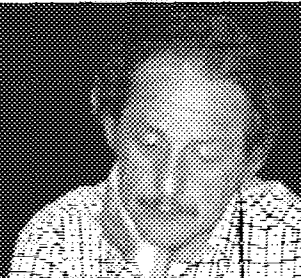


ISI's Chemical Information  
System Goes Marching On!



January 5, 1976

Number 1

It will come as a surprise to most readers of *Current Contents*<sup>®</sup> (*CC*<sup>®</sup>) to hear me say this. But there is an information service (which ISI<sup>®</sup> produces) whose readers are even more devoted and loyal than *CC*'s. When it first began in 1960, it was called *Index Chemicus*<sup>®</sup> (*IC*<sup>®</sup>). In 1970 we renamed it *Current Abstracts of Chemistry and Index Chemicus*<sup>™</sup> (*CAC/IC*<sup>™</sup>).

*Index Chemicus* has changed quite a bit in its sixteen years' existence. And in 1976 that evolution continues. Indeed, it is impossible to talk about *IC* alone anymore. One must talk about the entire ISI chemical information system, especially the *Index Chemicus Registry System*<sup>®</sup> (*ICRS*<sup>®</sup>). Today it is sometimes difficult to decide whether *ICRS* is a by-product of *CAC/IC* or vice versa. And the recent adoption of our *CHEMTRAN* software by the European Pharmaceutical Industry Documentation Ring makes that even more apparent! *CHEMTRAN* software converts Wiswesser Line Notation (WLN) into fragment codes.<sup>1</sup>

Sometime early in 1975, *CAC/IC* indexed its 2,000,000th compound. By the last issue of 1975, we had reached 2,135,086. Each of these new compounds was drawn from its 'primordial' article--the one that first de-

scribed its synthesis, isolation or whatever. Even the intermediates never indexed by *Chemical Abstracts* (*CA*) were picked up. It's a real compliment to us to learn that *CA* has now changed its policies to include such intermediates. We don't kid ourselves into thinking that any organization which needs *CAC/IC* will use it as a replacement for *CA*. So why shouldn't we applaud an improvement of this kind in a competitive service?

Why are *CAC/IC* readers so loyal? It takes a lot of loyalty to shell out over \$1000 each year for a service that is limited in its coverage to articles on *novum organum*--to be more exact, new organic compounds. We have no trouble finding 150,000 of them every year in about 20,000 articles.

But it is precisely this selectivity which is important to these readers--the rest is simply timing and format. Chemists like to read structural diagrams. *CAC/IC* is loaded with them. They are also interested in methods of synthesis and we cover all new ones, even if by chance the paper didn't include any new compounds.

We continually examine the usefulness of the way information is presented in *CAC/IC*. In 1976 we are introducing several changes. We are adding certain journals to its coverage

which are expected to yield several thousand new compounds each year. More important is a new *Instrumental Data Index*. Users of CAC/IC are, of course, familiar with the instrumental 'data disk' we have used to signal analytic methodology in our graphic abstracts. The new index recognizes the importance of easy and rapid access to instrumental technology and its application. This new index will appear in each weekly issue.

In addition, we are creating two new indexes by separating chemical and biological terms which previously appeared together in the subject index. Terms used as entries in the new separate biological-activity index will be of much the same type as before. Entries in the chemical index will be shorter so that it will be easier to identify new compounds as members of a particular, desired category. Although these index entries for chemical compounds will be shortened forms, the full description of a new compound will be seen in the abstract itself. All compounds will continue to be fully indexed by WLN in ISI's *Chemical Substructure Index*<sup>®</sup> (CSI). The CSI is also undergoing a significant change. In 1976 we are testing the use of microfilm as a means of providing continuous cumulations. Thus each new monthly *Chemical Substructure Index* will contain all compounds for the previous months until six months are cumulated. Each new batch of compounds for that month will be asterisked.

All of the WLN information to-date is available on magnetic tape.

We have already encoded 1.4 million compounds in this system and plan to finish our files in the next few years. We have not yet encoded compounds reported from 1960-65. These 550,000 compounds will increase the size of the file to 2,000,000. By the end of 1978 the file will exceed 2,500,000.

A vital part of this system is our CHEMTRAN software which enables you to convert line notations to particular fragment codes via our connectivity table. These tables can be used for conversion to other systems as well.

There is a large economic incentive on the part of the world's chemical and pharmaceutical industries to use chemical information systems. Finding one relevant compound can save thousands of dollars in research and may have significant patent implications. It is unfortunate that academic institutions do not yet use these systems as widely as they could--if for no other reason than as a means of training chemists who will go into industry.

I should like to suggest that this unique file could be made available to the American academic and industrial communities by cooperation between ISI, the American Chemical Society, the National Science Foundation, and appropriate private organizations. This cooperation could take many forms, too numerous--and advantageous--to mention here.

---

1. Garfield E. ISI's CHEMTRAN 'compatibilizes' files of encoded chemical structures. *Current Contents* No. 46, 15 November 1972, p. 5-6.