(91) Kirschner, Stanley "Chemical Literature Information Retrieval in Departments Without an Expert". "Abstracts of Papers", 184th National Meeting of the American Chemical Society, Kansas City, MO, Sept 15, 1982; American Chemical Society: Washington, DC, 1982; CHED 51.

(92) Skolnik, Herman "Relevancy of Chemical Literature in the Educational Process". *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 95-97.

(93) Wilen, Samuel H. "Unresolved Problems and Opportunities in Chemical Literature Teaching". *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 112-115.

Teaching of Chemical Information Science to Graduates

CHARLES H. DAVIS

Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign,
Urbana, Illinois 61801

Received March 13, 1985

An overview of graduate education in information science is presented, with special attention to the teaching of chemical information science in graduate schools of library and information science. These curricula are contrasted with those in computer science; master's programs are compared with doctoral studies, and a promising future is foreseen for research and development.

INTRODUCTION

Ten years before the appearance of this journal and at approximately the time that UNIVAC I was made commercially available, the American Library Association's Committee on Accreditation (COA) coincidentally decided that the master's degree would be the first professional degree in library science. For the past 35 years COA, acting under the authority of the Council on Post-Secondary Education, has maintained that only those individuals well grounded in the arts and sciences would be admitted to these master's degree programs. The purpose has been to defer the teaching of professional matters involving library and information science until after students have majored in traditional disciplines including, but obviously not limited to, chemistry. The principal models that inspired this approach were medicine and law.

From time to time, other professional associations and societies in information processing have expressed an interest in accrediting programs or certifying individuals, but the Council on Post-Secondary Education has ruled that the COA would have sole authority for accrediting professional programs in librarianship and closely related areas. To its credit, the COA has sought input from sister societies such as the American Society for Information Science (ASIS) and the Special Libraries Association (SLA), and members from these and other associations serve both on COA and on site teams that visit and assess graduate programs throughout the U.S. and Canada.

Some chemistry graduates subsequently enter programs in computer science; however, the majority of chemistry librarians and formally trained chemical information specialists graduate from schools of library and information science, the best of which encompass both computer-based systems and traditional methods of information handling. While departments of computer science sometimes have courses or programs in information retrieval, the majority continue to address problems of hardware and/or applied mathematics—usually depending on whether the parent unit is an engineering school or a department of mathematics. Not to be overlooked in all of this is the simple fact that many employers insist on a degree from a school whose program has been accredited by the COA.

A complete account of the historical reasons for the existing situation in education for information science is clearly beyond the scope of this paper. However, those readers interested in pursuing the topic are referred to Shera's classic work entitled *The Foundations of Education for Librarianship*, in which he discusses the interrelationships between and among traditional librarianship, documentation, and information science.[1]

Charles H. Davis received his B.S. degree in chemistry from Indiana University in 1960 and studied at the University of Munich as a German Government Fellow in 1960 and 1961. He worked in the Organic Index Editing department of Chemical Abstracts Service from 1962 until 1965, after which he resumed his graduate studies at Indiana, earning an M.A. and a Ph.D. in library and information science in 1966 and 1969 with a graduate minor in toxicology. He taught as an Assistant Professor at Drexel University (1969-1971) and as an Associate Professor at the University of Michigan (1971-1976) and has served as Dean and Professor at the University of Alberta, Canada (1976-1979), and at the University of Illinois at Urbana-Champaign (1979-1985). He plans to resume full-time teaching and research after a sabbatical leave during the coming academic year.

STRUCTURE OF THE PROGRAMS

Herman Skolnik, the first and long-time editor of this journal, once wrote that chemical information science "...is without a bridge to academic chemistry departments."[2] And it is true that few of the master's level programs described permit a real major in chemical information science, choosing rather to concentrate on scientific literature, science documentation, and general theoretical and practical matters of concern to everyone involved in information retrieval.

