History of Abstracting at Chemical Abstracts Service†

DALE B. BAKER,* JEAN W. HORISZNY, and WLADYSLAW V. METANOMSKI

Chemical Abstracts Service, Columbus, Ohio 43210

Received May 20, 1980

The evolution of abstracts and abstracting for *Chemical Abstracts* (*CA*) from 1907 to the present is traced. The growth of scientific publishing, changes in the content, form, format, national origin, and language of primary documents, the shift from a large corps of volunteer abstractors to in-house document analysts, and the transition to modern computer processing have all affected the preparation and publication of *CA* abstracts.

Chemical Abstracts was born out of the dissatisfaction American chemists felt with coverage of the American chemical literature in European abstracting journals. Faculty members at the Massachusetts Institute of Technology attempted to remedy this deficiency single-handedly by publishing the *Review of American Chemical Research* in 1895, issued both as a section of *Technology Quarterly* and separately. In 1907 this publication was superseded by the more extensive *Chemical Abstracts* (*CA*), sponsored by the American Chemical Society (ACS), with W. A. Noyes, Sr., as its first editor. While the history of *CA* includes the evolution of both abstracts and indexes, this paper will restrict itself to the history of abstracts and abstracting at *CA* from 1907 to the present.

EVOLVING DEFINITIONS

The American Chemical Society charged *Chemical Abstracts* with the mission of abstracting the complete world's literature of chemistry. This deceptively simple charter left to the small staff of the fledgling journal the task of deciding precisely what such terms as "complete", "abstract", and "chemistry" were to mean. The history of abstracting for *CA* is largely the history of defining these terms with ever-increasing precision.

Completeness. From the beginning, *Chemical Abstracts* included industrial and applied chemistry in its definition of "complete", despite some opposition from academic or "pure" chemists in the ACS membership. The German *Chemisches Zentralblatt* did not add industrial coverage until 1919, while in Britain papers on academic and industrial chemistry were abstracted in two separate journals until 1926.

By 1910, *CA*'s editor (Austin M. Patterson) was claiming that *CA* had become "as complete in its coverage of chemistry as any other similar publication".[1] While this may have been so, it managed to become even more complete over the next several years. In particular, W. Russell Stemen was influential in expanding *CA*'s coverage of the industrial application of chemical science and technology after he joined the editorial staff in 1922. Stemen also recognized the importance of patents as sources of new chemical information and pressed for broader coverage of the patent literature.

However, even after improvements made during Stemen's time, abstracts of patents were purposely kept brief until 1945. In that year, the war-time restrictions on paper eased, and patent abstracts were materially increased in length. Also, arrangements were made to obtain the patent specifications for abstracting by a subject expert, rather than publishing abstracts prepared by a volunteer at the patent office itself. Patent abstracts after 1945 have averaged one-third longer than formerly.

Chemistry. Even in 1907 it was not always easy to tell the chemical from the nonchemical literature. According to E. J. Crane, Editor of *CA* from 1915 to 1958 and Director of Chemical Abstracts Service (CAS) from 1956 to 1958:[2]

> It is not possible to establish criteria or to draw up inflexible rules that will easily and sharply classify published material into the two groups, chemical and nonchemical. This is so because of the indistinct borderline between chemical and nonchemical, and also because everything we abstract is new and cannot be readily forced into a system of classification that can be based only on what is already known.

However, Crane and the editors before and after him also recognized the necessity to define as precisely as possible the boundary between chemical and nonchemical for those who would be expected to choose which papers should be covered and which rejected. In this effort, as in the development of many of the basic technical and administrative policies that guided *CA*, he was assisted by Elmer Hockett, a chemist with a wide knowledge of the field and the ability to handle many languages, who was Crane's first hire in 1914 and continued to serve past retirement until 1970. He, Crane, and their colleagues agreed that a liberal and broad interpretation would be made in deciding whether or not a document was or was not chemical. As Crane wrote

> Suitable for *CA* are studies of new chemical reactions, new information on known reactions, chemical, physical, and biological properties of elements or pure compounds, preparation, purification, and analysis of elements or compounds, apparatus of particular interest to the chemist or the chemical engineer, and procedures that in themselves may not involve chemistry but are essential to an industry that is generally considered chemical.

As an illustration, he offered[3]

> There may be nothing more chemical about the construction of a sulfur burner than there is about the construction of the firebox for a steam boiler (each is designed to conduct a chemical reaction, and these reactions are indeed similar to each other), but the sulfur burner (reported in *CA*) is used exclusively in a chemical industry, the firebox (passed by) more often in a nonchemical industry.

The current instructions for coverage reflect the same ambiguities mentioned by Crane, and acknowledge the same

† Presented in part before the Division of Chemical Information, Symposium on History of Abstracting, American Chemical Society/Chemical Society of Japan Chemical Congress, Honolulu, HI, April 4, 1979.

0095-2338/80/1620-0193$01.00/0

Apparatus
General and Physical Chemistry
Radioactivity
Electrochemistry
Photography
Inorganic Chemistry
Analytical Chemistry
Mineralogical and Geological Chemistry
Metallurgy
Organic Chemistry
Biological Chemistry
Foods
Nutrition
Water, Sewage, Sanitation
Soils and Fertilizers
Fermented and Distilled Liquors
Pharmaceutical Chemistry
Acids, Alkalies, Salts
Glass and Pottery
Cements and Mortar
Fuels, Gas, Coke
Petroleum, Asphalt, Wood Products
Cellulose and Paper
Explosives
Dyes, Bleaching and Textile Fabrics
Pigments, Resins, Varnishes, India Rubber
Fats, Fatty Oils and Soaps
Sugar, Starch and Gums
Leather
Patents

Figure 1. Thirty sections of first *CA* (Vol. 1, No. 1, January 1, 1907).

necessity to make a decision, even if an arbitrary one, based on a broad definition of chemistry:[4]

> CAS defines chemistry to include all studies of elemental subatomic particles, elements, and compounds; their occurrence, composition, preparation, structure, and properties; their reactions and changes of state and the laws that govern these; their detection and determination; and apparatus designed for conducting these studies. Chemical engineering is defined to include the procedures usually classed as unit operations and unit processes, and the design and construction of plants and equipment for these procedures—CAS coverage does not include physical testing and physical shaping of materials nor does it extend to clinical studies at the chemical/medical borderline or to morphological taxonomy at the biological borderline.

