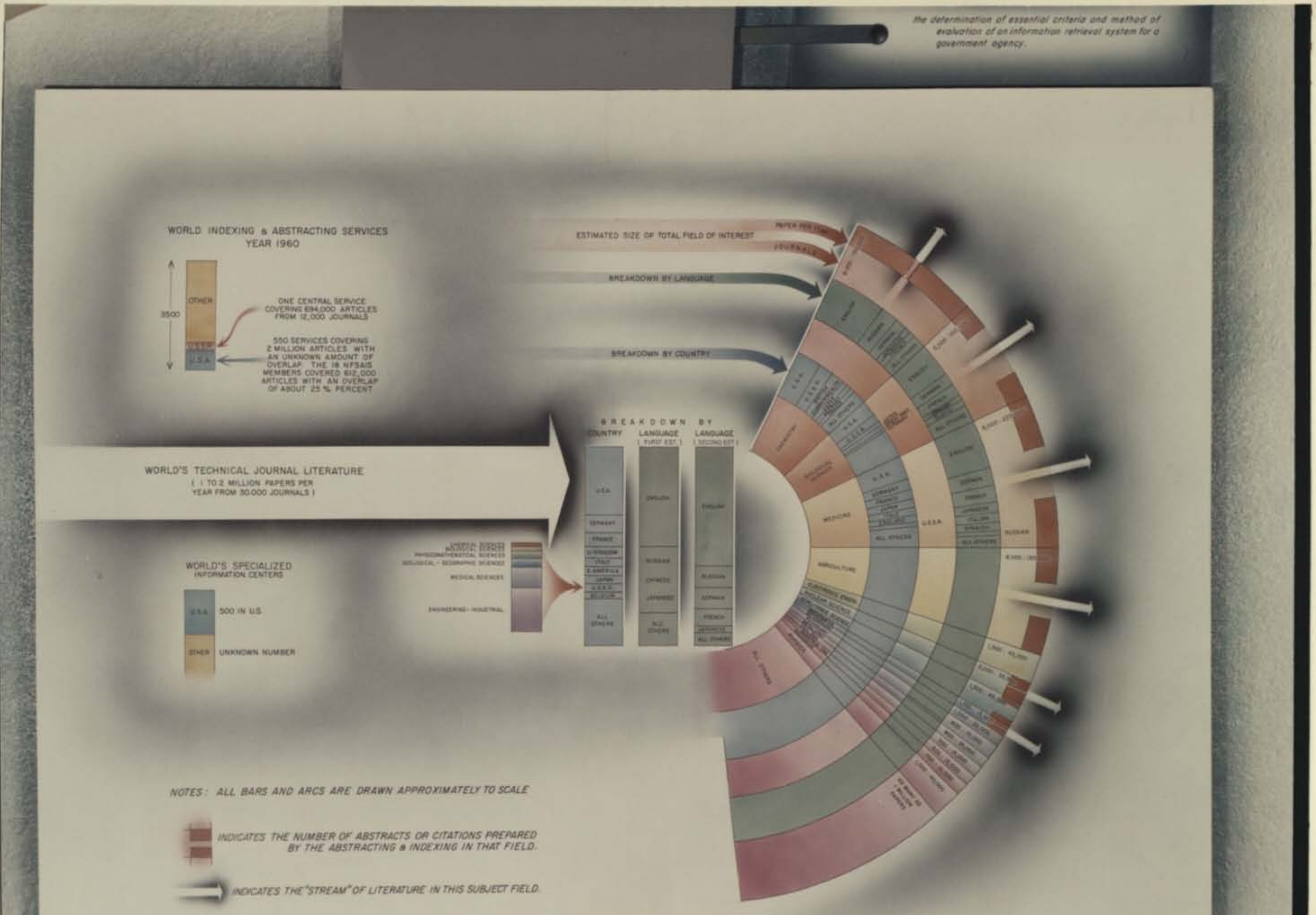
~~[illegible]~~ "World's Tech. J. Lit"

AD 13(2) pp. 159–168. Apr 1962.

appeared initially in IRE Trans. EWS
Aug 1962 (submitted in April 1960)



CURRENT COMPOSITION OF THE WORLD'S TECHNICAL JOURNAL LITERATURE

The above illustration describes the results of one of our recent studies in the storage, retrieval, and dissemination of information -- an estimate of the volume, origin, and language of each of many subject fields of the world's technical journal literature.

A quantitative estimate is made of the magnitude of the world's scientific and technical journal literature problem. Using a number of basic sources of statistical information, a composite picture is established to show such things as the total volume, linguistic and national origins, breakdown by subject field, and degree of coverage by the abstracting and indexing services.

The World's Technical Journal Literature: An Estimate of Volume, Origin, Language, Field, Indexing, and Abstracting

CHARLES P. BOURNE

Stanford Research Institute
Menlo Park, California

• I. Introduction

The world's scientific community is presently generating a flood of technical literature, and much of it is not getting into the hands of people who could use it. It is probably missing these people for two main reasons: only a fraction of the literature is covered by abstracting, indexing, or citation publications; and only a fraction of those people who could use the information are familiar with the literature or aware of the tools and facilities for locating the information. We can measure the amount of literature that is actually covered by the abstracting and indexing services, but this does not indicate what fraction of the important information is getting to the potential users. In many fields the abstracting and indexing services may be covering only a quarter or a third of the estimated total pertinent literature, but this fraction may include all the significant research advances or applications while the remaining fraction consists of rehashes, popularized writings, and trivial contributions. For this reason, the absolute fraction of coverage by the secondary publications is not a perfect measure of their performance—although it does give a good rule-of-thumb measure. Another complicating factor is the business of estimating the extent of coverage when the total production is unknown or extremely uncertain. The reader will see such estimates in the data that are used as source material in this paper. Another comment that must be made here is that because of variations in quality, the actual number of journals or articles is not a very good statistical unit. However, we are forced to use it because of the unavailability of other convenient quantitative measures. It is certainly true that a relatively small number of journals are responsible for the greatest percentage of significant or worthwhile information, and these are the ones that should be covered by the secondary publications and are included in the literature searching systems. However, there is still the chance that very worthwhile work may be presented in small, obscure publications.

The literature is being generated in a great many countries, in a great many languages, in a large range of quality, and—most important—it is being generated in ever-increasing quantities. This is indeed a serious situation, and a considerable effort is being made by many individuals and organizations to find methods or systems to alleviate it. It is doubtful that there can ever be a "solution" to the problem, but there surely must be ways to improve our position or obtain some relief.

One of the first steps that must be taken is to define and describe the problem in as complete and accurate a manner as possible. At best, this problem is an amorphous "bag of worms" that to date has resisted most efforts to describe or define it, or put some bounds on it. However, this does not preclude the possibility of obtaining some partial definition of it. Such a partial description might be obtained by measuring the terminal characteristics of the technical information system, as well as the rates and mixtures of material going in and out of the system. There seem to be these two main types of problem elements here:

1. Those that cannot be measured simply or in any convenient way. Examples of these are: the value of a located piece of information; the cost or penalty of not furnishing the right information at the right time; the relative value of various forms of media and communication channels; and the relation of a technical or scientific worker's productivity to the type, value, and amount of information he receives.
2. Those that can be measured in a straightforward manner, and will provide some indication of the nature of the problem. Examples of these are: the

national and linguistic origins and distributions of scientific information; the forms of publication; the characteristics of the literature in various fields; the volume and growth rates; and the degree of coverage by indexing and abstracting services.

This paper concentrates its attention on the more easily measured elements, and provides a composite estimate—or best guess—of what the sheer magnitude of the problem is, with respect to the periodical or journal literature. An examination was made first of the total number of journals and their annual production; the national and linguistic origin of the journals; and the subject breakdown by country, where possible. Consideration was then given to each of several recognized subject fields such as chemistry, medicine, and metallurgy for the same types of estimate of national and linguistic origins.

Similar studies should be extended to cover other important forms of publications, such as books, patents, reviews, and the "fugitive" document or report literature. Another very interesting study would be the examination of the journal literature traffic (e.g., composition, volume, origins, destinations) between the various countries.

Before getting into the details of this paper I must apologize for the fact that the supporting statistical information does not provide a complete and consistent picture. The statistics and data do not represent a snapshot in time, as they should, but instead, vary over a range of several years. The volume of literature and linguistic composition of one subject field may be compared, in this paper, with the corresponding data for another subject field, but taken three years earlier. There is currently no way in which this information can be easily compared on a common time scale. For many measurements, there are conflicting reports from different sources. Wherever possible, I have tried to reconcile this difference, although in some cases this was done with a flip of the coin. In summary, I must say that to a certain degree these assembled data are unreliable, out of date, incomplete, and inconsistent—and the resulting conclusions must be viewed with some suspicion. However, as bad as the data are, the estimates are probably our best current approximation to the magnitude of the problem.

• II. Discussion

A. TOTAL VOLUME OF PUBLICATION

Estimates of the total number of scientific and technical journals published throughout the world have run as high as 100,000. However, a very thorough and accurate world-wide journal inventory currently being performed by the Science and Technology Division of the Library of Congress for the National Science Foundation has given preliminary indications that there are about 30,000 to 35,000 journals. A more definitive figure should be available at the completion of the project, early in 1962.

Using 30,000 journals as a starting point, and using empirical factors of 30 to 70 articles per journal per year (see Supporting Data Section for the derivation of these factors), we can estimate that a total of one to two million scientific articles are published annually throughout the world. However, more realistic estimates seem to point to a world-wide publication of about 15,000 significant journals, and 1 million significant papers per year.[1]

B. NATIONAL AND LINGUISTIC ORIGINS OF THE TOTAL VOLUME

Some very crude estimates have been made of the national and linguistic origins of the total volume. However, most of these estimates are unclear as to the measure of the volume (e.g., total wordage, number of journals, number of articles, and so on). English is still the predominant language, comprising about one-half of the total production. There are indications that Russian may be coming abreast with the traditional French and German. From a nationalistic standpoint, the published reports indicate that the United States still produces the greatest volume of literature, followed by Germany, France, and the United Kingdom, in that order. However, recent impressions by persons knowledgeable in these statistics indicate that Japan may rank third in the production of all scientific and technical literature, and may be out in front of France and Germany in chemistry, biology, medicine, and agriculture. The relative proportions for each country appear to differ markedly in the various specialty fields. The Soviet literature, for example, seems to be very prominent in chemistry, but relatively light in other fields such as medicine.

Our academic and professional linguistic requirements, as well as our literature acquisition and abstracting policies, should consider these changing trends in the national and linguistic origins.

C. THE MAGNITUDE AND CHARACTERISTICS OF SEVERAL SUBJECT FIELDS

The world's scientific community has recognized and defined various groupings of subjects or disciplines into "fields" such as chemistry, medicine, and mathematics. There is little that is definite about the boundaries of any of these fields, and in many cases there is overlap or complete inclusion of one subject field into another field. There is a tendency, for example, to consider biology, agriculture, and medicine as separate sciences, even though agriculture and medicine are two applied aspects of biology. Figure 1 is subject-biased to show some of these "fields," and because of this must not be taken too seriously with regard to the subject "fields" shown. Nevertheless, a certain body of material or subject matter has been grouped together, rather arbitrarily, usually around an academic tradition, a professional society, or an indexing and abstracting service, to form a single body of material. Artificial as the grouping may be, it exists; and the literature of each single group represents the normal span of interest or information re-

[1] Most of the estimates in this discussion have been graphically portrayed in Figure 1. Supporting references and additional evidence for all of the estimates in this paper are provided at the end of this discussion.

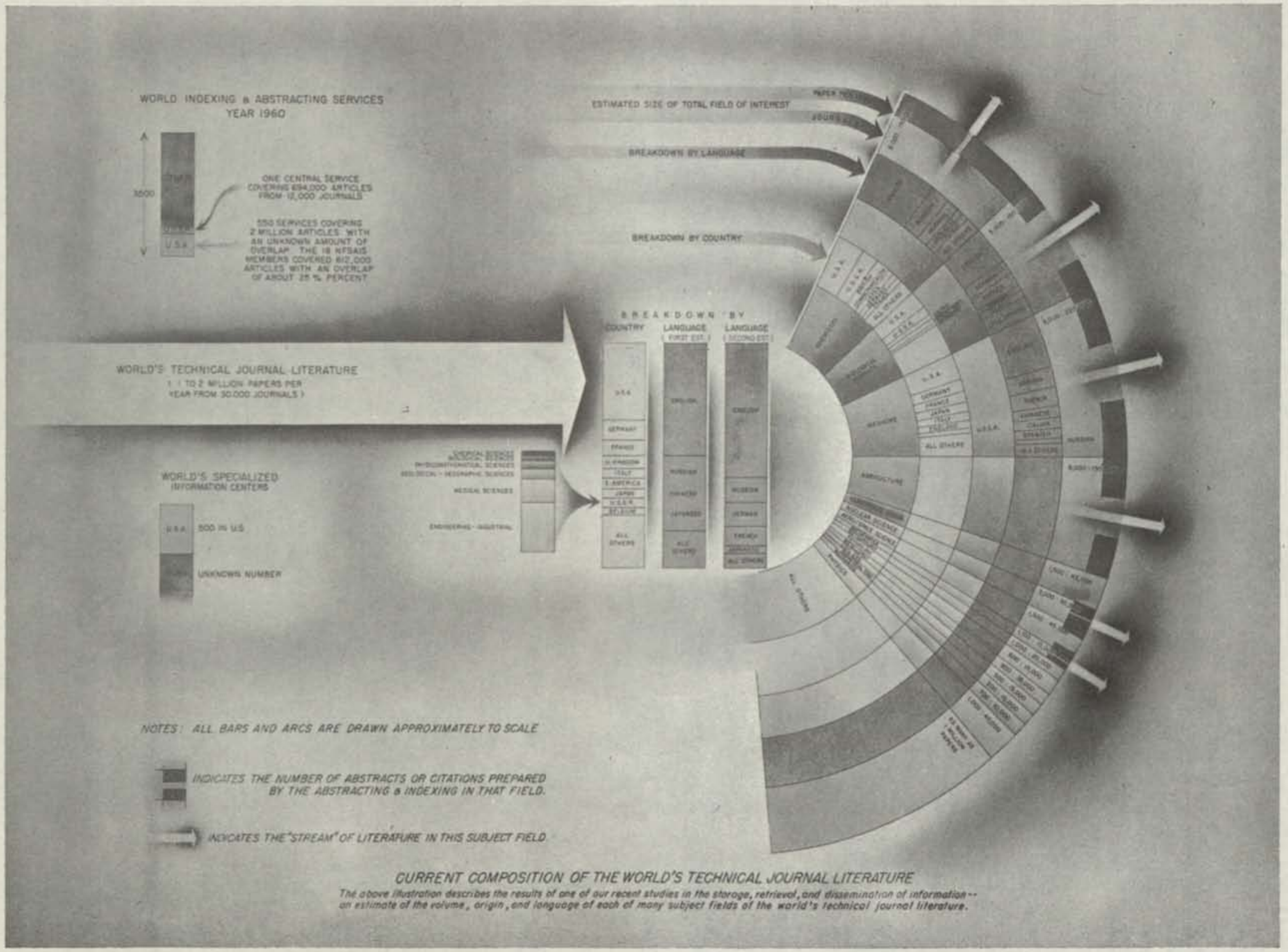


Fig. 1. Current Composition of the World's Technical Journal Literature

sources for an individual working in that field. That is, a worker in any of these subject fields will "see" a certain volume of literature of interest to him; to him, this volume represents his technical information. Two workers in two different subject fields may be looking at some of the same literature. To simplify the graphic portrayal, Figure 1 shows distinct boundaries between neighboring subject fields. However, there is no intention of suggesting, for example, that there is a sharp boundary between chemistry and the biological sciences. In fact, the arrangement of the specialty field segments in Figure 1 was made on a pseudo-random basis.

The subject fields of medicine, chemistry, agriculture, and biological sciences are the largest organized fields. They are separately represented by extensive indexing and abstracting services, even though there is a considerable amount of overlap between the fields.

Estimates of the relative sizes, number of journals, number of papers, national and linguistic origins, and the degree of coverage by the abstracting and indexing services—for each of several subject fields—are shown in Figure 1.

D. ABSTRACTING, INDEXING, AND REFERENCE SERVICES

There are currently an estimated 3,500 abstracting and indexing services throughout the world, about 550 of them in the United States. There are also an estimated 450 special information centers in the United States that maintain collections of information on special technical topics (e.g., Snow, Ice and Permafrost; Defense Metals; Air Pollution) for literature searching and reference services.

As illustrated by the shaded arcs in Figure 1, some specialty fields, such as chemistry, are well covered by abstracting and indexing services; other fields are badly covered. Where no shaded arcs are shown for a subject field, it is because of the lack of information or the amount of coverage in that field. With the possible exception of metallurgy and chemistry, there is currently no practical mechanism, nor is there likely to be one in the next few years, for searching the entire world's literature in any specified subject field to answer a search request. Workers in the field of metallurgy currently have ready access to an operational machine searching system sponsored by the American Society of Metals and the National Science Foundation and operated by the Western Reserve University. Within two years it is expected that workers in the field of medicine will also have ready access to a machine searching system at the National Library of Medicine, and workers in the field of chemistry may have access to a machine searching system at Chemical Abstracts Service. Some indication of how far we are from the indexing coverage of the world's literature may be obtained by noting that in 1961, all of the major U. S. scientific abstracting and indexing services *combined*—services that constitute the membership of the National Federation of Science Abstracting and Indexing Services (NFSAIS)—will cover only about 16,000 out of the world's estimated 30,000 scientific and technical journals, and produce a total of about 750,000 abstracts. The centralized Soviet *Referativnyi Zhurnal* had a coverage of about 694,000 articles in 1960.

E. THE GENERAL LACK OF RELIABLE CENSUS DATA

It is probably a fair assumption that the current pattern of disseminating scientific information by way of journals, books, reports, and proceedings will continue for many years to come. It is this system of distribution that must receive our immediate attention.[2] However, we must know more about the traditional dissemination pattern if we expect to improve it in a systematic way. Measurements and means for continually monitoring the composition, volume, and rates of growth of the literature must be established before any comprehensive plans or policy decisions can be made.

An accurate problem statement is a good start toward the solution of a difficult problem. In the case of the technical literature problem, there seems to be a general lack of good data to describe the problem. A good start has been made by the National Science Foundation, the Library of Congress, and the NFSAIS with their developments of a world inventory of scientific journals, a directory of U. S. abstracting and indexing services, a directory of U. S. special information centers, and a directory of the world's abstracting and indexing services. Some of the individual abstracting and indexing services have also been active in collecting data and analyzing their basic products.