0095-2338/85/1625-0323$01.50/0

At least one notable exception has evolved at Indiana University, where extensive cooperation has taken place between chemistry librarians, the chemistry faculty, and specialists in data processing. This program has been described in an interesting paper by Wiggins.[3] Nevertheless, the assertion by Skolnik is essentially accurate and represents a challenge to chemists and information specialists alike, who must cooperate if the special problems of chemical information retrieval are to be addressed properly and in depth. What is the nature, then, of these other graduate programs that obviously address at least some of the problems faced by specialists in chemical information science?

In the first place they are definitely generalist programs requiring all students to know the fundamentals of both traditional and contemporary techniques of information organization, storage, retrieval, and dissemination. Specialization comes afterward. It has been fairly well established that there is a core of intellectual problems that must be addressed regardless of the devices one employs to perform this set of tasks, and dealing with such matters is definitely the long suit of these master's degree programs. The intellectual problems include, but are certainly not limited to, establishing authorship of works, classifying and determining the subject content of documents, providing satisfactory physical descriptions of items for identification or archival purposes, determining the optimum method for storing and retrieving materials, performing question negotiation to determine the actual (as opposed to the stated) needs of clients, planning search strategies (often involving a mix of media such as print, audiovisual, and computer-based materials), and implementing techniques permitting evaluation of the services provided.

This turns out to be a nontrivial assignment. In fact, many schools are now exploring means of lengthening their programs so that these fundamentals can all be covered and so that a decent measure of specialization can be provided as well. The tradiational length has been 1 full year, typically consisting of two semesters and a summer session, or else four quarters, depending on the calendar of the parent university. Within the last few years, some schools have elected to require a 2-year master's degree program, while others have chosen to provide an optional second year and a second credential, often called a Certificate of Advanced Study. In any case, even though the curricula may differ in some ways, all of the master's programs accredited by the ALA provide an introduction to the foundations of library and information science, to the techniques related to the organization and storage of materials, and to information retrieval and the provision of traditional and contemporary information services.

SPECIALIZATION IN SCIENCE AND TECHNOLOGY

Although specialization in chemical information science per se is rare, specialization in the literature and information services of science and technology assuredly is not. In fact, because of its long history and reputation for excellence, chemical documentation usually represents a major component of such areas of concentration. Time did not permit a survey of all the schools; however, the University of Illinois course entitled "Scientific and Technical Literature and Reference Work" (LIS 412) is probably representative. Its stated aims are (a) to acquaint students with typical library materials in science and technology and (b) to develop proficiency in their selection, evaluation, and use for reference work. A prerequisite of this course is LIS 320, "Introduction to Information Sources and Services", a course required of all students that introduces information referral techniques, reader's advisory and on-line information services, and that examines representative printed and on-line sources as well as developing question negotiating skills and search strategies.

The text for LIS 412, which is recommended rather than required, is *Science and Engineering Literature: A Guide to Reference Sources* 3rd ed., by Malinowski and Richardson.[4] Assignments include evaluating encyclopedias and dictionaries as well as indexing and abstracting serials, tracing the scientific publication cycle, performing literature searches in three of six areas such as chemistry, a case study, an on-line search, and a term project.

"Chemistry" is defined in accordance with the McGraw-Hill *Dictionary of Science and Technical Terms* as "The scientific study of the properties, composition, and structure of matter, the changes in structure and composition of matter, and accompanying energy changes." Its scope includes biochemistry, chemical engineering, materials (adhesives, fuels, paints, and the like), and metallurgy, as well as analytical, inorganic, organic, and physical chemistry.

For the field of chemistry alone, the current *Study Guide* for the course lists 6 articles or journals for background reading (including the *Journal of Chemical Information and Computer Sciences*), 9 bibliographies or guides to the literature, 14 encyclopedias, 9 dictionaries, 19 handbooks, 1 special set of chemical tables, 6 directories, 4 abstracts or abstracting services, *Chemical Titles*, the American Chemical Society Primary Journal Database, CHEMDEX, CHEMLINE, and CHEMNAME.

A number of other elective courses are usually recommended to individuals expressing an interest in such areas as chemical information science. Typical of these are courses dealing with automation or data processing, government publications, information storage and retrieval, on-line information systems, library service to specialist users, medical literature and reference work, advanced bibliography, measurement and evaluation of library services, telecommunications, research methods, information management, bibliometrics, and techniques for managerial decision making in library and information science. Students are also counseled to take cognate courses in other schools and departments as appropriate—typically in computer science, business administration, communications, and linguistics.

