OPINION PAPER

Opportunities for Alternative Suppliers of Secondary Chemical Information†

PETER F. URBACH[1]

Pergamon International Information Corporation, McLean, Virginia 22101

Received September 26, 1983

INTRODUCTION

Chemical Abstracts Service (CAS) today is *the* database producer in the field of chemistry. There are no other producers of secondary chemical information of any consequence. This statement has very significant implications for the present and future international flow of chemical information. Jim Seals, CAS's Director of International Programs, is quoted[2] as pointing to a number of "competing" services, specifically Derwent, ISI, and Viniti. These are all no doubt useful services in their own right, but they are not realistic alternatives for Western chemists to use to search the worldwide chemical literature. There is in fact only one comprehensive source for secondary chemical information, CAS. The CAS bibliographic file is the third most widely used on-line bibliographic file in the world.[3] Following major files in law and medicine, the CAS files are the most used. Not just in the field of chemistry but in all fields. Essentially all chemical information services worldwide use the CAS file, link up to it, or build upon it to provide their own products and services. There can be no doubt that the information products and services of CAS are *the* source of secondary chemical information worldwide.

There are, of course, alternative means of access to this information. The chemical files on the Dialog Information System are perhaps the most widely used and incorporate a number of important features developed by Dialog, but they are still the Chemical Abstracts Service files, licensed by Dialog from CAS. Similarly, the innovative DARC Substructure Search System developed by Prof. J. Dubois and available on the Telesystemes Questel service is based on the registry files of Chemical Abstracts Service. The specialized files of the National Library of Medicine are licensed from CAS, and the Chemical Information System of the EPA is based in large part on registry information originally licensed by the EPA from CAS. The files of IDC in Germany are yet another example. There can be no doubt that CAS is *the* international database for bibliographic information on chemical literature, and therefore, one must look at CAS in some detail to explore the question of the international flow of chemical information.

GROWTH OF CAS

Chemical Abstracts Service evolved over a period of 75 years from a small American service that covered only the U.S. chemical literature to a worldwide service covering all the chemical literature of the world. Today, nearly 75% of the material processed by Chemical Abstracts Service is of non U.S. origin. Further, and this may be somewhat surprising, the majority of the users and most of the usage are outside of the U.S.[4]

How did this upstart service, which at the turn of the century had to compete with the much more influential German *Chemisches Zentralblatt* and other more established abstracting services, gain such a dominant position? Clearly, there are a number of complex factors that contribute to answer this question. We can identify at least four key factors.

First, there is the importance of information in the field of chemistry as contrasted with its importance in other fields. This gave Chemical Abstracts Service an edge over services in the fields of Physics, Psychology, and Mathematics in competing for federal research and development dollars.

Second, there is the acceptance of English as the universal language of science which gave Chemical Abstracts Service a decided edge over the German, French, and other non English services. A true worldwide scientific information service today must be an English-language service.

The third key factor was the massive post Sputnik investment of U.S. Government research and development money, primarily from the National Science Foundation, into Chemical Abstracts Service from the late 1950s through the mid 1970s. The government's investment of about 30 million dollars was nearly matched by CAS funds and used by CAS to build the foundation for the automated information systems that are the cornerstone of CAS leadership today. The research and development funds given to CAS were greater than similar investments made by the government in other subject-oriented secondary information services and greater than investments made by other countries in similar services.

The fourth and final key factor that cannot be underestimated is the quality of the CAS management and its dedication over the years to product quality and comprehensiveness.

The result of these four factors, the critical importance of information in the field of chemistry, the dominant role of the English language in science and technology, the post Sputnik NSF investment in CAS, and the quality of CAS management, contributed in the last 30 years to catapult CAS into the dominant position it has today in the field of secondary chemical information.

CENTRALIZATION VS. DECENTRALIZATION

CAS management has always expressed a desire for an international decentralized information system. However, as the computerization of Chemical Abstracts Service became

† This paper was presented as part of a symposium on International and Transborder Flow of Chemical Information at the 1983 Annual Meeting of the American Chemical Society, Washington, DC.

0095-2338/84/1624-0001$01.50/0

more pervasive, operations changed from the relatively decentralized approach envisioned by management to a highly centralized operation. This can be seen in two areas—the shift from external volunteer abstractors to internal document analysts and the shift from cooperation with sister societies abroad to internal processing.

Prior to the high level of internal computer processing, external volunteer abstractors did some selecting and the preparation of the bulk of the abstracts. The volunteer corps represented contributors worldwide who undertook the abstracting effort not for the nominal honorarium but as a labor of love to aid in keeping abreast of the literature and to make their own contribution to effective communication of chemical information. At the height of the volunteer abstractor program in the mid 1960s, there were nearly 3300 volunteer abstractors, more than half of whom were outside the U.S.[5] Thus, there was a true international network with substantial transborder data flow.

