

Information System, Data Bases, and On-Line Services of the Japan Information Center of Science and Technology (JICST)[†]

HISAKO UCHIDA

Technical Coordination Office, The Japan Information Center of Science and Technology, 5-2, Nagatacho 2 Chome, Chiyoda-ku, Tokyo, Japan

Received June 15, 1979

The JICST information processing system consists of the data-base production system, authority file management system, bibliographic retrieval system, and printed issue compiling system. The bibliographic retrieval service based on the JICST On-line Information System (JOIS-I) has been available through leased line since 1976 and now also through dial-up line, which covers five data bases: the JICST bibliographic and on-going research information files, *CA Condensates*, MEDLARS, and TOXLINE files. The on-line output in Japanese kanji is also available. The newly revised JOIS-II system is now being developed.

I. INTRODUCTION

The Japan Information Center of Science and Technology (JICST) is a governmental information service organization established in August 1957. Since then it has published regularly a series of Current Bibliography on Science and Technology as one of its activities to disseminate current worldwide bibliographic information throughout Japan. Since 1968, the JICST information system has been developed. Including about 350 000 citations processed in 1978, the JICST bibliographic data base is the largest data base now in Japan.

Since 1972, SDI service has been provided using the JICST bibliographic data base as well as CaCon and MEDLARS data bases introduced from the U.S.A. Then in 1976, JICST began to provide on-line information retrieval service through the JICST On-line Information Retrieval System-I (JOIS-I), and the JICST on-line network has been extended steadily throughout the country. This paper describes the present and future situations of the information processing system developed by JICST, the JICST data bases, and its on-line information retrieval service.

II. JICST INFORMATION PROCESSING SYSTEM

Figure 1 shows the present JICST information processing system and various kinds of products generated therefrom. This system consists of the following: (1) data-base production system, (2) authority file management system, (3) bibliographic retrieval system, and (4) printed issue compiling system.

1. Data-Base Production System. The manuscript of each citation (including bibliographic items, abstract, and index terms) is entered into the system and galley proofs of input data are checked against corresponding manuscripts. Some of the data are then checked automatically by using either authority files or check digits, and detected errors are corrected interactively. Data from the authority files are added to each checked and corrected citation to complete each data-base record. The data to be added are: the UDC numbers corresponding to the JICST classification codes; the journal abbreviation, CODEN, and ISSN corresponding to the JICST journal code; and the phonetic description and all broader terms of the descriptors assigned by indexers.

2. Authority File Management System. (a) Vocabulary Control System. Descriptors are registered into the vocabulary file, tree structures of descriptors are generated from the hierarchical relation between terms, and the frequency data of the descriptor usage are stored. The vocabulary file is updated by periodic corrections and additions. In the data-base production system, this file serves as an authority file for

Table I. Data Bases Available in the JOIS-I Service

name of data base	inclusive dates (to present)	file size (as of March 1979)
JICST (bibliographic)	1974-	1 460 000
JICST (on-going research information)	1977-	15 000
CaCon	1974-	2 010 000
MEDLARS	1974-	1 300 000
TOXLINE	1974-	540 000
Total		5 325 000

validity checking and data addition in relation to the assigned descriptors. The JICST thesaurus is compiled based upon this file.

(b) Source Document Control System. This system is designed to determine the state of overall procedures from the acquisition of document to the processing conditions, such as being abstracted, in bindery, or on loan.

The source document master file, the core of this system, serves as the authority file in order to add data on each document at the stage of data-base production. The JICST holding list is compiled based upon this file.

3. Bibliographic Retrieval System. This system consists of batch and on-line systems. All of the data bases, including the data bases introduced from other organizations, are converted into the file format to be used in this system. The present on-line information retrieval system (JOIS-I) is designed to be simple and easy to use, so it fills most of the users' demands using only 10 commands.