DOCTORAL PROGRAMS

Thus far, the programs described have been designed primarily for practitioners—those destined to work in government, business and industry, abstracting and indexing services, or a variety of special libraries and technical information centers. Some graduates, especially those with substantial prior experience, may also choose consulting or some kind of entrepreneurial venture.

For those interested in research or teaching at the university level, there are over 20 schools in North America that offer the Ph.D. in library and information science. In addition, one or two offer the D.L.S. (Doctorate in Library Science) or the D.A. (Doctor of Arts). The D.L.S. is technically a professional rather than an academic degree, but it often has quite similar requirements; the D.A. is intended primarily for those wishing to pursue careers in teaching, and it may therefore have a less rigorous research component. In any case, entrance to such advanced programs normally entails the possession of an accredited master's degree (or its equivalent in a closely related field) and sometime suitable professional experience as well.

Given the highly eclectic backgrounds of those drawn to library and information science, dissertation projects range from traditional historical and bibliographic works through social and behavioral studies to systems-analytic projects involving operations research. Depending on the strength of an individual's background and the nature of the dissertation project, approximately 3–7 years are required to complete most doctoral programs.

A NOTE ON THE NEED FOR INFORMATION SPECIALISTS

Chemists are smart people who can do much of their own information retrieval. However, everyone reaches that famous point of diminishing returns—the point at which any individual, particularly a person doing scholarly research, will be spending too much time searching the literature. Information specialists save time and provide more thorough service precisely because they know more about information sources and systems and how to use them. They are trained to get things done efficiently that others might accomplish only after an inordinate amount of "hacking". Moreover, the problems are not limited to understanding hardware or system protocols; they also include intellectual challenges such as those involving nomenclature, vocabulary control, and the intelligent selection of search strategy to optimize retrieval performance.

Scientists on the cutting edge of research may also become arrogant. This is implicit in what Derek Price called the "invisible college", that group of scholars to be found in most scientific fields—perhaps all fields—who typically satisfy their information needs by consulting each other. It is asserted that their knowledge of the literature is so great that they have no need for it whatsoever.[5] Regrettably, experts in one field are often little better than laymen in others—including some that may be highly pertinent to the research in question, especially in this era of interdisciplinary work. Moreover, the experts' dependency on the literature, whether in machine-readable form or not, becomes proportionately higher the farther they stray from their original areas of expertise.[6,7] A personal case history will help illustrate the point as well as indicate why I have strong feelings on this subject.

Case History 1. During the 1950s I was lucky enough to have secured a National Science Foundation grant for undergraduate research participation, and I was given laboratory space with both master's and doctoral students. One of the most frustrating problems these graduates were having involved synthesizing the methyl ketone of 1,2,5-thiadiazole, a five-member heterocycle not reported until the mid-1950s. It seemed that all efforts were directed at the standard method of using a Grignard reaction starting with the acyl chloride derived from the ring's monocarboxylic acid, a compound they had already synthesized and with whose properties they were familiar. Unfortunately, every time they ran the Grignard reaction they broke open the ring. They tried other methods and reagents, including dimethylcadmium—a common and more gentle substitute for the magnesium compounds normally associated with Grignard reactions—all to no avail. In my naivetê, I looked through *Chemical Abstracts* and discovered an old but interesting method that involved using the acyl halide with sodium diethyl malonate. The organic groups are joined, and sodium chloride is "split out"; one then heats the resulting compound in hydrochloric acid until didecarboxylation occurs, and—voila—one has the methyl ketone without having torn open the ring! The faculty and the graduate students had assumed wrongly that they were familiar with the literature. It was a lesson I never forgot, and it convinced me that there was a legitimate place in the world for chemical information specialists.

Good information specialists also have a way of capitalizing on serendipity. They have wide-ranging interests and are sensitive to the relevance of seemingly unrelated facts and materials. They are pattern recognizers who put together pieces that others might ignore, and they provide themselves with a wide array of source materials—in short, they browse. Another case history will illustrate the point.

