Facets of

The Technical Information Problem

Charles P. Bourne and Douglas C. Engelbart

STANFORD RESEARCH INSTITUTE

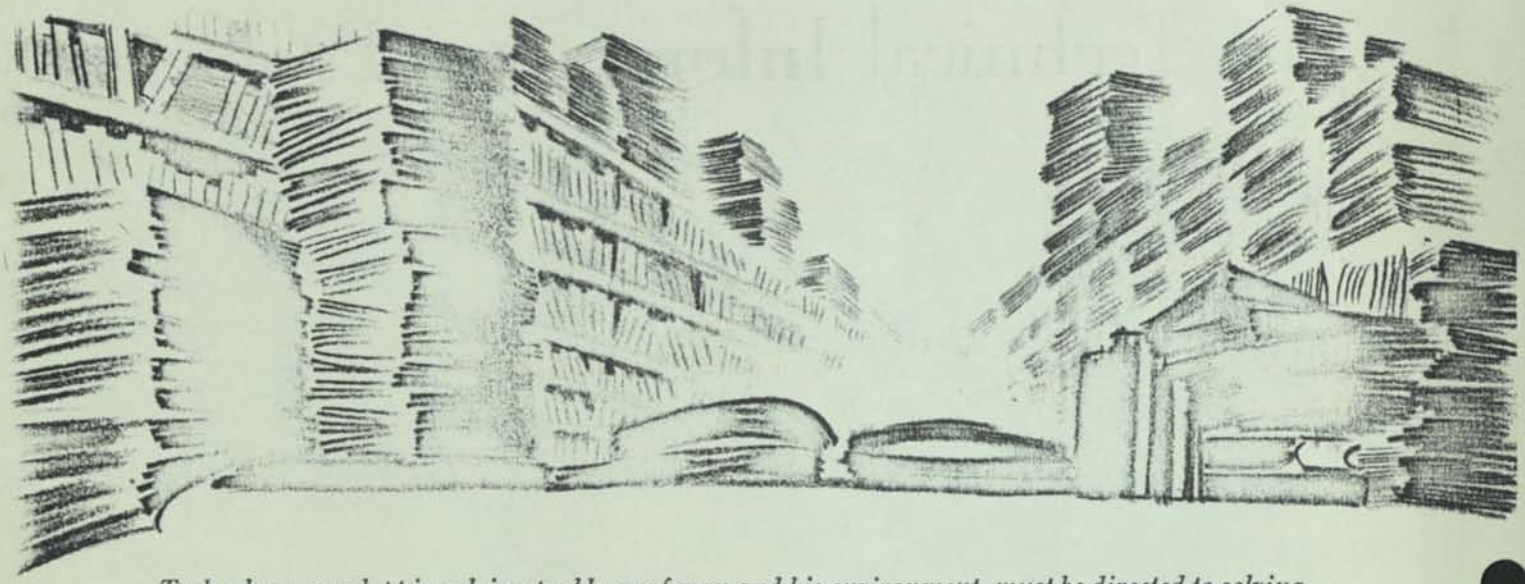
MENLO PARK, CALIFORNIA

Reprinted from *SRI Journal*, No. 1, 1958

Facets of

The Technical Information Problem

Charles P. Bourne and Douglas C. Engelbart



Technology, so adept in solving problems of man and his environment, must be directed to solving a gargantuan problem of its own creation. A mass of technical information has been accumulated and at a rate that has far outstripped means for making it available to those working in science and engineering. But first, the many concepts that must be considered in fashioning such a system and the needs to be served by it must be appraised. The complexities surrounding any approach to an integrated technical information system are suggested by the questions given here.

RECENT world events have catapulted the problem of the presently unmanageable mass of technical information from one that *should* be solved to one that *must* be solved. The question is receiving serious and thoughtful consideration in many places in government, industry, and in the scientific and technical community.

One of the most obvious characteristics of the situation is its complexity. A solution to the problem must serve a diversity of users ranging from academic scientists engaged in fundamental investigations to industrial and governmental executives faced with management decisions that must be based on technical considerations. The solution must accommodate an almost overwhelming quantity of technical and scientific information publicly available in many forms through many kinds of media and in many languages.

Some students of the problem, including men with many years' experience in various aspects of information handling, have viewed this complexity and concluded that the problem cannot be solved in its entirety. These authorities have recommended a piecemeal attack on components of the problem.

Stanford Research Institute believes that the techniques of systems analysis coupled with an understanding of the potentials of machines permit a powerful approach to the solution of this many-faceted problem. In fact, it may very well be that only by grappling with the problem as a single, integrated system can a realistic and lasting solution be attained.

However, to deal with the information system as a whole, it is necessary first to define its complexities with as great detail as possible. As an aid to the preliminary mapping of the system, a study group at SRI polled a portion of the Institute's own professional staff of engineers and scientists for questions they believe must be answered before an effective system can be designed. A representative list of the questions raised in this fashion is given in this article.

The list is impressive, but obviously not exhaustive. It

does confirm the multiplicity of points of view that must be appreciated before this problem can be attacked.

Many of the questions require simple factual answers (see "Data Needed About Information Sources and Services," p. 5). They can be answered by straightforward techniques of counting, surveying, sampling, and estimating. A few of the answers are already available, but the fact that most questions of this type cannot be answered from available sources emphasizes the pressing need for a much better quantitative assessment of the size and nature of the information problem before a rational attempt to solve it can be undertaken.

Another group of questions involves essentially matters of national and scientific policy that ultimately must be answered arbitrarily. Data and analysis can give guidance to the answers but the ultimate decision will be based on judgment of relative needs and relative values.

Questions Relating to Policy

What are the specific aims of the program?

Will the system start with only new information? Or will it process back literature, and, if so, how far back?

Will the Service process requests from allied countries? To what extent? Will it coordinate with the Soviet Union?

Can part of the operations be done abroad? What about translation?

Will an international classification, indexing, or retrieval system be adopted or promoted?

Will the system be designed to serve the brilliant, the sophisticated, as well as the more unsophisticated?

Will the Service be financially self-supporting?

Will big business have any better access than small businesses or individuals?

Would a private citizen or scholar afford to use the Service?

How will prices be established for the Service?

What is the range of subject matter to be included?

Will classified information be included?

Will safeguards be established to insure that classified information is kept under proper control?

What type of information should be included? Books (texts, tables)? Technical and trade journals? Conference proceedings and papers presented but not published? Industrial and government interim and final project reports, etc.? Operation and instruction manuals? Patents? Manufacturers' catalogs? Newspapers and general magazines?

Who will be responsible for selecting the material to be included?

What protection will be provided users who want their queries to remain confidential?

Should service be provided outside the technical community? To congressmen? Executives? Businessmen? High-school students?

Who will control the policy in the matter of designing, establishing, and/or operating the Service? An appointed committee, such as for the NACA? A civil servant? A political appointee? A committee elected by scientific organizations?

A Proposal for a National Technical Information Service

Members of Stanford Research Institute have long given thought to the increasing disparity between the accumulation of new knowledge and the means for organizing it for widespread utility. With this problem brought into sharp focus by recent events on the international scene, the Institute believed it appropriate to formalize its views on the magnitude of the problem and to suggest a possible solution. In January, a draft program for a National Technical Information Service was prepared and copies distributed to members of the President's staff, to selected members of Congress, to various agencies within the federal establishment, and to industrial leaders and technical societies, all known to be concerned over the state of technical information affairs.

This document describes a program to solve the nation's technical information problem through the establishment of a national service for the collection, processing, storing, retrieval, and dissemination of scientific and technical information from both foreign and domestic sources. The program comprises five phases, interrelated and partially concurrent:

1—Establish a central organizing and administering, federally constituted Agency.

2—Determine the gross dimensions of the problem.

3—Establish an interim information center using existing services and techniques.

4—Analyze the factors that determine the design and operation of an ultimate National Technical Information Service.

5—Encourage present and initiate additional research and engineering development programs leading to systems and equipment necessary to implement the ultimate National Technical Information Service.

This proposal, and others, for solution of the problem are currently under study by the interested bodies of the nation. Meanwhile, at the Institute study of various phases of the technical information problem, both in the gross, and of specialized aspects of data handling, storage, and retrieval, is continuing.

Would it be feasible to establish legal authority to speed up the standardization and coordination of existing facilities (such as the F.C.C.)?

Who is competent to design, establish, and/or operate the System? Would this be a civil-service organization?

Could the objectives of the Service be achieved by expanding existing government agencies (e.g. Bureau of Standards, the Library of Congress, Armed Services Technical Information Agency)?

If the Service were not directed by some existing government agency, would it not be best handled by some university?

Would it be economically feasible for any sort of commercial enterprise or non-profit corporation organized by the professional community, or by private industry, to establish and run a Service which would assure continued social and technical progress?

If we must look to the federal government for support, what residual responsibilities remain with the professional societies? Should private groups continue to sponsor special collections?

What economic and political limiting factors exist with respect to the freedom one would have in utilizing or changing those organizations already active in the documentation field, and whose existence could be over-shadowed by a national Service?

What about copyrights? Would royalties be forthcoming to the owner of the copyright if the Service distributes the material? What will be the impact on the technical publishing industry?

Should the Service act as a publisher for collections of papers (reprints) in very new and special fields?

How will the priority schedules be fixed for the Service?

How soon could the Service be initiated? With an immediate manual system? With an ultimate mechanized system?

What factors will determine the location? Can strategic dispersal considerations influence the location without adversely affecting efficiency?

Is the proposed Service simply an attempt to copy Russia?

CHARLES P. BOURNE and DOUGLAS C. ENGELBART are research engineers at Stanford Research Institute's computer laboratory. Mr. Bourne gained his first electronics experience in USN schools from 1950-51. From 1952 to 1953 he served as instructor of various aspects of guided missile operation and maintenance with Convair Guided Missile Division and as adult education instructor in electronics at Chaffy Junior College. After receiving his BS degree from the University of California in 1957, he was employed as a research engineer at SRI where he has been engaged in research on mechanization of information retrieval and logical design.

Dr. Engelbart received his BS degree in electrical engineering at Oregon State College in 1948, MEE in 1953, and PhD in 1955 at the University of California. His theses were concerned with design and programming of drum-type computers and special gas-discharge tubes for use in computers. He has worked as professor of electrical engineering at the University of California, as electrical engineer at Ames Aeronautical Laboratories, and as consultant. In October 1957 he joined the SRI staff. Information retrieval is one of his specialties.

Might not an interim solution be to translate and distribute the exhaustive Russian abstracts, thus leaving our interim energies free for other uses?

Might it not be better to reduce the amount of literature produced rather than go to the tremendous expense of providing super-service for all of it? Can a quality filter be applied to this output?

Why not allocate federal money to support more direct interchange between working scientists? Perhaps more meetings, special conventions, seminars, etc., would be more economical than better literature processing? Couldn't the money be better spent on education to achieve a given increase in scientific effectiveness?

Could a substantial portion of the information problem be solved by teaching the users more about present-day documentation techniques?

Questions Requiring Research

Some of the questions posed to the study group will require considerable study and research to produce valid answers. The research will be in many fields—in the social as well as in the natural sciences. Some of the study must be quite profound—even theoretical. Some will be more straightforward. Many of these questions must be answered before the policy decisions implied in the previous group can be made with confidence.

Can we separate apparent need, influenced by present concepts and experience, from real need? Lack of awareness of the potentialities of recently developed methods (or methods not yet developed) can easily result in an unimaginative formulation of the possibilities and opportunities for advantageously using recorded information.

How will users' habits and needs evolve as a good System becomes available?

How are the information needs of a user affected by his age, educational level, profession, type of position held, etc.?

What are the characteristic information needs of the basic (academic) scientist? The applied researcher? The engineer? The decision maker? Are they all equally critical, or is the "applier" of knowledge the one with the biggest problem?

What is the role of information retrieval, storage, etc. in the decision-making process of the research worker, engineer, scholar, administrator, etc.?

How much use does the scientist and engineer make of the facilities that are presently available?

By what processes does the scientist and engineer keep abreast of the advances in the art now? What are the relative importances of each of these processes?

How many scientists and engineers have a definite program of "keeping up with the literature"? How much time

Data Needed About Information Sources and Services

Before the designers of an overall information center can sketch in the outlines of the system problem, a large amount of data about the information input and the existing information services must be collected. Some of the kinds of essential data are suggested by the following.

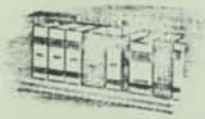
What subject fields are covered by the various journals, books, and reports? And in each case, in what depth?

What are the physical sizes of journals, books, and reports? Page size and number of pages? Frequency of publication? Kind and size of distribution? Cost or subscription price?

In what language(s) do the journals, books and reports appear?

Does each have an index? Are abstracts published, and where? Where is the information indexed?

Who, principally, are the contributors to the technical journals? Who selects or reviews papers for publication? How long, generally, between preparation and publication?



Are microfilm copies of books, journals, and reports available?

Who are the publishers of technical journals, books, and reports? Where is each located? And how long in operation?

How is each publishing operation financed?

What are the policies and objectives of the respective publishers in each field?

What fields of science and technology does each publisher operate in? In what fields does each concentrate or specialize?

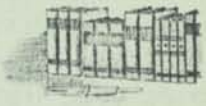
In what language(s) does each publisher produce his journal(s), books, or reports?

Could publishers of journals, books, and reports provide paper tape or other machine-readable copies of their works? At what cost?

How much has been produced to date in the various technical subject categories in journal, book, and report form? What is the physical mass of each? Are back copies available?

What libraries with technical collections, abstracting services, indexing services, and translating services are in existence? Where is each located? What is its organization? How is it financed?

What is the size and training of the staff of the various technical-information handling or processing organizations? In each case is the organization equipped to handle classified material?



In what field(s) does each information handling or processing unit operate?

What classification and indexing systems are in use?

What is the normal time between publication of a document and its appearance in the libraries? When is it abstracted? Indexed? Translated?

What are the types and numbers of scientific and technical people using libraries, and the abstracting, indexing, and translating services? In what ways does the technical community feel it is being adequately or inadequately aided by these services?

Would the various libraries and services be amenable to negotiation of changes or increase in area of coverage, or other changes of service, to fit a reasonable, overall system, if government controlled and subsidized?

What are the charges for service by libraries? Abstractors? Indexes? Translators? Which of these services are self-supporting?

Are special compilations of abstracts, bibliographies, or translations available? And for what fees? How long required to provide such special services?

would they "like to spend"? What keeps them from spending more time?

How much of the literature that would, with reasonably high probability, be useful to a scientist or engineer, is caught by him now by his own regular surveillance of the literature? How far out of his way will the average user go to be sure that he hasn't missed some possible information . . . considering the usual distracting pressures on him, his familiarity with the sources, etc.?

How many pages of literature in various categories relative to the level and interest-area of the user can we expect him to scan or search for his different information needs?

What are the relative merits of the different types of reference information services with regard to the user and his needs, desires, habits, and limitations?

What are the relative importances of the users' various informational needs? On one hand, he needs to know the newsy items such as who is working on what, what his current attack is, who disagrees with whom and basically why, etc.; and on the other hand, he also needs to be able to study in detail the carefully written treatises that may have bearing on his work. Can these different kinds of needs be met by a single system?

What are the special information requirements for different specialty fields?

Does the user, when he goes outside his special field for supporting information, want information in different

form or different levels than which he seeks in his own field? For instance, would he be looking more for "cookbook" techniques or for survey-type information?

How valuable would broad, multi-disciplinary searches be if they could be conducted effectively? How great is the problem of differences in nomenclature between fields?

What type of questions now go unanswered at the libraries?

Isn't the main problem of information retrieval one of identification—since people so seldom express satisfactorily their needs to the documentalist?

What are the major limitations in the various methods presently used in classifying and indexing scientific literature?

Is the problem that the information now is just not available at all, or is it that it is just hard to find?

Why aren't the existing services that process technical information satisfactory?

How many places does a user of each discipline have to look for index listings of a given special interest?

How can the processing of recorded information be planned so that it can be effective in spite of human limitations, or of limitations in numbers of human beings?

How much is missed by technical people leaning too heavily on librarians?

What relative gain in efficiency could be achieved by integration, merging, or better managing of existing documentation services?

What increase in efficiency of the scientist or engineer would result from improving the accessibility of recorded information?

What are the probable net benefits, short and long range, of an effective information Service to military, industrial, commercial, scholarly, government groups?

Can dollar costs be derived for reasonably well-proven delays and duplications, and can the total national loss rate due to this problem be realistically estimated? Can it be determined that the expense of delay and duplication now is greater than that of establishing and operating an information Service?

What is the lack of an information Service costing government agencies?

Can the savings in Federal money now spent on other information programs be diverted to a national information Service?

What are the relative costs and characteristics of different reproduction techniques that might be applicable to some of the dissemination and massive processing problems of an information service?

What are the techniques and costs involved in keeping up and in using large mailing lists in taking care of distribution of journals, etc.?

What are relative costs of providing the information in micro form as against making original-size photo copies?

Of the currently-operating abstracting services, how many are operating merely to satisfy an obligation of a professional society that would rather have somebody else do the abstracting?

What services does the Russian All-Union Institute really provide? What is the reaction of a Russian scientist to this information center?

How important is it to know what the rest of the world is doing?

Are any projects or areas of work reported almost exclusively in foreign literature?

What is the expected rate of growth of the system?

What are the potential information processing capabilities of existing mechanical devices?

What are the theoretical capabilities of existing or anticipated machine components which might be applied to the information processing problem?

The Soviet Approach to the Information Problem

The Soviet Union has a comprehensive technical information system in operation. In 1952 the Soviet All Union Institute of Scientific and Technical Information was established in Moscow. By 1957 the Institute had a permanent staff of 2300 translators, abstractors, and publishers. This staff is supplemented by more than 20,000 cooperating professional scientists and engineers throughout the U.S.S.R. who act as part-time translators and abstractors in their specialized fields. The Institute publishes 13 "abstract journals" which annually contain over 400,000 abstracts of technical articles from more than 10,000 journals originating in about 80 countries. It systematically translates, indexes, and abstracts about 1400 of the 1800 scientific journals published in the United States.

To reduce the time between the initial appearance of the more important information in any of the world's journals and its reaching the hands of Soviet scientists and engineers through the normal route of the abstract journals, "Express Information Journals" are also printed. These carry summary information on foreign technological developments within two or three weeks after their receipt. The work done is reported to be not only comprehensive but also of high quality.

The Institute provides numerous other technical information services, such as provision of bibliographies, micro and full size copies of original printed material, technical dictionaries, and foreign-language dictionaries.

The Institute maintains an extensive program aimed to introduce machine methods to information handling. This includes translating machines, and mechanisms for codifying, storing, and retrieving technical information. Significant progress by the Institute towards information mechanization methods and systems is reported.

How often will the system presumably be searched? How definitive will the search have to be? What volume of information should a search produce? How fast should the system respond?

Characteristics of the Information Service

As increasing data become available it will become possible to consider some of the last group of questions—those dealing with the desired or necessary operating characteristics of a comprehensive technical-information processing system. Certainly, the first system implemented would be of an interim nature using existing resources, which unfortunately employ largely manual techniques. However, ultimately it is inevitable, in view of the impressive advances made almost daily in information processing techniques, that a highly mechanized system will be possible.

How soon can an interim system be functioning?

How much can be done just by concentrating on abstract distribution and better dissemination techniques?

Would it be feasible for the abstracting publications to use a standard format and type font, such that mats (or something similar) could easily be distributed to other interested publishers, thus saving printing expenses?

What technical societies could cooperate to publish a single journal instead of numerous splinter journals?

What about the scale of the Service? Does it have to be a big system or nothing?

Does "having a large information Service" necessarily mean the physical collection of all activities at one central location?

Would a group of smaller centers, for specific fields, be of greater utility and more tractable?

Would a collection of special libraries be more useful?

What can a national service provide that is different than what is now available? Is this to be an entirely new type of service, a real advance in the state of the art, or is it to be just more and better of the same thing?

Will the System have a finite capacity? One system might work well with a few million entries, but be hopeless with a hundred million.

As the System grows in size, will it be possible to make changes easily in the classification scheme and bring the old coding into the new scheme?

If a private consultant, with "need to know" established, were to work on a government project, how would he locate and procure pertinent classified material?

Will financial filtering of requests by a uniform fee structure be desirable or effective, or would it be necessary to make non-uniform fee structure so that there is essentially some "priority" given?

What means can be used to pry loose useful information that customarily doesn't get into the published technical information channels?

Will the service include a positive program to declassify material under security restrictions?

What is an acceptable delay in getting information entered into this system?

Will all material in the subject fields be included or will there be an editor or a censor?

Will an attempt be made to standardize the form of the material before it gets into the center? Does the material have to be on standard-size sheets or forms?

What happens when the system becomes overloaded? Should service to users just be late, or should the service just be less complete?

How can we protect against freezing the specifications until enough systems work has been done to make clear what would be optimal?

Will the policy makers make sure that the final methods chosen for a retrieval system are not influenced too heavily by the requirement of compatibility with past systems?

Will abstractions be done? What kind? Descriptive? Critical? Informative? How can we get good-quality abstracts? Should the Service use volunteer abstractors directly or a staff of full-time abstractors? Or should it allow the various technical societies to organize their own volunteer abstracting services?

Will any effort be made to review old documents, and to remove or recode when necessary?

Is a standard (or artificial) vocabulary necessary? How much work will be required to design and institute such a vocabulary?

What techniques and devices can reasonably be developed and applied for facilitating such immediate requirements as printing, reproducing, storing, microfilming, billing, communicating, etc.?

What kind of a data-processing system will the Service need just to keep track of its operation?

Would the information Service keep a collection of the original documents?

What special precautions must be taken to store primary records? Would a duplicate file and collection be maintained to prevent disruption of service due to fires, or other catastrophes? How much would this cost?

What is the useful life of various forms of records? In use? In storage?

What will the information Service physically provide in response to information requests?

Will the output be in English, or a code that must be translated?

Will microform copies be acceptable to the users? If not, what improvements need be made in order to gain user acceptance?

Will the information Service output be in a form that the researcher can determine which of the documents are in a locally accessible collection?

Will the system give answers (e.g., "yes," "no," "5,000 tons in 1945," etc.) as well as references?

Why not periodically publish inventories of research in progress, to indicate what research projects are currently being undertaken in each specialty field, thus helping to eliminate duplication?

Will there be a "special communication network" in which workers in the various specialized fields can easily circulate working papers or "think pieces?" A central agency could maintain printing, listing, (in appropriate subject-interest categories), and mailing facilities for this sort of service.

Will the information Service be able to retain a file of questions to be asked of all new input material, thus providing up-to-the-minute data for standing questions?

Will it be possible to stimulate more writing of "review-the-literature" papers by qualified people in the various fields, in order to provide guides for other workers?

Can a partial search be made? (For example, can 1/10 of the file be searched and the results checked to determine if further searching is justified?)

Could the information Service operate on a "just search ½ the file for me; I don't need a comprehensive search" basis?

What kind of communications network will be needed for the operation of the interim information Service? Will it be accessible to anyone by telephone or other direct device, such that the searcher can interrogate the file directly and at will?

Would the Service be available for browsing?

What technical-manpower drain would the proposed information Service program have on other high-priority scientific programs?

What professional and educational background is needed for the personnel to operate the Service?

Could university science students be used part time and during summers to help with the various processing tasks, as a means of alleviating the shortage of people with adequate technical backgrounds?

Will there be special training for abstractors and translators or for documentation and information specialists, etc.?

How much research is needed? What research budget is reasonable?

If an information Service were established, how soon could present partial services by government agencies be terminated and funds diverted to the Service? Could some special activities in industrial libraries be eliminated?

These questions, by the very nature of their origin, are random and fragmentary. Even the full list from which they have been selected is far from comprehensive. However, we have found them a helpful stimulus as well as a disciplinary aid in viewing the technical-information problem in its broadest dimensions. We hope that others interested in this problem will be similarly served.

June 19, 1958

SURVEY OF TECHNICAL COMMUNICATIONS

I Introduction

The United States' military and economic future is critically dependent upon work done by its more than 250,000 scientists and engineers engaged in research and development. These people -- drawing salaries totalling several billion dollars a year -- exert a pivotal influence on government and private industry expenditures measured in tens of billions of dollars per year.

The almost explosive growth of the applied research and development programs after the war has caused the facilities for communicating the results of this very large effort to be outdated. Despite the expansion in library facilities and abstracting services and the introduction of systematic translation facilities, the gap between facilities and demand is widening every day.

The problems of technical communication are most acutely felt in applied research and development. The fundamental sciences are by their nature oriented towards quick and effective communication of results. A very high percentage of these results are published in scientific journals. The results are easily classified within the framework of scientific knowledge. Personal contacts between individual scientists are frequent. By contrast, the results of applied research and development are communicated principally in reports covering the specific areas of application.

Central to the communication of technical information is the problem of how best to extract from the accumulated documentary and published litera-

ture that information which is germane to a particular technical effort. Often it is not known in advance to what extent the pertinent information is available or even exists. Furthermore, the desired information is rarely well defined. A search based purely on conventional library indices tends to defeat itself either by turning up so much irrelevant material that nothing is gained or by being so selective that important material is excluded by the search. An alternative approach depends on a discriminating use of the bibliographies which form a part of nearly all scientific papers. This method has the essential advantage of incorporating expert judgement in the search. However, bibliographies are rarely appended to the document literature. Both the index search and the bibliographic-chain search, which are so useful in the fundamental sciences, are therefore very ineffective in the applied and developmental areas.

The inclusion of foreign technical literature in the picture further complicates the problem of information search. Such material in foreign languages is now receiving increased attention. Search procedures within this body of literature may be hampered by the fact that bibliographic trails flow through untranslated articles. The volume of translated articles must grow as journals from other countries are included and this will eventually complicate the search process.

