The Development of an Operating Information System*

By J. G. VAN OOT
Textile Fibers Department, E. I. du Pont de Nemours & Co., Wilmington, Delaware
Received June 19, 1962

The first part of this paper describes briefly the development of our information system: some of the intellectual problems involved in converting a punch card system to computer operation, with a few comments in retrospect as to what we might have done differently. The second part comments on the use of links and role indicators, and offers some thoughts on the usefulness of a thesaurus and on the handling of generic questions.

Several information centers in du Pont serve the various departments; my affiliation is with the one that serves primarily about 1500 technical people in the Textile Fibers and the Film Departments. The documents we process are primarily internal research reports.

HISTORICAL

Punch Card Operation.—In 1950, our research management recognized that the subject card catalog of our research reports was not adequate to handle our information retrieval needs. The Textile Fibers-Film Information Section was then established, and was the first major group in our company to index reports using concept coordination techniques. Indexing was done on a centralized basis. The Information Section prepared McBee Edge-Notched Punch Cards and distributed them to the various laboratories and plants. Searching for information was done on a decentralized basis by the librarians, file clerks, or the research men themselves at these several locations.

Details of this punch card operation and the coding system used have been published by Dinwiddie and Conrad.[1]

Bendix G-15D Operation.—In 1956, a study was begun on the practicality of converting our decentralized punch card searching files to centralized computer searching. We completed the programming and conversion to the Bendix G-15D computer by late 1958. Within a few months we had demonstrated that centralized computer operation was entirely feasible. After a year of trial operation, we discontinued the use of punch cards as a search tool, and began searching with the computer. Details of the Bendix operation were published by Grandine, Starr, and Putscher.[2]

Abstracts of all the reports indexed are still sent to the various plant and laboratory locations, but are now on 3 × 5″ cards rather than on the McBee 5 × 8″ punch cards. These cards are prepared by a "Multilith"-"Xerox" process from mats that are typed with the Document Writer attachment to the IBM 26 Card Punch, simultaneously with the punching of cards for updating the computer file. We now have just under 20,000 documents indexed, with a vocabulary of about 20,000 terms.

IBM 705 Computer Operation.—In early 1960, the use of the Bendix computer was discontinued, and our searches are now made on an IBM 705 computer. The file was converted from a serial one on the Bendix to an inverted file on the 705. The change to the 705 was made because the larger machine, operated by a completely staffed auxiliary organization, was available to us and was faster than the Bendix, because we were approaching the capacity of one tape on the Bendix, and because the Bendix computer was needed for full-time research use.

With the establishment of a centralized searching service, each of the four searchers in the Wilmington location has been provided with a trunk-hunting "report index" telephone. These four take turns of a week at a time in answering questions which come to us from locations in a seven state area. We answer at the present time just over a thousand questions per year. About half of them are answered directly by the computer, the other half by manual coordination of terms in the printout of the inverted file, prepared by the computer.

WHAT WE HAVE LEARNED

Conversion From Cards To Computer.—Now let us stop a moment and see what these changes have taught us. First there's the change from the hand-operated system to machine operation. The important message here is that a small information system which uses the concept coordination principle in indexing and retrieval, can be converted to machine manipulation when the system is large enough to warrant this. The use of hand-sorted, edge-notched cards is entirely feasible until the size of the file warrants mechanization.

Use of Information Specialist.—If we were to undertake the punch card operation now, there is certainly one thing we would do differently. This is to have the same people do *both* the indexing and the searching. As I mentioned, the indexing with the punch card system was done by a central group, and the searching was done at the individual locations by different groups of people. This did not provide adequate communication between the indexer and the searcher.

When our indexers began *answering* questions, they found that much of the information they were putting into the system was not the type that the questioners needed; and conversely, that some of the other information which questioners often asked was not being adequately

* Presented before the Division of Chemical Literature, American Chemical Society, Washington, D. C., March 23, 1962.

indexed into the system. We are thus convinced that the most efficient and effective method of both storage and retrieval of information in a deep index is through the intermediary of the information specialist.