Abstracts of documents determined to be both new and chemical were then arranged according to the principal subject of the new information. In Volume 1, this meant placing the abstract in one of 30 sections (Figure 1), the last of which was reserved for all patents. (This policy was changed in mid-1908, when patents were placed at the end of the most relevant section according to principal subject, like other scientific documents.) From the beginning, it was understood that this scheme—indeed, any subject classification scheme—was somewhat artificial. Again, it was Crane who later explained this perception:[5]

> Abstracts are often not definite units as to subject matter. An abstract may be of interest to more than one section of *CA* and it may not fit particularly well in any section.

Perhaps because the section arrangement was seen as a convenient aid to browsing rather than a definitive description of research areas in chemistry, the scheme was changed very little for 55 years. However, in 1962 CAS did reorganize its subject groupings to reflect the changing emphasis and growth of modern chemistry, and not incidentally, to produce sections of more workable size. The number of sections was expanded to 73 in 1962, 74 in 1963, and 80 in 1967, reflecting in particular the growth of applied chemistry and chemical engineering as well as the growth of research on nuclear phenomena. The most recent reorganization of the sections occurred in 1974 to accommodate the increasing number of abstracts appearing in the fields of fuels and energy, but the total number of sections was held to 80 at that time. The growth of a *CA* section related to nuclear chemistry and radiochemistry is illustrated in Figure 2. The coverage of that field grew from 2688 abstracts in 1948 to 64 850 abstracts in 1979.

The sections were in turn organized into five subject-related groupings: Biochemistry (sections 1–20), Organic Chemistry (21–34), Macromolecular Chemistry (34–46), Applied Chemistry and Chemical Engineering (47–64), and Physical and Analytical Chemistry (65–80). These section groupings have been available separately since their introduction in 1962. Abstract issues were changed from semimonthly to biweekly in 1961, and to a weekly schedule in 1967.

Abstract. Another definition that required considerable honing during *CA*'s development was the concept of "abstract" itself. In the early days of the publication, many abstracts were essentially shortened versions of the paper, and could for most purposes replace that document as a source of information for the chemist. Synthetic organic papers especially often yielded abstracts running over a column of text and including complete preparative information and descriptive details such as refractive indexes, melting or boiling points, and crystal form and color of the compounds reported as well as their derivatives. Analytical abstracts included complete, detailed procedures. Here is a typical example:[6]

> Heat 5 g. of sample with 50 ml. of 7.5 *N* HNO_3 on a steam bath. Boil off oxides of N on a hot plate and add 50 ml. of hot water and paper pulp. Boil 30 min., keeping the vol. at 80–100 ml. by adding hot water. Filter and wash the residue with hot 1% HNO_3. Treat the filter + residue with 15 ml. concd. H_2SO_4 + 5 ml. $HClO_4$ + 10 ml. of concd. HNO_3. Evap. to strong fumes, allow to cool, add 200 ml. of water + 75 ml. of concd. HCl, reduce Sn^{4+} with $SbCl_3$ + Pb, and titrate with standard I soln. . .

Other abstracts might be mere descriptions of the paper, never presenting the actual findings at all. While the *CA* editors strove to encourage informative abstracts for patents and journal articles (indicative abstracts are run to call attention to books, reviews, biographies and obituaries of chemists, and papers on chemical education and the history of chemistry), it was difficult to ensure consistency when the abstracts were actually written by hundreds—and later thousands—of volunteers in distant parts of the world.

While diversity in abstracts existed, the situation was far from chaos. Early in *CA*'s history two maxims were formalized in the "Guidelines for Abstractors", the combination training manual and style guide developed to encourage consistency and quality in the work of distant and disparate abstractors. They read

> 1910: The purpose of the abstract is to give *in the fewest words* the largest possible portion of the useful, *new* material in the paper.
>
> 1917: Judgment as to the importance of papers should be left largely to the present and future users of the journal.

As rephrased by Crane in 1952, this later maxim came to read "The inclusion of critical comments in abstracts has never been allowed. Authors have no opportunity to answer." This policy is illustrated by CAS reporting in the late 1960's of a series of papers describing the preparation of "water with