The main difficulties encountered in the communication of results within the applied research and development area may be summarized as follows:

1. The volume of documentation is so large that a search for information often becomes so lengthy as to lose its value.

2. The results are not easily classificable on the basis of immanent

ideas; the results are usually presented only with reference to the specific rather than the general area of application.

3. The novel results of applied research and development are often buried in large volumes of reporting pertinent only to particular applications.

4. Document literature, by and large, carries no bibliographies.

5. The search for information is often impeded or made impossible by artificial barriers imposed by military and commercial secrecy. As a consequence personal contacts among technical people are constrained.

6. The results of research and development in other countries are not readily available in the U. S.

II A Broad and Continuing Program

Within the restrictions of security, everything must be done to facilitate ready and free exchange of information among technical men working on our research and development projects. Further, if the means they use for this communication are fettered by excessive central executive control or lack adequate facilities, many technical lines of endeavor will be less productive than they might be.

The problem of providing means for technical communication that will adequately satisfy the demands of all technical men in the country is one of considerable magnitude. Above all, the problem must be tackled on a broad enough front to preclude arriving at sub-optimized solutions. Furthermore, the solution should be expected to involve a continuing effort to improve the communication system as new techniques become available and as fresh organizational possibilities open up.

Three main steps can be recognized as a logical way of working toward a solution. In the order of attention they are:

Phase 1 - An assessment in detail of the dimensions of the problem.

Phase 2 - Design and implementation of improvements that can be instituted immediately and that can be shown to be worthwhile.

Phase 3 - Longer range improvements made in planned evolutionary steps. System design and technique development programs that are appropriate at each step will be needed in this phase especially.

III Objectives of Proposed Work

This proposal is directed at accomplishing the job under phase 1 above. The assessment of the dimensions of the problem would have a twofold purpose:

1. to allow judgement to be made in quantitative terms about the degree to which the present system for technical communication is satisfying the demands of the users; and

2. to gather data that would materially assist in the planning and execution of whatever steps of improvement are needed and can be taken immediately.

Specifically, it is hoped that a detailed picture may be assembled of the full spectrum of demands of the users and a comparison may be drawn between it and the performance of the present system for technical communication. The characteristics of the present facilities would be presented in terms of the way in which they are organized, their number and magnitude, the volume and nature of the information they handle, and their resulting performance in the face of the demands as the operators of the facilities see them.

IV Method of Approach

A. Goals of Technical Communication

The first task of the project would be that of evolving in discussions with the sponsor a clear statement of the basic goals of technical communication. The initial statement of goals may well have to be augmented and clarified during the course of the study and especially during the preliminary investigation of the problem. It is most important, however, that at all times the work proceed under as definite a guide of this sort as possible.

Three main points should be settled on the basis of these goals. They are:

1. Criteria for choosing the lines of technical endeavor to be considered on the project. Such factors as the expenditures for each line of endeavor and the need for technical advances might well be among the criteria for selecting a given line of technical activity.

2. Organizational and economic boundary conditions that must be considered in any initial and early step to improve the system for technical communication.

3. Outline of the general course that improvements might take in terms of their effects on users. The survey could then more easily obtain data that would be of greatest possible use in designing such improvements.

B. Lines of Technical Endeavor

A choice of lines of technical endeavor would be made in accordance with the criteria established in conference with the sponsor so that those activities considered most important received greatest attention. Development

work in guided missile propellants, in inertial navigation systems, in high speed electronic computers might be among the lines of activity to which special importance should be given. Obviously the work in many technical fields contributes to the success of any specific endeavor. It is felt, however, that where possible, the scope of the project should be defined by those lines of technical endeavor that are in closest conformity with the goals set by the sponsor. The choice of a specific set of technical activities rather than a list of technical disciplines allows one to take account of the fact that in development efforts information may be needed in ways that do not necessarily conform well with presently established subject categorizations.

C. Preliminary Investigation

Following the determination of the subject scope of the project an investigation would be made of the past research and studies of the problems of technical communication. This would be done with an eye to three areas in particular. These are:

1. the elements of the problem on which adequate data exists or to which sufficient thought has already been given;

2. special techniques for sampling and surveying that are applicable to this problem; and

3. the correlations that have been found to exist between men in various groups within the technical population and the demands they make of the system for technical communication.

Discussion of the last point would be sought with leading men engaged in research and development in the lines of technical endeavor chosen. Contacts

would also be made with those currently carrying out research on information retrieval and with those organizations that have done work on various aspects of the problem of technical communication so the present state of research in these areas would be fully appreciated.

Exploratory meetings would be sought with those organizations that may be in a good position to gather or supply some of the data to be uncovered in the surveys. The meetings would serve the purpose of determining the nature of help that each group might be able and willing to contribute. They would also afford an opportunity of spreading knowledge about the program and of soliciting active interest in it. On every occasion possible the sympathetic understanding of the program and cooperation with it would be sought of all organizations now associated with the system of technical communication.

D. Survey Techniques

The four areas in which surveys are needed are those of the users demands, the volume of technical information handled, the organization of the system, and the performance of the system. The one that probably presents the greatest difficulty is that of the demands of the users. This is because the demands learned from the individual user can vary over all fields of knowledge and may be expressed in a wide variety of forms. Further, although the technical population that exercises the entire spectrum of demands may be known, the parts of the demand contributed by each man are known only roughly.

Many useful facts about the demands of the users may be obtained in a survey of the performance of the present facilities, but in exactly those areas where improvements are needed the results so obtained are likely to deviate most from the actual needs of the users. This means that these requirements must also be developed by a survey of the users themselves.

In all of the four areas trial surveys will be used to determine exactly what elements should be included in the total population, what sampling methods are applicable in each case, and to reveal how best to conduct the questioning of the chosen samples.

1. Survey of User Demands

Because any system for technical communication must allocate to each job a finite sum of resources in manpower, human talent, money, and facilities; and because the time of response of the system to each demand on it is a function both of these resources and of the number of similar demands that compete with one another, the magnitude of each type of demand is an essential element in the description of the demand. Further, the response of the system measured in its degree of sophistication and in time that each user experiences modulates the nature and number of demands he makes of it. For this reason the survey of users demands will seek to discover not only the nature of the elements in the spectrum of demand but also to attach to each a frequency or probability of occurrence.

As was mentioned earlier, the correlation between various segments of the technical population and the demands made of the system is not known exactly. Every attempt will be made to increase this knowledge so that, for a given sample of technical people, the contribution they make to the total demand can be ascertained with precision.

The demands of users are very complex because they reflect both their varied needs and the personal preferences of each in the search for and use of information. It is probably fairly simple to characterize these demands that correspond directly with the ways in which the present

facilities can now search easily, viz: by the titles of journal articles. It will undoubtedly be far more difficult to take into account the demands that involve searching by ideas or for solutions to problems stated, for instance, in mathematical terms. Because these searches require much more sophisticated reactions on the part of the system for technical communication, such demands are probably not being as well satisfied as the more conventional ones. An attempt will, therefore, be made to define such demands as clearly as the more orthodox ones. This is thought to be especially important because of the strong possibility that the demands of men in applied research and development often fall in just this category.

2. Input of Technical Information

The rate at which technical publication is occurring all over the world is an important parameter in the dimensions of technical communication. By survey methods and by consultation with those principally involved in publishing and storing the new data estimates will be made of this factor in the following terms:

rate of accumulation and growth of this rate
sources of origin
subject field
language of publication
printing or reproduction method
amount in machine-readable form
amount that might easily be put in machine-readable form
amount of data involving mathematical symbols,
chemical symbols, drawings, photographs, tables, etc.

percentage containing bibliographic material

percentage of bibliographies that refer to unavailable literature

percentage having good indexing, to articles, to main subsections, etc.

This survey of the volume and nature of current technical publication would be started only after a fairly good picture had been developed about the demands of the users in terms of subject matter.

3. Organization of Facilities

The survey of facilities would also be started as soon as the approximate spectrum of the users demands was known andit would be directed toward characterizing them in such terms as:

section of the technical community they purport to serve

administering body and avowed goals

length of existance and growth pattern

general subjects covered

type of service offered, i.e., translating, abstracting, special technical library, etc.

general size (perhaps by number of volumes or publications and by personnel employed)

mode of cooperation with other facilities

The object of this survey would be to develop data that might later be used chiefly in considering organizational and administrative questions pertaining to these facilities. It would, of course, also provide quantitative information about the facilities as a background for the figures on the mode of

operation and performance. Perhaps the most important product of this survey should be a statement by the operator of each facility of his avowed goals. It is against these that his performance should be measured, but it is the users demands to which it should finally be adjusted. Such a comparison may well reveal incompatibilities between the goals of some operators and the users demands.

4. Performance of Facilities

The survey of the performance of the present system for technical communication will occupy most of the time and effort of the team assigned to studying the facilities. This part of its effort will be by far the most difficult because its goal will be to describe the system performance as accurately as possible solely in quantitative terms without injecting weighting factors or value judgements. Further, it is hoped that, as the goals of some of the operators may turn out to be incompatible with the users' demands, some knowledge may be gained not only of present performance but also of that which could reasonably easily be achieved should the goals be altered.

The type of parameters that would be used to describe the performance of each facility may be illustrated by the following:

Type of facility

Amount of information available in main subject categories handled

Age, quality and completeness of information stored for subjects handled

Rate of accumulation of material in each category

Restrictions of access to information - how determined?

Distribution of material by language

Methods of storage

Methods of indexing

Methods of search employed

Frequency distribution of responses by type, by subject matter, by speed, unanswered queries, etc.

Method of generating abstracts

Method of performing translations

Means used for checking with user to assure that searches obtain information most useful to him

l. Relaxation factor applied to each type of information in making new data more readily accessible than old.

Pains will be taken to learn from each operator the shortcomings that he believes exist in his facility and how he thinks he can best improve his service.

E. Summarization

The results of the several surveys will be analyzed and summarized in convenient form for easy assimilation. From these summaries estimates will be made of the spectrum of demands of the entire population of users falling within the lines of technical endeavor chosen. Figures will be derived from the survey of facilities to indicate their total dimensions and a picture will be built up of the performance of the entire system for technical communication as it pertains to the endeavors studied.

Discussions would be held with the sponsor to determine the exact form in which the final results should be presented for easy comparison between

the system performance and the users' demands. Careful differentiation will be made between demands which the users impose on the present system and those that they might be expected to evidence if improvements were put in of the type that could be immediately implemented. At the same time, the improvements that the operators themselves believe possible will be summarized and the expected results of such changes will be analyzed.

V Time and Cost Estimates?

Should any be included?

VI Liaison and Reports

Throughout the study, close liaison will be maintained with the sponsor of the work. If the sponsor so desires, he could assign full time personnel to work with the Institute's team during the course of the research.

Throughout the course of the study, oral reports would keep the sponsor cognizant of the progress of the work. It is contemplated that reports would be submitted on the determination of significant findings were established, and, if desired, would be supplemented by written progress reports. These occasions would be used to modify or redirect the course of the work as mutually agreed.

At the conclusion of the study, a detailed written report would be submitted outlining the results of the work and the methods used in the research. If the sponsor desires, a summary report suitable for publication would be prepared.

VII Qualifications of Stanford Research Institute

Stanford Research Institute is a not-for-profit, non-endowed corporation engaged in supplying research services to industry and government in the fields of physical and biological sciences, engineering, and economics. The Institute has a staff of more than 1,300 persons of whom over 850 are professional researchers. Many of these professionals are skilled in the fields involved in the proposed research.

The fact that the Institute is engaged in technical work that ranges from basic research to applied research and development in a very wide range of fields will enable the research team on this project to bring to the work the point of view of the user.

The Institute's experience in systems design, operations research, information organization, and in the conception and development of information processing devices provides a unique perspective into the complex scope of this problem. The Institute maintains an active and expanding program of research in all phases of the information processing problem.

The capabilities of Stanford Research Institute personnel embrace a broad segment of the scientific disciplines. It is thus able to provide unique research services in complex problems of the type envisioned here. The Institute is well known for its team approach techniques in techno-economic studies.

This study would be carried out jointly by the Economics and Engineering Divisions. These two divisions of the Institute have staffs of about 600 persons including the specialists previously mentioned. Personnel from these

divisions have worked together on numerous inter-disciplinary studies, including .....

Previous studies by the Institute in the information processing field include .....

Bourne.

National Technical Information Service

Notated Outline

I Introduction

A. Roughly what is the problem?
- need to retrieve information on the basis of ideas.
- present mechanism allows one to do this in well documented fields with which the searcher is well versed.
- present system that relies heavily on a fixed classification method falls down when:

1. The searcher is not familiar with the field (across fields)
2. The literature is not well documented (untranslated foreign literature)
3. The ideas sought are diffuse and not really in shape for publication (classified reports resulting from work on government projects)

new fields or nomenclature develops — semantic shifts

Item 1 raises the question as to whether or not one should have associated with libraries persons expert in the fields in question to whom the searcher could turn for guidance. This move would tend to put the librarians in a service position only, but such a downgrading of their status might be thought as resulting from their inability to organize library searches on the basis of the association of ideas rather than according only to prearranged categories.

Item 2 can be taken care of by better indexing of foreign literature and by the provision of adequate facilities for translation.

Item 3 presents probably the greatest challenge especially if, as is believed, a very large proportion of all the scientific work now going on the world today is documented in this manner.

The overall problem is defined by the help that the scientific community can justify on an economic basis that it can profitably use, and that it can in fact secure from a National Technical Information Service. The problems that have yet to be tackled are set by the potential deficiencies of the present library system in meeting these needs.

II Main Factors to Consider in Assessing Dimensions of Problem

A. Objectives (partially affected by nature of economically and technically feasible solutions)
- basic objective of NTIS is not to connect a man with a document but with knowledge. For connection with documents the present library system probably works reasonably well. Achievement of the real objective calls for search by the association of ideas, which is done partially now through the use of bibliographies.

- are they different in any way from those of present-day libraries?
- how may they be modified by the recent accentuation of the needs of the scientific community?

B. Economic worth of an optimally effective NTIS
- value to be attached to lesser accomplishments

C. Evolution of the system
- how rapidly can improvements be brought about
- what is the most desirable time schedule for this evolution

D. Relative worth of handling various types of information
- probable cost

E. Relative worth of answering various types of questions
- worth dependent on nature of answer

F. Time of response in determining the worth of an answer

G. Responses most desirable for each class of user
- what value shall be placed on the time of each class of user
- relationship between time of response and time to obtain exact data desired from that supplied to the user by the system

H. Corrective Feedback Loops
- effect of communication between the user and the system in improving search efficiency and overall usefulness of the system
- methods of measuring system performance so as to apply corrections - small and large loops
- how rapid must the corrective measures be as the "question" environment changes when the users become aware of the improved performance of the system

I. Growth Factors
- to what factors is the "question" environment principally related?
- possible indicators - - GNP, volume of basic research, effort in applied research

J. Relaxation Factors
- how should the system response time or the probability that the answer is complete to any given degree be related to the probability of the question being asked?
- frequency of reference is undoubtedly a factor in determining the relaxation period used in relegating to less accessible storage areas

III Discussion Areas

A. What is wrong with the present information retrieval process?

B. How much help can an efficient information retrieval system be to various segments of the asking public?

C. How much is an information retrieval system dependent on trained people for its proper functioning?

D. How will the extent to which the using public is kept informed of the facilities available affect the "question" environment in which the system has to operate?

E. What proportion of the useful information is at any moment in an unpublished form? for users

F. Are there any strong factors of a psychological nature, or engendered by their education background that will influence the way in which they will use the system?

G. How can the users of an information retrieval system estimate the probability of their having got all the information they desire?

H. If methods for searching by the association of ideas are to be looked for, how will the difficulty of conflicting terminology in different fields and countries be handled?

IV Questions Needing Decisions Relating to the Proposal

A. Because of the complexity of the problem it is easy to ask questions about it and hard to provide answers to them. The reader of a document filled with questions about the problem is apt to feel confused and disquieted. Although most of these questions must in the end be answered, there is a good deal to be gained in having a firmly established jumping-off point from which to begin the investigation of a complex problem. are there any assumptions or basic principles that we agree upon?

B. Are there any basic principals on which we might expect general agreement to be found?

C. What can reasonably be expected of this project?

D. In the light of work by such people as Shaw for CIR what is an appropriate effort level and time scale for the project?

E. How much progress can be made before some of the policy questions that are central to the problem are tackled? Matters such as:

1. What material is to be included in the system
2. The relative worth of answering various types of questions
3. Organizational possibilities for the system as one of the boundary conditions within which the system must operate

must be decided, tentatively at least, before going too far in developing data on the problem.

F. How can any broad attack on the problem be made to include rather than exclude the efforts of present-day librarians without at the same time limiting the scope of the solution?

G. In what light should the need for the project, stated or implied, be put to those whose financial support is being sought, to those whose moral encouragement is needed, and to those whose cooperation is required for successful completion of the task?

H. Should this proposal be in more than one part - a large summary and the main body of the proposal or a proposal with appendices?

I. How heavily should the proposal dwell on the problem of evolving a system from the present network of libraries as opposed to the problem of a largely mechanized system of information retrieval?

J. What is our honest estimate of the time scale on which we are operating as far as mechanization is concerned?

K. Is the effective relaxation period for technical literature such that, in the time that it will take to mechanize, almost all of the useful material could be encompassed simply by starting now to assimilate only new material? (trails left by bibliographies)

L. Can we usefully tackle only one section of the problem, such as dealing with literature in engineering, physical sciences, and mathematics, at the outset or will such an approach be self-defeating either for political or technical reasons?

1:30 p. m.,
May 14, 1958,
Dr. Bolljahn's Conference Room
Room No. 24, Building 404-A

SRI Working Group on NTIS - Agenda

I An appreciation of the problem.

II Questions relating to the proposal.

III User needs as found at SRI

IV Time schedule for SRI efforts.

V Contacts with other investigators.

VI Covering letters for the various prospective clients.

NAS possible clients for this proposal (Bob Bruce's proposal) Bowman Brown
CLS

Bowman - Pres NAS. testimony

NRC should set up a working comm to explore need to see that the pub is aroused.
(funds from the govt.)

might vote extra funds for NSF, who then passes to NRC, who could
(NRC is an operating arm of NAS) contract to SRI.

direct to NSF who says "we have this problem in ~~some~~ broad - for
basic researchers." a difficult customer.

O.W. - instead of just asking questions - letters formulate basic principles
or ideas, to show that we have an awareness of the problem.

AOD - just find the scope of the problem.

MK - people are concerned - too long, too expensive, not reliable, ...

Bruce's Proposal - re-write to

remove pre-amble (assume that reader is familiar w/ the problem)

show our competence

our assets - we haven't shown any broad viewpoint yet.

March 27, 1958

Field of Interest

1. Consulting for organizations which have large files of information,
2. Computer programming and feasiblity studies for machine retrieval.
3. Development of technique for machine retrieval.
4. Evaluation of characteristics of machines to determine what features are needed.
5. Market study- determine needs for special information centers, show areas of duplicated effort and areas barren of coverage.
6. Prove, disprove, or develop fundamental ideas. e.g. structure vs content correlation.

STANFORD RESEARCH INSTITUTE
Menlo Park, California

To: National Technical Information Center Committee

From: Robert D. Bruce

Subject: Proposal Outline for Phase II of "A Draft Program for a National Technical Information Center."

Attached is a proposal outline for Phase II of "A Draft Program for a National Technical Information Center."

The objective of the research as outlined in this draft is somewhat broader than that contemplated in the original program description. In addition to market surveys to establish the gross dimensions of the problem, this outline proposes considerable work on systems analysis to establish a design for operation of the interim technical information center. The research contemplated is concerned basically with the establishment of the interim center, rather than with the ultimate center.

Distribution:

NTIC Committee

Ken Eldredge
Merritt Kastens
A. J. O'Donnell
Maurice Rappaport
Gene Ritter

Other

P. J. Lovewell
Bill Platt
R. W. Smith
Charles Bourne
Douglas Englebart
Mark Massie

RDB/jah

Enclosure

DRAFT

PROPOSAL OUTLINE

National Technical Information Center

I. Objectives

A. The over-all objective of the study isto establish a basis for the operation of an interim technical information center.

II. Scope of the Study

A. The study would be primarily concerned with determining the requirements and characteristics of an interim technical information center. However, it would lay the groundwork for future research on the factors which determine the design and operation of an ultimate National Technical Information Center.

B. Because of the importance of scientific research in the maintenance of our defense posture, the study would in large part be concernad with the technical information required to implement current scientific research efforts for national defense. Data which would be of value in facilitating technical information dissemination for the advancement of all scientific endeavor in the United States would not be excluded.

III. Method of Approach

A. The study would require an interdisciplinary approach to a problem of critical, immediate importance. The design of a basis for operation) of an interim technical information center cannot be accomplished without an understanding of the requirements of the users of the center as well as an appreciation of the sources of information and the facilities which handle it.

1. The principal research effort would be undertaken in two phases each to be conducted simultaneously. These two phases would be:

a. A survey of the requirements of the interim center (Phase A)

(1) Ascertain the nature and magnitude of the need for and service to be rendered by the center.

(2) Determine and appraise the existing sources of information and facilities which could contribute to the interim center.

b. Outline a system for immediate implementation of the center. (Phase B)

Survey of the

B. Phase A - Requirements of the Interim Center

1. Preliminary investigation

a. Thorough investigation of past and current studies in the use of technical information by scientists, total technical information output, and documentation and information gathering, storage and retrieval methods. to determine state of the art?

(1) Investigation would encompass both a thorough literature search as well as field interviews with leading authorities and organizations with current projects in documentation and information retrieval.

(2) Review and analysis of material gathered so as not to duplicate past efforts and bring significant material to bear on current problems.

2. Study of the need for and services to be rendered by the center.

a. The first step in this part of the study would be to provide information so as to determine the scope of the subject matter to be included in the interim center.

(1) Basic criteria would be for the center to handle that information most essential to scientific research for the defense effort, consistent with the limitations of time and money.

(2) In determining the subject matter to be included in the center the following methods of analysis would be used:

(a) An analysis of the location of current research and development expenditures, both by the federal government and by industry, by type of project. The importance of each of the disciplines in the national defense research effort could then be appraised on the basis of dollar investment. this is harder to research

(b) A similar analysis would be undertaken of planned research and development expenditures to 1960.

(c) Estimates would be made of probable future (beyond 1960) areas of research.

(d) Investigation of research areas where critical breakthroughs are needed for significant future development such as nuclear shielding, thermal barriers, and the like.

(e) These investigations would be accomplished through field surveys of personnel in government and industrial laboratories, leading scientists, members of the National Science Foundation and key personnel in the Department of Defense and Armed Forces.

(3) Consideration would be given to inclusion of complete information in a particular discipline or selected information of critical importance.

(4) The investigation would be followed by a review and analysis to establish the priority of importance of information to be handled by the center. Recommendations would be presented to the central Agency for final decision.

b. The next step in this part of the study would be to determine user requirements for scientific information in the subject matter areas selected.

(1) Field surveys would be made of research personnel and administrators in the selected disciplines to determine their needs and requirements for information. Factors to be determined would include but not be limited to:

(a) Current information gathering and use habits.
(b) Current availability of information in each discipline.
(c) Differences in information use practices and needs of pure and applied researchers.
(d) The volume of information, literature searches, etc., used and potentially needed.
(e) The age and level of information needed.
(f) Current costs of obtaining information.
(g) Opinions as to the value, potential use of, and dissemination of information from a national technical information center.

(2) A similar survey, although smaller in scope, would be conducted among librarians and documentalists of scientific research libraries and information centers on the use of information by researchers, methods of information dissemination, volume of use, and other characteristics.

(3) The information gathered from the surveys would be analyzed and interpreted to determine a framework of the final product of the information centers. Particular emphasis would be given to:

(a) Total volume and frequency of inquiries that might be anticipated, both now and in the future.

(b) Classification and indexing requirements for the most effective use of the center by researchers in the various disciplines.

(c) The most convenient form for location and dissemination of information.

(d) The age and level of sophistication of the information to be included.

(e) The potential revenue from fees paid to the center for its services.

2. Study to determine and appraise the existing sources of information and services and facilities which could contribute to the interim center.

a. An investigation would be made of all sources of information in the selected disciplines to be included in the center.

(1) Purpose of the investigation would be to estimate the total volume of information which would have to be processed by the center.

(2) The investigation would include but not be limited to the following sources of information: (advanced textbooks and monographs, research journals, trade publications, review publications, conference papers, handbooks, encyclopedias, mathematical and physical tables, theses, unclassified research reports, classified research reports, patents, standards and codes, supply catalogs.)

(3) The investigation would cover both domestic and foreign sources of information with emphasis particularly on British, German, French, Russian, and Italian documents.

b. An audit of existing information dissemination facilities would be made.

(1) This would include abstracting and indexing service, translating services, technical and professional societies, government agencies, and specialized library services. Centers of specialized knowledge which do or could act as question answering services would be determined.

(2) The technical facilities, coverage and depth of effort, current costs and fees, methods and form of output would be assessed from the point of view of incorporation of their output in the interim center.

(3) An analysis of existing facilities and services would serve to point up areas of duplication of effort and areas which are currently not serviced by such facilities.

c. A comparison of user needs for information in terms of subject matter, ease of location and dissemination with the audit of facilities and resources will delineate those segments for which the interim center will have to provide facilities. This may include additional abstracting and indexing services, translating services and other procedures.

(1) A study would be made of personnel available to staff the center. This would include specialists in documentation, information retrieval, translation abstracting, indexing, etc.

C. Planning and Design of the Interim Technical Information Center

1. a. Examine the existing systems of collection, retrieval and dissemination of technical information to determine their suitability for the interim center application. b. Some criteria must be established in order to compare the effectiveness of each of these systems. This information will be obtained by examining the merits and criticisms of various systems in those cases where this information has been published, and by interviewing the responsible individuals in a representative number of the existing large technical libraries or information centers. Consultations and group discussions among the people who are competent in this field will also be used to obtain information for this evaluation.

