Documentation

ALLEN KENT

It has been said that documentation can be defined as librarianship in high gear. The processes, or unit operations, of documentation may be considered to coincide completely or partially with those of librarianship, depending upon how narrowly or broadly a particular documentalist or librarian views his field.

A 1956 definition of documentation certainly reflects this viewpoint: "the group of techniques necessary for the ordered presentation, organization, and communication of recorded specialized knowledge, in order to give maximum accessibility and utility to the information contained." ¹

But despite evidence to the contrary, the author is convinced that the gap between documentation and library practice is narrowing and that the definition that in 1956 clearly distinguished documentation from librarianship no longer serves as well as it was once thought to. The narrowing of the gap is being caused by many forces. Two of the major ones seem to be these: (1) Many librarians have been awakening to the almost lost opportunity to serve the dynamic information requirements that they spurned for more than a decade. The ability of the documentalists to attract financial support and respectable salaries has caused many librarians to re-examine their stand and to begin to exploit in their day-to-day activities the fruits of documentation research. (2) The maturing of the documentation field has started to purge itself of the intemperates who, although recognizing the need for more dynamic information services, did not always take into account the well-learned lessons of librarianship in designing and operating systems for exploiting recorded knowledge. The more successful documentation systems have sound basic principles that are in common with those of sound library operations.

¹ The author is Professor of Library Science, and Associate Director, Center for Documentation and Communication Research, School of Library Science, Western Reserve University, Cleveland, Ohio.
Documentation

Of significant interest has been the increase in the periodical and report literature, to the point where, in the sciences alone, it is estimated that between 500,000 and 1,300,000 items are to be published in 1961. This number has been increasing dramatically with no sign of letup. The problems of selection and acquisition of this massive outpouring of literature are impressive.

Libraries acquire materials by communicating with a supplier who either maintains a stock of documents or is willing to prepare copies of them. The techniques of ordering, follow-up, payment of invoices, and checking in of materials received are functions that parallel those conducted in the handling of any commodity in the business community. As would be expected, various types of data processing machines have been applied usefully to facilitate the burden of record keeping in connection with the acquisition function.

More dramatic and more numerous developments have taken place in documentation with regard to the subject analysis of materials—some appropriate and useful, others inappropriate, naive, and not useful. These developments may be categorized under indexing, classifying, abstracting, and processing of full texts.

Traditionally, the subject indexing operation has involved the selection of words or ideas from a graphic record on the basis of well-defined rules; indexing has been carried out in order to facilitate the identification or selection of desired documents after they have been stored. As in any form of content analysis, some value judgments are made in the selection of aspects of subject matter as important for inclusion in an index (or in omitting aspects of subject matter as unimportant).

The availability of computers and computer-like devices which can perform repetitive tasks effectively and economically has led to the application of machines for subject indexing purposes. Examples of machine applications are as follows:

1. The concordance is an alphabetic index of words in a book in exact context. No discrimination is exercised in preparing this type of index. Each word that is present in the text is an index entry. Therefore, the decisions that must be made in conducting such an indexing operation are not very difficult and can be performed very well by machines.

2. The Key Word in Context (KWIC) Index (or Permutation Index) similarly requires little in the way of subject matter knowl-
edge and therefore is amenable to mechanization.\textsuperscript{6,7} The KWIC index is prepared by cyclic permutations of words in which each "substantive" term is brought to a predetermined position and alphabetized. This type of index is very much akin to the concordance; however, its applications have been limited to the preparation of KWIC indexes based on titles of papers only.

It should be obvious that the mere machine manipulation of words in this way does not increase the usefulness of the words as reference points in search. Therefore, if a conventional index which uses as entries only the substantive words in a title would not be considered useful, the use of computers to prepare the KWIC index will not result in a more useful product.

