features have dimensions of the order of 10^8 light-years. These features appear to be in a midway stage of evolution. Concentrated clusters of galaxies may be very recent formations.

The high angular momentum per unit mass in spiral galaxies indicates that at the time of their detachment—long after the ending of the fireball stage—the universe must have been in a vehemently turbulent state. After their detachment the protogalaxies first expanded; they became stellar systems only upon their recollapse, which may have taken place when t/t_0 was approximately between 0.1 and 0.2 (t_0 being the present age of the universe). This period may coincide with that of the high frequency of powerful radio sources.

In the last section a very cursory review is given of the large instability phenomena displayed by some galaxies. These appear to come from the nuclei. There are indications that the nuclei of most large galaxies are from time to time susceptible to enormous violent activity.

References

1. A. A. Penzias and R. W. Wilson, *Astrophys. J.* **142**, 419 (1965).
1a. G. Gamow, *Rev. Mod. Phys.* **21**, 367 (1949).
2. J. H. Oort, *Nature* **224**, 1158 (1969).
3. H. Shapley and A. Ames, *Ann. Harvard Obs.* **88**, 43 (1932).
4. J. H. Oort, in *La Structure et l'Evolution de l'Univers, Inst. Phys. Solvay* **11** (Brussels, 1958), p. 163.
5. V. Icke, personal communication.
6. F. Hoyle, in *Problems of Cosmical Aerodynamics*, I.U.T.A.M. and I.A.U. Symposium on the Motions of Masses of Cosmical Dimensions (Central Air Documentation Office, Dayton, Ohio, 1951), p. 195.
7. P. J. E. Peebles, *Astrophys. J.* **155**, 393 (1969).
8. J. H. Oort, *Astron. Astrophys.* **7**, 381 (1970).
9. L. M. Ozernoy and A. D. Chernin, *Sov. Astron.* **11**, 907 (1968); *ibid.* **12**, 901 (1969).
10. M. Ryle, *Ann. Rev. Astron. Astrophys.* **6**, 249 (1968).
11. J. H. Oort, in *The Nuclei of Galaxies* (Pontifical Academy, Rome, in press).
12. P. C. van der Kruit, *Astron. Astrophys.* **4**, 462 (1970).
13. G. W. Rougoor, *Bull. Astr. Inst. Netherlands* **17**, 381 (1964).
14. P. Morrison, *Astrophys. J.* **157**, L 73 (1969).
15. V. A. Ambartsumian, in *La Structure et l'Evolution de l'Univers, Inst. Phys. Solvay* **11** (Brussels, 1958), pp. 241–274; compare also *The Structure and Evolution of Galaxies, Inst. Phys. Solvay* **13** (Brussels, 1965), pp. 1–12.
16. R. Minkowski, in *Radio Astronomy* (I.A.U. Symposium No. 4), H. C. van de Hulst, Ed. (Cambridge Univ. Press, London, 1957), pp. 111–114.

Computer-Based Chemical Information Services

Some new aids for the research scientist are described.

Edward M. Arnett

It may come as a shock to many research chemists to realize that, from one point of view, all of their activities are concerned with chemical information—its acquisition, evaluation, storage, retrieval, and transmission. Unlike the chemical engineer or the chemist concerned with applications, the research chemist produces no material product. The gathering and manipulation of knowledge is his *raison d'être*. Enormously increased resources, in terms both of manpower and of automation, which have been given to research have produced such a flood of useful and relevant knowledge that we are all uncomfortably aware, not only of our intellectual and psychological limitations for assimilating this embarrassment of riches, but of the physical difficulties of sorting, storing, and retrieving the information that we might need to assimilate. Chemists (far in advance of workers in many other disciplines) have realized for some time that some of the manpower and automated facilities used for *gathering* information should be assigned to its efficient management.

In the forefront of this effort have been the professional chemical societies in North America, Western Europe, the Soviet Union, and Asia, working in collaboration with their respective governments. In the private sector, many of the big chemical companies have developed advanced computer-based systems for storing and retrieving internally generated research and development information and for patent searching. Lagging far behind have been the universities. One can identify at least three reasons for this gap.

1) University science libraries serve a broader spectrum of users, with more general goals, than do the libraries of mission-oriented government agencies or most industrial organizations, where the range of interest is, for the most part, clearly defined over a period of years. Furthermore, the libraries of agencies and industry are more accustomed to playing an active role in performing searches than university libraries are.

