INTERNATIONAL FEDERATION FOR DOCUMENTATION (FID)

WASHINGTON, D. C.
U. S. A.

10 - 15 OCTOBER 1965

Abstracts
1965 Congress
International Federation for Documentation (FID)

10 - 15 OCTOBER 1965

WASHINGTON, D. C.

Abstracts

SECRETARIAT, 1965 FID CONGRESS

9650 WISCONSIN AVENUE, WASHINGTON, D C 20014, U.S.A.
Best Available Copy
CONTENTS

SYMPOSIA

AREA I  Education and Training of Documentalists
  Symposium A  Present State of Education and Training  1
  Symposium B  Needs for the Next Ten Years and How to Meet Them  13

AREA II  Organization of Information for Documentation
  Symposium A  Transformation and Organization of Information Content  19
  Symposium B  Comparison and Evaluation of Transformation Techniques and Organizational Structures in Documentation Systems  26

AREA III  Information Needs of Science and Technology
  Symposium A  Methodology  33
  Symposium B  Increasing the Efficiency of Information  37

AREA IV  Information Needs of Society
  Symposium A  43
  Symposium B  Specific Knowledge Areas  46

AREA V  Principles of Documentation and Systems Design
  Symposium A  49
  Symposium B  53

CONTRIBUTED ABSTRACTS  57

INDEX  95

This booklet contains abstracts of both symposium papers and contributed papers, followed by an index of authors. Symposium abstracts are grouped sequentially under the five Area topics listed above, and appear in the order in which the papers will be presented at the Congress. Contributed abstracts are ordered alphabetically by the name of the first author. Numerals following author names in the index refer to the page number of the abstract.
AREA I SYMPOSIA

AREA I. EDUCATION AND TRAINING FOR DOCUMENTALISTS

Symposium A: Present State of Education and Training

THE TRENDS OF DOCUMENTALIST TRAINING IN EAST EUROPEAN COUNTRIES. Zygmunt Majewski. Committee for Science and Technology, Warsaw, Krakowskie Przedmiescie 1, Poland.

I understand the East European countries to be the socialist countries that mainly occupy this part of Europe.

The economic system which is specific for these countries is projected in a considerable measure onto other aspects of national life. It is projected onto scientific informational activities also in spite of the fact that it utilizes in its work methods and measures applied in other countries of the world.

The national economy in socialist countries is nationalized and the industry, communication equipment, and other economic branches are the property of the entire community and are centrally directed by appropriate economic organizations. The general outline of all the functions, production etc. are planned as a whole for several years in advance and the plans are expressed in respective state acts. Certain wide differences emerge among particular countries as regards agriculture which is sometimes wholly nationalized or just partly/as in Poland, for instance/. Scientific institutions also have their research plans. Such plans are coordinated by one state body irrespective of whom the given institution is subordinate to. Academy of Sciences - Higher Schools of Learning - Economic Departments/.

The influence of such a scientific and economic system on the work of the scientific information service is so reflected that there are no secrets between work establishments and scientific institutions and information can flow freely. Having centralized managements of entire branches of the national economy, central occupational information services can easily be formed. The same concerns the branches of the basic sciences. Having plans for the development of learning and economy for the future years, information centers can prepare in good time the necessary scientific information and be in a certain measure the coauthors of the plans mentioned.

How does documentalist training look and to what extent does the state of affairs delineated above reflect on documentalist training?

Please note that the question here is the training of documentalists exclusively those who are university graduates and not those who are merely graduates of a documentalist technical school /graduates of secondary schools and a professional documentation school as, for instance in Poland or assistant personnel-laboratory workers, for instance/.

Training university graduates was conducted until now at regular attendance courses not exceeding 100 lecture hours and exercises or at extramural courses of longer duration. Outside of a "classical" curriculum for documentation and scientific information these curricula took into greater consideration subjects facilitating cooperation between scientific information services and scientific or
economic institutions. Such subjects cover the organization of science and economy, cooperation between economy and science, planning methods etc. In as much as computers are used with increasing frequency both in planning and scientific information, the principles of their operation are one of the important subjects. Analytical and synthetic reports are a means of information used with increasing frequency; writing up such reports is also a subject of lectures. The same concerns using direct information /not by way of the document/ such as organizing scientific conferences, radio broadcasts and television.

