

Title	東大大型計算機センターにおけるオンライン情報検索
Author(s)	山本, 毅雄
Citation	大阪大学大型計算機センターニュース. 1976, 22, p. 67-73
Version Type	VoR
URL	<a href="https://hdl.handle.net/11094/65316">https://hdl.handle.net/11094/65316</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

Osaka University

# 東大大型計算機センターにおける オンライン情報検索

東京大学大型計算機センター 山本 毅 雄

## 1 TOOL-IR の概要

TOOL-IR システムは、主として学術文献データをオンライン（ディスク）データベースとし、これに対してTSS, リモートバッチ, オープンバッチ, クローズドバッチなど種々の方法で検索のできるシステムである。<sup>(1-3)</sup>

これは、東大センター研究開発部と東大理学部藤原研の共同研究として、科研費の補助をうけて研究・開発されてきた。現在、

- ① Chemical Abstracts Service (CAS) の化学文献情報
- ② Crystallographic Data Centre (XDC) の結晶学文献および数値情報
- ③ Institution of Electrical Engineers (IEE) の計算機・制御関係文献情報の三種類のデータベースをもち、東大センターの全利用者に公開されている。

これまで多くの文献情報システムに比較して、次のような点が特徴といえよう。

◦ TSS 処理・バッチ処理で、共通のコマンド・サブコマンドが使える

Table I に TOOL-IR の利用者用コマンド（主コマンド）、Table II に同じくサブコマンドを示す。Table III には、これらのコマンドを、どのような処理法で利用できるかを示す。これら種々の利用法を組合せて、たとえば検索質問の作成や、すぐに結果を得たい少量の検索は TSS で、大量の文献リストの出る検索はバッチ処理で、と使い分けることができる。

Table I TOOL-IR の主コマンド（データベースを選択する）

機能	名称	入力例	対象データ
データベース検索	CAS 遡及文献	CAS	CA Condensates 最新巻全部
		// CAS 80	CA Condensates 80巻全部
		// CAS 81	CA Condensates 81巻全部
	CAS 新着文献	NEWCAS	CA Condensates 最近2号
XDC 文献	XDC	// XDC	XDC 文献ファイル（最新版）全部

Table II TOOL-IRのサブコマンド(CAS, NEWCAS用)

名 称	機 能
SEARCH	一次検索。内容・著者名等で検索し、文献集合を作る。
AND	二次検索。文献集合間の積集合 ( $A \cap B \cap \dots$ ) を作る。
OR	二次検索。文献集合間の和集合 ( $A \cup B \cup \dots$ ) を作る。
DIF	二次検索。文献集合間の差集合 ( $A \cap \bar{B}$ ) を作る。
DISPLAY	文献集合の内容表示。
QSAVE	一連の質問を、ファイルにしまう。
QUSE	一連の質問を、ファイルからとりだして使う。
KEEP	文献集合をファイルにしまう。
ENTER	文献集合をファイルからとりだして使う。
EXTRACT	文献集合の一部を抽出して別の文献集合を作る。
DUMP	巻・号・文献番号を与えて内容表示。
REMINDE	文献集合に対応する質問の内容表示。
END	検索を終了する。

Table III TOOL-IR システムの利用

種 類	入出力用ハードウェア	利 用 の し か た
TSS 処理	固定端末 (タイプライタ端末, ビデオ端末)	端末のあるところへゆき, 会話形式で検索する。
	電話端末 (主にタイプライタ端末)	端末を音響カップラーで電話につなぎ, システムを呼出して会話形式で検索する。
リモート・ バッチ処理	リモート・バッチ端末 (カードリーダー, ラインプリンタ)	端末のあるところへゆき, カードリーダーから検索コマンドをよませる。処理がすんだ後, ラインプリンタから結果を受取る(コンソールタイプライタの操作が必要)。
オープン・ バッチ処理	カードリーダー, ラインプリンタ	センターのオープン入力室のカードリーダーから検索コマンドをよませる。処理がすんだ後, オープン出力室のラインプリンタから結果を受取る(トークンカードを差込めばよい)。
クローズド・ バッチ処理	カードリーダー, ラインプリンタ	検索コマンドの入ったデッキに計算依頼カードをそえて, センターに郵送する。結果はセンターから郵送される。