2) Academic chemists and teachers are primarily interested in the behavior of molecules and students, respectively, and are less interested than information scientists in manipulation of the symbols used to store material describing how molecules behave.

3) Up until the past 2 or 3 years, computer-based chemical information had relatively little to offer most academic chemists; they were correct in their decision, however unconscious it may have been, to wait until it did.

There have been considerable gaps, both in implementation and in credibility, in the chemical information field in the past, and many of the services which might, at first thought, seem to be ideal (for example, interactive searching, through a terminal, of the entire textual material in *Chemical Abstracts*) are so expensive or so impractical that they have not been developed and probably will not be attempted for quite a few years. I hope, here, to give a brief picture of the information needs of research chemists and to describe briefly the present availability of computer-based systems for handling them.

The author is professor of chemistry at the University of Pittsburgh, Pittsburgh, Pennsylvania; adjunct senior fellow of the Mellon Institute, Pittsburgh; and director of the Pittsburgh Chemical Information Center. This article is adapted from a paper presented in April 1970 at the Conference on Chemical Information, sponsored by the National Academy of Sciences, in Washington, D.C.

Organizing a Chemical Information Experiment Station

My experience has been obtained the hard way, through the Pittsburgh Chemical Information Center, which was developed with support from the National Science Foundation and the Pennsylvania Science and Engineering Foundation as a place where new chemical information systems could be implemented and evaluated, under competent professional direction, in an experimental mode, and where successful services could then be provided to outside users on a pay-as-you-go basis in a regional dissemination center. Figure 1 shows the organization of the Center. From this it is obvious that a complex, multidisciplinary effort has been required to meet our goals. Let me elaborate on this briefly.

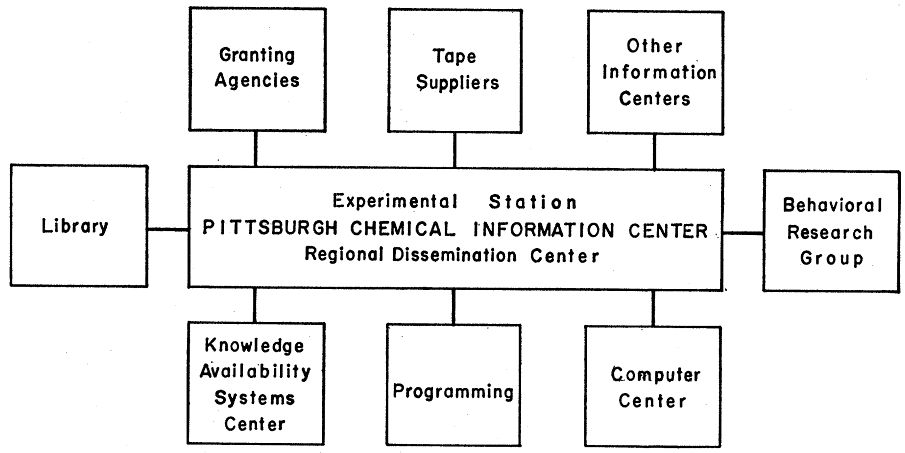

Fig. 1. Pittsburgh Chemical Information Center organization chart.

Development of computer-based chemical-information systems requires an application of the digital computer entirely different from the computational mode with which most chemists, especially physical chemists, are familiar. The programs and techniques used are much closer to those used for linguistic analysis, banking, or life insurance than they are to those used for "number crunching" in quantum mechanics. Data bases for broad-scale chemical information services are developed primarily at Chemical Abstracts Services (CAS), in Columbus, Ohio, and at the Institute for Scientific Information (ISI), in Philadelphia.

Tapes bearing the information in *Chemical Titles* or the header material (the article title, journal citation, and author's name) from *Chemical Abstracts* are delivered to the Center at weekly or biweekly intervals. A medium-sized digital computer (*1*) (of at least IBM 360/40 capacity) and a competent computer-center staff are needed to handle them. Although CAS supplies a workable search program for these two tape systems, considerable improvement in speed and efficiency for large-scale processing is possible through reprogramming. This is a formidable job of applications programming for which only a few university computer centers are prepared. The data processing group at the University of Pittsburgh includes three full-time applications programmers and a systems programmer. At the University of Nottingham, in England, there are four programmers. The IBM TEXT-PAC storage and retrieval system, with which we are at present experimenting at the Pittsburgh Center, requires about 40,000 cards in machine language and an estimated 12 man-years of programming time when it arrives from IBM to bring it to its present level of efficiency; in addition, considerable reprogramming is required to make the system compatible with the user's system and data base. Obviously, a large critical mass of facilities and people is required before the first research chemist can be served at all. This suggests that, now and in the immediate future, computer-based search services will be supplied to research chemists from a relatively small number of regional dissemination centers, which will cooperate with each other in the development, first of a national, and then of an international, network.