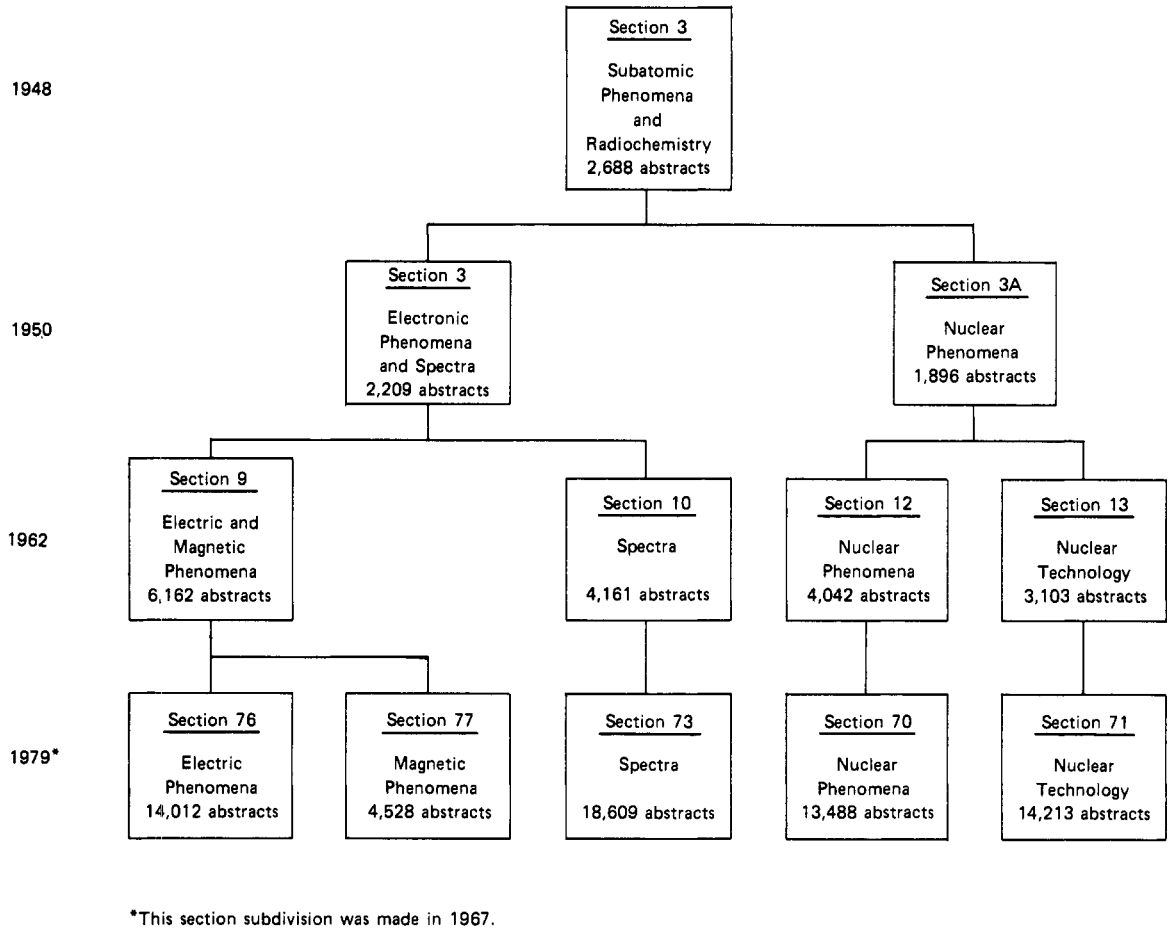

Figure 2. Growth of a *CA* section.

anomalous properties", dubbed by others "polywater". Later research showed the observed properties stemmed from silica impurities in the water, rather than any modification of its structure. Yet, the *CA* record reflects the state-of-the-art as it was believed to be true at the time of its reporting.

A third maxim—that abstracts are intended to lead to original documents, not replace them—was developed only gradually over *CA*'s first 40 years. The frequent need to shorten abstracts in the face of financial constraints, wars, paper shortages, and depression contributed to such a philosophy. But not until 1970 did a succinct statement of this view of the role of abstracts emerge, as CAS moved to a consistent "findings-oriented" abstract. As Russell J. Rowlett, Jr., Editor of CAS since 1967 and now also Director of Publications and Services, explained the change:[7]

> We are striving to present uniform abstracts, concisely written, with as much detail as needed to understand the author's main purpose. Our intention is to eliminate experimental detail which only repeats or enlarges on the research, and which is better obtained from the original document. *CA* will continue to provide access to the primary literature; it does not replace that literature.

Here is an example of a findings-oriented abstract of a process patent:

> Mg hydride is prepd. from Mg and H at 180–200° and 20–50 atm in the presence of an inert liq. in an autoclave filled with stainless steel balls and having a rolling and regular movement. Particularly reactive Mg hydride is obtained if an organometallic compd. of an element of Groups I, II, or III of the periodic system is added to the reagents and to the inert liq. Thus, Mg and H react in an autoclave at 210° and 28 atm for 17 h to give a yield of 15%. If anhyd. C_6H_6 is added to the reagents, the yield is 90%. If Et_2Mg is added to the reagents and to the C_6H_6, a particularly reactive Mg hydride is obtained.

A concise definition of a good abstract which corresponds closely to CAS's current standards is this from the American National Standards Institute:[8]

> A well-prepared abstract enables readers to identify the basic content of a document quickly and accurately, to determine its relevance to their interests, and thus to decide whether they need to read the document in its entirety.

In particular, the informative abstract favored by *CA* begins with a first sentence which is a brief statement of the major disclosures, results, and conclusions of the author. Like the lead paragraph of a news story, this sentence is designed to let the reader know whether or not the subject is of interest. The abstract text includes informative supporting statements on methodology, reactions, activities, properties, and applications.

Abstracting patents and patent applications is complicated by the fact that the primary purpose of the patent is to acquire certain legal rights, not the transmission of scientific or chemical information. As a result, the patent may contain general statements concerning possible applications of new products or possible alternate methods of carrying out the invention and possible related organic compounds that may be produced. The abstractor must sort out these generalizations (frequently referred to as paper examples) from scientific statements of what has been accomplished.

Each abstract is preceded by a heading, which (for journals) includes such information as the title, authors, location of work, journal title, year of publication, volume and issue numbers,

pagination, and language of the original document. For patents, the heading includes names of inventors and assignees, country and patent number, date of publication, classification, and date on patent application.

In 1972, highlighting by italic type of chemical substances significant in the studies reported was introduced into *CA* abstracts in 17 sections. For these sections, CAS Registry Numbers were also included following the names of these substances. By 1976, both practices had been extended to 40 sections.

The definition of the abstract developed at CAS did not arise purely from internal conversations, but rather in response to the needs of researchers who depend on abstract journals for access to the literature. H. Borko and S. Chatman explain this orientation:[9]

> The nature of an abstract reflects the editor's purposes, which in turn reflect the desires of the reader. Therefore, one cannot separate function from use: an editor's constant problem is to estimate what sort of presentation his readers really want.

The readers, according to R. E. Maizell and his colleagues, really want abstracts that provide "at least a partial solution to the almost universal lament of the scientist: 'I just don't have enough time to keep up with the literature'." They add[10]

> Abstractors and their abstracts promote current awareness, bridge language gaps, prevent duplication of previous work, and save time and dollars at almost all levels in the majority of organizations.... Abstractors and their abstracts help to conserve research time, hence, research costs. . . . In some cases, an abstractor can eliminate the cost of an entire research program by indicating that the work has already been done.

Reflecting on "the sort of presentation his readers really want", and on the characteristics not only of a good abstract, but of a good abstracting journal, Crane wrote in 1958:[11]

> Completeness, promptness, accuracy, clearness, and ready usability are considered to be the principal characteristics of a good abstract journal.... Such characteristics are essential if an abstract journal is to be used with confidence that significant existing information is not likely to be missed.

In a similar spirit, he wrote the following statement that has appeared on every *CA* masthead since 1937:[12]

> It is the careful endeavor of *Chemical Abstracts* to publish adequate and accurate abstracts of all scientific and technical papers containing new information of chemical and chemical engineering interest and to report new chemical information revealed in the patent literature.

The contents, form, and format of *CA* abstracts have been periodically examined, and various ways of improving them considered. Apart from user suggestions, comments have been specifically solicited from experts in relevant specialties, from section editors, and most recently from the members of the CAS Editorial Advisory Board, created in 1975.

GROWTH

The effort to maintain completeness in the face of continuing exponential growth in the number of documents to be covered has dominated much of *CA*'s history. The pressure of the enormous growth in the scientific literature was at times compounded by such events as depression, war, paper shortages, large deficits, fluctuations in the amount or character of the literature, and inflation. Under similar pressures many other abstracting journals have ceased publication or have limited the number of abstracts to be published or the number of journals to be covered. *CA* has resolutely resisted such outs, making it "the only business in the world with no control over its raw materials", according to current editor Rowlett.

