REFERENCES

(1) L. J. Stevens, Advances in Chemistry Series No. 10, American Chemical Society, Washington, D. C., 1954, pp. 494–499.
(2) L. J. Stevens, *J. Chem. Educ.*, **31**, 471 (1954).
(3) Electrochemical Society of Japan, "Denki Kagaku Benran," Maruzen Ltd., Tokyo, 1954, p. 5.
(4) M. Anzai, Ed. "Japan's Electrochemical Industry," Tokyo, 1958, p. 1 (in English).
(5) T. Aono, Ed. "Nippon no Denki Kagaku Kogyo no Hatten," Electrochemical Society of Japan, Tokyo, 1959.
(6) Group Discussion, *Denki Kagaku*, **21** (7), 327 (1953).
(7) S. Oka, *ibid.*, **22**, 50 (1954).
(8) E. Takagi, *ibid.*, **21** (7), 305 (1953).
(9) S. Godo, *ibid.*, **18** (7), 207 (1950).
(10) T. Kambara, *et al.*, *ibid.*, **20** (4), 602 (1952).
(11) M. Suzuki, *et al.*, *ibid.*, **21**, 164, 212, 264 (1953).
(12) Electrochemical Society of Japan, *ibid.*, **27** (2) (1959) (a special issue of 12 papers).
(13) T. Kambara, *Bull. Chem. Soc. Japan*, **27** (8), 523 (1954).
(14) Y. Yasumori, *ibid.*, **27** (8), 554 (1954).
(15) H. Shirai, *Nippon Kagaku Zasshi*, **81**, 1248, 1578 (1960).
(16) H. Shirai, *ibid.*, **82**, 339 (1961).
(17) A. Saito, *ibid.*, **82**, 212, 214, 715, 718 (1961).
(18) M. Suzuki, *Denki Kagaku*, **22**, 23, 112, 162, 220 (1954).
(19) T. Takahashi, *Kogyo Kagaku Zasshi*, **64**, 515 (1961).
(20) A. Tateisa, *Bunseki Kagaku*, **11** (1), 32 (1962).
(21) M. Saito, *Nippon Kagaku Zasshi*, **83**, 883 (1962).
(22) H. Hayagawa, *Denki Kagaku*, **19** (12), 374 (1951).
(23) S. Nagai, *ibid.*, **20** (4), 150 (1952).
(24) W. Sakai, *ibid.*, **21** 376, 442, 516 (1953).
(25) K. Anazawa, *ibid.*, **19** (11), 338 (1951).
(26) C. Okamoto, *et al.*, *ibid.*, **22**, 8, 56, 213 (1954).
(27) T. Mukaibo, *et al.*, *ibid.*, **22**, 11, 59, 516, (1954).
(28) T. Ishiyam, *ibid.*, **22**, 150 (1954).
(29) G. Okamoto, *ibid.*, **24**, 69 (1956); **25**, 166, 175 (1957).
(30) Corrosion Committee, *ibid.*, **25**, 218 (1957) (in English, 7 papers).
(31) Electrochemical Society of Japan, *ibid.*, **26** (12) (1958) (11 papers).
(32) G. Okamoto, *Chem. Chem. Ind.* (Japan), **14**, 648 (1961).
(33) T. Aono, *Denki Kagaku*, **11** (7), 16 (1943).
(34) Y. Nomura, *ibid.*, **21**, 389 (1953).
(35) T. Ishishiro, *ibid.*, **22**, 78 (1954).
(36) Electrochemical Society of Japan, "Special Issue for Carbide," *ibid.*, **26** (10) (1958).
(37) Electrochemical Society of Japan, "Special Issue for Primary Battery," *ibid.*, **22** (11) (1954) (14 papers).
(38) Electrochemical Society of Japan, "Special Issue for Batteries," *ibid.*, **26** (2) (1958).
(39) Chemical Abstracts Service, "Chemical Abstracts List of Periodicals," Columbus, Ohio, 1961, p. 20J.
(40) H. Kataoka, "Tokkyo Ho Kaisetsu" (Interpretation of Patent Laws), Nankodo, Tokyo, 1961.
(41) Electrochemical Society of Japan, "Denki Kagaku Benran" (Handbook of Electrochemisty), Maruzen Ltd., Tokyo, 1954.
(42) Chemical Society of Japan, "Kagaku Benran" (Handbook of Chemistry), Maruzen Ltd., Tokyo, 1952.
(43) Japan Institute of Metals, "Kinzoku Benran" (Handbook of Metals), Maruzen Ltd., Tokyo, 1952.
(44) H. Inouye and W. Nakajima, "Denki Kagaku Kogyo Taikei" (Outlines of Electrochemical Industry), Tokyodo, Tokyo, 1947.
(45) S. Oka and Y. Nagai, "Kagaku Kogyo Jiten" (Dictionary of Chemical Industry), Tokyodo, Tokyo, 1959.
(46) T. Inouye, *et al.*, "Rikagaku Jiten" (Dictionary of Physics and Chemistry), Enlarged Ed., Iwanami Shoden, Tokyo, 1958.
(47) L. C. Ku, *Hua Hsueh Shih Chieh*, **11**, 382 (1956).
(48) Anon., *ibid.*, **14**, 454 (1959).
(49) Y. C., T'ong, *Hua Hsueh Tung Pao*, (5) 41 (1959).
(50) J. Y. Yen, *ibid.*, (11) 495 (1959).
(51) T. C. Yuen, *Hua Hsueh Hsueh Pao*, **25** (5), 1 (1959).
(52) Anon., *Hua Hsueh Kung Yeh* (Chemical Industry), 11 (1959).
(53) Anon., *Hua Hsueh Shih Chieh* (Chemical World), **14**, 85 (1959).
(54) R. S. Du, *Chung-kuo K'o-hsueh* (Scientia Sinica), 8 (11), 1196 (1959).
(55) L. C. Yu, and C. C. Shih, *J. Peking Pedagogical Inst.*, (1) 29 (1959).
(56) S. H. Gould, Ed. "Sciences in Communist China," AAAS, Washington, D. C., 1961.
(57) K. Hibi, *Denki Kagaku*, **21**, 396 (1953).
(58) Anon., *Hua Hsueh Shih Chieh*, (2) 9 (1948).
(59) Ta-Hsung T'ung, private communication, July 10, 1963.