Information Handling Needs of Research Chemists

Now let us consider various types of information needs of research chemists. Chemistry is an experimental science, and it is indeed a rare chemist who does not at some time need to refer to the primary journal literature for information—whether it be in the form of numerical data, synthetic methods, or narrative descriptions of observations and theories. Most experimentalists working in a given field for a long period also develop a mass of data of their own, details of which become increasingly difficult to keep freshly in mind over a period of, say, 7 or 8 years. We have found that the information retrieval and storage needs of chemists depend to a considerable extent upon their particular fields of work and upon their particular research styles and personal habits; it would therefore be foolish to generalize or to impose on all chemists rigid standards or patterns of information use. Nonetheless, I have tried to present in Fig. 2 a schematic analysis of the general uses of information which applies to almost all chemists, from the most theoretical to those most concerned with applications. The central area of Fig. 2 represents the active interests of the group at a certain time; this encompasses the information that they are acquiring in the laboratory, are evaluating from the outside literature, and are discussing together. At the left of the diagram is the information they bring in from the library, the current literature, and other services. At the upper left is a block labeled "retrospective searches," the type of search through the older literature up to the present which a scientist makes when he is entering a new field. This would include an examination of the subject, formula, and author indexes of *Chemical Abstracts* (back to 1907, if the field were an old one) and examination of *Beilstein* (or perhaps *Landolt-Bornstein*), review journals, appropriate books, and then a deeper search through footnotes, finally augmented by the use of *Science Citation Index*. Such a search encompasses an enormous amount of information and is much in demand by users of the Pittsburgh Center. Some other centers have found it to be not economically feasible on a large scale at this time, but low-cost retrospective searching is an active area of research at the Pittsburgh Center.

Data Searches

Search for numerical data, such as refractive indexes, rate constants, and heat capacities, is a frequent activity of research chemists and can be very time-consuming. One might hope that all the reliable data in the literature will soon

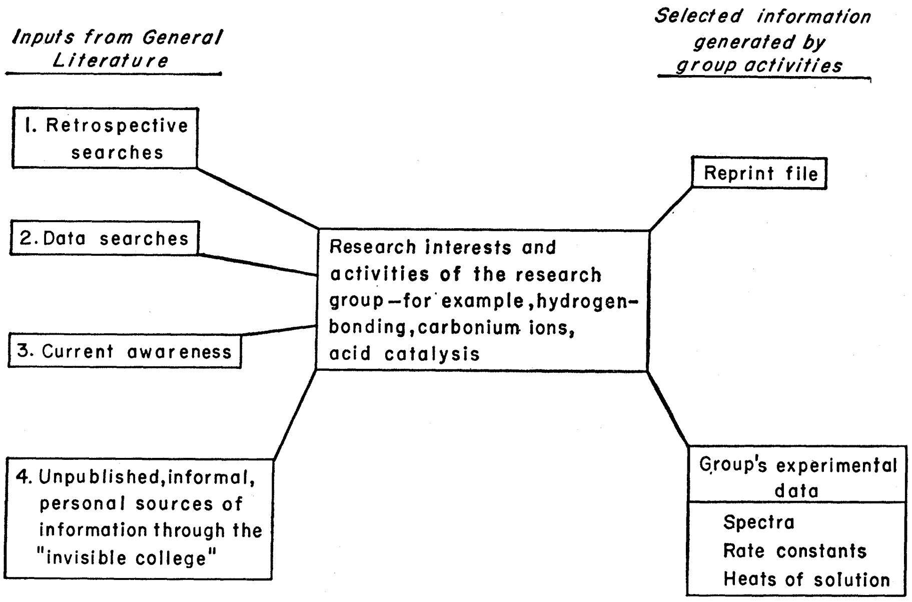

Fig. 2. Chemical information activities of research group.

be organized in computer-searchable form, along with structural information about the compounds to which they apply. By suitable programming of the computer the data might even be manipulated for running trial correlations and for other purposes. The task of storing and relating all this information in a useful manner places such demands on manpower and computer capacity (not to mention financial support) that we will probably never see exactly this type of system operative. However, the National Bureau of Standards already has under way a program of information centers devoted to various collections of data which have been subjected to critical evaluation by experts, and it is planned to link these together in a network so that such information will be on call for rapid access and dissemination. Within these limits, the most important data will probably be accessible in the next few years through such a system, but the use of human intermediaries in the centers and in the networks will certainly be important, and training for such work may become an important part of science teaching programs.

