: 1
⑤ データベース JICST JICST-COPYRIGHT
ハンゴウ 検索済 収録件数
S --- (1980.04) 32,555 ケン
0 --- (1980.04 - 1980.04) 32,555 ケン
1 --- (1979.04 - 1980.03) 361,297 ケン
2 --- (1978.04 - 1979.03) 351,400 ケン
3 --- (1977.04 - 1978.03) 390,072 ケン
4 --- (1976.04 - 1977.03) 333,654 ケン
5 --- (1975.04 - 1976.03) 354,792 ケン
S: データベース / ハンゴウ ハ? S,ハンゴウ=ハンゴウ
⑥ U: 0-5
⑦ S: サービス オ カイシシマス 1980.03.28 13:59:47
データベース JICST (1975.04 - 1980.04) 1,823,770 ケン
カイワ ハンゴウ 34

```

図 6 JOIS の検索例

①08D-0410 ユーザー ID, その次は使用者名, 1文字でもよい, ②パスワードは読めないように2度打ちをしている。③利用可能のデータベースが示される。④データベースの選択。⑤選定したデータベース（この例では JICST 理工学文献ファイル）の蓄積範囲, ⑥蓄積範囲の選定, ⑦サービスの開始

回答事項の主なもの、論文題名、著者名、著者の所属機関名とその所在地の略記、雑誌名、巻、号、ページ、発行年の書誌事項とその論文に付与されたキーワード(ディスクリプタ)とコード番号である。JOIS-I の場合には抄録はない。これらの書誌事項のうちのいくつかを選択して、そのみを回答させることもできる。この検索例を理工学文献ファイルと CAS ファイルについて一例を図7および8に示す。また本年3月よりサービスが開始された JOIS-K ではカナ文字で検索し、印字は漢字ひらがなの混じり文であり、要約も印字される。実施例を図9および図10に示す。図8には会話の終了のコマンドを与えて印字される部分も示した。質問時間(使用時間)と回答件数および使用料金が示されている。これで会話は終了する。図8の例は件数のみを CAS ファイルで調べたのみであるから、回答件数は0件、質問時間は2分間、料金が594円であった。1,734,219件の文献のなかから固定化酵素の報文がいくつあるかの調査が594円で2分間でできたが、途中Eコマンドを使ったりしているので、これのみでは1分以内で終了するはずである。

この経過をみると、こちらから入力するのは若干のコ

マンドとキーの打ち方、スペースの入れ方を覚えればよい。これは外国語の文法をマスターすること比べれば、われわれが知っておくべきコンピューターの文法はごくわずかでよいことが充分にわかっていただけたと思う。電子計算機を今まで使っていない人も、プログラムを全く知らない人でも、英文タイプが打てれば充分に使いこなせる。指一本でタイプを打てば、時間がかかって料金が高つくだけである。しかし、ここに落とし穴がある。初めてやってみると、論文が次々と打出されてくるのが面白くてたまらない。まして自分の論文が打出されるとやはりうれしい。そのために興味のおもむくままにあれもこれもとむだな検索をついやってしまう。麻薬中毒と同じで、あっという間に料金が増えてゆくからである。大事な格言を1つ。「コンピューターと遊ぶのはほどほどに」。

10. 検索例

(a) 酵素の例

理工学ファイルではソーラスに載っている酵素は1つのキーワードで検索できるが、CAS ファイルでは、英語が分ち書きのため、作用名の部分と基質名とに分けそれぞれの論理積で検索をする。そのためにどうしても

```

0048                                     JICST-COPYRIGHT
      KA031507      628.543:622.5
A77211412      G77210947      K77110557
THE ROTATING BIOLOGICAL CONTACTOR FOR BIOCHEMICAL (A1) (USA)
FERROUS IRON OXIDATION IN THE TREATMENT OF COAL
MINE DRAINAGE.
      OLEM H
P999AA      PB REP NO. PB264534 PAGE.85P (75) EN
カイクンハク; コウサンハク; サイクン; サンセイハク; セイカク;
セイフツカクテキミシヨリ; タンコウ; レツカコウフツ; ハイスイヨリ

```