Almost all the East European countries have prepared postgraduate curricula generally for one year of about 1000 hours of lectures and exercises. Such studies are arranged by colleges and universities which will be able to confer doctor's degrees.

Of the socialist countries only the USSR confers doctor's degrees in scientific information.

Training of users of documentation and scientific information is conducted in all the socialist countries both in colleges and universities /15 to 36 hours/ and at specialistic courses, as for instance training courses for managers.
Various problems involved in the education and training of persons required in different fields of documentation and information service are reviewed at length. Educational requirements, knowledge and training background in documentation necessitated by general, scientific, industrial documentation; documentation in arts, social sciences, literature and humanities; information management; economic and managerial information; translations and documentary reproduction have been exhaustively and systematically enumerated. A person with a basic degree would be more useful for documentation purposes than a specialist of a particular field. Library science qualifications or library training would be an additional advantage. Appropriate educational background with temperamental qualities would be more effective in the operation of the documentation services. To meet with the situation education in documentation at University level or by special institutions has been recommended. Training of documentalists by FID in its future program is also cross referenced. Training of the documentation personnel on Job basis, formal training, training in the home country or in foreign countries have been discussed in detail. Since training facilities in developing countries of Asia are very limited, foreign training is therefore highly desirable to improve the standards of documentation service. The role of Pakistan National Scientific & Technical Documentation Center (PANSDOC) with special reference to its training facilities offered to scientific and technical institutions and industries and providing courses on scientific and technical documentation to University students at post-graduate level have been briefly mentioned. It is suggested that a detailed study and analysis of the observations, experiences and results accrued from the existing systems of education and training in documentation would be conducive to improve and standardize the methods, systems, syllabi and practices regarding the education and training in documentation. Status of documentalists is evaluated from professional and social point of view; and it is pointed out that the documentalists have vital roles in the country's progress and hence their status should receive due recognition. In Pakistan, position and status of documentalists are at par with the scientific personnel.
AREA I. SYMPOSIA

THE PRESENT STATE OF EDUCATION AND TRAINING IN DOCUMENTATION, INFORMATION SCIENCE, AND SPECIAL LIBRARIANSHIP. Jesse H. Sher. School of Library Science, Western Reserve University, Cleveland, Ohio, U.S.A.

The origins of contemporary thought concerning the professional education of documentalists and other specialists in information science are probably to be found in speculation about the need for an adequate educational program for the professional preparation of special librarians. As early as 1911, only two years after the Special Libraries Association was established, A. G. S. Josephson of the John Crerar Library pointed out in a letter to the Special Libraries Association the need for specialization in library education. Seven years later, in 1918, J. G. Pearse argued in the pages of the British Library Association Record for training in documentation as a special branch of library education in general.

Thereafter, most of the argument concerning specialization in librarianship came from the special librarians, and little attention was paid to the field known today as information science. In 1925, Linda Morley offered at Columbia University the first course in special librarianship to be sponsored by any library school, and similar courses were slowly introduced into other library schools.

On this side of the Atlantic documentation was, during the 1930's, largely interpreted in terms of microphotographic and other photographic processes, and the concept was not extended to include non-conventional methods of information storage and retrieval until after the end of the Second World War. In 1950 Helen Focke of the School of Library Science at Western Reserve University was the first to offer a course in documentation, and five years later the program was greatly expanded with the establishment of the Center for Documentation and Communication Research as a division of the library school at Western Reserve.

The argument whether a special librarian should be trained primarily as a subject specialist or as a librarian was quickly transferred to the discussion of the professional education of the documentalists and information specialists. A study by Cohen and Craven (financed by the National Science Foundation) found from their sampling of opinions of practicing information specialists that such conventional subjects as information sources, cataloging and classification, and subject bibliography were given a high priority, but that computer technology, information science, and mechanized information techniques ranked low on the list. A number of conferences have failed really to clarify the problem of professional education, much less develop any authoritative conclusions or evolve a consensus.