◦ 検索質問を自分のファイルにたくわえておける

TOOL - IR コマンドでは、会話形式で複雑なロジックを含む一連の検索質問を作成しこれを QSAVE サブコマンドで自分のファイルに貯蔵することができる。

この質問は、ファイル編集コマンドを使って TSS 処理またはバッチ処理で修正でき、QUSE サブコマンドで呼び出して使うことができる。したがって、検索質問を、何回のセッションにもわたってゆっくり作成することができる。また、一度作成した検索質問でも、対象分野の進歩や自分の興味の発展に応じて、次第に作りかえてゆくことが容易である。

◦ 検索結果を自分のファイルにたくわえておける

KEEP サブコマンドによって、複雑な質問の結果えられた文献のリストを、自分のファイルにたくわえることができる。このリストは、ENTER サブコマンドによってよび出し、内容を印刷したり、さらにくわしい検索の出発材料として使うことができる。

◦ 検索質問ファイル、検索結果ファイルの貸し借り、やりとりができる

PERMIT コマンド、SHARE コマンド、TRANF コマンド、RECVF コマンドなど、システムに既に備わっているコマンドを使って、検索質問ファイルや検索結果ファイルを貸し借り、あるいはやりとりすることができる。これによって、たとえば遠隔地間の研究者同志の協力や協同作業が、ファイルを媒介にして簡単にできる。

◦ システムの質問ライブラリが使える

専門外の分野を検索するとき、検索質問の模範例あるいは参考として、その分野の専門家が作ってシステムに登録した検索質問を、QUSE サブコマンドで呼び出して使うことができる。

## 2 CAS データベース<sup>(8)</sup>

Chemical Abstracts Service (CAS) は米国化学会の下部機構であり、1907 年以来、化学および関連分野の国際的抄録誌 Chemical Abstracts を刊行してきた。1968 年以来、ここで抄録される全文献についての情報を、Chemical Abstracts Condensates 磁気テープの形で頒布するようになった。内容は、抄録された文献の書誌事項（論文・特許・レポートの表題、著者・特許出願者名およびその所属、掲載誌名及び巻・号・頁その他）、および抄録中から抽出されたキーワードフレーズである。抄録対象は、125 ケ国、14,000 種の科学・技術関係逐次刊行物、26 ケ国の特許、各国の会議議事録・学位論文・政府公報・単行本等で、これらの中から化学又は化学技術に関連した新しい情報を含むものを抄録している。しかし、“化学及び化学技術に関連する”という定義はたいへん広く解釈されており、素粒子のレベルから反応機器の製造までが含まれる。そのため、化学はもとより物理学・薬学・医学・生物学・農学・化学工学など、多くの隣接分野で基本的な

情報源となっている。頻度は週1回で、年52本の磁気テープが送られてくる。抄録文献数は30~40万件/年である。

TOOL-IRシステムでは、これからデータ切出しルーチンによって著者名、表題中の語、キーワードフレーズ中の語、CASセクション番号(CASで与える大分類の番号)等のデータを切り出し、これを編集して“内容語索引ファイル”と“著者名索引ファイル”を作る。一方、磁気テープ内容はそのまま再編集して“線型ファイル”を作る。これらのファイルを用いて、CASコマンド、NEWCASコマンドが実行される。

現在、東大センターでは、2週間に1回(テープ2本分)データベースの更新をおこない、1974年1月以来のChemical Abstracts文献約100件が検索可能になっている。