Our report authors provide abstracts and suggest indexing terms to what they feel are important in their reports. However, our indexers use these only as a guide, and do not rely on the authors to do an adequate job of indexing their own reports. The authors simply do not possess (nor do they need to possess) the intimate familiarity with all of the conventions, the ground rules, that have been adopted to ensure uniformity of indexing and vocabulary control. Without this uniformity, and I mean uniformity of indexing not only among the various indexers, but also uniformity for one indexer from one week to the next, there is no assurance that the information will ever be found again.

By this same reasoning, *we cannot expect the research man to do his own searching*, for the framing of a search question requires this same intimate familiarity with the indexing rules. If there are several ways used to enter a certain type of information into the system, all of these several ways must be used to retrieve it, or else some of it may be lost. Conventions must therefore be adopted to keep these several ways to a minimum, so that when a search is made of the file, these same indexing conventions can be used to assure complete retrieval. It is neither necessary nor desirable for the research man to learn these conventions in order to make his own searches.

In practice, there must be some redundancy in searching as well as in indexing. We ourselves ask, on the average, four questions of the computer to cover adequately all aspects of one inquiry from a research man. For the most part, however, this represents asking the same question in both broad and narrow ways. If there is a lot of information in the store, the narrow question will serve to isolate the documents of highest pertinence; if there are only a few documents pertinent to the question in the store, the broad question will serve to retrieve those that the narrower question might miss.

Depth of Indexing.—The change from punch cards to computer operation has had other consequences. In the punch card operation, the number of terms used per document was somewhat limited, to avoid the necessity of punching many cards for one report. Accordingly, the terms we used were rather broad in scope. Now that the computer is being used, there are essentially no limitations, as far as the computer is concerned, with regard to the number of terms per document. Therefore, we can index more deeply now than we did with the punch card system. This in itself increases the possibility of false correlation between terms and hence increases the amount of non-pertinent retrieval. In addition, we inherited the relatively broad terms used with the punch cards, and this aggravates the non-pertinent retrieval problem. These broad terms are being eliminated: as an example, our old indexing term "Gloves, hosiery, and sweaters" is now subdivided into the three individual terms. With the advent of computer operation, we found it desirable to scan the abstract cards of the reports listed by the computer in an answer, when this list was long. In this way, we avoided delivering to our client many non-pertinent references.

Systems Development Work Necessary.—The fact that we were getting many report numbers in answers, even when we used quite specific searching terms, indicated to us that some refinements to our techniques were in order. We therefore undertook an analysis of our system, and with the assistance of two documentation consultants, arrived at a list of potential changes designed to cut down on non-pertinent retrieval and give us greater power and flexibility in searching.

These comments are not intended to derogate unduly our present system; it is working well; feedback from our clients is generally quite favorable. Any refinements we might make in our techniques are aimed at being able to serve our clients equally well or better when our information inventory will be twice what it is now. A critical evaluation of these refined techniques was therefore in order before we modified our existing methods. The preliminary results of this evaluation constitute the second part of this paper.

TESTING OF REFINED TECHNIQUES

General Description of Test.—Our purpose in carrying out the test mentioned above was to evaluate vocabulary modifications, certain changes in mechanical operations, and the usefulness of links and roles (a link associates indexing terms for a single idea in a document; a role appended to a term states what the relationship of that term is to the other terms in the same link; these are discussed in more detail below). To do this, we re-indexed, using links and roles, 550 reports, representing all aspects of our technology, which had already been indexed by our conventional techniques.

