2. A program in information science to be incorporated as the minor in a traditional doctorate program in science or engineering. With such an educational background a researcher would be able to make the most efficient use of the literature for himself and for others. This program could be offered with no change from existing procedures by any institution which has a program in information science. It also would serve to train technical literature analysts.

3. A program leading to the degree of Doctor of Philosophy in Information Science, designed to train information scientists. This degree program would be research oriented, with the course work concentrated in information science, including linguistics, logic, and machine storage and retrieval.

For the curricula leading to the M. S. in Science Information we believe a reasonable division of the time among the various blocs is science—45%, information science—35%, language—20%. In the science bloc we include both the deepening studies in the student's own major field and the broadening studies in other technical fields. The Ph. D. subject specialists with a minor in information science would take approximately the same information courses as the technical literature analyst earning a master's degree. We have not yet come to a conclusion concerning the division of time in the Ph. D. program in Information Science except that its primary emphasis will be in information science.

We are in process of devising typical curricula which students with various technical backgrounds might take at Georgia Tech in pursuit of the M. S. in Technical Information. These are not yet sufficiently well worked out to warrent presentation in this report.

IN-SERVICE TRAINING

We have been concerned about the apparent demands in time and leadership of the instructing staff in order to maintain a satisfactory in-service training program. Furthermore, or perhaps I should say consequently, in-service training programs do not offer much opportunity for training substantial numbers of people. However, they do provide a means to develop an especially deep insight into the nature of the work responsibilities and the over-all picture of the operation. Consequently, these programs may be very desirable to train people who have high potential to be real leaders in their field. A number of libraries, especially medical libraries, have in-service training programs, often called internships. In-service training for technical literature analysts is common but only on an informal and relatively superficial basis.

There is a great deal of interest today in teaching as well as in studying the various facets of the information field. It is important that we not sell short the library schools simply because they have not been producing the type of talent which many of us think is needed. They are showing increasing signs of interest in modifying their programs to adapt to the demands which we feel. At the same time, there are opportunities for many types of educational institutions to provide training for science information personnel. Incentives for the students are needed. Finally, and perhaps most of all, there is a major need for education of the public, of college teachers, and especially of students to the existence of opportunities in this field. The work is attractive and rewarding and the opportunities are great, both for the students and the educational institutions.

The Literature Chemist Encouragement and a Curriculum Needed

By HOYLANDE D. YOUNG
Argonne National Laboratory, Argonne, Illinois
Received June 6, 1962

My contribution to this program has been arrived at by a circuitous route. In August of 1961, Bill Waldo offered an invitation to participate in a symposium on "Creating Literature Chemists" sponsored by the Division of Chemical Literature. By early October the symposium had become "Educating Litérature Chemists" and the Division of Chemical Education had become a co-sponsor. At that time it was suggested that my part in this symposium take its cue from a "Conference on Training Science Information Specialists" held at the Georgia Institute of Technology because that meeting was of national interest. In addition, I could inject as much of my personal experience in technical information work as possible, and I might express my views on whether education of literature chemists really constitutes an academic discipline.

When the abstract was submitted in December it was my intention to stress the need for a suitable curriculum to train "literature chemists," and the need for education of their associates to an appreciation of the contribution which the literature chemists make to the end that greater incentives and greater rewards would be proffered to entice more, and more able people into this work.

Finally, the program arrived. There were no abstracts of the proposed papers but the titles suggested to me that by 10:45 a.m. Messers Waddington, Gordon, Langham, and Kyle undoubtedly would have covered the matters of curriculum and training quite adequately. Hence, in the few minutes allocated to me, I decided to make just a few suggestions in regard to a suitable curriculum, to urge adoption of standards, and to stress the desirability of greater rewards both in appreciation and in money. Most of all, I decided to avail myself of the opportunity to express some personal views.

* Presented before the Divisions of Chemical Literature and Chemical Education, American Chemical Society 141st National Meeting, Washington, D. C., March 21, 1962.