2. The information obtained from the critical study of existing systems would be used to suggest what sort of system should be used for the interim center and the ultimate center. The chosen system must be flexible enough to be altered as the system design matures. Modifications of the chosen existing system would probably have to be made at this point to make it more suitable for this particular application.

a. Decisions will be made as to what areas and level of subject matter will be included, the classification or coding scheme to be used, the method for handling inquiries, the form of the final product and the method of dissemination, the type of storage to be used, the procedures for determining priority and fixing the fees, the degree of translating service to be provided and how this may be implemented.

3. The proposed NTIC design would then be subjected to simulation techniques, theoretical trial operations, and objective examination by technically competent groups to uncover any weaknesses or unfavorable features which the system may encompass. This trial and examination period will reveal some areas where the initial design can be modified to obtain a more optimum performance. This period will also indicate areas where more research effort should be applied. When these modifications are incorporated, the new proposed system could undergo another trial and examination period if necessary to further optimize its performance.

4. At the point where the governing agency decides that the planning and design of the NTIC has proceeded far enough to guarantee good performance, the existing plans will be considered to be the final design of the NTIC and the specifications for all the necessary equipment, personnel, money and facilities may be written.

D. Administration of the Study

1. Staffing

a. The study would be under the direction of a project manager who would have over-all responsibility for coordination of the study.

b. Project leaders would be in charge of Phase A, determining the requirements of the center, and Phase B, establishing a design for operation of the center.

c. The interdisciplinary approach envisioned would use specialists in the fields of systems analysis, documentation, data processing, market analysis, economics, psychology, and sociology. The staff would be composed of Stanford Research Institute's personnel who have extensive experience in interdisciplinary (techno-economic) studies, and would be supplemented by outside consultants and experts.

this is not synonymous w/ inter-disc. studies →

↖ To what degree?

2. Progress of the Study

a. As the study progresses, information obtained would be disseminated among all staff members so as to coordinate and channel efforts in the direction of greatest potential benefits. Modifications, deletions or additions to the work may be necessary as the study progresses. This would occur only with approval of the governing Agency.

b. To supplement and expedite the work, segments of the study may be subcontracted to qualified research organizations. However, over-all responsibility would remain with Stanford Research Institute.

could be stronger →

c. Throughout the study, close liaison will be maintained with the governing Agency and progress reports will be submitted on the determination of significant findings.

1-24-1958

Charlie:

Very interesting! and I think not without merit. It seems to me that such a pragmatic patchwork approach tends to be more likely to succeed than SRI's "let's start all over again" approach. From the standpoint of practical politics, Are the President, Congress and the Public so deeply concerned about this problem to provide the energies to start all over again? Can they be convinced it is necessary to start all over again? I doubt a positive answer to

Western Reserve University
Cleveland, Ohio

Charlie:

Very interesting! and I think not without merit. It seems to me that such a pragmatic patchwork approach tends to be more likely to succeed than SRI's "let's start all over again" approach. From the standpoint of practical politics, Are the President, Congress and the Public so deeply concerned about this problem to provide the energies to start all over again? Can they be convinced it is necessary to start all over again? I doubt a positive answer to

either question. In other words SRI's program is ideal, but is it practical?

A PLAN

for the Creation of a National Center

for the Coordination of

Scientific and Technical Information

To be presented at

a Special Invitational Meeting
of the
Council on Documentation Research

February 3 - 4, 1958

Western Reserve University
Cleveland, Ohio

F O R E W O R D

This working paper has been prepared at Western Reserve University by the School of Library Science and its Center for Documentation and Communication Research, for advance distribution to registrants for a special invitational meeting of the Council on Documentation Research to be held February 3 and 4, 1958.

The special session has been called in the hope that a concrete and workable plan can be developed by representatives of professional societies, government, education, and industry for the coordination of the world's scientific and technical information.

The plan here proposed is the product of only a few minds. It is set forth not as a finished product, but as a point of departure from which we can begin the formulation of a better program. We therefore commend it to you not for what it is, but for what it may become.

Jesse H. Shera
Executive Secretary
Council on Documentation Research

School of Library Science
Western Reserve University
Cleveland, Ohio

January 15, 1958

THE CHALLENGE TO AMERICAN SCIENCE

The February 3-4 meeting of the Council on Documentation Research and its invited guests reflects the sense of urgency which we all share. We must accept the obvious truth that both fundamental and applied research, as well as technological development and production, are becoming increasingly important in the new science-based world into which mankind is emerging; research power will, in the future, increasingly determine economic and political, as well as military, power. Unfortunately, however, we must also acknowledge the fact that in the promotion of research, the "decline of the West" has begun. Alexander King, writing editorially in Science, asserts that Western Europe's share of the world's total research output has fallen from the pre-war figure of approximately 70 per cent to less than 40 per cent today. But serious as this situation is, our primary concern is with the position of the United States in the promotion of research, which we know has seriously deteriorated in the past five years.

The Gaither and Rockefeller reports seem to agree that the United States is facing the most serious crisis in its history. The reasons for our scientific lag are doubtless many and complex, but we must admit that an obsolete and inefficient system for the dissemination and utilization of scientific information must bear its share of responsibility. So bankrupt and outmoded are our techniques for the organization of scientific literature that a vice president in charge of research for one of the country's largest corporations stated recently that "if a research task costs less than one hundred thousand dollars, it is cheaper to do it than to find out if it has been done before and reported in the literature." Admittedly, no body of scientific literature, however well organized, can create scientists where none exist; but a good information system can save the time of competent scientists for laboratory research and can mean the difference between superiority and mediocrity in national scientific achievement.

The Soviet Institute of Scientific and Technical Information

We are only now beginning to realize the impact upon scientific progress that the total mobilization of a nation's resources can achieve. Not until we saw, high in the night sky, the tangible evidence of Russian scientific achievement did our nation become aware of the potentially disastrous consequences of anti-intellectualism and the strangulation of research. The total mobilization of Russian manpower in the service of science, and their exploitation of every possible resource for scientific ends, is doubtless unparalleled in the history of the world. Certainly we, who

for so long have assumed American superiority in librarianship and documentation, have been given ample notice that we can no longer afford conservatism and complacency. For the past two years some of us have been urging greater professional attention to the work of the Soviet All-Union Institute of Scientific and Technical Information. We have talked to deaf ears of a bibliographic system in which a trained staff of two thousand produced some 360,000 abstracts in 1956 (abstracts published in a series of journals covering the physical, biological, and earth sciences); carries out a continuing research program in scientific information; is making considerable progress in the use of mechanized techniques for organizing bibliographic information; prepares reviews of scientific progress; and translates an imposing number of journals into the Russian language. And Robert Wallace, writing "First Hard Facts on All Russian Sciences" for LIFE (December 16, 1957), quotes a British zoologist who visited Moscow last year as saying that the "really shattering thing" he saw was this Institute of Scientific and Technical Information, where 8,000 scientific journals from all parts of the world are processed by a staff of senior scientists, and the resulting indexes published in an annual set of volumes larger than the Encyclopedia Britannica.

Planning for Action

But we shall not be meeting just to wring our hands. In our own country, during the past decade, many conferences have been held (some of them in Cleveland) at which scores of competent people have discussed at length a variety of new techniques for the organization and dissemination of recorded information. As a result, we have a reasonably adequate picture of the present state of existing methods available for improving the accessibility of scientific information. The reploughing of this ground would be of little value. The future, therefore, must bring a two-fold program:

1. An effective course of action which will make possible the fullest utilization of existing knowledge and techniques; and

2. An active and well-supported program of fundamental research which will ultimately result in new techniques for the effective use of recorded information.

This meeting has been called in the hope of bringing together, for fruitful discussion and constructive action, a group of people who are not only deeply concerned about the seriousness of the problem, but who are also in a position, because of the resources at their command, to take bold and courageous action. We are meeting, not at the request of a superior authority, but entirely on our own initiative and of our own volition. We represent no corporate existence, and we have no power other than the sum total of our individual resources--which, it might be added, are certainly not insubstantial. Finally, we are obligated only to our own consciences. Our only immediate objective is to create the solid foundation for a rational program of constructive action.

There may be those who will argue that progress must be made slowly, that there are dangers inherent in haste—and they are right. But we can no longer afford the luxury of protracted deliberation. Time is no longer on our side.

Criteria for an American Program

It is important to emphasize that the program which we envisage should not follow slavishly the pattern established by Moscow. A government monopoly of the dissemination of information may meet the needs of a scientist in the U. S. S. R., but it is not necessarily appropriate to the American scientist. We must work out this problem for ourselves, in our own way.

We believe that whatever program is developed must meet certain specific criteria:

1. It must serve the information needs of a wide variety of groups and interests, at both the theoretical and the operational levels.

2. Though the active leadership of the federal government is desirable, the plan must represent a truly cooperative effort of private industry and business, education, the professions, the federal government, and the foundations.

3. It must break down the barriers to a free flow of information throughout the several fields which comprise "the sciences."

4. It must be implemented by superior personnel, familiar with methods of research and trained not only in the techniques of librarianship and documentation, but also in one or more subject disciplines. To do this, it must present not only economic rewards which are attractive, but also an intellectual challenge. This is no task for hacks. At least in the recruiting of personnel we would do well to follow the Russian example.

5. Finally, the program must be adequately, even generously, financed, for the great danger is not that we will spend too much, but that we will spend too little and too late.

The budget envisaged seems astronomical, but good bibliography is always a bargain and, as a vice president of Colgate-Palmolive-Peet has said, "We shall soon be spending much of our time in either repeating experiments we didn't know existed, or searching so hard for data that we've no time for fresh experimentations. It seems that the time is right for centralized information centers from which subscribers can receive transmitted facsimile reproductions or teletyped abstracts that have been obtained electronically by scanning the literature. Perhaps abstract

services themselves should cease to exist as printed circulating publications and instead become central libraries of stored, processed, and carefully indexed information. If we could save only one per cent of the professional time at our disposal by better literature searching, it would be worth fifteen to twenty million dollars a year today, and much more in years to come. This would be a fair return on a capital investment of, say, a hundred and fifty million dollars that might be required for such facilities."

Though our situation is grave, it is certainly far from hopeless; even this modest assembly has tremendous potential if we but have the will to use it. We must, by every possible means, seek to sharpen the bibliographic tools of science. To this end we must drastically reshape our fiscal thinking just as we must re-assess our traditional library and bibliographic methods.

It is no easy task to which we invite you to address yourselves; it demands thoughtful planning, courageous and venturesome leadership, and dedication. But even these are not enough. We shall need money---a great deal of money by the traditional standards for bibliographic projects. Perhaps this should be the first serious problem for us to discuss. Certainly it is one of the most basic.

BASIS FOR THE PROPOSED PLAN

The plan proposed is directed toward achieving the greatest possible advance in information services to the scientific and engineering community of the United States consistent with present-day resources in human skills, equipment, and systems engineering, within the framework imposed by the necessity for a major improvement within a reasonable period of time.

Until quite recently it has been axiomatic that recorded knowledge was the most important single tool of the investigator. But today, because of the sheer volume of new knowledge generated by the research facilities of government agencies, universities, and industrial concerns, the investigator is faced with a problem having three dimensions of frustration:

1. The individual engineer or scientist cannot read and remember all of the world's literature that has a reasonable probability of being of later use to him.

2. The individual scientist or engineer—or his organization—cannot afford to finance the thorough processing for later retrieval of the majority of the world's literature of probable pertinent interest.

3. The individual can no longer rely on traditional library tools to deliver to him within a reasonable period of time the detailed information he needs.

Many of the major advances in science have been achieved by making use of methods and facts drawn from more than one scientific field. The application of findings in one field to problems in another becomes increasingly difficult as the continuing expansion of human knowledge makes narrow specialization an apparent necessity. Ready accessibility to knowledge in diverse fields of science and technology must be the goal.

It is now proposed that a national center for the coordination of scientific and technical information be created, which will achieve an effective balance among centralized processing of the world's published literature, cooperative activity with other processing centers, and provision of direct or indirect services to individuals, research groups, or specialized information centers.

Although various information-processing activities exist (e.g., abstracting, indexing, information retrieval, translating) in various

subject fields, there is still a considerable demand for additional services:

1. Where there is an actual gap in service;

2. Where an existing service may not be sufficiently prompt;

3. Where an existing service may not provide information in a form convenient for use;

4. Where an existing service, theoretically able to meet user needs effectively, may not make imaginative use of available communication facilities;

5. Where the mechanism of an existing service, because of the bulk and complexity of information involved, may not respond with sufficient effectiveness to user requirements.

In recent years we have seen a shift from competition between research libraries to cooperation. However, cooperation among libraries has been primarily in the area of acquisitions. The Library of Congress cooperative acquisitions program, the Midwest Inter-library Center, and the Association of Research Libraries' Farmington plan are outstanding examples. This type of activity has done much to strengthen the position of our libraries and is indicative of soumd planning in a long-range program. Nonetheless, it must be regarded as a first phase, and the time is overdue for the next logical step--the cooperative exploitation of acquired materials.

The major objectives of the proposed national center would be as follows:

1. To provide effective interdisciplinary information services;

2. To provide such information services as may be required to fill existing gaps in various fields, until such time as new agencies may be established;

3. To foster the development of new information services by existing agencies;

4. To provide advice and guidance to cooperating information centers in order to ensure economy of effort;

5. To act as a clearinghouse for information requirements that can be referred to existing agencies;

6. To foster and to conduct research and development in methods, equipment, and procedures in order that better services may be provided and costs may be reduced.

The following would be the principal services to be provided by the proposed national center:

1. To provide, from the world's scientific and technical literature:

a. Direct information, on demand or on a continuing basis, in response to specific requests from individuals or organizations;

b. An interdisciplinary literature searching service;

c. Published abstracts of material not available from existing agencies.

2. To distribute copies of the coded media prepared at the Center so that machine searching may be performed wherever facilities are available;

3. To stimulate the preparation of periodic bibliographic reviews by subject specialists;

4. To work toward complete translation of the world's scientific and technical literature by:

a. Support of existing services;

b. Translation services at the Center.

THE PROPOSED PLAN

I. Phasing

A fully operational program for the national science information coordination center could be achieved in phases during the next five years. However, partial operation might be initiated more quickly, with many services available in nine to twelve months. A proposed development schedule follows:

Phase 1: First two months.

Headquarters staff organized and working space acquired. (An initial organization chart is given as Figure 1.)

Phase 2: Third and fourth months.

Initial contacts, negotiations, and acquisitions operations:

- Potential clientele contacted through a series of national conferences bringing together the various levels of users of scientific and engineering information in government, education, industry; e.g.,

  research management

  research personnel

  information specialists

  librarians.

  At these conferences the projected services would be explained, and the opportunity given to the potential users to shape the development of the Center.

- Arrangements made with professional societies and other groups considered competent to perform the extra services required by the Center.

- Operating procedures prepared for the information-processing and exploitation functions of the Center.

- Previously processed information (especially abstracted or indexed material) identified. (Examples of abstracting services reported in a recent survey are listed in Appendix A.) Translating services also identified.

Phase 3: Fifth and sixth months.

First-year plans completed for division of processing responsibilities between the Coordination Center and other organizations. Any required subsidy negotiated.

Personnel hired and training started for operations to be performed at the Center.

Orders placed for automatic equipment required for reproduction, microfilming, searching, translation, transmission, etc.

Phase 4: Seventh to twelfth months.

Pilot operations initiated for the processing, reprocessing, and exploitation of recorded knowledge--leading to full-scale operation as soon as possible.

Phase 5: Operation.

Continuing collection and definition of terminology to support searching and translating operations, particularly those employing automatic aids.

II. Systems for Processing Information

In considering the systems of information processing and retrieval to be employed at the proposed Coordination Center, it is important to recall its projected services:

1. To provide, from the world's scientific and technical literature:

   a. Direct information, on demand or on a continuing basis, in response to specific requests from individuals or organizations;

   b. An interdisciplinary literature searching service;

   c. Published abstracts of material not available from existing agencies.

2. To distribute copies of the coded media prepared at the Center so that machine searching may be performed wherever facilities are available;

3. To stimulate the preparation of periodic bibliographic reviews by subject specialists;

4. To work toward complete translation of the world's scientific and technical literature by support of existing services and translation services at the Center.

Since it is to be expected that demands which can be fulfilled by traditional service will tend to be handled by the currently available tools and will be referred elsewhere when they come to the attention of the proposed Coordination Center, the interdisciplinary type of inquiry will be the Center's chief concern. In servicing this type of inquiry, the traditional classification and indexing systems tend to prove inadequate, even when dealing with relatively small document collections. Therefore, for the necessary searches, retrieval and correlating systems of the highest sophistication are required, to provide the required detailed penetration into the subject matter of the documents, as well as extreme flexibility for controlling the retrieval of unwanted material.[1]* With interdisciplinary searches, it is to be expected that useful information may be found in almost any document in the file. Therefore, a total search of each document in the file, or a "fail-safe" method of searching for any document of potential importance to any given search, is a basic requirement for the Coordination Center. Although initial operations of the Center would (of necessity) depend upon presently available processing, a highly sophisticated coordinated retrieval system would be phased into operation at the earliest possible time.

An extensive description of a system designed to meet such needs as these has already been made available in various publications.[2] Such a system is already undergoing pilot operational test in the field of metallurgy, and experimental equipment for performing the required searches is already in pilot operation. (See Appendix B.) Plans for the construction of a high-speed version of the equipment are in an advanced state; meanwhile, equipment of various sorts is being programmed for searches. (See section III, Tools for the National Center, page 19.)

It is expected that facilities will be provided not only for answering questions, maintaining a current awareness service, etc., but also for generating special-purpose files arranged according to specialized points of view by merely running a comprehensive high-speed search of the files and preparing sub-files ready for access or further searching at distant locations. A special-purpose file might be needed if a new metal (beryllium, for example) were to become of extreme interest to the Department of Defense, so that an information center would be required to serve new users of this metal and its compounds and alloys. A comprehensive search of the total files of the world's literature available at the Coordination Center (based on previous detailed analysis and translation) would yield a specialized file on beryllium—itself ready for specialized searches.

The scientific and technical information to be encoded for machine search falls into three general classes: Documents or papers not previously processed; those previously indexed or classified; and those in foreign languages. (See Figures 2 through 7.)

For material in English not processed before, normal abstracting operations may be combined with those steps needed to prepare a "telegraphic" abstract.[3] (Figure 3.) For foreign-language material, the abstracting may be done either before or after translation. (Figures 4a and

* References will be found on page 28.

4b.) Encoding of the telegraphic abstract for machine search is then accomplished, using automatic techniques. (See Appendix C.) For material previously processed (indexed or classified), methods have already been developed for automatic encoding. (See Appendix D.)

In order to determine where opportunities exist for coordinating various types of information-processing systems, various information-analysis methods have been investigated, and a start has been made toward devising methods whereby the results of analysis accomplished in one system may be re-processed for transfer to another system. An inventory of analysis methods and a qualitative approximation to a common language for systems interchange is given in Appendix E.[4]

It is proposed that mechanical aids for the translation of languages be employed at the Center as soon as possible. (See Figure 8.)

III. Tools for the National Center

In light of the basic services planned for the Center (See page 7.), the provision of rapid service will require the installation as quickly as possible of the most versatile equipment now technologically possible. Therefore, the following tools would be provided as early in the life of the Center as possible:

1. For Searching and Correlating. In order to achieve the versatility in searching provided by the encoding system referred to in the foregoing section, equipment for searching based on the "free field" principle is required. Examples of this type of equipment are the W. R. U. Searching Selector (Appendix B) and the Minicard Selector.[5] Other equipment may be programmed to conduct such searches; e.g., the I. B. M. 700 computers, Remington Rand Univac, I. B. M. 650, etc.

2. For Encoding and Translating. In order to maintain low-cost operations, both in encoding for searching and in translating from various target languages into English, an automatic dictionary would be the primary tool. (See Figure 9.) The functions required may be performed using standard punched card equipment or digital computers. However, it seems likely that the lowest-cost operation would be achieved through the use of specialized drum memory equipment.

3. For Reproduction. In order to supply copies of full papers, modern reproduction equipment would be utilized. When available, the Minicard equipment would be considered for the convenient reproduction of files and ready preparation of copies of published papers or translations.

4. For Transmission. The U. S. postal system would be used for transmitting information which is not of sufficient urgency to require more rapid action.

Depending upon urgency requirements, there will be various opportunities for more rapid transmission, particularly of information identified by machine runs:

- Teletype: for transmitting bibliographic information and short abstracts. The bibliographic material will have been punched onto tape during the initial typing, so that transmission of selected information would require no additional keyboarding operation. The use of teletype in filling a typical information request is charted in Figure 10.

- Ultrafax: for transmittal of larger numbers of abstracts or short papers, this medium would be considered.

- Television: (during non-broadcasting hours) for transmittal of large amounts of information might be feasible with careful planning, especially for material for which the most rapid transmittal is not required.

For the transmission of large quantities of information to form sub-libraries, the use of Minicard equipment would be considered.

Before any information would be transmitted in any form, a union list of holdings would be checked to determine whether or not the material identified might be available on the premises of the customer. If this should prove to be the case, only serial-number identification would be transmitted.

IV. Financing the Program

The cost of operating the proposed Center cannot be determined precisely, because of the obvious difficulty of predicting:

1. The amount of literature to be handled;

2. The ultimate division of processing responsibilities between the Center and other information services (e.g., professional societies).

However, it is possible to prepare estimates based on, say, the present U. S. S. R. activity (an estimated 360,000 papers processed annually).

In Figure 11 is given an estimate of the first ten years of total expense (Curve A), assuming pilot operations to be conducted at the Center during its first two years. The area between Curves A and B represents anticipated subsidies to processing organizations as cooperative activity increases. The area under Curve C represents income from the Center's services, leading to a self-supporting position by the end of the ten-year period.

The capital equipment required for operation of the Center is estimated to cost $5,000,000, including the cost of searching, encoding, translating, reproduction, transmission, and support equipment.

The total financial support envisaged to bring the Coordination Center to full operation and to self-supporting status is estimated at $36,000,000:

$26,500,000 - processing

5,000,000 - capital equipment

4,500,000 - quarters and maintenance.

At the present rate of research expenditure, this cost would be approximately 0.06 per cent of the total cost of research conducted in the United States during the ten-year period.

It is anticipated that the annual income of $5,000,000 after ten years would be derived as follows:

2,000 current-awareness customers, @ $500	$1,000,000
10,000 questions (about 30 per day) for at least 1/10 of the file, @ $250	2,500,000
Sale of processed information ready for machine search	100,000
100 special files, @ $10,000	1,000,000
Special services	400,000
	$5,000,000

The cost of service is predicated on a basic searching cost of 300 references searched (on a high-speed selector) per 1¢. A more detailed tabulation of searching costs is given in the table on page 26.

V. Future Developments

In the preceding discussion, a number of problems requiring further developmental work have been only mentioned, and solutions to others have been only implied. It was not intended that these problems be glossed over. Rather, it is expected that a documentation research and development department would be an active division of the proposed Coordination Center. Examples of the problems to be considered by such a group are:

1. Study of information-processing systems, to determine ways in which various kinds of systems and processing may be coordinated. (See Appendix F.)

2. Investigation of the strategy of searching files in order to ensure optimum value from the information stored at the Center. (An example of such work is given in Appendix G.)

3. Methods of creating "fail-safe" file divisions to facilitate searching. Initial work has already indicated that division along broad code characteristics will ensure total search, while reducing significantly the size of file to be searched, on the average, per question.

4. Document retirement programs, to consider methods for retiring or discarding obsolete information from the files of the Coordination Center.

5. Continuing study and collection of terminology to anticipate translating and encoding problems involving new terminology.

6. Study of the use of notation systems in such fields as chemistry, the engineering disciplines, etc., for the convenient encoding for search of scientific information.

7. Investigation of means for continuing decentralization of operations of the Coordination Center so that the various professional societies may assume more and more of the processing and specialized service tasks for the customers in their fields.

VI. The Next Step

In order that the creation of a coordination center should proceed at a rapid pace consistent with urgent national requirements, it is proposed that two committees be organized to carry forward development plans:

1. An Executive Committee: to investigate the sources of financial support, select a core management group for the Center, and recommend a site;

2. A Plans Committee: to prepare detailed estimates of costs and personnel requirements, and to conduct preliminary meetings with the various professional societies in order to establish contact and lay the groundwork for cooperative effort.

December 30, 1957

To: Bialik, Burch, Crane, Engelbart, Melville, Miller, Rosen

From: C. Bourne

Subject: Recent Discussions on Information Retrieval

The following paragraphs are some of the written by-products of discussions which were held by members of the Computer Laboratory in critically examining the NTIC proposal. This collection is being circulated in order to keep the group members aware of the efforts that had been expended, and to provide the groundwork for future discussions. It is hoped that this collection will be augmented by additional articles.

Summary of a Group Discussion on Information Retrieval Techniques, December 16, 1957

Present were Bialik, Bourne, Burch, Crane, Engelbart, Melville, Miller, and Rosen.

The purpose of the discussion was to stimulate group and individual thinking in the field of information retrieval, with special emphasis on very large scale information retrieval systems.

We briefly summarized the various classification systems, i.e., techniques for putting "handles" on documents so as late to be able to retrieve the information contained therein. We decided that these all seemed to promise that those techniques realized in the near future would probably contain numerically coded classification terms to be associated with each stored item in the large file and that retrieval search would probably involve the search for some logical combination of a given number of coded description terms. We chose to attempt to evaluate the magnitude of the problem involved if we were to use the uniterm system on all the technical information that the control technical information agency would probably contain.

We began by assuming that there would be of the order of 10^8 items to be stored in the retrieval system per year, each item we guessed to have an average of 10 pages. We decided to allow for 10^{10} possible such items ultimately in the storage system. To identify the serial number of each item would require 33 bits as an accession number. We further allowed such a system to contain 10^5 uniterms in the central dictionary. Storage of each uniterm code would require then 17 bits. If we assume that each item in the retrieval file would be associated with an average of 17 uniterms, there would be required some 300 bits of such uniterm identification associated with each stored item. This, together with the 33 bit accession number would give a total of 333 bits, on the average, for each stored item in the retrieval file.

