with advanced projects, the information team insures against prolonged repetition of old work and nourishes genuine new leads with relevant information.

Fruitful information sources for such research projects include many that are frequently by-passed in ordinary literature searches. Typical are university theses, trade journals, pamphlets, technical bulletins, house organs, newspapers, and securities prospectuses. Sometimes the files of the Securities and Exchange Commission are the only public source of a particular bit of information. Special libraries can be very helpful. Publications of U. S. government agencies are valuable, as are conversations with some of their people. File histories of U. S. patents may provide well-considered technical evaluations. French and Belgian patents often provide an early hint of coming technical developments.

Such information research has been productive, and we should have more of it. Few of its results are dramatic, but one will illustrate a rare situation. Our company has long had an interest in the separation of para-xylene from its isomers by selective crystallization. Early in this development, a xylenes-ethylbenzene phase diagram was found showing 92 experimental points defining the phase relationships. These data were contained in a defunct Russian journal, available only in the New York public library. Obviously, experimental determination of these points would have been slow and costly.

CONCLUSION

The digestion of technical information has grown increasingly difficult and complex as a consequence of the startling growth of research in industry, government, and university laboratories. Needless duplication of past research is a luxury none can afford in this era of intense competition. Information services can play an important role by reporting new information, assembling pertinent information whether new or old, and evaluating specialized information. Consultation with key research people will help establish important needs; continuing liaison will assure effective development and maintenance of useful information services.

The information needs of people directly concerned with laboratory work can largely be met with services that provide a scanning of new scientific and trade literature, make journals available promptly, provide translations where necessary, inform concerning other internal activities, and execute searches, whether to provide only a bibliography or a broad, interpretive survey of a field of research. If too much service is supplied or if a particular service receives too much emphasis, there is danger that the researcher will lean too heavily on information services as a crutch. There is no substitute for the researcher's personal use of the library where he can interpret what he reads in the light of his own experiences. An important by-product is accidental exposure to other journal articles that may even prove so informative and provocative that the original quest is completely ignored or momentarily forgotten.

There exists a need for more information specialists of the highest order—creative scientists in their own right—who can cope with the special information problems that frequently beset a research staff. Machines can be devised to store and retrieve information; they cannot exercise judgment or provide technical insight and interpretation.

Information services are justified whenever a valid need can be served economically. Needs change and economic bases often require modification. Each information service should be evaluated periodically to assure its continued ability to catalyze effective research.

The petroleum researcher can only extend sympathy as he views the many problems besetting the information expert today. He is anxious for success in this area because he values highly the information services that have contributed significantly to his own advances in research.

Technical Information Needs of the Chemical Processing Industry*

R. P. GENEREAUX
Engineering Department, E.I. du Pont de Nemours Company, Wilmington, Delaware
Received September 22, 1961

During the time we are spending this morning, the world will add about one thousand scientific and technical articles to its storehouse of published knowledge. You and I know that the organizations for which we work—and mankind in general—will use those thousand articles inefficiently. My objective today is to gain support for some practical steps that should be taken promptly to improve the use of technical information yet to be published.

To be meaningful and specific, I shall talk about the technical information needs of the chemical processing industry. There are two reasons for doing this. First, my entire career has been associated with that industry. And second, you as ACS members are primarily interested in the chemical industry.

We will start by reviewing broadly the technical information needs of people in various functions in the chemical industry. Then we will focus on technical information problems of engineers and scientists in the industry. I will outline certain improvements in technical information services which I believe should be widely and immediately applied. Then, in conclusion, we will take a look at longer-range objectives and actions that should be undertaken.

Figure 1 is a simplified representation of the technical information needs of the chemical processing industry. It relates the source of information on the left to the function

*Presented before the Division of Chemical Literature, ACS National Meeting, Chicago, Ill., Sept. 6, 1961.