I came into technical information work through circumstance. There have been periods when taking this road has benefited me more than the laboratory road would have done; there have been other periods when the reverse was true. There have been times when I have found the work deeply satisfying; other times when I have greatly regretted not staying with experimental research, not proving myself in that area. The chemist who accomplishes in the field of experimental research, be it organic, inorganic, physical or other, usually earns the respect of a reasonable number of his fellow chemists even if he fails to get a high financial return. They understand what has been involved and some at least appreciate the significance of his findings. The chemist who concerns himself or herself exclusively with literature which reports the findings of others has until quite recently encountered few other chemists who considered the work or the contribution worthy of much recognition. What really counts with most scientists are creativity and basic science. The literature chemist is not considered to be creative nor is he expected to contribute much push to the frontiers of knowledge. How can he hope to win acclaim? In spite of this there have been very capable individuals who dedicated themselves to work of this kind and achieved so outstandingly that no other chemist could fail to appreciate the contribution.

Some able chemists recognized years ago how important the contributions of literature chemists could be. I recall one who told another highly-placed chemist-executive that his company would locate three people qualified to serve the company as vice-presidents before the company would find more than one who could organize the technical files and keep the company scientists abreast of pertinent developments in fields of interest. I also recall an instance when one of my co-workers, a tall, good-looking man with a B.S. degree, a couple of years of experience, and a bright future ahead, came over to my desk, planted one foot on it, and said, "Take a letter." My co-worker knew that I was a chemist with training and experience superior to his own. No doubt his income was greater—he was a young man on the way up—but I think that his error resulted primarily from the opinion shared with many another that no red-blooded scientist would be a "literature scientist"; a person who worked with literature might be any one of a great variety of people, but he could not be a *scientist*.

This misconception has been and continues to be supported, and in some cases supported by literature chemists. Have you ever heard one with training and experience remark that a clerical assistant "can do everything that I can do?" I have and I have heard this from enough persons over the years to convince me that one thing that literature chemists need is boosted morale and more self respect.

What is required to accomplish this? Fate and the inevitable course of human events appear to be helping. The rapid advance in scientific knowledge and technology has resulted in a "publication explosion" that has made it increasingly difficult if not impossible for an experimental chemist to keep up without aid. Some chemists are not situated so that they can get help from literature specialists, some fail to recognize the value of such assistance, but more and more persons concerned with the management of chemical companies have recognized the advantages and have sought to complement their experimentalists with literature scientists, men and women who combine knowledge of literature searching techniques, languages and science.

A couple of years ago, one of my co-workers visited a number of university science departments and a number of chemical companies to learn what he could of that which would be helpful to us in providing our people with better technical information services. Significant was the finding that the scientists at the universities were provided with library collections and the freedom to use them, but the industrial scientist was not given much chance to miss something in the current literature which might be of importance to his work. Quite a variety of devices were used to bring the literature to the experimental scientists—title lists, abstracts, circulation of journals, briefings, etc. One method which can be very effective and which you have heard described in this symposium, is that the literature chemist is made a member of the research team.

Have you seen the Directory of Selected Information Services in the Physical and Biological Sciences recently issued by the National Science Foundation? (NSF-61-68 Nov. 1961 GPO $1.75.) Peruse this directory and you will see that taken together there is extensive activity in specialized scientific information services in the United States.

The U. S. Government is big business and several government departments operate laboratories with science programs comparable in size with those of the large industrial laboratories. With responsibility for so much of what some have called "Big Science," it is not surprising that several of the government departments have pioneered in the use of literature specialists. The Atomic Energy Commission is one; ASTIA is another. A quite extensive technical information service is being set up for NASA.

Another assistance to the literature chemist in the United States has come from the U.S.S.R. The publicity and attention which have been given to The All-Union Institute of Scientific and Technical Information (VINITI) of the USSR Academy of Sciences and the State Technological Commission of the USSR Council of Ministers have served to make many people in the United States aware of the importance of the literature to the advancement of science and technology and to appreciate the need for scientists who have specialized in storing, retrieving, translating if necessary, and circulating the knowledge which is recorded in the literature. We may not believe (few of us have seen enough first hand to *know*) that the Russians have the ideal procedure for maximum return from the literature but there must be quite a few U.S.A. literature specialists who are pleased that the U.S.S.R. has helped to focus attention upon the important contribution of the literature specialists.

Evidence of the increasing attention to technical information work appears almost every day. Here are examples from one issue of *Science*:

1. The Air Force Office of Science Research has established a Directorate of Information Sciences to reduce the time lag between discovery of new knowledge and its useful application. The Directorate, consisting of an information sciences research division and a technical

information division, is headed by Harold Wooster of AFOSR's Directorate of Mathematical Sciences.

2. A two-week summer course in the use of *information theory in science and engineering* will be conducted at Dartmouth June 25 to July 6. The course is designed to acquaint engineers, scientists, and business adminstrators with a new approach to probability and statistical inference.