4. Printed Issue Compiling System. This system compiles the data to make ready for photocomposition or printout in appropriate formats for each product, such as abstracts journals, holding list, thesaurus, and SDI announcements. About 2800 characters including kanji, hiragana, katakana, alphanumerics, and punctuations are available in the system for those products.

III. DATA BASES FOR ON-LINE SERVICE

The data bases available in the JOIS-I service, the number of citations, and inclusive dates of each data base are shown in Table I.

The JICST bibliographic data base covers about 8000 titles of domestic and foreign journals and other source documents in the fields of science and technology. About 40% of the citations come from the fields of chemistry and chemical industry.

Index terms are controlled by the JICST thesaurus, which contains about 30 000 descriptors (subject terms in Japanese). Common chemical substances are indexed by their specific names as descriptors, but most chemical substances are indexed

[†] Presented at the Division of Chemical Information, ACS/CSJ Chemical Congress, Honolulu, Hawaii, April 5, 1979.

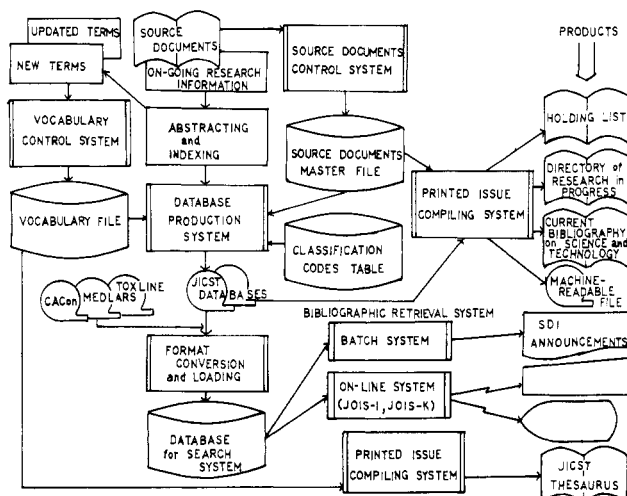
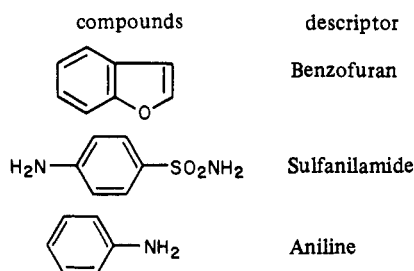
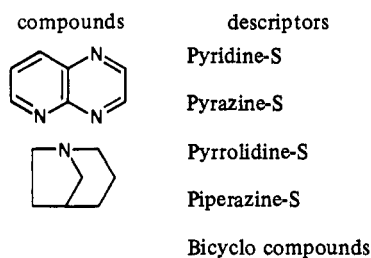


Figure 1. JICST information processing system.

by combining more than one descriptor representing their chemical features as shown in the following examples. Examples of organic compounds indexed with specific names:



Examples of organic compounds indexed by combining more than one descriptors:



(The terminal "S" of a descriptor indicates that the substance is one of the derivatives of the compound which the descriptor represents).

One of the characteristics of the JICST data base is that each citation contains all the broader terms added automatically to each specific descriptor assigned by indexers. Therefore, the broader term only will be sufficient for executing a generic subject search. It means that there will be no need of combining more than one narrower term subjected to that term by "OR" operators. For specific names of chemical compounds as the descriptors, their broader terms representing their structural features are added automatically; therefore, a certain substructure search is possible by using these broader terms. "Sulfanilamide", an example shown above, will be retrieved when "sulfonamide", its broader term, is used as a search term.

The JICST on-going research information data base contains the information on about 15 000 on-going research projects of governmental and other public research institutions in Japan. It covers the same scope of fields as the bibliographic data base mentioned above.

The data bases are stored by yearly accumulation and the most current data can also be used separately for SDI service.