One approach that would be very useful would be to establish a permanent mechanism for obtaining current and accurate information about the parameters of the literature problem—that is, a continuing documentation census that keeps a finger on the pulse of the literature flow to note the volume, character, and trends of the various mixes in the flow, in much the same manner as the continuous efforts of the U. S. Census and Commerce Bureaus and other government agencies. The results of such a continuing census would help to define the problem areas and provide guidance for librarians, curriculum planners, professional societies, publishers, translation services, many government agencies, and researchers in scientific documentation.

• III. Supporting Data

A. NOTES ON GLOBAL AND NATIONAL PARAMETERS

1. *Number of Journals and Journal Articles*

In 1957, the Conference of the American Scientific and Technical Abstracting and Indexing Services estimated an annual world-wide production of 790,000 research papers (2, p. 242).

[2] However, for long-range planning, we must also give serious attention to the many proposals and attempts to modify and augment the present patterns of dissemination. For present and long-range purposes we can also study such things as better methods of birth control for the literature.

In 1960, Dr. B. W. Adkinson, Head, Office of Scientific Information, National Science Foundation estimated conservatively that 1¼ million original papers would be published throughout the world in technical journals in 1960 (1, p. 99).

In 1961, Professor A. Mikhaylov, Director of the Institute of Scientific Information of the USSR Academy of Sciences estimated that there were more than 40,000 journals in the natural and technical sciences, publishing an annual total of 3.5 million articles (28).

In 1956, Mr. G. Miles Conrad, Director, *Biological Abstracts* estimated that there might be as many as 1½ million articles per year of biological interest (11). If biological articles account for 1½ million articles, then it could be inferred that the total number of articles for all fields would be considerably larger, at least on the order of two to three million articles per year.

In 1958, The National Library of Medicine, in a survey of over 31,000 articles in over 2,000 medical journals, found that there was an over-all average of 58.1 articles published per year in each medical journal, and that the titles then indexed in the *Current List of Medical Literature* yielded 71.3 articles per year per journal title (4). *Biological Abstracts* has at different times suggested an estimate of 70 articles per year per journal title (11, p. 2132), and 50 articles per year per journal title (25, p. 5). Since the count of the total number of periodicals throughout the world (30,000) is believed to be a relatively accurate figure, we might use this as a base and assume that each journal produces from 30 to 70 articles per year. This yields a total of approximately one to two million articles per year, and would seem to be our best estimate.

Dr. L. Quincy Mumford, Librarian of Congress, estimated in 1958 that the world's output of a total of 50,000 scientific journals was about two million articles per year (3, p. 367).

Mr. G. Miles Conrad, *Biological Abstracts*, stated that in 1952 there were 50,000 scientific and technical journals in existence throughout the world (2, p. 255). Mr. Conrad cites two references for this estimate (12, 13).

These following facts about 1960 publications seem to support the estimate that throughout the world there are currently about 15,000 important journals publishing approximately one million significant papers:

1) The major U. S. abstracting and indexing services, members of the NFSAIS, covered a total of about 16,000 journals.
2) The centralized Soviet reference service, VINITI, covered about 12,000 journals.
3) The 492 U. S. indexing and abstracting services described in Ref. 7 listed a total of about two million items, with an unknown amount of overlap.
4) The members of the NFSAIS produced a total of 612,000 abstracts or citations with overlaps of perhaps 25 per cent between services.
5) The centralized Soviet reference service had a coverage of 694,000 articles with some intentional overlap.

2. *Volume of Production by Country*

In 1961, Mr. Charles M. Gottschalk of the Library of Congress, Science and Technology Division, provided in a private communication the following preliminary estimates and rankings of the number of scientific and technical journals published by each country or group of countries (excluding "house organs"):

Originating Country	No. of Published Journals
USA	5,950*
Germany (East and West combined)	2,950
France	2,660
United Kingdom	2,090
Italy	1,520
South America (combined)	1,500
Japan	1,470
USSR	1,450
Belgium	1,250
Red China	650
Spain	300
Taiwan	150
North and South Korea (combined)	150

* U.S. "house organs" are unofficially estimated to number approximately 6000 journals.

Mr. Paul W. Howerton, Deputy Assistant Director, Central Intelligence Agency, in a study of Soviet reference work of the Foreign Language Publishing House, published in Moscow, estimated in a private communication that there were a total of 3,824 Soviet journals of all types, with 2,295 of these in science and technology.

In 1959, The National Science Foundation quoted some statistics from a Soviet reference work which indicated that in 1955 there was a total of 2,026 Soviet journals of all types published in the USSR with perhaps 1,100 of these being scientific or technical journals (14).

The *Japanese Periodicals Index*, published by the Japanese National Diet Library, lists titles from about 500 Japanese periodicals in science and technology, and about 300 Japanese periodicals in the medical sciences.

3. *Soviet Literature*

Mr. Paul Howerton, Central Intelligence Agency, recently did a study of the scientific literature of the USSR for 1958 (19). He obtained the following breakdown by subject field:

Subject Field	Per cent of Total
Physicomathematical Sciences	9
Chemical Sciences	3
Biological Sciences	5
Geological-Geographical Sciences	10
Medical Sciences	17
Engineering-Industrial	56

The Central Intelligence Agency stated in 1960 that Soviet publication in science and technology had been estimated to be in excess of 700 million words per year, with about 7 per cent of this being translated into English (1, p. 65).

The Library of Congress and the National Science Foundation estimate that currently about 120 Soviet journals are being translated cover-to-cover into English (6).

4. *Total Volume of Production by Language—First Estimate*

Dr. Adkinson, National Science Foundation, estimates that more than one third of the world's scientific and technical literature is produced in the USSR, China, and Japan; and that about 50 per cent of all technical and scientific literature appears in English (1, p. 127). (I have assumed that the papers originating in the USSR, China, and Japan were written in Russian, Chinese, and Japanese, respectively.)

5. *Total Volume of Production by Language—Second Estimate*

A UNESCO report provides a summary of several statistical studies of percentages of publications in various languages (8). The report suggests that all of the estimates vary within wide limits and are probably subject to significant errors. The report provides the following composite estimates, which are meant to show the order of magnitude only, and do not represent accurate statistics:

Language	Per cent of Total
English	60
Russian	11
German	11
French	9
Japanese	3
Spanish	2
All Other	4

6. *Abstracting and Indexing Services*

The National Federation of Science Abstracting and Indexing Services (NFSAIS) stated that more than 612,000 abstracts and title listings were prepared in 1960 by its member organizations and more than 753,000 would probably be prepared in 1961 (9).

Mr. G. Miles Conrad, Director of *Biological Abstracts* estimated that the U.S. abstracting and indexing services covered virtually 75 per cent of the total literature in 1958 (2, p. 243).

The Library of Congress directory of the world's abstracting and indexing services in science and technology is expected to be published early in 1962.

The NFSAIS is currently compiling a list of periodicals covered by the major U.S. abstracting and indexing services. A preliminary estimate of a total of 16,000 periodicals covered by these services has been made by Mr. Jensen, Executive Secretary of the NFSAIS in a private communication.

Professor A. Mikhaylov, Director of the Institute of Scientific Information of the USSR Academy of Sciences, estimates that more than 300 different abstract journals are published throughout the world, with each journal specializing in one subject field. The same reference indicates that over 90 abstract journals are published in the USA, more than 40 are published in England, and 27 are published in West Germany (28).

7. *Soviet Information Services*

A 1960 description of the Soviet All-Union Institute of Scientific and Technical Information (Moscow) stated that the 13 abstract journals which it was then publishing contained abstracts selected from more than 12,000 journals (plus 3,000 special Soviet publications) originating in more than 95 countries and appearing in 65 different languages (21). The Institute also publishes 48 Express Information Journals, each in a general subject field, to provide quick summaries of reported information. A *New York Times* article credits the Institute with coverage of 16,000 scientific publications from 88 countries (20).

8. *Special Information Centers*

Dr. Adkinson, National Science Foundation, estimated that there are about 3,000 specialized science data and information centers in the United States (1, p. 122), with more than 50 such centers operated or supported by the U. S. Dept. of Defense.

A study recently conducted by Battelle Memorial Institute indicates that there may be about 450 specialized information services in the physical and biological sciences in the United States (26). A detailed study of these services indicates that there are perhaps 200 that might be defined as services that are staffed in part with scientists or engineers, conduct a selected data acquisition and processing program, and exist for the primary purpose of providing analytic and interpretive reports (27).

B. NOTES ON SUBJECT FIELD PARAMETERS

9. *Medicine*

The NFSAIS stated that *Index Medicus* listed 125,000 articles in 1960, and will probably list 140,000 articles in 1961 (9).

The National Library of Medicine estimated in 1958 that at that time there were about 13,000 medical journals, and a total of about 220,000 indexible articles per year which could be considered medical literature (4). However, this number probably includes many journals which are included in the natural sciences. The same reference cited similar studies that indicated a smaller number of periodicals (4, p. 445). A total of about 9,000 journals, embracing all fields of medicine, would seem to be a good compromise number.

In 1958, the National Library of Medicine provided the following estimates and rankings of the language breakdown of medical literature (4):

Language	Per cent of Total
English	37.3
German	13.0
French	12.9 (corrected figure, Ref. 5)
Japanese	7.8
Italian	7.5
Spanish	6.2
(Polylingual)	3.3
Russian	3.1
All Other	8.9

In the same reference, NLM also mentioned that they currently receive periodicals from at least 85 different countries, and provided the following rankings and estimates of the origin of medical journal articles by country:

Originating Country	Per cent of Total
USA	24.5
Germany	11.7
France	10.2
Japan	8.9
Italy	7.8
England	7.4
USSR	3.2
All Other	26.3

10. *Chemistry*

In 1959, the Chemical Abstracts Services provided the following estimates of the percentage of journal articles in chemistry produced in the following languages (5):

Language	Per cent of Total
English	50.5
Russian	16.8
German	9.7
Japanese	6.1
French	5.5
Italian	3.6
(Other)	7.8

In 1961, Dale B. Baker, Director, Chemical Abstracts Service, stated that the articles abstracted by *Chemical Abstracts* in 1960 came from 97 countries, and were published in 52 languages (18). In the same reference, Dr. Baker also made the following estimates and rankings of the national origin of the literature in chemistry and related fields:

Originating Country	Per cent of Total
USA	27.1
USSR	19.1
British Commonwealth	13.8
Japan	7.8
Germany	7.8
France	5.0
Italy	3.2
Czechoslovakia	2.0
Poland	1.5
Switzerland	1.3
Hungary	1.2
Sweden	1.1
The Netherlands	0.9
All Others	8.2

Dr. E. J. Crane, Chemical Abstracts Service, stated that new chemical information appears originally in 8,000 scientific and technical publications (2, p. 167).

In 1961, NFSAIS stated that in 1960, *Chemical Abstracts* printed 132,159 abstracts, and will probably produce 145,200 abstracts in 1961 (9). Dr. Baker states that the 132,159 abstracts were generated from 104,484 articles and 26,675 patents (18).

Dr. E. J. Crane, Editor, *Chemical Abstracts*, estimated that in 1957 *Chemical Abstracts* covered about 98 per cent of the world's chemical literature (2, p. 242). A projection of the degree of coverage to the present situation would indicate a current potential of about 150,000 articles in this field.

Chemical Abstracts stated in 1957 that they covered 7,500 journals, 2,000 books, and patents from 22 countries (3, p. 314).

11. *Biological Sciences*

In 1960, Dr. B. W. Adkinson, National Science Foundation, stated that *Biological Abstracts* attempts to cover a field almost as voluminous as that of chemistry, and that *Biological Abstracts* will have achieved almost 50 per cent coverage by 1960 (1, p. 119).

NFSAIS stated that *Biological Abstracts* produced 72,000 abstracts in 1960, and would probably print 87,000 abstracts in 1961 (9). Assuming that 72,000 abstracts represents about 50 per cent coverage, then there are currently about 150,000 papers per year in this field.

Dr. H. T. Cox, Executive Director of the American Institute of Biological Sciences, stated that in 1958 there were at least 30,000 journals which published original biological research papers—with a conservative estimate of a total output of $1\frac{1}{2}$ million articles per year (1, p. 150).

Mr. G. Miles Conrad, Director of *Biological Abstracts*, notes that in 1953 there were 2,175 U. S. journals in the biology field, producing an estimated 155,000 articles per year (2, p. 239). He also notes that over a 200-year period the number of the world's biological research journals, exclusive of medicine and agriculture, has doubled on an average of every 18 years (24, p. 711).

The Soviet biological abstract journal, *Referativnyi Zhurnal: Biologiia*, published 104,000 abstracts in 1957 (2, p. 239).

Mr. Conrad also estimates that *Biological Abstracts* had a potential of 155,000 articles in 1957 (2, p. 243), and that the number of biological research journals, excluding medicine and agriculture, in 1958 was 3,500 (2, p. 248).

In 1956, Mr. Conrad estimated that at that time there were between 20,000 and 23,000 journals of biological interest (11). Using an estimate of 70 articles per year for each journal, this suggests a total of about 1½ million articles per year. However, he goes on to state that, "... although there is much published of biological interest, it is not always of professional interest. When we come down to earth it seems more likely that there are from 3,000 to 4,000 journals throughout the world that consistently publish reports of original research in biology. By multiplication with our factor of 70 articles per journal per year we find that the maximum abstracting load per year is more nearly 250,000." In a private communication in 1961, Mr. Conrad also mentions that the estimate of 20,000 journals of biological interest represent biology in all of its pure and applied ramifications, and that to say that there are 4,000 biology journals, 13,000 medical journals, and 8,000 agricultural journals and suggest that these are mutually exclusive is most inaccurate and misleading. *Biological Abstracts* covers most of the basic material of a biological nature in these three fields, and leaves much of the clinical or application reporting to other services. Because of this fact, the indicated gap of coverage of the biological sciences in Fig. 1 is not an accurate indication of the total coverage of the biological literature. However, this is the only data that was readily available. Mr. Conrad also mentioned that *Biological Abstracts* was currently scanning slightly more than 5,000 journals for inclusion in that publication.

In late 1960, a reference publication (Ref. 15) listed approximately 5,000 journals from 84 countries that contain articles in biology that are abstracted by *Biological Abstracts*. The same reference provides the following information about the production of biological research literature in various countries:

Originating Region	Per cent of Total
Europe and the Middle East	47
North America	29
Asia and Australasia	13
Central and South America	8½
Africa	2½

The same reference also gave the percentage breakdown by country. The leading contributors are:

Originating Country	Per cent of Total
USA	26.75
USSR	11.40
Japan	5.85
Great Britain	5.83
Germany	4.22
France	4.10
Italy	3.15
Brazil	2.64
India	2.60
Canada	2.13

These data were based on an analysis of a list of biological sciences serials that are scanned for abstracting by *Biological Abstracts*.

Biological Abstracts provided the following information about the production of biological research literature in various languages (11):

Language	Per cent of Total
English	39
German	13
French	13
Spanish	6
Italian	3
Slavic Tongues (Several)	8
Others	18

These data were based on the source cited in (15).

12. *Physics*

The American Institute of Physics stated that in 1959 *Physics Abstracts* contained approximately 14,000 abstracts, and the Soviet *Referativnyi Zhurnal Fizika* contained nearly 28,000 abstracts (16).

In 1958, Dr. Crane, Chemical Abstracts Service, estimated that *Chemical Abstracts* covered at least ¾ of the papers appearing in the field of physics (2, p. 170).

I have assumed that there are currently 1,000 journals and 40,000 articles per year in this field.

13. *Mechanics*

In 1957, *Applied Mechanics Review* estimated that there was a potential of 6,000 articles per year from 700 journals in that field (2, p. 243).

In 1961, the NFSAIS stated that *Applied Mechanics Review* published 6,700 abstracts in 1960, and would probably print 7,000 abstracts in 1961 (9). I have assumed that there are about 10,000 papers per year in this field.

14. *Psychology*

In 1957, *Psychological Abstracts* estimated that they produced 9,000 abstracts from a potential of 12,000 articles (2, p. 243), and covered 500 to 600 journals (3, p. 318).

The NFSAIS stated that *Psychological Abstracts* printed 8,500 abstracts in 1960, and would probably print 7,000 in 1961 (9). I have assumed that there are about 15,000 papers per year in this field.

In 1952, a study of the production of the psychological literature in various languages provided the following estimates (17):

Language	Per cent of Total
English	77
French	8
German	6
All Other	9

15. *Nuclear Sciences*

In 1957, *Nuclear Science Abstracts* estimated that they covered 14,000 abstracts of a potential 14,500 articles (2, p. 243). However, they also stated at another time that their coverage of books and journals was only 60 per cent (3, p. 317).

The NFSAIS stated that *Nuclear Science Abstracts* printed 26,500 abstracts in 1960, and will probably produce 32,000 abstracts in 1961 (9). The coverage of NSA includes patents, reports, books, and translations, in addition to journal papers (1, p. 38). I have assumed that there are 2,000 journals and 35,000 papers per year in the field of nuclear science.

16. *Agriculture*

In 1957, the *Bibliography of Agriculture* estimated that there was a potential of 150,000 articles in their field (2, p. 243).

The NFSAIS stated that the *Bibliography of Agriculture* contained almost 97,000 titles in 1960, and would probably list the same number in 1961 (9).

Since the estimates of total number of articles (150,000) are the same for the field of agriculture as for chemistry, I have assumed that the number of journals for agriculture is the same as for chemistry (i.e., 8,000).

17. *Electronics and Electrical Engineering*

No information has been assembled to show the magnitude of the electronics and electrical engineering literature. There are preliminary indications that this amounts to a total of about 30,000 to 70,000 articles per year from about 1,000 journals.

18. *Aero-Space Sciences*

Aeronautical Reviews in 1957 estimated that there was a potential 13,000 articles per year in this field (2, p. 243).

The NFSAIS stated that *Aero-Space Reviews* (now *International Aero-Space Abstracts*) printed 9,500 abstracts in 1960 and would probably produce 10,000 abstracts in 1961 (9).

A 1960 Special Libraries Association publication lists 676 periodicals (22). Assuming 30 articles per periodical per year, this would amount to about 20,000 papers per year in this field.

Mr. Charles M. Gottschalk, Library of Congress, stated in a recent private communication that a forthcoming bibliography of aeronautical and space periodicals to be published by the Library of Congreses reveals approximately 1,500 journals exclusively in the aero-space sciences. Assuming 30 articles per periodical per year, this would amount to about 45,000 papers per year in this field.