```

0049                                     JICST-COPYRIGHT
      KA031105      628.543:667
A77211388      J77112158      K77110519
ホリヒニルアルコール ノ ヒセイフツフンカイ ト ソ カンコウハイスイヨリ ヘ ノ (B2) (JPN)
オウヨウ
      ススキ; タクイ; フクカ
F231AA      NIPPON NOGEI KAGAKU KAISHI VOL.51 NO.7 PAGE.R53-R58
(77) FIG.6 TAB.2 REF.24 JA
カツセイラツシホウ; ハイスイヨリ; ホリヒニルアルコール; PSEUDOMONASソク

```

```

0050                                     JICST-COPYRIGHT
      KA031105      628.543:636/637
      KA031600      628.543:667
A77211415      J77111907      K77100561
PHOTOSYNTHETIC BACTERIA IN WASTE TREATMENT. ROLE OF (A1) (JPN)
RHODOPSEUDOMONAS CAPSULATA WITH
AGRICULTURAL/INDUSTRIAL EFFLUENTS.
      SAWADA H; PARR R C; ROGERS P L
G5358A      J FERMENT TECHNOL VOL.55 NO.4 PAGE.326-333 (77)
      FIG.5 TAB.3 REF.10 EN
サンキヨウ; ノウキヨウ; ハイスイヨリ; リョウ; RHODOPSEUDOMONASソク

```

図7 理工学ファイルの検索例

この検索例は、[1] ハイスイヨリ 8214 件、[2] サイクン*1 122 件のオフライン出力したものの一部分である。3つの報文は、PB レポート、農化誌、発工誌である。


```

S: ツキノ シツメン カ アリマスカ ?   AI=Y, IAI=N
U: Y
S: テーザース マリハ ルイ オ カイマスカ ?   IAI=N, テーザース=0, ルイ=R
U: 0
JICST オンライン クソク サーフス
NO   テーザース   サーフス   ルイ   クソク   ショウリョク   ユーイ
1 --- JICST   ( 1975.04 - 1980.04 )   1,823,770   ㌥   17:00
2 --- CAS   (V.084N.01-V.092N.08)   1,734,219   ㌥   17:00
4 --- CLEARING   (ON GOING - HISTORY )   19,208   ㌥   17:00
5 --- SSIE   ( 1978.01 - 1978.12 )   104,511   ㌥   17:00
S: テーザース ノ ルイ カ ?   1,2,4,5,7=N
U: 2
テーザース CAS   ACS-COPYRIGHT
ルイ   サーフス   クソク
S --- (V.092N.08)   9,866   ㌥
0 --- (V.092N.01-V.092N.08)   69,040   ㌥
1 --- (V.090N.01-V.091N.26)   436,682   ㌥
2 --- (V.088N.01-V.089N.26)   428,174   ㌥
3 --- (V.086N.01-V.087N.26)   409,742   ㌥
4 --- (V.084N.01-V.085N.26)   390,581   ㌥
S: テーザース ノ ルイ カ ?   S: ルイ   ルイ   ルイ
U: 0-4
S: シツメン オ カシ シマズ   1980.03.28   14:23:43
テーザース CAS   (V.084N.01-V.092N.08)   1,734,219   ㌥
カワ ルイ   34
[ 13 U: IMMOBILIZATION
S: 2087   ㌥
[ 23 U: ENZYME
S: 23034   ㌥
[ 33 U: 1*2
S: 887   ㌥
[ 43 U: YE IMMOBILIZE
EXPAND IMMOBILIZE
NO   EOI   EAI(0)   ㌥
/01   3   1 IMMOBILIZE
/02   1755   895 IMMOBILIZED
/03   1   0 IMMOBILIZES
S: ショウリョク   オカシマシ
[ 43 U: /02*2
S: 746   ㌥
[ 53 U: ¥0

***   テーザース   シツメン   シマカ   カワク   クソク   ¥OFF   カスク   ¥OFF   クソク   リョウケン
***   CAS   00:02:00   0   ㌥   0   ㌥   0   ㌥   ¥594

***                                     ㌥474   ¥4,734

S: ツキノ シツメン カ アリマスカ ?   AI=Y, IAI=N
U: N

***   テーザース   シツメン   シマカ   カワク   クソク   ¥OFF   カスク   ¥OFF   クソク   リョウケン
***   JICST   00:23:00   27   ㌥   0   ㌥   0   ㌥   ¥4,140
***   CAS   00:02:00   0   ㌥   0   ㌥   0   ㌥   ¥594