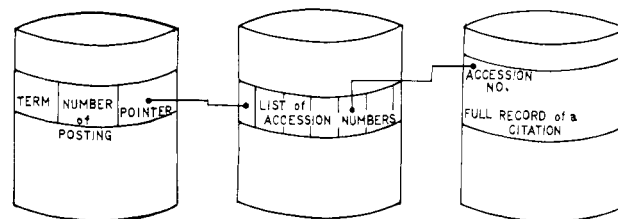


Figure 2. File organization of the JOIS-I data base.

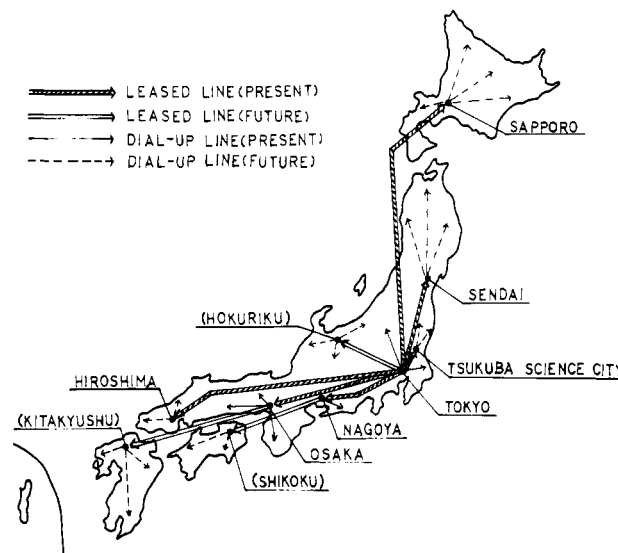


Figure 3. JICST on-line network.

```

CC= 541002, (CA010407)
CN= B78040886, C78100014
TI= AN INTERACTIVE COMPUTER GRAPHICS SYSTEM FOR PROCESSING CHEMICAL
    STRUCTURE DIAGRAMS.
AU= BLAKE J E; FARMER N A; HAINES R C
JN= A294AB, J CHEM INF COMPUT SCI
VN= VOL.17, NO.4, PAGE.223-228, 77
CI= (1) (A1) (EN) (USA) (9, 12)
KW= 化学情報システム; 化学情報システム; 化学情報システム; CAS
  
```

Figure 4. Output sample of JOIS-I.

When starting a search, the users can specify the range of years they prefer.

The file organization of the JICST on-line data base is shown in Figure 2. Directly searchable data elements and indirectly searchable elements of the data base are listed in Table II.

IV. PRESENT SITUATIONS OF ON-LINE INFORMATION RETRIEVAL SERVICE

Three years have passed since JOIS-I service started in April 1976. Besides the services through the leased line, dial-up service became available in July 1978, and the number of calls per month has been greatly increased. The number of terminals connected to JOIS-I is about 270 as of April 1979. As shown in Figure 3, the JICST on-line network covers almost all the districts of Japan, i.e., from Sapporo to Hiroshima, and is expected to extend to Kyushu within 1979.

The on-line service using kanji characters (JOIS-K) through the leased line is also available now. Figures 4 and 5 show the samples of output from JOIS-I and JOIS-K, respectively.

V. FUTURE OUTLOOK FOR THE JICST DATA BASES AND ON-LINE SERVICE

The present JICST data bases (bibliographic and on-going research information) cover the fields of science and technology, but it is planned to expand its coverage to medical science, agriculture, and biology in the future. If should be mentioned here that for the Japanese users information services

Table II. Searchable Data Elements of the JOIS-I Data Base

data elements	directly search- able	indirectly search- able
keyword	X	X
title and/or abstract word (CACON, TOXLINE)	X	
classification code or CAS section number or MeSH tree number	X	X
author's (investigator's) name	X	X
author's affiliation or location of work		X
name of research institution	X	
code indicating the type of research institution	X	X
citation number		X
publication type		X
JICST journal code or CODEN or ISSN		X
type of document (e.g., review, patent)		X
year of publication or ending date of investigation		X
country of publication or country of patent application or location of institution		X
language		X
entry date		X
type of budget		X

in Japanese language are preferable particularly when the output of abstract texts becomes available on-line, because users will be required to read the abstract texts on the display of the terminal quickly and to proceed to the next step, judging whether the given citation is relevant or not. Therefore, JICST is required to give efforts to meet various kinds of the Japanese users' needs relating the JICST data bases. The JICST thesaurus is updated every year and better quality of indexing will be expected. New data elements, such as affiliations of authors and other data, are to be added to the data base in the near future.