The growth of computerized processing within CAS, the need to maintain stricter quality control on the work product, and the desire to produce a more timely product than the international network of volunteers would permit led to the gradual dismantling of the network and its replacement with corps of internal document analysts at CAS. Today, 300 document analysts in Columbus, OH, perform the bulk of the work previously performed by the international network. Thus, the processing of the raw material for CAS has gone from a decentralized international system to a highly centralized local operation.

A similar trend can be seen in the relations between CAS and its sister societies in other countries. In the early 1960s, the managers of CAS recognized the importance of international cooperation in collecting and processing worldwide chemical information. They sincerely sought to establish an international network, making arrangements with The Chemical Society (now the Royal Society of Chemistry) in the U.K., the Chemical Society in Germany, and an organization in Japan that was established expressly for the purpose of working with CAS. They also sought to make similar arrangements with CNIC and its predecessors in France. The notion was that each of these national institutions would be responsible for collecting, screening, and processing the chemical information in its own area and forwarding the resulting processed information to Columbus. In return, the participating national institutions were to have exclusive marketing rights for the CAS products and services in their territories.

Unfortunately, the far-sighted and statesman-like plan never worked properly. It is now essentially dismantled with worldwide processing responsibility once again centralized in Columbus. The growth of computerized processing systems in Columbus, the economies of scale resulting from the huge automated processing factory developed at Columbus, and the desire to maintain high quality and timely service led to further centralization rather than the planned and hoped for decentralization. With the exception of the cooperation with Japan, where because of special problems with the language cooperation is still essential, the notion of decentralized input is no longer being pursued.

There are some among the European sister agencies of CAS who believe that their decision to participate in the unsuccessful international cooperation, providing their input to the CAS product, led to the decline of their own national services. The present domination of the American service in world chemical information is a matter of considerable concern to the Europeans, whose own chemical information systems and services are now completely dependent upon the products provided to them by CAS. We see some of the CAS "partners" getting nervous about their continued access to CAS files, and we have seen in at least two areas the shift from decentralized international cooperation to centralized processing with a single central information service dominating the secondary chemical information field, worldwide.

ON-LINE ACTIVITIES

The last 3 years have seen one further significant shift in the role of Chemical Abstracts Service that will radically affect the international flow of secondary chemical information. In the last 3 years, CAS has become an on-line vendor for CAS files. This move is a logical extension of CAS activities and one that was debated within the councils of CAS for at least 12 years. It is a major step. CAS is no longer just a database producer licensing its electronic information product to a variety of vendors to provide public access to the information. It is now a vendor in its own right, competing with its own licensees to provide users with computerized access to its files.

This action on the part of CAS has obviously made the other on-line vendors of bibliographic information very nervous. They now have a new competitor, a competitor who is in a special privileged position. This uniquely positioned vendor, CAS, has announced that its offering will be more complete than that of its licensee competitors. Abstracts are available on the CAS service and are not available to others. Thus, the established on-line vendors who have been providing access to the CAS File for the past decade now face competition from the database owner, who will retain for itself a superior product. The concern of the other on-line vendors in both the U.S. and Europe is certainly understandable.

The Amerian Chemical Society explains that there are very good reasons for the actions being taken by CAS. They are, to quote from a letter written earlier this year by John Crum, the Executive Director of the American Chemical Society, "... to provide a single, comprehensive, online system that would be superior in its utility for scientists and engineers worldwide to any online system available today; and, to increase the probability that CAS can maintain its economic viability during the uncomfortably rapid migration of use from printed to online services".[6] A clear statement of purpose, but hardly one to reassure sister chemical societies around the world or competing vendors that a true cooperative international service is evolving.

The pattern for the international flow of chemical information appears now to be well established. Primary literature flows from all over the world to Columbus. From there, after being processed in the highly automated computer factory, the secondary information flows out throughout the world in the form of printed product and magnetic tape from Columbus or via computerized access to the computers in Columbus.

Chemical Abstracts Service is now making plans to expand beyond bibliographic and chemical structure information into chemical data. CAS hopes to cooperate with a number of international data organizations to develop decentralized computerized information systems, an international network of scientific and technical information utilizing the software and capabilities developed by CAS. It is too early to predict how these plans for a decentralized system will evolve. We have only the history of the previous decentralized cooperative efforts to guide us in suggesting that the end result may prove to be less decentralized and less cooperative than the present planners hope.