At the present time the professional education of documentalists and information specialists in the United States falls into two major groups, those attached to conventional library schools, as at Western Reserve, the University of Minnesota, and the Drexel Institute of Technology, and those which are not based upon librarianship at all as at the Georgia Institute of Technology, Lehigh University, and, most recently, at the University of North Carolina. In a few instances special degrees in information science have been established, but as yet none has been recognized by any authoritative accrediting agency.

There are many unresolved problems which must be solved before a definitive program, or programs, can be established. The field itself must achieve a higher degree of standardization, and its terminology must be stabilized. Effective studies of the intellectual needs and skills of successful practitioners of
documentation are desperately needed. All of the programs are, at present, too young to have produced graduates whose competence has been tested over a sufficiently long period of time. Programs have been predicated on assumptions of what is and should be rather than on any precise knowledge of what the information specialist does and needs to know in order to do it. The role of automation and computer technology in this newly emerging field of documentation is only vaguely understood, if indeed it is understood at all. Much cautious and well-founded exploratory work and experimentation must be done if the training of a substantial, perhaps even disastrous, number of misfits is to be avoided.
One of the most remarkable consequences of progress, both technical and scientific is an immediate production of "documents" in all their forms - intellectual and physical - and the problem of preparation, classification, retrieval and locating of the information contained in them. This phenomenon, in a greater or lesser degree, directly or indirectly, happens in every country in the world and has aroused a demand for specialists in the use of documentary techniques. As is well known the cultural, economic, social, scientific and technical development of the Latin American countries included approximately between latitude north 30° and latitude south 50° is unequal. Undoubtedly these aspects should be studied in their relation to the continent as a whole but, in every case taking into account the individual conditions and necessities of each country.

To introduce the teaching of documentation it is essential to know beforehand the most urgent requirement that would contribute to the industrial development or the solution of the economic, social, educational or health problems (1). We know that courses in documentation are held in Argentina (Curso regional de orientación a la documentación científica y técnica-UNESCO-C.N.I.C. y T.), Brazil (I.B.B.D.) and Chile (CENID) and according to a recent report from the Medellin Seminar on Library Science studies, we know that out of 33 schools, 12 include this subject. The programs of study are not sufficient to give us a clear conception of what the classes comprise. As Gaston Julia used to say: "What is important are the ideas more than the subjects which are taught, the manner of working more than the deeds". The great majority of librarians in Latin America - as in other more developed countries - have not had the opportunity of acquiring a documentary conscience, a fact which would have no importance if they would limit themselves to the typical librarian function of an "educational" nature (2); but due to the lack of human resources they must face the organizing of documentation centers, special and specialised services of information, etc. It is not worthwhile to repeat here the sterile discussions which occur between librarians and documentalists, let us say however, that there is a new profession which is no more than ten years old and which we shall call "A", which must tend to such a variety of requirements; as some of the ones which are mentioned here: preparation of the documentation and automatic processing of information in public administration (3); in most countries the same or almost the same limitations and deficiencies can be found which affect the adequate functioning of the administrative machinery as an instrument of public service and general progress) (4); statistics, census and sociological surveys, scientific and technical documentation of: firms, banks and insurance companies, market surveys, registration of persons, register of patents, legislative and fiscal, information for the armed forces, newspaper and publishing enterprises (as much for the information they gather as for the means of dissemination) (5); if we put aside the principle that documentation is part of library science, inasmuch as, besides having different methods and techniques, it is a mental attitude, it will be easier to determine the characteristics of the professional training. In Latin America undoubtedly there exist common problems, but it should be established which are the most urgent necessities of each country and specify what appointment will be given to the trained people (public administration, type of industry, research center).