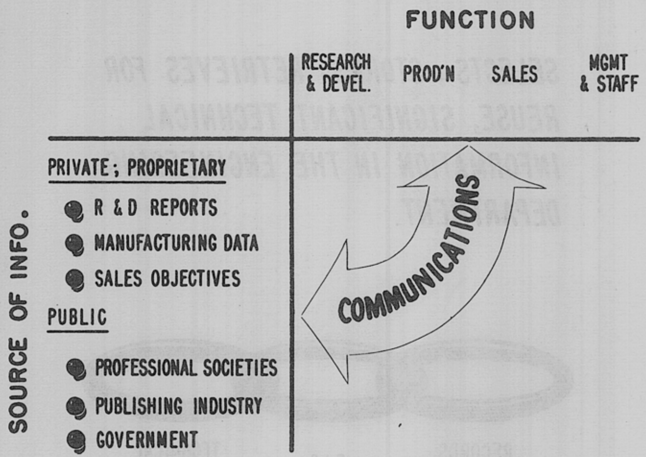

Fig. 1

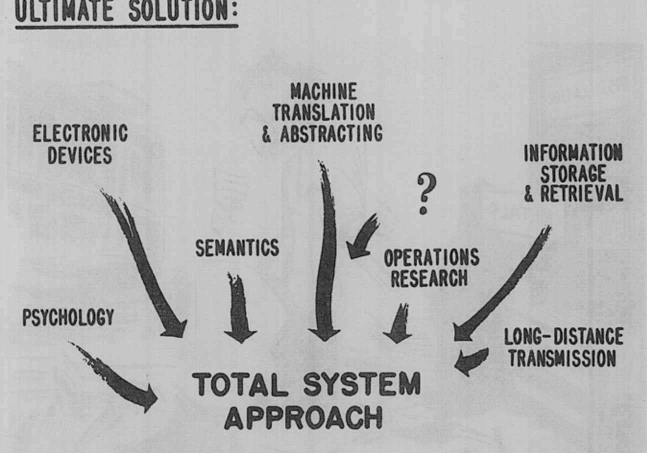

Fig. 2

on the right. It portrays the fact that variously trained people, in the many functions in the industry, should use various types of technical information from numerous sources. Looking at it another way, it is a supply and demand problem: supplying the right kinds of information from the right sources, to meet the demands and stimulate the minds of managerial people as well as professional.

Most of us here are concerned primarily with the information needs of engineers and chemists engaged in research and development work and allied activities. This is only a fraction of the total needs of the chemical industry for technical information. People engaged in production, sales, staff functions, and top management, require appropriate technical information to perform their jobs effectively. I think we do a disservice if we view the problem as confined to supplying technical information to people engaged in purely technical activities. Persons in other responsible positions throughout the industry deserve full consideration of their information needs.

When viewed in this light, the technical information needs of the chemical industry are a problem of communications in the fullest sense of the word. The problem is beset with many complex relationships which must be considered: the various people and sources of information, expected improvements in information processing and devices, timing, and economics.

The problem will be solved best when we are able to apply a total system approach to its solution, Fig. 2. It will require integration of knowledge and latest developments from such diverse fields as psychology, semantics, machine translation and abstracting, logic of information storage and retrieval, long distance transmission, electronic computers and other devices, and Operations Research. One other thing is required: continued production of good technical information to advance these fields and others, and to make the entire effort worth while.

However, I promised to be specific. Therefore, let us examine more closely the technical information problems of engineers engaged in research and development work. Engineers are living in a vastly different information environment than did those a generation ago. I shall use my personal experiences to illustrate the information morass we have allowed our engineers to get into.

When I started in chemical engineering research in 1929, Fig. 3, my principal assignment was in the field of fluid flow. The objective was to determine basic relationships among fluid flow phenomena pertinent to du Pont's manufacturing operations.