Current Awareness

Once a chemist has become familiar with the background literature in a field and decides to stay with that field, he is faced with the problem of keeping up with new developments as they are published. This is often a formidable task. There are at least 600 primary journals in which most of the chemical literature is published. Abstracts from over 11,000 more journals related to chemistry are occasionally included in *Chemical Abstracts*. An active academic chemist, with broad-ranging interests in organic chemistry or in biochemistry, may find that information commonly of use to him appears regularly in 20 or 30 journals and that another 50 or 60 occasionally contain articles of interest to him.

It is in this area of current awareness that the greatest advance in computer-based service for chemists has been made in the last 5 years. Inexpensive and remarkably effective current-awareness services are now available from a number of chemical information centers in the United States and abroad. Pioneering work is being done by Anthony Kent at the United Kingdom Chemical Information Center. In the United States, the University of Georgia, the Illinois Institute of Technology Research Institute, the University of Louisville, and our own group at the University of Pittsburgh are spinning *Chemical Abstracts* tapes to give current awareness service on *CA Condensates, Chemical Titles*, and other systems in fields related to chemistry. The Institute for Scientific Information, in Philadelphia, provides a number of alerting services such as *Current Abstracts of Chemistry, Index Chemicus*, and a special computer search system, *ASCA*, which is based on citation footnotes as well as on key words and authors—the usual search mode for *CA Condensates* and *Chemical Titles*.

The average price for a year's service for any of these services runs between $100 and $200; this represents a small part of an average research budget and saves from 1 to 10 hours a week of valuable time which might otherwise be spent in flipping pages in journals or scanning hard copy of *Chemical Titles*. We have found that very complete coverage is obtained through a combined approach based on *Chemical Titles* and *CA Condensates* searched in the University of Pittsburgh Chemical Information Center and the *ASCA* service based on citation footnotes obtained from ISI. All of the graduate students and postdoctoral fellows who have worked with it have told me that they are "hooked on the service" and have found it valuable in saving their time. We strongly recommend such a service to other chemists, provided it is used as a backup for their normal browsing of the four to six key journals which they have found, through experience or by examination of their computer printouts, to be the best hunting ground for useful papers.

Current-awareness tapes may be searched nowadays through the use of several different types of search terms common to the header material in *Chemical Abstracts*—author's name, key words in the title, author's country or place of work, coden designation of the journal, and citation. In addition, through the *Index Chemicus Registry System*, tapes produced by ISI, or a special CAS subfile—the Common Data Base (at present under study at the University of Georgia)—it is possible to make a search in which structures or even substructures of important compounds are used as search terms. Since over 90 percent of the chemical literature is structure-oriented, and since most structures (unlike names) are absolutely unambiguous, structure searching is the most attractive searching tool for the chemist. The coding and searching involved for the 4 million compounds believed to be named or discussed in the literature is a formidable task. Although structure searching is rapidly gaining acceptance in a number of special industries (for example, the pharmaceutical chemicals industry), it is not, here and now, a valuable tool for most academic chemists. Because of the enormous power of this method of information handling, we may expect that the coming decade will see the rapid growth of its use within universities, where it may revolutionize the interaction of chemists with most kinds of chemical information.

The foregoing comments suggest that some of the factors which go into the design and operation of a computer-based chemical information center are (i) the size of the data base, (ii) the number of potential users, (iii) the importance of speed in delivering the results of the search, (iv) the cost of developing the system and searching it, and (v) the information needs of the users and the ability of the system to fulfill them. In terms of these factors, current-awareness and retrospective search systems obviously require enormous data bases, involving great developmental costs, and expensive search facilities. This implies that a large number of potential users is needed in order to recover the cost of implementing and operating the system.

From the boxes at the right of Fig. 2 we see that there are other types of information systems of considerable importance to the working scientist which do not require a large data base or many potential users. My own research group of physical-organic chemists, working with our colleagues in the Pittsburgh Chemical Information Center, has tried to attack some of these problems on a scale general enough so that the results would be of value to scholars working in almost any field within the university.