All of the indexing terms then were edited. In this process we examined each of the terms, decided whether or not that term should be included in our experimental vocabulary, wrote down all of the terms in our vocabulary related to that term, and established generic trees including chemical structure aspects. This information was punched on IBM cards which were used, in conjunction with IBM 1401 and IBM 705 computers, to print out an inverted file, term-role abstracts, and a thesaurus. These documents are similar to the ones described by Costello,[3,4,5] Morse,[6] and Holm.[7]

A term-role abstract (Fig. 1) is a listing of all of the terms, along with their roles, used to index a document or link. These term-role combinations for each link form

Document Number 123A

Roles	Interpreted Roles	Terms
8	Info. on	Precipitation
7	Product	Barium sulfate
3	Starting with	Barium chloride
3	Starting with	Sodium sulfate
5	in/on	Water
6	By-product	Sodium chloride
2	Effect of	Temperature
9	Effect on	Size
11	of	Crystals

Fig. 1.—Term-role abstract

a so-called "telegraphic style" abstract which can be read by anyone familiar with, or having a key to, the meaning of the roles. It is possible to have a computer interpret the meanings of the roles so that these telegraphic style abstracts can be read by anyone, but we have not done this yet. If one is satisfied with this type of abstract, it can be produced automatically by computer. The extra time that is required to append roles to the terms is thus offset by not having to prepare a short descriptive abstract when indexing without roles.

The inverted file is the usual listing, under each term-role combination, of the documents with their links that were indexed under that term-role combination (Fig. 6). A detailed discussion of the thesaurus is given below. The searching phase of this experiment is only now being evaluated. Therefore I can present here only our tentative conclusions.

Links.—A link is a symbol employed to associate terms used to index one idea, and to dissociate these from the terms used to index another idea in the same document, to prevent the false correlation of indexing terms between the two different ideas. The link symbol is appended to the document number. Links, as we use them, show only that terms *are* associated with each other. They do not show *what* the relationships are between terms; this is the function of roles. The word "link" is also used to mean the group of indexing terms associated by the link symbol and/or that portion of the document indexed by such associated terms.

We find links valuable, easy to use, and we will probably adopt them. The question is: to what extent should they be used?

Let me illustrate our problem. Suppose we have a report that deals with the Driving of Cars, the Beating of Rugs, and the Washing of Dogs. This might logically be divided into three links, as in Fig. 2. Term-role abstracts for these three links would be clear and understandable. On the other hand, no false correlation would result if we combined links A and B. That is to say, no one would be looking for information on the Driving of Rugs, or the Beating of Cars. (This assumes that there are no flying carpet operators or stock car racers in your organization.) If you use term-role abstracts, however, it might be better to keep the two ideas separate, because the term-role abstract with two ideas combined might not be clear. This is particularly true for a deeper index with more terms per link. In the example, the separation of the indexing terms in link C from those of link A or B is desirable to prevent the false retrieval of information on the Washing of Cars, the Washing of Rugs, or even the Beating of Dogs.

An additional problem with a deeper index is the question as to whether or not one should separate subordinate details into several links. For example, let us say that the part of the above document dealing with the washing of dogs actually discussed the washing of poodles with soap, and the washing of collies with detergents. This could be indexed with either one or two links, as in Fig. 3. With the one link, this document would be retrieved if one were looking for information on the washing of collies with soap; this is a false retrieval, since the document stated that the collies were washed with detergents, not soap. If it is important to make this distinction in the details, using the two links would be entirely appropriate. You may be able to see the problem better if you were to substitute here your own process or product with subordinate details.

The effect of using the same term in several different links is the accompanying increase in the number of postings, and this will lengthen the search tape or increase the number of punched cards. Subdividing the document into too many links may also result in loss of information (I might have been interested in the washing of collies with detergents, even though I asked for information on washing them with soap).

We have therefore concluded that links are valuable to prevent false retrieval, and to facilitate the reading of term-role abstracts. But as far as our index is concerned, we feel that links should be used only to separate relatively major parts of a document. The technology in certain documents may warrant the use of *numerous* links. It becomes a matter of judgment, then, as to what extent the *details* in a document need to be separated from each other by links, and it should be recognized that a practical degree of use of links will not eliminate all possible false retrieval.