Fig. 3. Yesterday's Engineer

At that time, the current literature on fluid flow consisted of a relatively small number of journals, mostly American, some German, French, and English. It was not difficult for a person to review the pertinent literature, maintain awareness of latest information in the field, and still have time left to conduct a laboratory research program.

There was a manageable number of books available, some classical, some textbooks, and a few handbooks. As I recall, the most valuable classics were Lamb's "Hydrodynamics," Stodola's "Steam and Gas Turbines," and Gibson's "Hydraulics and Its Applications," a few textbooks such as Walker, Lewis, and McAdams' "Principles of Chemical Engineering," and compendiums such as the report of the National Research Council Committee by Dryden, Murnaghan, and Bateman.

My principal technique in accumulating information useful to my research program consisted of following up conscientiously on references to past information noted in current publications. Each article of interest was abstracted on cards and arranged in subfields of technological classification. Articles not available in our own library were obtained as photocopies. People who knew that I was interested in fluid flow helped by supplying references to articles they thought I might not see. Thus was built up an information collection in fluid flow which proved

Fig. 4. Today's Engineer

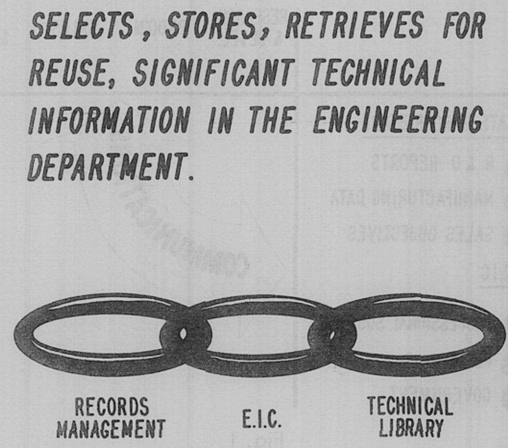

Fig. 5. E.I.C. Objective

completely adequate for my research needs in this field during the thirties. This storehouse of information was also of considerable value when it came time to participate in preparing the Fluid Flow section of Perry's Chemical Engineers' Handbook, the first edition of which was published in 1934.

My best guess is that I spent about one-third of my time on literature search. I am sure that better methods of information storage and retrieval have made my engineering research in this field more fruitful.

Today's engineer in the chemical industry is almost literally the hapless fellow shown in Fig. 4. You are acquainted with data on how the storehouse of technical information is increasing by leaps and bounds—and how difficult it is to locate information pertinent to a specific problem.

Instead of reciting more such generalized data, I will summarize how the du Pont Engineering Department took action on the information problem of our associate in the picture.

In 1956, we initiated a study to evaluate possible applications of then-arising management science techniques to improving technical, manpower, and cost effectiveness in the Engineering Department. We were aware that the effectiveness of today's engineers is dependent on many factors which management can influence. We were interested in determining precisely how the accelerating application of science to the management field could improve our department's effectiveness. Distinguishing features of the new management science included: applications of mathematical techniques, electronic computer technology, integrated data processing, performance measurement and evaluation methods, and practical use of information theory, namely, "documentation" as it has come to be known.

The reason we vigorously supported this study, and have pursued its findings, was an economic one. The engineering costs of a construction project had risen gradually and we were convinced that a sharp look at our practices would reveal opportunities for our engineers to eliminate unnecessary costs.

In the years since initiation of the study, the du Pont Engineering Department has applied to its business each of the techniques mentioned in modern management science. I would enjoy telling you the entire story, but shall restrict it to a review of what we did on documentation.

In 1956 and 1957, a portion of the study group concentrated on distilling some order, as they saw it, from the seeming chaos in the field of information storage and retrieval. To the distillate they added ingredients of their own creation. The resulting procedures for information retrieval were tested on several kinds of technical documents in the Engineering Department. It was demonstrated clearly that these new procedures were significantly more effective in retrieving information than the methods traditionally used. At that point, in late 1957, the du Pont Engineering Information Center was created to apply the procedures in a production way, while the study group continued development work in the field of documentation.