We should insist that the documentalists as well as the librarians should undertake activities eminently practical and not speculative. As a result of a recent trip through various Latin American countries (Chile, Peru, Colombia, Venezuela,
Mexico and Uruguay) we gathered the impression that in those countries there exists interest for courses in documentation. At this time it does not seem opportune to found schools for long term studies; it would be sufficient to establish the following courses, choosing those most apt according to the country and the circumstances: A) Lectures or short elementary courses of information sponsored by scientific institutions, school for librarians, professional associations, documentation centers, universities...; B) Higher level courses of from 4 to 6 months for young researchers, engineers, technicians, post-graduates, government officials in technical posts, graduates from higher level schools for librarians. The course could be divided into 1) general documentation and 2) specialized documentation, which in turn could be subdivided into: a) social, economic and humanistic sciences and b) pure and applied science, technology. These courses would be sponsored by Councils or National Centers of research or documentation or higher institutes that play the role of receiver of information problems. It would be advisable to be able to count for these classes with the participation of professors of international repute in order that the very latest advances in the field of documentation be included as well as experimental work and prospects for the future. C) Complementary courses and/or documentary subjects included in the program of study in schools for librarians.

The basic formation we have outlined could be used for distinct categories of personnel which the services require, some of which have been mentioned by Koblitz (6). The furthering of the courses in each country could be accomplished through the Secretary of FID/CLA/.

   — v.5. Situación cultural.
(2) Shera, Jesse H. y Egan, Margaret E. "Examen del estado actual de la bibliote-
    conía y de la documentación". En Revista Universidad. Universidad nacional
    del Litoral (Santa Fe), nº 61, 1964, p. 279-321.
(3) Sadosky, Manuel y Gianetto, Guillermima G. "La documentación científica y la
    automatización". En Seminario latinoamericano sobre documentación científi-
    ca, Lima, 3-8 setiembre 1962 (Montevideo). Centro de cooperación científica
(4) "La administración pública en América Latina". Resumen de la publicación OEA/
    ser. H/14-CIES/348/español, 27 set. 63. En Revista de administración públi-
    ca, (Buenos Aires), nº 12, mar. 1964, p. 80-113.
(5) Taube, Mortimer and others, Studies in coordinate indexing /s.l./ Documentation
(6) Koblitz, Josef. "Factores humanos", en Otto Frank y otros, Técnicas modernas
A SWEDISH COURSE IN SCIENTIFIC INFORMATION TRANSFER AND RETRIEVAL. BJÖRN V. TELL. Royal Institute of Technology, Stockholm 70, Sweden.

During the Fall and Spring semesters 1964/65 a course called "Transmission of Scientific Information" has been given in Sweden. It was offered to university students and to graduates with experience in documentation, and was held at the Tekniska Högskolan (Royal Institute of Technology), and was sponsored by the Swedish National Committee on Documentation. Lectures and demonstrations comprised over 100 hours. The purpose was to give a general introduction to documentation focused, however, around information storage and retrieval on computers.


Of the sixty participants from industries, universities, and libraries, 46 fulfilled the course with high attendance at the lectures. Of the students over 40 were graduated from various universities. However, the difference between the graduates and undergraduates didn't seem to be so cumbersome as did the difference between those with training in science and those without. Therefore, the conceived short introduction into mathematical and logical basis for information transmission had to be expanded. As example of systems in use the Hoffmann-La Roche and Derwent Ringdoc at Karolinska Institute were demonstrated on IBM 1401. The construction of a union catalogue was shown on a Ferranti Mercury by use of paper-tape input.

An evaluation of the course was made by means of a questionnaire to the students at the end of the course. It was shown that about half of the students wanted a better background in mathematics and statistics, that more emphasis should have been given to exercises, that the text-book (C P Bourne, Methods of Information Handling) supplemented by compendia by the teachers was a good choice. For the time being it was obvious that this course had to depend on a team of specialists as teachers. Some overlapping and differences in terminology were, thus, unavoidable.

Plans for the immediate future is to start a similar course in the Fall at Uppsala University. This course will fit into the regular university programme. In order to avoid the introductory training in mathematics, the students missing such training will be requested to attend a special course offered by the statistics department. The course will include about the same amount of lectures, and 60 hours of exercises and special work will be added. The students have to choose between: A) Documentation within the library network, B) Scientific-technical
documentation, C) Managerial and economic documentation, and D) Medical documentation.