To orient ourselves further, we assumed that this information could be stored on glass disk photo optical system for which we knew operational storage densities of three million bits per square inch were realizable. The retrieval information for one year's accumulation of literature would then require approximately 11,000 square inches of such storage. Assuming storage disks of two feet diameter with with approximately 200 square inches of storage area in each disk, it would require about 55 disks to store a year's retrieval information. At a search rate of about 1 megacycle, it would require approximately ten minutes to scan a given disk.

It seemed evident from these figures that the retrieval system in itself could be realized with present-day techniques without undue strain and that the search time for any given list of accession numbers in accordance with some specified list of uniterms could be accomplished in a relatively short time. It was hypothesized that each disk could in itself be a search station and that a relatively large number of parallel searches could be going on at each disk station simultaneously. It was further proposed that several different types of search stations be outfitted at each disk search station. Some of these search stations could merely watch for certain uniterm combinations and provide tallies at the end for the number of coincidences upon certain specified uniterm combinations that were found in one complete search of the disk. Other stations could be outfitted so as to print out the accession numbers of those file items for which the uniterm combinations coincided with those specified in any of various categories stated by the searcher.

It was assumed that, at this point, a searcher would be provided with lists in various categories and uniterm combinations accession numbers which corresponded to the particular category. Realizing that the job of storing the actual information contained in the 10^8 total filed items in such a storage system would be too large to expect local information centers to contain, we then considered the problem of storing and finding the particular reference articles corresponding to particular accession numbers. If we were to assume that these 10^8 documents representing 10^9 pages of text were to be stored on 16 mm microfilm at perhaps thirty pages per foot, it would consume some 3×10^7 feet of film per year. It seemed reasonable to assume that once the searcher had located a list of particular accession numbers in which he was interested, that a delay of from a day to a week in obtaining the actual copies of the information would not be serious. Therefore, we assumed that we could use such as a 10 foot reel of this microfilm which could hold approximately one million linear feet of film for the storage at a given search station. At a search rate of only twelve inches per second, it would take just about 24 hours to search this complete reel. If thirty of these stations held the entire text of a year's information accumulation there would of course be some problem in parallel access to this information by a number of searchers. It seemed, however, that the task of looking simultaneously for the information required by a number of searchers during any one pass of a given large reel would not be a particularly hard technical problem. The hardest part of the problem would be to keep track of whose information was whose and of compiling later the newly exposed frames of information for each particular searcher.

The consensus of the group at this point was that the physical problem of storing and retrieving information for such a large scale information retrieval system was not particularly insurmountable even with present day techniques. A few minutes consideration regarding the problem of coding and preparing 10^8 documents per year for such a system convinced us that this part of the operations involved in establishing the large scale information center would be probably the hardest to evolve. Certainly some automation help for the people who must do this job is necessary. We realize that newer components and techniques would certainly make the job of storage and retrieval of this information probably even more feasible than as outlined above. On the other hand

though, careful study of the system requirements for this large scale information center would undoubtedly ask for considerable more flexibility in the performance of such a system than we had assumed and therefore the performance requirements upon the retrieval and storage system would undoubtedly be quite a bit higher. It is apparent, however, that we would be in a fairly good position technically to consider the construction of a very large scale information retrieval center.

Our time was up at this point and we adjourned the meeting.

D. C. Engelbart

Miscellaneous Contributions, NTIC Pre-proposal Effort

1. Copyright royalties will generate their own share of problems in the total bookkeeping which NTIC must cope with. There will either have to be a drastic revision of copyright law, or else a procedure for crediting any of a very large number of publishers' accounts for each document delivered to a customer. This will be quite a task in itself.

2. It probably should be pointed out that, large though the cost of a NTIC may be, it could save a great deal of money for the rest of the nation's (world's) documentation centers. An equitable assessment system could possibly cover the financial burden of NTIC at a net savings to all.

3. The entire system of technical-information dissemination would be subject to drastic evolutionary changes. A careful study of techniques, problems and needs, coordinated with the system evolution of NTIC, could guide these changes and do much more to facilitate dissemination than merely to provide search and compilation of articles published in the conventional manner. Perhaps intermediate publication will tend to vanish, and rapid appearance of information in the NTIC can be accomplished through an automated, fast-acting central editorial system. Published for circulation to professional-society membership may perhaps be only lists of pertinent abstracts....on newspaper-like format.

This whole area very logically seems to deserve serious research within the NTIC-development effort. (I think that some very provocative discussions along this vein have been published in the past.)

4. The automation of the document processing, putting the information into the retrieval system in the first place, will be an essential factor in economical feasibility of a NTIC. If one considers the large and constant problems of (1) finding and acquiring all the world's useful technical literature, (2) filming and storing every page for high-quality storage and reproduction availability, (3) inspecting and assigning retrieval codes to each document, and (4) integrating these data into the retrieval system, and realizes that $100,000,000 a year would allow on the order of 10¢ per page for this total job, he can appreciate the first statement.

Allowing some skilled "coder" five minutes per document would cost on the order of 5¢ per page right there. This is perhaps the only individual attention a document may recieve, but the rest of the handling will certainly have to be efficient to keep costs down.

D. C. Engelbart

Analysis of Retrieval Systems

Up to the present time, a great deal of effort has been expended in the analysis and development of coding and retrieval systems. However, this effort has been directed in many different directions with many different objectives and requirements. Several systems have been set up for particular fields of knowledge and are effective within that narrow field. If the NTIC is to operate effectively, some leadership must initially be provided to consolidate and coordinate the development of coding techniques, and to direct the group efforts to the immediate problem of developing a coding and retrieval system which would satisfy the requirements of the NTIC system.

The prime objective of any retrieval scheme is to locate documents which may be pertinent to that subject or idea which originally initiated the searcher's request for information. Aside from the technique of mechanization, the coding and retrieval processes are intimately related. Coding may be looked upon as a method of attaching some tags on the document in such a manner that a searcher may, at some later time, locate this document by the tags on the document. Every coding system in use today (Dewey Decimal System, Universal Decimal Classification, coordinate indexing, semantic structures or synthetic languages, etc.) has inherent weaknesses which limits its effectiveness in applications to very large and dynamic collections of documents. A more satisfactory method of "ordering" a large and diverse collection of technical documents for effective retrieval has to be developed.

Future coding and retrieval systems may be radically different from present schemes and may operate on such principles as the characteristics of word frequencies and probabilities, or correlation techniques. Work should begin on the investigation of such techniques.

Some techniques for the mechanization of the retrieval process are well advanced and are independent of the particular coding or classification scheme chosen for the system. This is fortunate because it allows a parallel and nearly independent development of mechanization techniques and techniques of coding and retrieval.

The urgency of the project dictates that parallel development work be initiated to (1) consolidate and coordinate present efforts in developing coding and retrieval techniques which are suitable for large and dynamic document collections, and (2) to develop and improve the mechanization of the coding and retrieval process.

C. P. Bourne

Hardware Problems

Ones that need detailed investigation with no fixed system solution assumed.

1. Scanning or Reading - If printed matter is to be accepted by the system a means of automatically reading it and converting it to an electric signal is needed. This is independent of what definition or specifications are set on the type of electric signal. Many partial solutions and attempts have been made at reading, such as MICR check reader and Solaritron.

2. Storage and Recognition - If the information is in an electric signal form similar to the teletype code, it is quite possible to store in digital memories and recognize a fixed vocabular. There are no hardware problems here, it is merely necessary to balance complexity and budget to limit the size of the memorized vocabulary to a reasonable dollar value.

3. Data Transmission - All present systems of transmitting data require that human beings monitor either the data itself at many points (not permitting passage beyond a point if distorted) or the operation of unreliable component parts and stop data flow on failure. If the system requires large amounts of data to be transmitted from point to point inside the system where human monitoring is impossible a great deal of work is required yet on reliability. It is conceivable that the system could be envisioned such that the data flow had check points that could be monitored and if need be corrected by humans.

4. Size - Bigness alone can place a strain on present technology. The curve of reliability vs. number of parts seems (at present at least) to fall very rapidly as the number of parts gets large. Almost as if there were an upper bound beyond which we could not build an operating system. More work is still required to either increase this upper bound or disprove the present shape of the curve.

Also as we increase physical size inductance and capacitive effects play a great role with just a straight piece of wire, all of which could be neglected in small size.

5. Print Out - Currently available hardware for printing out or otherwise getting the pertinent available documents into the hands of the inquirer are entirely too slow. Some new concepts would need to be developed here.

If the method of describing a given page from a document could be made less precise and still prove precise enough for the problem at hand the demands of size and complexity could be greatly relaxed. For example an assumption of 6 bits/character, 6 characters/word, and 300 words/page means storing and scanning for recognition 10^3 bits/page. Any reduction in this number decreases the size of the memory by a proportionate amount.

S. W. Miller

Machine Translation

It is now possible to effect word-for-word translations on high-speed general purpose digital computers, but translations of this type could be made more economically on special computers built specifically for that purpose. It is technically feasible to build such machines.

Word-for-word translations done by machine are very crude but can be produced much more rapidly and at a much lower cost than those done by human translators. Such rough translations probably could suffice for rapid survey purposes, while those items of interest culled during the survey could be given more thorough attention.

For more complete translations from one language to another a procedure would be needed which would allow a machine to take account of the grammatical and syntactical rules of both languages. It would probably operate on a sentence-by-sentence, or structural, basis and would produce grammatically correct translations. Although the necessary machinery is available for such a scheme, a great amount of techno-linguistic research would be needed to insure the correct transfer of meaning from one language to another.

Output printing devices are available for any language. At present the development of an input counterpart, capable of automatically reading common printed characters, is being attempted in several quarters. Successful completion of such an effort would greatly increase the value of a mechanical translating scheme of any type.

G. A. Barnard

Information Retrieval

I wish to express one concept which I believe is pertinent to the general Information Retrieval problem, which I will call the "quick look". For convenience I will use the analogy of the human brain as the central storage file, visual data as input, and pattern recognition as output. As a specific example of "quick-look" consider the human ability to scan faces at a very rapid rate, recognizing almost instantaneously a "familiar" face about which more information is available. The subsequent retrieval process, which calls out a specific identity, and supplies details on circumstances of prior encounters may take a relatively long time.

So, in information retrieval, if we could get a "relatively instantaneous" answer to the question "Does the central file contain any information on this subject?" we can call this the quick look. If the answer is "yes" we may then be very willing to wait an appreciably longer time for the information itself.

A further step in this direction, of course, is some measure of "how much" information is available as an answer to the quick-look instead of just yes-no. This would permit the inquirer to rephrase his question in more or less specific terms if the amount of information available was more or less than he wished to accept.

R. W. Melville

STANFORD RESEARCH INSTITUTE

To: Those listed below Date: 11/26/57

From: M. L. Kastens

Subject: Attached draft

Attached is a first draft of the "thinking piece" on a technical information center.

Please let me have your comments at your earliest convenience. I am interested in your reactions to the length, scope, the material included, the material omitted, and any other observations you care to make.

The piece was intentionally written to a five-page length, which, of course, precluded a detailed discussion of the problem or of the suggested program. However, it was our belief that a longer document would not be read by the type of people to whom this is primarily addressed.

You will also notice that there is no "hook" in the presentation. SRI's name does not appear at all. Presumably, this exhibit would always be transmitted with either an oral or written statement of SRI's profound interest in participating in the program outlined.

Thank you in advance for your advice.

MLK:em

Blumberg
Bourne
Carter
Clark
Ebey
Eldredge
Gibson
Hutchison
Johnson
Krause
McGuigan
McLachlan
Nielsen
Noe
O'Donnell
Platt
Ritter
Steffens
Strickland
White

A NATIONAL TECHNICAL INFORMATION CENTER

Introduction

The apparent ability of the USSR to effect technological innovations, particularly in the military area, with significantly shorter "lead" time than the United States has concerned our strategists, intelligence experts, and technical men since the end of World War II. The superiority of the Soviet performance in this respect was implicit in the development time-tables of the intercontinental bomber, the jet transport, the fission bomb, and other instances. It was demonstrated with disconcerting finality in the cases of the thermonuclear weapons; long-range ballistic missiles; and, most recently, satellite development. Todate there has been no satisfactory explanation for the difference in development time-tables in the USSR and the U.S.

Without detracting from what has recently been emphasized with great force and pertinence regarding the efficacy of the Soviet educational system in producing technically-trained personnel, it is most improbable that during the past ten years the Soviets have had either more or better scientists and engineers than the United States. Unless we are prepared to grant to the authoritarian, hyper-disciplined approach of the Soviets a major inherent advantage over the democratic, individual-initiative management philosophy which prevails in this country, it is impossible to attribute the speed of Soviet developments simply to superior administration. Similarly, it would be false modesty to consider that the Soviets have surpassed the U.S. to any substantial degree in the techniques of building and operating production facilities.

What, then, is the missing factor which is working to the Soviet's advantage? There is an aspect of the Soviet system of technological development which is absent from our own. There is a field of science and technique which has been the subject of earnest study and careful development in Russia, but which has been relatively neglected in the U.S. This unique present asset of Soviet technology demands our most scrutinizing attention.

Perhaps because of early dependence on foreign technology, perhaps because of Marxist preoccupation with propaganda techniques--for whatever reason, the Soviets have given great importance to the development of systems for the collection, organization, and dissemination of technical knowledge--the field that has recently come to be called in this country, information processing. The culmination of this development was the establishment a few years ago of the All Union Institute of Scientific and Technical Information, staffed by

more than 2300 specialists, supplemented by 20,000 scientists and engineers who act as part-time abstractors and translaters. The function of this Institute is to take in the published technical literature of the world (10,000 journals, 80 countries), process it, and make available the information it contains in the most useful form possible. In the few short years of its public existence this Institute has proven its effectiveness to the extent that prominent American scientists have said that the best way to find out what American science is doing is to read the Russian literature. It is painfully apparent that no such facility exists in the Western world.

The operation of this technical information center may provide the margin that has permitted the Soviets on occasion to outstrip us in the development race. Examining the entire span of any technical development from the theoretical discovery or unique concept through to finished hardware, it is immediately apparent that the major time interval is between the initial discovery and the design of a prototype device. In historical fact, this period may be as long as 50 or 100 years. Even the most involved engineering development seldom requires more than five or ten years. There is more time lost--or there is more time to be gained--between the laboratory and the drawing board than there is between the drawing board and the production line. And it is in this period where ideas are involved, rather than physical materials, that speedup is most feasible. Ideas have no inertia--they can be accelerated infinitely--subject only to the limits of our skill and facilities for information processing.

A Proposal

The technical community has suspected for some time that it was drowning in the flood of information and literature it was producing. The continued vitality of the scientific progress may demand an organized solution to the information problem. However, in an era in which international competition is increasingly determined on the basis of technological accomplishment, the problem takes on compelling urgency. We can no longer afford the piecemeal efforts toward fragmentary solutions which have been the only kind of efforts this problem has enjoyed in this country until now. The information system itself has become a weapon system. It is the "weapon system" on which all military devices as well as our peacetime progress depend.

In the present world environment the solution to the technical information problem requires two steps. First, we must build an immediate, interim facility to balance an existing military potential which has already proven its ability to work to our disadvantage in the present power struggle. Having this interim capacity in hand, we must work

with all urgency to seek an ultimate, efficient solution to the problem of full and rapid utilization of the knowledge mankind has gained, and is gaining, through thought and experiment. We know our competitor in the great struggle has long since mounted his attack on this ultimate problem. We do not know how far behind we are, but there is ample evidence, both direct and indirect, that he has made substantial progress.

The course to the interim solution is clear. Conventional survey techniques can define the magnitude and the scope of present technical information processing activity. For this phase, there is neither time nor need for a profound systems analysis of information processing operations, much less an intensive exploration of the mechanism of information transmittal. A valid statement of the size and general nature of the requirement will satisfy this preliminary phase.

With survey data in hand on the amount and kind of demand that exists for collection, translation, abstracting, codification, storage, dissemination, the first National Technical Information Center can be designed and established. In order to attain operation quickly, it must employ the best of available techniques and equipment. These techniques are essentially manual. It must utilize established information operations which have achieved partial solutions to the general problem, but which have not been previously coordinated. The initial operation will be awkward; it will be inefficient; above all, it will require many people. But it will be an operation. It will provide an American center where the knowledge of the world can be marshalled and made available to American scientists and engineers, managers and planners. The inadequacies of this stopgap system must be anticipated and accepted. The major goal in the design of the system must be to make it sufficiently general and flexible so that it can be evolved into a more rational and efficient system as new techniques are conceived and new equipment developed.

Once steps are underway to meet the immediate challenge, the more profound campaign for a true understanding and solution of the ultimate problem can be begun. This campaign must enlist the skills of many branches of modern science. Many are now applied to the information problem, but usually in isolation or without generalized coordination.

Before the systems problem can be tackled, we must ascertain the state of development of the components. Psychology will contribute its knowledge of the learning process and problem solving mechanisms. Library science has data about information patterns and inquiry characteristics and their interrelation. Linguists, semanticists, and lexicographers must provide the understanding of language structure necessary

for machine translation and codification. The dramatic new discoveries in electronics--memory and storage devices, data-processing systems, remote reading and printing equipment--must be exploited and adapted to provide the mechanisms which will make a true information processing system possible. Information theorists and statisticians have techniques which may be adapted to coding and searching information. Finally, the powerful techniques of systems analysis and operations research may be used to correlate these various factors and define the characteristics of the complete system.

To coordinate the contributions of these varied fields to this problem and determine the relative states of their varied arts will be a demanding task, but it is critical to the proper planning of a program to develop the ultimate information system. The development of that ultimate system is undoubtedly beyond the capacity of any one organization anywhere. It will require the participation of many teams from industry, universities, and research institutes. However, as has been learned in the development of other weapons systems, these various teams will contribute most effectively if their activities are coordinated and their general orientation planned by a central group.

Of course, the optimal system will never be developed. This is an open-ended job which will continue to achieve refinements for all foreseeable time. However, properly integrated with an operating information center, the development program can test and then install its equipment and benefit from continual feedback, which, in turn, will stimulate further discoveries.

The true value to the nation's economy of a comprehensive and effective technical information center using high-speed automatic equipment is difficult to calculate, but it is of large magnitude. It has been estimated that such a facility would have the effect of increasing by at least 25 percent the effective supply of creative scientists and engineers by saving time now spent in laborious literature searching and from the recovery of time now wasted in repeating work already done and reported, but inaccessible in the present chaotic mass of technical literature.

Beyond these savings would be the elimination of much of the effort and expenditure now put into collecting and maintaining libraries by traditional methods--expenditures that are duplicated thousands of times throughout the country to perform nearly identical operations.

Finally, there are the benefits which can be foreseen only generally. It is exciting to contemplate the stimulation and innovation that may be achieved through a facility which can bring to bear on a specific problem virtually all past recorded thought and data, cheaply and quickly.

There is a justifiable analogy with the introduction of high-capacity computers during the past decade. These devices have made possible entry in areas of research, and even areas of thought, that were inaccessible or inconceivable before these far-reaching extensions of human capacities were available. Similar new potentialities can be anticipated from high-capacity information processing equipment.

The rewards for the successful solution to the information processing problem, both economic and intellectual, are great. The political and military incentives for seeking this solution immediately, with high concentration of effort, are also great. Failure to grasp the information problem in its entirety and to seek a general and positive solution will increasingly entangle our technological society with unread literature and unused knowledge until we cannot fail to lose our present position of technological advantage. In a period in which both military might and international prestige are heavily dependent upon scientific and technical achievement, we cannot afford to fail to meet this challenge.

M. L. Kastens:erm
Stanford Research Institute
November 26, 1957

DRAFT
C. P. Bourne
January 23, 1958

Questions to be Edited for Considerations

I Format used for this listing:

NTIC Market Survey

What's available?
User's Needs and Habits
NTIC Justification
Retrieval theory
Realization
Operation
Prime Agency Problems

Retrieval Theory
Operation
Prime Agency Problems

II Listing:

NTIC Market Survey - what's available?

What are the total number of technical journals in each language?

What are the circulation figures for each of the technical journals?

What journals are not covered by an abstracting service?

What journals have duplicate coverage by an abstracting service?

How much recorded information is available for utilization?

How many books, journals?

In how many fields and in what languages?

What is the distribution of quality of this material?

To what class of user is the material directed?

What technical societies could cooperate to publish a single journal instead of numerous splinter journals?

Which of the technical publications furnish an annual index?

What would be the physical dimensions, weight and cost of a year's crop of all the technical publications? (bound and unbound)

What potential audience is currently exposed to the abstracts, of articles which appear in each of the particular journals? (e.g., abstracts of an article in the IRE may be published in 8 different publications. What is the combined circulation of the people who could see this abstract?)

Of the currently-operating abstracting services, how many are operating merely to satisfy an obligation of a professional society and would rather have somebody else do the abstracting?

What are the total number of abstracting services in each language?

In cases where abstracts of a particular article appeared in several publications, how many of these abstracts were only copied, and how many were written by different people?

Are you implying that the listed group of centers which serve to process technical information aren't satisfactory? Why aren't they? -- They don't cover the field well enough, Senator - duplication, gaps, lack of cross-fertilization.

What technical libraries are currently staffed to handle a large number of information requests? (e.g., NRL, Battelle, ...)

What organizations or publications provide any special information services? (e.g., edge-punched cards, microfilm)

Are there any useful information services which aren't being provided now?

Why not periodically publish inventories of research in progress, to indicate what research projects are currently being undertaken in each specialty field, thus helping to eliminate duplication, as well as to generate interest in other scientists. Another service would be a listing of ideas for future development or thesis work.

What material is not now, but should be incorporated in an NTIC file?

What is the time interval now between the publication and the arrival into the possession of an interested researcher of pertinent documents?

Certainly some work on this problem has been done in the U. S. before now. Who has done it and how much have they done?

Are there enough Russian-speaking scientists, or qualified people, to translate Russian scientific literature right now?

What kind of effort is being duplicated? Do you have any examples? At what magnitude of expense? To whom?

NTIC Market Survey - User's needs and habits

How are the information needs of a scientist affected by his age, educational level, profession, type of position held, etc.?

What is the distribution of document age vs. demand, for each of the specialty fields?

What is the role of information retrieval, storage, etc., in the decision-making process for such as the research worker, scholar, administrator, etc?

What increase in efficiency would result from improving the accessability of recorded information?

What characteristics must a documentation system possess if it is to meet the requirements of a particular discipline, subject field...?

By what processes does the scientist keep abreast of the advances in the art? What are the relative importances of each of these processes?

How much time does it take a typical researcher to scan all of the literature which might be of value to him? Is it well enough organized for him to be able to scan a minimum amount of material not of immediate interest in order that he doesn't miss that which is of value?

How can the information services be made more attractive to the users?

How much need is there for foreign literature?

Is any particular type of publication any more efficient or useful for information purposes?

What type of questions go unanswered at the libraries?

If abstracts are desired, what is the best form? (critical, informative)

Could we suspend publication of journals and just furnish reprints upon request?

How many scientists work without adequate access to a library or document center?

What is the main problem - the time delay, or the fact that many documents just won't reach the reader?

Market Survey - NTIC Justification

Just why do you think we need an NTIC?

How about some more explicit data about the Russian center? What services do they really provide? Have you ever asked a Russian scientist how much good this might do him?

What are the relative benefits, short and long range, likely to be among military, industrial, scholarly, commercial, government activities, of having an effective NTIC through the years?

Why wasn't this proposed five years ago?

Why is so much research needed?

What research costs are justifiable?
What type and depth of research is needed?
What would you do now if you received a grant of $_____ ?

What is it costing us in appropriation money not to have an NTIC? Can we estimate how much savings in federal money will be available (by NTIC) to spend on NTIC?

Can you prove that the expense of delay and duplication now is greater than that of establishing and operating your NTIC?

If all of this money is forthcoming, just what services will be then provided that are different and better than what is now available? Is this to be entirely new types of service, a real advance in the state of the art, or is it to be just more and better of the same thing?

You say that we are using "processes and services that are over 50 years old." Surely something new has been added in the last 50 years - I find your statement hard to believe. What have our ASTIA appropriations been doing?

How can a Congressman use a NTIC? (Who around here might be familiar with the special information retrieval problems of a Congressman?)

What could this do for lawyers and business men, such as many Congressmen were? What could it have done for them if they had had such?

Just how might this be helpful to my home community?

Has it been tried before? What kind of real benefit can you say will accrue from the establishment of NTIC? Can you back this up with figures?

How important is it to know what the rest of the world is doing? English is good enough for me, why isn't it good enough for the scientist?

By "serious delays in technological progress," do you mean that this would have helped us beat Sputnik? How? Can you really give us any evidence that we were beaten on Sputnik, or on the bombs or missiles, by such as this? Can you concrete examples or evidence of serious delays of any kind caused by the problem you expound?

I don't understand what you mean by "paralyzing glut" and "ultimate collapse" of our technical communications. Haven't people always had to hunt for the information they want? Isn't this just a matter of degree?

You say "We have not reached our full potential." Come now, do you ever really expect to reach the full potential when humans are concerned? Can you define for me "full potential"? How much better do you think we can get?

How can we compare the effectiveness of current machine searching and retrieval systems?

How many retrieval problems could be solved by teaching the researchers more about present-day documentation techniques? How much is missed by their leaning too heavily upon librarians?

Is our problem that the information now is just not available at all, or is it that it is just a little hard to find? How much can be done just by concentrating on abstract distribution and better dissemination techniques? Is it going to be worth it to have push-button magic do what a few hours of searching could do?

How much more is the push-button system going to give the researcher, in the way of valuable information, that he could not have found himself after a reasonably exhaustive personal search through up-to-date index systems and abstract compendia?

Do you know of any instances in which companies with similar interests cooperatively operated an information center?

Market Survey - Retrieval Theory

Is the system available for browsing?

