Repackaging of Scientific and Technical Information*

By JULIUS FROME and JOSEPH F. CAPONIO
Defense Documentation Center for Scientific and Technical Information (DDC) (formerly ASTIA), Arlington Hall Station, Arlington 12, Va.
Received April 15, 1963

One of the major problems of today's Aerospace Age is controlling, or more precisely exploiting, the information revolution catalyzed by the advances of the "new frontiers" in science and technology. The successful nations of the future will be those which will best learn to effectively organize and manage the knowledge emerging from the many intellectual enterprises.

The spectacular increase in the number of scientific discoveries and their technological applications is reflected in an equally impressive quantity of relevant literature. Because so much has been said and written about the exponential character of the development of science and the resulting information problem, there is little to be gained by elaborating on the statistical analysis of growth curves of scientific and allied publications. Suffice it to say that the individual research worker, bewildered by the overabundance of publications in his subject field, is also dismayed by the wide array of services and conflicting claims for systems to solve his information problem. The seriousness of the problem is further aggravated by the rapid emergence of new fields of knowledge and interdisciplinary sciences, whose growth has been accelerated by the technological pressures of modern warfare. Witness, for example, the rate of advances in the developments of such fields as radiobiology, chemical physics, bioastronautics, geochemistry, psycophysiology, bionics, and the like within recent years. Although these new fields of endeavor presage era of challenging and rewarding research, the researcher, assigned to a project of considerable priority, finds conventional systems of retrieving information inadequate. The problem stems not from a lack of scientific information, but rather from the abundance of material dispersed in countless journals, monographs, technical reports, and other allied publications. As a consequence, a substantial part of the productive researcher's time is expended in acquiring, searching, and evaluating the literature.

This problem is being vigorously attacked by professional societies, privately sponsored organizations and institutions, federal agencies, individual documentalists, and information scientists. For example, many abstracting, indexing, and reviewing services, as well as new specialized information centers, in both mission-oriented and discipline-oriented fields, have been established during the past several years. Although work is progressing on the problems of mechanizing the handling of technical information, recent emphasis on large-scale computers, machine-generated indexes, auto-abstracting, etc., has tended to focus too much attention on only a part of the over-all science communications problem. It should be remembered that current efforts to improve and expand abstracting, indexing, and related information services of the United States have generally resulted in wider abstract coverage, and larger subject and author indexes. However, these efforts have done little to eliminate or lessen the researcher's dilemma. On the contrary, he now has more papers to read, more indexes to check, more abstracts to scan and more bibliographies to review. Indeed, the wonder of it all is that the scientist or engineer, frustrated in his abortive attempts to tame the torrent of information, doesn't now spend even more time determining what research has already been completed. It has been said that for the sake of one genuine fact, a thousand tons of verbal ore must be examined.

Thus, the task of making readily available and conveniently accessible the vast storehouse of knowledge on a given subject or problem requires the development of novel and speedier techniques with regard to repackaging and disseminating scientific and technical information. This paper describes a specialized information package designed to meet these specific requirements. It limits itself to those aspects of the information retrieval problem of most direct concern to the scientist, engineer, or technical administrator: the location and acquisition of scientific information pertinent to his field of work. The specific information package discussed here pertains to the interdisciplinary science of radiobiology, but the package system can be applied to other scientific areas as well.

This information service system combines in one publication, and its attendant services, the past, current, and projected work efforts on a specialized area in science or technology. The principal elements of this new service system are: a custom bibliography (with abstracts), listing the material available in a collection of documents; an index to the bibliography section indicating current descriptive terms used in the area of the discipline or technology being covered; a listing containing information on current research projects; and a supplementary detailed vocabulary section listing terms by which detailed information can be retrieved, with some specific guidelines for obtaining the desired information quickly.

The Information Package. (A) Custom Bibliography (with Abstracts).—The custom bibliography would be an up-to-date, comprehensive collection of pertinent abstracts taken from all publications in a document collection.

* Presented before the Division of Chemical Literature, 144th National Meeting of the American Chemical Society, Los Angeles, Calif., April, 1963.

This abstract search would provide the background and state-of-the-art information the researcher needs to begin his investigation. If he needs additional details from any of these reports, they could be quickly retrieved from the collection which has been indexed in depth according to the terms in the Supplementary Vocabulary Section. The specific features of the custom bibliography would be the informative abstracts it contains and the ease with which the researcher may scan them.