Case History 2. During the mid 1960s, after I had left Chemical Abstracts Service to resume graduate studies, I was faced with the task of supporting myself and elected to work for the Aerospace Research Applications Center at Indiana University. We were doing all manner of information retrieval with the old NASA, AEC, and Commerce tapes, with an eye toward showing their importance to business—especially small businesses in Indiana. It was this sort of spin-off that the government wanted to help justify the large expenditures it was making on the space program at that time. In any case, every once in a while an off-the-wall question would arrive, one of which was a request for information on plastic horseshoes—an assignment that I received and did not particularly welcome. I did perfunctory searches through the NASA files, hoping, I suppose, that they would have some off-beat "Project Pegasus" or something, but the result was always the null set. Then, one day, I was browsing through *Chemical & Engineering News*—an activity I still engage in and recommend—when I suddenly spotted an advertisement showing two absolutely adorable kittens peering out of what were unmistakably plastic horseshoes! The shoes were made by a manufacturer who catered to mounted police in major metropolitan areas where park patrol by equestrian police was common. As it happened, the client was interested in lighter shoes for horse racing and could not use this particular information, but he was deeply impressed that we had found anything at all. As a reward, I think I was given every off-beat question during the remainder of my tenure at the Center—a flattering but unwanted result.

Insufficient Information Can Be Dangerous. More serious are the consequences when chemists perform incomplete searches of the literature, a situation exacerbated by having substantial newer material in machine-readable form while much older and still valuable documents remain in printed form. The temptation is to assume that only the most current information is of value, and chemists may content themselves with on-line searches when they should also examine older sources for the sake of thoroughness.

This point was driven home by Nair in a recent letter to the editor of *Chemical & Engineering News*, in which he responded to earlier warnings about the danger of acetone peroxides:

> "[T]he apparent reluctance to do thorough literature surveys to discover safety considerations is quite pertinent. This is one aspect of research that the younger generation...may easily neglect. ...Given the track record of the chemical industry on safety of chemicals or at least the public perception of it, neglect of older literature can lead to unfavorable consequences when least anticipated. Modern investigators need an awareness that chemical safety data do not always lie in the modern computer data banks; if you don't find [them], do not assume all is well, that no safety data had ever been developed. It takes a little longer to do it the old-fashioned way, but the results may be valuable and future mistakes may be avoided."[8]

As more and more data are converted to machine-readable form, our databases will of course become more reliable. For the time being, however, one would probably be well advised to enlist the help of a professional chemical information specialist when doing such crucial searches of the literature. Moreover, it is highly probable that such an individual will do a better job than the average researcher even when the stakes are not so high, and the researcher will have more time to spend on those things for which he or she may be uniquely qualified.

INFORMATION RESEARCH AND DEVELOPMENT

On the other hand, there is considerable hope for the future in designing systems that are more user friendly and that will allow the end users to do more of their own information work

326 J. Chem. Inf. Comput. Sci. 1985, 25, 326-333

if they wish to do so. Changes in hardware and software—especially hardware—hold out the prospect for faster machines with remarkable storage capacity. We have already seen the introduction of microcomputers that have gone from 8- to 16- and even 32-bit words, thus blurring the distinctions between and among micros, minis, and mainframes. Among other things, this means that we can already have personal computers for home, school, and business that represent most of what we need for stand-alone jobs, and which can also be used as intelligent terminals for interfacing with other computers through existing networks. Many of the inconveniences associated with complicated protocols and sign-on procedures can be eliminated by using these micros intelligently. The gradual replacement of copper wire with fiber optics will allow us to transmit vastly greater quantities of data while also providing greater security for the transmitted messages. And laser disks hold out the promise of extremely dense storage for both pictorial and textual material, perhaps permitting us to download entire sections of on-line encyclopedias or other information sources.

In addition, there is hope emanating from the field of artificial intelligence, which has already introduced so-called "expert systems" into a number of areas—including medicine, where routine preliminary diagnoses can now be initiated by computers. And there is also progress in voice recognition, which would release us from the constraints imposed by keyboards—still the most efficient and cost-effective method for communicating on-line with computer-based systems.