At present, almost all the data in the JICST data bases are given in the Japanese language; therefore, it is rather difficult for non-Japanese people to utilize the data bases. To solve this problem and make the information produced in Japan available for overseas as well, bilingual processing is being

#001 54:002, (CA010407)

B78040886, C78100014

化学構造の図を処理するための相互作用型計算機グラフィックスシステム (aΦ)
EN

An interactive computer graphics system for processing chemical structure diagrams. BLAKE J E, FARMER N A, HAINES R C
:A294A J Chem Inf Comput Sci (USA) 17 (4)
(223-226) ('77) 1

CASの出版物に使用できる品質の化学構造の図を作り得る表題グラフィックスシステムを開発した。IBM370/168ホスト計算機に連結したデジタルエキップメントコ-ポレーション(DEC)PDP-15グラフィックスシステムに基づき、利用者は端末からライトペンによりメニューから原子または環を選択し、結合を指示すると、構造図の優先位置が計算される。入力された構造は蓄積され、テキストと共に出版物用にAutologic APS-4で写真製版される。このシステムは2年以上CA出版物の作成に使用されている。基本的環系と構造図との二つの参照ファイルが作られている。写図9参12

グラフィックスシステム、分子構造、CAS、図形処理、化学構造検索

Figure 5. Output sample of JOIS-K.

considered as a future plan. The JICST data bases adopt the distribution format which conforms to ISO 2709 to make it internationally exchangeable. The descriptions of bibliographic data elements such as journal abbreviations, ISSN, country code, and language code also conform to respective ISO standards. The description of author names is also standardized in accordance with the recommendations in the UNISIST Reference Manual.

JOIS-II, the revised system of JOIS-I, is now being developed. It will have following features: (1) to extend the coverage of data bases to the maximum 99; (2) to make authority files (JICST vocabulary file, MeSH file, CHEMLINE file, etc.) searchable; (3) to offer services to locate the source journals using the source documents master file; (4) to increase the directly searchable data elements, such as journal codes, names of organizations, and segments of chemical substance names; (5) to make the output of abstract texts available and to provide more output format options; (6) to increase the number of commands to three times as many as JOIS-I; (7) to store search questions for later uses; (8) to make the output in kanji characters available through dial-up connection of terminals; (9) to provide the JICST on-line service news and to receive users' comments; and (10) to make the output of system messages in English available.

Adapting the Gmelin Handbook to Modern Information Requirements[†]

WALTER LIPPERT

The Gmelin Institute, 6 Frankfurt/Main, West Germany

Received April 26, 1979

The current 8th Edition of the "Gmelin Handbook of Inorganic Chemistry" comprises 380 volumes, devoted to a systematic description of the entire knowledge of all chemical elements and compounds. Continuation of the Handbook in its historic form is no longer feasible; to preserve the Handbook's timeliness a number of editorial changes have taken place. Supplement volumes devoted to fields of chemistry where important developments are occurring, expanded use of English, and a Formula Index to allow quick user access and to prepare for computer-storage and -retrieval ensure that the Gmelin Handbook will continue to hold its place as an indispensable source of critically evaluated, basic information.

What should a chemist do when information is needed on a specific chemical element or a given inorganic compound,

such as its occurrence in nature—its geochemistry or cosmochemistry—or whether a given compound really exists at all, and, if it does, what are its structural properties? Many possibilities are available for answering these questions. One

[†] Presented at CSJ/ACS Chemical Congress, Honolulu, Hawaii, April 1-6, 1979.