A note of caution regarding "selective" bibliographies—the "selective" label, according to Jacobius, "is often a euphemism intended to conceal the fact that the authors were either unable or unwilling to achieve comprehensiveness."[1] He further states that "by being incomplete under the pretext of being selective, it may mislead the user into accepting what is fragmentary for what he believes is essential." To ensure the comprehensiveness of coverage, marginal and analogous information is included in the definitive bibliography.

The term abstract is familiar to all scientists, engineers, and technical administrators, but there is not sufficient uniformity as to its explicit meaning. DDC's definition of abstract has been accepted by a large portion of the scientific and engineering community. By this definition, an abstract is an *informative* summary of a given work, article, or technical report which presents the purpose or objective, methods, results, and conclusions of the author. The informative abstract is distinguished from an *indicative* abstract or annotation whose purpose is to tell what the report is about. Indicative abstracts are provided only when no summary is possible, that is, when the report, itself, is a bibliography, review article, or a collection of papers such as a proceedings or symposium record.

The customized bibliography also plays a prominent role in the dissemination of information to the graduate student entering a specialized field of research for the first time. The rapid advances in today's scientific research and technology make much of the knowledge acquired in college and graduate school somewhat obsolete by the time the student is ready to pursue research. Moreover, many of the nation's recent graduate specialists, lacking the prerequisite knowledge of bibliography and information science, are not aware of the numerous and rich storehouses of information available for his use. As a consequence, the student does not take advantage of all the latest information pertinent to his specialty, and may duplicate work which has already been performed.

The literature searching phase should be a continuing one for the life of a program, rather than just a prerequisite to its initiation. Information in the literature and that being generated by the experimental scientist or technical representative should be integrated to enhance the research phase by receiving the productivity of the researcher. Such literature-user feed-back control has already gained considerable acceptance as evidenced by the development of science information departments in many of the nation's leading scientific and industrial organizations.

(B) List of Current Scientific Research Projects.—One of the greatest deterrents to a productive research force, and a factor that threatens the effectiveness of Western science, is the lack of a central integrated system for promoting the interchange of information on current scientific research.

As stated earlier, the information retrieval problem is widely recognized, has inspired the quest for new and improved techniques for communicating scientific information, and has led to the creation of a new scientific discipline. Yet, according to a report by the Subcommittee on Government Operations, "no comprehensive studies have been undertaken of the possible salubrious effects of a unified project index on the management and conduct of research by which both science administrators and investigators could readily inform themselves on current research activities being conducted nationally or even internationally."[2]

No central register or index exists for the Nation's research and development program. However, considerable progress has been made in providing an inventory or central index on who is doing what, where, and with what funds in federally sponsored research and development. The Science Information Exchange, established in 1960 within the Smithsonian Institution, is acquiring, organizing, analyzing, and disseminating information and data on research which has not reached the publication stage. The Exchange fills an important gap in disseminating information in the life and physical sciences. Originally designed as an administrative tool to prevent unintentional duplication of research support, the central index of the SIE is rapidly becoming a necessary tool for the researcher who requires advanced knowledge of unpublished results and the identification of experts and authorities in these fields.

The Department of Defense also maintains several project indexes on the objectives and substantive content of its RDT&E programs. The principal one is the Department of Defense Research and Development Project Card Index, DD-613. It is important to point out, however, that although this file is available to technical administrators, engineers, and scientists within DOD, it is not generally available to the over-all scientific and industrial community.

The system discussed is designed to present an "end product" which is representative of the total system of science communication rather than of any of its parts. In this connection, one section of this package is a list of brief summary statements on current research projects in the field of radiobiology. It should be remembered that available sources do not provide information on the progress of results of research; its basic objective is to register current research problems and the investigators and contractors working on them. The progress and final reports are prepared for the contracting agency and submitted for storage and retrieval *via* the regular reporting mechanism. To the extent that prompt dissemination of current research information stimulates new research ideas, expedites the accomplishment of research programs, and prevents unintentional duplication of research effort, the DD-613, used by DDC, represents a powerful tool, if properly used by the research administrator, engineer, and scientist, and if the information in the DD-613's can be kept up-to-date.

(C) Index of Descriptive Terms.—The index attached to the bibliography (in this case, Radiobiology) has been prepared from a code sheet which is also described in a

later section of this paper. The subject input terms of the index consist of several major descriptors grouped according to their generic relationships and, in some cases, sub-grouped to achieve a greater degree of specificity. It should be pointed out that these index entries refer to concepts contained *in the document* rather than in the abstracts alone.

