

- (3) Crane, E. J., "Words and Sentences in Science and Industry," *Science* **86**, 549-53 (1937).
- (4) Crane, E. J., "Abstracting and Indexing," *J. Chem. Educ.* **20**, No. 11 (1943).
- (5) Crane, E. J., "CA Today," *Chem. Eng. News* **26**, 2190 (1948).
- (6) Crane, E. J., "Scientists Share and Serve," *Chem. Eng. News* **29**, 4250 (1951).
- (7) Crane, E. J., "Chemical Abstracts Yesterday, Today, and Tomorrow," *Proc. Chem. Soc.*, p. 334, Dec. 1957.
- (8) Bottle, R. T., "Use of the Chemical Literature," 231 pp., Butterworths, London, 1962.
- (9) Burman, C. R., "How to Find Out in Chemistry," 220 pp., Pergamon, Oxford, 1965.
- (10) Mellon, M. G., "Searching the Chemical Literature," 40 pp., ACS, Washington, D. C., 1964.
- (11) Mellon, M. G., "Chemical Publications," 4th ed., 324 pp., McGraw-Hill, New York, 1965.
- (12) Gould, Robert F., Ed., "Searching the Chemical Literature," *Advan. Chem. Ser.* **30**, 326 pp. (1961).
- (13) Bottle, R. T., "Use of The Chemical Literature," 231 pp., Butterworths, London, 1962.
- (14) Tate, F. A., "Progress Toward A Computer-Based Chemical Information System," *Chem. Eng. News* **45**, 78-90 (1967).
- (15) Davenport, W. C., "CAS Computer-Based Information Services," *Datamation*, pp. 33-8, March 1968.

Retrieving Chemical Information with *Index Chemicus**

G. S. REVESZ and ANDREA WARNER

Institute for Scientific Information, 325 Chestnut St., Philadelphia, Pa. 19106

Received October 23, 1968

By communicating in the international language of structural diagrams, *Index Chemicus* provides for rapid scanning and retrieval of information about chemical compounds via subject, author, and molecular formula indexes. In addition, data on instrumental analyses and proposed biological and other uses of compounds are presented. *Index Chemicus* has been extended to provide for computer retrieval and identification of compounds via line notations of structural formulas, and coverage of newly reported reactions of old compounds is now provided.

The basic objective of *Index Chemicus* is to present information about new chemical compounds and reactions in a simple and quickly understandable form, and to make this information easy to retrieve through a variety of indexing methods.

Published weekly, *Index Chemicus* contains an average of 350 abstracts which describe about 3000 new compounds, gathered from the world's chemical literature. Averaging 80 pages per issue, the publication functions as both a secondary journal and a unique registry of new chemical compounds.

Index Chemicus, *IC*, *Index Chemicus Registry System*, *ASCA*, *OATS*, and *ISI* are registered trade marks of the Institute for Scientific Information. Articles published in *IC* are selected from a basic list of 180 regularly scanned journals and meet one criterion: each contains newly synthesized or elucidated chemical compounds. If an article meets this criterion, it is abstracted, regardless of the country of origin, language, or specific field of chemistry.

The journals regularly processed by *Index Chemicus* are scanned page by page by a team of chemists as soon as they appear in print, some even as they appear in page proofs.¹ These are primary chemical journals, which probably include close to 99% of all published new compounds.²

In addition to these regularly-indexed journals, the *Automatic Subject Citation Alert (ASCA)*, another service of the Institute for Scientific Information, provides articles

from 2000 additional journals in allied fields. This input yields the remaining few compounds indexed. The articles from a particular journal appear in *Index Chemicus* between 35 and 45 days after the journal is received.

Each abstract of *Index Chemicus* contains the following information:

- a. Bibliographic data (author's name and affiliation, mailing address, journal citation, received date, and language)
- b. Summary (if supplied by author)
- c. Identification of new compounds by structural diagrams and molecular formulas; reaction flows are included when relevant
- d. Proposed uses of compounds, if tests and accompanying data are presented
- e. Instrumental data and indication of new reaction, when appropriate
- f. Author's approval of abstract

In addition, a monthly index is provided, which is cumulated semiannually and annually for retrospective searching by:

- g. Author
- h. Subject
- i. Molecular formula
- j. Journal
- k. Rotaform³

Optional services available to the reader include

- l. Magnetic tape search
- m. Tearsheet or reprint of original journal article
- n. Magnetic tapes with structural diagrams in line notation form

*Presented before the Division of Chemical Literature, 156th Meeting, ACS, Atlantic City, N. J., September 13, 1968.