そのうち、線型ファイルにデータがあり、くわしい内容表示が可能なものは約30万件である。

Fig.1にはTSS端末からの検索例(下線部が入力)、Fig.2には検索結果の内容表示の一部を示す。

### 3 XDC データベース<sup>(4)</sup>

Cambridge Crystallographic Data Centreはイギリスのケンブリッジ大学化学教室に1965年に設立されたもので、科学教育省の科学技術情報オフィスOSTIより財政援助を受けている。また世界的にも結晶学データセンターとして認められており、干渉法による(X線、中性子線を用いる)結晶構造解析結果の集成、検索、評価、合成および流通を行っている。

TOOL-IRで使用しているXDCデータベースはこのデータセンターで作成された磁気テープによるもので次の内容を持ち、同所との契約により、日本における学術研究の目的のための使用を認められている。

(a) 文献ファイル 1935年より現在までの有機および有機金属化合物の結晶構造解析の文献を次の内容で記載してある。内容は6週間毎に更新され、絶えず最新の状態に保たれている。件数は現在約14,000件で、6週間に約150件ずつ増加する。内容は化合物名、

```
TYPE IN COMMAND
@SEARCH RAT
```

```
SEARCH RAT
*OR1*
KEYWORD='RAT
DOCUMENT SET #1 CREATED
CONTAINS 272 DOCUMENTS
```

```
TYPE IN COMMAND
@SEA MOUSE, MICE
```

```
SEA MOUSE, MICE
*OR1*
KEYWORD='MOUSE
KEYWORD 'MICE
DOCUMENT SET #2 CREATED
CONTAINS 56 DOCUMENTS
```

```
TYPE IN COMMAND
@AND 1, 2
```

```
AND 1, 2
DOCUMENT SET #3 CREATED
CONTAINS 7 DOCUMENTS
```

Fig. 1 TSS 端末からの検索例

TYPE IN COMMAND  
@DIS M. D

DIS M. D  
DISPLAY DOCUMENT SET #8  
SET CONSISTS OF 20 DOCUMENTS  
DIGESTS OF LAST 20 WILL BE DISPLAYED

V/I/N: CAS VOLUME/ISSUE/ABSTRACT-NO  
A: AUTHOR C: CODEN V:VOLUME P: PAGE Y:YEAR  
T: TITLE  
K: KEYWORD PHRASE

V/I/N: 79/25/142847M A: PUJIC, D. C: PRTIAV(J) V: 54 P: 200, 202\* Y: 73  
T: EFFECT OF QUAT ON PRODUCING HIGHER WEIGHT INCREASES AND  
BETTER FOOD VALUE OF PIGS  
K: QUAT PIG FATTENING/BROMOBENZYLAMINE PIG RESPIRATION  
PROPHYLACTIC/FEED ADDITIVE PIG QUAT

V/I/N: 79/25/143110C A: OSTERMANN, L. ET AL. C: DTTIAF(J) V: 80 P: 254-7 Y:73  
T: COURSE AND CONTROL OF THE SEXUAL DEVELOPMENT OF FEMALE  
PIGS. III.  
EXPERIMENTS TO INDUCE OVULATION AND TO OBTAIN OVA IN  
15-WEEK-OLD GERMAN LANDRACE SOWS  
K: OVULATION GONADOTROPIN SOW/OVARY DEVELOPMENT SOW  
GONADOTROPIN/ESTRADIOL OVULATION SOW

#### Fig. 2 結果表示の例

同義名，分子式，著者名，掲載ジャーナル巻，頁，年，化合物分類コードからなり，1960年以前の文献についてはデータが“Structure Report”，Suttonの“Interatomic Distance”の表に掲載されているかどうか，1960～1965の文献についてはLandolt - Bornsteinの表に掲載されているかを記してある。

- (b) 数値データ・ファイル 1960年以後現在までに発表された有機化合物・有機金属化合物の結晶構造データを記載してあり，6ヶ月毎に更新される。件数は現在約11,000件である。内容は結晶の単位セルの大きさ（およびその標準偏差），スペースグループ，非標準の対称操作，単位セル中の分子数（分子式数），密度の計算値，密度の測定値（およびその標準偏差），測定温度，融点，原子座標，結合の長さ，結合データ等である。なお，内容がそれ自身 consistent であるかどうかをUNIMOLというプログラムを用いる計算でチェックしてある。この点で，データファイルの内容は文献に記載されたものよりも更に信頼できるものになっている。（ミスプリントを除いてある）