(D) Supplementary Vocabulary (Microthesaurus) of Descriptive Terms.—One of the primary problems in the design of an information retrieval system is the development of an effective, detailed subject-indexing system. Although the conventional techniques of classification and subject headings are still used to some extent, both in this country and abroad, the advent of punched cards, computers, and other special devices has led to the deemphasis of such classification schemes and to the increased use of coordinate or correlative systems of indexing in some places. During the transition from a highly developed manual system of subject heading control to a partially mechanized system of indexing, a most powerful information retrieval tool has been the *Thesaurus of ASTIA Descriptors*. The development of this thesaurus has been described in several publications.[3-6] It is a dynamic, but controlled, authoritative vocabulary of descriptors that are displayed, together with their interrelationships and notes, to indicate their usage. Despite the tremendous advance represented by the creation of the first edition, and reflected in the refinements of the second edition of the *Thesaurus of ASTIA Descriptors*, it became apparent that a new program was needed to meet the needs of the scientific and industrial community. In recognition of the need to bolster the Thesaurus approach and alleviate the problems within the DDC system, we established a microthesauri program.

The primary objective of the microthesauri program is to supplement the master *Thesaurus of ASTIA Descriptors* with a series of controlled vocabularies covering very specialized areas of a particular discipline or technology. The program is designed to (1) permit greater depth of indexing, (2) offer a high degree of searching selectivity, and (3) provide the basis for a high-speed (1-hr.) search service. The microthesauri are developed around a nucleus of terms taken from the basic Thesaurus, vocabularies of federal and nonfederal agencies, and thesauri of professional societies and other organizations. These terms and the associated vocabulary terms encountered most frequently in the report literature are grouped according to their generic relationships and displayed on a "deep-indexing" code sheet. These descriptors are coded and punched into one or more cards and stored in the conventional index file. Since quick and convenient access to the search file is an integral part of the system, it was natural for us to explore the feasibility of utilizing the microthesaurus file as a keystone of a possible 1-hr. telephone search service. A scientist or engineer can now phone in his request and receive a specific answer within the hour. The answer would be either a list of pertinent documents which will be mailed immediately if the requester desires, or a telephone reply by a DDC science specialist.

Another significant improvement accruing from the development of the microthesauri is the increase in the depth of indexing. In the past, indexing practices at DDC generated descriptor sets of 10 to 15 terms per document. With the implementation of the microthesauri program, the descriptor sets will average 40 to 50 terms per document. Inasmuch as the microthesauri or satellite systems are direct descendants of the second edition of the *Thesaurus of ASTIA Descriptors*, each microthesaurus, upon thorough evaluation and acceptance by the scientific community, will be merged into the over-all system to form one master thesaurus. Thus, each document entering the DDC system will be descriptorized on the basis of a deep-indexing code sheet and will be entered both in the specialized microthesauri file and master search file. In keeping with the thesaurus program, each microthesaurus code sheet will be used in conjunction with a code manual which contains scope notes, definitions, and instructions as to constraints and special usage of the terms comprising the vocabulary.

In the development of a mechanized system for literature searching in a discipline-oriented or mission-oriented field, there is often a natural inclination to focus too much attention on the machine aspects at the expense of the human element. Since the key to successful dissemination of scientific information lies primarily in the strength of indexing, it becomes apparent that the microthesaurus, more than any other tool, will determine the degree of effectiveness of our indexing system for document retrieval. It is obvious then that the prerequisite to effective information dissemination is the creation of an authoritative microthesaurus. To this end, scientists, engineers, and documentalists from all parts of the country are now actively participating in the development and refinements of these specialized machine-searchable vocabularies. Moreover, these same individuals, through cooperative efforts, are building into the over-all system a high degree of compatibility so as to make possible the prompt and efficient interchange of magnetic tapes and related search files. Before any microthesaurus is adopted, however, it will be formally submitted to review panels consisting of scientists and engineers engaged in the specific field of the microthesaurus.

Summary.—We have attempted to discuss briefly the major objectives and scope of a new service package, to trace its genesis and development, and to enumerate a few of its more salient features. While the need for long-range solutions to the over-all information problem persists more than ever, we have limited our efforts in this endeavor to some of the more pressing short-range counterparts. In this area, a system package has been designed which we feel answers the needs of the scientist, engineer, and technical administrator. Several classes of information services have been integrated and are now readily available to the researcher. They are: (1) a broad collection of informative abstracts for scanning and browsing; (2) a current research project or prepublication service providing information on "who" is "doing what" and "where"; (3) a descriptor index to the annotated comprehensive bibliography; and (4) a microthesaurus and 1-hr. search service which provides information on past, current, and projected work efforts.