USE OF INDEX CHEMICUS: AN EXAMPLE

Consider a typical article to illustrate what information can be derived from it (Figure 1). In the upper left hand corner appears the number 91041. This is a unique identification number through which compounds, authors, and subjects can be identified in the respective cumulative indexes which appear monthly and are cumulated semi-annually and annually. This is a serial number which started with 00001, when the first *Index Chemicus* issue appeared, and reached the 100,000 mark in issue 250, in August, 1968.

Next to the article number appears the title of the primary article in English. All foreign language titles are translated by the indexer or by the translating department of the Institute. The author and coauthors' address at which the work was performed, the mailing address of the first author if different from the first address, journal citation, date when the original paper was received by the primary journal, and the language of the original article complete the bibliographic data. Whenever an article provides an author's summary, it is reproduced in its original language.

The most important part of the abstract is the identification of new compounds. Every new compound has a running number, which, together with the article number, uniquely identifies it in the cumulative indexes. Each compound also has a diagram reference number which ties it to a given structural diagram, thus defining the compound at a glance without complicated chemical nomenclature. Diagrams can be easily scanned, making *Index Chemicus* not only a very effective searching tool,

- - 91040	BENZODI-CYCLOHEPTENES (4.1.0), SYN	91147
91710	BENZODI-THIAZOL-5-ONES, SYN	91140
- - 91106	CYANINE	91141
91420	BENZODI-THIAZOL-5-ONES, ADDN IR + CL	91210
- - 91420	BENZODI-THIAZOL-5-ONES, SUBST	90848
- - 91664	BENZODIAZEPINES, 5-DM-1H-1,4-, SYN + REARR	91231
NO3) 91504	7-CL-5-CYCLOHEXYL-1-ME-2-OXO-	- - 91041
91587	2,4-DI-H-	- - 91041
- - 91926	BENZODIAZEPINES, 7-CL-5-CYCLOHEXYL-1-ME-2-OXO-	- - 91041
ES 91663	DI-H-2H-1,4-, SYN	91787
- - 91712	BENZODI-THIAZOL-5-ONES, 1,3,2-	- - 91932
91713	BENZODI-THIAZOL-5-ONES, AMINOKETONYL-1,4-	90782
	BENZODI-THIAZOL-5-ONES, AMINOKETONYL-1,4-	- - 90924
	COUMARINS	- - 90924

Figure 2. Subject Index

but also a quick scanning tool for the chemist with limited time for browsing.

In this particular example, the researcher scanning for new developments in benzodiazepines can immediately spot what is of interest to him by looking at the structural diagrams without reading a single printed word. Conversely, he can look in the subject index under the heading "Benzodiazepines" (Figure 2), which will, through the number 91041, lead him back to the article of interest. From the given flow diagram, he can also ascertain the method by which these compounds were prepared. In the same way, he can look up the molecular formula of a specific compound of interest (Figure 3), which, again through the number 91041, will lead him to this article.

Index Chemicus indexes all intermediates, many of which are not isolated or analyzed by the author. Many compounds which would otherwise be missed are thus brought to the reader's attention.

The Use Profile, Data Alert, and New Reaction Alert (Figure 1) are refinements which were added to *Index Chemicus* after the original concept had become well

UNIQUE NUMERICAL IDENTIFICATION SERIAL NUMBER: **91041**

AFFILIATION AND ADDRESS: SCHMITT J., COMDY P., SUQUET M., BOITARD J., LEMEUR J., BASSELIER J. J., BRUNAUD M., SALLE J., CENTRE RECH, ETABLISSEMENT CLIN-BYLA, PARIS, 5.

ORIGINAL JOURNAL CITATION: CHIM THER 2(4), 254-9 (1967). RECD JUNE 3, 1967. IN FRENCH.