TOOL - IR システムでは，切出しルーチンを使ってこれから化合物名，データ入力年月日，分子式，著者名，文献の雑誌コード・巻・号・頁・発行年，化合物分類コード等のデータを切り出し，これを転置ファイルとしてXDCコマンドで検索できるようにしてある。検索の結果得られる化合物 - 文献コード（REFCODE）を使って，さらにくわしい

計算結果を得るための XCALC コマンド, 分子構造を端末やプロッターに出力するための XPLOT コマンドなどを利用することもできる。

Fig. 3 に, XPLOT コマンドを使ってプロットさせたグルタミン酸の立体投影図を示す。

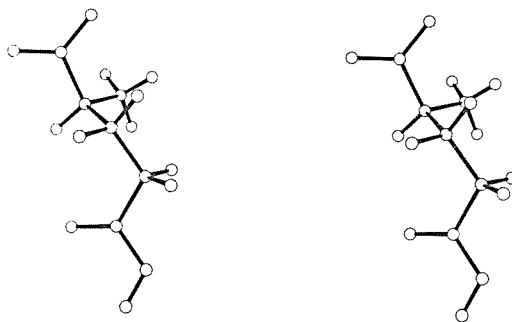


Fig. 3 検索結果を使ったグルタミン酸の立体投影図 (右眼で右図を、左眼で左図をみると立体的に見える)。

#### 4 IEEE データベース

英国電気工学者協会 Institution of Electrical Engineers は,

INSPEC 文献データベース (Electric and Electronic Abstracts, Physics Abstracts, Computer and Control Abstracts) を発行している。このうち, 東大センターでは現在 Computer and Control Abstracts を購入し, 1976 年 1 月以降の分をオンライン検索可能としている。このデータベースは, 抄録を含んでおり, TOOL - IR では抄録中の語もふくめて検索の対象となる。

#### 5 おわりに

TOOL - IR システムは, 特定研究 “広域・大量情報の高次処理 (総括代表者: 島内東大センター長)” の “化学における情報処理” 班 (代表者: 藤原鎮男教授) 及び開発 - C 班 (代表者: 藤原教授) によって研究・開発された。プログラムの大部分は, 根岸正光 (東大・経・大学院), 牛丸 守 (東大センター), 野上法正 (東大・理・大学院) 及び筆者の手になるが, その他多数の方々にプログラミング, 辞書の検討, システム評価, 助言, ドキュメンテーション, 質問ライブラリ作成等さまざまな方面で御協力頂いている。われわれの主な目的のひとつは, なるべく多数の研究者にオンライン情報検索システムの利用経験をもって頂き, その間に自然発生的に利用者間での知識の共用・交換が行われるようにすることであったが, この目的も次第に達成されつつあるように思う。特に阪大には, 遠方にもかかわらず熱心な利用者の方々がおられ, 活潑なグループをなして活動されているのはシステムを運営するものにとって何よりも嬉しいことである。今後とも, より多くの利用者の方々に, このコミュニティーに参加して頂ければ幸いである。

## 参 考 文 献

- (1) T. Yamamoto, M. Negishi, M. Ushimaru, Y. Tozawa, K. Okabe, S. Fujiwara, "TOOL-IR: An On-Line Information Retrieval System at an Inter-University Computer Center", Proc and US-Japan Computer Conf., p. 159(1975)。
- (2) 根岸正光, 山本毅雄, "オンライン文献情報検索システム TOOL-IRにおけるマン・マシン・インターフェース", 情報処理 17, p. 402(1976)。
- (3) 東大センター TOOL-IRプロジェクト, "TOOL-IR利用者マニュアル(上(第2版), 下)", 1976。
- (4) 島内武彦, 山本毅雄, "結晶構造データとその検索", 特定(1)広域大量情報開発C班, 1975。