For those students who want to proceed in their studies to an intermediate or higher level opportunities will also be open. From July 1965 the Swedish universities will implement a new curriculum in information handling. The earlier disciplines: Numerical analysis, and Data processing will merge under a new umbrella: Information handling. The second term curriculum includes information retrieval and the processing of verbal information. In this way the students will become machine oriented.

References:
THE EDUCATION AND TRAINING OF DOCUMENTALISTS IN CANADA. Samuel Rothstein.
University of British Columbia School of Librarianship, Vancouver, Canada.

The Canadian pattern of education for librarianship and documentation closely resembles that of the United States. The agencies responsible for such education are graduate library schools. These are departments of universities, analogous in organization and function to other professional schools, such as social work, within those universities. There are now five Canadian library schools, located at the University of British Columbia (Vancouver), the University of Toronto (Toronto), McGill University (Montreal), the University of Ottawa (Ottawa), and the University of Montreal (Montreal). The first three serve the English-language libraries and the University of Montreal the French-language group; the University of Ottawa Library School is a bilingual institution. All five library schools demand a bachelor's degree as prerequisite to admission.

The program of instruction in U.B.C., Toronto, Ottawa and Montreal is of one academic year's duration and leads to the Bachelor of Library Science degree (B.L.S.), which is roughly equivalent to the Master of Library Science (M.L.S.), the first professional degree commonly awarded in the United States. The library science program at McGill has recently been extended to two years and leads to the (sixth-year) Master of Library Science degree. An extension of the University of Montreal program to two years is now contemplated. The University of Toronto also offers a (sixth-year) Master of Library Science degree course and the University of British Columbia will offer a similar course in 1966.

Documentation is generally regarded in Canada as a branch of or as an advanced specialization within librarianship and the training of documentalists is therefore almost entirely a function of the library schools. Most of the library schools offer some orientation or beginning courses in documentation as part of the B.L.S. degree course, but they reserve more intensive and specialized study of this field for the M.L.S. program, where it is likely, indeed, to be a major feature of the course. In addition to such degree programs in documentation, informal workshops and short courses on various aspects of this subject are frequently presented by library schools, library associations and by individual libraries.

The next decade will certainly see the establishment of more library schools in Canada (two are, in fact, already announced as to begin in 1966--at the University of Alberta and at the University of Western Ontario) and large-scale changes in the curriculum. There is considerable pressure to give documentation a greater place in the library school program and it is likely that the present limited offering within the B.L.S. degree course will be expanded. It is also possible that Canadian library schools, in conjunction with computer centres and other interested university departments, may offer degrees in information science. Nevertheless, Canadians will probably retain the view that documentation does not appropriately constitute a field separate from librarianship and that the training of documentalists should therefore be based upon and follow a thorough preparation within librarianship.
Bibliography


2. Bulletins of the library schools at the University of British Columbia, University of Toronto, McGill University, University of Ottawa, University of Montreal


Elle est assurée depuis 1950 par l'Institut national des techniques de la documentation (INTD), Conservatoire national des Arts et Métiers. Recrutement direct pour licenciés et ingénieurs. Examen probatoire pour bacheliers. Deux années d'enseignement (72 + 61 heures). 71 diplômés en 1964 sur 125 entrées en 1962. Augmentation croissante des candidats. Insuffisance du nombre des candidats: a) masculins, b) du niveau de la licence, c) ayant une formation scientifique. Collection "Documentation et Information" (Gauthier-Villars) - Pour favoriser la formation technique supérieure dans les Facultés des lettres et des sciences humaines un certificat de technologie (Documentation) a été créé en 1962 à Toulouse et en 1963 à Nancy. Projets pour Lyon - Cours à l'Ecole pratique des Hautes-Etudes sur "Sémiologie et documentation" (J.C. Gardin et B. Jaulin), "Bibliographie et méthodes de documentation dans les sciences sociales" (J. Meyriat) - L'Union française des organismes de documentation (UFOD) assure la formation des aides-documentalistes (2 sessions par an) et organise des sessions (5 jours) d'initiation ou de recyclage - Des études en cours (Commission de la recherche scientifique et technique du Ve Plan), à propos des problèmes de documentation de la propriété industrielle, ont porté sur la formation d'un personnel chargé de concevoir et de mettre en œuvre les méthodes et machines nouvelles nécessaires aux applications automatiques dans le domaine de la documentation. Si les crédits sont accordés, comme on l'espère, il conviendra de déterminer l'établissement ou les établissements d'enseignements compétents, soit organismes nouveaux, soit transformation d'institutions existantes comme l'INTD ou certaines universités.
AREA I. EDUCATION AND TRAINING OF DOCUMENTALISTS