As the number of scientific and technical papers and reports continues to grow, the need for developing novel techniques for their organization and effective control becomes imperative. In its role as the Defense Documen-

tation Center, DDC has accepted the challenge by launching both a long-range and short-range improvement program, one aspect of which has been described in this paper.

BIBLIOGRAPHY

(1) A. Jacobius, *Aerospace Med.* **30**, 507 (1959).
(2) "Coordination of Information on Current Scientific Research and Development Supported by the U. S. Government," Committee on Government Operations, U. S. Senate, Report No. 263, 1961.
(3) W. A. Barden, W. Hammond, and J. H. Heald, "Automation of ASTIA—A Preliminary Report," Armed Services Technical Information Agency, AD 227 000, 1959.
(4) W. Hammond, and Associates, "Controlling Literature by Automation," Armed Services Technical Information Agency, AD 247 000, 1960.
(5) E. Wall, Final Report—First Revision of the *Thesaurus of ASTIA Descriptors*. Armed Services Technical Information Agency, AD 278 168, 1962.
(6) J. F. Caponio, and T. L. Gillum, "Practical Aspects Concerning the Development and Use of ASTIA's Thesaurus in Information Retrieval," Presented at a meeting of the Division of Chemical Literature, 142nd National Meeting of the American Chemical Society, September, 1962.

Literature Research for a Space Materials Research Program*

By E. G. KENDALL, EDYTHE MOORE, and CHARLES HAYS
Aerospace Corp., El Segundo, California
Received May 15, 1963

Ballistic and skip glide reentry as well as solid propellant propulsion requirements have surpassed the capabilities of the common engineering materials. Extremely high temperature regimes have caused the missile and space systems designer to turn his attention to the relatively unknown refractory materials. A refractory materials applied research and development program was initiated at the Materials Sciences Laboratory of the Aerospace Corporation during 1962. This paper will describe the experiences involved in searching the literature in this field.

The project was initiated by defining the classes of materials of interest; these included the refractory metals, their carbides, borides, oxides, and beryllides. Specific problems within these classes of materials were then singled out to receive a concentrated effort. In order to improve the potential of these materials for operational use in missile and space systems, it was ascertained that improvements must be made towards their development as engineering materials. To accomplish this, programs were outlined with emphasis on the synthesis and fabrication of a well characterized material. The material would then be evaluated for its mechanical and physical properties. The specific properties to be determined are summarized in Table I.

The mechanical and physical properties of any solid material are extremely sensitive to its characterization and method of fabrication. Thus, it is critical that any tabulation of this sort of data be accompanied with the properties of the solid which have been listed within the characterization column. This has been a major disadvantage of many of the tabulations of data, source books, and critical tables, which have been recently published in the materials field. With the direction of the program outlined, attention was turned to a review of the literature.

Table I

Characterization	Mechanical properties	Physical properties
Chemical composition (purity)	Tensile strength	Electrical resistivity
	Compressive strength	Thermal conductivity
Stoichiometry	Modulus of elasticity	Coefficient of linear expansion
Microstructure	Modulus of rupture	Emissivity
Atom structure	Hardness	Melting point
Density (% of theoretical)	Elongation (%)	
	Thermal shock	
Grain size	Poisson's ratio	
Preferred orientation (texture)		

As is generally the case, a materials research group is moderately up-to-date on subjects of personal interest. Within the Metallurgy and Ceramics Group at Aerospace, several of the research scientists had considerable background in the field of refractory materials. This resulted in an immediate collection of papers, reports, books, and reference works.

An initial check was made of the literature generated from the Defense Metals Information Center (DMIC). The materials research scientist has found DMIC to be a valuable assistant in keeping him up-to-date. It is difficult to find suggestions for improvement in its operation with respect to delivering information to where it is important—the user. DMIC has done an excellent job in reviewing the literature over the past two years with respect to all phases of the refractory metals. But what of the other classes of refractory materials mentioned previously? Only one document was related,[1] since they are not within the DMIC charter.

A step in the right direction to amend this deficiency was the creation, last year, of the Ceramics and Graphite Technical Evaluation Section of the AFSC Directorate of Materials and Processes. It has not yet been able to

* Presented before the Division of Chemical Literature, 144th National ACS Meeting, Los Angeles, Calif., April 1, 1963.