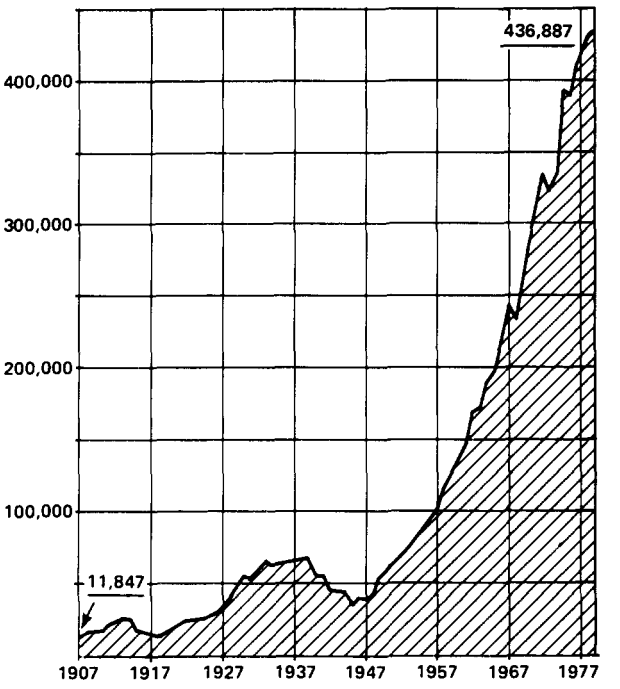

Figure 3. Growth of number of *CA* abstracts published (1907–1979).

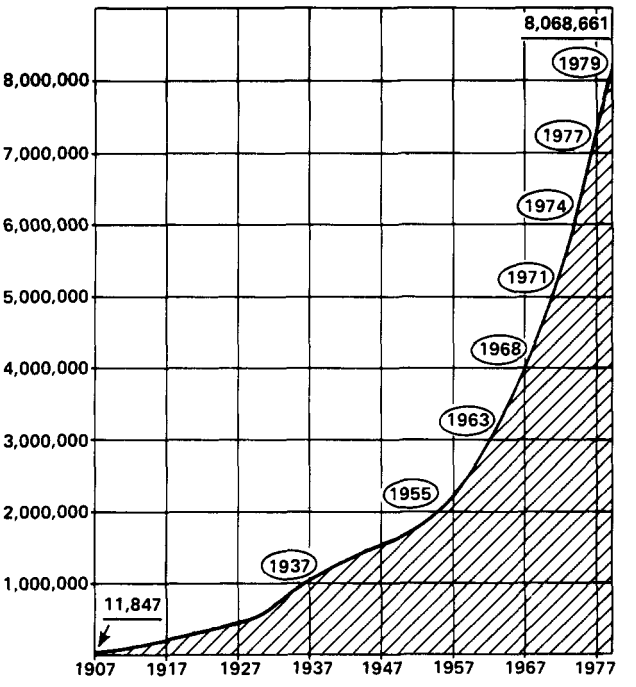

Figure 4. Cumulative growth of *CA* abstracts (1907–1979).

The graph in Figure 3 summarizes the growth of abstracts published by *CA* each year. Just under 12 000 were published in 1907; in 1979, nearly half a million were published. Figure 4 shows the cumulative total of *CA* abstracts, clearly illustrating the exponential nature of the growth. While it took *CA* 30 years to publish the first million abstracts, it now takes only 2 years to add a million abstracts to the total record.

This striking growth in the literature of chemistry, reflected in the number of abstracts published in *CA*, has stimulated a parallel growth in almost every aspect of CAS information services and the organization that produces them. From the number of journals monitored, to the amount of money needed to sustain the operation, the number of employees, the size of *CA* itself and the indexes which provide access to it, the story has been one of uncomfortably rapid expansion.

Predictably, the most persistent problem for most of *CA*'s history has been money. In 1907, *CA* required a staff of only two part-time editors and a secretary (supplemented by 129 unpaid volunteer abstractors), and had an annual budget of $15 500. Today the daily cost of producing this world digest of chemistry is over seven times that amount.

Through 1933, *CA* had been financed solely by ACS member dues, and all members who wished received the publication free. After 1933, a small subscription fee was imposed which did not begin to defray expenses. By 1955, operation expenses exceeded $1 000 000 for the first time, and the deficit over subscription revenues was almost $500 000. In that first year, the ACS Board of Directors changed subscription policies to place *CA* on a self-supporting basis, with a dramatic rise in the price for *CA*.

During the many years when limited funds had to be continually stretched to cover mushrooming literature, *CA* attempted to balance the budget not by cutting back on the number of journals analyzed or placing a limit on the number of abstracts published, but by shortening those abstracts which were published. As the literature continued to expand, *CA* also had trouble meeting its commitment to promptness. However, this difficulty generally affected indexes much more seriously than the abstracts themselves.

Other crises which made it difficult for *CA* to maintain its commitment to completeness included the depression, the paper shortage of 1943–44, and, of course, two world wars. During both wars, the total number of papers abstracted dropped markedly—partly because less scientific work was being done; much work that was being done was classified and therefore not published, and the war seriously disrupted *CA*'s normal channels for obtaining papers. To cover completely that literature which was still being published, *CA* developed several ingenious schemes to obtain papers from war-torn areas. The following is an account of several such schemes from World War II:[13]

> 1. Direct mail arrangements were made to obtain journals from publishers and to have these journals sent at first down through Italy, whose ships were not being blockaded for a while, and then across Siberia as Italy became involved, but before Russia did.
>
> 2. A staff of abstractors in Europe (75 of them at one time in Switzerland) was organized with the effective aid of W. Nowacki of Bern. Clipper plane mail brought abstracts typed on thin paper to the United States, where *CA* workers translated them.
>
> 3. Eventually papers were obtained for abstracting, often in microfilm form, with the aid of the Office of Scientific Research and Development.
>
> 4. Abstracts made for other purposes in Europe were obtained and used.
>
> 5. Advantage was taken of the readiness to cooperate of some American abstractors interned in Germany (in China also).

Abstracts from Japanese journals, which were unavailable for many months after World War II, were obtained through cooperation with the U.S. Army of Occupation in Tokyo.