(3) Uniterm indexing, as described a number of years ago, involves the analysis of contents of graphic records in terms of key words which represent the content of the record that is being indexed.\textsuperscript{8} These key words include not only single common English words, but also serial numbers and other symbols, if they are found in the text, and if, in the judgment of the indexer, they represent the content of the record.

In pursuing rules of "word" indexing, as in other word indexing procedures described earlier, one should recognize that certain assumptions are being made, primarily that the user of such an index is sufficiently familiar with the subject matter of his search that he can provide the control over the words used in the Uniterm index that the indexer has been instructed to overlook (or was not instructed to take into account).

(4) Controlled indexing, as opposed to "word" indexing, implies a careful selection of terminology used in indexes in order to avoid, to the extent possible, the scattering of related subjects under different headings.

Although proponents of indexing approaches to exploiting the literature believe that the most attractive feature of the index is its ability to identify specific aspects of information that may be discussed in a document, there is likewise a desire to combine some of the advantages of classification approaches, i.e., to group related subjects.

Another method for controlling the subjects chosen during analysis is based upon the frequency of occurrence of key words in running text—those occurring most frequently being considered, prima facie, to be most significant for analysis of the texts. This type of analysis
Documentation

may be conducted by keypunching the entire text of a document so that it may be scanned by machine, and compiling statistics as to the frequency of occurrence of various key words.

An independent variable in controlled indexing is the language used to record the results of analysis of graphic records. In many ways this variable is completely analogous to the control of subjects chosen. In the latter, various means are used to limit the subject matter that is made explicit during indexing. This may be considered analogous to providing an index with a special "point of view."

Independent of the control of subjects chosen, or "point of view" of the indexer, it is considered helpful to regularize the manner in which index entries are expressed. Some of these methods are identical with those used to control choice of subjects (the subject authority list).

Another control technique that has been used in modern documentation systems is the "role indicator," or "role factors" or "role directors." These techniques have been used in various machine systems and are quite analogous to the "modifications" used with index entries in conventional retrieval systems. These indicators are useful for limiting the area of meaning of the index entry, according to the "role" that this entry plays in a particular context.9-11

The major developments in the area of classification during the past decade or so have been concerned with the multidimensional classification. This type of classification involves the characterization of each graphic record from more than one point of view. This classification can be accomplished for the physical placement of records only when more than one copy is available for filing in more than one location in the classification system.

However, many classification systems are employed as guides to a physical collection of records, as in a classified catalog, where copies of catalog cards, each one representing a graphic record, are filed within the classification system. This system makes it convenient to have as many cards per record available as there are "dimensions" to the classification.

Although this type of classification has been very popular in Europe, there have been few applications of multidimensional classifications, such as the Universal Decimal Classification, in the United States. One of the multidimensional classification systems that have been proposed recently has been in the field of education media.12

Perhaps the documentation activity that has spread most extensively
Allen Kent

in the past two decades has been abstracting. Traditionally, an abstract has been considered "that which comprises or concentrates in itself the essential qualities of a larger thing, or of several things"—a summary of a publication or article accompanied by an adequate bibliographical description to enable the publication or article to be traced. In recent documentation literature, it is possible to identify three types of abstracts: (1) traditional abstracts, (2) extracts, and (3) telegraphic abstracts.

The traditional abstracts are of two general types: descriptive and informative. The descriptive abstract embodies a general statement of the nature and scope of a document; it is not pretended that this type of abstract can serve as a substitute for reading the original document; it merely presents several clues as to whether or not the information being sought might be contained in the original record. The informative abstract, on the other hand, has the purpose of presenting (concisely) information of probable high significance contained in the original record and in the ideal case to obviate the necessity of referring to the original.

The three functions that traditional abstracts have served are (a) current awareness: to aid a reader in keeping informed concerning new developments and in acquiring new technical information; (b) reference: to provide a back file of accumulated information which may be consulted as required; and (c) indexing (or classifying): to serve as a basis for indexing a record; the abstract is prepared after a decision has been made regarding the important aspects of a record. This same analysis can be used to provide those important aspects that should be incorporated into an index or other searching tool.