Symposium B: Needs for the Next Ten Years and How to Meet Them

THE PRESENT STATE OF EDUCATION AND TRAINING FOR DOCUMENTATION IN JAPAN AND SOME FUTURE PROBLEMS. Masao Kotani. Faculty of Science, The University of Tokyo, and Yosoji Ito. The University of Tokyo Library, No. 7 Hongo, Bunkyo-ku, Tokyo, Japan.

The status of the regular training courses of library science especially for research or special librarians in Japan is very poor. Although two undergraduate library schools are opened in Japan, the objectives of these schools are to train professional librarians for various types of libraries, so the curriculum is widely generalized and the training for research librarians and special librarians is insufficient. No graduate courses for librarianship and the regular courses for documentalists are founded in Japan. The Japan Library Association, the Japan Documentation Society, and the Ministry of Education frequently open each short term training courses for documentation. Unfortunately, these short term training courses are planned independently by each organization and are not in liaison with each other. It is gradually recognized among scientists of different disciplines, particularly among chemists, that the basic training of literature search, art of writing primary papers, etc., need be included in the curricula of university courses. In this connection the present situation in some universities of Japan will be reported.
INTEGRATED TRAINING POLICY FOR DOCUMENTALISTS ON NATIONAL LEVEL. NEW TRENDS IN THE FEDERAL REPUBLIC OF GERMANY. Dr. Martin Cremer. Institut für Dokumentationswesen, Frankfurt am Main, Germany.

The rapid development of documentation and information in the past ten years is based - apart from organizational and financial demands - on two conditions: effective methods have to be found and applied in order to evaluate the continuously increasing flood of knowledge and to comply with the growing demands of information. Secondly, a sufficient number of specialists have to be trained, who will be charged with these tasks. Both conditions cannot be complied without certain planning activities. With regard to the different educational systems these activities in the field of training may be realized on a national level.

Therefore, in the Federal Republic of Germany the enlargement and integration of training possibilities is prepared. The Deutsche Gesellschaft für Dokumentation has organized until now 8 courses for documentation - each course lasting about 10 months and comprising nearly 400 lessons. This has been a first and successful step, but it became evident that for present and future needs, these courses are not sufficient. Therefore, the creation of a training centre for Documentation in Frankfurt is planned. This centre will train mainly professional documentalists with or without academic degree. Besides this, special courses for medicine, chemistry, technology, social sciences, sciences and humanities will be arranged, and special fields (indexing, classification, reprography and application of mechanized techniques) will be taught. The centre will also arrange special courses for further education of documentalists giving them the opportunity to inform themselves about the most recent development.

All necessary elements of the science of information and documentation as well as important sections of library sciences with particular regard to the needs of special libraries should be included in the plans for instruction. Thus, a not always existing but desirable integration of documentation and special libraries by common training opportunities is intended.

Therefore, it is to be expected that the number of participants of the different courses will be high, especially the number of participants of the courses for professional documentalists.

The instruction will be done by a small number of full-time teachers and a big number of part-time experts coming from different fields of documentation and library science.