Over the years, *CA* has been affected by the growth in sheer bulk of the literature as well as by rapid changes in its composition. One such change was the rapid explosion of technological innovation that swamped the patent offices of many countries after World War II. In addition to the growth in inventions, the proportion of patents having chemical significance increased. Further, inventors increasingly sought patent protection for the same invention in numerous countries, increasing the number of patent duplicates to be examined. To clear their own backlogs, patent offices of many countries began publishing patent applications in various stages of examination, contributing further to the amount of paperwork to be examined and referenced by CAS for each invention. For example, *CAStings* (a newsletter for volunteer abstractors) in 1972 ran this message:

> Help! We need more Japanese language abstractors. Beginning soon after July 1, we must abstract and index the new Japanese patent application publications, the *Kokai*... It is estimated that there is a backlog of about three-quarters of a million of these publications to be issued.

A similar message was issued in 1968, when the West Germans began publishing patent application documents at the rate of 2500 a month to clear out a backlog of about 284 000 applications. Finding and training the subject specialists and language experts to cope with such sudden and large-scale changes in the literature presented major problems for an already large and complex business.

	1937	1961	1965	1970	1975	1978
English	40%	43.3%	52.0%	56.4%	59.7%	62.8%
Russian	7	18.4	20.0	22.6	23.3	19.5
German	15	12.3	9.8	6.6	4.8	5.0
Japanese	5	6.3	4.0	3.4	3.0	4.7
French	6	5.2	5.1	4.0	3.0	2.4
Other	27	14.5	9.1	7.0	6.2	5.6
	100%	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 5. Languages of original papers.

Changes in chemical specialty alter the number of abstracts appearing within various *CA* sections, and, of course, alter the organization's need to maintain experts in the various subspecialties of chemistry. In 1971, Rowlett commented:[14]

> Total published abstracts in *CA* are running more than 15 percent ahead of last year. We do not have complete analyses, but much of this growth is in the Biochemistry and Physical Chemistry Sections. Growth appears to be related to biomedical subjects and to applications studies of environmental factors. Recently one Section Editor suggested facetiously that all professors who are having trouble finding support for their new research are at home writing up their past research.

At the same time, CAS Director Dale B. Baker was commenting on the change from discipline-oriented science—reflected in the classification scheme used in *CA*—to mission-oriented science. Mission-oriented science, he pointed out, tended to involve teams of scientists from different fields to solve particular problems rather than to extend basic knowledge within a particular field. How to organize and present science with such a focus created problems for a journal organized largely according to a discipline-oriented plan.

Long-term changes in the literature of science have also affected CAS, but more slowly and therefore less painfully. For instance, almost half of the abstracts published in *CA* in 1907 came from papers contributed from Germany. By 1913, that number was down to one-third, by 1929 to one-fourth, and in 1979 it was down to 7.2% (includes both the Federal Republic of Germany and the German Democratic Republic). Meanwhile, the literature originating in the U.S.S.R. rose from 6.3% right after World War II to over 20% of the literature abstracted in the 1970s. Another change that has greatly affected CAS is the growth of science outside the six most industrialized nations (U.S.A., U.S.S.R., Japan, West Germany, United Kingdom, and France). The contributions of

1907	1917	1927	1937	1947	1957	1967	1977	1979
129	172	261	449	486	1,283	3,245	1,406	1,029

Figure 6. Number of volunteer abstractors.

these 150 or so nations have quadrupled from 6.8% in 1907 to 26.5% in 1979. Despite this growth, however, the number of papers in languages other than English has declined steadily through the years (Figure 5). Today, almost 95% of the literature is published in one of five languages (English, Russian, German, Japanese, or French), with almost two-thirds appearing originally in English. Among the other 52 languages encountered, CAS has abstracted papers published in such languages as Esperanto, Gaelic, Interlingua, Sindhi, Swahili, and Welsh.

VOLUNTEER ABSTRACTORS

For most of its history, *CA* has relied on volunteer chemists to prepare abstracts for publication in *CA*. In addition, each subject section was overseen by one or more volunteer section editors, subject-matter authorities who monitored the coverage of each section and the quality of the abstracts. The journal began in 1907 with a network of 129 volunteers, who contributed their time and skill with no monetary reward whatsoever. In 1929, the decision was made to offer minimal pay, and the following elaborate statement was published in the "Directions for Assistant Editors and Abstractors of *Chemical Abstracts*":

> It is realized that abstracting is done very largely in a spirit of service to chemistry and that some reward is to be found in the satisfaction and advantage of taking part in and being associated with a worth-while project. Some consider abstracting as an interesting and altogether satisfactory way of doing reading which would be or at least should be done anyway. Nevertheless it has been arranged to pay abstractors at the rate of $3.00 per printed page of *Chemical Abstracts*.

This rate was raised to $6 per printed page in 1939. It was periodically increased over the years and in 1979 had reached 18¢ per line (or about $32 per page). However, it was clearly recognized by abstractors and *CA* editors alike that the small amount of money involved would never in itself serve as motivation for the task of abstracting.

As the chemical literature grew, the number of volunteer abstractors had to grow commensurately (Figure 6), so that by 1954 the "*CA* family" totaled over 1000, and by 1961 over 3000. In some countries, notably in Japan and Poland, associations or working groups of *CA* abstractors were formed to facilitate common understanding of abstracting rules and to improve communication with the Columbus office.

Coordinating the activities of this large group of geographically separated people fell to the Columbus-based editors, and from the years 1915 to 1958, to *CA*'s Editor, E. J. Crane. There have been many individuals who abstracted for 50 years or more. Crane referred to them as the "iron men of *CA*". To promote a feeling of solidarity, transmit informational tidbits, and generally maintain communication with the large volunteer network, he began publishing a newsletter called the *Little CA* in 1930. This small publication featured humor, motivational material, reminders of deadlines, poems, and personal statements from the editor. It continued publication after Crane's retirement as Director in 1958, ceasing with his death in 1966.

Samples include[15,16]

> The M used in connection with author names in French journals to stand for Monsieur is sometimes misleading. It is taken by abstractors to be an initial M.
>
> When density and refractive index data are given in abstracts it is desirable always to indicate the temperature at which determinations were made.

The publication also featured poems titled "Chirps" by Mildred Bird, the office librarian:[17]

> For years poor number twenty-two
> Has been so oversize
> So abstracts, come in early, please,
> And help it slenderize.