An extract is analogous to an abstract in that it represents what is considered the important subject matter of a graphic record that has been selected for quotation. It is felt by some that the use of direct quotations or extracts from a record provides more effective service to a reader than does an abstract.

Extracts may be selected by human analysts or by the application of machine techniques. When machines have been used for extracting, the resulting product has been called an "auto-abstract."[18]

The techniques used by humans to prepare extracts are subjective and involve the exercising of judgment by an analyst in order to determine which portion(s) of a document is of sufficient potential significance to warrant recording. When a machine is used for ex-
Documentation

tracting, the entire text of a record is converted to machine-readable form and is scanned by a digital computer. When these methods are applied, it is assumed that the frequency and distribution of key words in the text can be used as the basis for determining the relative significance of sentences in text. In accord with this assumption, the sentences which are highest in "significance" (as determined by their containing the greatest frequency of key words) are printed out to produce an extract (or "auto-abstract").

A telegraphic abstract is a detailed index to a graphic record. It is composed of (1) significant words selected from the articles; (2) code symbols called "role indicators" which fit the selected words into context; and (3) punctuation symbols which separate and group the words and role indicators into various units in somewhat the same fashion as conventional punctuation does.

The telegraphic abstract differs from a regular index in that: (a) it contains more words; (b) it has a prescribed arrangement along quite different lines; (c) it is, when abstracters have finished it, still only a semi-finished product. The finished product is a reel of tape with the information, which the abstracters have partly furnished, translated into a computer code ready for searching by machine.

Telegraphic abstracting, then, is a method for recording important characteristics of information contained in documents so that the information may be further processed for machine input where it will serve the function of an index enabling the document to be identified by machine in answer to requests for information. The design of the telegraphic abstract is part of an overall logic which also dictates the design of the code into which the words are encoded and the design of the strategy and program by which the information is searched. The telegraphic abstract, the encoded terms, and the search program taken together comprise a machine information retrieval system. The purpose of the telegraphic abstract is to provide "input" to the machine in a consistent and predictable form so that the machine can be programmed to search for certain predictable forms of information within this input.

Another "analysis" technique involves the "processing" of a full text for retrieval purposes. In general, this type of processing implies that certain subjects and points of view have been selected from the text by human or machine analysis in order to record decisions as to what in the text of a document is of greatest probable importance for retrieval purposes.
However, some research is now in progress which has as its aim to record for retrieval purposes essentially everything that is available in the full text of a source document. The rationale behind this approach is that the indexing policies required to serve all potential needs are so diverse that essentially only the recording of an entire document would provide adequate services.

It should be recognized that the approaches to recording of all texts for retrieval purposes make certain assumptions that should be made explicit:

(1) That the text will be “read” by machine, or that the text will be made available by publishers in machine-readable form (e.g., monotype tape). It is not yet evident that the full range of type fonts, styles, and reproduction techniques will permit effective and economical programs for machine recognition of text. Otherwise, it is necessary to “keyboard” all text into machine-readable form in order to make it amenable for further processing.

(2) That it will be economical to search by machines the tremendous full texts of documents produced even in restricted fields, without attempts to “compress” the available material by a “probabilistic” approach to analysis of the document.

(3) That most questions can be analyzed with sufficient precision, that a precise selection of matching information in the full text of a document will provide useful results. It is anticipated that the major problem to be faced is that “normalization” or syntactical analysis of running text by machine may well yield answers to questions from trivial mentions of a “subject” as well as to those which are more significant.

Also, it is tacitly assumed that requests for searches will be based upon recognition memory rather than recall memory. In the former, a request for a search is based upon the way in which the text has been recorded in the remembered record. In the latter, a request for a search is based upon recalling some subject that is of interest, whether or not it may have been seen before. In this latter case, the resources of natural language are sufficiently great that the alternate clues which would have to be provided are sometimes staggering.