The Engineering Information Center—we call it EIC—has a basic charter to select, store, and retrieve for reuse, significant technical information in the Engineering Department. The charter, Fig. 5, is designed to enable EIC to join with other traditional agencies in offering a complete spectrum of information services to the technical force. The Records Management function, for example, is responsible for customary filing of all Company documents. The Technical Library of the Company has a branch for the Engineering Department and offers principally non-du Pont information. Thus, today's engineer in our department is provided three major agencies to serve his information needs: Records Management for traditionally filed Company information; EIC for selected Engineering Department information processed via modern documentation; and the Technical Library for non-du Pont information. This tripartite division has a few practical exceptions, of course.

The engineering department has accumulated over 4,000,000 technical documents on its past efforts on engineering research, applications of new technical knowledge, solutions to operating problems, and the design

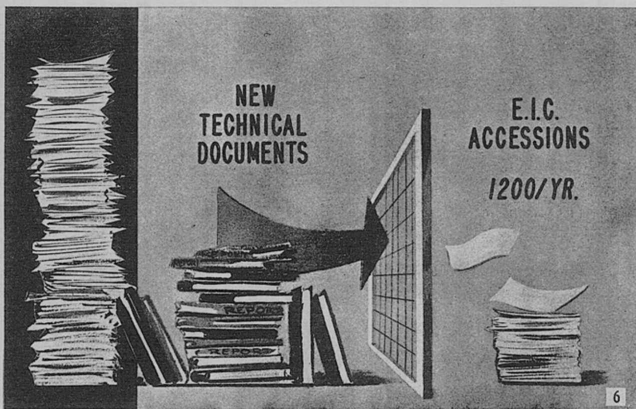

Fig. 6. E.I.C. System

and construction of physical facilities. Each year some 3,000 studies and 1,500 construction projects contribute further to this storehouse of recorded information, the creation of about 2,000 technical people.

The system of the Engineering Information Center, Fig. 6, selects from the annual avalanche of new technical documents a small percentage, say about 1%, which is judged to have maximum long-term reusability. This selection process results in the identification of around 1,200 EIC accessions a year. An accession number may refer to a research report, along with significant file papers; to the drawings, calculations, purchase orders, *etc.*, of a portion of a multi-million dollar construction project; or to any collection of documents which interrelate to record a specific technical accomplishment.

How we summarize and index the information covered by an accession number has been told elsewhere. The important point I wish to make is that a *decision* was made to create information packages, or accessions, to record all outstanding technical efforts, and to process the packages for maximum future use *via* the power of modern documentation techniques.

Once this decision was made it determined the basic level of personnel and budget for EIC. Approximately twenty people, representing an annual budget over $200,000, are required to process the selected information from the technical force of 2,000.

These 2,000 in the technical force are engaged in engineering work ranging from fundamental, basic, or pioneering research, through application research and development, the solution of plant operating and maintenance problems, and the pursuit of improvement programs, to the complete design and construction of all plant facilities.

Now in its fourth year of operation, EIC has an inventory of 10,000 accessions including several thousand technical reports processed during the early development period. We have statistical evidence, and occasional clear cases, which demonstrate that EIC inquiry services save engineers' time in aggregate at least equal to the EIC annual budget. It is handling 1,200 inquiries a year, which is less than the rate should be. A continuing educational program is required to convince all of today's engineers that modern documentation services can provide useful information they are not likely to find themselves or even attempt to look for.