The training centre for documentation should be in close local and technical connexion with a documentation-library, including all instruction fields, and an information centre for literature on documentation, information and library science, with a division of reprography as well for instruction as for practical use, and with the Zentralstelle für maschinelle Dokumentation (ZMD), which will take care of the training on its particular field.

This integrated training system will be completed in a non-centralized way by education of users of documentation at universities, in industry and in administration. Here planning and promoting activities are necessary, too.
EDUCATION FOR SYSTEMS PLANNING. Don R. Swanson. Graduate Library School, University of Chicago, Chicago, Illinois.

The training of documentalists and librarians is usually considered in terms of skills required to operate libraries and document centers. It does not necessarily follow that those who acquire such skills will also be able to competently and imaginatively plan future libraries and document centers; it is education for this latter purpose that constitutes the topic of this paper.

The planning of almost anything involves, first, identification of goals or objectives, and secondly an understanding of the resources for accomplishing those objectives. Resources for accomplishing goals consist of people, machines, information, and techniques of indexing, classification, and other means of "intellectual access" to information. The task of systems planning is one of the optimal allocation of limited resources.

A graduate curriculum addressed primarily to the planning of information systems might reasonably be divided into four categories of courses: (1) The use and users of information, (2) systems planning and computer technology, (3) a study of existing library and information resources, and (4) a study of indexing, classification, reference and bibliography. Further discussion of these and related categories is given in reference 1.

Undergraduate preparation for such training is a matter to which a great deal more serious thought should be given by educators. There are certain areas of knowledge which I believe provide an especially good foundation for the planning of information systems, and, at least tentatively, would propose the following for consideration:

(1) The language of elementary mathematics and logic.

The student of systems planning should have good preparation in elementary mathematics at the level of first year college, and to include algebra, simple functions, graphs, and systems of equations and inequalities. Mathematics is here regarded not so much as a specialty but as extension of our powers of communication. Communication in terms of numbers, symbols, graphs, and simple statements of formal logic are prerequisite to the study of the technology of information handling.

(2) Linguistics.

Syntactic and semantic relationships of words in subject headings and titles constitute an important foundation in any study of indexing and classification. The science of linguistics has been largely detached from studies of indexing but should be regarded as an essential component of educational preparation in this area.

(3) Empirics and statistics.

Those who plan information systems of the future should be able to understand and interpret the results of research studies in
the information sciences. In many of these studies, data are collected and interpreted. Thus, an understanding of research procedures, methodology, empirical reasoning, and elementary statistics is necessary.

Undergraduate preparation should above all involve the student deeply in the pursuit of knowledge and scholarly inquiry, and in the solving of problems. To this end, the more substantive the curriculum the better, regardless of the area of specialization. In addition to mathematics, logic, and linguistics, those conventional disciplines probably best suited to this type of purpose would include most sciences, especially physics, psychology, econometrics and the more quantitative aspects of the social sciences.

FUTURE TASKS IN EDUCATION. Dr. J. Toman. Czechoslovak Academy of Sciences. Prague, Czechoslovakia.

In the field of documentation it is necessary to devote attention to two great problems:

1. **The education of specialists in the art of personal documentation.**

   The "explosion of information" cannot be mastered by the efforts of the information specialist only. Every future scientist, technician and specialist must be educated at school already in the art of personal documentation.

   There must come about a revolution in the minds of those who prepare the high school and university curricula. The education at schools in the whole world is based on the assumption that the student must acquire as much knowledge as possible by storing it in his memory. But the capacity of the memory is limited and cannot expand. This problem can be solved only by admitting the fact of limitation of the memory and by teaching the students how to store information in other media than memory. The art of personal documentation must be included in the curricula of schools.

   General understanding for the problems of scientific information will be spread only when every student will master the fundamental principles of information technique and when most of them will experience their personal documentation.

   We see from experience especially in the pure research that the personal documentation of individual scientists can be well coordinated with the effort of information centres.