Another area of documentation development has been with regard to the storage of documents. This area has benefited from the wide availability of inexpensive office copying machines, which has resulted
in the ability to store single copies of documents, with extra copies for demand distribution being prepared as required.

The larger centers either have used similar approaches or have converted the document collections to microform, with a variety of techniques being available for the preparation of disposable copies or of additional generations of microform for perusal at distant locations with suitable reading devices.

Another documentation operation that has been the subject of a considerable amount of activity in the documentation field is the identification of documents from collections and delivery of documents.15

The basic developments have resulted from the following considerations. Retrieval systems—especially machine systems—have been developed or are under consideration because it is “impractical,” inconvenient, or too expensive to locate desired records from a file by other means. Usually the “size” of the file (as measured by the number of records, depth of analysis of each record, and complexity of subject matter covered) has reached the point where an existing retrieval system is not expected to provide adequate service in response to the average inquiries. When such a situation exists, it may be assumed that the number of records being incorporated into the file exceeds the ability of a potential user of the information it contains to read and to remember the contents of every record that is incorporated. When this situation has been reached, then for any record that is to be analyzed for retrieval purposes, it cannot be assumed that a person who may later want to see the record will have read it previously or even have seen it before. In these cases, requests for records containing desired information will be based upon clues, subjects, or verbalization of subjects which are probably drawn from the requester’s background and not necessarily from the text of records stored in the file.

Now, when graphic records are to be stored away for potential retrieval at a later time, two decisions must be made by the analyst of the record:

1. Which aspects of the record are of probable importance to potential users?

2. How should these aspects be expressed in the retrieval system so that there will be a good probability of matching the way potential users will think of and express their requests?
ALLEN KENT

The first decision is discussed earlier. The second decision is a particularly difficult one to make because an analyst reading a document is tempted to use words found in that document to record the results of his analysis. However, as was pointed out in previous discussion, these words are not necessarily those which are most likely to be used in information requests. So the analyst (or designer of codes, or subject authority lists) is faced with the task of providing a number of clues with regard to the subject matter of the document in an attempt to anticipate any way in which a searcher's point of view might be expressed.

Now, how is the analyst to predict which words will come to a searcher's mind when he desires information on some subject?

The following approaches may be considered:

1. If the decision is made to operate the retrieval activity as a "closed" system, this relieves the analyst of the responsibility of predicting which words will be used by "outsiders" in searching the file. All searches must then be performed by operating personnel of the retrieval system who are in a position to interpret requests in terms of the language used by analysts.

2. If the decision is made to operate the retrieval activity as an "open" system, attempts must be made to control the use of terminology in one of two ways: (a) Analysts and systems designers may feel that they are in sufficiently good intellectual contact with their potential clientele to permit them to predict terminology that will likewise be chosen by their clientele in formulating questions. (b) It is possible to submit a potential clientele to association tests, which would provide some basis for using terminology according to its probable significance to information systems users.

The associations of words to other words, whether predicted by systems designers or derived from empirical or experimental data, are based partially upon personal points of view with regard to the significance of words and partially upon the inherent meaning of words.

The developments with regard to the identification of documents from collections have involved the use of subject authority lists, glossaries, classifications, thesauri, and semantic and other code dictionaries. In their development these various tools for controlling the scatter in meaning of terminology during searching operations have much in common. The first steps involve the collection and definition
Documentation

of the terminology expected to be encountered in the searching system. The resultant "language" may differ with regard to symbolization but may have many features in common—particularly the concepts covered by the "language."

Various devices have been used for identifying relevant documents in response to questions. When a question is asked of an information retrieval system the language of the original inquirer must be interpreted in terms of the language of the system. With the newer retrieval systems involving the recording of more and more detail relating to the contents of documents, there appears to be more advantage to systematizing the methods of exploiting the system—the strategy of searching—so that consistent results may be assured even when different users attempt to operate the system.