19. *Mathematics*

In 1957, *Mathematical Reviews* estimated that there were a potential 10,000 articles in this field (2, p. 243).

Mathematical Reviews stated in 1957 that they covered 1,036 journals and 300 books (3, p. 316). These figures appear to be high since that would amount to an average of less than one article per issue.

The NFSAIS stated that *Mathematical Reviews* printed 7,800 abstracts in 1960, and would probably produce 10,000 abstracts in 1961 (Ref. 9). I have assumed that there are about 15,000 papers per year in this field.

20. *Meteorology*

In 1957, *Meteorological Abstracts* estimated that they produced 5,000 abstracts of a potential 7,200 articles (2, p. 243), and covered 2,500 to 3,000 journals (3, p. 316). As in mathematics, this must include a large number of journals which contribute very little to the total number of articles.

The NFSAIS stated that *Meteorological Abstracts and Bibliography* (now *Meteorological and Astrogeophysical Abstracts*) printed 12,800 abstracts in 1960, and would probably print 15,900 abstracts in 1961 (9).

Mr. Charles Gottschalk, Library of Congress, stated in a recent private communication that a count of the publications abstracted in the 1959 *Meteorological Abstracts and Bibliography* was 980 and included many journals of general scientific interest. I have assumed that there are about 1,000 journals and 20,000 papers per year in this field.

21. *Metallurgy*

The NFSAIS stated that the *Review of Metal Literature* printed 12,000 abstracts in 1960, and would probably print 12,000 in 1961 (9).

The British Iron and Steel Institute has established an *Abstract and Books Title Index Card Service* and expects to list approximately 9,000 abstracts per year. The British *Metallurgical Abstracts* is currently publishing about 7,500 items per year from 1,000 journals (23). The American Society for Metals now subscribes to 900 journals for their literature searching service, and are indexing about 36,000 documents per year, including books and reports.

I have assumed that there are about 600 to 900 journals and 30,000 to 35,000 papers per year in this field.

22. *Civil Engineering*

No information has been assembled to show the magnitude of the civil engineering literature. I have assumed that there are about 500 journals and 15,000 papers per year in this field.

23. *Industrial Engineering*

No information has been assembled to show the magnitude of the industrial engineering literature. I have assumed that there are about 500 journals and 15,000 papers per year in this field.

24. *Other Specialty Fields*

No information has been assembled to describe other specialty fields. However, considering the overlap and the duplication of articles between the fields shown, then there may be as many as one million articles in fields not specifically shown on Figure 1 in order to add up to the suggested total of one to two million papers. Another way to organize some of the specialty fields on Figure 1 into a relatively large grouping would be to group all of the fields of engineering into a single field of engineering. The *Engineering Index*, for example, published over 35,000 abstracts for this field in 1960.

References

1. U. S. Senate Report, "Documentation, Indexing, and Retrieval of Scientific Information," 86th Congress, 2nd Session, Document 113 (1960).
2. U. S. Senate Report, "Hearings before the Subcommittee of the Committee on Government Operations of the U. S. Senate 85th Congress, 2nd Session on S.3126 (Science and Technology Act of 1958)," Part I (May 2, 6, and 7, 1958).
3. U. S. Senate Report, "Hearings before a Subcommittee of the Committee on Government Operations of the United States Senate 85th Congress, 2nd Session on S. 3126 and S. 4039 Part 2 (June 25, 26, 1958).
4. Brodman, E., and S. Taine. Current medical literature: a quantitative survey of articles and journals, *Proc. Int'l. Conf. on Scientific Information*, (1958).
5. Heumann, K. F., and P. M. Bernays. Fifty Foreign Languages at *Chemical Abstracts, Journal of Chemical Education*, **36,** No. 10: 478. (October 1959).
6. *List of Russian Serials Being Translated into English and Other Western Languages*, 3rd Revised Edition, Library of Congress, 1961, and *List of Russian Scientific Journals Available in English*, National Science Foundation Report NSF 61-66. August 1961.
7. *A Guide to U. S. Indexing and Abstracting Services in Science and Technology*, prepared by the Science and Technology Division of the Library of Congress, and published by the NFSAIS. This is a directory of 492 scientific abstracting, indexing, and title announcement services in the U. S.
8. "Scientific and Technical Translating and Other Aspects of the Language Problem," 1957. UNESCO.
9. Science information services in the United States, *Biological Abstracts*, p. xxi, Vol. 36 (14), (July 15, 1961).
10. How good is Soviet science? *Fortune*, **LV,** No. 2 (February 1957).
11. Conrad, G. M. 1956. The Present Status of Biological Abstracts, *Biological Abstracts*, **30,** No. 8.
12. Smith, W. A. *et al., World List of Scientific Periodicals Published in the Years 1900–1950. 3rd Edition* (New York, Academy Press, 1952).
13. Brown, C. H. *Scientific Serials*, ACRL Monograph No. 16 (Chicago, Association of College and Reference Libraries, 1956).
14. *Providing U. S. Scientists with Soviet Scientific Information*, Revised Edition, National Science Foundation (May 1959).
15. *Biological Abstracts List of Serials*. 1960. Reprinted from *Biol. Abstr.*, **35,** No. 20.
16. Maizell, R. E. AIP Documentation Research Project, *Physics Today* (October 1960).
17. Louttit, C. M. 1955. The use of foreign languages by psychologists, *Am. J. of Psychology*, **68:** 484–486.
18. Baker, D. B. Growth of Chemical Literature—Past, Present, and Future, *Chem. Eng. News*, 78–81 (July 17, 1961).
19. Howerton, P. W. The parameters of an operational machine translation system, Paper presented at the National Conference of the American Documentation Institute, Berkeley, California (27 October 1960). To be published in *Machine Translation*.
20. Frankel, M. Soviet amasses scientific data, *New York Times* (15 May 1958).
21. Baker, D. B., and M. Hoseh, Soviet science information services, *Chem. Eng. News*, 70–73 (11 January 1960).
22. *Aeronautics and Astronautics: A World List of Current Periodicals*, published by the Science Technology Division of the Special Libraries Assoc. (15 November 1960).
23. Bristow, J. S. Metallurgical abstracts. Paper presented at the Annual Convention of the Special Libraries Assoc. in San Francisco on 29 May 1961.
24. Conrad, G. M. Growth of biological literature and the future of Biological Abstracts, *Federation Proceedings*, **16,** No. 3 (September 1957).
25. *Proc. of a Symposium on Biological Communications*, Philadelphia, October 6–7, 1960, *Biological Abstracts*.
26. *Specialized Science Information Services in the United States*, Report NFS 61–68 published by the National Science Foundation (November 1961).
27. Simpson, G. S., Jr. Information Centers. Paper presented at the Annual Convention of the American Documentation Institute, Boston (November 5–8, 1961).
28. Mikhaylov, A. Reply to an enemy, in the Russian-language periodical *Ekonomiche skaya Gazeta* (Economic News), Moscow, p. 4 (20 June 1961). The translation is available at Report JPRS:8865, OTS:61–31.604 *Foreign Developments in Machine Translation and Information Processing, No. 33* (14 September 1961) Distributed by the Office of Technical Services, U. S. Dept. of Commerce, Washington, D. C.

an AID publication

Information Systems Workshop

THE DESIGNER'S RESPONSIBILITY AND HIS METHODOLOGY

an publication

Information Systems Workshop

THE DESIGNER'S RESPONSIBILITY AND HIS METHODOLOGY

(Based on a conference sponsored by the American Documentation Institute and the University of California at Los Angeles under support of a grant from the National Science Foundation and the National Aeronautical and Space Administration, May 29-June 1, 1962.)

SPARTAN BOOKS

6411 CHILLUM PLACE, N.W. • WASHINGTON 12, D.C.

CON

an publication

Information Systems Workshop

THE DESIGNER'S RESPONSIBILITY AND HIS METHODOLOGY

(Based on a conference sponsored by the American Documentation Institute and the University of California at Los Angeles under support of a grant from the National Science Foundation and the National Aeronautical and Space Administration, May 29-June 1, 1962.)

SPARTAN BOOKS

6411 CHILLUM PLACE, N.W. • WASHINGTON 12, D. C.

Library of Congress Catalog Card Number: 62-22240.
Manufactured in the United States of America.

ACKNOWLEDGEMENT

Grateful acknowledgement is made of the services of Charles P. Bourne, Stanford Research Institute, in editing the transcript of the original Lake Arrowhead Conference Papers and again the revised papers appearing in this volume.

The Lake Arrowhead Workshop Committee, representing both U.C.L.A. and A.D.I., acknowledges with appreciation financial assistance from the National Science Foundation and the National Aeronautic and Space Administration in making both the Workshop and this publication possible.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS

Page

INTRODUCTION

R. M. HAYES
Advanced Information Systems, Inc.
Los Angeles 34, California

On May 29, 30, 31, and June 1 an Information Systems workshop was held at UCLA's Arrowhead Conference Center. It was sponsored jointly by the American Documentation Institute and the University of California at Los Angeles under support of a grant from the National Science Foundation and the National Aeronautical and Space Administration. This workshop had the purpose of bringing together the nation's experts in the solution of information systems problems, particularly those systems in which mechanization has made evident the need for precision and efficiency.

Although numerous conferences and symposia have been held in the past in the general area covered by the term "information systems", they have generally dealt with details of specific approaches to solutions in each of the problem areas involved in implementation. To our knowledge, no major symposium had dealt with the problem of how the information systems designer does his work.

Thus, there exists a large and ever growing set of techniques and approaches applicable to the implementation of individual, discrete areas of the information systems problem. These include the great variety of indexing schemes, classification schedules, data handling programs and procedures, compliments of equipment, etc. On the other hand, the application of these techniques and approaches to information systems problems is still an extremely difficult and ill-defined task. There is no well-defined systems methodology or set of tools for selecting which of the potential answers should be combined for the optimal solution of any specific complex problem.

In part, this is a result of the lack of recognition of system design as a major area of responsibility. The lack of recognition of the nature and even, in some cases, the existence of the problems facing the information systems designer has meant that there has been little or no orderly development of generally agreed upon system methodology. There exists only the diverse experiences of individual workers.

It is for this reason that ADI and UCLA proposed that this Information Systems Workshop on the Designer's Responsibility and his Methodology be held. It brought together prominent people in the field of information systems design within a well defined framework in which they could coherently discuss the state-of-the-art of their work. The prime purposes of this proposed workshop were: (1) to identify this problem area, (2) to structure it, (3) to show where needs exist, (4) to point the way for further development effort, and (5) to define the educational needs pertaining thereto.

The plan of the workshop recognized that the methodology of the information systems designer has different functional aspects which themselves vary in significance with different types of information systems. Therefore, the structure of the proposed program was based on discussing each of four functional aspects of systems methodology as exemplified in two different types of information systems—the technical information system and the business data system.

It was planned that, following the workshop, a set of papers based on the discussion would be "commissioned" for presentation at a later conference of ADI and for publication for wide-scale distribution. In this way, although the workshop provided a vehicle by which the participants could formulate their ideas in an atmosphere of open discussion, the discipline necessary to put these ideas into a formal paper has insured that concrete ideas will result. This volume represents the results of this approach.

THE NEED

The workshop was based on the recognition by a number of independent groups of the need for definition of the problems and techniques involved in system design. An "ad hoc committee" of individuals in the Los Angeles area was formed in late 1961 to formulate the structure of a symposium to be concerned with this problem. Because of the close association of these individuals with the

American Documentation Institute and because of the concern of ADI itself with this type of activity, the intent of this ad hoc committee was to have the contemplated conference sponsored by ADI. As a result, the ADI appointed an official committee to coordinate this effort.

Independently of this, UCLA had been developing plans to explore some possible solutions to the educational needs in this field. Various departments of the University share with the ADI a concern about scientific information systems. Among these departments are: The University Library, itself a highly complex information storage and retrieval operation which is currently planning further mechanization, automation, and systems analysis to increase operating efficiency; the Medical Center, engaged in both teaching and research which draws upon and depends upon scientific data; and the School of Library Service, which has responsibility for the professional education of librarians and information specialists, a new department (opened in September, 1960) which will be adding research and specialized instructional activities to its present generalized teaching function. There are other departments and facilities on the campus concerned with various aspects of information sciences, but these three illustrate the breadth of the University's activity in the documentation fields, encompassing source, use, training, and research. The development of information specialists and of researchers in the field of information systems to meet the needs of our modern society makes it imperative that the universities develop strong academic programs in these areas.

The close relationship of these independent efforts made it natural to combine purposes, and the conference was therefore presented jointly by ADI and UCLA.

The problem area itself can best be seen in relationship to other aspects of the general information systems problem. The development of an information system in any specific situation requires recognition of the user—scholar, engineer, technician, or business manager—must clearly be recognized, since they represent the basic reason for existence of the information system. On the other hand, the information specialist and librarian who provide the communication link between the information system and the user are professional individuals with roles of vital importance; their needs must be recognized.

That an information system exists as an operating unit, that the librarian has day-to-day problems in administration, that there are difficulties in meeting these problems—all must be recognized. The information system must be designed and put together in such a way as to meet the conflicting requirements placed upon it, and it is the job of the system designer to accomplish this. In doing so he must draw upon the capabilities provided by technology and industry in the development and manufacture of various devices. Thus, the user, the information specialist, the librarian, the designer, and the engineer each has his role to play and, in each case, his needs and capabilities must be recognized.

Past conferences concerned with the information system problem, such as the International Conference on Scientific Information in November of 1958, have examined various aspects of these different areas of interest. However, they have been concerned, usually, with detailed results, such as specific "user studies", or details of different techniques for the organization of information, or the detailed characteristics of different types of devices. There has been no well-defined effort to lay out a general approach to the problems in tying together the needs and capabilities from all the diverse areas involved.

It was the purpose of the proposed workshop to initiate such an investigation. It was hoped in this way, not only that the task and responsibility of the systems designer would be defined but, in addition, that some picture of the needs for education and research in the field of systems design would be made evident.

In large part, the documentalists, the data processing system designers, the information scientists in general are concerned with an analytical process which must result in an operable system meeting the needs of all parties concerned in an optimal manner. Thus, this represents a professional activity of extreme complexity and importance. It is vital that it be recognized as such.

THE PROGRAM

The structure of the workshop itself is of particular interest in its approach to meeting the needs outlined in the previous section.

There was an introductory session held in the evening of Tuesday, May 29. Wednesday and Thursday, May 30 and 31, then were devoted to four technical sessions, each concerned with a different area of responsibility for the system designer. The sixth and concluding session was held Friday morning, June 1; it was devoted to discussion of the requirements for education and research in this field itself—a topic transcending in importance all other aspects of the conference.

The four technical sessions on Wednesday and Thursday were the intellectual heart of the program, since it was here that the details of methodology were presented. Each session consisted of a session chairman, two invited speakers, and two additional participants to form, with speakers and the chairman, a panel of five. Each of the four sessions therefore involved some introductory remarks by the chairman, followed by the presentation of the two papers. This, in turn, was followed by discussion among the five participants on a less formal basis, discussion from the floor was encouraged.

In each of the four technical sessions, the two invited speakers were drawn from different areas of application, with emphasis on the technical information center on the one hand, and on the business information system on the other. Since, in general, it appears to be true that the business information system designer has a firmer foundation on which to work, it was hoped that this should provide an illuminating contrast on the comparative status of development and the different problems and difficulties in the two areas. These differences were indeed illuminating and pinpoint a significant consideration to which I will return later.

The four technical sessions were concerned with the designer's responsibility and his methodology in each of four problem areas:

In System Usage Analysis

The intent here was to have each of the invited speakers describe, in terms of his own concrete experience in the actual implementation of specific information systems, his impression of the designer's responsibility for determining the needs of usage and his method for working with the potential users in so doing. He was asked to present details, not on what the results were, but rather on how he obtained them. For example, did he use an interview, a questionnaire, analysis of past statistics, or some other approach? He should have presented some evaluation of the success, or lack of it, connected with the approach he used; hopefully, he would be very candid.

In Meeting Operational Problems

In this session, the concern was with such problems as the position of the information system in the organizational structure, the difficulties in obtaining and training adequate personnel, the problems in meeting budgetary limitation, and similar operational problems. Again, the intent was that each of the two speakers would present, in terms of his actual experience in meeting the needs of an information system with whose implementation he was concerned, his view of the designer's responsibility, and a description of the techniques which he used in meeting them.

In His Own Area of Technical Work

In this technical session, the interest was in how the designer selected the particular solution to the technical problems of system design itself. These technical problems include definition of file organization, item format, coding, information flow procedures, machine programming, report formats, etc. Again, however, it should be emphasized that we were not concerned with the details of a particular file organizational structure, for example, but rather with the criteria by which the one utilized was selected. It must be emphasized that the approach was to present this in terms of actual experience in selection.

In Equipment Functional Specifications and Selection

The intent of this session was to highlight the techniques by which the designer determines what equipment capabilities are required and the criteria he utilizes for the selection of the equipment used for implementation. These criteria may be, for example, cost, equipment sophistication, manufacturer reliability, or other considerations. The important point, for the purpose of this technical session, was not the details of what equipment was selected, but rather the details of the processes by which the equipment was specified or selected. Again, it was important that the papers presented actual experience in performing this function, and that the speaker should be extremely candid about the success or failure in the techniques which he used. The sixth and final session represented the culmination of the workshop, intellectually speaking. No profession is worthy of the name, unless it takes strong and effective steps to perpetuate and enhance itself through the education of new members. A new profession such as that of the information systems designer, especially must look to the future and put the education of its future members on a sound footing.

Two speakers were presented at this session: one, to discuss the needs for education and to delineate the various patterns of curricula now extant; two, to examine the needs for research of the academic level, particularly as made evident by this workshop, and especially as it might be coupled with an educational enterprise.

THE RESULTS

As has been mentioned, it was the intent that a set of papers be developed from the workshop proceedings. As a result, a total of 10 papers were developed and are included in this volume.

These ten papers represent the tangible results of the workshop. To assure the most effective comunication of these results to the attendees at the conference, they are being distributed in preprint form at the ADI conference and presented only in summary so as to allow the maximum time for discussion between the speakers and the conference attendees.

The intangible results of this conference were the development of an awareness of why research and education is necessary in this field, where that education should be done as opposed to where it is now being done, and what problems must be overcome.

We have been faced over the last fifteen to twenty years with a technological revolution in the field of information handling which can not really be compared to any previous technological revolution, at least in terms of the speed with which it has taken place. From the birth of a concept—at least of a technology which could fulfill that concept—twenty years ago to the fruition of it in successful engineering, marketing, and functional operation in that period has never occurred before. The automobile certainly did not arrive at the same state of fruition. Atomic power as a realizable and usable product hasn't yet arrived at this stage, although we do see a few installations of atomic power plants now.