To complete the du Pont picture, there are now several other department information centers in the Company

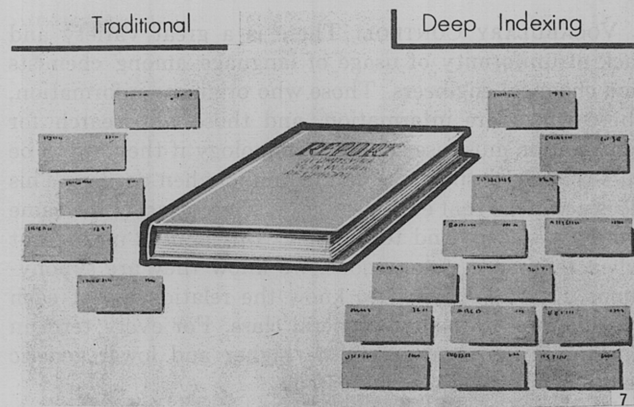

Fig. 7. Deep Indexing

which practice modern documentation. One center handles research information from the Textile Fibers and the Film Departments, and is directed by Dr. Carl Conrad, who is Chairman of the Chemical Documentation Committee of the ACS Division of Chemical Literature.

Six years of documentation experience have resulted in our having firm views on the present and future technical information needs of the du Pont Company, and even of the chemical processing industry and of the nation. I shall summarize them in the remainder on my remarks.

The present information needs include: Service, Deep Indexing, Vocabulary Control, Concept Co-ordination, Source Indexing, and Machine Considerations. Let us look at these in some detail.

SERVICE. Information must be provided as a service to scientists and engineers, in contrast to an often occurring philosophy of information as a record storage and maintenance activity.

DEEP INDEXING. We need to provide the means for a person with a specific problem to be able to retrieve the particular information needed in a practical length of time without wading through mountains of literature.

In 1936, the single indexing term "*nylon*" might have been sufficient to describe reports on that material. There weren't many references. However, the use of a single indexing term to cover information on "nylon" today would result in the inquirer being deluged with probably more material than he would be able to read in a lifetime. We need to be more specific. We need to index the subject more deeply, Fig. 7.

- **DISTILLATION**
  (SYNONYMS)
  FRACTIONATION
  RECTIFICATION (CHEMICAL)
- **METALS**
  (INCLUDES)
  COPPER
  IRON
  LEAD
  STAINLESS STEELS
  TIN
- **FURNACES**
  (RELATED TERMS)
  BURNERS
  CALCINERS
  HEATERS
  KILNS
  OVENS
  ROASTERS

Fig. 8

VOCABULARY CONTROL. There is a great variety and lack of uniformity of usage of language among chemists and chemical engineers. Those who originate information, those who store information, and those who search for information must use similar terminology if they are to be sure that the desired idea can be found when needed. This does not mean that each person must use exactly the same words in science and technology. But it does mean that when terms are synonymous, we know they are synonymous, Fig. 8. We need to know the relationship of each specific term to its more general class. For every term in an information system all the higher and lower generic relationships should be indicated.

In addition, it is helpful to indicate other words, words that are related to the primary concept but not necessarily always synonymous nor which always have a generic relationship with the main term. These would be related terms or reminder words to consider. Vocabulary control of this type is a necessity in an effective information system. A technical thesaurus is the most efficient device to accomplish vocabulary control. The American Institute of Chemical Engineers recently published a "Chemical Engineering Thesaurus" of engineering terms.

CONCEPT CO-ORDINATION. Most of the newer systems installed today in industry and in the U. S. Government are based upon the simple principle of concept co-ordination. In this approach, as most of you know, a document is indexed by single words representing the concepts expressed in the document. These may be uniterms, descriptors, keywords, or other single word representations, depending upon the individual's needs. In the example shown in Fig. 9, the report discusses the oxidation of nitric oxide to nitrogen dioxide using activated carbon as a catalyst. References to information are stored by indicating for each concept the document numbers that have information in them on these concepts. References to information are retrieved by co-ordinating the terms the searcher desires to find information on, then noting matching document numbers on those terms.