AUTHOR'S SUMMARY: Preparation of 7-chloro-5-cyclohexenyl-1-methyl-2-oxo-2,3-dihydro-1(1H)-1,4-benzodiazepine (tetrazepam), a compound which is different from Diazepam because of the substitution of the phenyl radical by the cyclohexenyl radical. The synthesis is effected in four steps from the corresponding cyclohexenyl compound, such as N-chlorination, intern rearrangement of chlorine at the top of connection of cyclohexenyl, deshydrohalogenation and finally methylation of lactam nitrogen. The structure of the final product and intermediates is established without ambiguity. The value of the product, chiefly as a myorelaxant one, comes out from the results obtained on the animals, results that are given in comparison with those supplied by the diazepam and from the clinical experimentation, undertaken on a large scale, for which an abstract is given.

TESTED COMPOUND USES: **USE PROFILE**

COMPOUND NUMBER: 1) C15 H16 CL2 N2 O (II), 2) C15 H16 CL2 N2 O (III), 3) C15 H15 CL N2 O (IV), 4) C16 H17 CL N2 O (IVB)

MOLECULAR FORMULA: CNS ACTIVITY SPASMODIC ACTIVITY

DIAGRAM REFERENCE: (I), (II), (III), (IV), (IVB)

TITLE OF ARTICLE: A NEW MYORELAXANT IN THE BENZODIAZEPINE CLASS: TETRAZEPAM.

AUTHOR AND CO-AUTHORS: SCHMITT J., COMDY P., SUQUET M., BOITARD J., LEMEUR J., BASSELIER J. J., BRUNAUD M., SALLE J.

INSTRUMENTAL DATA ALERT: UV, IR, NMR

INDICATION OF LANGUAGE OF ORIGINAL ARTICLE: IN FRENCH

DATE RECEIVED BY ORIGINAL JOURNAL: RECD JUNE 3, 1967

STRUCTURAL DIAGRAMS: (I), (II), (III), (IV), (IVB)

INDICATION OF AUTHOR'S APPROVAL OF ABSTRACTS: author reviewed

SUBSTITUENTS: A) R H, B) R CH3

Figure 1. Typical entry in the *Index Chemicus*

N2 0.....91787-5 N 06.....91781-35
 07.....90826-1 N3.....91757-55
 N 0.....90778-9 N3 0.....91948-20
 N 5.....91257-11,19 N2.....91796-32
 2 05.....91751-6 N3 03.....91481-25/91940-18
 N4 04.....90923-36 N3 06 52.....91755-26,27,28
91063-6 N5.....91765-8,13
 5.....90951-1 N5 04 52.....91493-16
 52.....91498-14,15,16,17 N5 05 5.....91493-14,15
 53.....91498-30,31,32 N5 06.....91493-16
 2 5.....91498-7,8,
 3 5.....91498-11,12
 1 5.....91574-
91571-
 1.....91498-23,25,26,2
 22.....90808-3/91948-14
 34 5.....91002-
 17 52.....91002-
 14.....91364-3,6,11,1
 16.....91720-1,2,4,
 C15 H12.....91469-
 N2 02.....91489-34/91935-10
 1 0.....90849-11 N2 02 5.....91137-2,12
 02.....91422-7 N2 03.....90777-8/91750-22/
 3 0.....91852-6/91948-8 91757-3/91798-8,9,
 02.....91422-4 10/91822-3
 06.....91422-5,9 N2 03 5.....91779-2
 5.....91257-3 N2 04.....91749-20,21
 3.....91948-18 N2 04 52.....90773-1
 3 0.....91948-6 N2 06.....91654-19

Figure 3. Molecular formula index

established. Many people posed the question as to whether compounds prepared were found useful for some kind of practical application, such as dyes, detergents, drugs, etc. Thus, the Use Profile was conceived to alert the reader that test data for some kind of activity, whether negative or positive, had been disclosed by the author. Such activities are not only listed in the article itself, but can also be found as subject entries in the subject index (Figure 4). When compounds are considered to have possible biological activity, but no tests were performed, the Use Profile is not indicated. However, there will appear in the subject index an entry for that activity followed by the word "Potential." A very complete search can thus be performed by a quick reference to the subject index.