Reprint File

Most of us who have been active in research in a given area for a long time accumulate a large file of reprints and photocopies. From time to time we may try different methods of cross-indexing them for convenient retrieval. At one time or another most of us have tried developing systems of edge-sorted punch cards and have found, to our dismay, that when such a file gets much larger than 100 or 200 punched cards, we lose the will either to punch cards or to carry through the multiple sorting with needles which is needed to recover them when they must be searched.

At the University of Pittsburgh our approach toward dealing with this problem has been the development of a list of words which have been found to appear repeatedly in the papers in our field of interest. In Fig. 3 such a list is presented in the form of an indexing sheet. When a member of the group reads a paper, he simply checks off the items on the list that are covered in the paper, the structures and the compounds that are important, and the various solvents that were used, and the article is then given an identification number. A storage and retrieval system based on the Pitt Time-Shared System has been developed: an article is filed under all of the items checked on the list and is retrievable in terms of any combination

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr>
<tr><td colspan="9">Accession Number</td><td colspan="22">Principal Author (abstractor, circle)</td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>47</td><td>48</td><td>49</td><td>50</td><td>51</td><td>52</td><td>53</td></tr>
<tr><td colspan="6">Journal</td><td colspan="6">Volume</td><td colspan="6">Page</td><td colspan="4">Year</td></tr>
</table>

ABSTRACTOR ______________________________

I. Authors

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

II. Solvents

1. Acetic acid
2. Acetone
3. Acetonitrile
4. C_5-C_8
5. Benzene
6. n-Butanol
7. t-Butanol
8. Carbon Disulfide
9. Carbon Tetrachloride
10. Chloroform
11. Diethyl Ether
12. Dimet[illegible] For[illegible]de
13. D[illegible]

26. Tetrahydrofuran
27. Toluene
28. Water
29. Hexamethylphosphoramide
30. D_2O
31. H_2SO_4
32. HSO_3F
33. HX

III. Properties & Processes

1. Acidity
2. Acidity, Function
3. Acid, Strong
4. Acid, Weak
5. Activation, Energy
6. Activation, Entropy
7. Activation, Parameter
8. Activation, Volume
9. Activity Coefficient
10. Aprotic
11. Aromaticity
12. Basicity
13. Base, Strong
14. Base, Weak
15. Binary
16. Bioenergetics
17. Bond Parameter
18. Calorimetric
19. Catalysis
20. Charge Transfer
21. Computer Calculation
22. Conformation
23. Density
24. Deuteri[illegible]

38. Electrolyte (see V-salt)
39. Empirical Relationship
40. Enthalpy
41. Entropy
42. Equil. const.
43. Free Energy
44. Heat Capacity
45. Hydration
46. Hydrogen Bond
47. Hard-Soft
48. Hybridization
49. Hydrophobic
50. Hyperconjugation - (see Effect, Baker-Nathan)
51. Ion Pair
52. Ionic Conductance
53. Ionic Mobility
54. Ionic Strength
55. Ionizing Power
56. Kinetic
57. Leaving Group
58. Method, Instrumental
59. Method, Preparative
60. Method, Purification
61. Method, Separation
62. Neighboring Group
63. Nonelectrolyte
64. Optical Properties (e.g. [α], ORD)
65. pKa
66. Polar
67. Polarizability
68. Polarography
[illegible] Pressure

4. Ultraviolet-Visible
5. Ultrasonic
6. Stereochemical
7. Theory
8. Thermodynamic
9. Titration
10. Volume
11. Viscosity
12. Water Structure
13. Transition Energy
14. Frequency Shift

IV. Reaction Types

1. Addition
2. Association
3. Asymmetric Synthesis
4. Alkylation
5. Carbene
6. Cleavage
7. Complexing (not association)
8. Coupling
9. Cycloaddition
10. Electrophilic
11. Elimination
12. Exchange
13. Fragmentation
14. Hydrolysis
15. Intermolecular
16. Intramolecular
17. Nucleophilic
[illegible]tion

Fig. 3. Key word indexing sheet.

of them. A typical search could be initiated by requesting, for example, all of the papers which mention the terms *hydrogen-bonding*, *ionization*, *amides*, and *nuclear magnetic resonance* in combination. Access to the system may be through a console with a typewriter or through a cathode-ray display system. After a preliminary search, the user is first told the number of papers that will answer his question. If it is unacceptably large, he may qualify the question further. He may then request the accession numbers of the related papers and then easily locate them in the file. Microfilming is not expensive, and it would be quite feasible to have the papers stored on microfilm beside the terminal for rapid scanning after the file had been searched in the computer. We are at present developing a data base from our literature in physical-organic chemistry and are investigating its value as a retrieval system.