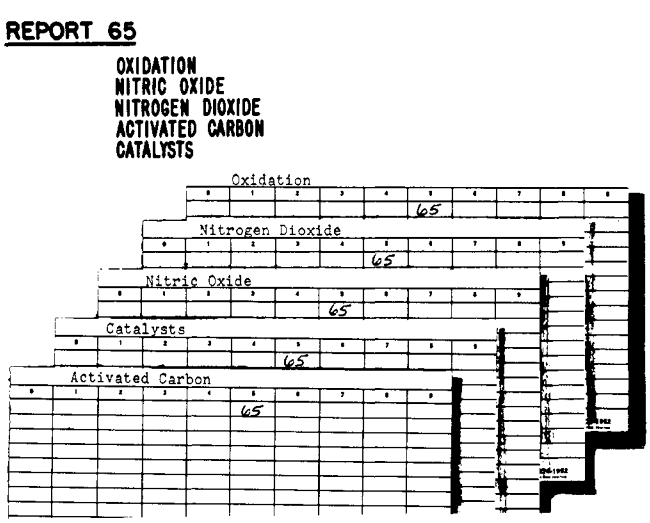

Fig. 9. Concept Coordination

SOURCE INDEXING. The most effective point at which to create references to information is at the time of its generation, Fig. 10. The author of a technical article and the editor already have digested the technical content of

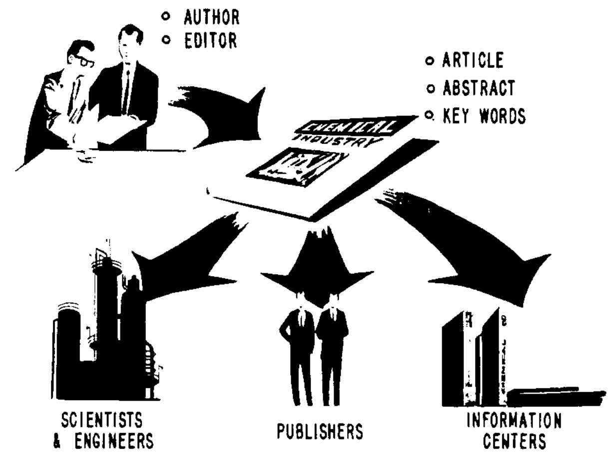

Fig. 10. Source Indexing

the material and are best able to select the major concepts which should be referenced for storage for future use. Under ideal conditions the author would supply key access words and an abstract for his articles. Publishers would co-operate by publishing these references and by assuring compatibility between authors and other publications.

It would not be practical to go back through the world of literature to improve its retrievability. But we can do a better job on present and future information. The time to begin is now.

The American Institute of Chemical Engineers now publishes with every article in *Chemical Engineering Progress, A.I.Ch.E. Journal,* and the *Symposium Series* an abstract along with a list of indexing terms. Thus, the A.I.ChE., as one generating source of information in the chemical engineering field, provides improved access for the information it publishes. The *Petroleum Refiner* is cooperating in this program by the publication of indexing terms with their articles.

MACHINE CONSIDERATIONS. Machines are only a tool in the solution of the information problem. The solution itself is intellectual. Our first problem and the major one in the field of information retrieval is to define the system requirements. The needs of the users must be defined and the policies of operation established for each system. Equipment should only be used where it is justified on the bases of faster service, better service, and less expensive service. In the evolving developments in digital computers, equipment for storage and retrieval is developing faster than our ability to use it effectively. So much for the short-range. Now let us look at some long-range opportunities, as listed in Fig. 11.

INFORMATION NEEDS OF SCIENTIFIC PERSONNEL. There has never been, to my knowledge, an adequate study of the information *needs* of scientific personnel. There have been studies on how much time scientists spend in getting information and studies on where they obtain this information. But no one has established what information a scientist needs in the conduct of his program. I submit that such a study would be invaluable in outlining requirements for information storage.

VALUE OF INFORMATION. Along with this, we need to determine the value of information. How much can we afford to spend to provide information services to our personnel and what is this information worth to them?