To what extent will past literature have to be digested and included in the NTIC? If NTIC services are to be saleable to researchers, how much of the past literature will have to be included?

What constitutes an information retrieval problem? How many of the problems are problems because people might not know of present techniques which could help them, and how many are beyond the ability of present techniques to help?

Some might think that information retrieval problem is try ng to find all of the pertinent information there is pertaining to a given special information need. One might have a narrower picture of the "problem" - e.g. that it is merely one of being able to store and search - or merely to store large amounts of information, or codes. In a sense, one might say that the "problem" is one of understanding the general and specific needs of researchers, and when this is understood, solution of the rest will be easy.

Will you just be able to add years and years of information into the system or will it all become too bulky and obsolete in a few years so we'll have new appropriations to argue about?

How about utilizing, instead of a giant national center, a collection of special libraries?

Couldn't a group of smaller centers, for specific fields, be of greater utility - more tractable?

Can't we do it on a smaller scale? Does it have to be a big institute or nothing?

What are the advantages and disadvantages of a centralized organization such as NTIC?

Is there some other way of doing it?

The nature of the material to be stored for retrieval must be considered from the aspects of (1) how often will it be searched, (2) how sophisticated a search process is needed - i.e., would simple linear classification do? (3) how much information is likely to be involved in the different categories.

Considerations, operational - Should make sure that the final methods chosen for a retrieval system are not influenced too heavily by the requirement of compatibility with past systems.

A given theory of retrieving must know the relationship between time and complexity in a search and the amount of material to be searched. ie, a system might work well with a few thousand entries, but be hopeless with a hundred thousand.

We mustconsider how the processing of recorded information may be conducted so that it can be used effectively in spite of human limitations, or of limitations in numbers of human beings.

On the basis of current information-processing trends, manpower requirements will soon make it i possible to continue to apply present methods of processing.

Will the machine give qualitative answers (e.g., yes, no, 5,000 tons in 1945, etc.) as well as references?

Will you be able to perform a searching operation in parts? (For example, search 1/10 of the file and check the results to see if you need further searching, in which case you resume searching in the other9/10 of the file?)

What type of information should be contained in the NTIC? :books (texts, tables), technical and trade journals, conference proceedings, papers presented at conferences but not published, industrial and government interim and final project reports, etc. (similar to the type handled by ASTIA), operation and instruction manuals, patents, manufacturer's catalogs, abstract publications, maps, films and photographs

What are the cataloging schemes for the principal abstract services?

What are the special information requirements for different specialty fields?

Market Survey - Realization

What is the present state of our methods, systems, techniques, equipment, and procedures for meeting modern information requirements?

Market Survey - Operation

Will the machine be accessible to anyone by telephone or other direct device, such that the searcher can interrogate the file directly and at will?

Are the output results going to be in English, or a Zeidler Binarital code that we'll have to have translated?

Are there going to be any significant delays in getting information entered into this system?

Can you really be sure that Secret documents can't get out through your center? Who is going to fix the price for the service? Who is going to determine the priority of the service?

What is the expected range of inquiries to be answered, or services to be provided?

What is the expected yearly document input?

What cost data is available on the cost of preparing an abstract?

Can you cope with the Russian language? Do we have enough translators?

What will I get in my hand after asking your center? and how do I put the questions to it? Do I mail it, phone it?

Will the NTIC output be in a form such that the researcher can take it to his own librarian and find which of the documents on the list are in his own company's collection?

Results (600) of questionnaire on abstracting plots of: (1) no. abstracts prepared by companies for their employees, (2) abstracting personnel, (3) abstract distribution, (4) annual budget for abstracting and indexing (5) time lag: paper to abstract, (6) attitudes of companies toward services by professional societies.

Extra desired service: (1) provision of records, ready for machine searching, (2) question-answering services, (3) monthly indexes to chem. abstracts, (4) overall indexing of all fields, with abstracts available on request to cover given subject ranges, (5) speedup of existing services

There are 312 abstracting and indexing services. Average organization uses 10.

Information on services provided by prof. societies (and abstracting budgets).

Will the system be able to retain a file of questions so that it can ask these questions of all new input material, thus providing up-to-the-minute data for standing questions?

What type of abstract to use? (descriptive, critical, indicative)

If we must look to the National Government for support, what residual responsibilities remain with the professional societies?

How will you pay loose the information that might be useful, but held?

Would this be a civil service organization?

Market-operational-considerations: With the foreseeable exponential increase in the quantity of information that has to be processed by information centers, would it be economically feasible for any sort of commercial enterprise or internally-sponsored professional group, to establish and run a center/? Could these centers be self-supporting and still furnish the type of service which continued social advancement requires?

Would it be feasible to obtain legal powers to speed up the standardisation and coordination of existing facilities (such as the FCC)

Would it be feasible for the abstracting publications to use a standard format and type font such that cuts (or some similar technique) could easily be distributed to other interested publishers, thus saving printing expenses?

With present-day methods, what would it cost to microfilm a year's crop of technical journals? What would be the physical dimensions of such a file?

What about copyrighted material? Would royalties be forthcoming to the owner of the copyright if your #TIC uses the article?

How are we sure that it wouldn't be better to reduce the number of or amount of literature rather than go to the tremendous expense of providing super-service for all of it? Why can't we put a quality filter on this output, weeding out what won't have significant future value?

How can we get good-quality abstracts?

Facsimile for Federal Libraries Adams, Public Service, 5/53, p.169

Inter-library loans: 100,000 vols. for Library of Congress in 1951; 2/3 of this is scientific periodicals; average length of required article is 10 pages; None government libraries estimate inter-library loans cost at 1-7 dollars/loan.

Inter-library loans by facsimile violate copyright laws.

Is the taxpayer going to be paying for something that will just give more advantages to the big-corporation research centers? Is big business going to have any better access than small businesses or individuals?

Would a private citizen (scholar) be able to afford the service?

Will you merely convert new information into the system, or will you go back and put in all the old information?

Who do you feel is competent to do the job?

1958 Intl. Conference on Scientific Information, Area 2, pg. 2 "Russian Inst. states that it furnishes complete coverage of science on the basis of 10,000 journals. From these 10,000 journals they prepare 400,000 abstracts, to be compared with Chem. Abstr. checking 1000 journals for 80,000 abstracts.

The Russians agree with the C. A. view that professional abstractors should not be used. The abstracts appear in 13 abstract journals."

How many part-time voluntary abstractors are presently working for the professional societies, and what is their average monthly production and the average time spent per abstract?

What are the present translation facilities in this country for each of the important technical languages? (Russian, German, Japanese, etc.).

How much delay can the user tolerate?

Market Survey - Prime Agency Problems

Table of Abstracting and Indexing Services (312)
"In the U.S., the tendency seems to have been in the direction of fragmenting the information services in many professional fields" vs. In the USSR, trend is toward a central agency.

Couldn't the money be spent better on something else? How about spending it on education instead?

What technical-manpower drain would the proposed NTIC program have on our other high-priority scientific programs?

What are the specific aims of the program?

Any prediction on when you could include such fields as Law and Medicine in your center?

Where will the Library of Congress and ASTIA fit into the picture? Can they be expanded to do the job?

Where would this center be?

Shall manufacturer's files (as in Batelle) be considered a part of the system?

Any idea of what the situation will be ten years from now?

How soon will we get results?

What is the range of subject matter to be included?

What is the expected rate of growth of the system?

How many things, such as ASTIA, can we cease spending money on, to help fund this program?

If we start with something like this, how soon could we drop ASTIA?

Isn't this exactly what ASTIA has been doing? What is the difference between what you propose and ASTIA?

What if we gave $60 million a year more to ASTIA, wouldn't this solve lots of our problems?

Could this possibly be sponsored by private industry? Why go to the government? Could private groups, professional societies, sponsor important special collections?

Can't the implementation of the solution for this be done right within Government agencies, e.g. Bureau of Standards and the Library of Congress, working together?

If not handled by some existing Government agency, would it not be best handled by some University, e.g. MIT?

Retrieval Theory

Could the system operate on a "Just search 1/2 the file for me, I don't need a comprehensive search" basis?

Will you make any effort to remove outdated or obsolete data from the file?

Are you going to record everything or will there be an editor - or a censor?

As the system grows in size, will it be possible to make changes easily in the classification scheme and bring the old coding into the new scheme?

Kauts feels that main problem of information retrieval is one of classification - since people so seldom can express in a usable way (to the documentalist) what it is he wants.

How much work must be done to standardize a vocabulary? To use for coding and searching.

Retrieval theory - self that from 12/16/57 - Wouldn't compilations of previous researchers be very valuable to later searchers? Could the record of the query structure and the retrieved-data results be entered into the retrieval system (probably in special classified section) as a new accession item?

This might get complicated with respect to security requirements but could like the efficiency a good deal.

Could a more flexible, yet comprehensive and efficient retrieval system be built if the search query be "compiled" and coded by a special-purpose computer before the search machine takes over? Perhaps "compiler" could use random samplings of the storage to help it predict in optimum query.

How much updating of active search files can be managed (how often, how completely, how easily?), and how much might this possibility be of advantage (theoretically) in developing the system? If retrieval coding of the stored items can be modified as usage and relative-importance shifts may point up need, could give lots more flexibility, and relieve initial coding of a great deal of burden.

Operation

What professional and educational background is needed for the personnel to operate this system?

Should a review board screen the subject fields and select the most significant papers from that year's crop, and only pass these papers to the retrieval system?

Does "having a large Central Information Center" mean the physical collection of all these services at one central location?

If translating is to be done, why not have it done in Europe?

Is there going to be any attempt to standardize the form of the material before it gets into the center? i.e., does the material have to be on standard-sized sheets or forms?

Areas for Cooperation in Information Processing: (1) Intellectual tasks - (a) writing papers and abstracts, (b) indexing, classifying, encoding, for machine search; (2) Production and distribution of (1) above - (a) printing and publication, (b) distribution of card catalogs, punched card files, etc.
e.g. Punched Card Service for Corrosion Engineers
e.g. Am. Petroleum Inst. Abstract Bulletin.

Would a duplicate file and collection be maintained to prevent disruption of service due to fires, etc.?

Would the information center keep a collection of the original documents on hand?

Should a review board be set up to periodically review and re-code the old information in the file?

Should the NTIC also act as a publishing agency to periodically publish collections of papers (reprints) in a very new and special fields (e.g., info retrieval)

Prime Agency Problems

What factors are going to determine the location - can strategic dispersal considerations influence the location without adversely affecting efficiency?

For an interim solution, can we translate the exhaustive Russian abstracts and provide distribution of these for our own people so as to leave our interim energies more free for other uses?

Should there be special training for abstractors and translators (similar to what Calhoun reports Russia is giving for translators)?

Are we just trying to copy Russia in establishing this center?

Is this just a face-saving move to keep up with Russia?

Is this really a significant thing to do - will it be popular, meaningful, or is it just another proposal by an outfit that wants to get itself funded?

66. Are you implying that the listed group of center which serve to process technical information aren't satisfactory? Why aren't they?

they don't cover the field well enough, Senator....duplication, gaps, lack of cross-fertiliaation...

67. Any prediction on when you could include such fields as Law and Medicine in your center?

depends upon the money available, Senator.

How much, and when?

68. I don'd understand what you mean by "paralyzing glut" and "ultimate collapsE" of our technical communications. Haven't people always had to hunt for the information they want? Isn't this this just a matter of degree?

69. You say that we are using "processes and services that are over fifty years old." Surely something new hasbeen added in the last fifty years...I find your statement hard to believe. What haveour ASTIA appropriations been doing?

70. What if we gave $60 million a year more to ASTIA, wouldn't this solve lots of our pro lems?

71. You say that no rapid access is a ailable to special small collections of information....what about IBM's RAM?

72. Couldn't a group of smaller centers, for specific fields, be of greater utility...more tractable?

73.

~~Do people who say~~

X73 What type of questions go unanswered at the libraries?

X74 What type of abstract to use? (descriptive, critical, indicative)

X75 What are the ~~adv~~ advantages & dis-advantages of a centralized organization such as NTIC?

X76 What cost data is avail. on the cost of preparing an abstract.

X77 How much work must be done to standardize a vocabulary?

X78 Is the system available for browsing?

X79 Will you make any effort to remove outdated or obsolete data from the file?

X80 As the system grows in size, will it be possible & make changes easily in the classification scheme & bring the old documents into the new scheme?

X81 How can we ~~effectively~~ searching & retrieval compare the effectiveness of current working systems?

X82 What is the expected yearly document input?

X83 " " " rate of growth of the system?

X84 " " " range of inquiries to be answered, or services to be provided?

X85 " " range of subject matter to be included?

~~" " expected probability~~

X86 Could the system operate on a "just search ½ the ~~yearly~~ file for me, I don't ~~if~~ need a comprehensive search" basis?

X87 If we must look to the Natl. Govt. for support, what residual responsibilities remain with the professional societies?

53. What about copyrighted material? Would royalties be forthcoming to the owner of the copyright if your NTIC uses the article?

Kautz.

54. How are we sure that it wouldn't be better to reduce the number or amount of literature rather than go to the tremendous expense of providing super-service for all of it? Why can't we put a quality filter on this output, weeding out what won't have significant future value?

54. For an interim solution, can we translate the exhaustive Russian abstracts, and provide distribution of these for our own people, so as to leave our interim energies more free for other uses?

55. To what extent will past literature have to be digested and included in the NTIC? If NTIC services are to be saleable to researchers, how much of the past literature will have to be included?

56. How can we get good=quality abstracts.

57. Are we just trying to copy Russia in establishing this center? and translatora.

58. Should there be special training for abstractors (similar to what Calhoun reports Russia as giving for translators)

59. Are there enough Russian-speaking scientists, or qualified people, to translate Russian scientific literature right now?

60. As a potential user of an info. center, Kautz would like mainly to recieve periodically the lists of abstracts of recent publications which bear in a pertinent sense upon fields of his interest.

61. Kautz feels that main problem of info. retrieval is one of classification.... since people so seldom can express in a useable way (to the documentalist) what it is he wants.

Miller again

62. What kind of effort is being duplicated? Do you have any examples? What At what magnitude of expense? To whom?

63. By "serious delays in technological progress", do you mean that this would have helped us beat Sputnik? How? Can you really give us any evidence that we were beaten on Sputnik, or on the bombs or missles, by such as this? Can you concrete examples or evidence of serious delays of any kind, caused by the problem you expound?

64. You say, "We have not reached our full potential." Come now, do you ever really expect to reach the full potential, when humans are concerned? Can you define for me, "full potential?" How much better do you think we can get?

65. As an example, in Steve's experience; He wanted to build a flash X-ray unit. Found only two articles pertinent in available literature, through use of a specialized bibliographer (Jean Furnish). Took them 18 mos., three men about half time. Afterwards, they found, in articles on shaped charges, etc, very detailed literature on the subject which could have allowed completeion probably in three months. These reports were in SRI Library....a matter of retrieval. Just want the facts, Mam!

Rosen, cont'd.

42. What could this do for lawyers and business men...such as many Congressmen here? What could it have done for them if they had had such?

43. How will you pry loose the information that might be useful, but held. ...?

1-8-58

44. What constitutes an information retrival problem? How many of the problems are problems because people might not know of present techniques which could help them, and how many are beyond the ability of present techniques to help?

Some might think that an information retrieval problem is trying to find all of the pertinent information there is pertaining to a given special information need. Some might have a narrower picture of the "problem", e.g. that it is merely one of being able to store and search...or merely to store..large amounts of information, or codes. In a sense, one might say that the "problem" is one of understanding the general and specific needs of researchers, and when this is understood, solution of the rest will be easy.

45. How many retrieval problems could be solved by teaching the researchers more about present-day documentation techniques...how much is missed by their leaning too heavily upon librarians?

Is our problem that the information now is just not available at all, or is it that it is just a little hard to find? How much can be done just by concentrating on abstract distribution, and better dissemination techniques? Is it going to be worth it to have push-button magic to do what a few hours of searching could do?
How much more is the push-button system going to give the researcher, in the way of valuable information, than he could not have found himself after a reasonably exhaustive personal search through up-to-date index systems and abstract compendia?

46. If all of this money is forthcoming, just what services will be then provided that are different and better than what is now available? Is this to be entirely new types of service, a real advance in the state ofthe art, or is it to be just more and better of the same thing?

Steve Miller...

47. Certainly some work on this problem has been done in the U.S. before now. Who has done it, and how much have they done?

48. Can't the implementation of the solution for this be done right within Gov. agencies, e.g. Bureau of Standards and the Library of Congress, working together?

49. Isn't this exactly what ASTIA has been doing? What's the difference between what you propose and ASTIA?

50. How many things, such as ASTIA, can we cease spending money on, to help fund this program?

51. If we start with something like this...how soon could we drop ASTIA?

52. If not handled by some existing Gov. agency, would it not be best handled by some University? e.g. W.R.U.

Ziedler

Dev. 23. Are theregoing to be any significant delays in getting information entered into this system?

Mkt. 24. What is the time interval now between the publication and the arrival into the possession of an interested researcher of pertinent documents?

Mkt. 25. What is the main problem,,, the time delay, or the fact that many documents just wont reach the reader?

Dev. 26. Are the output results going to be in English , or a Zeidler Binarital code that we'll have to have translated?

Dev. 27. Will you merely convert new information into the system, or will you go back and put in all the old information?

Mkt. Dev. 28. Will yOu just be able to add years and years of information into the system, or will it all become too bulky and obsolete in a few years, so we'll have neww appropriations to argue about?

Mkt. 29. Any idea of what the situation will be ten years from now?

Mkt. 3.0 Can you prove that the expense of delay and duplication now is greater than that of establishing and operating your NTIC./

Dev. 31. Are you going to record everything, or will there be an editor...or a censor?

Mkt Hist. 32. How about some more explicit data about the Russian center ? What services do they really provide? Have you ever asked a Russian scientist how much good this might do him?

Political Dev 33. Is the taxpayer going to be paying for something that will just give more advantages to the big-corporation research ceners? Is big business goi ng to have any better ac ess than small business or individuals?

Leo

Mkt. Dev., Op. 34. Can you cope with the Russian language? Do we have enough translators?

Mkt. - Dev. 35. What are the specific aims of the program?

Mkt. Dev. 36. What will I get in my hand after asking your center? And how do I put the questions to it? Do I mail it, phone it, ...?

37. Is there going to be any attempt to standardize the form of the material before it gets into the censer? i.e. does the material have to be on standard-sized sheets or forms?

Rosen

38. Is this really a significant thing to do..will it be popular, meaningfull, o r is it just another proposal by an outfit that wants to get itself funded?

39. What is it costing us in appropriation money not to have an NTIC...CAN WE estimate how much savings in Fed money will be available (by NTIC) to spend on NTIC?

40. How important is it to know what the rest of the world is doing? English is good enough for me, why isn't it good enought for the scientist?

41. What are the relative benefits, short and long range, likely to be among military, industrial, scholarly, commercial, Government activities, of having an effedtive NTIC through the years?

What if we were members of a Senate committee, querying F. Carter...

1. just why do you think we need such as NTIC?

2. Why wasn't this proposed five years ago?

3. What technical-manpower drain would the proposed NTIC program have on our other high-priority scientific programs?

4. Is there some other way of doing it?

5. Can't we do it on a smaller scale? Does it have to be a big institute or nothing?

6. Couldn't the money be better spent on something else? How about spending it on education, instead.

7. Who do you feel is competent to do the job?

8. Has it been tried before? What kind of real benefit can you say will accrue from the establishment of NTIC? Can you back this up with figures?

9. Where did you get the $50 million figure?

10. How about utilizing, instead of a giant national center, a collection of special libraries.

11. Could this possibly be sponsored by private industry? Why ~~~~ to the government? Could private groups, professional societies, sponsor important specill collections?

12. Why is so much research needed?

13. Where would this center be?

14. Where will the Library of Congress fit into the picture? Can it be expanded to do the job?

15. How soon will we get results?

16. Is this just a face-saving move to keep up with Russia?

17. Would a private citizen (scholar) be able to afford the service?

18. How can a Congressman use a NTIC? (Who around here might be familiar with the special information retrieval problems of a Congressman?)

19. Just how might this be helpful to my home community?

20. Would this be a civil service organization?

21. Can you really be sure that secret documents can't get out through your center? Who ~~the hell~~ is going to fix the price for the service? Who ~~the hell~~ is going to determine the priority of the service?

22. What factors are going to determine the location...can strategic dispersal considerations influence the location without adversely affecting efficientcy?

DRAFT
C.B.
Jan 23,'58

The Systems Group would:

Coordinate and direct all the information-gathering activities of the concurrent study phase.

Collect all the data from the information-gathering activities and examine it to determine where the duplication of efforts exist, and to determine the areas which need added or improved coverage.

Show how to incorporate the existing facilities into a homogenous working agency.

Determine what new services should be introduced, and what services should be discontinued.

Evaluate the different methods of classification and possible modifications of these methods, to see what would be best suited for the NTIC. The chosen system would have to be flexible enough to be altered as the system matures.

Examine the overall NTIC operation to determine the areas which would lend themselves to mechanization techniques.

Indicate the areas where more research or development work is currently needed, and where future work should be directed.

Determine a general organizational structure and method of operation for the NTIC.

Develop a means of obtaining a performance "figure of merit" for the system so that the usefullness of the NTIC may be quantitatively evaluated as the system grows and as experiments are performed upon it.

Develop methods for obtaining information from the NTIC so that operating and performance data will be readily available for future studies and planning toward a long range system.

Recommend the areas and level of subject matter to be included in the NTIC, the classification or coding scheme to be used, the method for handling inquiries, the form of the final product and its system of dissemination, the type of storage to be used, the procedures for determining priority and fixing the fees, the degree of translating service to be provided and how they may be implemented.

Continue to observe the operating NTIC for stated period in order to make suggestions for improving the performance of the NTIC and to obtain more data for recommendations on the ultimate information center.

WRU, Jan 21, 1958

~~WRU~~: A Plan for the Creation of a National Center for the Coordination of Scientific & Technical Information

The major objectives of the proposed natl. center would be:

1. To provide effective interdisciplinary information services.
2. To provide such information services as may be required to fill existing gaps in various fields, until such time as new agencies may be established.
3. To foster the development of new information services by existing agencies.
4. To provide advice and guidance to cooperating information centers in order to ensure economy of effort.
5. To act as a clearinghouse for information requirements that can be referred to existing agencies.
6. To foster and to conduct research and development in methods, equipment, and procedures in order that better services may be provided and costs may be reduced.

The principal services to be provided by the proposed national center:

1. To provide, from the world's scientific and technical literature:
   a) direct information, on demand or on a continuing basis, in response to specific requests from individuals or organizations.
   b) an interdisciplinary literature searching service.
   c) published abstracts of material not available from existing agencies.
2. To distribute copies of the coded media prepared at the Center so that machine searching may be performed wherever facilities are available.
3. To stimulate the preparation of periodic bibliographic reviews by subject specialists.
4. To work toward complete translation of the world's scientific and technical literature by support of existing ~~services~~ ~~support of existing services~~ and translation services at the Center.

estimated WSE activity: 360,000 papers processed/yr.

Phasing - Pilot operation initiated in 7th & 12th mos., Full scale operation after this time.

"With interdisciplinary searches, it is to be expected that useful information may be found in almost any document in the file. Therefore, a total search of each document in the file, as a "fail-safe" method of searching for any document of potential importance to any given search, is a basic requirement for the Coordination Center." Although initial operations of the Center would (of necessity) depend upon presently available processing, a highly sophisticated coordinated retrieval system would be phased into operation at the earliest possible time."

"The cost of service is predicated on a basic searching cost of 300 references searched per 1 cent.

quoted figures: estimated scanning speed of 300,000 doc./hr.

No. yrs. of World's sci. info.	10	25	75
No. documents	4 mill.	10 mill.	30 mill.

reference: A. Kent, "Study of Centralized Information Services" report to NSF (Grant G-4105) 12/31/57 (190 p.)

POSSIBLE SPONSORS FOR THE SOLUTION OF COMMERCIAL INFORMATION RETRIEVAL PROBLEMS

1. Catalogs and customer information requests.
   eg. (quoted at IBM seminar) Norton Tool Grinding Co. handles ___ thousand requests per month concerning what grinding wheels are applicable to each customer inquiry.

2. Listings and Directories.
   eg. selection from a mailing list of people who are coded for a special characterists. An example might be the extracting of a list of names from the IRE Directory of people who have an interest in components, etc..

3. Commercial or Industrial Intelligence Services.
   eg. Dow Chem. would like to keep informed (via journals, newspapers, contacts, etc) of all the investments, business transactions, expansion plans, technical developments, etc. of the competition.

   eg. Any of the large national magazines or newspapers (TIME, N.Y. TIMES) could benefit by a more efficient morgue or information file.

   eg. Govt. agencies (intelligence and otherwise)

   eg. Police Dept - how to search fingerprint files with less than a complete set of prints?

   eg. Drafting of legislation, etc.

   eg. Maintenance of mailing lists (additions, change of address, deletions, etc) for large installations. (eg. Readers Digest, Time-Life, Tax Agencies)

4. Large abstracting, collecting, and disseminating serviced
   eg. ASTIA , NSF

5. Classical Info. Retrieval Problems
   eg) mechanizing or implementing legal search for precedents, legislation, etc.
   eg) " " scientific " .

D. C. Engelbart
January 2, 1958

Expanded Sequence List

A. NTIC "Market" survey ... since we, now, are oriented to consideration of NTIC

1. user requirements?

2. economics, how much would various kinds of service be worth to the consumer market? (limits) ... speed, completeness, volume, final results, limits

3. who has access needs? (limits)

4. security requirements?

5. priority considerations?

6. integrate into No. 1 the influence of present index-abstract services and scientific-society efforts.

7. where would information boundaries be if it is to be a "technical" file?

8. vertically, what technical level to be included?