***   カワ   シマカ   00:25:00                                     ㌥474   ¥4,734

S: カワ   ショウリョク   シマズ   1980.03.28   14:25:54

```

図 8 CAS ファイルの検索例

この検索例は、[2] ENZYME 23,034 件と [1] IMMOBILIZATION 2,087 件および [4] IMMOBILIZE の expand (E コマンド) で検索された IMMOBILIZED のそれぞれの組合せの結果と、会話の終了部分を示した。

ノイズの入ることは避けられないのが欠点である。

④ Methanol dehydrogenase

⑤ データベース : CAS

検索式

[1] METHANOL 5167 件

[2] DEHYDROGENASE 9582 件

[3] BACTERIA 8610 件

[4] 1*2*3 3 件

[5] 1*2 30 件

この 30 件のなかのノイズは、脱水素酵素作用に対するメタノールの影響の報文で、飼料に与えたメタノールが動物臓器の各種脱水素酵素活性の変動への影響の報文

```

*****      1980/03/27  10:15:01      *****
          サ - ビスを開始します
          (サ - ビス終了予定時刻   12:00:00)
****      J I C S T 漢字オンライン検索サ - ビス      ****
¥IN JOIS,08J-0200,TEST
S: パスワードは ?
U:
      J I C S T 漢字オンライン検索サ - ビス
      記号 データベース      蓄積範囲      蓄積件数      終了予定時刻
      A      理工学文献 (1978.04 -1980.04)  745,252 件      12:00
      B      研究案内 (ON-GOING-HISTORY)  19,208 件      12:00
S: データベースの記号は ? A,B,無し=N
U: A
      データベースは理工学文献
      番号      蓄積範囲      蓄積件数
      S (1980.04)      32,555 件
      0 (1980.04-1980.04)  32,555 件
      1 (1979.04-1980.03)  361,297 件
      2 (1978.04-1979.03)  351,400 件
S: データベースの範囲は ? S,番号-番号
U: 0-2
S: サ - ビスを開始します 1980.03.27  10:16:36
      データベースは理工学文献(1978.04-1980.04)  745,252 件
      会話番号      1

```

図 9 JOIS-K の 検 索 例 (1)

図 6 と同じであるが、漢字ひらがなの混じり文でわれわれの日常用語と同じであるために、カタカナ文よりもはるかに見やすい。活字の大きさも大きい。

が含まれた。

(b) 生体成分の例

◦ Acyl CoA

データベース : CAS

[1]	ACETATE	24,986 件
[2]	PROPIONATE	1,010 件
[3]	BUTYRATE	488 件
[4]	COENZYMEA	タグガアリマセン
[5]	COA	1,021 件
[6]	(1+2+3)*5	133 件

(c) 代謝の例

◦ リグニンの分解

データベース : CAS

[1]	BIODEGRADATION	354 件
[2]	LIGNIN	2,280 件
[3]	BENZOATE	4,672 件
[4]	PHTHALATE	1,460 件
[5]	1*2	17 件
[6]	1*3	1 件
[7]	1*4	6 件

以上の例はデータベースとして CAS ファイルを用いた生化学的分野のものであるが、CAS では化学の全分野の項目について検索できるし、理工学文献ファイルではさらに広い範囲にわたって検索できる。

[5] U: 3*4

S: 1 件

[6] U: ¥P

#001 577.1:576.8.095.1:((546+547, [LC150508]

C79322923

Rhodopseudomonas capsulataのニトロゲナ-ゼ生成に及ぼすL-メチオニン-D,
L-スルホキシイミン(MSX)およびβ-N-オキサリル-L-α,β-ジアミノ
プロピオン酸(ODAP)の影響 (a②) EN

Effects of L-methionine-DL-sulfoximine and β-N-oxalyl-L-α,β-diaminopropionic
acid on nitrogenase biosynthesis and activity in Rhodopseudomonas capsulata.