These and other exciting possibilities are being explored by a variety of people in a number of settings, sometimes in industry, but often in our universities, where those teaching information science at the graduate level are also engaging in research.

REFERENCES AND NOTES

(1) Shera, J. H. "The Foundations of Education for Librarianship"; Wiley: New York, 1972.

(2) Skolnik, H. "Chemical Information Science in Academe". *J. Chem. Inf. Comput. Sci.* **1980**, *20* (2), 2A.

(3) Wiggins, G. "The Indiana University Chemical Information Specialist Program: Training the Library User and the Librarian". *Sci. Technol. Libr.* **1981**, *1* (3), 5-11.

(4) Malinowski, H. R.; Richardson, J. M. "Science and Engineering Literature: A Guide to Reference Sources", 3rd ed.; Libraries Unlimited: Littleton, CO, 1980.

(5) Price, D. J. de S. "Little Science, Big Science"; Columbia University Press: New York, 1963; pp 90-91.

(6) Davis, C. H.; Rush, J. E. "Information Retrieval and Documentation in Chemistry"; Greenwood Press: Westport, CT, 1974; pp 9-10.

(7) Davis, C. H.; Rush, J. E. "Guide to Information Science". Greenwood Press: Westport, CT, 1979; pp 7-14.

(8) Nair, J. H. "Chemical Safety Literature in Data Banks". *Chem. Eng. News* **1985**, *63* (5), 4.

DARC System for Documentation and Artificial Intelligence in Chemistry

JACQUES-EMILE DUBOIS* and YVES SOBEL

Association pour la Recherche et le Dêveloppement en Information Chimique, 75005 Paris, France, and
Institut de Topologie et de Dynamique des Systêmes, associê au CNRS, Universitê Paris 7,
75005 Paris, France

Received June 4, 1985

The DARC System deals with structural information both for documentation and for artificial intelligence (AI) endeavors in chemistry. Its topological concepts are briefly reviewed in conjunction with the creative data needs in knowledge information processing systems (KIPS). Knowledge base, inference engine, and user interface are discussed with reference to the DARC potential in the field of AI and expert systems. AI methodology and its impact on knowledge production are reviewed. New chemical computer-aided design (CAD) tools to develop more creative and innovative research in synthesis planning, structure elucidation, and prediction in drug design are no longer pure prospective challenges.

INTRODUCTION

Chemistry is an excellent testing ground for serious development in the years ahead, both in cognitive science and in artificial intelligence. It has a long association with graph and combinatorial theories. Chemical facts, well structured and diverse, can be managed easily by a class of structural languages adapted to various needs ranging from bibliographic data to the most advanced correlations and to modelization connecting morphological facts to properties, reactivities, and synthesis. Moreover, it is, by tradition, very user friendly, and its ideograph presentation is naturally adapted to graphics display.

This paper was conceived as a follow-up of "French National Policy for Chemical Information and the DARC System as a Potential Tool of This Policy". Born about 20 years ago, the DARC System was presented in the *Journal of Chemical Documentation* in 1973 for its documentation potential, but its second facet, Automatic Research of Correlations (ARC), was, at that time, left in the dark!! Indeed, concepts and tools of artificial intelligence were still in their infancy. Artificial intelligence was then less than 17 years old and struggling for recognition. The concept of "decision support systems" that we have used in chemistry has also been developed only very recently.

DARC Concepts. The *Journal of Chemical Information and Computer Science* (JCICS) Silver Anniversary Issue is an excellent occasion on which to sum up a field whose recent progress has been significant, with the development of much basic formalization. DARC concepts and tools for artificial intelligence are based on more structured knowledge than is used by either yesterday's or today's expert systems.

The same concepts and tools reflect the choice of an associative approach, one that relates factual documentation to achievements of artificial intelligence. In many fields, the passage from one to the other is still difficult, and communication between documentation data and inference-creating information usually remains a problem. In chemistry, factual data banks (FDB) must be linked to knowledge banks, since

0095-2338/85/1625-0326$01.50/0