This has meant that the people who could control this technology have had to grow up with it over a very short period of time. Along the way, they have had to develop the methodology to control this technological explosion. There is a very close relationship between the formalization in a discipline and the applicability of the computer to it. In order to use the computer, formalization must be introduced.

This historical development has seen the use of data processing equipment in numerical analysis, in business data processing, and now in the library type of problem. In numerical analysis we were fortunate in applying computers and computer-like devices to a field where there was a large amount of formalized knowledge on which to draw. The whole of mathematics and physical science—all of the problems to which the computer was applied—were well defined and well formulated problems. The tools which were available to the person trying to understand the computer were thus well defined tools.

In comparison with mathematics, business data processing is a relatively unformalized field, and yet the knowledge was there in basic accounting principles and the use of the punch card for many years in business data processing. Therefore, those concerned with the introduction of the computer, although their job was difficult and still has not been completely resolved, could draw on the knowledge available.

We are faced with a rather different type of situation in information retrieval. We are trying to apply this technology to a field which seems to have little formalized background on which to draw.

Robert Fairthorne, during the conference, made an observation: "Let's suppose that we had a computer which could read every single document in the library and do it instantaneously. Would this solve the retrieval problem? The answer is, obviously no, it wouldn't because we have completely ignored the intellectual problems".

On the other hand, and this is a point which Robert Fairthorne has previously made over many, many years, let's think of a different type of monster. Let's think of a monster which could absolutely accurately make the decisions as to whether a document was desired or not, and then ask the question, "Has this solved the retrieval problem?", and the answer is again, "No". The reason is that a library is not only a means for communication, it is also a means of storing physical things in a physical form. There are not only the intellectual problems of communication with the library but the physical problems of handling the physical forms in the library.

This was brought out very clearly during the conference because the participants were talking essentially about these two aspects of the same problem, the physical aspects and the intellectual aspect.

The dichotomy was that those concerned with the physical aspects were talking about the tools, the technologies, the methodologies in mathematical models which are applicable to the physical characteristics of equipment. Those who were approaching it with other tools such as the documentation methodologies and tradition library tools were concerned with the intellectual problem.

Some come into this type of situation from an equipment standpoint, with the problem of determining how to use equipment in terms of its limitations, its speeds, its capacities, and its costs. Others come into the problem of system design from the problems of communication, of using the device, of communicating with the device, of communicating with the device and with the people who need the information.

The necessity then is to bring these two people together, to provide a leavening effect upon the quantified approach of the mathematician and the engineer, and a comparable leavening effect upon the librarian. This, then, gives rise to the very problem we are concerned with; namely, the communication of both aspects of the problem and the research into their interrelationships. This is why education and research is necessary in this field: to build up a capability of handling both aspects of the problem, both the physical considerations and the intellectual ones.

The methods of systems analysis as Bourne has outlined in his paper, which have been traditional in the field of business data processing for many years, are immediately applicable here, too. But something more is needed, and this is an understanding of the intellectual problems.

These two approaches have got to come together, and this, then, is the role of education and research.

be achieved, in that those concerned with the physical aspects are talking about the tools, the technologies, the methodologies, the mathematical models which are applicable to the physical characteristics of equipment. Those who were approaching it with other tools are the documentation methodologies and tradition [illegible] tools are concerned with the intellectual problems.

[illegible] the type of situation about an equipment [illegible] with the problem of determining how to use equipment in [illegible] its speed, its capacities and the costs. Others [illegible] the problem of system design from the problems of [illegible] of using the devices of communications with the devices of communicating with the device and with the people who need the results.

The necessity, then is to bring these two groups together, to have a tempering effect upon the quantified approach of the engineer and the engineer, and a comparable tempering effect upon [illegible]. This, then, gives rise to the very problem [illegible] namely, the communication of both aspects of [illegible] [illegible] into their own relationships. This is the [illegible] [illegible] in this field, to build up a [illegible] of [illegible] of the problem, both the physical and the [illegible] [illegible] ones.

The [illegible] of [illegible] [illegible] as Becker has outlined in [illegible] which have been [illegible] in the field of [illegible] data [illegible] for many years, are not [illegible] applicable here [illegible] [illegible] is needed, and this is an understanding of the [illegible] [illegible].

[illegible] have got to come together, and share [illegible] of education and research.

A REVIEW OF THE METHODOLOGY OF INFORMATION SYSTEM DESIGN

CHARLES P. BOURNE
Stanford Research Institute
Menlo Park, California

INTRODUCTION

THERE is very little recorded information that describes *how* a new information system is designed. It is certainly true that systems are being designed, but the designers usually have a very difficult time describing exactly how they go about achieving the design. Essentially nothing exists in the way of written rules or procedures to follow, and the methodology has certainly not progressed to the point where there is any "handbook" design approach. It is the intent of this paper to note and record, in a summary fashion, some of the tools and methodologies that have been used or proposed for the general and idealized specification and design of information systems. The approach is idealized to such an extent that it may be difficult to follow it in practice. However, it should at least serve as a guide or goal. An information system is defined for the purposes of this discussion as a complex of people, equipment, and procedures—working together to provide needed information to a group of users. The discussion assumes the context of a library type of information system, but the approach can certainly be generalized to many other types of information systems.

In general, the design of any information system usually proceeds through the three-step procedure shown below.

(1) Problem Definition and Determination of User Requirements
(2) Synthesis and Design
(3) Evaluation.

The designer may proceed through these steps in sequence from the first step through the third step, and then make repeated passes through the cycle, incorporating additional information gained at each step. For example, the result of the evaluation step may suggest a better way to synthesize the system, and so on. There may be variations in the sequencing through the three steps. Instead of a single pass or iterations through the three-step sequence, there may also be some minor cycling between any two of the steps. The designer, for example, may start the problem with some predetermined system design. This would probably be the approach taken by the salesman for an equipment manufacturer who is directly or indirectly trying to design the system to utilize a specific type of equipment. In effect, the salesman has taken some short cuts through the design process, and has immediately discarded some feasible designs. Similarly, the analyst who is determining user requirements may have different kinds of systems in mind as he starts determining the requirements. Consequently, he may execute several iterations between the first two steps before any appreciable effort is spent on the third step. Similarly, during the evaluation process, some iterations are again executed in a minor loop. Trade-offs and balances are continually made between all the considerations and constraints imposed upon the system.

As one further point, it should be noted that the design cycle can be done either on an idea plane (i.e., a paper study), or it can be done on a physical plane (i.e., the development of some prototype equipment or pilot installations).

The three design steps are described in some detail in the following sections, along with a discussion of the tools or methodologies that can be used to assist the designer at particular points.

PROBLEM DEFINITION AND DETERMINATION OF USER REQUIREMENTS

There are four different approaches that can be of use at this point. Each of them will cover some different facet of the problem, and they will all serve to provide the information to describe and define the problem. Depending upon the circumstances, any number of these four approaches can be taken:

(1) Statistical studies of the quantitative parameters

(2) Functional analysis of the job processes
(3) Identification of relevant external influences
(4) Determination of specific user requirements.

Each of these approaches is described in more detail in the following sections. The term "user" refers to the ultimate user of the information, and not the system operator.

STATISTICAL STUDIES OF THE QUANTITATIVE PARAMETERS.

The first of these approaches is to conduct statistical studies of the quantitative problem parameters where this is possible. This means determining the file size, file mix or file composition, file media, traffic patterns and volumes of the paper forms and other materials that go through the system, and obtaining measures of as many important quantitative problem parameters as possible. This also includes measures of such things as the amount of time allocated by the users for satisfying their information needs.[21] Determining quantitative parameters is usually a straightforward arithmetic process that can be helped by a number of special sampling techniques, such as reference counting to find out what journals and what kinds of material are of interest to the users. There is nothing especially elegant about this approach, and it usually simply requires brute labor. Some of the parameters that people have felt it important to measure are listed in Table I.

Table I

SOME QUANTITATIVE MEASURES OF INFORMATION SYSTEMS

File Size
File Growth Rate
Number of Service Requests
Number and Type of Personnel Served
Number and Type of Operating Personnel
Total Cost and Per Unit Cost of Providing Service
Response Times to Obtain First, Final, and Major Group of Search Products
Distribution of Ages of File Items
Physical Form and Format of Search Products
Amount of Irrelevant Search Products Furnished
Number of Concepts per Question

Functional Analysis of the Job Processes

The second approach is a description of the processes being performed, or contemplated—that is, a functional analysis or job description. Flow charting of the operations and of the material or information traffic through the system is one useful approach here.

Some preliminary efforts have been made to devise a flow chart for the general operations performed in an information system.[7,42] The recording of detailed descriptions of the activities of the personnel in the system as determined by observations is another way to determine what is being done, and what possible duplication or mis-application of effort exists. The recording of the utilization of major units of equipment or facilities in the system (i.e., a machine load chart) provides the same type of information as the personnel activity chart, but for equipment rather than people. A forms analysis chart (i.e., a display of all the activities of an organization and the paper forms used at each step) can also provide useful information about the processes being performed. A number of well-developed techniques are available to help conduct the functional analysis.[3]

Identification of Relevant External Influences

The third approach is to identify the external forces or ground rules that might influence the design and operation of the system. In most instances the designer has to take into account some factors over which he has virtually no control, and which place a definite constraint on his freedom to design the system. Examples of such factors that must be considered are: (1) the number and type of personnel that will be available to operate and install the system; (2) the amount of time and money available for the installation and operation; (3) the requirement to be compatible with other systems; and (4) the requirement to use some equipment or services that already exist in other parts of the organization. These are considerations or constraints that face the designer. It is important that these external constraints be identified early in the design process.

Determination of Specific User Requirements

Requirements that Stem from External Sources. Some user information requirements, possibly unknown or unrecognized by the user, stem from sources other than the immediate user. For example, plans and schedules by research managers for the work of their research personnel generate some special information needs for the research staff, regardless of the personal interests of the staff members, and perhaps even unknown to the staff members. This is an example of the generation of a user requirement from sources external to the user.

A school director's change in the school curriculum—made with or without notification to the library—would impose upon the library some new requirements for information because of the new subject material or increased demands for service.

Practicing physicians should have timely information about new vaccines, clinical tests, or medical treatments that have significant relevance to the health of their patients, regardless of their own personal interests or self-stated information requirements. These, again, are requirements that stem from sources other than the immediate users; as indicated, they may be unknown or unrecognized by the user.

Identification of User Habits, Preferences, and Idiosyncracies. Another kind of requirement that should be examined, or at least identified, are the habits, preferences, and idiosyncracies of the users—those things that would make the system more comfortable to the user but are not crucial to the operation of the system. This includes such things as: (1) the user's preferred terminology; (2) the form and media of the search product that he is used to working with; (3) the user's prior knowledge of reference tools and workings of the information system; (4) and various habits or gimmicks that make the system more comfortable for the user. These special preferences need not be absolute design goals, but they should be identified so that they may be included in the design where possible.[22]

Identification of Specific Requirements. The third kind of requirement, the kind which has received most of the attention to date, are those requirements that make up the main design goals or specifications of the information system.* These are the goals that the system must try to meet in order to satisfactorily serve its users. Some of these requirements may be recognized and easily stated by the users (e.g., "I want my searches back within one day."), while others may be implied or obtained from the user without relying on his statement or opinion

* Some good state-of-the-art and summary papers have been written on this topic.[35,48,50]

of what his needs are. Examples of specific and stated requirements are: tolerable search response time; tolerable amount of irrelevant material provided; preferred intervals for current awareness reporting; literature time span for retrospective searches; and the preferred form of search product. These are examples of some of the parameters or factors that can be determined directly from the user in one or more of the methods listed in Table II and described in more detail below. Most studies of this type could be classed as opinion surveys, rather than surveys for the collection of quantitative data.[28] The problem with this approach is that this type of observation gives little indication of what the optimal or ideal media might be.

Another factor that should be considered here is the basic limitation that the user has in his ability to read and absorb different amounts of information. The human's "input channel capacity and processing rates" are factors that might be considered in the scheduling of the delivery of information.

Traffic Observation. A direct observation of the human and material traffic in the information system can indicate such things as what information or reference sources are used, and at what times, and at what places.[9,20,36,43] Several studies have been made, for example, of withdrawal records of libraries to determine the fluctuations of the demand volume on a seasonal or daily basis in order to learn

Table II
METHODS OF DETERMINING SPECIFIC USER REQUIREMENTS

- Traffic Observation
- Direct Questioning of User
- Indirect Questioning of User
- Questioning of People Who Have Provided Information Services to the Users
- Diary Keeping by the Users
- Postulate a New System and Invite Service Requests
- Pose System Specifications for Users' Debate
- Monitor the Genesis of Information Needs of an Active Group of Users
- Controlled Experiment
- Perturb the Present System

when the demands for service originate, and what the trend and ranges of expected traffic will be for the system. Detailed examinations of the circulation records and reference lists (i.e., citation counts) can indicate which journals, documents, or file items are of the most interest to the users.[17,30,51] Such studies could also indicate what types of information are required at different stages in a researcher's project. One point of caution here is the fact that rankings by reference counts may not correctly indicate the ranking of utility to the users.[8]

Direct Questioning of User. The most common approach used to determine user requirements is to ask the user, by a personal interview or a questionnaire, direct and specific questions about his needs and preferences.[23,28,31,52] This is certainly the easiest and most obvious approach to take, but it is full of pitfalls and inaccuracies, and must be done with extreme care and skill in order to obtain meaningful results. It does have a definite value in providing a first estimate or approximation of what the true needs are. The questionnaire or interview must be carefully designed and administered to keep from suggesting any specific answers to the interviewer. In order to maintain the interest and cooperation of the user being queried, the interview should ordinarily not require more than about 45 minutes to complete. This means that the questions should be concise, unambiguous, and limited to gathering essential data that could not easily be obtained by other means. Examples of some questions that have been asked with this approach are given in Table III.

A major argument against the direct questioning of the user is that it may be extremely difficult for him to distinguish between *need* and *habit*. What he feels that he would like to have may not actually be what he needs.[5] In some instances it has been argued that the user should not be asked to state his requirements because he is not qualified to do so—and that the needs or requirements should be prescribed by the expert in information system design.[48]

Indirect Questioning of the User. One other direct questioning technique is to probe in depth the circumstances surrounding the users' requirements for information.[7] For example, sit down with a user and ask him detailed questions about the last search that he conducted—that is, reconstruct a real situation. What was the time schedule that he was really trying to meet? What was the breadth of coverage that he really had to have, and what were his special needs

Table III

REPRESENTATIVE DIRECT QUESTIONS ASKED WITH INTERVIEWS OR QUESTIONNAIRES

1. During the last time that you performed a search, or had a search performed, approximately how long was it from the time you made your request until you had received the major group of relevant references? Was this adequate, or did you really need the material sooner? (If needed sooner, how much sooner?)
2. How old were the most recent references turned up by your last search? Was this adequate, or did you really need more recent material?
3. Here is a list of performance measures by which retrieval systems can be judged. Would you please number them in the order that you feel is most important to you.
4. On the type of search we've been discussing, how long from the time you make your request can you generally wait for the results of a search which covers 50% of the potential sources?
5. An ideal reference retrieval system would be able to provide you with references to the current journal and report literature on a current basis. In terms of your current work, please list the requests you would make for such lists of references, specifying the desired time interval between searches.
6. Assuming that this information center will be organized to contain a fairly comprehensive collection of information in your subject field, which three foreign languages should be covered, and in which order of priority?
7. Are there any subject areas that you would like to have covered more completely by the existing information center? If so, what are these subject areas?

---

for that particular search? If this is done for enough people it is possible to get a composite picture of what some of the requirements might be. And if it is done carefully, this approach will perhaps provide more meaningful information than asking the user to state his needs directly. A suggestion box in a library may yield some clues on desires of users, as well as an indication of their ignorance of the operation of the present system.

Questioning of People Who Have Provided Information Services to the Users. Some useful information can be obtained by asking questions of the operators of the information system (i.e., librarians or information specialists) who serve the information needs of the ultimate users.[52] There are two main problems with this approach. First, the operators can only describe their experiences with the current users of the system, and cannot speak for the great number of people (e.g., the dissatisfied user) who, for many reasons, including inadequacy of the information service, do not visit the information center or contact the person who is providing the service. Thus they can only speak for a part of the potential user population. Second, there may be a tendency for the information personnel to be somewhat biased, consciously or unconsciously, in their responses to the questions. There is a natural inertia and tendency to keep from suggesting needs or services that are extreme departures from what is traditional or currently being done, or that could require a considerable amount of extra effort. It is also natural for the people in any service organization to be less critical and find less fault with their own service than would be the case with an outside observer.

Some useful information can be obtained from monitoring and analyzing the written records generated or collected by these service people (e.g., collections of reference questions). Studies of such records have provided information on such things as: (1) the complexity and form of logic used in the reference questions;[24,25,54] and (2) the nature of the information requested (conceptual structure or viewpoint).[25] The danger of studying the query rather than the original motivating requirement lies in the fact that the users' questions represent a translation of his request into a form to which he thinks the information center has a capability of responding.

Diary Keeping by the Users. One extension of the direct questioning technique is to require the users to maintain diaries to record their actions and needs for information.[4,26,44] The user does not answer specific questions, but records things as they happen and as they occur to him. In the diary he can record incidents of the need for information and the action taken to get it. In this sense, the diary will contain an open-ended monologue rather than a dialogue. The recording may be done on a comprehensive basis to record everything that occurs, or it may be done on a sampling basis, with entries recorded oniy at specified intervals. The diary methods have been used for test periods ranging from a week to several months.

A large amount of information may be obtained by this method, with perhaps a sizeable fraction of it being of little value for determining specific requirements. At the end of the test period a large volume of material must be examined, correlated, and interpreted, since the test is an unstructured affair that collects all information for later study. One problem with this method is that it interferes with the daily work of the user, and this may be intolerable. In addition, it may be difficult to obtain full reporting, even under controlled conditions. Another important consideration is the fact that the act of recording and formulating the diary statements does in fact influence the way in which the user will satisfy some of his information needs. The note-taking may make him consciously think about what he is doing, and cause a change from his normal behavior. In this case, the instrumentation set-up probably disturbs or changes the process to be measured.