Instrumental analysis has been advancing from its infancy into adulthood. It is becoming more and more important to the average reader; thus, the Instrumental Data Alert (Figure 1) indicates to the reader which types of analyses were performed during the course of research for a particular article. In this typical article, infrared, ultraviolet, and nuclear magnetic resonance data were presented, as shown by the unscreened portions of the Instrumental Data Alert.

The center of the disc is reserved for alerting the reader to a new chemical reaction. In this article, no new reaction was used for the preparation of compounds, and the center of the disc remains screened.

ACCURACY OF THE INFORMATION

Every article indexed undergoes several checking stages during the production phase.⁵ Each molecular formula is checked by the indexer and the number of double bonds ascertained. Sometimes this is facilitated by the HydroBond computer, a circular slide rule.⁴ This is a rapid method for determining whether a formula is correct. Before the copy is printed, every article is proofread for accurate title translation and correct keypunching of all other bibliographic data. Every formula is compared to the diagram it represents. All additional information, including activities and analyses used, is verified directly from the journal.

CINEOLIC ACIDS, RXNS - - - - -91898
 CINNAMANILIDES, HYDRO-, C-14 LABELED 91326
 CINNAMIC ACIDS, CMEZCN ESTER - - -91197
 CINNOLINECARBOXYLIC ACIDS, ARYL-, 91160
 CINNOLINES, - - - - -90897
 ALKYL-NH- - - - -91895
 REDN - - - - -91160
 CINNOLINONES, CARBALKOXY-, TRI-F- - -91620
 CLATHRATE ADS, SILVER OXIDE-HF2, - - -90854
 D LABELED - - - - -91620
 CLAUSENA BREVIYSTILA, AMIDES - - -90854
 CNS ACTIVITY - - - - -91041
 - - - - -91791
 CNS ACTIVITY, - - - - -91411
 POTENTIAL - - - - -91489
 CNS DEPRESSANT ACTIVITY - - - - -91763
 - - - - -91769
 - - - - -91787
 - - - - -91795
 CNS DEPRESSANT ACTIVITY, POTENTIAL - - -91446
 CNS STIMULANT ACTIVITY - - - - -91741
 COCLAURINE, 3',5'-DI-TERT-BU-, - - -90931
 DEHYDROGENATN - - - - -90969
 COENZYME Q1, F- - - - -91456
 COMPLEXES, - - - - -91456
 AG-THIOUREA - - - - -91456
 AG-AU, CD, CO, CU, FE, HG, NI, PD, PT, W, - - -91456
 ZN-PENTA-F-BENZENETHIOL - - -91456
 AL-(HALO)-ACAC - - - - -91518
 AL-PYROCATECHOL VIOLET - - - - -91719

Figure 4. Biological activity shown in subject index

As a final assurance that the information presented is correct, a proof of each article is sent to the first author, together with multilingual instruction sheets asking him to check everything in the abstract and return it to the editor. He is asked on the envelope to return the proof with his comments, corrections, and signature immediately. If he complies with this request, an "Author Reviewed" label is affixed to the article.

Requesting the author to proofread the abstract provides editorial control, the advantages of which are not immediately obvious. Mistakes or omissions by the author and errors in printing in the primary journal are often discovered only after the receipt of an author proof. Typical errors of this type include incorrectly drawn structural diagrams, errors in molecular formulas, or inconsistencies between formulas and their corresponding names or diagrams. Thus, *Index Chemicus* can claim the distinction of often being more accurate than the original article indexed. Standardizing the method of presentation provides also the additional benefit of eliminating all confusion resulting from the diverse and often conflicting editorial policies (or lack of such policies) followed by some journals.

If an author proof is delayed in the mail and does not reach the editorial office before publication, any corrections contained therein which were not made during the editing stage can still be applied to the cumulative indexes.

ADDITIONAL INFORMATION SERVICES AVAILABLE

All of the information contained in *Index Chemicus*, except summaries and structural diagrams, is keypunched on punched tab cards and is available on magnetic tape as well as in printed form. A scientist who has the tapes and who wants to initiate a search, for example, on all articles containing new compounds tested for spasmolytic activity which were also analyzed by infrared spectroscopy, can run a computer search on *Index Chemicus*' magnetic tape file.