It is suggested that some of the principles of selecting a strategy of searching may be common to various retrieval systems. The steps to be taken are (1) analysis of the question with regard to aspects of importance and desired relationships among the aspects; (2) analysis of the alternate logical configurations of the question which would lead to optimum results.

The results of a search, regardless of retrieval systems, are being presented in the form of sets of (1) pertinent bibliographic references, (2) abstracts, or (3) full documents, which are presumed to be pertinent to the question. The means used to deliver responses are dependent upon the means used for storage of documents, as discussed earlier.

And now to review documentation activities of the past decade. The various literature searching and documentation developments, although masquerading under differing labels, have at least a single common purpose—to facilitate the communication of knowledge across barriers of time, space, and language. A subsidiary common purpose involves the desire to conduct in the most economical and effective manner the various unit operations involved in achieving this communication.

In attempting to achieve a balance between economy and effectiveness, one must weigh the importance of "custom" processing of the literature to take into account the peculiar point of view of a particular requirement, as against the economy of using a "standard" processing which is performed by a "centralized" organization. The more complete the standard processing, the less is the advantage for a "custom" processing to be performed.
ALLEN KENT

For searching systems, there are usually required an analysis of the document and the selection of various aspects of probable interest and their recording in the form of abstracts, index entries, subject headings, and classification headings. It has not been considered feasible during the selection process to take into account most of the potential special points of view from which the document might be of interest during retrieval operations. However, some of the machine searching systems show promise of taking into account a sufficiently wide spectrum of potential needs that agreements on a common language, or the establishment of standards for interconvertibility among various machine languages, appear to be in order. The potential savings in time, effort, and money in avoiding the unnecessary reprocessing of documents, or in avoiding the effort involved in the development of minor alterations of already existing systems, appear to warrant a determined effort to achieve agreements.

The possibilities of developing a common language for machine searching offer attractive economies if automatic abstracting or indexing should prove to yield sufficiently promising results for certain requirements. The use of a common language would then permit the translation and also the indexing of the same text with only a single scanning.

Now let us consider documentation in relation to demography. The study of demographic trends reveals many cross currents that may influence documentation developments. As would be expected, these trends reveal the concentration of documentation centers, public libraries, and special information centers in areas of high population concentration, where they can serve the largest numbers of individuals and organizations. However, individuals and organizations, although tending to concentrate in certain areas, also decentralize in many cases.

In order for one to discern meaningful patterns, one way in which the documentation situation may be viewed is to consider the postures of large and small organizations and individuals in areas of high and low population concentration.

The large organizations tend toward broader, interdisciplinary interests. Their requirements are rather broad and encompass various subject specialties, from specific as well as from interdisciplinary points of view. Almost no organization, regardless of size, has the economic resources to acquire and to process the total available literature that will be of possible interest. However, it can establish a
special library to acquire and process the core literature in the subjects of probable interest and rely upon the resources of public libraries, universities, and other special libraries which are generally available in areas of high population concentration.

In low population areas, there is less of a tendency for large public and university libraries to be readily available; nor are there many special libraries that can serve to backstop the considerable requirements for literature. Therefore, these organizations are faced with the alternative of building up literature resources in anticipation of possible needs, or of relying upon distant special information centers for demand service in response to specific questions. Examples of reliance upon both of these alternatives may be found.

Another matter to be studied is the posture of small organizations in areas of high population concentration. Although the smaller organizations tend to have more restricted subject interests, their documentation needs are not more restricted. This is the case because they are confronted with the same massive number of publications, and they cannot afford to devote effort to the perusal of this material and the selection of material of possible or probable interest. Since the resources of these organizations are limited, they usually cannot establish an effective special library and must rely upon the local resources available in their area, backstopped by the special documentation centers throughout the country.