An Analysis of the Designing, Installation, and Operation of a Coordinate Indexing System Using Links and Roles for the Plastics Department of the Du Pont Company*

By BARBARA A. MONTAGUE

Research and Development Division, Plastics Department, E. I. du Pont de Nemours and Company, Inc., Wilmington, Delaware

Received December 9, 1963

This paper discusses the decisions involved in the planning and operation of a technical information system incorporating internal research reports, patents, and sales reports dealing with polymers and polymer intermediates. References to pertinent documents are located by a manual inverted coordinate index or by an IBM 650 computer.

The Need for an Information System.—Several years ago the development of one of our new plastics was nearing completion and plant start-up was in process. The research technology which led to the development of this plastic had been documented in some 200 internal research reports

* Presented before the Division of Chemical Literature, 145th National Meeting of the American Chemical Society, New York, N. Y., Sept. 10, 1963.

by members of the task force assigned to the project. The success of plant operation would depend heavily on the prompt application of the experience recorded in these documents and in the memory of the men who did the work. Classification techniques had been used for years for the storage and retrieval of research information contained in the internal reports of the Research Division. This system was becoming more and more unsatisfactory with the build up of documents in the collection because sequential searching for lightly indexed documents resulted in time-consuming, incomplete retrieval. The problem at hand demanded faster, more complete, selective retrieval.