(possibly in D) ~~9. duplication of file needed?~~

10. Effect of separation (geographic)? Should a number of centers be established?

11. volume of "business"?(limits)

12. nature of service...i.e. ponderous, exhaustive...fast, flexible, more limited? (limits)

(D?) ~~13. how determine billing?~~

14. effects of future developments in dissemination, research techniques, etc. on future needs?

Who will benefit by NTIC? What is the need now? Predict future growth, needs? How much can be saved? What do Russians provide?

B. Retrieval Theory

1. What schemes are applicable to large, dynamic systems?

2. Must work with

a. high capacity
b. be realizable
c. interdisciplinary language differences, other-country language.
d. semantic shifts (need a standard dictionary?)

3. list theories to date, abstracting but not classifying.

4. how would the totality of information be handled if effective systems could only work comprehensively with fragments?

C. Realization

1. personnel requirements?

2. what's been done? (abstract)

3. what theories do they fit (no.?)?

4. what's been proposed?

5. analyze each "theory" possibility for realization techniques - anticipate bottlenecks.

6. consider these bottlenecks.

7. cost of each of 5 (initial, operational)

8. automatic input processing, billing

9. transmission

10. mechanize translation?

D. Operation

1. by whom - agency, corporation

2. self supporting?

3. controlled by? (political)

4. duplication of file?

5. How determine billing

6. Coordination with industry, professional societities, publishers, armed services, libraries.

7. individual working scientists for abstractors

8. procurement (borrow? buy? store originals?

9. copyright problems.

PRIME AGENCY PROBLEMS (problems of agency which is responsible for NTIC development)

E. ~~NPIC Developmental problems (project problems)~~

1. who controls ~~it~~ NTIC?
2. who manages ~~it~~ NTIC?
3. manpower requirements for NTIC - number, skills, phasing?
4. schedules for NTIC
5. integration of existing facilities into NTIC
6. intermediate solutions for temp. NTIC
7. what system will be used?
8. How much will it cost

F. Subsidiary developmental pressures

1. Information dissemination systems should evolve to shorten time - perhaps allow for intermediate reports with less formality - reduce trivia
2. publishers provide digital translation
3. research programs include NTIC query more as a matter of course
4. printing techniques should become such that cheap copies can be made of photo-storage - or master storage, so small-lot documents may be obtained for price comparable to present books.
5. Individual, and small-group, techniques for automatizing direct utilization of retrieval and generated information.

STANFORD RESEARCH INSTITUTE
MENLO PARK, CALIFORNIA

January, 1958

Baurne

A DRAFT PROGRAM FOR A NATIONAL TECHNICAL INFORMATION CENTER

THE PROBLEM

The preeminent position of the United States as a world power is dependent upon continued leadership in science and technology. Of vital importance to technical progress is the ready availability and applicability of discoveries, concepts, and data from all sources--past and present, foreign, and domestic.

Existing mechanisms for processing technical information are unable to cope with the present deluge of publications. The result has been enormous duplication of effort and expense, serious delay of technological progress, and failure to realize the full potential of a rapidly expanding military and civilian research effort.

THE PROGRAM

This document describes a program to solve the nation's technical information problem through the establishment of a national center for the collection, processing, storing, retrieval, and dissemination of scientific and technical information from both foreign and domestic sources. The program comprises the following actions:

1. Establish a central organizing and administering Agency.
2. Determine the gross dimensions of the problem.
3. Establish an interim information center using existing services and techniques.
4. Analyze the factors which determine the design and operation of an ultimate National Technical Information Center.
5. Encourage present and initiate additional research and engineering development programs leading to systems and equipment necessary to implement the ultimate National Technical Information Center.

NATIONAL TECHNICAL INFORMATION CENTER

Because of the size and nature of the National Technical Information Center, it should be administered by a federally constituted Agency.

It is strongly recommended that the resources and services of the Center be equally accessible to all research, industrial, academic, and government organizations throughout the country, consistent with security requirements.

In view of the breadth of service anticipated and the demonstrable economic value of these services, the Center should be at least partially supported through service fees from its various users.

There is no inherent limitation on the scope of subject matter which can be processed by an operation such as described here. However, in order to develop an integrated system in a reasonable time, it is desirable to place arbitrary boundaries on the information to be included. At present the most highly organized and easily defined body of information is the literature of the physical and biological sciences. There are compelling practical incentives to include the slightly less well organized literature of the various fields of engineering and medicine. It is recommended that the initial planning of the system be limited to not more than these categories. Once the Center is in operation, a well-conceived system can be extended to encompass the literature of law, the behavioral sciences, and other fields of importance, but which are presently less easily organized for systematic processing. This expansion should be undertaken at the earliest possible time.

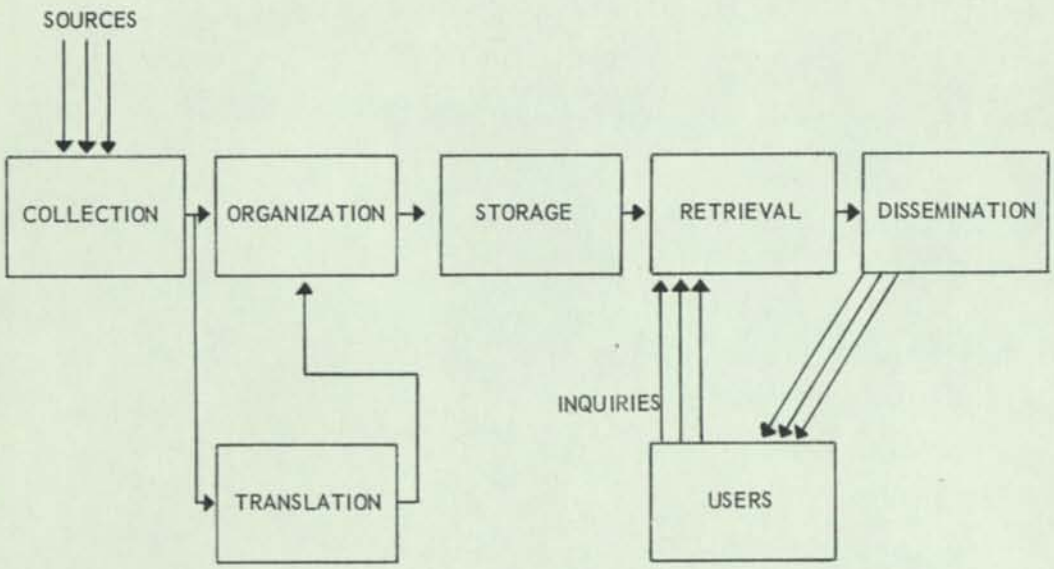
An inherent function of the central organizing and administering body is to establish a strong program of research. This research effort must probe into the origins, transmittal, and use of information. It must also develop the principles and mechanics of new devices for use by the Center. It is not intended that the Center itself conduct all or even a major part of such research and development programs, but rather that the Agency contract for and coordinate such efforts wherever they can be carried out most effectively.

The Center is expected to function as a coordinating and communications focus in its information processing operations rather than as a completely self-contained mechanism. Undoubtedly, a large volume of the operation will be conducted at a single location, but modern communication facilities make it unneccessary for the entire operation to be physically centralized. Furthermore, dispersed operation will permit exploitation of already existing capacities and thus make possible earlier and more economical operation.

The Center in its interim form will utilize large numbers of specialized personnel employing manual techniques similar to those presently employed by large technical abstracting services, such as that of the American Chemical Society and the Soviet Technical Information Institute. However, as the system evolves and as mechanized equipment is developed and becomes available, the operation should become increasingly mechanized. With this evolution the service should become speedier, more comprehensive, and more creative.

INFORMATION FLOW

The functions of the National Technical Information Center, in both its initial manual form and its ultimate highly mechanized development, can be represented by the following flow diagram:



A DRAFT PROGRAM FOR A
NATIONAL TECHNICAL INFORMATION CENTER

BACKGROUND

Man's most persistent medium for recording his progress and for communicating his ideas and discoveries to others is the written word. In the world of science this communication is formalized in periodicals, reports, and books. In recent years the rapidly growing volume of such publications threatens technical communications with a paralyzing glut and ultimate collapse. Inability to deal with the rising flood of technical publications has already seriously handicapped this country's industrial and scientific development.

Furthermore, increasing concern has been felt over the time for new products--particularly new military devices--to reach the manufacturing stage. History of any technical development from conception to the finished product shows that the major time interval lies between discovery of the basic principle and demonstration of its practical usefulness in the laboratory. This period may extend for decades. On the other hand, even the most involved development engineering seldom requires more than five or ten years. More time is lost--or more time is to be gained--between idea and demonstration than between demonstration and production. It is in this period, where ideas rather than physical materials are involved, that speedup is most feasible. Information has no inertia; it can be accelerated infinitely--subject only to the limits of our skill and facilities for transmission and correlation.

Not so!!

Existing information handling processes and library services are based largely on techniques developed more than fifty years ago. Recent research and development efforts have produced fragments of solutions at best. The continued vitality of scientific progress in this country demands a generalized solution to the information problem. In an era in which scientific and technological accomplishment is the measure of a nation's international stature, the problem is of compelling urgency.

MAGNITUDE OF REQUIRED EFFORT

The magnitude of an undertaking designed to process a major part of the world's technical information cannot be measured by any single parameter. However, certain figures emphasize the massive nature of the required effort. Informed estimates provide the following data on the world's ouptut of technical literature:

	Total	Annual Growth percent
Responsible technical journals	30,000 to 50,000	
Articles published annually	500,000 to 2,000,000	10 to 15
Technical books published annually	60,000	
Documents issued annually (U.S. Government only)	100,000 to 150,000	30 to 40

The heaviest users of technical literature in the country are the 200,000 professional research specialists who spend from 5 to 25 percent of their time seeking useful published information. The engineering profession makes much less use of such information by reason of its difficulty of access. The resulting duplication of development effort increases costs and injects unwarranted delays in product availability. Several thousand research libraries representing millions of dollars of annual operating costs and tens of millions of reference items utilize some 500 abstracting and indexing services in a valiant but futile effort to meet these informational needs. Coordination and integration of these services and operations would reduce the waste of money and scientific manpower.

In addition to struggling with the sheer volume of the literature, the American scientific community encounters an increasing complication in dealing with the technical information published in languages other than English. There is a serious lack of awareness in the United States of the growing significance of foreign technical publications. Limited data indicate that "pure" scientists refer to foreign sources (including those published in English) about 30% of the time; "applied" scientists 10%; and engineers even less. Apparently the information extracted from foreign language literature by U.S. scientists and engineers is very small,

yet this literature as a whole is probably expanding much more rapidly than that in English. Furthermore, the number of languages in which technical information is being generated is increasing.

No complete information processing facility exists in this country or abroad. However, beginnings toward this objective have been made by the Soviet Union through the Soviet All-Union Institute of Scientific and Technical Information founded in 1952. Its function is to accumulate the published technical literature of the world, process it and distribute abstracts, reviews, and compendia as quickly as possible. The full-time staff of 2,300 specialists is supplemented by 20,000 scientists and engineers who serve as abstractors and translators on a part-time basis.* The relative effectiveness of this Institute is attested to by prominent American scientists who have said that the best way to determine what American science is doing is to read the Russian abstract literature.

Soviet scientists have undoubtedly found their technical information Institute an important factor in the scientific race with the West. Only a few months intervene between publication of Western research information and when it is placed in the hands of a Soviet worker in his own language. Duplication of research and library effort is reduced, thereby permitting more effective use of the scarce and valuable trained manpower, and all investigators enjoy the stimulation resulting from access to the thinking of colleagues throughout the world.

The Soviet Institute does not appear to provide extensive facilities for systematic searching of the literature. It is known to have a strong and far-advanced research and development program aimed at creating mechanized search and retrieval devices. It is reported that the Soviet Institute has an operational English-Russian electronic translator. Furthermore, the Russian operation seems to limit its scope to the physical sciences, biology, and some branches of engineering; excluding medicine,

---

* An operation of this size in the U.S. would require an annual operating budget of about $60 million, excluding cost of journals and books; printing and distribution; and part-time abstractors and translators.

law, economics, civil engineering, and other valuable areas of information. However, having developed the procedures to handle some categories of technical information, presumably the Institute will ultimately extend its activities into additional subject fields.

METHOD OF APPROACH

The actions proposed below have two prime objectives, which must be pursued simultaneously. First, to meet the nation's most immediate needs, it is considered necessary that an interim technical information center be brought into being as quickly as possible to collect, process, and disseminate technical information. Second, to establish, following analysis and research, an ultimate National Technical Information Center, utilizing maximum mechanization for accelerating the flow of technical information from all parts of the world into the laboratories of the United States.

The success of the program depends inherently upon the coordination and continuity of the five actions outlined below:

1. Establish a central organizing and administering Agency.
   An Agency of the United States should be established with the responsibility and authority to make arrangements for the conduct of research, development, and operations relating to the interim and ultimate National Technical Information Center. The Agency will employ directly a small staff consistent with its policy making, contracting, and administrative functions.

   An extensive and continuing systems analysis of the nation's technical information problems and programs will provide the basis upon which the Agency will contract for new research and development projects. Privately supported research and development efforts will be encouraged when they are consistent with the Agency's long-range goals. Existing and new technical information projects conducted by other government agencies will be subject to the review and concurrence of this Agency. In this way all aspects of the national technical information problem will be dealt with in an expeditious, systematic, and well-integrated manner.

   Operation of the National Technical Information Center may be conducted under contract to the Agency. Such an operation will involve the coordinated participation of many supporting services, in conjunction with a central information processing and storage facility.

   Program details would be formulated by the Agency as the requirements develop.

2. Determine the gross dimensions of the problem.
   Before any steps can be taken to establish even an interim technical information center, it will be necessary to achieve a quantitative appreciation of the size and detail of the

work to be done and the service to be rendered. A survey to this end should be initiated promptly. The study would utilize known techniques, and would provide reliable estimates of the magnitude of various factors critical to entering into the interim operation. Among these are the volume of technical publications produced in various subject categories; the distribution of technical information sources by country and language; the general nature, frequency, and number of user inquiries; the numbers and types of personnel required to staff an interim center; the size of facility and types of equipment required.

Many organizations have already recognized the importance of the information problem and have undertaken studies of several of these factors. (See Appendix A.) These studies can contribute much valuable information for the design of the interim operation.

A large reservoir of facilities and experience in individual components of the information center's program is already in operation. Extensive but uncoordinated abstracting services exist. (See Appendix B.) In addition, an estimated 500-700 industrial research groups maintain abstracting services in their own limited fields of interest. Translating services are less fully developed, but some are in operation and probably could be expanded. (See Appendix C.) These and similar existing services will be inventoried and factored into the interim system to assure maximum economy and earliest possible operation.

These surveys can be completed in about six months if maximum practical concentration of effort is applied.

3. Establish an interim information center using existing services and techniques.

The nation need not await completion of extensive analyses or development programs before taking positive action to meet the immediate needs for better access to technical information. Immediate steps can be taken to establish an interim operation to perform at least part of the desired information processing functions using available techniques. The interim operation can provide at least translating, abstracting, and conventional indexing operations following procedures already developed by the major abstracting journals.

Unfortunately, the readily available techniques are largely manual. Fully developed machines are available only for small segments of the total operations. The available machines should be used in standard or modified form wherever possible. However, the interim operation will unavoidably require large numbers of specialized personnel. The supply of trained information specialists is small. Hence recruiting and training must be given careful thought in devising a plan for the immediate center if the plan is to be practicable.

The surveys outlined under action 2 will provide the basic parameters for the design of the interim center. Other studies of organization, operational analysis, machine feasibility, and short-term machine development programs will be undertaken as the interim center is brought into being.

The interim center will suffer from many inadequacies and inefficiencies, but an immediate partial solution to the nation's information problem will have been provided. Better organization and dissemination of technical information will significantly reduce the volume of literature which the individual scientist or engineer must use. Our research and engineering laboratories will have improved access to the findings of colleagues in foreign lands, particularly those lands which are beginning to contribute important technical and scientific discoveries. Furthermore, existence of a center in operation will provide an insight into the systems problem that would be difficult, if not impossible, to gain completely through theoretical analysis.

Fortuitously, we are assured that this interim approach is feasible. The initial operation of the Soviet Technical Information Institute was undertaken on a similar basis--in fact, Russian spokesmen have said that it was a conscious extension of the methods of the American Chemical Abstracts. The Soviet Institute is in operation and seems to be effective. Its research program of the past five years is reportedly now beginning to produce unique advanced equipment for the refinement and extension of the Institute's services.

An aggressively and realistically pursued program in the United States could result in a similar operating center within two years.

4. Analyze the factors which determine the design and operation of an ultimate National Technical Information Center.
There is a critical need to identify and evaluate the importance of the many factors pertinent to the design and operation of an ultimate National Technical Information Center. These include organizational and operational considerations. The organizational factor is particularly important for the encouragement of research necessary to implement successfully an ultimate information center, and it will be discussed separately later.

Human factors and machine factors constitute the two major operational considerations that must be systematically examined, and which will lead to a host of research problems. It is anticipated that systematic study of user needs and characteristics as well as research on the logic of information systems will have to be undertaken. Methods of classifying, abstracting, indexing, storing, retrieving and disseminating increasingly formidable amounts of information will have to be studied in detail. No scheme for accomplishing these

operations adequate to the needs of the scientific community is yet at hand.

It will be necessary to consider many machine factors and potential advances in the state of the art. No present machine is equal to the task of storing the immense quantities of information available. Even further from solution is the problem of achieving rapid access to a specific small fraction of a very large body of stored information. The complexity of the situation is further increased by the need for machine translation and organization. Partial solutions exist. Printed matter can be read by machines to a limited extent; machine translation from one language to another has been demonstrated on a small scale; high-speed handling of printed documents and turning of book pages are under development; high-speed printing mechanisms are now in operation. Major improvement in these machine capabilities is certain. When current machine performance is considered in the course of the system analysis, improvements will be identified which can be pursued in the research program.

Certain of the boundaries of the system analysis will be determined by policy decisions to be made by the governing Agency. Who will pay for the service and on what basis? What kinds of information will be handled? Who will have access to the information?

The systems analysis will provide a basis for an objective design of an information retrieval structure that can meet the nation's information needs with least time, money, and effort.

Experience gained from the interim technical information center will also contribute to the programming of specific research and development projects and to the design of the ultimate Center.

5. Encourage present and initiate additional research and engineering development programs leading to systems and equipment necessary to implement an ultimate National Technical Information Center.

The Agency will guide the total research effort. It will encourage and support necessary research and development programs, prevent duplication of effort, and integrate and coordinate pertinent activities being undertaken throughout the country. It should have control over the rate of development of the Center, primarily through its control of funding for research and development work.

This Agency will be responsible for aggressively thinking through and supporting an extensive and uninterrupted long-term research and development effort. The magnitude of this effort will determine the rate at which the ultimate Center is designed and becomes operational. The Agency will be responsible for obtaining the services of individuals and

laboratories possessing both capabilities and facilities for effectively contributing to the total program. Only through such an arrangement can an ultimate National Technical Information Center be established which will, in time, contribute markedly to the scientific and technical world leadership of this country.

BUDGETARY CONSIDERATIONS

This presentation provides only the broad outlines of a suggested program. Many refinements will have to be injected. Many details are ignored and many remain to be uncovered.

The intent here is primarily to indicate the magnitude of the program, something of the state of the art, and to suggest an approach to the solution of the compelling national need for better use of technical information. However, from these broad outlines and what is known about implementing programs of this size, order-of-magnitude estimates of immediate budget requirements can be made.

The compelling nature of the information problem suggests that funds be authorized immediately to pursue the initial investigation of the gross dimensions of the problem and to initiate the establishment of an interim center. Subsequent authorization of funds should also be sought to ensure an aggressive and continuing program. In addition, it may be assumed that private enterprise will continue to invest large sums in research and development, particularly on equipment necessary to mechanize information processes.

For the remainder of fiscal year 1958 a supplemental authorization of $2.0 million would meet the estimated needs of the governing body to support $500,000 for surveys and initial investigations, and $1.5 million to begin procurement of equipment and facilities to accommodate the interim technical information center.

It is too early to provide a definitive estimate for the fiscal year 1959 budget, but obligational authority should be sought for $50 million. Included in this figure are support of short-range research and engineering development necessary for the equipping and operation of the interim facility and for the beginning of the long-range systems analysis, initial staffing and operation of the interim center. Support of certain long-range research and development programs of obvious pertinence to any system for the ultimate National Technical Information Center should be initiated during this period. It is questionable whether an optimum long-range research program could be attained within the figure given here, but a valid estimate is impossible at this time.

For the fiscal year 1960 and subsequent years the annual budget can be expected to exceed $100 million. The upper limit will be determined by the aggressiveness with which the research and development effort is pursued. It will also be affected by the extent to which the operation of the National Technical Information Center will be made self-supporting through charges for services. Such income can be substantial in view of the large present expenditures of government and industrial organizations for individual services which may be relieved by the operation of the Center.

CONCLUSION

The technical information needs of the nation can be met by an intelligently conceived and earnestly pursued research, development, and operational program. Such a program will be difficult--it will be costly. However, it is neither more difficult nor more costly than other systems development programs which the United States has undertaken and completed to achieve and retain our internal and international status. The cost in lost prestige and technical stagnation which may result from failure to solve this problem might well prove catastrophic.

Stanford Research Institute has prepared the Draft Program for a National Technical Information Center as a public contribution to the solution of a national problem which it believes to be of critical importance.

The Institute's experience in systems design; operations research; information organization; and in the conception and development of information processing devices, we believe, provides a unique perspective into the complex scope of this problem.

The Institute maintains an active and expanding program of research in all phases of the information processing problem. The Institute invites consideration for a role in planning and coordinating, as well as execution of appropriate research phases of any national technical information program, such as the one outlined herein.

APPENDIX A

A PARTIAL LIST OF ORGANIZATIONS CONDUCTING RESEARCH AND DEVELOPMENT IN INFORMATION PROCESSING

All-Union Institute of Scientific and Technical Information (USSR)
Association of Special Libraries and Information Bureaux (England)
Battelle Memorial Institute
Birkbeck College (England)
Burroughs Corporation
California Institute of Technology
Cambridge Language Research Unit (England)
Case Institute of Technology
Chemical Abstracts Service
Columbia University
Documentation, Inc.
Dow Chemical Company
Eastman Kodak Company
Federal Telecommunications Laboratories
Eugene Garfield Associates
Georgetown University
Computation Laboratory of Harvard University
Herner, Meyer and Company
International Business Machines Co.
International Telemeter Corporation
Lehigh University Library
Librascape
Arthur D. Little, Inc.
Low Temperature Research Station (England)
Magnavox Co.
Massachusetts Institute of Technology
National Academy of Sciences-National Research Council
National Bureau of Standards
National Cash Register Co.
North American Aviation Inc.
Ramo-Wooldridge Corporation
RAND Corporation
Rutgers University
Sperry Rand Corp.
Stanford Research Institute
Teleregister Corporation
U. S. Patent Office
University of Michigan
University of Pennsylvania
University of Virginia
University of Washington
Western Reserve University
Zator Company

APPENDIX B

A PARTIAL LIST OF TECHNICAL ABSTRACTING PUBLICATIONS*

Accountant's Index
Acoustical Society of America
Acoustical Society Contemporary Papers, Journal
Aeronautical Engineers Index and Review Abstracts
Aeronautics, Index
Agricultural Index
Agriculture, Bibliography of
Agronomy Abstracts
Air Pollution
Air University Periodical Index
Allergy and Applied Immunology, Quarterly Review
Alloy Digest
Aluminum Abstracts Bulletin
American Concrete Inst. Proc.
American Documentation
American Journal of Nursing
American News of Books
Analytical Abstracts
Analyticus Cancer, Index
Anco
Anesthesia and Analgesia Current Index
Animal Breeding
Annotated Bibliography of Economic Geology
API Technical Abstracts
Applied Chemical Abstracts, Journal
Applied Entomology, Review of
Applied Mechanical Review
Applied Physics, Journal of
Applied Spectroscopy
Art Index
Asian Studies, Journal of
ASTIA Card Service
ASTIA Unclassified Reports
Astronomical Newsletter
Atomic Energy Reporter
Battelle Technical Review
B.C.U.R.A. Bulletin
Best's Bulletin Service
Bibliographic Index
Biography Index
Biological Abstracts
Biological Science, International Abstracts of
Book Review, International
British Abstracts
B.S.E.A. Abstracts
Building Science Index
Business Service Checklist
Buttersworth Card System

Canada, Bibliography of
Canadian Index
Canadian Patent Office Record
Cancer Current Literature
Card Service, Advance Abstracts
Ceramic Abstracts
Ceramic Abstracts, British
Chemical Abstracts
Chemical Abstracts, Analytical
Chemical Abstracts, British
Chemical Engineering News
Chemical Literature
Chemical Market Abstracts
Chemical and Physics Abstracts, British
Chemical Spotlight
Chemisches Zentrablatt
Child Development Abstracts
Cinnotalid Bibliography of North American Geology
Classified Reports (ABC), Abstracts of
Computers and Automation
Corrosion Abstracts

---

* From a survey by Western Reserve University

Corrosion et anti-Corrosion Abstracts
Corrosion Engineers Abstracts, National Society of
Crerar Metal Abstracts
Crippled Children Bulletin of Current Literature, National Society for
Cumulative Book Index
Current Chemical Papers
Current Technical Literature, Index to

Dairy Science Abstracts
Data
Dental Abstracts
Dental Literature, Index to
Dirivent Patent Service Dissertation Abstracts
Distillation Literature and Abstracts
Document Index
Dyers and Colourists, Journal Society

Eastman Kodak
Eastman Kodak Library Reports
Education Index
Electrical Engineering Abstracts
Electronics Abstracts
Engineering Index
Engineer's Digest
EPA Technical Digest
Essay and General Literature Index
European Technical Digest
Excerpta Medica