Roles.—Now let us consider roles. A role appended to a term states the function of that term in the document or link. A role appended to a chemical term, for example, may show that this chemical is a reactant, or a product, or a solvent, or a catalyst, *etc.*

An analysis of the searches we have made with the computer has shown that the searches that have taken longer than usual were lengthened by repeated comparisons of document numbers listed under the very heavily posted terms. The term "nylon fiber" has, for example, been used more than 6,000 times in our operating system.

If the block of reports now under the term "nylon fibers" had been indexed with roles to show which of these dealt with nylon as a product, which as a starting material, *etc.*, our file would have several smaller blocks of report numbers under "nylon fiber" in its various roles. Because some of the reports would cover this term in more

Link A	Link B	Link C
DRIVING	BEATING	WASHING
CARS	RUGS	DOGS

Fig. 2.—Use of Links

Using One Link:

Link D
WASHING
POODLES
COLLIES
SOAPS
DETERGENTS

Using Two Links:

Link E	Link F
WASHING	WASHING
POODLES	COLLIES
SOAPS	DETERGENTS

Fig. 3.—Use of Links

AMINOBENZENE, SEE ANILINE

ANILINE

SF AMINOBENZENE
PO AROMATIC COMPOUNDS
PO AMINES, PRIMARY
PO AMINES

AMINES, PRIMARY

GT ANILINE
GT ANILINE, 2-CHLORO-
GT ANILINE, 2,4-DICHLORO-
GT BUTYLAMINE
PO AMINES

Fig. 4.—Thesaurus Entries, Showing Generic Classes

than one role, the total number of postings would be larger than it is now. But searching would, more often than not, be limited to the comparison of report numbers in only one of the smaller (term-role) blocks, thus reducing searching time.

In addition, the number of non-pertinent documents retrieved would be reduced, since, for example, reports in which nylon was a starting material would not be retrieved when reports with nylon as a product were asked for.

On the other hand, over half of the terms in our vocabulary have been used to index only one or two reports each, and three-quarters of our terms are indexing terms for fewer than ten documents. Questions involving these terms can quite easily be answered by manual look-up in the printout of the computer's inverted file. Here the role gives the searcher an immediate idea as to the nature of the information about this term that will be found in the documents listed. However, the remaining one-quarter of the terms, or about 5,000 of them, are the ones we are primarily concerned with, as they constitute about 90% of the postings in our system.

Our preliminary analysis has shown that roles can be helpful. They must be few in number, however, and their definitions broad and mutually exclusive. I believe that roles, to be most effective, should indicate little more than "who does what to which, and with what." For chemical reactions, the set of role definitions recently adopted by the American Institute of Chemical Engineers is reasonably satisfactory. Any attempt to make the roles more specific in meaning than this, will result in increasing the number of roles. Numerous roles require highly specific definitions and complex rules for their use. This increases the time required to index documents and the possibility of losing information in searching. Therefore, while roles can be useful, they must be used judiciously and in moderation, or else information will be lost.

The Thesaurus.—According to Vickery,[8] "Thesaurus" is a new word in documentation, and has been applied to several different types of word listings. Our thesaurus is closest to that described by Costello[4] and Wall.[9] It is simply a listing of each term in our experimental vocabulary, and under each term, a sub-listing of all other terms related to that term. These relationships may be generic, specific, or associative.

The associative relationships are simply "see also" references to synonyms and near synonyms. For example, under "Agitation" might be listed as related terms (identified with the letters RT) "Mixing," "Slurrying," "Tumbling," *etc.*

Fig. 4 shows other types of thesaurus entries: "See" references are handled in a straightforward manner. After "Aminobenzene," for instance, appears the instruction "see "Aniline." Under "Aniline" is the designation "SF (Seen From) Aminobenzene." Note also that under "Aniline" appear the higher generics, keyed with the letters "PO" (for Post On), "Aromatic compounds," "Amines, primary," and "Amines." Other generics might have also been listed here, *e.g.*, "6-carbon compounds," "Nitrogen compounds," *etc.*

In a similar manner, the higher generics have the specific terms listed under them. In this example, "Amines, primary" has listed under it the term "Aniline," and all of the other primary amines in the system, with the key letters GT (Generic To) printed out in front of each specific term. In this figure, the code numbers for each of the terms have been omitted; they actually appear in the thesaurus in front of each term.