Postulate a New System and Invite Service Requests. Some information can be obtained by postulating a particular system design, and inviting potential users to request service from such a system. This allows the users to ask for service that cannot be provided now with the present system, and removes some of the constraints on his requests. The main difficulty with this approach would seem to be the fact that it may be difficult for the user to project himself into this pretended environment and examine what services he might like, or what types of responses he would prefer. This may be somewhat like asking a person in the early 1900's what features and performance he would like to have in an automobile. The American Institute of Physics is currently using a modification of this method. They have invited physicists to submit search questions to a postulated system. These search requests are made in terms of the physicist's current work, which he also describes. Analysis of these search questions provides guidance in developing indexing systems for physicists and for testing experimental systems.

A more concrete version of this approach is to actually develop an experimental or pilot system, and allow it to be used by the prospective users, while noting their comments about the good and bad features of the system.

Pose System Specifications for Operators Debate. Instead of inviting service requests for a hypothetical system, an alternate approach would be to let the operators debate the arguments for and against each of the proposed features. This would help to bring out considerations that might otherwise go unnoticed. This approach has been used with some success in one case to help define the objectives of a university library.[42]

Monitor the Genesis of Information Needs of an Active Group of Users. Another approach is to find a group of people who are working on a project or an experiment, and then stand off to the side and monitor their activities from the inception of the project through its final conclusion, to try to find out when and where the information needs arise, what type they are, and what the circumstances were surrounding their needs at that time.

Controlled Experiment. One other approach is to use a controlled experiment, where several groups of people work on the same or identical problems, and are given different degrees and types of information resources, to see how their productivity and effectiveness is affected by the type and amount of information provided. This could serve to indicate the relative importance and value of different types of service.

Perturb the Present System. One final approach is to perturb the present system with changes in service or procedures to observe the effect upon the user, and to see how he reacts to these changes in the system.

SYNTHESIS AND DESIGN

There seem to be four different approaches that are used for the actual design or synthesis of an information system:

(1) Copy successful systems
(2) Develop from existing building blocks
(3) Develop a brand new or "invented" system
(4) No conscious design.

Each of these four approaches is described in some detail in the following sections.

COPY SUCCESSFUL SYSTEMS

One of the most common methods used to design an information system is to copy, extrapolate, or modify some existing and successful system. The repeated use of the Uniterm indexing system in almost identical manners in many installations represents such an approach. A manual Uniterm system in one company looks very much like a manual Uniterm system in almost any other company—the whole idea seems to be lifted and copied into many different installations, with little design or analysis work by the individual installations. Repeated use of the L.C., U.D.C., and Dewey classification systems in hundreds of libraries is probably another good example of this approach where—instead of inventing your own system—you simply copy, extrapolate, or slightly modify a system that is apparently successful somewhere else. This is a useful approach, especially for the initial system design, with further re-design being done to make a closer match to the specific problem.

Develop from Existing Building Blocks

Another common design approach is to develop new systems from existing building blocks. In this case, the designer looks around to find all the techniques, tricks, equipment, and gimmicks that are available—pulls a piece from here and a piece from there, and then patches these together in some coordinated manner while taking care of the loose ends—and this becomes the system design. Most computer searching systems, for example, are mechanical equivalents of their corresponding manual systems, with many of the operations, functions, and activities done in the computer system being identical to those of the manual system. The designer has simply changed a few blocks in the system diagram in order to replace some manual operations with their machine equivalents. This building block design approach is used by people in many fields (e.g., architects, mechanical engineers, etc.) and its success depends in large part on the designer's awareness of the availability and capability of thousands of different kinds of building blocks.

Most designs are made without following any comprehensive sets of design rules, and without proposing all possible alternatives for subsequent evaluation. As a rule, no checklists or rules are followed in the design process to ensure that all points and alternatives have been considered. It is usually not necessary to check all possible combinations or designs, since you can often take advantage of the fact that for some range of functions, one component is always better than all the other alternatives possible (e.g., one best duplicating machine, one best typewriter, etc.), and can be suggested without considering all the other possible alternatives. In any case, design methodology has not progressed to the point where a handbook on automatic design can be made. However, some systematic design approaches have been developed, for application in limited areas, that combine some of the design and evaluation functions in such a way as to optimize the design of some sub-components of an information system.[1,32,41]

Develop A New Or "Invented" System

A third approach is original invention. That is, the designer develops a new system whose major components or technique have no obvious ancestral background. However, this is an idealized situation, and very few entirely new systems are developed, since the designers usually build by simple extension and extrapolation on the work of others.

No Conscious Design

The last design procedure is *no* design. This is the situation when systems come into being by historical accident, or by edict, or for some other reason, with no conscious design effort. There are possibly some systems in operation at this time that have been developed in this manner.

EVALUATION

Evaluation can be separated into two areas: (1) a performance evaluation, in which we consider the basic capability of a system and say nothing about the cost; and (2) an economic evaluation. The two should really be talked about together; however, it is procedurally simpler to determine and discuss them separately.

Performance Evaluation

The performance measurements can be categorized into two approaches: (1) single-criterion evaluations on specified parameters; and (2) composite figure of merit evaluations.

Single Criterion. For the single-criterion evaluation the experimenter chooses a single operating parameter—such as the relevancy of search products—and then evaluates this single system or compares it with other systems on the basis of this one criterion or performance parameter. Many different criteria could be selected as the basis for evaluation.[7,27,29]

The simplest single-criterion evaluation technique, and the one that is used most commonly, is to pose some sample questions to one or more systems, and observe the *volume* of references furnished. If the search seems to provide an adequate number of references, then the performance is judged to be satisfactory. This is the type of subjective test conducted consciously or unconsciously by many users and operators of current systems. It may be adequate for some purposes, but it does not provide a controlled and critical test of the system, since the volume of references furnished is not a complete measure of a system's performance.

A more effective test is to compare one system with one or more other systems, real or idealized. There have been several experiments of this type, but the test usually provides only measures of relative performance that cannot be extended or extrapolated to the comparison of other systems. A better test would be to compare the performance of a system against some standard, such as a perfect or ideal system. In addition, it would be more useful to obtain absolute rather than relative performance measures.

One example of a single criterion evaluation is the work underway by Cleverdon to compare the performance of five indexing and classification systems: (1) the Universal Decimal Classification; (2) a faceted classification developed for this experiment by the Classification Research Group in England; (3) a conventional subject-heading system; (4) a coordinate indexing system; and (5) a semantic factors code developed at Western Reserve University.[11-16,39] The first four of these systems have been studied under controlled conditions with a collection of 18,000 documents in the field of aeronautics. A number of factors, besides the indexing method, which might influence the test results (e.g., the indexer's training, time allowed for indexing, etc.) were indentified in the early design of the experiment, and were varied in a controlled way so that their effects could be noted separately from the effects of variations in the indexing method. However, there are still so many variables that might influence the performance that it is extremely difficult for a single experiment to take all of them into account. However, some useful results can still be obtained. In this case, the test procedure was to: (1) select a document from the 18,000 indexed in the collection; (2) formulate a question that could be satisfactorily answered by that document; (3) formulate each question into the search language of each of the five systems; and (4) note whether or not the search located the original source document that generated the question. Several hundred different questions were used in the test. The basic performance measure was the over-all fraction of searches that were successful in locating the source document. The preliminary results indicated a range of about 70 to 80 percent recall of source documents for the first four indexing systems mentioned above. A secondary performance measure was the over-all relevance of the documents produced by the search. That is, the relevance or lack of relevance of all of the search products were assessed, to provide a single relevancy figure for each indexing system.

Another example of a single criterion is the work by Swanson to compare machine and manual searching.[46] In this case, the first part of the test procedure was to develop, by the following three steps, a standard for later comparisons: (1) form a collection of articles to serve as the library for this specific subject field; (2) form a large set of detailed search questions in a specific subject field with each question inspired by some particular article in the collection; (3) using experts in the subject field, examine each article in the collection in the light of each question, and assign a weighting factor to each question-article combination to indicate the degree of relevance. This amounts to a prior determination of all revelant responses in this collection of questions and articles. For comparison of the indexing or searching system under test against this standard, another group of subject experts transformed each of the previous search questions into a search instruction for the system under test, performed the search, and scored the retrieval results on the basis of a scoring algorithm that depends on both the fraction of relevant material retrieved, and the amount of irrelevant material retrieved. The relative weight given to each of these two factors is arbitrary and represents a judgment of the worth of a successful retrieval, and the penalty for irrelevant material. The retrieval score is given by the expression

$$\text{Score} = R - pI$$

$$\text{where } R = \frac{\text{Sum of relevance weights of retrieval articles}}{\text{Sum of relevance weights (for a given question) of all articles in the file}}$$

p = Irrelevance penalty factor
I = N–LR, is the effective amount of irrelevant material
N = Total number of articles retrieved
L = Total number of relevant articles in the collection

Borko has used similar criteria in his studies of the performance of machine indexing techniques. His expression for a retrieval score is also:

$$\text{Score} = R - pI$$

but his factor I is defined differently than Swanson's. In this case, I is defined as the fraction of retrieved documents that were irrelevant. That is, I equals the number of irrelevant documents retrieved, divided by the total number of documents retrieved. For this particular experiment, test questions were used that had actually been used previously as the basis for real searches by experts in this subject field.

The work of Cleverdon and Swanson has been criticized on the basis that the questions posed to the systems are synthetic and artificial, having been fabricated after reading an item in the collection. A more valid approach would be to use real requests made by active workers in the subject field represented by the file.

Mooers has suggested an evaluation technique similar to that actually used by Swanson.[37] Because the technique uses a sampling approach, and because the samples are studied intensively, the method is called the "intensive sample" test. It is suggested that this technique would yield an absolute measure of performance, rather than a performance measure relative to some other system. The performance is measured against information known to be in the file prior to the retrieval experiment. The test procedure requires a set of typical users to read assigned sets of typical file items, and formulate queries that would be answered in full or in part by one or more of the file items in their set. These queries are then posed to the file from which the sample was taken. Estimates of the relevancy of the sample documents to the queries, expressed as "Not Relevant(N)," "Relevant(R)," and "Crucial(C)" rather than by a numeric score, are made before the test by the users who framed the test questions. The list of test questions is then given to the system operators, who may process the request in any way they desire, in order to search the contents of the entire file. After searching, the results are examined to see which of the *sample* items have been retrieved, and for which question. Attention is paid only to the previously determined sample items and their relevance to the questions. Retrieved documents that were *not* in the samples are *not* considered in the scoring. Three numeric performance scores are computed for the entire list of questions or fraction thereof:

$$\text{Crucial Documents Ratio} = \frac{\text{No. of documents with C relevance cited}}{\text{No. of documents with C relevance in sample}}$$

$$\text{Relevant Documents Ratio} = \frac{\text{No. of documents with R or C relevance cited}}{\text{No. of documents with R or C relevance in sample}}$$

$$\text{Non-Relevant Documents Ratio} = \frac{\text{No. of documents with N relevance cited}}{\text{No. of documents with N relevance in sample}}$$

All the ratios take a value between zero and one. Good performance is indicated by a high Crucial Documents Ratio, which suggests that there is a high percentage of retrieval.

Bornstein has suggested an evaluation technique that yields a measure of the relative efficiency of different indexing systems using the following criteria: (1) the amount of relevant, partially relevant, peripherally relevant, and non-relevant information given in response to a query; and (2) the amount of time spent by the user in examining (screening, scanning, and/or reading) the responses.[6] In this technique, the test search questions are generated independently of the files to be studied, and are based on actual experiences and needs for information encountered by representative users.

One evaluative criterion suggested by Maizell is a "performance index" which registers the fraction of titles actually delivered to a user in a reasonable amount of time in response to requests.[33]

Composite Figure of Merit

One way to derive a composite figure of merit is to describe both the requirements and the performance in some quantitative form, compute a numeric measure of agreement between the requirements and the performance, weigh this measure according to the relative importance of that requirement, and then sum up these figures to get a single composite figure of merit for the performance of the system.[7]

Another way to derive a single figure of merit is to use a time-cost model in which the major performance parameters are described in terms of a time or cost penalty to the user.[7,34] For example, a system's performance in terms of search speed means something in terms of time to the user. Likewise, the false drop rate of a particular system also means something in terms of time to the user. Hopefully, these operating parameters of a system can be analyzed and related in some way to a penalty in time and cost to the user. This, then, could be another composite figure of merit that would be useful for a performance evaluation.

Cost Evaluation

Much has been said about the difficulties of determining the economic benefits of information systems.[10] However, even if we don't know the *value* (i.e., the equivalent revenue or income from providing the service of an information system), we can still determine the costs or expenses of alternative systems to do the same specified task, and choose the system with the least expense. That is, even if we can't compute a pay-off, or rate-of-return on the information system, we can still design to minimize the expenses for a given task.[2,7,47] There is certainly a practical value in performing an evaluation from a cost viewpoint.

There are at least three approaches that can be used to help provide a cost analysis or evaluation of an information system:

(1) Good Cost Accounting Procedures
(2) Simulation and Intensive Economic Analysis
(3) Cost-Performance Index.

Good Cost Accounting Procedures. The first aproach is to use good cost accounting procedures so that accurate cost estimates can be made of different parts of the system operation. These measured costs can then be compared with: (1) the costs to do the same job in another type of system; or (2) previously determined standard times and costs for similar or identical operations. This would provide very useful data and would be relatively useful to apply. However, data to describe either operating costs or measured time and cost standards is conspicuously absent from the information systems field. Some data has been reported, but often in such a way as to make it extremely difficult or awkward to extrapolate or compare with other systems.[38,45]

Simulation and Intensive Economic Analysis. A second approach is to make some comprehensive economic studies, possibly utilizing a computer simulation of the proposed information system.

in order to estimate the costs of such a proposed system. Such analysis and simulation procedures have been developed at SRI which take basic time and cost data for all the elemental tasks to be performed in a proposed information system, and then simulate on the computer the operation of the system over a time period of several years.[7] In this way, it is possible to estimate the system operating costs over wide ranges in such operating parameters as file size, accession rate, and volume of search requests. This tool makes it possible for a designer to take a paper description of a system and estimate, to a fair degree of accuracy, the operating costs of that system over wide variations of operating parameters.

If the designer is comparing alternative systems, then each information system can be put through the analysis and simulation procedure to find the operating regions (e.g., file size, accession rate, volume of search requests) where one system is economically more attractive than the others. This same approach can also provide estimates of the number and types of people and equipment required. Simulation can also be used as a product design tool, since once the proposed system has been simulated, the designer can go back and suggest changes in the performance specifications or characteristics of the system, re-simulate, and see what these changes have done to your system in terms of performance and cost.

Cost-Performance Index. The third cost evaluation approach is to develop a simple cost-performance index. Examples of such indices are productivity and effectiveness ratios, which describe the per unit cost (in terms of time or dollar costs) of doing different jobs. Other simple cost indices can be derived to provide similar evaluations on a cost basis.

ADDITIONAL COMMENTS

This concludes a very brief description of some of the tools that have been used or proposed for the design of information systems. Little has been said here about the pros and cons of each of the techniques, or where they can be used to best advantage. However, it is appropriate at this point to mention some of the weak points, and the work that probably should be done to improve these tools.

Problem definition (i.e., statistical studies, functional analysis, etc.) has not been adequately utilized to date because the people with the problems were not trained in the techniques that would have been useful and applicable. The tools and techniques that are standard in the fields of what might be called industrial engineering, systems and procedures, functional analysis, systems analysis, and so forth are available and well developed and could be used to good advantage. The reason they have not been used to date is that few of the people who have the problems have been trained with these particular tools. But there is no inherent or basic reason why good results cannot be achieved for this first step. That is, there is no fundamental problem here, but merely a problem of applying the right kind of energy and talent to this particular problem.

The determination of user requirements presents difficult fundamental problems, and some basic work needs to be done in this area. The tools and the techniques are not well developed for this work. Some work has been done to improve the methodology, but much remains to be done for working on this part of the problem.

The synthesis procedure can probably be categorized today as arts and crafts. It is not a scientific process—at least not as practiced to date. It requires a rather complete knowledge of many equipments, procedures, gimmicks, costs, capabilities, and limitations. And although considerable intelligence is required to design a useful system, this intelligence and ingenuity is not used in a particularly well defined manner. The best system design often depends upon the best knowledge of the building blocks or the gimmicks and tricks for putting things together, and not on the best knowledge of the mathematics, information theory, or some other special subject skill. The design procedure may not be able to use any new theoretical procedures or groundwork for all applications, although some mathematical techniques have been utilized for the design and organization of a file.[32,41]

A good design or synthesis requires a complete understanding of the characteristics and availability of all the component parts that might make up a system. At present, these data are not readily available in any coherent assemblage, nor is there adequate time and cost data to aid in the judicious selection of the building blocks. There is no real basic problem or fundamental limitation here. The problem can be helped simply by a better collection and dissemination of information about operating systems and their parts and components and capabilities. The major item needed to help the general synthesis is a better collection and availability of information to describe the potentially useful building blocks.

With regard to evaluation procedures, it seems that essentially nothing is available for the practical evaluation of system performance other than some very crude measures. Some preliminary work is being done, but nothing for universal and wide-spread application. Much work needs to be done in this area to develop better tools.

REFERENCES

1. Advanced Information Systems Co., Los Angeles, California, "Report on the Organization of Large Files with Self-Organizing Capability," an un-numbered Interim Report on an NSF Project (1961).
2. Arthur Anderson & Co., New York, "Research Study of Criteria and Procedures for Evaluating Scientific Information Retrieval Systems," Final Report on NSF contract NSF-C218 (March 1962).
3. Barnes, R. M. *Motion and Time Study,* 4th edition (Wiley, New York, 1958).
4. Bernal, J. D. "Preliminary Analysis of Pilot Questionnaire on the Use of Scientific Literature," In the *Royal Society Scientific Information Conference Report* (1948), pp. 101-102, 589-637.
5. Bernal, J. D. "The Transmission of Scientific Information: A User's Analysis," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 77-96 (NAS-NRC), Washington, D.C., 1959).
6. Bornstein, H. "A Paradigm for a Retrieval Effectiveness Experiment," *American Documentation,* Vol. 12, No. 4, pp. 254-259 (October 1961).
7. Bourne, C. P., et al., "Requirements, Criteria, and Measures of Performance of Information Storage and Retrieval Systems," Final Report for NSF on SRI Project 3741, Stanford Research Institute, Menlo Park, California (December 1961), AD-270 942, OTS price $10.50.
8. Brodman, E. "Choosing Physiology Journals," *Medical Library Assoc. Bull.,* Vol. 32, pp. 479-483 (1944).
9. Bush, G. C. et al., "Attendance and Use of the Science Library at Massachusetts Institute of Technology," *American Documentation,* Vol. 7, pp. 87-107 (1956).
10. Case Institute of Technology, Cleveland, Ohio, "Measurement of Value of Recorded Scientific Information," Final report on an NSF project, by the Department of Management, Operations Research Group (July 12, 1961).