The decision was made to try a new approach, namely, concept coordination. There were two reasons for this selection: first, the Engineering Department of the Du Pont Company had successfully improved its information retrieval by using this new approach. Second, the management of the Plastics Department was progressive and receptive to new concepts to solve this aspect of its plant start-up problem. It was this need on the one hand and the availability of a solution on the other which led to the introduction of a new concept in information handling in the Plastics Department.

System Design Considerations.—The system comprised deep indexing coupled with links and roles to offset the potential of false retrieval resulting from deep indexing. Knowledge from experience of other systems indicated that vocabulary control was needed to eliminate synonyms, solve semantic problems inherent in the English language, and handle class relationships (generics).

It was felt that a manual search device was satisfactory for such a small system. Reports were retrieved using a dual dictionary by comparing accession numbers of documents with links under term-roles representing information desired. A companion document called a Title List consisted of a list of documents in the system arranged in accession number order along with the title, authors, and the informative abstract written by the author. The dual dictionary and title list provided each chemist or chemical engineer at plant start-up with a "Do It Yourself" information kit which gave him the means to immediate location of documents containing answers to his questions. Documents were procured by a telephone call or trip to the file room.

Plant start-up began and the value of immediate access to past work soon became obvious. Many of the men who were members of the product development task force in research followed start-up by assignment to the plant. Problems were efficiently and quickly solved through personal contact of experience personnel. As these experienced men were replaced by new people, the technical know-how was no longer accessible through interpersonal communication. There was more and more reliance placed upon research reports as information sources. The fast access to this technology through use of the coordinate index more than paid for itself in time and effort saved.

Extension of System to All Research Division Reports.—Attention turned next to the internal reports which recorded the results of research activities of the entire Research Division. Management decided to finance the installation of an information system based upon an extension of the original Coordinate Index.[1] The collection consisted of 2100 reports covering the Department's research from 1950 to 1960.

The 2100 research reports were indexed by people selected from the user group, *i.e.*, 125 chemists, engineers, and supervisors working in laboratories, semiworks, and offices located at the main research center in Wilmington, as well as at the three plants. These people were selected to do the indexing for two reasons: first, to complete a sizeable piece of work, and second, to familiarize a large percentage of the users with the new concept in information handling and, in so doing, encourage them to use the system. It was thought that the users would develop a thorough understanding with confidence and hence would be more motivated to utilize a system in whose installation they had actively participated.

The indexers were trained in groups of 30 to 40 by systems engineers in two eight-hour sessions which consisted of thorough explanation of the principles of coordinate indexing along with practice drill in indexing techniques. Ground rules for indexing were distributed and reviewed, their purpose being to maintain consistency among indexers. Indexing consistency was one of the most difficult and critical aspects of the program owing to the naturally wide variance in viewpoint among people—for herein lies the secret of success in any information storage approach. Unless all documents are consistently indexed according to a common set of rules which in turn are rigorously followed when searching for the documents, relevant references will be missed. Ground rules guided the indexing initially and were followed by careful review of all 2100 documents by the two systems engineers to assure as consistent input as feasible. Variance in viewpoint is especially critical in classification or alphabetical indexing where documents are indexed lightly with few, if any, cross references. One of the advantages of coordinate indexing with links and roles is the ability to index deeply and thereby provide many access points to retrieve information. The probability of missing references indexed by different points of view is thereby reduced.

The use of 125 indexers made it possible to process a large backlog of reports in a short time; however, a selected group of experienced indexers can process documents more proficiently because they are working with the rules of the system day by day. It is for this reason that having authors index their reports at time of writing is not practiced.