It might be noted here that, the more complete such a thesaurus is, the more it approaches a classification index to the terms. There is no requirement that a single idea or thing be classified in only one class. A house cat, for instance, might be classified as a mammal, vertebrate, animal, pet, nuisance or mousetrap, and be retrieved from any of these points of view on search. The thesaurus is, however, an index to the terms in the vocabulary, and not a classification of documents.

The preparation of such a thesaurus may seem a formidable task. Without implying that it is not, I might say that the use of a computer (we used the IBM 705 and the IBM 1401) to sort and print out the material in the desired format simplifies the job enormously.

Chemical Structure and Second-Level Index to the Vocabulary.—The listing under each generic term of all its specific terms in our vocabulary provides us with a chemical structure index, and what we call a second-level index to our experimental vocabulary. This idea was suggested to us by Taube,[10] and is very similar to the "variable scope patent searching" technique of Leibowitz, Frome, and Andrews.[11] It will become, I am convinced, a valuable tool when we apply it to our operating vocabulary. Fig. 5 illustrates how it works. The specific terms in the thesaurus can be thought of as *items*, under the generic terms as *terms*, in a concept coordination inverted file. It is a simple matter then to coordinate all the specific terms under the generics "amines, primary," "aromatic compounds," and "chloro-organic compounds" to obtain the names of all chlorinated aromatic primary amines there are in the system. In the example, orthochloroaniline would be one of such compounds so identified.

The searcher can then take all of these, or select some of them in which he has an interest, and then consult the term-document inverted file to obtain the accession numbers of the documents with information on these particular compounds, in the usual manner. In Fig. 6, the documents listed (*i.e.*, document number plus link letter) under chloroaniline in role 3 are the ones where this chemical is used as a starting material. If the searcher is interested in the synthesis of chloroaniline, he would

look in the documents under this term in role 7, the "product" role.

We have applied this idea not only to things, such as chemical compounds, but also to ideas, such as physical properties. Thus under the generic "physical properties," we have listed the specific ideas "melting point," "refractive index," "tensile strength," *etc.* (This is not a data retrieval system as we use it now, however, that is, no numbers, *e.g.*, actual melting points, are given.) The many experimental and commercial fibers and films we have made have also been listed this way. We can locate, for example, the codes for fibers we have made that meet certain specifications as to tenacity, luster, yarn count, composition, etc.

We are now considering the practicality of putting this second-level index on tape. Since the logic of search of this file is the same as that of our document file, we believe that computer searches of this generic tape could be made.

Generic Searching.—The combination of computer search of the generic tape and the document tape presents the interesting possibility of handling generic questions *during the search.* This is apparently the technique used by the Saint-Gobain Company in France,[12] and was suggested by Yngve.[14]

Costello[4] and Montague[13] have described a system to handle generics by posting, with the computer, of specific terms to the generics *on input.* In this method, the numbers of the documents indexed under, say, "chlorine" would be automatically listed by the computer under the term "halogens" in the inverted file.

Handling generics during the search, however, has certain advantages over the techniques of posting generics on input. First, the second-level index to the terms in our vocabulary permits us to determine directly, without consulting the documents, what specific chemicals, fibers, films, or ideas that meet the desired specifications, have been indexed in our file. Second, the combined length of the generic magnetic tape and document tape will be shorter than the document tape with all generic postings. Search time, and hence computer rental cost for searches, may be thus reduced by searching generically. Third, the level of specificity of the entries on the document

