

Chemical Information for the
Man Who Has Everything

Number 16, April 17, 1978

Most *Current Contents*[®] readers would be surprised to learn that chemical information retrieval was, so to speak, my first love. Many associate me exclusively with *CC*[®] or the *Science Citation Index*[®] and do not know that I have long been preoccupied with problems of chemical documentation. Indeed, in a sense, ISI[®] has been "competing" with *Chemical Abstracts* for almost twenty years.

This may help explain the mixed emotions I felt about the Herman Skolnik award I received about a year ago. This award is presented annually by the American Chemical Society's Division of Chemical Information for "outstanding and sustained service to chemical information." While I'm pleased to have this recognition, I'm sure that those who are not chemical information "insiders" believe the award reflects only my work on *CC* and the *SCI*[®]. On the other hand, chemical information specialists have never fully appreciated the importance of *CC*!

ISI has been building its chemical data base since 1960.¹ It was originally designed primarily to alert chemists to new organic compounds reported in the journal literature. Later on, new reactions

and new syntheses were also covered. About 150,000 new compounds are added to the data base each year. This is estimated to include 90% of all the new compounds reported in the literature since 1960. Although neither ISI's nor any other system available is absolutely complete, I doubt that much of *significance* is missing.

Over the years, the ISI chemical data base has been continually improved and made available in an increasing variety of formats. It is now accessible through print and microform publications, computer-generated reports, and magnetic tapes—all of which are functionally integrated and cross-referenced to each other. This flexibility means that a wide variety of organizations, ranging from national institutions to multi-national companies to comparatively small institutes, are able to make use of the data base in the form best suited to their needs.

One of the problems we have in promoting the data base is dealing with the names of the different services we extract from it. The names tend to be long, not self-explanatory, and easily mistaken for other services produced by ISI and other organizations. We made

things even more complicated by twice changing the name of one service. So, here for the first time in one place in *CC* are the correct names of our chemical information services with very brief explanations of what they do.

Current Abstracts of Chemistry™ (CAC™) is a weekly printed service that contains graphic abstracts of the articles reporting new compounds (including intermediates), reactions, or syntheses. Emphasis is on ease of browsing, through liberal use of structural diagrams, and rapid coverage of the literature. Most compounds are reported in *CAC* within 45 days after their initial appearance in the journal literature.

Index Chemicus® (IC®) is now the index section to the abstracts. It was the original title of the entire service. *IC* is included as part of the weekly *CAC*. It is also issued separately in quarterly and annual cumulations. *IC* provides a variety of ways to search the file of abstracts—by molecular formula, author, journal, organization, and subject term searches. The quarterly and annual cumulations also include a rotated molecular formula index.

The next three services I will describe involve the use of Wiswesser Line-Notation (WLN).² We have pioneered in the use of this linear notation since 1968. WLN describes compounds unambiguously. It also shows how the substituents are arranged in relation to each other. Since WLN uses standard alpha-numeric characters, it is ideal for input to computerized

"substructure searching" systems is also compatible with manual searching. A substructure search enables you to locate all compounds reported that contain specific molecular substructure fragment.

Chemical Substructure Index (CSI) is another index to the compounds reported in *CAC*. *CSI* is an alphabetic list of permuted WLN descriptions of compounds.³ Issued monthly with a semi-annual and annual cumulation, *CSI* has been in microfilm format since 1976. The years 1966-1975 are covered in printed indexes. It is shocking how few chemists make use of this incredibly simple and useful tool covering over a million compounds.

Automatic New Structure Alert® (ANSA®) is a computerized personal alerting service similar to our *ASCA®* system. *ANSA* automatically selects the articles reporting new compounds that match a pre-established interest profile. In addition to the familiar types of profile terms which can be used (e.g., author, key words, molecular formula, etc.), WLN descriptions of substructures can also be employed. Subscribers receive a printout or a month providing the bibliographic descriptions of articles reporting the compounds matching the profile. Substructure searching is done by direct computer scanning of the linear WLN notations and/or scanning connectivity tables generated from the WLN.

Index Chemicus Registry System® (ICRS®) is the complete chemical data base supplied

magnetic tapes. Issued monthly with an annual cumulation, the tapes can be used for current awareness or retrospective searching. In addition to WLN descriptions of compounds, *ICRS* tapes contain, as an option, fragment code descriptions based on the so-called Ring Code of the *Pharmadokumentationsring* group of companies.⁶ A subscription to *ICRS* includes search software at no extra cost.⁷

The reprint of "Where is Chemical Information Going," which appears on the following pages, is the acceptance talk I gave when I received the ACS award.⁸ I originally entitled the talk, "For the Man Who Has Everything." This was my way of acknowledging the irony of receiving such an award from the American Chemical Society—an organization with which my disagreements have been long and loud. But if the man who has everything could *really* have his way, he would want ISI's chemical

services to come out from under the shadow of their more famous siblings, *Current Contents* and the *Science Citation Index*.

In closing I might add that while they are a small and select group, those chemists and pharmacologists who do make use of *Current Abstracts of Chemistry and Index Chemicus*TM are as strongly addicted to it as most of you are to *CC*. This is partly because they realize that searching *Index Chemicus* for current information is a logical step before or after using *Chemical Abstracts* or *Beilstein*. They also know that a search of the *SCI* is an important way to supplement searches done with the other tools.

In the near future ISI is going to make a concerted effort to expand the use of our chemical services. We hope to accomplish this through substantive changes that will make the services better able to meet the increased need for the retrieval of specific chemical data.

REFERENCES

1. Garfield E. ISI's chemical information system goes marching on! *Current Contents* (1):5-6, 5 January 1976.*
2. Gibson G W & Granito C E. Wiswesser chemical line-notation. *Amer. Lab.* 14:27-37, 1972.
3. Granito C E & Rosenberg M D. *Chemical Substructure Index (CSI)*. A new research tool. *J. Chem. Doc.* 11(4):251-6, 1971.
4. Garfield E. Introducing ANSA—ISI's Automatic New Structure Alert—a compound-retrieval service for people more interested in compounds than retrieval! *Current Contents* (19):5-10, 9 May 1973.*
5. Garfield E & Sim M. The *Index Chemicus Registry System*—past, present and future. *Pure Appl. Chem.* 49:1803-5, 1977.
6. Garfield E. ISI's *CHEMTRAN* 'compatibilizes' files of encoded chemical structures. *Current Contents* (46):5-6, 15 November 1972.*
7. The retrieval and dissemination of chemical information. 4. *ICRS RADIICAL* software. *Current Contents* (31):M1-2, 5 August 1970.*
8. Where is chemical information going? *J. Chem. Inform. Comput. Sci.* 18(1):1-4, 1978.

*Reprinted in: *Essays of an Information Scientist*. Philadelphia: ISI Press, 1977. 2 vols.