If we now consider the situation of small organizations in areas of low population concentration, we can discern that these organizations are faced with especially difficult problems, since their documentation needs usually cannot be satisfied by resources and services in their immediate vicinity. And since their resources do not permit the buildup of sizable collections, they are generally forced to obtain service from distantly located documentation centers and libraries. Although their requirements for service may not be frequent, their need for effective and rapid aid is nonetheless critical.

Another situation to consider is the posture of individuals in areas of high population concentration. The problem of the individual who approaches the mounting literature is perhaps the most devastating. He is faced with the dilemma of (1) finding the total literature of potential interest, (2) being unable to acquire and process for his own use only a small fraction of the available literature, or (3) being unable to afford the fees of those special documentation centers when requiring service.
ALLEN KENT

The public and university libraries are the sole source of sizable collections for the individual. However, the individual usually does not have sufficient time available to devote to the task of attempting to exploit these collections. Therefore, he must usually be content with a superficial penetration of the available literature.

Last, let us look at the posture of individuals in areas of low population concentration. Here the situation is almost hopeless. The absence of adequate library resources in his vicinity and the lack of sufficient financial resources to permit travel to distant centers or to pay the fees of special documentation centers usually force the individual investigator into organizational situations or into an urban area.

Summary

It may be evident from the foregoing that the trends in the documentation field are these: (1) for the published literature to increase in quantity and complexity as a consequence of the expected continuous expansion of research activity; (2) for the need for effective exploitation of the literature to increase because of the increasing pressures for avoidance of repetition of research effort; (3) for the demands on documentation centers to increase because of the decreasing ability of individuals and organizations to cope with the literature on an individual basis; (4) for the continuing breakdown of all but mechanized methods of information retrieval.

The ability to record information on magnetic tape, punched cards, and punched tape, and to produce copies of these media automatically and at low cost represents a revolution in the documentation field as significant and with as far-reaching consequences as those of the invention of printing. Thus is provided the opportunity to establish a new type of library, or a new aspect of traditional librarianship, which will make possible the collection and storage of analyzed and encoded published literature in a form ready for mechanized exploitation. The new types of libraries will be the customers of, or subscribers to, the centralized processing of this "machine feed." This centralized processing agency will acquire, analyze, and encode the scientific literature received from the entire world.

The processor of "machine feed" can operate under the direct control of a central agency, or independently, with permissive cooperation, through a coordinating agency. The same would be true for libraries of other information centers. In any case, the character of
the new information retrieval system should be analogous to that of the Bell Telephone System, which provides a communication network for the entire country, without a central office through which all communications must pass; rather, each point of the network is a potential substitute for a central office in that all the facilities of the entire system are available upon demand at any one point.

The analogy for a mechanized analysis-encoding-searching system for science is depicted in Figure 1. This configuration of activities provides the economic and technical advantages of centralization, but also permits any degree of decentralization that may be desirable in a particular locality or for a particular subject field.

It is the conclusion here that this type of central and coordinating agency will be in the future and will provide the next major alleviation of documentation problems.

The main function of this agency will be to conduct a continuing program of thorough acquisition and analysis of the world's literature and to make the resulting material available in many forms to a wide variety of users.

Some of the major services contemplated are charted in Figure 2. Specifically, these services involve (1) searching the accumulating machine record to provide at regular intervals current information on specific questions that are of continuing interest to subscribers; (2) searching the store of machine record retrospectively to provide information on a specific problem or make available the raw material for a review of the literature or other compendia; (3) copying of the machine record and transmittal to a local installation for search and other exploitation; (4) providing copies of abstracts and source documents for special subject fields in order to "stock" special libraries with the materials required to serve their clientele; (5) serving as a proving ground for the testing of new documentation and bibliographic methods, and promoting in other appropriate ways the improvement of the exploitation of recorded information through research.