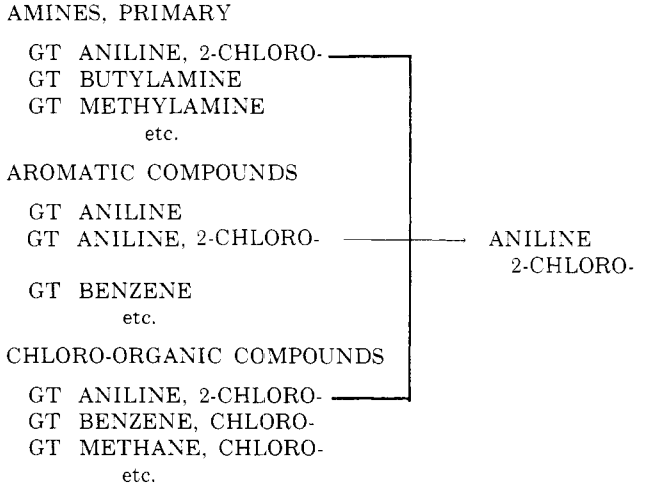

Fig. 5.—Generic File

ANILINE, 2-CHLORO- (Role 3)
67B 93A 104D 605A
ANILINE, 2-CHLORO- (Role 7)
23B 97A 167D

Fig. 6.—Term-Document Inverted File

tape is the same as that of the document; this is particularly important when searching the patent literature using such a system.

There are techniques that can be used with the generic posting route that can overcome some of its disadvantages; likewise, there are some disadvantages to the generic treatment during the search. We are now studying the pros and cons of both of these methods.

SUMMARY

In summary, an information system, using edge-notched punch cards, can be converted to a mechanized system using concept coordination techniques. In a deep index, it is very desirable to have the same people who search the file also put the information into the file. Vocabulary control, including a careful definition of terms, is a necessity. Ground rules for using these terms must be comprehensive to assure uniformity and consistency of indexing and searching; without this consistency, information is lost. We are convinced that the provision of a second-level index and chemical structure index to our vocabulary will be a valuable addition to our information system. In our opinion, links and roles are helpful, but the extent of application of these techniques is a matter of judgment and practicality, and depends to a large degree on the depth of indexing. Finally, we are evaluating carefully the potential of retrieval of generic information during the search, rather than providing generic postings on input.

Acknowledgement.—Most of the work in the evaluation of new documentation techniques reported here was done by Miss Ruth E. McFarlane, Mrs. Florence H. Kvalnes and Dr. John L. Schultz. I am indebted to them, and to Dr. C. C. Conrad and Mr. K. H. Zabriskie for the many animated and stimulating conversations in developing most of the opinions and ideas presented in this paper. Our thanks are also given to Mr. T.B. Oglethorpe and his staff for programming and running the special computer operations necessary for the experiment. Miss Louise Rushmer also assisted in the editing of the terms in the test vocabulary.

REFERENCES

(1) Dinwiddie, S. W., and Conrad, C. C., "Report Indexing by Hand-Sorted Punched Cards," in "The Technical Report," edited by B. H. Weil, Reinhold Publishing Corp., New York, N. Y., 1954, pp. 303–316.

(2) Grandine, J. D., Starr, E. M., and Putscher, R. E., "Report Index Searching on the Bendix G-15D Computer," *J. Chem. Doc.*, **2**, 79 (1961).

(3) Costello, J. C., Jr., "Uniterm Indexing Principles, Problems, and Solutions," *American Doc.*, **12**, 20 (1961).