The above "Chirps" refers to a long tradition called "The Roundup", or "Remember November", an effort on the part of Editor Crane to call in those abstracts which may have been left over from summer vacation and the start of the new academic year, so that they could be published in the last issue of the year—No. 22 appearing in November. The December volumes were then devoted entirely to indexes.

Crane's favorite poem which he often repeated was[18]

> BREVITY COMMENDS ITSELF
> In speeches, yes,
> In skirts, I guess,
> In hair, perhaps,
> In office naps,
> And now, anew,
> In abstracts, too.

After Crane's death, Rowlett chose not to continue the *Little CA*—so closely identified with Crane's particular style and personality—but began a more formal newsletter called *CAStings*. This was replaced in 1973 with *CAS Report*, a publication whose audience includes the librarians and information specialists who have become so important in making *CA* accessible to the scientific community. The Editor's column "Sidenotes" is specifically directed to abstractors, section editors, and advisors.

Publishing a journal with input from such far-flung and disparate talent put a heavy load of responsibility on the Columbus staff. Their task, according to Crane, included editing:[19]

> with a dozen different considerations in mind, chemistry as well as newness of information, English as well as technical language, word abbreviation as well as possible abstract abbreviation or lengthening, style as well as standard form, classification as well as the making of cross references, etc.

The editors were also responsible for placing abstracts in one or another of the *CA* sections, and indicating appropriate cross-references. Because they dealt with abstracts made from documents originally published in many languages, most editors were expert in several of the more frequently appearing languages, and in addition "acquire the ability surprisingly often to settle questions in abstracts by examining original papers appearing in a great variety of languages".[20]

After 1966 the fraction of abstracts done by volunteers began to decrease, declining from close to 100% to 8.6% in 1979. The number of volunteer abstractors and section editors has declined from a peak of 3292 in 1966 to just over 1000 in 1979, but is not expected to fall much lower in coming years. The function of section editor has been absorbed gradually into the Columbus office, and to reflect different responsibility, the title was changed to section advisor. In addition to the abstracts prepared by volunteers, the Royal Society of Chemistry in the United Kingdom provides abstracts and index entries for papers and patents published in the U.K. (about 20 000 per year) under an agreement negotiated with the American Chemical Society in 1969. The West German organization Gesellschaft Deutscher Chemiker also contributed abstracts and index entries under a similar agreement signed the same year. In 1975 another organization, Internationale Dokumentationsgesellschaft fuer Chemie (IDC), representing the West German chemical community, signed a new agreement with the ACS and chose not to provide abstracts but

instead pay a larger share of the cost of producing the data base. Inflation and the devaluation of the American dollar made input of abstracts and index entries to the CAS base financially unattractive for IDC.

The shift away from abstracts prepared by volunteer subject experts to those prepared by full-time, on-site specialists has occurred for three basic reasons. First, the continuing expansion of the chemical literature made larger groups of volunteers necessary, and the administrative effort of identifying, training, coordinating, and editing abstracts from such a large and widely scattered group became increasingly difficult. As CAS looked at the further growth likely in the early 1960s, doubts were expressed about whether such a process could be continued, or could be managed quickly enough to deal with continued expansion of the chemical literature in any timely fashion.

Second, computer processing was making possible increasingly rapid and efficient work-flow patterns throughout the late 60s. As CAS began the task of eliminating redundancy in the clerical tasks associated with producing *CA*, the central redundancy in the intellectual tasks became more glaring—two chemists had to read and understand each document, one to index and one to abstract. In most cases, one or two editors would also read the document in order to proof both efforts. In 1970, the concept of Unified Document Analysis was adopted, which called for a single document analyst to prepare both the abstract and all index entries. This program depends on highly trained experts and would not be compatible with a largely volunteer organization.

Third, the financial situation had improved to the point that CAS was able to pay full-time chemists to abstract and index. Increases in the price of *CA* had brought the organization self-sufficiency and the freedom to put itself on an improved business footing.

As the computer has taken on more and more of the routine tasks associated with the production of *CA*, questions have arisen as to whether the marvelous machines will ever be able to analyze documents and prepare abstracts all by themselves, with perhaps only editing assistance from humans. Several programs have been written—some by CAS staff in the Research and Development Division—which do produce abstracts from computer-readable documents. However, to date, only a tiny fraction of documents handled by CAS is available in computer-readable form. Abstracts and extracts derived automatically are not as good as those prepared by document analysts. Two CAS information scientists have concluded: "It will not be possible to produce abstracts of manual quality by computer without a great breakthrough in linguistics, especially in the area of semantics".[21]

DIRECTIONS FOR ABSTRACTING

In order to ensure some consistency in abstracts obtained from volunteer abstractors of diverse nationalities and educational backgrounds, uniform standards were agreed upon early in the life of *CA* and distributed to the volunteers in the form of brief pamphlets. The first was published with Volume 2 of *Chemical Abstracts* itself. Called "Organization and Directions for Abstractors", it included a list of *CA*'s 30 sections, followed by a general discussion of the responsibilities of section editors and abstractors, 4.5-page-long directions organized into 35 numbered paragraphs, and a list of journals abstracted. About a third of the directions was devoted to nomenclature standards, and another third to stylistic conventions.

By 1910 the pamphlet had grown to 8 pages and 122 points, and was published separately under the title "Directions for Assistant Editors and Abstractors". Additions included a list of 74 standard abbreviations, a list which had grown steadily with each new edition. In 1939 Crane altered the pamphlet to a pocket format which could easily be carried about with the volunteer as he traveled from home to office. He described its 1952 edition as[22]

> A 46-page booklet containing information on assignments, procedures, style, forms, abbreviations, nomenclature, notation, and numerous miscellaneous details, as well as full discussion of the nature of the abstract desired. This booklet is indexed by subject. *Chemical Abstracts* has a special 14-page set of "Instructions for Abstracting Patents", and, for the sake of emphasis, the more important directions for abstracting are frequently distributed as a separate group in so-called "Abstracted Directions".
>
> It is considered (1) that abstracts should be informational rather than descriptive, (2) that as much brevity as is consistent with adequate clear reporting is a virtue, (3) that abstracts should not be critical, (4) that their length should be influenced somewhat by the accessibility of the source, (5) that new data should be given with adequate precision, (6) that an author's purpose should be made clear and his conclusions reported, (7) that new terms should be defined, and (8) that abstracts should be helpfully correlated by the inclusion of *CA* references to definitely related earlier abstracts.