The reviewed indexing forms called tracing sheets were keypunched by term description, role, accession number, and link letter. Each term, in addition to being assigned a link and role, was given a category number. The purpose of these categories was to sort the punched cards into groups of chemical and nonchemical terms to facilitate editing. The chemical and nonchemical terms were listed separately and each list sorted alphabetically. The chemical terms were edited to conform with *Chemical Abstracts* rules of nomenclature explained in "The Naming and Indexing of Chemical Compounds by *Chemical Abstracts*"[2] because the system was designed primarily for a user group of chemists who were familiar with *Chemical Abstracts*. This editing was performed by a Ph.D. organic chemist and one of the systems engineers to provide chemical accuracy coupled with systems consistency. Some chemical terms were edited to conform to departmental usage

and conventions instead of C.A. because searching was to be done initially by the users. In retrospect, this was not a wise decision because the searching is now performed entirely by a staff of technical information specialists who have had to learn a set of systematic rules, that is, C.A., plus a set of exceptions, namely, Plastics Department conventions. Operation would be greatly simplified if the system followed *Chemical Abstracts* exclusively.

Generic classes were generated by selecting broad classes of chemicals, such as in the following: acids (carboxylic), alcohols, aromatic hydrocarbons, fluoro organic compounds, olefins, and sodium compounds. Each specific chemical was posted to one or more of 1200 generic classes according to the functional groups and other characteristics of the compound. This generic posting was in essence a fragmentation of each chemical compound. Searching one generic class such as alcohols results in retrieval of all compounds containing the hydroxyl moiety. Coordination of two or more generic classes such as carboxylic acids with alcohols retrieves hydroxy acids by means of fragmentation.

Nonchemical terms such as the names of properties, techniques, equipment, and processes were edited by a chemist and a systems engineer. Used as guides and authorities were the Engineering Department information system's thesaurus [3] and members of the Plastics Research Division who were specialists in the fields of concepts being edited. Nonchemical terms were also posted to higher generic classes where appropriate. Vocabulary control is one of the most time consuming building blocks in installation of an information system and at the same time one of the most important. It presents the third opportunity to maintain input consistency; the other two are indexing ground rules and review indexing. The investment in time required by editing is well justified and consistent with the basic philosophy of coordinate indexing, namely, a careful, detailed, one-time analysis of information which permits fast, inexpensive, selective access on demand.

Processing of the edited input to convert it to an inverted coordinate index was accomplished by means of an IBM 650 computer. This particular equipment was selected because it was available in the Research Division. The input tape was processed to prepare a print-out of the coordinate index by means of an IBM 705 computer for the editing of alphabetic information. The computer listing was photographically reproduced and bound in dual dictionary form for manual searching. The inverted index was prepared by computer because of its accuracy and speed plus the elimination of human errors due to clerical transposition of numbers. Typing of the index alone would have required 2000 hours compared with the 2.5 hours by 705 computer. Another reason for using a computer was the product of an inverted index on magnetic tape which offered the potential for future mechanized searching.

Title lists were again prepared as companion documents for the dual dictionaries. The title list consisted of three sections, the first, a list of the reports in accession number order containing title, authors, and indicative abstracts prepared by a term-on-item sort of the term-roles generated from indexing the documents. The second section was an alphabetical inverted index of authors with accession numbers of reports posted under the names. The third section was a title and author list of the reports ordered by departmental issue number with corresponding accession number for conversion purposes.

The progress of system installation was announced by statements in monthly reports published within the Research Division. Final announcement of start-up of the Information System was made by means of a series of publicity presentations throughout the Wilmington area, as well as at the three plant locations. At this time copies of the dual dictionary and title list were distributed to technical personnel accompanied by an explanation of the system and drill in using these search tools. Dual dictionaries and title lists were distributed for two reasons: first, to provide each technical person with an in-hand mechanism for obtaining quick access to research information; and second, to provide him with an educational tool which promoted his understanding and confidence in the system and thus encouraged him to use it.

System Operation.—The first problem encountered in putting the new system into operation was selection of a proper location. The decision was made to combine the new activity with the existing record center to facilitate communication between the people who process information inquiries and the people responsible for storage and loan of documents. The combined operation was called the Plastics Information Center. The old system of storing documents and correspondence in file cabinets was replaced by open-shelf filing because of ease of access and saving in space.