Fats, Oils, Detergents Abstracts
Field Crop
Files Coping the Technical News
Fisheries, Commercial, Abstracts
Fisheries Abstracts, F.A.O. World
Food, Current Abstracts
Food Science Abstracts (British)
Food, Index and the Literature Journal of Food
Forestry Abstracts
Fuel Abstracts

Gas Abstracts
Gaylor's Technical Service
General Electric Internal Bulletins and Reports
Geological Abstracts
Geographical Publications, Current
Geophysical Abstracts
Glass Technical Abstracts, Journal Society
Government Classification Cards
Great Britain Dept. of Science and Industries Research, Translation of Russian

Herbage Abstracts
Highway Research Abstracts
Historical Abstracts
Horticultural Abstracts
Hosiery Abstracts
Hospital Abstract Service
Hospital Literature, Cumulated List of
Hospital Literature, Index to Current
Hygiene, Bulletin of

IGT Abstracts
IMM Abstracts
Industrial Arts Index
Industrial Diamond Bibliography
Industrial Health, Achives of
Industrial Hygiene Digest
Industrial Labs
Information Service Bulletin
Ink Maker
Instrument Society of America
International Index
International Institute for the Conservation of Museum Objects
Interplanetary, British
Iodine Abstracts and Reviews
IPC Library Notes
IRE Abstracts
Iron and Steel Institute Abstracts, Journal

Labor--Personnel Index
Lectrodex
Leukemia Abstracts
Library Bulletin of Abstracts
Library of Congress Air Pollution
Library of Congress Bibliography of Translation from Russian
Library of Congress, Books; Subjects

Library of Congress Central Europe Accession
Library Literature
Light Metals Bulletin
Literary Notes

Magnesium Review and Abstracts
Management Abstracts
Market Abstracts
Material Construction, Review of
Mathematical Abstracts
Medical Abstracts
Medical Digest, International
Medical Literature, Current List
Medical Science, International Abstracts of
Medical and Veterinarian Mycology Review of
Medical and Veterinarian Zoology, Index-Catalogue of
Medicus, Index
Medicine, Modern
Metal Finishing
Metal Literature, ASM Review
Metals Journal, Institute of
Metallurgical Abstracts
Metallurgical Abstracts (British)
Metals Review
Meteorological Abstracts
Military Affairs
Milk and Milk Products, Abstracts of Literature
Mond Nickel Bulletin
Monitor

National Defense
NACA Research Abstracts (Card File)
Natural and Synthetic Fibers
Naval Research Laboratories
Neurology and Psychology, Digest of
New Testament Abstracts
New York Times Index
Non-Destructive Testing
Non-Ferrous Metals, British
Nuclear Science Abstracts
Nutrition Abstracts
Nutrition Reviews
NYSEM Bibliography of Electr. Micr.
Obstetrics and Gynecology Survey
Official Gazette of U.S. Patent Office
Olin Mathieson Chemical Corp.
Opthamalic Abstracts and Research

Packaging Abstracts
Paint, Colour, Review of Current Literature Relating to
Paint, Varnish and Lacquer Assoc., British
Paint, Varnish and Lacquer Assoc. National
P.A.I.S.
Paper Chemists, Bulletin of the Institute
Paper Making and U.S. Patents, Bibliography of
Pasteur, Bulletin Institute
Patent Journal (Union of S. Africa)
Patent Research Office Patent Lists, International
Patents Abstract Journal (British)
Pediatric Current Literature, Abstracts from
P. B. Reports, Index to
Personnel Administration
Personnel--Management Abstracts
Pesticides Abstracts
Petroleum, Journal of the Institute (Abstract section)
Petroleum Technology
Philips Technical Review
Philippine Periodicals, Index to
Photographic Abstracts
Physics Abstracts
Plant Breeding Abstracts
Plant Disease Reporter
Plastics Fed, Abstracts, British
Plastics Technology
Poliomyelitis Current Literature
Prevention of Deterioration Abstracts
Printing and Lithographic Abstracts
Psychological Abstracts
Public Affairs Information Service Bulletin
Public Health Engineering Abstracts
Publisher's Weekly

Quality Control and Applied Statistics
Quality Control and Industrial Statistics Abstracts
Quarterly Cumulative Index Medicus

Radio Engineers, Institute of
Reader's Guide
Refrigeration Abstracts
Rehabilitation Literature
Report of NRL Progress
Research Information Service
Resins, Rubbers, Plastics
Revue d'Aluminum
Road Abstracts
Rubber Abstracts
Rubber Formulary
Russian Accessions, Monthly Accessions

Schnelidienst--Literature
Science Abstracts
Science Instr. Research Assoc., British
Science and Technical Papers, Abstracts of
Science and Technical Reports, Bibliography of
Semiconductor Electronics
SIPRE Bibliography
Sociological Abstracts
Soil Science, Bibliography
Soils and Fertilizers
Solid Propellants
Special Library Consultants
Specifications and Rel. Pub. Air Force, Index of
Specifications and Standards, U.S. Army
Specifications and Standards, U.S. Navy
Spectrographic
Stanford Handbook
Statistical Methods in Industry, International Journal of
Subject Index to Periodicals
Subject Index to Periodicals, British
Sugar Industries Abstracts
Surgery, International Abstracts of
Surgical Digest, International
Surgery and Gynecology, Quarterly Review
Synthetic Liquid Fuel Abstracts
Synthetic Methods of Organic Chemistry (Theilheimer)

Tappi
Technical Abstracts
Technical Book Review Index
Technical Digest Service (new)
Technical Journals
Technical and Electrical Industries-File
Technical Reports
Technical Survey
Textile Industry, Journal of the
Textile Technical Digest
Textracts
Tobacco Abstracts
Toilet Goods Association
Translated Contents List of Russian Periodicals
Translation Monthly
Tropical Disease Bulletin
Tuberculosis Index and Abstracts

United Nations Documents Index
Uniterm Index of Chemical Patents
Universal Oil Products Abstracts
Unlisted Drugs
U.S. Bureau of Census Cataloguing
U.S. Dept. of Commerce (Office of Technical Service)
U.S. Government Monthly Publication Catalogue
U.S. Government Research Reports
U.S.D.A. Medical and Veterinarian Zoology
U.S.D.A. Soil-Water Conservation
U.S.D.A. Fish and Wildlife

Vacuum Abstracts
Veterinary Bulletin
Veterinarius Index
Vide
Vitamin Abstracts

Water Pollution Abstracts
Weed Investigation, Bibliography of
Welding, Bibliographical Bulletin for
Welding Journal, British
Welding and Metal Fabrication
Wilson Indexes
Wireless Engineer's Abstracts
Wistar Institute
World Medicine, Abstracts of
World Veterinary Abstracts Journal

Z.D.A. Abstracts
Zoological Record

APPENDIX C

A REPRESENTATIVE LIST OF PUBLIC AND COMMERCIAL TECHNICAL TRANSLATING SERVICES

American Institute of Biological Sciences - Publishes four Russian biology journals in complete English translation.

American Institute of Physics - Publishes eight Russian physics journals in complete English translation.

American Mathematical Society - Publishes annual volumes of Russian mathematical papers.

Associated Technical Services - A technical translating service providing material in 25 languages. Also issues periodic lists of articles translated in chemistry, physics, electronics, medicine, etc., predominantly Russian, but including many languages especially among Soviet countries.

Berlitz Translation Service

Columbia Technical Translations - Publishes Physics Series of the Bulletin of the Academy of Sciences of the USSR.

Consultants Bureau, Inc. - Publishes 19 journals in the fields of chemistry, physics, electronics, geology, biology. Also issues translated tables of contents free and from time to time publishes collections of important papers on subjects of current interest.

Engineering Societies Library

Far Eastern and Russian Institute

Henry Brutcher, Technical Translations--Issues monthly list of translated articles available, with abstracts, from Russian, Polish, Japanese, German, and occasionally other languages.

Institute of the Aeronautical Sciences

International Physical Index Inc. - Translates, abstracts, and indexes over 60 leading Russian electronics and automation journals monthly. Complete bibliographic and translation services covering Russian language publications of the physical sciences also available.

Morris D. Friedman, Russian Translations - Issues frequent lists of translated articles.

Pergamon Institute - Publishes an English translation of the Russian journal "Elektrichestvo," and translates three leading Russian journals in the fields of radio engineering, communications, and electronics. Also translates and publishes Russian books in English. (Pergamon Press)

Primary Sources - Translates and publishes selected Russian technical papers from four Soviet journals dealing with research in metallurgy--to be published quarterly beginning with first issue (Quarterly Review) March, 1958.

The Technical Information Problem

authors

The attached list is composed of questions which are designed to present as many of the considerations as we could collect from miscellaneous sources in a limited time, which must be taken into account in the planning and design of any large information system. The primary emphasis was placed upon considerations especially pertinent to an interim National Technical Information Center (NTIC).

The primary objective in preparing and distributing such a list as this is to provide a perspective view into the nature of the problems to be expected in planning, designing, and operating a NTIC. These questions (considerations) are representative of those which will have to be faced, and they should serve to give some appreciation of the nature and the magnitude of the job. Auxiliary benefit in providing stimulus and orientation for research in this area is expected, also.

These "considerations" are grouped in rough categories according to the topic outline given below. Designation to a particular category in many cases is quite difficult, owing to the interaction of many of the aspects of the total job, and so final placement was in some cases rather arbitrary. We have each of these "considerations", as well as a miscellaneous group of facts, etc., written one each on edge-punched cards which are notch-coded in both the categorical division outlined below and a multi-aspect subject code. Regrouping of these "considerations" to meet special desires should be relatively easy, and the cards are available to those who might wish to do so.

Outline for "considerations" listing

1. User requirements and characteristics.

   Can we separate apparent need, influenced by present concepts, from real needs which can be met with an end system?

   In defining the variety and scope of information requirements,

it is dangerous practice to take too literally statements that may be made by the clientele to be served. Such statements can be expected to reflect their expectations of what can be accomplished and such expectations are usually limited by their experiences with conventional procedures. Lack of awareness of the potentialities of recently developed methods can easily result in an unimaginative formulation of the possibilities and opportunities for advantageously using recorded information. How will users' habits and needs evolve as a good system becomes available?

What type of questions go unanswered at the libraries?

What is the role of information retrieval, storage, etc. in the decision-making process for such as the research worker, scholar, administrator..etc?

What increase in efficiency would result from improving the accessability of recorded information?

By what processes does the scientist keep abreast of the advances in the art? What are the relative importances of each of these processes?

How many scientists work without adequate access to a library or document center?

How much use does the scientist make of the facilities that are presently available?

How are the information needs of a scientist effected by his age, educational level, profession, type of position held, etc.?

What are the relative merits of the different types of information services with regard to the scientist, and his needs, desires, habits and limitations?

How many scientist schedule, or regiment, or have a definite program of "keeping up with the literature"?

How much time per week does the average scientist spend now in "keeping up with the literature"? How much time would he "like to spend"? What keeps him from spending more time?

How much time would it take the scientist to maintain surveillance of the literature which would be adequate to insure that he catches an optimum percentage of the useful literature? The time requirement would seem to vary widely with people in different levels and in different areas.

How much of the literature which would, with reasonably high probability, be useful to a scientist, is caught by him now by his own regular surveillance of the literature?

How far out of his way will the average scientist go to be sure that he hasn't missed some possible information..considering the different types of information he needs, the usual distracting pressures on him, his familiarity with the sources, etc.?

What are the relative importances with regard to this "information problem", of the different aspects of the scientists' informational needs? On one hand, a scientist needs to know the newsy items such as who is working on what, what his current attack is, who disagrees with whom and basically why, etc., and on the other hand he also needs to be able to _____ the carefully-written treatises which may have bearing on his work, and which serve to build this theoretical structure.

Does the scientist in a given field perhaps want, when he goes outside his special field for supporting information, information in different form or different level than which he seeks in his own

field? For instance, would he be looking more for "cookbook" techniques, or for survey-type information?

How many pages of literature, in various categories relative to the level and interest-area of the scientist, can we expect a scientist to be able to scan or search for his different information needs? What happens to his efficiency, his attitude.. how much time does he have for that sort of job, etc.?

What are the special information requirements for different specialty fields?

Is any particular type of publication any more efficient or useful for information purposes?

If abstracts are desired, what is the best form? (crticial, telegraphic, informative)

What is the distribution of document age vs. demand, for each of the specialty fields?

If NTIC services are to be saleable to researchers, how much of the past literature will have to be included?

What users possibley prefer to receive periodically, instead of the usual bound journal of restricted papers, a larger list of abstracts from which his journal subscription may allow him to choose a selection of papers which suits him most?

How much need is there for foreign literature?

Are there projects or areas of work which are reported almost exclusively in foreign literature?

Can a scientist of each discipline expect to be able to scan all of the periodical literature in which he would find directly pertinent papers without taking a disproportionately large amount of time, or without having to go through too much subject matter which is not of interest to him?

How many places does a scientist of each discipline have to look for index listing of a given special interest topic in order to be reasonably sure that he has not missed too many useful documents?

What type of information should be contained in the NTIC?

books (texts, tables)
Technical and trade journals
conference proceedings
papers presented at conferences but not published
industrial and government interim and final project reports, etc.
(similar to the type handled by ASTIA)
operation and instruction manuals
patents
manufacturer's catalogs
abstract publications
maps
films and photographs
newspapers
magnzines

How much delay can the user tolerate?

What is the main problem - the time delay, or the fact that many documents just won't reach the reader?

How many of the problems are problems because people might not know of present techniques which could help them, and how many are beyond the ability of present techniques to help?

How can the information services be made more attractive to the users?

→ Insert

2. a. Literature

What potential audience is currently exposed to the abstracts, of articles which appear in each of the particular journals? (e.g., abstracts of an article in the IRE may be published in 8 different publications. What is the combined circulation of the people who could see this abstract?)

For each serious technical journal in the world,

how often is an issue published,
what is the subscription cost,
who publishes it, and where,

Insert 2. Present dissemination system

how big is each issue, generally, (no. papers, text pages)
what subject field or fields does it cover,
how big is its circulation,
who are its readers,
in what language, principally, is it written,
who, principally, are its contributors,
how it it supported,
what is its usual backlog of papers, and what is the usual publication delay for a paper
does it have an annual index,
do all of its published papers include abstracts,
where do abstracts of its articles appear,
how many different abstracts are written for its articles,
where are its papers generally indexed,
what is its relative nature of its coverage within its subject field or fields,
who reviews or selects the papers for publication,
what is its financial health,
how many pages of advertising compared with pages of text,
what are its format characteristics,
what would be the physical dimensions and weight of a years' publication, bound and unbound,
into what languages it is translated,
does anyone provide microfilm copy of it as a regular service,
how long has it been publishing,
how many papers, total pages, have been published in its life,
what are its publishing policies, objectives,
could the publisher provide paper tape copy (punched) without duplicate typing,
does it publish a paper if the information is known to have been published elsewhere,
and what would it charge the NTIC for its needed copy or copies.

What are the summaries of the above single-journal questions for all of the pertinent journals?

For each technical-book publisher,

how many different books does he publish in a year,
what does their size average,
what fields does he specialize in,
what is his book distribution over different specialties or fields,
~~what is his book distribution~~
does he solicit writers, and among whom,
what would he charge the NTIC for its needed copy or copies,
what is his financial health,
average sale price of his books,
does he publish paper-backed books,
how cheaply ~~can~~ could he publish his books,
how many books does he sell in the different topical categories (sales figures per book), and to whom,
how much does it cost him to publish his books,
in what language or languages does he publish,

How long has he been publishing in the different technical fields,
How many books in the different fields and categories has he published to date...how many of these does he feel are out of date,
what are his publishing policies and objectives,
could he provide paper-tape copy (punched), or magnetic-tape copy, without duplicate typing,

For each organization that publishes technical reports,

how many does it publish,
in what special fields,
in what language or languages,
at what technical level,
does it provide abstracts for its reports,
who else might write abstracts of its reports,
where do abstracts of its reports appear,
how many different abstracts for it reports are written,
where are its reports generally indexed,
to whom are the reports circulated,
how many copies of each report are published,
how many pages are the reports,
who finances (or what finances) the publication of the reports,
what is the financial health of the organization,
what are the format characteristics of the reports,
how long has it been publishing reports,
how mang reports, total pages, total weight, has it published altogether,
what would it cost NTIC to obtain needed copies of their reports,
are back copies available,
into what language or languages are his reports translated,
does anyone provide microfilm copy of his reports as a regular service,

How much recorded information is available for utilization?

How many book, journals?
In how many fields and in what languages?
What is the distribution of quality of this material?
To what class of user is the material directed?

What journals are not covered by an abstracting service?

What journals have duplicate coverage by an abstracting service?

2. b. Services

For each library which directly services research scientists,

what types, numbers, and fields of coverage do you have in your collected documents,
what is the normal time delay between publication of various kinds of documents and integration into your system,
who finances your operation,

what are the types and numbers of scientists served by you,
what types of services do you provide,
what is the number and training of your staff,
what classification and indexing systems do you use,
how much traffic could you handle,
what suggestions would you have regarding improved retrieval and dissemination of scientific literature,
are you especially well qualified in any particular field of scientific literature,
would you be capable of and willing to do research in the field of scientific documentation,
what complaints to your scientific ~~enx~~ customers have,
do you have facilities for making microfilms,
how many microfilm readers do you have,
how much use do you make of microfilmed documents,
what are your comments regarding wide usage of microfilm documents,
how do you select your acquisitions,
what journal, abstracting ~~xd~~ services, and ~~xxxnx~~ translating services do you subscribe to or utilize,
are you set up to ~~ahdnle~~ handle classified material, and do you handle it,

For each organization which regularly publishes abstracts of papers that they themselves do not publish,

who prepares the abstracts,
what sources do you cover completely with abstracts,
what other sources do you cover with abstracts,
what kind of abstracts to you publish,
how soon after primary publication does your abstract get published,
in what language or languages are your abstracts written,
how much does it cost to prepare your abstracts,
how long does it take (actual working time) for your abstractors to abstract typical papers,
how frequently do you publish your periodical abstracts,
how many copies do you publish,
to whom do you ~~dkx~~ distribute your abstracts,
do you make special compilations of abstracts for customers,
what are the regular and special costs of your services,
do your service fees pay for the operation,
what are your policies of growth, coverage, service, etc.,
would you be amenable to negotiate a change or increase of area of coverage, or other changes in your service, to fit a reasonable over-all NTIC plan, if government controlled and subsidized,
what cataloging scheme do you use,
what is the size and training of your staff,
do you treat classified material, and how is it handled,

For each organization which regularly published indices to literature which they do not themselves publish,

who classifies the index entries,
what sources do you cover completely,
what other sources do you cover,
do you publish abstracts as well, and if so how are they integrated with the indices,
how soon after primary publication does an index containing reference to the item get published,
in what language or languages are your indices published,
into what language or languages are they translated,
how much does it cost to prepare these indices,
how much would your service cost to the NTIC,
how frequently do you publish your indices,
how many copies do you publish,
to whom do you distribute your indices,
do you make special indices for customers,
how are your operations financed,
what are your policies of growth, coverage, service, etc.,
would you be amenable to negotiate a change or increase of area of coverage or other changes in your service, to fit a reasonable over-all NTIC plan, if government controlled and subsidized,
if you publish abstracts, who does the abstracting,
do you make special searches for customers,
how healthy is your financial situation,
what is the size and special training of your staff,
do you treat classified material, and how is this handled

For each organization which provides scientific translation services,

between what languages do you translate,
what technical specialties do you translate, if any,
to whom do you distribute the translations,
what are your service fees,
what material do you usually translate,
do you do contract work, subscription work, or both,
who does your translating, what are their backgrounds,
what kind of a backlog do you usually have,
how much service could you provide,
how long does it normally take for a special-request translation,
what other languages and scientific areas could you give service in if need be,
are you equipped to handle classified material,
would you be amenable to negotiate a change or increase of area of coverage or other changes in your service, to fit a reasonable over-all NTIC plan, if it were government controlled and subsidized,
do you do abstracting as well as translating,

For each university with scientific staff,

what size and coverage does your scientific library have,

what number and type of scientific personnel within your university,
what number and type of scientific personnel must your information facilities service,
what is the number and training of your library staff,
are you set up to handle classified material,
what abstracting, indexing, and translating service do you utilize,
what are the service complaints of your scientific users,
do you utilize microfilmed documents,
how many microfilm readers do you have,
do you have the facilities for making microfilms, at what rate,
do you use ASTIA,
what is the usual time delay between publication of a document and its normal availability in your library,

For representative companies which utilize professional technical personnel,

what special information needs do your people have,
what services do you provide them with, your technical people
to what extent do ~~make use of outside facilities~~,
what number and type of professional technical people do you have,

If a private consultant were to work on a government project for which "need to know" could be established, how would he locate and procure pertinent classified material?

What professional organizations or publications provide any special information services? (e.g., edge-punched card, microfilm)

What technical libraries are currently staffed to handle a large number of information requests? (e.g., WRV, Batelle,. . .)

Why aren't the listed group of centers which serve to process technical information satisfactory?

What does ASTIA do?

Are there any useful information services which aren't being provided now?

Is there any historical data which could help to judge (ahead of time) the relative gain to be had in efficiency of different documentation services by integration, merging, or ~~better~~ better managing?

Do you know of any instances in which companies with similar interests cooperatively operated an information center?

What technical societies could cooperated to publish a single journal instead of numerous splinter journals?

What economic and political limiting factors exist with respect to the freedom one would have in ignoring or utilizing or changing those organizations which are at present already active in the documentation field, and whose existence will be overshadowed by the national center?

What are the total number of abstracting services in each language?

In cases where abstracts of a particular article appeared in several publications, how many of these abstracts were only copied, and how many were written by different people?

Of the currently-operating abstracting services, how many are operating merely to satisfy an obligation of a professional society and would rather have somebody else do the abstracting?

How many part-time voluntary abstractors are presently working for the professional societies, and what is their average monthly production and the average time spent per abstract?

What is the time interval now between the publication and the arrival into the possession of an interested researcher of pertinent documents?

Are there enough Russian-speaking scientists, or qualified people, to translate Russian scientific literature right now?

2. c. Techniques

What are the relative costs and advantages of the various different printing styles and techniques, and of the various kinds of paper, which might possibley be applicable to the dissemination of

scientific information (e.g. such as paper-backed books, newspapers, conventional journals, etc.)?

What are the relative costs and characteristics of different reproduction techniques which might be applicable to some of the dissemination and massive processing problems of a NTIC?

What special precautions must be taken to store primary records? Cost figures for the storage alone if these primary records are to be kept.

How long will microfilm stand up? Will a relatively seldom-used main store of microfilm records last as long as we might wish?

How long would microflim in use last? How often will we have to make new copies from the master?

What possibility is there that user's will accept microfilm or microprint copies of the literature which the system will have retrieved for them?

What improvements need be made in microprint or microfilm handling and viewing equipment in order that this media might find user acceptance?

What are relative costs of providing the information in microfilm or microprint form as against making photo-copies at original size?

What seem to be the major limitations in the various methods presently used in classifying and indexing scientific literature?

What are the techniques and costs involved in keeping up and in using large mailing lists in taking care of distribution of journals, etc.?

What is the present state of our methods, systems, techniques, equipment, and procedures for meeting modern information requirements?

3. NTIC Justification

a. Direct Questions

What are the net benefits, short and long range, likely to be among military, industrial, commercial, scholarly, Government activities, of having an effective NTIC through the years?

Can you prove that the expense of delay and duplication now is greater than that of establishing and operating your NTIC?

What is (it) costing us in appropriation money not to have an NTIC? Can we estimate how much savings in Federal money will be available (by NTIC) to spend on NTIC?

What facts and evidence can be shown to prove that duplication and delay have existed because of the lack of a reasonably good scientific information dissemination system? Can we derive some dollar costs for reasonably well proven delays and duplications, and can we make a realistic estimate as to your total national loss rate due to this ~~xp~~ problem? Can we put a value of some sort on the "tactical" worth ~~a~~ to us of reducing these delays and duplications by an amount which we could reasonably expect to be derived by the establishment of a NTIC?

What services will be provided that are different and better than what is now available? Is this to be entirely new types of services, a real advance in the state of the art, or is (it) to be ~~must~~ just more and better of the same thing?

Has it been tried before? What kind of real benefit can you say will accrue from the establishment of NTIC? Can you back this up with figures?

How about some more explicit data about the Russian center? What services do they really provide? Have you ever asked a Russian

scientist how much good this might do him?

Can't we do it on a smaller scale?

Does it have to be a big institute or nothing?

What technical-manpower drain would the proposed NTIC program have on our other high-priority scientific programs?

Why is so much research needed?

What research costs are justifiable?

What type and depth of research is needed?

What would you do now if you received a grant of $_____?

Certainly some work on this problem has been done in the U.S. before now. Who has done it? and how much have they done?

Why wasn't this proposed five years ago?

Isn't this exactly what ASTIA has been doing? What is the difference between what you propose and ASTIA?

Is our problem that the information now is just not available at all, or is it that it is just a little hard to find?

Just why do you think we need an NTIC?

Is this really a significant thing to do - will it be popular, meangingful, or is it just another proposal by an outfit that wants to get itself funded?

Is the taxpayer going to be paying for something that will give more advantages to the big-corporation research centers then to the small businesses or individuals?

How could a good national informations system help a Congressman in his legislative work, or in his political work? How could it have helped him before he entered politics, in the practical life that he knew as a business man, lawyer, etc.?

With reference to you "Draft Proposal for a NTIC", you say that

we are using "processes and services that are over fifty years old." Surely something new has been added in the last fifty years. What processes and services have our ASTIA appropriations been buying?

How important is it to know what the rest of the world is doing? English is good enough for me, why isn't it good enough for the scientist?

Just how might this be helpful to my home community?

Are we just trying to copy Russia in establishing this center?

Is this just a face-saving move up to keep/with Russia?

b. Alternative Suggestions

Is there some other way of doing it?

What if we gave $60 million a year more to ASTIA, wouldn't this solve lots of our problems?'

How many retrieval problems could be solved by teaching the researchers more about present-day documentation techniques?