Since the computer places no value on the types of terms that go into it, this technique of storing and retrieving scholarly information could be applied equally well to any other area of research, be it sociology, Byzantine art, biology, or classics.

Storage and Retrieval of Internally Generated Data

When any research group has continuing interest in a given field, a point is reached after several years where considerable effort is required to determine just what systems have been studied by the group, and under what conditions. In planning future research or in preparing reports and papers on the group's activities, considerable searching through notebooks and reports several years old may be required to discover whether solvolysis of a certain chloride (or was it a bromide?) had been carried out in 70 percent aqueous ethanol (or was it 80 percent?) and at 25°C (or was it 40°?). We have accordingly developed a data storage system for the thermochemical data which the group has generated over the last 6 years. A worksheet is filled in by each experimenter after a run in which the calorimeter is used, for example; and the results are punched and stored, so that they may be searched by structure fragments, by solvent, by solute, and by each of the terms shown. Since there are relatively few ways in which such a file might be searched, we simply use a listing program to write an updated book from time to time in which the results obtained by the group are neatly indexed in the best of all retrievable forms—hard copy on paper.

Table 1 summarizes the status of the various types of files we have discussed in terms of size, number of users, cost, and search made.

Now let us consider the response we have had to one of these files—current-awareness service on *CA Condensates*, *Chemical Titles*, and *ASCA*.

Who Wants What Chemical Information Services?

Chemical information centers at present exist primarily to serve scientists, technicians, and engineers. The Pittsburgh Chemical Information Center is, to our knowledge, unique in having originated and been developed entirely through a university chemistry department. The basic experimental group of users was formed 2 or 3 years ago from members of the University of Pittsburgh chemistry department; the group has since expanded to include Mellon Institute; Carnegie-Mellon University; 21 carefully chosen industries in Pennsylvania, New Jersey, and Delaware; and research groups and individual chemists at a number of other universities, including Harvard, Columbia, and Princeton. Since the acceptance and use of chemical information is a question of human behavior rather than of chemistry, and since we are responsible for the collection and evaluation of feedback data on the use of our systems, a behavioral science research group has been involved from the very beginning in gathering and analyzing such data.

Many chemists, like other groups of people, have a strong bias against behavioral scientists. Like most prejudices, this is based either on a lack of understanding of the other man's job or on some unpleasant experiences with some of the less-competent members of the field. The prejudice of sociologists against physical scientists is generally based on the same foundations. Recalling the ancient adage "He who would his own lawyer be, hath a fool for a client," I was glad to find several competent sociologists who were eager to become deeply involved in studying the reactions of chemists to computer-based services and who could be entrusted far more reasonably than could any chemist with the construction of objective interviews, questionnaires, and feedback forms, and with statistical analysis of the results. Their interest in the project arises mostly from the realization that chemistry has the largest and best-organized data base in science and is so central in serving technology and other sciences that its information base may serve as a model for computerizing other disciplines in the future. Furthermore, the importance of chemistry which results from its central position implies that there is much incentive for its support, so it seems likely that chemical information services will be in demand for a considerable period, and will not be dropped in a few years, as has happened in some other fields.

As of 1 January 1970, our first experimental evaluation of *CA Condensates* came to an end. A part of the statistical analysis of the behavioral data is now complete, and some of the conclusions can be given.

The results described here are based on large samples (100 to 300) of our user group of chemists. They have been studied for almost a year in order to determine, if possible, whether the research style and behavior of the users was modified in any way as a result of the new tools which were provided them. Demographic analysis of the user group indicated that its composition was essentially the same, in terms of field of chemistry, as the national distribution established through the NSF manpower study. However, the group was almost evenly divided between academic and industrial chemists.