Most of these points remain applicable today, even through the major changes in abstracting which occurred around 1970. However, the practice of correlating abstracts to related earlier material was discontinued in 1967.

Pocket editions of the "Directions" continued past Crane's term as editor. The 1960 edition had grown to 43 pages followed by a 7-page subject index and including a 9-page section devoted to abstracting patents.

The 1970 edition of the "Directions for Abstractors" moved away from the pocket form in favor of an 8.5 × 11 in. folder that could be updated easily. The introductory statement described the folder as "designed for flexibility and for adaptability to changes". The 1971 edition included instructions for the new findings-oriented abstract adopted as the norm in 1970. While such abstracts had been fairly common in *CA* for a long time, they had been interspersed with others—particularly in the organic sections—which were long, cookbook abstracts. The 1971 "Directions" described the new philosophy:[23]

> The abstracts . . . are not meant to serve as laboratory manuals or chemistry handbooks. Their primary purpose is to give the reader accurately and quickly enough information on the chemical content of the document abstracted to allow him to determine whether he wants to consult the original publication.

With the increasing reliance on in-house abstractors and continued growth of the literature, larger and more detailed directions had to be issued. The 1978 edition is the gigantic four-volume "Editorial Analysis Manual", distributed in 3-in.-thick red loose-leaf covers. Only the first of the four volumes applies primarily to abstracting, and even that one also contains sections on "Structure Drawing for *CA* Issues" and "Keyword Indexing", reflecting some new responsibilities taken on by in-house chemists who are now called "document analysts" rather than "abstractors".

THE PROCESSING SYSTEM

Although the abstract itself has undergone relatively few changes over the almost 75 years of *CA*'s publication lifetime, what happens to an abstract once it is written has been thoroughly revolutionized.

Throughout the early history of *CA*, abstracts, abstractors, index entries, journals, bibliographic headings, and every other

conceivable bit of information had to be kept track of by hand—generally through the mechanism of cards in filing cabinets. For instance, here is Crane's description of the file used to keep track of volunteer abstractors:[24]

> The abstractor is selected by consulting three files of cards, one arranged according to subjects, one according to abstractors' names, and one according to language (for all languages except English, German, and French).
>
> When an abstractor has been selected, a temporary debit card carrying the terminal date for the approximately two-week period ending on the next 1st or 15th of the month is inserted behind the abstractor's name in the assignment file.
>
> A geographical file of all abstractors is maintained; it is classified by states and principal cities. . . This is useful in making assignments from journals not available to us.

Similarly elaborate filing systems were required for keeping track of authors, index entries, and abstracts. Abstracts were recorded on "manuscript cards" and filed by author, to guard against the publication of more than one abstract per document and as an aid in referencing an abstract to earlier related abstracts.

As mentioned before, the growth of the chemical literature after World War II put severe strains on this complex and redundant system in a variety of ways. For one thing, as the number of volunteer abstractors tripled, the effort of assigning and editing increased more than proportionately. The multiple efforts required in abstracting, editing abstracts, indexing, editing index entries, and processing accompanied by several proofing stages became overloaded. In all, some 14 professional handlings and 21 clerical operations took place between the selection of a paper or patent for coverage and final publication of its abstract and all of its index entries.

Maintaining this processing stream as the literature increased steadily at about 9% per year meant huge increases in staff. From 1961 through 1964 the CAS staff doubled. Projections made at that time indicated it would double again by 1970 using processing standards then current. Fred A. Tate, Associate Director of CAS, wrote in 1967:[25]

> This rate of increase cannot continue indefinitely; we simply cannot recruit the highly qualified personnel we need and integrate them into our production operation fast enough.

In the early 1960s, CAS began studying the application of modern data-processing technology to the handling of chemical and chemical engineering information. In light of the eventual triumph of computer technology in so many areas of modern life—and most particularly in information processing—it is difficult to recall how speculative and risky such an investment seemed with the relatively clumsy and expensive machines then available. Crane remained to the end of his life highly skeptical about the role computers could play in information processing, as he expressed in the following poem, entitled "Thoughts on a Computer" and printed in the *Little CA* of May 5, 1960:

> A thing like Deuce
> Is not much euse
> Unless one siques
> For right techniques
> To get the sort of data
> For which the thing can cata.

By 1965 (and with assistance from the National Science Foundation), CAS's research program had yielded some promising techniques, and the organization embarked on a 5-year program to implement an operating computer-based information system. This effort was not without its pitfalls, as Rowlett noted in a 1967 *CAStings*:[26]

> In 1967 we learned how to enter the output of men's minds into computers—how to store it in machine language—and how to retrieve it in forms useful for different purposes than those for which it was originally developed.
>
> Because complex data processing is a still-developing science, we were not always successful. Regretably, some of our products are not as perfect as we wish; schedules have been missed, and one, *POST*, had to be suspended for four months.

The effect of computer processing has been most evident in simplifying and speeding the publication of indexes and in making index entries more flexible through their availability as search terms in computer-readable files. But the effect on abstracting has also been significant, most recently in the simplification of the abstract-editing process made possible through the introduction of the Distributed On-Line Edit System (DOLE) in 1978.

Since mid-1975 all abstracts have been entered into the computer data store by keyboarders whose typing at a terminal transcribes the written or taped abstract into computer-readable form. However, editing these abstracts for publication then required that the printed abstract be checked by hand, corrections keyed (with the accompanying possibility that new errors might be introduced), and then the corrected abstract checked yet again. Simply putting the resulting pieces of paper into envelopes and taking them out again consumed major portions of staff time. In 1978, a system for editing these abstracts on the screen of a terminal tied to a minicomputer was perfected, so that the abstract could be called by the chemist, corrected, and returned to the data base without the necessity of papers flowing through the building and without the possibility of errors creeping in between the correction and publication process.

The speed of computer processing has also dramatically improved the currency of both abstracts and indexes. The computer cut the average time needed to process an item for *CA* from 28 weeks to 14 weeks, and over the past 10 years this figure has fallen even further to 10 weeks. This time includes acquisition, document analysis, processing, and publication.