The next problem to be solved was selection of a staff to operate the new information system. It was decided to use four B.A./B.S. chemists because they would be indexing technical information as well as processing searches for chemists who needed assistance in the use of the dual dictionaries or who preferred to have searches made for them. The title given to the position was scientific literature analyst.

The staff was given a four-day training course after which they were able to index documents. Their work was carefully reviewed to maintain consistency in indexing viewpoint. Consistency and growth were also encouraged by arranging the indexer's desks to permit discussion and communication. Provision was made for the staff to attend research reviews to broaden their knowledge of departmental technology and to attend professional society meetings to acquaint them with current research in the field of documentation.

With a permanent staff of trained literature analysts available, the users of the information system were encouraged to call the Information Center for search service instead of using the dual dictionaries which in time get out of date as new reports are indexed. It was decided not to publish updated issues of the dual dictionary and that all searching would be done by the literature analysts. To this end, a telephone extension was reserved for search requests and connected to the phone of each of the analysts who rotate searching duty. The reason for having searches run by the literature analysts is that the most accurate information retrieval is obtained by those doing the indexing. The second reason for having the same people search and index is the educational benefit derived from seeing the product of their indexing put to use. The more one uses a system to search for information, the better understanding one has for indexing and *vice versa*.

Searching is done both manually from printed indexes and mechanically by the 650 computer. The ability to perform fast searches manually on terms which are lightly posted is invaluable, whereas the saving in time permitted by mechanized searching of densely posted terms is also necessary.

The search strategy employed is to select all related terms to minimize differences in indexing viewpoint. This strategy is possible with deep coordinate indexing using links and roles which eliminate a great deal of false retrieval. Answers still must be screened for direct relevance to the question asked. Analysts are urged not to screen documents too rigorously but to leave the final judgment to the inquirer. Documents themselves are sent in answer to information inquiries. The system locates documents containing pertinent information but does not answer questions or process information.

One means of assessing the value of an information system is by its use. Figure 1 shows a plot of search activity over the past three years where the continued increase in user response is notable. This increase is the result of system publicity and continued effort to provide prompt, effective service. Another means of measuring a system's value is in terms of technical man hours saved. The classification system used formerly by the department required about 24 hours per search compared with an average of one hour per search by the Coordinate Index. This difference of 23 hours per search at 800 searches per year represents an annual saving of 18,400 man hours which can be invested more profitably in technical work.

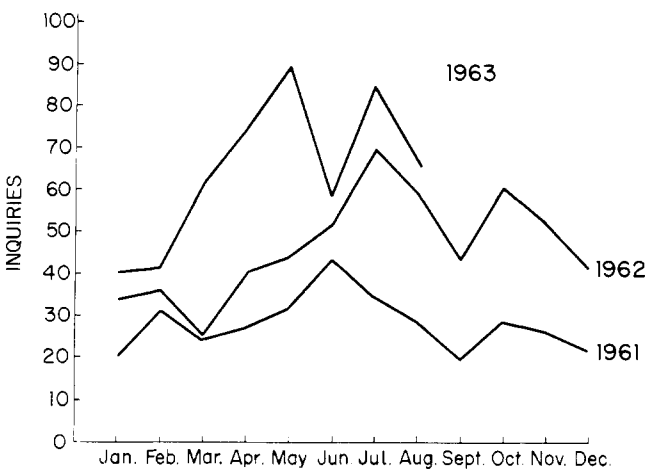

Fig. 1.—System inquiry activity.

One technique which has been successful in encouraging use of the system is the distribution of a publicity flyer shown in Fig. 2. So much reading material crosses the technical man's desk that it was felt that a graphical message would be more effective than the written word alone. Response to this flyer has been very gratifying.