3. A conference on self-organizing information systems was held in Chicago from May 22–24. The conference, jointly sponsored by the Office of Naval Research and the Armour Research Foundation, covered the evolution of self-organizing systems with particular emphasis on research progress during the past three years.

This is just a sample. Meetings on this subject area are becoming so numerous as to be a hazard; if you work in Chicago and would like to attend a meeting in San Francisco, you are quite likely to be told to be patient because there certainly will be a similar meeting in Chicago soon. And to your disappointment such a meeting is announced in a few days.

Communication is one of the biggest, if not the biggest, need today. This is true from many points of view but here I am not speaking of communication with men in outer space, fantastically exciting as that may be; I am not speaking of communication between parent and child, husband and wife, teacher and pupil, specialists with laymen, Kennedy and Khrushchev. It is with exchange between scientist and scientist that the literature chemist concerns himself and with suitable qualifications he can make an invaluable contribution. Most scientists would prefer to exchange first hand, but time and money, more often than not, are too limited to permit the scientists to get together as often as they would like. Publication may be a partial answer; publications plus the literature chemist are even better.

It is difficult for me to talk in terms of a literature chemist. I think that limits the area of discussion too greatly. I prefer to talk about chemists who are also technical information specialists. But whichever term is used, there is need for definition. I would like to see some group such as this one undertake to define the area of activity and to set up standards. What work is appropriate? What is the value of the work? What constitutes quality of performance?

Recently, I noticed a short article by Leslie J. Nason of the University of Southern California attempting to explain why so many top students flunk out of law school. In selecting students, great reliance is placed on grades, entrance examinations, aptitude tests and intensive interviewing. It appears that a student may have a brilliant undergraduate record because he excels in areas of high creativity and little competition but he is unable to handle a more organized body of knowledge. The author stated that he had never known a top grade engineering student to flunk out of law school. Apparently, in order to be successful, a law student must be able to (1) organize knowledge as he reads; (2) organize and summarize knowledge as he listens; (3) present knowledge in an organized fashion in speaking.

It appears to me that to be successful in certain areas of the literature work requires abilities not too different from those which the author considered to be desirable for a good law student—abilities to absorb information quickly, to organize, to communicate. Creativity need not be a dominant characteristic for this work but there are areas where it is extremely important. How else will break-throughs be achieved for ways of organizing the flood of new technical information, storing it compactly, and recalling what you want when you want it. Some attention should be given to consideration of the qualities which are of most value for accomplishment in this field.

I would like to see a set of values for this work which makes clear that the information specialist himself recognizes the difference in qualifications and training required for the person who revises a table and the person who devises a satisfactory scheme for storing and retrieving recorded knowledge on, for example, the metabolism of proteins or the metallurgy of iron alloys. There should be appreciation for an editing job well done when the editor replaces each six-syllable word (exclusive of chemical names, of course) by six one-syllable words, but there should be recognition as well of the greater competence required to envision exciting possibilities for successfully treating a given type of cancer by the use of a given chemical compound in the course of coordinating the literature on the properties of a series of compounds. It can be done.

As long as the emphasis is on technical information people who are performing simple tasks of technical editing, straight abstracting, compiling of bibliographies and the like, technical information specialists will not achieve the status which brings generous rewards and high praise. Thelarger tasks are waiting. It is good that more and more people are recognizing them and recognizing that these will be handled best by people with special talents and special training for this work.

A documented account of some of the many occasions when a literature chemist has brought about greater overall accomplishment, saved money by preventing duplication, economized on the time of a brilliant experimentalist, sparked an outstanding idea, brought accomplishment more speedily, found something important that the non-specialist would overlook might be worthwhile. I do not have a copy of Harold Wooster's study but it appears that he has done just that. There should be more similar studies.

For the literature work, and especially for the most challenging kind, understanding of the fundamentals of communication is indispensable—linguistics, languages, mathematics; how to read with understanding; how to write to convey information precisely, concisely; knowledge of the newest techniques for abstracting; indexing, translating, storing and retrieving; and beyond the newest, sufficient understanding of the problems to suggest the possibility that you may be the one who provides the breakthrough to a better approach, one which no one has thought of to date. You will need training in science, in the special languages of science, in the habits of scientists, how they reason, how they experiment, how they communicate. You should be trained to observe every bit as accurately as the experimentalist, although he may be observing the reactions first hand and you, his record of observations. You should be trained to see relationships accurately. There is no room for the sloppy reader, nor the sloppy reporter. It is well to remember that all the

gain in speed and efficiency from mechanization and the use of machines is lost if the material which goes into the machine has not been skillfully and dependably selected to present the facts accurately and fully.