MEYER J, UIGNAIS P M :B118A Biochem Biophys Res
Commun (USA) 89 (2) (353-359) ('79) 1

標記光合成細菌では、グルタミシンテターゼ(I)の阻害剤は、NH₄⁺の存在下で
ニトロゲナ-ゼ(II)の生成を低下させる。II生成の新しい低下剤ODAPを、
広く用いられるMSXと比較した。両方の化合物とも、Iの阻害とII生成の低下の
間に定量的な相関関係がみられた。両化合物は、標記細菌のin vivoのII活性を阻害す
るが、C. pasteurianumのIIは阻害しない； 写図2表2参18

Rhodopseudomonas属、ニトロゲナ-ゼ、生成、代謝阻害、
グルタミシンテターゼ、酵素阻害剤、メチオニン、プロピオン酸S

S: 終了しました

JICST-COPYRIGHT

図 10 JOIS-K の 検 索 例 (2)

[1]	アミノサン	9,840 件
[2]	オキシドレダクターゼ	5,974 件
[3]	1*2	498 件
[4]	RHODOPSEUDOMONAS ゾク	171 件
[5]	3*4	1 件

の検索結果をオンライン出力したもの、最後の2行に示された論文に付与されたディスクリプタは
8個。これらと検索に用いたキーワード(ディスクリプタ)とを比較するとおもしろい。

第1回に述べたように、本講座の目的はオンライン検索の実際を解説することであった。よし、俺(私)もこの機会にぜひ一度やってみようと思っていたけれど、目的の大半は達せられたことになる。皆さんの所属機関に端末機がなかったり、登録していないときには、最寄りのJICST支所(サブセンター、表8)へ電話をかけてみて下さい(注3)。そこにある特定端末機を利用して一度検索を御自分でぜひやってみて下さい。もし端末機があれば、JICST、東大、筑波大学の学術情報センター(次回以降にて説明する)、あるいは丸善、紀伊国屋へ電話をされて手続方法を確かめて下さい。取扱機関によりシステムが全く異なるので、それぞれのマニュアルを読んで下さい。

次回に、より詳しい検索法、検索式の検討等とともに利用者の専門分野からみた問題点、システムへの要望等

(注3) JICSTに登録された大学関係機関は、表9に示したとおりであるが、地区別にみると、まだ全国的に普及していない。55年4月現在で、北海道地区2大学(5端末、東京支所)、東北地区10大学(44端末、東京支所)、関東地区29大学(41端末、東京支所、筑波支部)、中部地区16大学(18端末、名古屋支所)、関西地区16大学(18端末、大阪支所)、中国地区5大学(6端末、中国支所)、四国地区2大学(2端末、大阪支所)、九州地区9大学(12端末、九州支所)となっている。東北地区にとくに端末機が多いのは、東北大学に31台登録されているためである。

について述べる.

付 記

1. データファイルの特徴, 利用の仕方等については説明を省略した.
2. 本稿で用いた下記の用語は本文中でゴシック体で示したところに簡単な解説を入れた.

データベース, データファイル, アップロスティング, プロファイル, コマンド, ディスクリプタ, 非ディスクリプタ, 端末装置, 検索タグ.

次の用語は説明をしていない. それは MEDLARS ファイルの検索に関係したもので, 今説明の必要は少ないと考えたためである.

MeSH トリー番号 (MeSH Tree Number), qualifier.

その他の解説のしていない特殊な用語は次回に説明をする予定である.

謝辞 JOIS に関してご親切に説明をしていただいた日本科学技術情報センター筑波支部長三橋宏氏および支部長代理梶原康敬氏に深く感謝をいたします.

- (6) 国立大学図書館協議会図書館機械化調査研究班編: 大学図書館の機械化, 紀伊国屋書店, 1979, p.157.
- (7) 武田敬一: ドクメンテーション研究, **29**, 473 (1979)
- (8) 磯本征雄, 角戸正夫: 情報管理, **21**, 105 (1978).
- (9) F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Bice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tatsumi: *J. Mol. Biol.*, **112**, 535 (1977).
- (10) 日本科学技術情報センター: JOIS 利用のてびき, 第2版 (1978).
- (11) 荒木啓介, 内田尚子, 森田 朗: 情報管理, **19**, 508 (1976).
- (12) 稲葉安養子, 情報管理, **20**, 49 (1977).
- (13) 斎藤和男, 諸岡 徹, 情報管理, **20**, 143 (1977).