Three different instruments for gathering data were used, and checked against each other wherever possible. About 100 of our users responded to a structured interview which was developed from a series of pilot discussions with members of the University of Pittsburgh chemistry department. The interview was designed to gather information on the communication patterns, research style, interests, and information needs of the chemists. A second method which was used extensively but has many shortcomings was that of sending out questionnaires on a regular basis. Some questionnaires were sent several times to the same individuals at regular intervals in order to detect changes in the response of the users. Although questionnaires are a favorite tool for gathering behavioral data, their interpretation and administration are fraught with pitfalls, and their use is looked upon by most sociologists as a poor second to the interview technique. A third method of data gathering involved objective observation of the user's behavior vis-à-vis the computer in terms of the type of profile he constructed, how often he modified it, his search strategy, and his use of the information given him, as determined by means of a series of special feedback return cards on which his output was printed. The voluminous results of the study were analyzed by several statistical methods where appropriate, primarily by multivariate discriminant analysis and factor analysis.

The service was given free in return for the feedback information and was applied primarily to a study of the use of *CA Condensates* and *Chemical Titles*. (A separate study on the use of the *ASCA* citation service is currently being made.) The results of the study are, of course, rigorously applicable only to our user group, and no inferences can be properly drawn as to why our users displayed different types of behavior. Most of our conclusions are in no way surprising, but they provide strong statistical documentation for easily recognizable patterns. One should note, however, that many patterns of human behavior which would not have surprised us would not be supported by a careful study of this kind.

Of principal relevance to the development of a national information system is the fact that over 90 percent of our users concluded that the current-awareness services which they had used had been of definite value to them. By far the majority thought that the service had been of greatest value to them when used as a supplement to reading their favorite journals; in other words, they could rely on the information system to scan journals of peripheral interest, thereby gaining time for more intensive browsing of the journals which gave them the highest literature yield. The users were asked to estimate what percentages of various research budgets of different sizes they felt should be dedicated to purchasing services of this type (this is naturally quite a different question from that of how much they would *actually* pay). A logarithmic relationship was found between the size of the research budget and the suggested amounts to be allocated for chemical information services. One could estimate crudely that a research budget of $10,000 to $20,000 per year would generally support the standard $100 to $200 annual cost for an alerting service.

Table 1. Summary of the present status of computer-based systems for storing and retrieving chemical information.

Search or file type	Data base	Stage and cost	Desirable search time	Search mode
Literature, current awareness	600–1200 journals; *ASCA; CBAC*; Chemical Titles; CA Condensates*	Here, cheap	Weekly or biweekly alerts	Batch
Literature, retrospective	*Chemical Titles*, 8 years; *CA Condensates*, 18 months	Here, expensive	Fast; need to limit output in early stages of searches	Batch; interactive, at present too expensive
Literature, data (spectra) search	Dow Chemical Co.; Sadtler; ASTM	Infrared, some mass and nuclear-magnetic-resonance spectra; here	Fast	Interactive
Reprint file	Searcher's reprints	Here	Fast	Interactive
Group data storage	Searcher's data	Here	Fast	Hard copy

* Chemical-Biological Activity.

Scientists are well aware that there are wide divergences in research styles. This fact, which most of us arrive at intuitively, is completely supported by the very high degree of variance which was found in the responses of our group to practically any question having to do with their use of the research literature or communication of research results. Furthermore, the longitudinal aspect of the study indicated clearly that, over a period of 9 months to a year, there was little or no change in the response of our users to questions about their research style. One may conclude from this that designers of information systems must be very careful to avoid assuming monolithic patterns of behavior, and must lean over backward to avoid imposing particular ways of doing things on the chemists they wish to serve. The use of computer-based information unavoidably imposes some new structure on the research environment of scientists. This could naturally lead to resentment and frustration on the part of the recipients, and we were continually on the lookout for such reactions among our users. There was no indication, as the experiment progressed, of increasing frustration or apprehension on the part of the users or of any less enjoyment in doing research as a result of this new type of man-machine interaction.

An important behavioral question which has not received much attention in previous studies of the sociology of scientists is the use to which people put the information they get. We found that most of our users decided on the relevancy of a given citation when they received the "alert" from the computer. This is to say, in most instances they did not bother to go to the library to follow up all of the alerts but were generally satisfied to make the decision on the basis of the title and author as seen in the citation. The type of information considered most valuable by all members of our user group, including industrial and academic users, was that needed for the formulation or testing of scientific theories, concepts, or models, or that suggestive of new projects or programs. The type of information that our users felt was needed but could not be obtained easily was information about methods and procedures. No simple pattern of information needs emerged from our group of chemists. Some found the use of journals of particular value; others found their needs best met by conferences or face-to-face discussion.