OPPORTUNITIES

Now, with all that has gone before as background, what are the opportunities for alternative suppliers of secondary chemical information? Given that there is a single worldwide service that dominates the scene, that it is capable of operating in-

dependently, without cooperative links internationally, and that it has recently become an aggresive on-line marketer of chemical information, a competitor to all other on-line chemical information services, what opportunities remain? The answer is rather bleak. Clearly, no one, with the possible exception of the Russians, can afford to abstract and index the world's chemical literature anew. Therefore, there are only two choices—operate under license from CAS, which itself intends to operate a "superior" competing service, or develop an information system that does not depend upon the files of CAS. This is not very easy to do given the nature of chemical information and the files of CAS.

Perhaps some entrepreneurs will be able to carve out small areas, specialized files, that CAS does not dominate. We have examples of some successes such as ISI and DARC, some specialized services providing access to patent information or chemical data files. However, in each instance, we can probably expect CAS to develop competing products to maintain their worldwide market dominance.

The opportunities may appear bleak for alternative suppliers of secondary information, but is that bad? Is the user worldwide best served by the ACS-sponsored nonprofit service that seeks only to provide "... a single, comprehensive, online system that would be superior in its utility for scientists and engineers worldwide"?[6]

Is this a field where we do not need the pressure of competition to ensure innovation and reasonably priced services? Only time will tell.

REFERENCES AND NOTES

(1) P.F.U. served on the Advisory Board of Chemical Abstracts Service from 1974 through 1978.
(2) Collier, H. "Whose Afraid of Chemical Abstracts? Part II". *Monitor* **1982**, *22*, 5.
(3) Williams, M. *Invest. Med. Int. Rep.* **1983**, *1*, 1, 2.
(4) Dickman, J. T.; Platou, G. O. "Foreign Literature Challenges to Chemical Abstracts Service". *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 70.
(5) Reference 4, page 72.
(6) Crum, J. Letter dictated 28 Feb 1983 to H. Collier, Chairman European Association of Information Services.

ARTICLES

Vapor Pressure Data of Simple Organic Substances: Their Availability and Reliability†

ELIAHU HOFFMAN

National Center of Scientific and Technological Information, Tel Aviv 61201, Israel

Received August 16, 1983

The literature and other sources of information were searched for vapor pressure data of organic substances. Some compilations were found and compared with the data recorded in the *Handbook of Chemistry and Physics*. Discrepancies were checked and, where possible, analyzed. The reliability of the recorded data has been discussed.

INTRODUCTION

The compilation of vapor pressure data for organic substances in the *Handbook of Chemistry and Physics* is taken from the paper of D. R. Stull.[1] It is reproduced in every edition up to the last one, the 63rd[2]—except for the 52nd, 53rd, and 54th editions, 1971–1974—as well as in the *American Institute of Physics Handbook*[3] and in the *Chemical Engineers Handbook*.[14] In addition, its data are cited in numerous other compilations.[4,5] It is thus a most accessible and useful source of information and is probably used extensively in spite of its early publication in 1947. The vapor pressures of approximately 1200 organic substances are reported therein as a function of the temperature, within a pressure range of up to 1 atm. For 60.8% of these compounds, only single references sources are cited; 19.7%—that is, almost one-third of those single citations—date back beyond the year 1900; 14.8%—i.e., almost one-fourth of them—are actually one and the same—pre 1900—reference source.[6,7] These facts by themselves need not necessarily reflect on the reliability of the data. It stands to reason, however, that confirmation from several independant sources would enhance their reliability. In addition, it should be remembered that the experimental methods of measurement have undergone significant improvements in the last 80 years. Thus, the question arose whether there exist more recent, independant compilations that would verify these older data or, alternatively, whether one could obtain verification for specific compounds via the common literature or other sources.

AVAILABILITY OF DATA

In order to answer these questions, some seven simple organic compounds were chosen from Stull's compilation, the reference citation of which was this very same single—pre 1900—source.[6] Four of these substances, methyl dichloroacetate, 1,2,3-trichloropropane, ethyl carbamate, and ethyl trichloroacetate, are commercially available, and inquiries were made as to whether the manufacturers had any vapor pressure data. The results were disappointing. Data could be obtained for only two—ethyl carbamate and 1,2,3-trichloropropane—of the four substances, and this was only from two of the four firms that were approached. One firm provided a couple of values for 1,2,3-trichloropropane. The other one did quote continuous values for 1,2,3-trichloropropane and ethyl carbamate, presumably from the literature, but these did not tally

† Presented in part at the 7th International CODATA Conference, Kyoto, Japan, Oct 8–11, 1980.

0095-2338/84/1624-0003$01.50/0