11. Cleverdon, C. W., "ASLIB Cranfield Research Project on the Comparative Efficiency of Indexing Systems," (with discussion) *ASLIB Proc.*, Vol. 12, pp. 421-431 (December 1960).
12. Cleverdon, C. W., "An Investigation into the Comparative Efficiency of Information Retrieval Systems," *UNESCO Bull. Libr.* Vol. 12, No. 11-12, pp. 267-270 (November-December 1958).
13. Cleverdon, C. W., "ASLIB Cranfield Research Project: Report on the First Stage of an Investigation Into the Comparative Efficiency of Indexing Systems," an un-numbered report by the College of Aeronautics, Cranfield, England, of work supported by NSF (September 1960).
14. Cleverdon, C. W., "ASLIB Cranfield Research Project: Interim Report on the Test Programme of an Investigation into the Comparative Efficiency of Indexing Systems," an un-numbered report by the College of Aeronautics, Cranfield, England, of work supported by NSF (November 1960).
15. Cleverdon, C. W., "The Evaluation of Systems Used in Information Retrieval," *Proc. of the Int'l Conf. on Scientific Information*, Vol. 1, pp. 687-698 (NAS-NRC, Washington, D.C., 1959).
16. Cleverdon, C. W., et al., "The Cleverdon-WRU Experiment," papers presented by Cleverdon and others at a Western Reserve University conference in Cleveland, Ohio, April 1962.
17. Curtis, G. A. "A Statistical Survey of the Services of the John Crerar Library," M. A. Thesis, University of Chicago (1951).
18. Fishenden, R. M., "Methods by Which Research Workers Find Information," *Proc. of the Int'l Conf. on Scientific Information*, Vol. 1, pp. 163-180 (NAS-NRC, Washington, D.C., 1959).
19. Frarey, C. J. "Studies of Use of the Subject Catalog: Summary and Evaluation," in *The Subject Analysis of Library Materials*, edited by M. Tauber. New York Columbia University School of Library Service, 1953, pp. 147-166. This report describes 27 studies of the use of subject catalogs.
20. Glass, B. and Norwood, S. H., "How Scientists Actually Learn of Work Important to Them," *Proc. of the Int'l Conf. on Scientific Information*, Vol. 1, pp. 195-198 (NAS-NRC, Washington, D.C., 1959).
21. Halbert, M. H. and Ackoff, R. L., "An Operations Research Study of the Dissemination of Scientific Information," *Proc. of the Int'l Conf. on Scientific Information*, Vol. 1, pp. 97-130 (NAS-NRC, Washington, D.C., 1959).
22. Herner, S. "The Information Gathering Habits of American Medical Scientists," *Proc. of the Int'l Conf. on Scientific Information*, Vol. 1, pp. 277-286 (NAS-NRC, Washington, D.C., 1959).

23. Herner, S. "Information Gathering Habits of Workers in Pure and Applied Science," *Industrial and Engineering Chemistry,* Vol. 46. pp. 228-236 (1954).

24. Herner, S. "The Relationship of Information-Use Studies and the Design of Information Storage and Retrieval Systems," Technical Note 1 by Herner & Co., Washington, D.C., under contract AF 30(602)-1857, RADC-TN-59-136, AD-213 781 (December 8, 1958).

25. Herner, S. and Herner, M., "Determining Requirements for Atomic Energy Information From Reference Questions," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 181-188 (NAS-NRC, Washington, D.C., 1959).

26. Hogg, I. H. and Smith, J. R., "Information and Literature Use in a Research and Development Organization," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 131-162 (NAS-NRC, Washington, D.C., 1959).

27. Jaster, J. J., et al., "Evaluating Coordinate Indexing Systems," Chapter 7 in "The State of the Art of Coordinate Indexing," a report by Documentation Inc., Washington, D.C., under contract NSF-C-147 (February 1962).

28. Kee, W. A. "Must Library Surveys be Classics in Statistics?" *Special Libraries,* pp. 433-436 (October 1960).

29. Kent, A. "Minimum Criteria for a Coordinated Information Service," Tech. Note 10, by Western Reserve University, under contract AF 49(638)-357, AD-229 882 (October 1959). This report was described briefly in *American Documentation,* Vol. 11, No. 1, pp. 84-87 (January 1960).

30. Kilgour, F. G., "Recorded Use of Books in the Yale Medical Library," *American Documentation,* Vol. 12, No. 4, pp. 266-269 (October 1961).

31. Lane, W. A. "Staff Study on Dissemination of Technological Information About Materials and Materials Research, Part I Interim Report," Report MAB-51-SM of the Materials Advisory Board of the NAS-NRC, Washington, D.C. (September 1959).

32. Magnavox Research Laboratory, Torrance, California, "Mathematical Models for Information Systems Design and a Calculus of Operations," Final Report R-51 on Contract AF 30(602)-2111, RADC Report RADC-TR-61-196 (27 October 1961).

33. Maizell, R. E., "Standards for Measuring the Effectiveness of Technical Library Performance," *IRE Trans. on Engineering Management,* Vol. PGEM-7, No. 2, pp. 69-72 (June 1960).

34. Meier, R. L. "Efficiency Criteria for the Operation of Large Libraries," *The Library Quarterly,* Vol. 31, No. 3, pp. 215-234 (July 1961).
35. Menzel, H., et al., "Review of Studies in the Flow of Information Among Scientists," an un-numbered report of the Bureau of Applied Social Research, Columbia University (September 1960).
36. Menzel, H. "Planned and Unplanned Scientific Communication," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 199-244 (NAS-NRC, Washington, D.C., 1959).
37. Mooers, C. N. "The Intensive Sample Test for the Objective Evaluation of the Performance of Information Retrieval Systems." Report ZTB-132 of the Zator Co., Cambridge, Massachusetts, RADC Report RADC-TN-59-160 under contract AF 30(602)-1900 (August 1959).
38. Mueller, M. W., "An Evaluation of Information Retrieval Systems," Memo Report 7170 of Lockheed Aircraft Corporation, Burbank, California (September 30, 1959).
39. O'Connor, J. A review of the Cleverdon reports. *J. Documentation,* Vol. 17, No. 4, pp. 252-261 (December 1961).
40. Overmeyer, L. "Test Program for Evaluating Procedures for the Exploitation of Literature of Interest to Metallurgists," *American Documentation,* Vol. 13, No. 2, pp. 210-222 (April 1962).
41. Schoffner, R. M., "A Technique for the Organization of Large Files," *American Documentation,* Vol. 13, No. 1, pp. 95-103 (January 1962).
42. Schultheiss, L. A., et al., *Advanced Data Processing in the University Library* (Scarecrow Press, Inc., New York, 1962).
43. Scott, C. "The Use of Technical Literature by Industrial Technologists," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 245-266 (NAS-NRC, Washington, D.C., 1959).
44. Shaw, R. R. "Pilot Study on the Use of Scientific Literature by Scientists," an un-numbered report to NSF (November 21, 1956). This report abstracted and analyzed all major use studies performed up to that time.
45. Stevens, N. D., *A Comparative Study of Three Systems of Information Retrieval* (Rutgers University Press, New Brunswick, New Jersey, 1961). Reviewed by P. A. Zaphyr in *Computing Reviews,* Vol. 3, No. 4, p. 159 (July-August 1962).
46. Swanson, D. R., "Searching Natural Language Text by Computer," *Science,* Vol. 132, No. 3434, pp. 1099-1104 (21 October 1960).
47. Taube, M. "Cost as a Measure of Efficiency of Storage and Retrieval Systems," Tech. Report 13, by Documentation, Inc., Washington,

D.C. on Contract NONR-1305(00), AD-137 399 (December 1955).

48. Taube, M. "An Evaluation of Use Studies of Scientific Information," AFOSR Tech. Note 58-1050 on Contract AF 49(638)-91 to Documentation Inc., Washington, D.C., AD-206 987 (December 1958). Available from OTS for $4.80 as PB 138-260.
49. Thorne, R. G. "The Efficiency of Subject Catalogues and the Cost of Information Searches," *J. Documentation,* Vol. 2, No. 3, pp. 130-148 (September 1955).
50. Tornudd, E. "Study on the Use of Scientific Literature and Reference Services of Scandinavian Scientists and Engineers Engaged in Research and Development," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 19-76 (NAS-NRC, Washington, D.C., 1959). This paper provides brief summaries and references to 69 previous user studies.
51. Urquhart, D. J., "Use of Scientific Periodicals," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 287-300 (NAS-NRC, Washington, D.C., 1959).
52. Voigt, M. J., "Scientists' Approaches to Information," ACRL Monograph No. 24, (American Library Association, 1961).
53. Wayne, I., Questionnaire for the Evaluation of the ASM Information Searching Service, prepared and administered by the Bureau of Social Science Research, Inc., Washington, D.C., on Project 352 (1962).
54. Whaley, F. R., "Retrieval Questions from the Use of Linde's Indexing and Retrieval System," *Proc. of the Int'l Conf. on Scientific Information,* Vol. 1, pp. 763-769 (NAS-NRC, Washington, D.C., 1959).

Proceedings of NATIONAL MICROFILM ASSOCIATION

(11th Annual Meeting and Convention)

Edited by VERNON D. TATE

CONTENTS INCLUDE:

cloth 360 pp. illustrated $9.00

SPARTAN BOOKS

6411 CHILLUM PLACE N W • WASHINGTON 12 D C

NATIONAL SCIENCE FOUNDATION
WASHINGTON 25, D.C.
20550

Office of Science Information Service

November 27, 1963

VIA AIR MAIL

Dr. Charles P. Bourne
Stanford Research Institute
Menlo Park, California

Dear Dr. Bourne:

When looking through the Polish periodical, Aktualne problemy informacyi i dokumentacyi, vol 8, no. 2, 1963, I found an article on pages 13-20 which will be of interest to you, if you have not yet seen it. I am enclosing a copy of it. It appears to be an almost word-for-word translation of your article which appeared in American Documentation, vol. 13, no. 2, April, 1962, pages 159-168.

Sincerely,

Inez C. O'Brien

Inez C. O'Brien
Professional Assistant
Eastern Europe
Foreign Science Information
Program

Enclosure

[illegible] I have had no correspondence w/ anyone in Poland [illegible] that it would ever be published in this manner [illegible] AD article? [illegible] Am. Doc.! I can see it now, "Am. Doc., the world's most widely translated [illegible]!"

December 3, 1963

Miss Inez C. O'Brien
Professional Assistant
Eastern Europe
Foreign Science Information Program
National Science Foundation
Washington 25, D. C.

Dear Miss O'Brien:

Thank you very much for the letter and for sending a copy of the documentation article that appeared in the Polish publication. It came as a great surprise. It seemed to be, as you suggested, a literal translation of a previous paper of mine in American Documentation. I have had no recent correspondence with anyone in Poland and have had no reason to believe that this article would ever be published in this manner. Have you seen any other American Documentation articles published in a translated version. Perhaps we can use this as an endorsement for American Documentation! I can see it now, "American Documentation, the world's most widely translated documentation journal."

Thanks again.

Sincerely,

Charles P. Bourne
Research Engineer

CPB/na

002.2:62:05(100)

ŚWIATOWA LITERATURA TECHNICZNA W CZASOPISMACH

Ilość, pochodzenie, język, dziedzina wiedzy, indeksowanie, dokumentowanie *)

WSTĘP

Świat nauki w dobie obecnej tworzy ogromną ilość literatury technicznej (fachowej), której znaczna część nie trafia niestety do rąk tych, którzy mogliby z niej korzystać. Dzieje się tak prawdopodobnie z dwóch głównych przyczyn: tylko pewien wycinek literatury jest objęty sporządzaniem przeglądów dokumentacyjnych, indeksów tytułowych lub innych publikacji tego typu i tylko pewna liczba ludzi, którzy mogliby korzystać z informacji, zna technikę posługiwania się literaturą. Ilość literatury, która jest opracowywana w formie opracowań dokumentacyjnych i indeksów, może być wymierzona, ale na tej podstawie nie można ustalić, jaki procent potrzebnej informacji dociera do potencjalnych użytkowników. W wielu dziedzinach sporządzanie opracowań dokumentacyjnych i indeksów obejmuje tylko czwartą lub trzecią część literatury, lecz ta część zawiera informacje o istotnych osiągnięciach w badaniach i ich zastosowaniu, podczas gdy pozostała część obejmuje powtórzenia (przeróbki), opracowania popularnonaukowe i zwykłe przyczynki. Dlatego też liczba opracowań wtórnych (dokumentacyjnych) nie jest idealną miarą opracowań dokonanych, jakkolwiek daje dobrą miarę opracowań wartościowych. Innym trudnym zagadnieniem jest sprawa oceny rozmiarów opracowań literatury, dokonanych w abstraktach, w sytuacji, gdy całość produkcji tej literatury jest nieznana lub wysoce niepewna. W artykule Ch. P. Bourne'a jest zawarta taka ocena na podstawie danych szacunkowych, które zostały tu użyte jako materiał źródłowy. Należy zastrzec, że ze względu na różną wartość czasopism lub artykułów liczba ich nie jest dobrą jednostką statystyczną. Jednakże autor był zmuszony posługiwać się nią ze względu na brak innych odpowiednich danych ilościowych. Jest rzeczą niewątpliwą, że stosunkowo mało czasopism zawiera bardzo duży procent wartościowych informacji, a tylko te czasopisma powinny być objęte opracowaniami dokumentacyjnymi i włączone do systemów informacji naukowo-badawczej.

Literatura naukowa i techniczna powstaje w wielu krajach, wielu językach, w dużej rozpiętości, jeżeli chodzi o jej wartość, i — co najważniejsze — jej ilość stale wzrasta. Wiele osób indywidualnie oraz wiele organizacji czyni wiele wysiłków, ażeby znaleźć metody lub systemy, które by powstrzymały zalew wydawnictw. Należy wątpić, czy możliwe jest rozwiązanie kiedykolwiek tego problemu, na pewno jednak istnieją sposoby poprawienia tej sytuacji.

Jednymi z pierwszych kroków, jakie powinny być podjęte, jest określenie i opisanie problemu w sposób możliwie kompletny i dokładny. Ten cel przyświeca autorowi omawianego artykułu.

Ch. P. Bourne widzi w omawianym zagadnieniu elementy dwojakiego rodzaju:

1. Elementy, które nie dadzą się w żaden sposób wymierzyć. Na przykład: wartość zamieszczonego fragmentu informacji, koszt lub szkoda wynikła z niedostatarczenia właściwej informacji we właściwym czasie, względna wartość różnych form, środków i kanałów przekazywania, wreszcie stosunek wydajności pracownika technicznego lub naukowego do rodzaju, wartości i ilości otrzymywanych przez niego informacji.

2. Elementy, które można w prosty sposób wymierzyć i które dostarczają pewnych wskazówek dotyczących charakteru problemu. Można tu wymienić na przykład: źródła oraz rozdział informacji naukowej według krajów i języków, formy publikacji, cechy charakterystyczne literatury z różnych dziedzin, dane dotyczące ilości i wzrostu literatury, stopień ujęcia literatury w abstraktach i indeksach.

Artykuł Bourne'a koncentruje się na tych bardziej wymiernych elementach i daje — jeżeli chodzi o literaturę zawartą w czasopismach — syntetyczną ocenę samej tylko ilościowej strony problemu.

Najpierw poddano badaniom ogólną liczbę wychodzących na świecie czasopism i ich roczną produkcję; ustalono, w jakich krajach i w jakich językach są one wydawane; następnie, tam gdzie to możliwe, przeprowadzono podział rzeczowy według krajów. Uwzględniono przy tym takie dziedziny, jak chemia, medycyna, metalurgia — dla tych samych typów źródeł językowych i według krajów.

Przed przystąpieniem do bardziej szczegółowego omówienia należy stwierdzić, że przytoczone dane statystyczne nie przedstawiają kompletnego i zwartego obrazu. Dane te dotyczą wyrywkowo różnych lat.

Reasumując należy stwierdzić, że zebrane dane są do pewnego stopnia niekompletne i nieaktualne i dlatego końcowe wnioski powinny być przyjmowane krytycznie. Jednakże pomimo nie-

*) Na podstawie artykułu Charles P. Bourne.: The world's technical journal literature: an estimate of volume, origin, language, field, indexing and abstracting. Amer. Docum., 1962, nr 2.

doskonałości tych danych zawarta w artykule ocena daje — w warunkach obecnych — najlepsze pojęcie o ogromie omawianego problemu.

OMÓWIENIE PROBLEMU

A. Ogólna liczba publikacji

Ogólna liczba publikowanych na świecie czasopism naukowych i technicznych oceniana jest na 100 tys.. Jednakże ewidencja czasopism, bardzo starannie i dokładnie prowdzona przez Oddział Nauki i Techniki Biblioteki Kongresu dla Narodowej Fundacji Nauki, wykazała, że istnieje ok. 30 — 35 tys. czasopism.

Biorąc za punkt wyjścia liczbę 30 tys. czasopism i zakładając, że każde czasopismo zawiera rocznie od 30 do 70 artykułów *) można przyjąć, że rocznie publikuje się na całym świecie od 1 do 2 mln artykułów naukowych. Jednakże bardziej realistyczna ocena wskazuje na 15 tys. ważniejszych czasopism i na milion bardziej wartościowych artykułów rocznie.

B. Kraj i język wydawnictw

Usiłowano z grubsza dokonać oceny pochodzenia całości publikowanych czasopism z punktu widzenia kraju i języka. Jednakże nie zdołano jasno ustalić rozmiarów tej grupy publikacji (tj. liczby czasopism, liczby artykułów itd.). Językowo ma w dalszym ciągu przewagę angielski, obejmując około połowy całkowitej produkcji. Pewne dane wskazują, że język rosyjski zaczyna dorównywać tradycyjnym francuskiemu i niemieckiemu. Z punktu widzenia kraju (według ogłoszonych sprawozdań) pierwsze miejsce co do ilości wyprodukowanej literatury zajmują Stany Zjednoczone przed Niemcami, Francją i Wielką Brytanią. Najnowsze dane statystyczne pochodzące od znawców przedmiotu wykazują, że Japonia może ubiegać się już o trzecie miejsce w produkcji literatury naukowej i technicznej, a prawdopodobnie prześcignęła Niemcy i Francję w chemii, biologii, medycynie i rolnictwie. Względne proporcje odnoszące się do każdego kraju zaczynają się znacznie różnicować w zakresie poszczególnych specjalności.