The literature chemist will gain, in my opinion, by recognizing clearly the opportunity which is open to him, bringing to the work all the intelligence, ingenuity and insight of which he is capable, doing his work with meticulous detail when meticulous detail is needed but never losing sight of the over-all canvas, and coming together to discuss problems, objectives, ways and means as we are doing here. This has been a great help and will I am sure continue to bring status and stature to those in this work.

As for compensation, the fastest way in which to get higher pay will be to advance to an executive. From that position, one can have a voice in determining salaries. When your turn comes, don't forget.

The literature chemist has interesting work, demanding work, important work. The problems awaiting solution are formidable; he should stand up and be counted.

Information Retrieval and the College Chemistry Curriculum*

By ELBERT G. SMITH
Mills College, Oakland 13, California
Received June 6, 1962

In recent years there have been geologic upheavals in chemical literature. The hills are being thrust up into mountains and these will soon become mountain ranges. The horse and buggy methods that were once good enough for getting around in gentler terrain are becoming less useful and soon will become impossible. There seems to be considerable recognition of this in chemical industry where something is being done about it by developing new methods for retrieving chemical information. As yet there seems to be little recognition of, or interest in, this problem in most of the colleges and universities in this country. Less than half of them even offer a course in chemical literature and only a very small proportion of the publications in the chemical information retrieval field is coming from academic institutions. I should like to discuss some of the reasons for this situation and what might be done about it, but first it seems desirable to review for my academic colleagues some current developments in chemical information retrieval that are taking place in industry.

As a chemical company's file of compounds grows, it becomes increasingly evident that the difficulties of nomenclature and the catch-all nature of molecular formula indexes make indexes based on these principles increasingly helpless, or expansive, or both, in dealing with the company's requirements for information. New kinds of indexes are being tried, such as coordinate indexes or permuted indexes. Machines are being used to do information retrieval jobs that were formerly impossible and even unimaginable. In at least one company a chemical notation system is being used for indexing company compounds and this index has completely displaced both nomenclature and molecular formula indexes. I shall illustrate this general movement within industry with examples from a small area of chemical information retrieval with which I happen to be most familiar.

One pressing problem in the pharmaceutical industry and in a growing number of others is that of generic chemical structure retrieval—locating all those compounds in the company's file that have the same functional groups, or ring structures, or chain structures, or logical combinations of these features. Indexes based on nomenclature are utterly powerless to cope with this problem and even inverted molecular formula indexes, once thought to be a solution, aren't really of much use for most structural features and in any event bog down when an index grows to any considerable size. Files of 10,000 to 50,000 compounds are not at all unusual in these industrial concerns and at least one company forsees a file of 100,000 compounds.

The most common approach in industry to this problem of generic structure searching has been the fragmentation of structural formulas into bits and pieces which can be assigned code numbers which in turn can be manipulated with punch card machines of varying complexity to produce lists of the compounds that contain the particular combinations of structural units required in a given problem. Most of these codes trace their descent, one way or another, from the pioneering work of Donald Frear, Karl Heumann and others who developed the CBCC code[1] more than 15 years ago, and from the work of Fred Whaley[2] who simultaneously simplified these practices and developed advanced ideas of logical manipulation. Starker and Cordero[3] recently have published a typical example of one of these fragmentation codes and Geer and Howard[4] have listed other examples that have been proposed and sometimes used.

Another approach to this problem was pioneered at Monsanto Chemical Co. by W. H. Waldo,[5] who has devised a way of putting edited structural formulas directly into a computer and of searching these stored structures with a suitable computer program to print out the structures of those compounds meeting the search requirements. This approach is being developed in at least two other companies.

A third approach has been pioneered by Howard Bonnett[6] at G. D. Searle & Co. and was described by him a year ago before the Division of Chemical Literature at the St. Louis meeting. In this approach structural formulas are translated into Wiswesser notations which then can be arranged, by the simpler kinds of punched card machines,

* Presented before the Divisions of Chemical Literature and Chemical Eduction, American Chemical Society 141st National Meeting, Washington, D. C., March 21, 1962