The great flexibility of CAS's information store has been again of greatest benefit in the production of indexes. For instance, the names of the authors of a paper need to be recorded only once to appear in the heading of the abstract and in Issue, Volume, and Collective Author Indexes. The same flexibility has begun to make possible creative new services based on *CA* abstracts as well. In 1965 CAS began producing a subset of abstracts concerning the biological activity of chemical substances called *Chemical-Biological Activities* (*CBAC*). This was followed in 1967 by *Polymer Science and Technology* (*POST*). Abstracts continue to be available for computer display. While many computer services make available the full range of *CA* index entries for on-line retrieval of references and bibliographic headings, no one has yet solved the problems which would allow the full text of all *CA* abstracts to be stored and displayed. Only in some specialized subsets are full *CA* abstracts displayable on-line, e.g., in TOXLINE of the National Library of Medicine (NLM).

In 1976 CAS introduced a new printed information service based on the flexibility offered by computer processing—*CA Selects*. Begun with 6 topics, the service reached 110 titles in January 1980. Each bulletin is designed to help specialists in a narrowly defined interest area keep up with the latest developments in the field by providing references and abstracts retrieved by a computer search of the entire CAS data base. Search profiles are refined to produce complete biweekly bulletins that can be easily scanned and are of interest to a

significant number of researchers.

Today, CAS is experimenting with similar biweekly bulletins designed especially for the research interests of industrial customers, who pay not only for each bulletin purchased but support the developmental costs of preparing the computer profiles as well. Such alerting bulletins can be especially useful where traditional library support is not available.

The abstract is today the great untapped resource of the CAS data base, and much interest has been expressed in making this valuable source of information more available in computer-readable form and a basis for retrieval through computer searching of words contained in the abstract text. While all abstracts containing new chemical and chemical engineering information will continue to be published in *Chemical Abstracts* as the complete, archival record of the world's chemical literature, more emphasis will be placed in the future on finding new and more individually tailored services that make abstracts more accessible through the speed, accuracy, and flexibility of computer processing.

REFERENCES AND NOTES

(1) "A Brief History of Chemical Abstracts Service", in "A Century of Chemistry"; American Chemical Society: Washington, DC, 1976, p 3.
(2) "CA Today: The Production of Chemical Abstracts"; American Chemical Society: Washington, DC, 1958, p 22.
(3) Reference 2, pp 21-22.
(4) "CAS Editorial Analysis Manual"; **1976**, Vol. I, Chapter 2, p 4.
(5) Reference 2, p 32.
(6) *Chem. Abstr.* **1950**, *44*, 72b.
(7) *CAStings*, **1970**, *4* (1).
(8) Weil, B. H. "Standards for Writing Abstracts", *J. Am. Soc. Inf. Sci.*, **1970**, *21*, 351-357.
(9) Borko, H.; Chatman, S. "Criteria for Acceptable Abstracts: A Survey of Abstractors' Instructions", *Am. Doc.*, **1963**, *14*, 149-160.
(10) Maizell, R. E.; Smith, J. F.; Singer, T. E. R. "Abstracting Scientific and Technical Literature"; Wiley-Interscience: New York, 1971, p 2.
(11) Reference 2, p 2.
(12) *CA* Masthead.
(13) Crane, E. J. "Chemical Abstracts", in "A History of the American Chemical Society"; Browne, C. A.; Weeks, M. E. Eds.; American Chemical Society: Washington, DC, 1952, pp 344-346.
(14) *CAStings*, **1971**, *5* (1).
(15) *Little CA*; No. 12, Christmas, 1933.
(16) *Little CA*; No. 20, Aug 17, 1936.
(17) *Little CA*; No. 6, Sept 21, 1931.
(18) *Little CA*; No. 8, May 1, 1932; No. 57, Christmas, 1948; No. 100, Aug 20, 1963.
(19) Reference 2, p 29.
(20) Crane, E. J.; Langham, C. C. "On-the-Job Training at Chemical Abstracts Service", *J. Chem. Doc.* **1962**, *2*, 199-204.
(21) Pollock, J. J.; Zamora, A. "An Automatic Abstracting Research at Chemical Abstracts Service", *J. Chem. Inf. Comput. Sci.* **1975**, *15*, 226-232.
(22) Crane, E. J. "The Training of Chemists for Abstracting and Indexing", in "Training of Literature Chemists", *Adv. Chem. Ser.* **1956**, *No. 17*, 18.
(23) "Directions for Abstractors"; Chemical Abstracts Service: Columbus, Ohio, 1971, p 1-1.
(24) Reference 2, pp 16, 17, 20.
(25) Tate, F. A. "Progress Toward a Computer-Based Chemical Information System", *Chem. Eng. News*, **1967** (Jan 23), 78-90.
(26) *CAStings*, **1967**, *1* (2).

A Comprehensive Audio Course on the Use of the Chemical Literature†

SAMUEL H. WILEN

Department of Chemistry, The City University of New York, City College, New York, New York 10031

Received March 17, 1980

The rationale for, development of, and organization of a comprehensive audio course on the chemical literature under the auspices of the Americal Chemical Society are described. The structure of the course is based upon an analysis of chemical information activities of individuals by Arnett (1970) in which types of searches (retrospective—exhaustive or reconnaissance—searches, data searches, current awareness) are matched to appropriate printed as well as computer-readable chemical information sources. The course features visual presentations of typical publication content, detailed descriptions particularly of secondary literature sources, and identification of user aids.

INTRODUCTION

An audio course entitled "Use of the Chemical Literature. An Introduction to Chemical Information Retrieval" (1978) has been released by the American Chemical Society, Department of Educational Activities. This paper gives the rationale for the preparation of this course, describes its genesis, and explains its organization.

The course physically consists of a set of eight audio tapes totaling 8.8 h of lecture, a manual consisting of 300 visuals and illustrations, and a set of 60 exercises to give students facility in using search tools. The course manual contains an index and may have some utility as an independent chemical literature reference tool.

The course is designed to prepare individuals to use the major search tools and even most minor ones required in chemical literature searches. It is intended to do this for individuals ranging all the way from inexperienced students to experienced users. A specific goal of the course is to provide a means of updating the literature searching skills of the latter group.

The course subtitle in part reflects the effort which was made not only to describe search tools and show how to use them but also to guide individuals in the development of search strategies.

BACKGROUND

In recent years, much of the academic instruction on the use of chemical literature in the United States and abroad has been informal. Typically, a small group of students is taken to the library, given a brief tour and some rudimentary instruction on secondary sources, and directly launched on searches.

Although I have conducted such introductions myself, I know how deficient they are. Since they are too often the only

† Presented to the Division of Chemical Information, Symposium on the Teaching and Use of Chemical Information in Academia, 178th National Meeting of the American Chemical Society, Washington, DC, Sept 10, 1979.

0095-2338/80/1620-0201$01.00/0