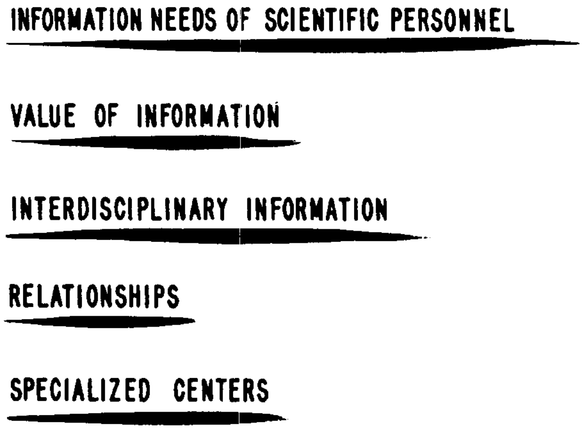

Fig. 11. Future Information Needs

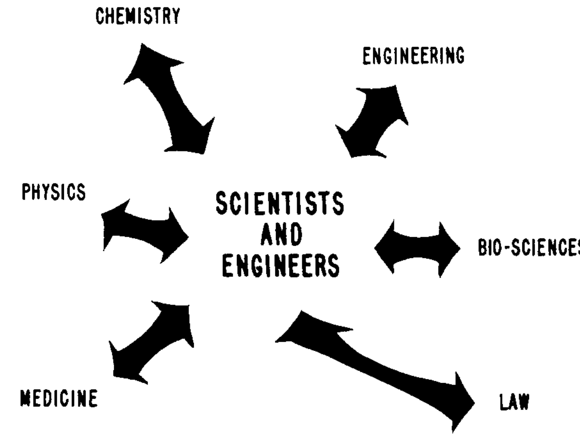

Fig. 12

If we knew the value of information, we could justify and program the right amount and kinds of development work on the problem.

INTERDISCIPLINARY INFORMATION. It is no longer practical for every scientist to try to survey all the current literature that might be of interest to him. One of the major reasons for this is that it is the boundary regions of science which offer the richest opportunities to the qualified investigator. Disciplines in science overlap and interreact. We need a retrieval mechanism so that individuals can obtain information from all important sources. In addition, a current awareness philosophy is needed which will provide a scientist, automatically, with selected articles in his areas of interest from those sources which he does not normally review. Each scientist will normally review a select number of journals of particular interest to him. He should not have to review all the peripheral publications which only rarely will contain an article of interest to him.

RELATIONSHIPS. We need to establish optimum relationships among authors, publishers, and information centers. As discussed perviously, we need to encourage the generators of information to provide with that information the access to it. Publishers should co-operate by publishing these indexes. Guidance, stimulation, and co-ordination should come from the various professional societies to encourage this development within their disciplines.

SPECIALIZED CENTERS. We will see in the future the establishment of specialized centers, Fig. 12, serving the technology within a discipline, and a communications network between these centers. Each center would concentrate on information of interest to that discipline and when the need arose would communicate with centers maintaining other types of desired information. Thus, a person with a need for information would go to the appropriate center which could be expected to supply that information. The center would search its references and supply, if desired, copies of the required information.

It is not practical or possible for every company or individual to maintain information systems containing all the references in the world's literature of interest to that system. Rather, each individual or company should maintain information systems of only the most pertinent and most reusable information with appropriate access to outside centers which will maintain wide coverage in depth. In my opinion, industry would prefer to pay a specialized information center to search through their voluminous references rather than the company try to maintain all these references itself.

CONCLUSIONS

I am convinced that information is a valuable commodity. I am concerned because the retrievability of this information does not match today's needs. Professional societies should continue to stimulate their members in adopting improved information processing techniques. Professional societies should continue to interchange developments which promise to improve information service. It will take courage to depart from familiar ways. To be effective, information service must meet the needs of the user of information for he supplies the economic justification for it.