C. Wielkość i cechy charakterystyczne dziedzin nauki

Świat nauki wyodrębnił i ujął różne grupy dyscyplin naukowych w dziedziny takie, jak chemia, medycyna i matematyka. Granice między tymi dziedzinami nie są ściśle określone i w wielu wypadkach następuje przenikanie lub wchłanianie jednej dziedziny przez drugą. Na przykład istnieje tendencja do traktowania biologii, rolnictwa i medycyny jako odrębnych nauk, pomimo że rolnictwo i medycyna są biologią stosowaną. Podany rysunek ma za zadanie przedstawić niektóre z tych dziedzin. Jednakże pewna część materiału lub dziedzin nauki została zgrupowana w sposób raczej dowolny, zwykle zgodnie z tradycją akademicką, stowarzyszenia zawodowego lub służby informacyjnej, w celu utworzenia wyodrębnionego zestawienia materiału. Podział ten, jakkolwiek może sztuczny, niemniej jednak istnieje. Literatura każdej grupy stanowi przedmiot czyjegoś zainteresowania lub źródło informacji dla indywidualnych opracowań z tej dziedziny. To znaczy, że osoba pracująca w każdej z tych dziedzin będzie widziała pewną ilość literatury, która ją interesuje. Literatura ta stanowi jej techniczną informację. Przy tym pracownicy różnych dziedzin mogą być zainteresowani tą samą literaturą. Dla uproszczenia obrazu graficznego na rysunku przeprowadzone są wyraźne granice pomiędzy sąsiadującymi dziedzinami. Jednakże nie oznacza to, że istnieje np. wyraźna granica między chemią a naukami biologicznymi. W rzeczywistości podział na dziedziny specjalistyczne na rysunku jest dość przypadkowy.

Najlepiej zorganizowanymi dziedzinami — jeśli chodzi o informację — są: medycyna, chemia, rolnictwo i nauki biologiczne.

Szacunkowe dane dotyczące względnych wielkości w zakresie liczby czasopism, liczby artykułów, kraju pochodzenia i języka oraz stopnia objęcia przez służbę informacyjną są pokazane na rysunku dla każdej z właściwych dziedzin.

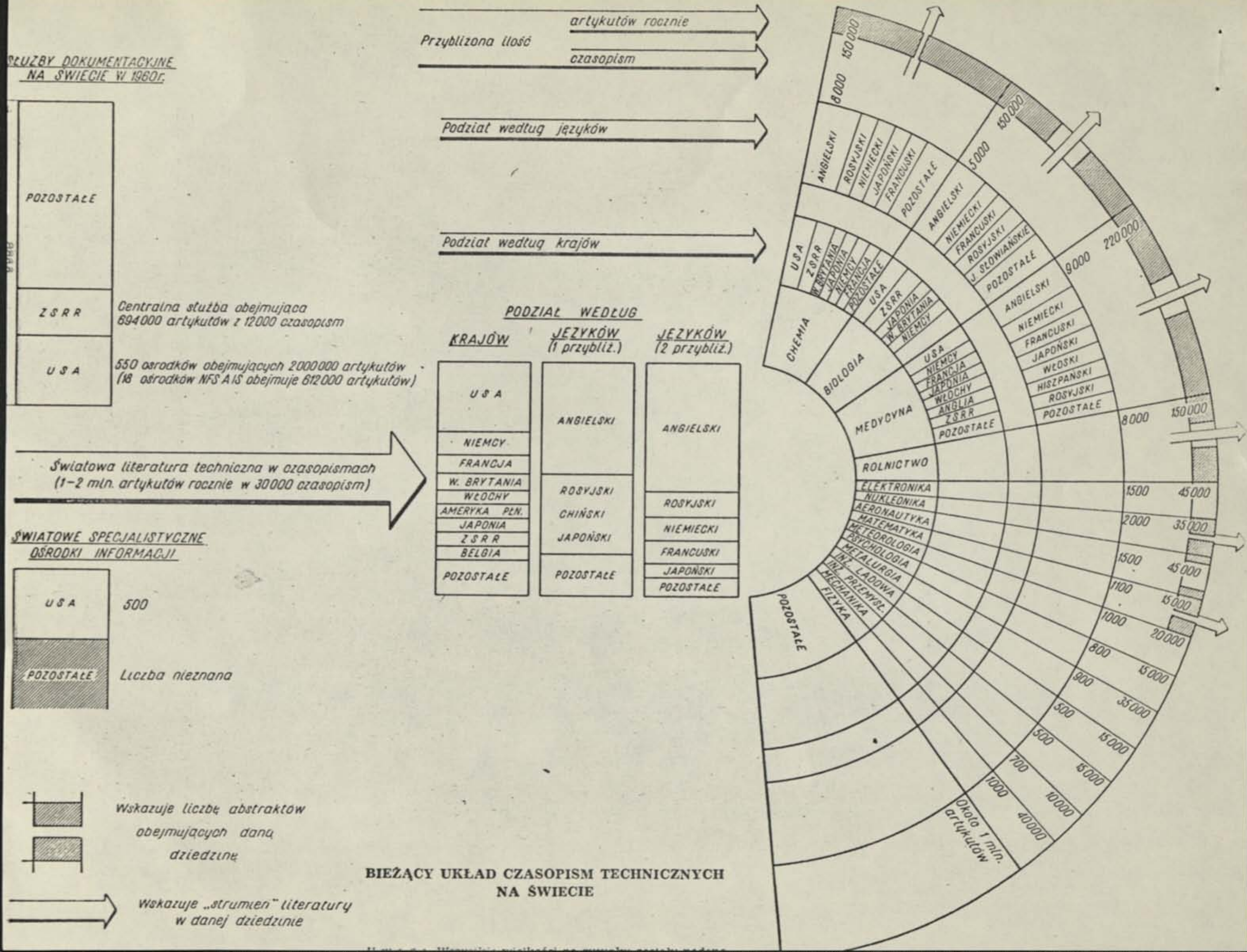
D. Służba informacji naukowej i technicznej

Na świecie istnieje obecnie 3500 ośrodków sporządzających przeglądy dokumentacyjne, z których 550 znajduje się na terenie Stanów Zjednoczonych. Ponadto w Stanach Zjednoczonych istnieje ok. 450 specjalistycznych ośrodków informacji, które utrzymują zbiory informacji na specjalne tematy techniczne (jak np. śnieg, lód i szron, wytrzymałość metali; zanieczyszczenie powietrza) dla służby zajmującej się wyszukiwaniem literatury.

Jak widać na rysunku, niektóre specjalności, np. chemia, są dobrze objęte służbą informacji naukowej; inne dziedziny są słabo obsługiwane w tym zakresie. Luki bez zakreskowania na rysunku oznaczają brak informacji w ogóle lub brak danych co do rozmiarów objęcia służbą informacji naukowej czy technicznej danej dziedziny specjalistycznej. Obecnie nie istnieje praktyczny zmechanizowany system wyszukiwania w całej literaturze światowej żądanych informacji z zakresu jakiejś specjalności, poza — może — metalurgią i chemią. Jedynie pracownicy dziedziny metalurgii mają obecnie możność korzystania z maszynowego systemu wyszukiwania informacji, utrzymywanego przez American Society of Metals i National Science

*) W rozdziale „Dane źródłowe" omówiono pochodzenie tych danych.

*) W rozdziale „Dane źródłowe" omówiono pochodzenie tych danych.



BIEŻĄCY UKŁAD CZASOPISM TECHNICZNYCH NA ŚWIECIE

Foundation, działającego przy Western Reserve University. Prawdopodobnie w najbliższym czasie również pracownicy medycyny będą mogli korzystać z maszynowego systemu wyszukiwania w National Library of Medicine, a pracownicy chemii — mogą go wykorzystywać w Chemical Abstracts Service.

O tym, jak daleko jest w USA do pelnego objęcia indeksami literatury światowej, może świadczyć fakt, że w 1961 r. wszystkie poważniejsze połączone ośrodki zajmujące się sporządzaniem abstraktów i indeksów — będące członkami Narodowej Federacji Służb Sporządzania Abstraktów i Indeksów (NSFAIS) — opracowały spośród 30 tys. tylko ok. 16 tys. czasopism naukowych i technicznych, sporządzając ok. 750 tys. abstraktów. Scentralizowany radziecki Riefieratiwnyj Żurnal opracował w 1960 r. ok. 694 tys. artykułów.

DANE CYFROWE

A. Informacje o wielkościach ogólnych i według krajów

1. Liczba czasopism i artykułów w czasopismach

C. P. Bourne przytacza w swym artykule wypowiedzi różnych specjalistów, dotyczące ogólnej liczby czasopism naukowych i technicznych na świecie. Liczba ta według fachowców waha się w granicach od 30 do 50 tys. czasopism zawierających ok. 2 mln artykułów rocznie.

W konkluzji autor stwierdza, że na świecie istnieje obecnie ok. 15 tys. poważnych czasopism publikujących w przybliżeniu milion artykułów, uzasadniając to następująco:

1) podsawowe służby informacji naukowej i technicznej w USA, członkowie NFSAIS, objęły swymi pracami 16 tys. czasopism;

2) scentralizowana radziecka służba referatowa, WINITI, objęła ok. 12 tys. czasopism;

3) 492 ośrodki służby informacji naukowej i technicznej Stanów Zjednoczonych zanotowały ok. 2 mln pozycji, przy czym nieznana jest ilość zbieżnych (zdublowanych) opracowań;

4) członkowie NFSAIS sporządzili 612 tys. abstraktów lub analiz, przy 25% zbieżności między poszczególnymi służbami;

5) scentralizowana radziecka służba referatowa objęła 694 tys. artykułów, zakładając pewne zbieżności.

2. Wielkość produkcji w poszczególnych krajach

W 1961 r. Charles M. Gottschalk z Biblioteki Kongresu, Oddział Nauki i Techniki, podał w komunikacie naukowym następujące wstępne dane szacunkowe i klasyfikacyjne, dotyczące czasopism naukowych i technicznych, publikowanych przez dany kraj czy grupę krajów (z wyjątkiem druków nie będących w sprzedaży).

Kraj pochodzenia	Liczba publikowanych czasopism
Stany Zjednoczone	5 950
Niemcy (NRF i NRD łącznie)	2 950
Francja	2 660
Wielka Brytania	2 090
Włochy	1 520
Ameryka Płd. (łącznie)	1 500
Japonia	1 470
ZSRR	1 450
Belgia	1 250
Chiny Ludowe	650
Hiszpania	300
Taiwan	150
Korea Płd. i Płn. (łącznie)	150

Pau W. Howerton, dyrektor Central Intelligence Agency na podstawie studium nad radzieckimi pracami Wydawnictwa Obcojęzycznego, opublikowanego w Moskwie, ocenił w prywatnej informacji, że ogółem ukazują się 3824 czasopisma radzieckie różnego typu, w tym 2295 z dziedziny nauk ścisłych i techniki.

W 1959 r. Narodowa Fundacja Nauki zacytowała pewne dane statystyczne z radzieckiej pracy referatowej, które wskazywały, że w 1955 r. publikowano w Związku Radzieckim ogółem 2026 czasopism wszelkiego typu, w tym prawdopodobnie 1100 naukowych i technicznych.

„Japonese Periodicals Index", opublikowany przez Japońską Bibliotekę Narodową, wymienia tytuły ok. 500 japońskich czasopism naukowych i technicznych oraz ok. 300 japońskich czasopism z zakresu nauk medycznych.

3. Literatura radziecka

Paul Howerton z Central Intelligence Agency ostatnio opracowywał literaturę naukową Związku Radzieckiego z roku 1958. Otrzymał on następujący podział przedmiotowy.

Przedmiot	Procent całości
Nauki fizyko-matematyczne	9
Nauki chemiczne	3
Nauki biologiczne	5
Nauki geologiczno-geograficzne	10
Nauki medyczne	17
Nauki inżynieryjno-przemysłowe	56

Central Intelligence Agency stwierdziła w 1960 r., że publikacje radzieckie z zakresu nauki i techniki przekraczają 700 mln słów rocznie, przy czym 7% tej ilości tłumaczy się na język angielski.

Biblioteka Kongresu i Narodowa Fundacja Nauki oceniają, że aktualnie ok. 120 radzieckich czasopism jest tłumaczonych „od deski do deski" na język angielski.

4. Całkowita ilość produkcji literatury naukowej i technicznej według języków. Pierwsza ocena

Dr Adkinson z Narodowej Fundacji Nauki ocenia, że ponad 1/3 światowej literatury naukowej i technicznej powstaje w ZSRR, Chinach i Japonii oraz że ok. 50% całej literatury technicznej i naukowej ukazuje się w języku angielskim. (Należy założyć, że artykuły pochodzące z ZSRR, Chin i Japonii pisane są odpowiednio w języku rosyjskim, chińskim i japońskim).

5. *Wielkość produkcji według języków. Druga ocena*

Jedno ze sprawozdań UNESCO zawiera wyciąg z szeregu opracowań statystycznych, dotyczących publikacji w różnych językach, w ujęciu procentowym. Ze sprawozdania wynika, że wszystkie dane szacunkowe wykazują znaczne rozbieżności i prawdopodobnie zawierają błędy. Sprawozdanie zawiera następujące zestawienie zbiorcze, które nie przedstawia dokładnych danych statystycznych, lecz ma jedynie ukazać rząd wielkości.

Język	Procent całości
Angielski	60
Rosyjski	11
Niemiecki	11
Francuski	9
Japoński	3
Hiszpański	2
Pozostałe	4

6. *Służba informacji naukowej i technicznej*

National Federation of Science Abstracting and Indexing Services (NFSAIS) stwierdziła, że ponad 612 tys. abstraktów i zestawień tytułów w 1960 r. przygotowały odpowiednie organizacje członkowskie i ponad 753 tys. miały przygotować w 1961 r.

G. Miles Conrad, dyrektor „Biological Abstracts", ocenił, że służba abstraktowa i indeksowa Stanów Zjednoczonych dokumentowała w 1958 r. w rzeczywistości 75% całości literatury.

Prof. A. Michajłow, Dyrektor Instytutu Informacji Naukowej Akademii Nauk Związku Radzieckiego, ocenia, że na świecie wychodzi ponad 300 różnych czasopism zawierających opracowania dokmentacyjne, przy czym każde z tych czasopism specjalizuje się w określonej dziedzinie. Prof. Michajłow ponadto wskazuje, że ponad 90 takich czasopism publikuje się w Stanach Zjednoczonych, ponad 40 — w Anglii i 27 — w NRF.

7. *Radziecka służba informacyjna* *)

Pochodzący z 1960 r. raport Radzieckiego Wszechzwiązkowego Instytutu Informacji Naukowej i Technicznej (Moskwa) stwierdzał, że 13 publikowanych wówczas przez ten Instytut czasopism informacyjnych zawierało opracowania dokumentacyjne wybrane z ponad 12 tys. czasopism (plus 3 tys. specjalnych publikacji radzieckich), pochodzących z przeszło 95 krajów i ukazujących się w 65 różnych językach. Instytut wydaje również 48 czasopism informacyjnych „Expres informacja" (każde obejmujące inny zasadniczy przedmiot), ażeby zapewnić szybkie streszczenia z podawanych informacji. Opublikowany w „New York Times" artykuł przypisuje omawianemu instytutowi opracowanie 16 tys. publikacji naukowych z 88 krajów*).

8. *Specjalistyczne Ośrodki Informacji*

Dr Adkinson z National Science Foundation ocenia, że w Stanach Zjednoczonych istnieje ok. 3 tys. wyspecjalizowanych ośrodków informacji naukowej, z czego ponad 50 jest na usługach Ministerstwa Obrony Stanów Zjednoczonych lub przez nie finansowanych. Badania prowadzone ostatnio przez Battelle Memorial Institute wykazały, że w Stanach Zjednoczonych istnieje ok. 450 wyspecjalizowanych służb informacyjnych w dziedzinie nauk fizycznych i biologicznych. Bardziej szczegółowe badania wskazują, że mniej więcej 200 spośród tych służb częściowo zatrudnia naukowców i inżynierów, realizuje program zdobywania wybranych danych i ich opracowywania oraz istnieje przede wszystkim po to, żeby dostarczać opracowań analitycznych i interpretacyjnych.

B. Dane o literaturze z poszczególnych dziedzin

9. *Medycyna*

NFSAIS stwierdziła, że w Index Medicus" w 1960 r. zebrano i wymieniono 125 tys. artykułów, a 140 tys. artykułów miało być ujętych w 1961 r.

National Library of Medicine oceniła w 1958 r., że w tym czasie istniało na świecie 13 tys. czasopism medycznych i rocznie ukazywało się 220 tys. artykułów kwalifikujących się do indeksowania, które można byłoby uznać za literaturę medyczną. Jednakże w tej liczbie 13 tys. prawdopodobnie mieściły się czasopisma, które zostały uwzględnione w grupie nauk przyrodniczych. To samo źródło cytowało podobne opracowania, które wskazywały mniejszą liczbę czasopism medycznych. Wydaje się, że liczba 9 tys. czasopism, obejmujących wszystkie dziedziny medycyny, stanowi właściwy kompromis.

W 1958 r. National Library of Medicine dostarczyła następujących danych szacunkowych dotyczących literatury medycznej z podziałem na grupy językowe.

Język	Procent całości
Angielski	37,3
Niemiecki	13,0
Francuski	12,9
Japoński	7,8
Włoski	7,5
Hiszpański	6,2
Wielojęzyczne	3,3
Rosyjski	3,1
Inne	8,9

W tym samym opracowaniu National Library of Medicine podaje, że otrzymuje czasopisma medyczne z przynajmniej 85 różnych krajów, co przed-

*) Dokładne dane dotyczące wydawnictw informacyjnych WINITI opublikowano w APID, 1962, nr 6.

stawia się w ujęciu procentowym z podziałem na kraje następująco:

Kraj pochodzenia	Procent całości
Stany Zjednoczone	24,5
Niemcy	11,7
Francja	10,2
Japonia	8,9
Włochy	7,8
Anglia	7,4
ZSRR	3,2
Pozostałe	26,3

10. Chemia

W 1959 r. „Chemical Abstracts Service" dostarczył wymienione poniżej dane procentowe dotyczące artykułów na temat chemii, publikowanych w czasopismach w następujących językach.

Język	Procent całości
Angielski	50,5
Rosyjski	16,8
Niemiecki	9,7
Japoński	6,1
Francuski	5,5
Włoski	3,6
Inne	7,8

W1961 r. Dale B. Baker, dyrektor „chemical Abstracts Service", stwierdził, że artykuły, z których „Chemical Abstracts" sporządziły w 1960 r. abstrakty pochodziły z 97 krajów i były opublikowane w 52 językach. W tym samym opracowaniu dr Baker podaje następujące dane procentowe, dotyczące literatury z zakresu chemii i dziedzin pokrewnych według krajów pochodzenia.