The information system originated within the Plastics Research Division as a solution to its information problem. Consideration has been given from time to time regarding its proper location. The system remains in the Research Division because the literature analysts deal with technical information and thrive on the stimulation afforded by a research environment. Further, the majority of requests to date come from research personnel, and it is important to

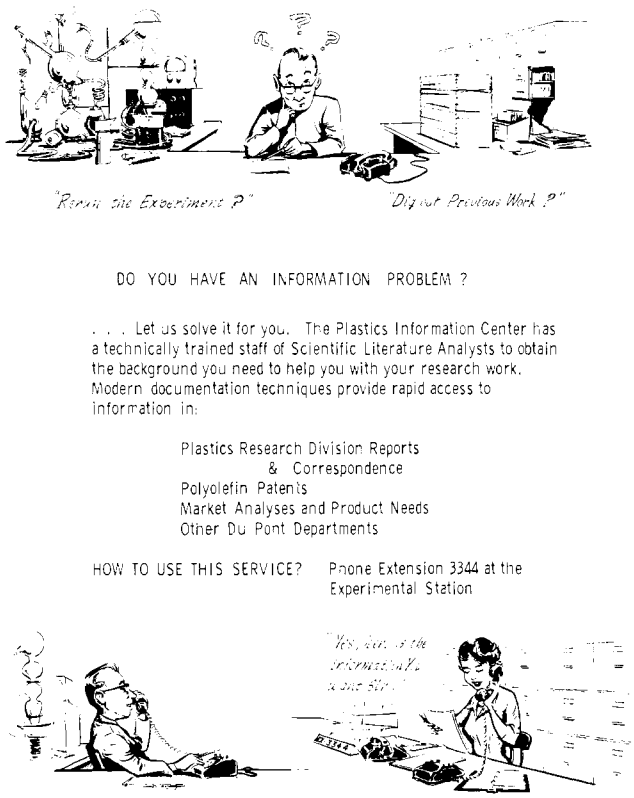

Figure 2.

be physically located near the clients. The close connection of the analyst to the center of research activity improves her understanding and familiarity with the technology, whereas, the personal interaction between analyst and user improves communication and thus effectiveness of service. One of the most critical elements of retrieval of stored information is the ability of the analyst to negotiate the question with the inquirer to obtain a clear picture of the information he desires. This is a skill not to be underestimated, for the success or failure of the entire system rests on this key ability. Comprehension of the user's needs is then followed by the ability to translate this query into system language. The success of a technical information system measured by the frequency of its use depends upon the ability of the staff to create an image which commands confidence and respect in the minds of the users.

Extension of the System to Include Other Divisions.—With the research report system successfully installed and operating, a new problem was presented. The attorneys in the Patent Division of the Plastics Department asked whether the documentation techniques used to solve the research information problem might provide a solution to a similar need regarding patents. To answer this question, a research effort was undertaken by indexing 5000 U. S. patents comprising the Patent Division's collection of polyolefin art.[4] These patents were selected because they represent coverage of one well-defined area of technology and provided a wide range of indexing challenges in patent literature covering composition of matter, process, apparatus, and articles of manufacture. An alternate route would have been the selection of all patents covering some specified time period. It was felt that limiting the collection to a single body of art would provide a more useable package for searching.

Twenty per cent of the patents, specifically the articles of manufacture, were indexed by summer employees hired as additions to the regular staff. It was found possible to use these people effectively by drawing up indexing ground rules for them to follow. This proved to be an inexpensive means for processing a large number of patents, and the ground rules provided a written record of indexing for future reference. Installation of the system was successfully completed using the same principles and elements of the report index.[5]

An analysis of the Plastics Patent Index is currently under way by comparing it with three other systems, two information systems based on concept coordination and the classification system used by the Plastics Department's Patent and Contracts Division. The analyses are being performed by addressing a series of the same questions to all four systems and comparing the patents retrieved by the Plastics Patent Index with each of the other three systems. The comparison reveals reasons why some patents were missed and others retrieved falsely by each system. Analysis of the results permits insight into the value of system components such as concept coordination, deep indexing, links, roles, and vocabulary control. Moreover, it can demonstrate quantitatively the effectiveness of coordinate indexing by comparing this approach with a classification system.