A reader may be interested in knowing what a certain author has recently published in a particular field of interest. For this purpose, he can search the author index in conjunction with the subject index and find all articles by an author on any given subject. In the article described, one could search for any one of the authors, e.g., Lemeur,

J., and determine the article numbers associated with his name. He would then look under "Benzodiazepines" in the subject index. A comparison of article numbers would lead him to the article in question.

If a reader desires even more information about a particular article, an OATS postcard, available from ISI upon request, with the ISI accession number, can be sent to the Institute for Scientific Information, and a copy of the original article will be furnished to the reader by return mail.

INDEX CHEMICUS REGISTRY SYSTEM

As mentioned above, structural diagrams are not contained on tab cards or magnetic tape. Because demand is growing for substructure searching by computer, it has become imperative to adopt a computer-usable language to depict structural diagrams. The Wiswesser Line Notation was chosen by ISI, from among several such languages proposed for the *Index Chemicus Registry System*, because of the relative simplicity of this notation, its adaptability for computer searching, and the ease by which these notations can be retranslated into molecular formulas and even structural diagrams.

The more than 150,000 new compounds indexed annually in *Index Chemicus* are available in the line notation form on magnetic tape and computer printouts from the beginning of 1968. We plan to convert the *Index*

Chemicus back file of 800,000 compounds into this linear notation in the near future.

A major new feature has been added to *IC* for 1969. Starting with the first issue of the year, all articles containing reports of new chemical reactions are being indexed, even though these reactions are for the synthesis of old rather than new compounds. Wherever possible, the synthetic flow diagram will be shown, and the end products of a reaction which are old compounds will be clearly indicated in the *IC* indexes.

The new reaction code will also appear in the magnetic tapes for the *Index Chemicus*, along with the codes for instrumental and analytical methods.

LITERATURE CITED

- (1) Elias, A. W., G. S. Revesz, and G. H. Foeman, "Effects of Mechanization on a Chemical Information Service," *J. CHEM. DOC.* 8, 2 (1968).
- (2) Garfield, E., and G. H. Foeman, "Statistical Analyses of International Chemical Research by Individual Chemists, Languages and Countries," Presented at the Div. Chem. Lit., ACS, Chicago, Ill., September, 1964.
- (3) Sher, I., J. O'Connor, and E. Garfield, "Rotaform—A New Index for Chemical Searching of Chemical Compounds," *J. CHEM. DOC.* 4, 49 (1964).
- (4) Sher, I., A. W. Elias, and E. Garfield, "Control and Elimination of Error in ISI Services," *J. CHEM. DOC.* 6, 132 (1966).
- (5) Sher, I., G. H. Foeman, E. H. Baus, "A Slide Rule for Calculation of the Number of Double Bonds and Hydrogen Atoms," Presented at the 145th National Meeting, ACS, September 1963.

The IDC System for Chemical Documentation

ERNST MEYER

Badische Anilin- & Soda-Fabrik AG, 67 Ludwigshafen am Rhein, Germany

Received November 12, 1968

After a decade of systems development, firms of the European chemical industry founded a corporation (IDC) to make the chemical journal and patent literature accessible by efficient computer methods. A comprehensive and critical review of these methods accommodated to the four most important types of chemical data is given.

Chemistry probably more than any other science builds on results which are experimentally reproducible and important even after decades. Compounds described many years ago may be today's starting material and have to be synthesized again; their properties, use, and preparation are continually of interest.

The chemical industry, therefore, requires especially good documentation systems and has been willing to invest monies to avoid unnecessary duplication of experimental effort and to have easy access to the information needed. Numerous documentation services are offered and bought; yet they cannot satisfy all needs and have to be expanded and improved constantly.

The chemical industry not only purchases commercially available services, but widely indexes journal and patent literature in addition to their internal reports. The individ-

ual companies, however, in being confronted with the task of coping with the rapidly growing amount of literature, have been considering far-reaching improvements.

This was accomplished by developing new methods of documentation, making full use of modern computer techniques, and by consolidating firms of similar interests—i.e., by founding a joint corporation for centralized documentation to avoid duplication of input effort. After having developed and tested methods covering important areas of chemistry, the International Documentation in Chemistry (IDC) was founded in the spring of 1967¹ to index important areas of the journal and patent literature and to supply the corresponding magnetic tapes for searches. This paper reviews the documentation approach since, up to now, only excerpts have been published.

A really efficient documentation system has to meet