Kraj pochodzenia	Procent całości
Stany Zjednoczone	27,1
ZSRR	19,1
Wspólnota Brytyjska	13,8
Japonia	7,8
Niemcy	7,8
Francja	5,0
Włochy	3,2
Czechosłowacja	2,0
Polska	1,5
Szwajcaria	1,3
Węgry	1,2
Szwecja	1,1
Holandia	0,9
Pozostałe	8,2

Dr E. J. Crane z „Chemical Abstracts Service" stwierdził, że oryginalne informacje z dziedziny chemii ukazują się w 8 tys. publikacji naukowych i technicznych

„Chemical Abstracts", w 1957 r. obejmowały ok. 98% światowej literatury z dziedziny chemii z 7,5 tys. czasopism, 2 tys. książek oraz patentów z 22 krajów.

11. Nauki biologiczne

W końcu roku 1960 źródłowe wydawnictwo *) wymieniło w przybliżeniu 5 tys. czasopism z 84 krajów, zawierających artykuły z zakresu biologii, będące przedmiotem abstraktów w „Biological Abstracts". To samo źródło dostarcza następujących informacji na temat litreatury naukowej z zakresu biologii w różnych krajach.

Rejon geograficzny	Procent całości
Europa i Bliski Wschód	47
Ameryka Północna	29
Azja i Australia	13
Ameryka Środkowa i Płd.	8,5
Afryka	2,5

Wymienione źródło podaje również procentowy udział według krajów. Głównymi dostawcami tej literatury są:

Kraj pochodzenia	Procent całości
USA	26,75
ZSRR	11,40
Japonia	5,85
W. Brytania	4,22
Niemcy	4,10
Francja	5,83
Włochy	3,15
Brazylia	2,64
India	2,60
Kanada	2,13

Powyższe dane są oparte na analizie listy periodyków z zakresu nauk biologicznych, z których „Biological Abstracts" czerpią materiały do abstraktów.

„Biological Abstracts" dostarczyły następujących informacji na temat produkcji literatury naukowej z zakresu biologii według języków.

Język	Procent całości
Angielski	39
Niemiecki	13
Francuski	13
Hiszpański	6
Włoski	3
Języki słowiańskie	8
Pozostałe	18

12. Fizyka

American Institute of Physics stwierdził, że w 1959 r. „Physics Abstracts" zawierały w przybliżeniu 14 tys. abstraktów, a radziecki „Riefieratiwnyj żurnał: Fizyka" podawał ich ok. 28 tys.

W 1958 r. dr Crane z Chemical Abstracts Service ocenił, że „Chemical Abstracts" obejmowały przynajmniej ¾ artykułów ukazujących się w dziedzinie fizyki.

W 1961 r. wychodziło na świecie 1000 czasopism i 40 tys. artykułów z tej dziedziny nauki.

13. Mechanika

W 1957 r. „Applied Mechanics Review" ocenił, że rocznie ukazuje się 6 tys. artykułów w 700 czasopismach z dziedziny mechaniki.

W 1961 r. NESAIS stwierdziła, że „Applied Mechanices Review" opublikował w 1960 r. 6700 ab-

*) „Biological Abstoracts List of Serials" 1960. Przedruk z „Biol. Abstr." 35, nr 20.

straktów i miał wydać drukiem 7 tys. abstraktów w 1961 r. Przypuszczalnie w tej dziedzinie wiedzy ukazuje się rocznie 10 tys. artykułów.

14. Psychologia

W 1957 r. „Psychological Abstracts" oceniły, że opracowanych zostało 9 tys. abstraktów z 12 tys. artykułów, pochodzących z 500—600 czasopism.

NFSAIS stwierdziła, że w 1960 r. „Psychological Abstracts" wydały drukiem 8,5 tys. abstraktów i miały ich wydać 7 tys. w roku 1961. Należy przypuszczać, że rocznie ukazuje się z tej dziedziny 15 tys. artykułów.

Przeprowadzone w 1952 r. studium dotyczące ukazującej się na świecie literatury na tematy psychologiczne — według języków — dostarczyło następujących danych:

Język	Procent całości
Angielski	77
Francuski	8
Niemiecki	6
Pozostałe	9

15. Fizyka jądrowa

„Nuclear Science Abstracts" oceniły w 1957 r., że opracowano i opublikowano 14 tys. abstraktów z 14,5 tys. artykułów. Jednakże przy innej okazji padła liczba 60% opracowanych abstraktów w stosunku do liczby opublikowanych artykułów.

Według danych NESAIS — „Nuclear Science Abstracts" wydały drukiem 26,5 tys. abstraktów w 1960 r. i 32 tys. w 1961 r. Przedmiotem opracowań publikowanych w NSA są — poza oryginalnymi artykułami — patenty, sprawozdania, książki i tłumaczenia. Należy przypuszczać, że w zakresie fizyki jądrowej ukazuje się rocznie 2 tys. czasopism zawierających 35 tys. artykułów.

16. Rolnictwo

„Bibliography of Agriculture" oceniła w 1957 r., że w dziedzinie rolnictwa publikuje się rocznie 150 tys. artykułów.

NESAIS stwierdziła, że w 1960 r. „Bibliography of Agriculture" zawierała prawie 87 tys. tytułów i spodziewano się takiej samej liczby tytułów w 1961 r.

Ponieważ liczba publikowanych artykułów z dziedziny rolnictwa (150 tys.) jest taka sama jak w dziedzinie chemii, można w przybliżeniu przyjąć dla rolnictwa taką samą liczbę czasopism (8 tys.) jak dla chemii.

17. Elektronika i elektrotechnika

Dotychczas nie zgromadzono wystarczających informacji, ażeby pokazać rozmiary literatury z zakresu elektroniki i elektrotechniki. Według wstępnych danych ocenia się liczbę artykułów na 30 do 70 tys. rocznie, a liczbę czasopism na tysiąc.

18. Aeronautyka

„Aeronautical Reviews" oceniły w 1957 r., że w dziedzinie aeronautyki ukazuje się rocznie 13 tys. artykułów.

Według danych NESAIS — „Aero-Space Reviews" (obecnie „International Aero-Space Abstracts") wydały drukiem w 1960 r. 9,5 tys. abstraktów; przewidywano, że w 1961 r. wydadzą 10 tys. abstraktów.

Jedno z wydawnictw Special Libraries Association z 1960 r. wymienia 676 czasopism z tej dziedziny. Zakładając przeciętnie 30 artykułów rocznie dla każdego czasopisma, można przyjąć z tej dziedziny liczbę 20 tys. artykułów rocznie.

Charles M. Gottschalk z Biblioteki Kongresu stwierdził niedawno w prywatnej informacji, że będąca w opracowaniu bibliografia czasopism z dziedziny nauk lotniczych, która będzie wydana przez Library of Congress, wymienia w przybliżeniu 1,5 tys. czasopism wyłącznie z zakresu aeronautyki. Zakładając rocznie 30 artykułów dla czasopisma, można przyjąć 45 tys. artykułów rocznie z tej dziedziny.

19. Matematyka

W 1957 r. „Mathematical Reviews" oceniły, że w dziedzinie matematyki ukazywało się 10 tys. artykułów rocznie. Pismo to stwierdziło ponadto w 1957 r., że abstraktami objętych zostało 1036 czasopism i 300 książek.

NESOIS podaje, że „Mathematical Reviews" wydrukowały w 1960 r 7,8 tys. abstraktów i miały wydać 10 tys. abstraktów w 1961 r. Należy przypuszczać, że w tej dziedzinie ukazuje się rocznie 15 tys. artykułów.

20. Meteorolgia

W 1957 r. „Meteorological Abstracts" oceniły, że opracowały i opublikowały 5 tys. abstraktów z ok. 7,2 tys. artykułów z 2,5 — 3 tys. czasopism. Podobnie jak w dziedzinie matematyki ta ostatnia liczba zawiera dużą ilość czasopism, których udział w ogólnej liczbie dokumentowanych artykułów jest bardzo nikły.

NESAIS stwierdziła, że „Meteorological Abstracts and Bibliography" (obecnie: „Meteorological and Astrogeophysical Abstracts") wydrukowały w 1960 r. 12,8 tys. obstraktów i zamierzały wydać 15,9 abstraktów w 1961 r.

Charles Gottschalk z Biblioteki Kongresu stwierdził niedawno, że liczba publikacji, z których „Meteorological Abstracts" opracowały w 1959 r. abstrakty, obejmowała 980 pozycji, wśród których znajdowało się wiele czasopism naukowych. W tej dziedzinie ukazuje się orientacyjnie tysiąc czasopism rocznie, zawierających 20 tys. artykułów.

21. Metalurgia

NESAIS stwierdziła, że „Review of Metal Literature" wydał drukiem w 1960 r. 12 tys. abstraktów i w 1961 r. zamierzał wydać 12 tys. Brytyjski Instytut Żelaza i Stali założył „Abstract and Books Title

Index Card Service" i spodziewa się rejestrować rocznie około 9 tys. abstraktów. Brytyjski „Metallurgical Abstracts" aktualnie publikuje rocznie ok. 7,5 tys. tytułów z tysiąca czasopism.

American Society for Metals prenumeruje obecnie ok. 900 czasopism dla swojej służby badawczej i indeksuje ok. 36 tys. dokumentów rocznie, włączając w to książki i sprawozdania. Należy przypuszczać, że w tej dziedzinie wychodzi od 600 do 900 czasopism i ukazuje się od 30 do 35 tys. artykułów rocznie.

22. Inżynieria lądowa

Nie zebrano jeszcze dostatecznych informacji, dotyczących wielkości literatury na temat inżynierii lądowej. Należy przypuszczać, że w tej dziedzinie ukazuje się 500 czasopism zawierających 15 tys. artykułów rocznie.

23. Inżynieria przemysłowa

Brak informacji na temat wielkości literatury w tej dziedzinie. Należy przypuszczać, że w zakresie inżynierii przemysłowej ukazuje się 500 czasopism publikujących rocznie 15 tys. artykułów.

24. Inne specjalności

Na temat wielkości literatury w innych dziedzinach brak informacji. Jednakże biorąc pod uwagę zbieżności i dublowanie się artykułów między wyodrębnionymi dziedzinami, można założyć, że istnieje milion artykułów z dziedzin nie wyszczególnionych, aby dojść do ogólnej liczby dwóch milionów artykułów.

ADAM GÓRSKI

025.4:608.3

KLASYFIKACJA PATENTÓW I DOŚWIADCZENIA OŚRODKA INFORMACJI PRZEMYSŁOWEGO INSTYTUTU MASZYN ROLNICZYCH

Dokumentacja patentowa jest dotychczas mało docenianym w Polsce źródłem informacji o najnowszych wynalazkach i postępie techniki na świecie. Brak dotąd również odpowiedniej sieci informacji patentowej. Ośrodek informacji Przemysłowego Instytutu Maszyn Rolniczych zorganizował — na wzór zagranicy — branżowy zbiór patentów (krajowych i zagranicznych). Prowadzi w szerokim zakresie działalność informacyjną. Zapoczątkował wydawanie własnych wydawnictw patentowych oraz przystąpił do prac w dziedzinie ochrony praw własności przemysłowej. Wykonuje m.in. analityczno-syntetyczne opracowania — w formie ekspertyz — na tematy ochrony patentowej poszczególnych wyrobów w kraju i za granicą.

Bardzo szybki postęp techniczny na świecie, coraz większa rywalizacja w opanowywaniu zagranicznych rynków zbytu i pojawiające się sygnały o konfliktach na tle naruszenia patentów spowodowały konieczność zainteresowania się przez placówki naukowo-badawcze i zakłady produkcyjne w znacznie szerszym niż dotąd zakresie zagadnieniami ochrony praw własności przemysłowej.

Ogólnie stwierdzić należy, że znajomość tych zagadnień w naszym społeczeństwie jest niedostateczna. Niedocenianie problemów wynalazczości i jej wpływu na postęp techniczny przejawia się również w zbyt małym wykorzystaniu dokumentacji patentowej, a szczególnie opisów patentowych, które jako publikacje treści zgłoszonych wynalazków są przecież pierwszymi informacjami na ten temat.

Z drugiej strony przyznać trzeba, że zagadnienie ochrony prawa własności przemysłowej w dalszym ciągu natrafia na szereg trudności organizacyjnych, a bogate zbiory patentów zagranicznych — z uwagi na brak centralnej informacji w tym względzie — są wykorzystywane w stopniu minimalnym. Trudności występują głównie z powodu braku odpowiednich komórek organizacyjnych zajmujących się zagadnieniami ochrony praw własności przemysłowej. Nie rozwiązano również zastępstwa prawnego w tej dziedzinie. W dalszym ciągu nie mamy etatowych rzeczników patentowych — specjalistów odpowiednich branż przy odpowiednich zakładach, którzy z tej pozycji powinni czuwać nad ochroną praw własności przemysłowej, jak to jest zresztą za granicą. Nawet u naszych najbliższych sąsiadów (NRD, CSRS) istnieją specjalne biura do spraw wynalazczości i patentów, w których pracują tzw. inżynierowie patentowi, mający te same uprawnienia co rzecznicy.

Czy tych inżynierów patentowych mogliby w naszych warunkach zastąpić rzecznicy patentowi, prowadzący własne kancelarie i zajmujący się wszystkimi dziedzinami techniki? Raczej nie. Rzecznicy nie będący specjalistami branżowymi ograniczać się muszą głównie do pośrednictwa w korespondencji i do udzielania pomocy tylko formalnej, a to przecież naszym zakładom nie wystarcza.

Nie rozwiązanych spraw jest znacznie więcej. Z tych też względów problem ten nie jest łatwy ani prosty. Wiąże się on z wieloma czynnikami, a wśród nich również z właściwym ustawieniem dokumentacji i informacji o patentach.

Jednym z pierwszych przemysłów, które postanowiły zagadnienie ochrony praw własności przemy-

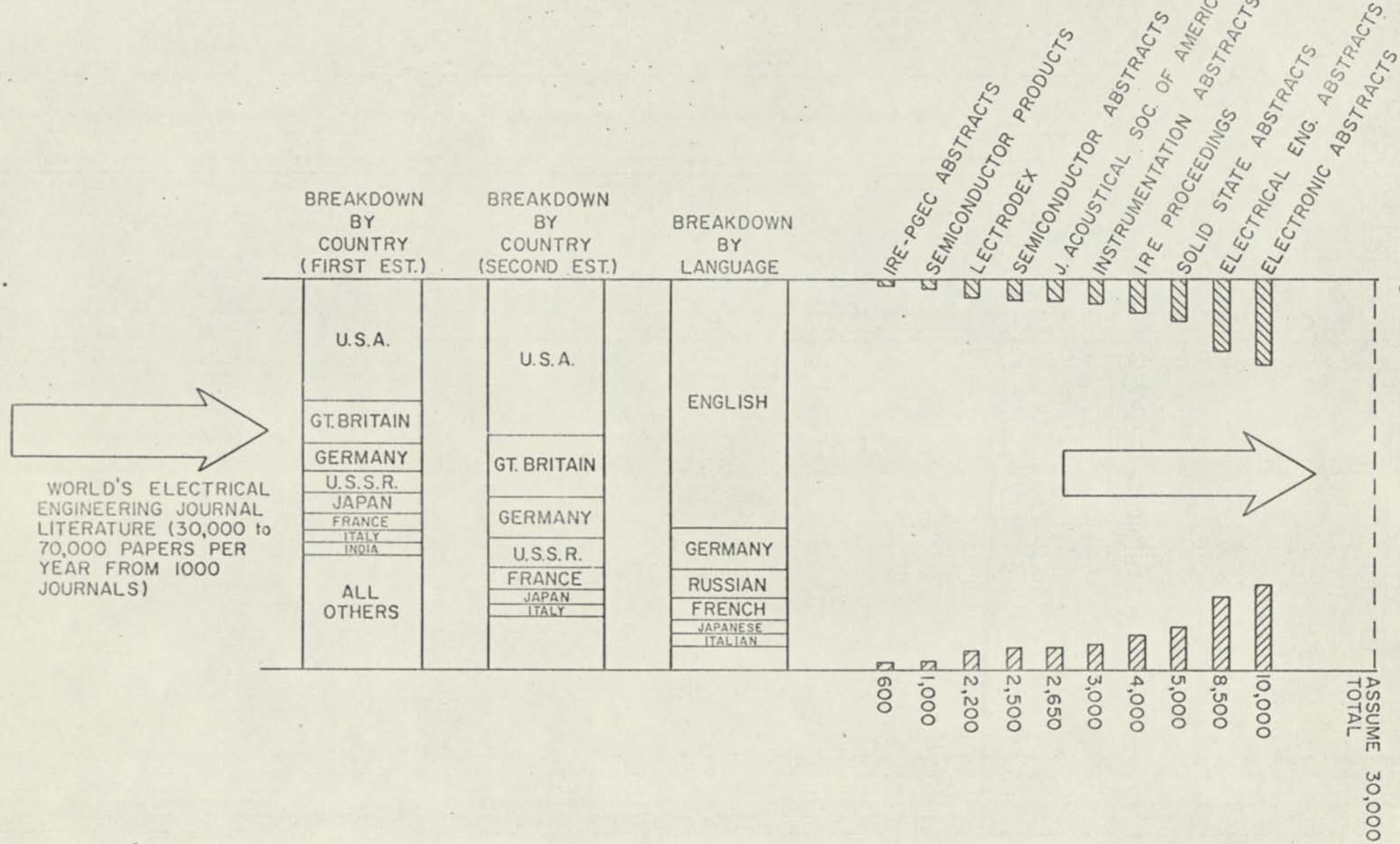


FIG. 3

CURRENT COMPOSITION OF THE WORLD'S ELECTRICAL ENGINEERING JOURNAL LITERATURE

SRI REPORTS PROJECT CARD

101 ACC/CLASS	0023287	U	DO NOT ANNOUNCE
201 CORP AUTH			
301 TITLE	REVIEW OF THE METHODOLOGY OF INFORMATION SYSTEM DESIGN.		
302	*PROJ 4229		
303			
304			
305			
306			
307			
308			
309			
310			
401 AUTHORS			
402	BOURNE, CP		
403			
404			
501 PROJ/DATE	PROJ ESD-4229	SEP 62	
502 PROJ			
601 CLIENT			
602			
603			
604			
605			
606			
701 ADDL NOS			
702			
703			
704			

PROJECT 4229

DATE SEP 62

COPY NO(S) 1 COPY

CONTRACT

O/P

SRI 2369