The next division which expressed an interest in utilizing concept coordination techniques for information handling was the Sales Division, which requested the extension of the information system to include its market analysis reports, correspondence, and salemen's call reports. Like the research chemist and the patent attorney, the market analyst found their information techniques were failing to meet his needs.

The system approach used for research and patent information proved to be completely applicable to and compatible with marketing information. The same vocabulary, roles, and links are used for all three types of documents. This enables searches to be performed simultaneously for all three types of information if so desired. Another advantage in using one approach throughout is return on prior systems investment.

Such requests for extension of the system to include new types of information are accompanied by questions regarding cost. Considerable experience in designing information systems has generated valuable data needed to prepare such cost estimates. Figure 3 presents a graph illustrative of the kind of data used to assess the cost of proposed information systems. From past experience the documents to be included can be evaluated as to necessary depth of indexing, and from such a graph the indexing time, and hence costs, can be determined. This has proven invaluable for questioners desiring a quick answer regarding costs.

The information service provided to the Plastics Department is rounded out by access to the technology of other departments by means of an intracompany exchange among information centers. Access to the outside literature is available through the technical libraries of the company.

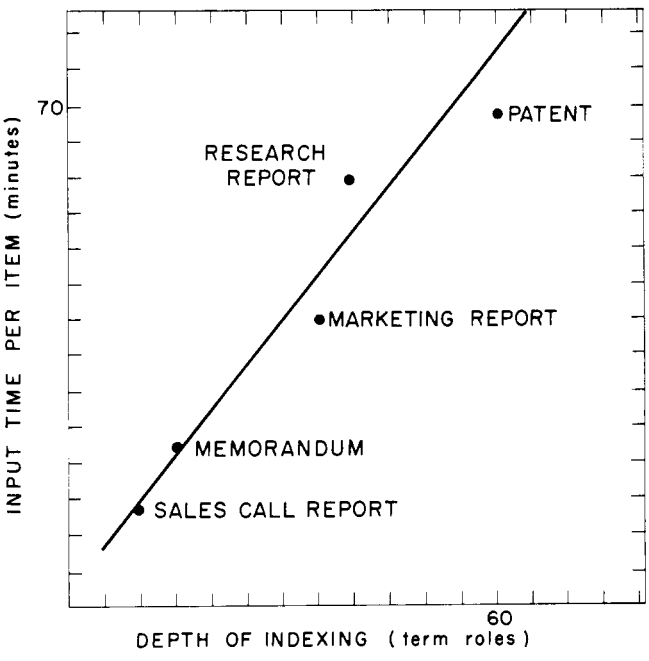

Fig. 3.—Estimating systems installation costs.

SUMMARY

A technical information system has been designed for the research, patent, and sales personnel of the Plastics Department. It has been possible to realize a return on the initial investment in system design by extending it to include other types of the Department's technical information. Future areas to be considered for inclusion in the system are technical sales information, storage of property data, technical correspondence, and extension of the patent index to other arts. An analysis of the system is in progress, and user interviews are planned in order to provide improved service.

Information systems can be justified by time saved. Practical systems tailored to meet the needs of the users are valuable tools for a technical research organization.

REFERENCES

(1) J. C. Costello, Jr., *Am. Doc.*, **12**, 111 (1961).
(2) "The Naming and Indexing of Chemical Compounds by Chemical Abstracts," Introduction to the Subject Index of Vol. 56 (Jan.–June, 1962) of *Chemical Abstracts*.
(3) B. E. Holm and L. E. Rasmussen, *Am. Doc.*, **12**, 184 (1961).
(4) B. A. Montague, *ibid.*, **13**, 104 (1962).
(5) "Nonconventional Technical Information Systems in Current Use," National Science Foundation, No. 3, Oct., 1962, pp. 108–110.