In order to provide and utilize the machine record required for the information facility of the future, the following processing steps will be required:

(1) Acquisition of materials: selection and acquisition of source documents, or informative abstracts suitable for reprocessing.

(2) Analysis: review, by subject specialties, of the source documents or abstracts, for the preparation of stylized abstracts. (Alter-
FIGURE 1
MECHANIZED INFORMATION SERVICE NETWORK

- geographic area to be served
- area service and processing stations
- regional repositories and service centers
- regional processing centers by subject fields
- headquarters coordinating operation

[238]
Documentation

FIGURE 2
SPECIFIC SERVICES OFFERED BY MECHANIZED INFORMATION SYSTEM

Centralized and Decentralized Acquisition and Processing of Documents

Centralized and Decentralized Acquisition and Processing of Documents

Centralized and Decentralized Acquisition and Processing of Documents

Creation of Machine-Readable and Machine Reproducible Record (backed up by reproducible microprint file of conventional abstracts and original documents)

Centralized Exploitation of File for Many Purposes

Decentralized Exploitation of Reproduced File at Various Locations and for Many Purposes

Specific Machine Searches

Current Awareness Searches

Retrospective Searches

Copying of All or Portions of Machine Record for Local Exploitation

"Stocking" of Special Libraries with Abstracts and Source Documents Based on Statement of Continuous Interests

Provision of:

Bibliographies

Collections of Pertinent Conventional Abstracts

Copies of Pertinent Source Documents

[239]
nate, compatible methods of analysis, possibly based upon machine processing of source text, could be substituted if this proved to be an economical and technically feasible procedure.)

(3) Encoding: transformation of the stylized abstract into code, using automatic procedures. This code must provide a ready means for controlling synonyms and partial synonyms, and for identifying the inherent meaning of terms used in the system.

(4) Recording: transfer of encoded abstracts (or other form of compatible document analysis) to a machine record (e.g., magnetic tape, punched tape, punched cards).

(5) Storing: microstorage of conventional abstracts and source documents for each item processed.

(6) Reference services:
   (a) receipt and analysis of questions;
   (b) programming and operation of searching device;
   (c) selecting and copying of abstract or source document to subscriber.

(7) Auxiliary services:
   (a) copying of machine records or searchable file for delivery to subscriber;
   (b) copying or reproducing of all or portions of microfile of abstracts and documents for delivery to subscriber;
   (c) copying or reproducing of machine record of code books for up-dating or initial delivery to subscriber who is conducting analysis of confidential documents for private use.

Several forms of cooperative activity are desirable, and one is essential to provide the fullest and most flexible service. The desirable cooperative activities are (1) the provision of appropriately informative abstracts already being prepared in various subject fields in order that duplication of analysis of source documents may be avoided; and (2) the coordination of machine languages for storage of the machine record, in order that similar work undertaken by other groups may be used in the proposed mechanized system.

The one essential area of cooperation relates to the acquisition of rights to reproduce copies of copyrighted informative abstracts and source documents. To permit copying and delivery of such material to subscribers, the payment of a royalty or other equitable reimbursement would be necessary.

The volume of such service that could be absorbed by the American
Documentation

economy has been estimated from a market survey of representative organizations in government, industry, and other agencies. The market survey assumed the existence of a comprehensive information system at various cost levels—service to be offered either from a central organization or within an individual organization. The market sample numbered 3,000, with a response of 20 per cent. Without extrapolation from the sample to the entire population, the replies to the questionnaire suggested that more than 2,000,000 questions, either retrospective or continuing, would be asked each year (either centrally or through a decentralized activity), and that the annual worth of such a service would be slightly more than $2,900,000,000.

The "product mix" of services described in the foregoing will be necessary to satisfy the needs of various types of individuals and organizations in a changing demographic picture.

References

12. Tauber, M. F., and Lilley, O. L.: Feasibility Study Regarding the Estab-
ALLEN KENT


16. Ibid., Ch. 61.