Another observation which has considerable practical implication is that chemists procure information from the sources which are closest at hand. The matter of availability has an extremely important bearing on the degree to which chemists will follow through after being alerted to a certain type of information. One consequence of this is a high rate of dropout, from our system, of users in small companies with poor library facilities. Accessibility of the library is of great importance in institutions endowed with large libraries. It is also important to recognize that computer-based or other special services in large institutions should not be established at the expense of the conventional library system. Otherwise, people may find themselves in the peculiar situation of being alerted to articles in journals which should be available in their library but can no longer be afforded by the library because of the cost of the expensive new services.

A final matter related to library operations is the importance of fruitful personal interactions between an information specialist in the library or in an information center and the users who are constructing profiles. By far the best results are obtained from those institutions where a well-informed and imaginative librarian or information expert can work with a chemist in developing a profile and then help him follow through on improving it to the point where it really serves his intended purpose. Many users isolated in small departments or small companies where face-to-face contact with an information expert is impossible gradually get poorer and poorer results until they are alienated from the entire system. As is true with many other services, there is often a direct relationship between the initial effort of developing a good profile and the amount of value that is received from it later.

Notes

1. The University of Pittsburgh has an IBM 360/50 computer; the University of Georgia, an IBM 360/65.
2. I acknowledge support through NSF grant GN 738 and Pennsylvania Science and Engineering Foundation contract No. 60 for the development of the information systems described.

Education beyond the Horizon

Uses of communications satellites for education in developing nations are considered.

Lawrence P. Grayson

Communications satellites are not going to solve the world's problems. No technology can. Human understanding, an awareness of the problems that exist, and a desire to help are needed. However, satellites can be a mechanism for promoting understanding, for providing high-quality education to people in remote and underdeveloped parts of the world, and for assisting nations less developed than ours to advance. A satellite offers a country that lacks ground communications a system that can cover a geographic area of a million or more square kilometers. It can broadcast to the entire region or beam its presentations selectively to specific areas for particular users. It can reach isolated, mobile, and dispersed populations with an ease and total cost equal to that involved in reaching dense groupings of people. Unobstructed by mountains and rivers and "impassable" terrain, it offers easy access to regions that would be extremely difficult or very expensive to reach by ground systems. Truly, satellites are a man-made resource having a potential to reach and affect everyone.

The author is the acting deputy director of the Division of Educational Technology in the Bureau of Libraries and Educational Technology, U.S. Office of Education, Washington, D.C. 20202.

The Division of Nations

Social evolution has divided the nations of our world into two broad categories, which may be classified as the rich and the poor, the developed and the less developed, those that are technologically advanced and those that depend on intensive manual labor. One of the chief bases for classification in one or the other of these two groups is the average level of formal education of the nation's population. In a country with a high level of technological development, such as the United States, most of the population receives some form of post-secondary school education. On the other hand, in many nations of the world not more than 10 percent of the people are literate and perhaps less than 1 percent ever receive any secondary schooling.

The influence of Greek law, the adoption of a Judeo-Christian ethic, the founding of city-states and mercantile societies, the industrialization of the West, and the development of science and technology and their uses to improve the health of the people (*1*) have all been cited as factors that have led to this disparity between nations. Irrespective of the causes, however, the uneven distribution of wealth and the lack of opportunity in many nations are considered to be at the root of much of the national unrest and the international concern in the world today.

There is no simple way to reduce the gulf between nations and no rapid, easy way to bring the emerging countries on a par with countries that are more advanced. If this can ever be accomplished, it will be only through multifaceted developments. Agriculture must be modernized, industry must be expanded, outside capital must be brought in, and ways must be developed to effectively utilize resources existing in emerging countries. However, of all the needs, perhaps the most prevalent and urgent is the transformation of education. It has been suggested that from 50 to 60 percent of the gains in productivity in the West in the past half century can be attributed to education and the systematic application of knowledge. More specifically, it has been estimated that in the United States in the period 1929 to 1957, 21 percent of the growth of real national income per person employed is attributable to increased education of the labor force, while another 36 percent is attributable to the general advancement of knowledge (*2*). Yet, despite the importance of education for economic growth, developing countries have only the rudiments of a national educational system, and what does exist is often of doubtful quality.

How does one create an educational system to reach all the people in a country where teachers and educational materials are scarce, where the roads and mail systems are inadequate, where the transportation system is underdeveloped, and where school buildings and electricity are not widely available?

While there is no mix that will produce "instant education," we now have a potential capability to reach people, in a continuous, sustained, rapid manner,