How much is missed by their leaning too heavily upon librarians?

How much more is the push-button system going to give the researcher in the way of valuable information, then he could have found himself after a reasonably exhaustive personnal search through up-to-date index systems and abstract compendia?

How much can be done just by concentrating on abstract distribution and better dissemination techniques?

Can't the implementation of the solution for this be done right within Government agencies, e.g. Bureau of Standards and the Library of Congress, working together?

Why not allocate federal money to support more direct interchange between working scientists? More meetings, special conventions, seminars. Perhaps this would be a more effective say to spend money than for automated retrieval?

Couldn't the money be better spent on education, instead, to achieve a given increase in scientific effectiveness?

For an interim solution, can we translate the exhaustive Russian abstracts and provide distribution of these for our own people so as to leave our interim energies more free for other uses?

Couldn't a group of smaller centers, for specific fields, be of greater utility - more tractable?

How about utilizing, instead of a giant national center, a collection of special libraries?

How are we sure that it wouldn't be better to reduce the number or amount of literature rather than go to the tremendous expense of providing super-service for all of it? Why can't we put a quality filter on this output, weeding out what won't have significant future value?

4. NTIC System Design

a. Policy

Who will have the final say in establishing policy in the matter of designing, establishing, and/or operating the NTIC? Could it be appointed committee, such as for the NACA, a civil servant, or a ~~pix~~ political appointee, or a committee elected by scientific organization..?

What are the specific aims of the program?

What type of information should be contained in the NTIC?

books (texts, tables)
technical and trade journals
conference proceedings
papers presented at conferences but not published
industrial and government interim and final project reports, etc. (similar to the type handled by ASTIA)
operation and instruction manuals
patents
manufacturers' catalogs

abstract publications
maps
films and photographs
newspapers and magazines

Who will be responsible for selecting the actual material to be included in the NTIC system?

Who do you feel is competent to design, establish, and/or operate the NTIC?

Would this be a ~~xic~~ civil service organization?

How far ahead should we set the jumps between an interim and a "final" system, so that the final isn't just a modification of the interim?

What factors are going to determine the location? Can strategic dispersal considerations influence the location without adversely affecting efficiency?

Who is going to fix the priority schedules for the system?

Who is going to fix the price of the service?

Would it be feasible to obtain legal powers to speed up the standardization and coordination of existing facilities? (Such as the F.C.C.)

Will this center also service requests from allied countries? To what extent? Will this center try to cooperate in any way with the Soviet Institute? Will we ~~xi~~ adopt and/or promote an international classification, indexing, or retrieval system?

Will you merely ~~x~~ convert new information into the system, or will you go back and put in all the old information?

What provisions will you make for the needs of industrial researchers who would want the details of their queries to remain confidential?

If translating is to be done, why not have it done in Europe?

Would a duplicate file and collection be maintained to prevent disruption of service due to fires, etc?

The over-all organization which has responsibility for establishing the NTIC should be very carefull not to try to freeze the specs which are given to the systems people until enough systems work has been done to make it clear what would be optimum.

This of course extends all the way down the line, between systems and hardware people, etc.

Should make sure that the final methods chosen for a retrieval system are not influenced too heavily by the requirement of compatibility with past systems.

Your system will have to be useful for both the brilliant, sophisticated person, and the more unsophisticated person.

Is big business going to have any better access than small businesses or individuals?

Would a private citizen (scholar) be able to afford the service?

4. b. Operation

What is the range of subject matter to be included?

What is the expected range of inquiries to be answered, or services to be provided?

Should a review board screen the subject fields and select the most significant papers from that year's crop, and only pass these papers to the retrieval system?

Will you make any effort to review old documents, and to remove or recode when necessary?

Are you going to record everything or will there be an editor--or a censor?

Is there going to be any attempt to standardize the form of the material before it gets into the center? Does the material have to be on standard-sized sheets or forms?

Will you try to stimulate more writing of "review-the-literature" papers by qualified people in the various fields, in order to provide guides for other workers?

Will you provide "special communication networks" in which workers in the various specialized fields can easily circulate working papers or "think pieces." A central agency could maintain printing, listing (in appropraite subject-interest categories), and mailing facilities for this sort of service.

Will there be special training for abstractors and translators (similar to what Calhoun reports Russia is giving for translators), or for documentation and information specialists, etc.?

What is the expected yearly document input?

Would the information center keep a collection of the original documents on hand?

How will you pry loose the information that might be useful, but held?

Are there going to be any significant delays in getting information entered into this system?

Is the system available for browsing?

What should happen when the system gets overloaded? Should service just be late in coming (backlog developing), or should the service just be less complete?

Will financial filtering of requests by a uniform fee structure be desirable or effective, or would it be necessary to make a non-uniform fee structure so that there is essentially some

"priority" given?

Will you be able to perform a searching operation in parts? (For example, search 1/10 of the file and check the results to see if you need further searching, in which case you resume searching in the other 9/10 of the file?)

Could the system operate on a "Just search 1/2 the file for me, I don't need a comprehensive search" basis?

Will the system be able to retain a file of questions so that it can ask these questions of all new input material, thus providing up-to-the-minute data for standing questions?

Will the system give qualitative answers (e.g., yes, no, 5,000 tons in 1945, etc.) as wll as references?

Are the output results going to be in English, or a code that we'll have to have translated?

What will I get in my hand after asking your center? And how do I put the questions to it? Do I mail it, phone it....?

Will the NTIC output be in a form such that the researcher can take it to his own librarian and find which of the documents on the list are in his own company's collection?

Will the machine be accessible to anyone by telephone or other direct device, such that the searcher can interrogate the file directly and at will?

What about copyrighted material? Would royalties be forthcoming to the owner of the copyright if your NTIC uses the article?

Should the NTIC also act as a publishing agency to publish periodically collections of papers (reprints) in very new and special fields? (e.g., information retrieval)

Why not periodically publish inventories of research in progress,

to indicate what research projects are currently being undertaken in each specialty field, thus helping to eliminate duplication, as well as to generate interest in other scientists. XXXa Another service would be a listing of ideas for future development of thesis work.

c. Technical

The nature of the material to be stored for retrieval must be considered from the aspects of,1) how often will it be searched 2) how sophisticated a search process is neededXXXXX (i.e. would simple linear classification do?) 3) how much information is likely to be involved in the different categories.

A given theory of retrieving must know the relationship between time and complexity in a search and the amount of material to be searched, i.e. a system might work well with a few thousand entries, but be hopeless with a hundred thousand.

What characteristic must a documentation system possess if it is to meet the requirements of a particular discipline, subject field...?

As the system grows in size, will it be possible to make changes easily in the classification scheme and bring the old coding into the new scheme?

How much work must be done to standardize a vocabulary? (To use for coding & searching.)

What techniques and devices can reasonaly be developed and applied for facilitating such interim-NTIC tasks as printing, reproducing, data processing, retrieving, storing, microfilming, billing, communicating, etc.?

What kind of a XX data-processing system will the interim NTIC

need just to keep track of its operation.

What kind of communications network will be ~~xxx~~ needed for the operation of the interim NTIC?

With present-day methods, what would it cost to microfilm a year's crop of technical journals? What would be the physical dimensions of such a file?

Would it be feasible for the abstracting publications to use a standard format and type font such that mats (or some similar technique) could easily be distributed to other interested publishers, thus saving printing expenses?

What type of abstract to use? (descriptive, critical, indicative)

How can we get good-quality abstracts?

Should the NTIC make use of volunteer abstractors directly, or will it allow the various technical societies to organize their own volunteer abstracting services?

Could you make good use of university science students at part time and summer work to help with the various processing tasks, as a means of alleviating the shortage of people with good enough technical background?

Can you really be sure that classified documents can't get out through your center?

What professional & educational background is needed for the personnel to operate this system?

What is the expected rate of growth of the system?

How can the processing of recorded information be conducted so that it can be used effecitvely in spite of human limitations, or of limitations in numbers of human beings?

On the basis of current information--processing trends, man-

power requirements will soon make it impossible to continue to apply present methods of processing.

Isn't the main problem of information retrieval one of classification - since people so seldom can express in a usable way (to the documentalist) what it is they want?

Can we assess the relative importance of the following information sources, and perhaps devise ways of augmenting the more effective ones?

> Advanced publications, research perodicals, reports, handbooks, trade journals, conversations, indices, abstracts, library card catalogs, conferences.....

"almanac" facts, "how to do it" facts, "catalog" facts will be the needs of some people. How can these be best reated?

Can we devise an "index of performance," which can be determined by some form of sophisticated sampling among users, and which could be used to rate the utility or efficiency of an activity in the information retrieval or documentation field?

Could we suspend publication of journals and just furnish reprints upon request?

d. Miscellaneous

If we must look to the Nat'l. Govt. for support, what residual responsibilities remain with the professional societies?

Could this possibly be sponsored by private industry? Why go to the government? Could private groups, professional societies, sponsor important special collections?

With the forseeable exponential increase in the ~~quality~~ quantity of information that has to be processed by information centers, would it be economically feasible for any sort of commerical enterprise, or internally-sponsored professional group, to

establish and run a center? Could these centers be self supporting and still furnish the type of service which continued social advancement requires?

Where will the Library of Congress and ASTIA fit into the picture? Can they be expanded to do the job?

If not handled by some existing Government agency, would it not be best handled by some University, e.g. WRU

How soon will we get results?

Will you just be able to add years and years of information into the system or will it all become too bulky and obsolete in a few years so we'll have new appropriations to argue about?

Any prediction on when you could include such fields as Law and Medicine in your center?

Any idea of what the situation will be ten years from now?

How many things, such as ASTIA, can we cease spending money on, to help fund this program?

If we start with something like this, how soon ~~XX~~ could we ~~XX~~ drop ASTIA?

Does "having a large Central Information Center" mean the physical ~~XXXX~~ collection of all these services at one cental location?

"A paper by Perry and co-workers makes the important point that quantitative evaluation of the performance of an information retrieval system should take ~~in~~ into account the fact that identification of a considerable number of pertinent documents is no guarantee that all of them have been identified even when attention is directed to a considerable number of items of secondary or marginal interst. In fact, a large number of non-pertinent items may be selected while at the same time pertinent items are

not being selected." [Kent, Allen, Berry, M.M., Luehrs, F.J., Jr. and Perry, J.W. Machine literature searching. VIII. Operational Criteria for designing information retrieval systems, American Documentation, 6, 93-101, 1955.]

One may stand the chance of getting many marginally pertinent items from the file, and still miss some of the first-order pertinent ones. "Such a situation may arise, for example, when indexing is based on using the words found in a document rather than on the ideas or concepts to which the subject content of the document pertains."

Personnel requirements...points out that background knowledge in the field of activity is necessary for adequate classifying, coding, abstracting and searching...but that a second requirement of a good knowledge and understadningXXX of ~~a~~ documentation system is needed by the specialist if he is to serve well in helping in the system, or if he is to be able to conduct searches well.

"Personnel requirements and their key importance have been pointed out in order to emphasize the fact that the introduction of new methods, systems and machines ~~XXXX~~ does not guarantee more effective use of recorded knowledge. Methods, systems and machines are tools whose effective use requires skill and imagination in applying them. The goal in developing ~~X~~ and applying various methods, systems and machines must be to enable people_____the requisite qualifications to accomplish results that otherwise would be impossible. It is well to add that, in many information systems at least, important tasks must also be performed by clerical personnel and technicians.

Scientists working in a very limited subject area, may be too few to support a journal. They publish in a variety of journals, none of which are devoted to their subject discipline. They then have a problem of how to keep mutually informed on what is going on in their field.

Importance of human memory. "The well-integrated knowledge of the expert coupled with seasoned judgment and creative imagination is an information source that cannot be equalled by a collection of documents or other records. Rather, such collections must serve as an adjunct to human memory--as a source of knowledge needed in making decisions but not already in mind because of various human limitations."

(We shouldn't neglect possible capitalization of the human expert's abilities when we design a retrieval system.)

"The problem of information retrieval is that of rapidly locating sources pertinent to the questions asked when no a priori knowledge identifying the source is known.

If we know that a pertinent source was prepared at the MIT Radiation Lab under Contract 10567A, entitled "MIT in MEW Systems," written by J. Glaser--or any combination of these facts--then we are placed with no real problem. Excellent methods of rapid location by such indices have long been in existence.

If, however, we are searching for information about the use of MIT in MEW radar systems, then we are definitely faced with a problem."

We invite suggestions, comments, criticisms, additions and questions pertinent to the attached list or to any other feature of the "information

retrieval" problems and prospects.

Doublas C. Engelbart
Charles P. Bourne

Engineering Division, 410B

Work and time schedule for three-week period ending 1-24-58.

Objective is essentially to produce two papers...

a. Problem Expansion (for NTIC in particular), for the consumption of representatives of SRI, to serve to give them more working data, and to serve to unify the attitudes and mpp efforts within SRI.

b. Historical review of information-retrieval development...for much the same purpose as "a" above, and to help give perspective and to avoid making public blunders such as failing to give recognition to the pertinent and valuable work which others have already done.

The organization of "a" is tentatively set to be as follows:

1. Under appropriate headings (such as have been established in our "sequence list") list all of the pertinent (and perhaps impertinent) considerations which can be derived.
2. Give proposed solutions to these, where available.
3. Give criticisms to these solutions, where available.
4. Provide discussion to give orientation, coherence, evaluations (pertaining to relevance, difficulty, size, etc.) where possible.

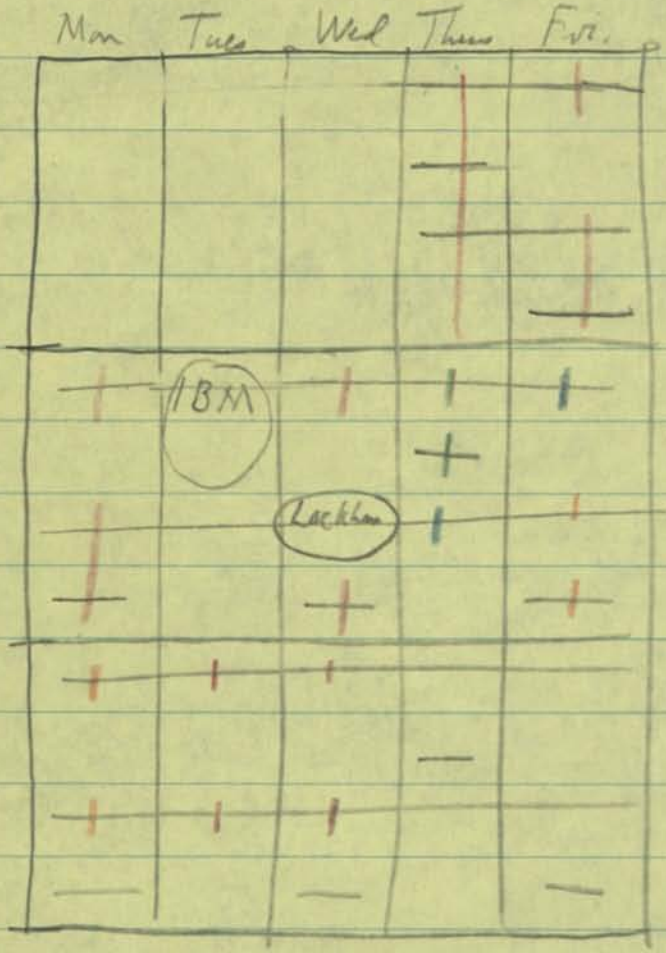
The question of how to work these different things together can await further clarification of the use to which ~~him~~ the end paper will most likely be put. For instance, if it is to provide a basis for future study primarily, then it would seem that, following an appropriate introduction, the considerations, solutions, criticisms, and discussions should be blended within the appropriate headings. If, on the other hand, it is to be used as ammunition stores, or fact finder, then they could very well be kept as three or four separate lists so that the differentiation between problem consideration and solutions etc. can be more ~~mmmmmm~~ easily observed.

TIME SCHEDULE: Giving ourselves till Jan 24 to finish this, we plan to divide our time as follows: We have essentially 27, one-quarter(two-man)days, i.e. two of us will be available ~~for~~ for 27/4= 6.75 days (13 man days). Using the one-quarter(two-man)-day units, we will use them as follows:

Units.	Task.
10	Read all pertinent literature, purposefully, to abstract on cards(with appropriate coding, all considerations, historical data, solutions, criticisms facts, newly-stimulated personal thoughts, etc.)
2	Organize these, together with the collection of questions and considerations, etc. which others have contributed, into the logical groups of the end paper. Integrate and rephrase a coordinated set of considerations, with uniform wording style, so that all of the considerations appear therein, but redundancy is absent and mm the main points are salient. Arrange these in order or in sub-groups under the main headings.
2	Relate solutions, criticisms, discussions, etc. to this list.
4	Compose final form of paper "a".
1	Follow through publishing and dissemination.
4	Compose the historical review, paper "b".
1	Follow through publishing and dissemination.
2	IBM visit
1	Lockheed visit

Total = 27

1-8-57



6

14

14

34 total
- 4 for visits
31 hrs work

—— assemble data

—— organize, rephrase, re-order

—— relate solutions, criteria, decisions, etc

—— compose paper "a"

compose paper "b"

NOTE DECK UTILIZATION:

Objective of the note-deck utilization is to avail ourselves of an efficient way for two or more people to contribute efficiently to the gathering of pertinent information, as well as to establish an optimum technique for extracting all of the useful thoughts, facts, considerations, etc. , for a given limited purpose, in such a way that the total job of extraction and later utilization is most efficient.

1-6-58

MODIFIED OBJECTIVES OF THREE-WEEK PROJECT: (after talk with Eldredge)

Prime objective is to put out a paper which will serve to unify SRI's efforts in furthering the NTIC. A quasi-technical dissertation which can present the essentials of the background and of the solution difficulties of the total problem in such a way as to serve perhaps a dual purpose: 1) to let the men in one spcialty area become aware of the nature of the problems which will be faced in by those of another area, and 2) to let our representatives have a condensed but fairly comprehensive presentation of the entire problem (including SRI's potential and background) to use to back them up when they are grilled by VIP's of one sort or another.

It also seems important, for the benefit of all interested SRI personnel, to prepare a historical summary which can give coherent perspective of the present information-retrieval status with relation to past developments. This may or may not want to be included with a paper which is intended to satisfy the prime objective listed above.

An immediate objective would be to make a reasonably detailed outline of our (Bourne and Engelbart) idea of the finished paper, together with a detailed statement of the objective of the paper, and circulate these about the Institute. Reaction to this should produce helpful suggestions, as well as (it is hoped) recognition of the need for a little formal help (in filling out the outline to make a complete paper) from other divisions.

Eldredge emphasises that finished style of writing might be inefficient, in view of the relative certainty of modifications required and editing done.

superceded

Jan 6, 1958

Objectives of three-week project?

Prime objective is apparently to put out a paper which is to serve as a followup to the initial SRI-NTIC proposal. To be distributed outside the Institute, to be technical in nature, to be directed explicitly towards NTIC, to be written (by DE and CB, at least) in technical language.

We could put out a list of the pertinent considerations which will be involved in the establishment of the project and in the realization of the NTIC. Could perhaps break up the total problem into apparently logical areas, each to be an interdependent entity. Could delineate the limits and boundaries of the problems...try to make upper-bound estimates .

To show that we are competent in the field, and that we have a background of interest and information in this field.

Can include, with lists of considerations, the partial solutions which various workers have proposed, and the objections to these which have been raised by themselves and others.

Would seem safe, for the first week or so of the project, to abstract in a comprehensive fashion all of the pertinent considerations, all of the proposed solutions, and all of the criticisms of such solutions, that can be either generated by us or found in the literature. We would then be ready to write up a final report in one of several different manners, since it seams that this above type of data would be vitally important to almost any writeup we may be asked to produce.

authority to commandeer publications | Bob Way [illegible]

Charge number

Feb. 3-9, Western Reserve will have meeting to discuss essentially the "gross dimension" .. Kesten is going

Similar work -- hand-writing study, among other things

Verbar copies of work lev mtg

& outline of Feb. meeting.

1-24-58

NOTE TO SRI COMMITTEE ON INFORMATION RETRIEVAL

Closer analysis of the contents of the Draft Proposal, reveal to us that the proposal which is now being prepared, if it is to be concluded in the same train of thought that our two meetings have established, will not be representable as satisfying item 2 of the Method of Approach section of the Draft Proposal.

First, the "gross dimensions of the problem" are being limited by us to the problems pertinent to an intereim center, and not to the total problem.

Wasn't furnished or submitted

1-28-58

"Introduction"to the proposal. (in which we state the objective of the project which we are trying to sell, and in which we explain the inter-relation of the various parts of the outline with respect to our course of action)

The objective of the proposed program will be to provide a plan, or a set of alternative plans, for an interim NTIC which meets the spirit of the recommendation in the SRI draft proposal. The program which is to produce this plan is conceived to be in two co-existing parts (aspects?): 1) a data-gathering study and 2) a planning or design study. (there probably is a better word to use here than "study"). The first study will serve to collect all of the data which is necessary to design a plan for the interim NTIC. ~~The second-named study will serve to co-ordinate the findings of the first study with the practical considerations of time, money and human resources to devise an optimum plan, with perhaps alternative plans, for the interim NTIC.~~ The second-named study will produce a plan, or alternative plans, for the interim NTIC, which co-ordinate in an optimum fashion the findings of the first study and the existing limitations of time, money, and manpower resources.

These two studies will be carried on simultaneously, under a single program manager, who will coordinate the activity of the two parts of the program, in the light of the prime objective,.

Copied by Bob Bruce, possibly used in his proposal.

1-24-58

Outline (rough) of Bonus proposal-to-be.

✓ Background.

✓ I. Objective

✓ II. Scope

a. for interim cents, essentially

b. nat'l-defense oriented security

c. U.S. needs, total-world coverage

✓ III. Method of Approach

A. General plan - two coord. aspects.

B. Marketing Study.

1 User needs

2. Sources & facilities

C. Planning & Design of interim NTIC

IV Project organization

✓ V Qualif. of SRI

✓ VI Time & Cost.

IRF

Data entered in VF.

600 Kent, "Trends in Info. Services - US vs USSR Developments in Scientific & Eng. Fields. WRU, 4/57 (in VF. Eng. Lib. under WRU)

1/20/58 601 Herner, "An O.R. approach to the Resurrection of Knowledge" presented at A.C.S. 3/55

602 British Standards Institution, "Universal Decimal Classification", (Z696) Abridged English Edition 025.45
British Standard 1000A :1957
(F.I.D. No. 289)
2nd. ed., revised 1957

1/20/58 603 Current Research & Dev. in Scientific Documentation
Compiled by Natl. Sci. Foundation, July 1957

1/20 604 "Fosdic II - Reads Microfilmed Punched Cards"
NBS Tech. News Bul., May 1957

1/20 605 "Taped Code Operates Microfilm Data File" (NBS photomatrix)
Automatic Control, Aug. 1957

1/20 606 Filmorex advertising brochure w/translation, Bob Lee's report, & samples.

1/20 607 Kuipers, "Microcards & Microfilm for a Central Reference File"
Ind. Eng. Chem. 1950, 42: p. 1463

1/20 608 Brenner, "Experience in Setting up & using the Zatocoding System"
WRU, 4/57

(OVER)

1/20 609 Sass. "Industrial Libraries Come of Age"
the Management Review, Mar. 1957

1/30 610 Jungreisen, "the Role of a Science Info. ~~Center~~ Dept. in Industrial R&D."
J. Chem. Ed. 12/56

1/30 611 "Building a Functional Library"
Chem. Eng News, 12/13/54, p. 4980

1/30 612 Herner, "Tech. Info. – Too much or Too Little?"
the Scientific Monthly, 8/56, p. 82

613 "Scientific and Technical Translating and other aspects of the language Problem." UNESCO, 1957 (SRI Library)

1/30 614 Gray, "General Prob. of Info. Services"
the Percolator (the Chemists' Club, New York) 11/57
also Spec. Lib. 9/57

615 Tanbe, "Communication theory & Storage & Retrieval Systems"
Tech. Rep. # 12 prepared under contract NONR-1305(00)
10/55 ASTIA # AD 93970 (SRI-E-6375)

2-6 616 Recordak Corp. "The Signal Corps Microfilming Project at Ft. Monmouth, New Jersey".

617 Haagdorn, A. "Description of the N.L.L. Hollerith card system Catalogue of aerodynamic measurements." Oct. 1957, National Luchtvaartlaboratorium, Amsterdam, the Netherlands. (E-6360)

Kreithen, A. "Mathematical Foundations for a Storage & Retrieval Theory", June 1957, Documentation, Inc., AD-132 475, E-7147, AF 49(638)-91

Ramo-Wooldridge Corp, "Design Study for an Integrated USAF Intelligence Data Handling System. Appendix A - Machine Translation", AF 30(635)-2867, AD-131 179, E-7138 (U)

Perry, J., "Mincard Indexing Study. - third report." 15 Apr 56 to 14 May 56, AF 30(602)-1498 AD-96 889, E-7155 (U)
(this include a $150,000 proposal to do a study of AF. information systems).

Bedford, G. "Systems of Information Storage & Retrieval", Final Rep. on Phase I of Project Eye Bee, 1 May 55, Univ. of Penn. (Confidential), AD-66 477, E-7148 AF 18(600)-125

Herner, Meyer, & Co., Research, Problem-Solving, & the Use of Technical Information in Small & Medium Sized Manufacturing Firms. 1958. Distributed by U.S. Dept. of Commerce, Office of Technical Services, PB-131 578, E-7350

"Proceedings of the Int'l. Study Conference on Classification for Information Retrieval" (London, May 1957), Avail. at Stanford. 025.4 I59 pi

In our Info-Ref-File (IRF) 1-8-58

No.:	Item:	
001	SRI "Draft program" for NTIC, Jan. 1958, 25p.	
002	D.C.E.	
003	C.P.B.	
004		
005		
6		
7		
8		
9		
010		Correspondence
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		