(4) Costello, J. C., Jr., "Storage and Retrieval of Chemical Research and Patent Information by Links and Roles in du Pont," *American Doc.*, **12**, 111 (1961).
(5) Costello, J. C., Jr., "Some Solutions to Operational Problems in Concept Coordination," *American Doc.*, **12**, 191 (1961).
(6) Morse, R., "Information Retrieval," *Chem. Eng. Prog.*, **57**, 55 (1961).
(7) Holm, B. E., "Information Retrieval—A Solution," *Chem. Eng. Prog.*, **57**, 73 (1961).
(8) Vickery, B. C., "Thesaurus—A New Word in Documentation," *J. Doc.*, **16**, 181 (1960).
(9) Wall, E., "A Practical System for Documenting Building Research," in "Documentation of Building Science Literature," publn. 791 (1960), p. 15; U. S. Senate Report, 86th Congress, 2nd Session, Document No. 113, 1960.
(10) Taube, M., private communication. Also "Final Report: Feasibility Study for Coordinate Indexing and Machine Retrieval of Chemical Warfare Laboratory Library Documents." Prepared under Contract Number EA18-108-405-CML-676 for the Chemical Research and Development Laboratory, U. S. Chemical Corps, Ded. 1960, and "Handbook of Procedures and Comprehensive Report" prepared in conjunction with above feasibility study.
(11) Leibowitz, J., Frome, J., and Andrews, D. D., "Variable Scope Patent Searching by an inverted File Technique," U. S. Patent Office, Research and Development Report Number 14, Nov. 17, 1958.
(12) Levery, F., Reibell, B., Cohendet, R., de Picciotto, S., Biet, F., Picot, G., Deribere-Desgardes, M. L., Slove, D., Escaravage, R., and Fournier, M., "Une Experience de Selection Automatique de Documentation," published by IBM France and the Compagnie de Saint-Gobain, Dec., 1960.
(13) Montague, Barbara A., "Patent Indexing in du Pont by Concept Coordination Using Links and Roles," *American Doc.*, **13**, 104 (1962).
(14) Yngve, V. H., "The Feasibility of Machine Searching of English Texts," in "Proceedings of the International Conference on Scientific Information," Washington, D. C., Nov. 16–21, 1958, Vol. 2, p. 975.

Five Operational Years of Inverted Index Manipulation and Abstract Retrieval by an Electronic Computer*

B. K. DENNIS, J. J. BRADY, and J. A. DOVEL, Jr.
Technical Information Center, Flight Propulsion Division,
General Electric Company, Cincinnati 15, Ohio
Received May 9, 1962

In this presentation we have attempted to record the salient facts about the conception, development, operation and continued improvement of our Automatic Information Retrieval System (AIRS). We have had a great many experiences in our information retrieval work, both good and bad. The path we have followed over the past five years has been rocky. We have stumbled and fallen flat on our faces more times than we care to recall, but we are still moving forward. We are happy to note that our progress has accelerated during the past year, showing specific pay-off in terms of a growing body of customers satisfied with value received from our services.

As our discussion proceeds, you should bear in mind that the General Electric Automatic Information Retrieval System has not been developed and operated as an experiment under research laboratory or "hothouse" conditions. We have not had a staff of information system theoreticians and specialists along the way to tend it, nurture it and keep it progressing toward carefully established and well defined experimental objectives with adequate financing to reach those objectives. To the contrary, our information retrieval system was developed in response to a felt but not clearly defined need. It has been considered a production tool from the outset, available to its users on a specific charge basis. Its design, development and operation has been solely in the hands of information service and computer programming practitioners. Its rate of development has been dictated by a widely fluctuating level of financial support.

Since the unveiling of our retrieval system in 1958, many people have visited or contacted our Center to learn more about our automation work, with many visitors providing much valuable advice. There have been those who have taken exception to our approach. Early in our work with the IBM 704, one outspoken critic accused us of "using a sledge hammer to drive a carpet tack." That comment always came to mind every time we were required to break up a machine search and re-run it because the memory of our 704 could not handle it. There have been times when our searchers felt that they were attempting to drive a railroad spike with a tack hammer and would have been pleased to inform the "sledge hammer" critic accordingly.

During this discussion, we will attempt to provide answers for many of the questions which have been asked by visitors during the past few years. Of necessity, much of what we have to say will be of an intuitive or "educated guess" nature. One of the unfortunate aspects of attempting to advance the state-of-the-art in a service organization is the fact that the more sophisticated scientific or experimental objectives may often be compromised by the immediate necessity to provide a service. Therefore, much experience data which would be

*Presented before the Division of Chemical Literature, ACS National Meeting, Washington, D. C., March 23, 1962.