   The methods of personal documentation must differ from those of information centres. The author of this paper first wrote a manual of documentation technique and soon came to the conclusion that there is a great need for a book about personal documentation. Shortly before his book about the technique of personal documentation /"How to collect knowledge"/ had been published in 1960, he was impressed by the fact that the American scientist Engelbart /Stanford Research Institute, Menlo Park, California/ wrote in the American Documentation about the importance of devoting attention to the elaboration of special methods of personal documentation /"microdocumentation"/ in the same way as the author in Czechoslovakia.

   The experience of the author shows that a general book can satisfy a large circle of readers. The scientists only would need a special book containing a higher degree of knowledge, in accordance with their needs.

2. **Accumulation of practical experience with different methods of documentation.**

   The author was entrusted by the Committee FID/TD /Training of Documentalists/ to compare existing manuals of traditional documentation technique. It is interesting to state that most of the newly edited books in the field of information
are devoted to special subjects and to the methods of mechanical retrieval and that there exist only very few books in the world comprising the whole complex of documentation methods, the knowledge of which is so important for the majority of information centres.

The author came to the conviction that we should give our energy to the accumulation of experience with different documentation methods used for different stages of the information process.

It would seem that under these circumstances it is rather difficult to accumulate descriptions of thousands of information systems. But when analyzing many of the information systems we realize that although we do not encounter two similar systems, we find the same methods in them. We come to the conclusion that each information system consists of a number of activities, which can be performed by different methods. As there are on the average no more than 10 different methods used for each of these stages /activities/ we find that the total number of methods, of which the information system can be built does not exceed 100-120 methods.

Thus an information system is only a combination of these different units /methods/. The effectiveness of an information system depends on the choice of suitable methods for different stages of the information process according to the needs and conditions of the organization where the information center works.

The accumulation of experience with all these methods, which form the different information systems, would not demand much time and money and still it would result in a survey of the units from which the suitable information system should be built.

We should devote our attention to the project of this kind. The accumulated experience published in a manual could form the basis of an important subject in the curricula of schools and courses for the training of documentalists.
The RAND Corporation, Santa Monica, California, U.S.A.

Sentences occurring in ordinary text are often extremely efficient forms for storing information and a single sentence may be capable of providing answers to many questions. Their surface structures reflect the preoccupation of a writer with his immediate purposes for organizing his data, for emphasizing some aspects of it and subordinating others. Various stylistic choices are also open to him. As a result, semantically equivalent content may be embedded in a great variety of syntactic structures. Moreover, automated parsing grammars assigning structural descriptions directly to sentences in a text encounter great difficulty. If applied heuristically, they may miss a valid structural assignment that correctly correlates an expression with equivalent paraphrases and relevant questions. If applied algorithmically, they tend to produce an unmanageable number of parses, of which a surprising proportion correspond to possible ambiguities in interpretation and are therefore not eliminable on syntactic grounds.

Recent advances in grammatical theory provide a framework for constructing grammars with a base component to produce or recognize a relatively small set of simple, unambiguous "deep" structures from which the more complex surface structures of sentences are derived by a transformational component which rearranges, conjoins, and embeds. The transformations do not alter meaning, although they are frequently sources of ambiguity. Consequently, the unmanageably large number of multiple analyses produced by loosely constructed grammars applied to surface structures may be reduced by subjecting them to inverse transformations and comparing the remaining structures with a tightly constructed grammar for the simpler deep structures from which any valid surface structures must be derived. Since these deep structures are unambiguous, they can provide appropriate canonical forms for semantic analysis and structural matching in question-answering and deductive systems for information retrieval.

In the last six years, several research groups have attacked the problem of designing automated question-answering systems based on text rather than on highly structured data bases. They have used various techniques of syntactic and semantic analysis, in varying combinations. The view proposed here is that semantic (and other) techniques will prove more effective if they are applied after a syntactic analysis that makes the deep structures explicit.

The major linguistic task is to provide analytic recognition grammars with transformational components adequate to deal with the complexities of a full syntactic analysis, so that the necessarily ad hoc simplifying assumptions of previous question-answering systems can be largely dispensed with. Until quite recently, transformational grammars have been written to generate rather than to analyze and recognize the sentences of a language, although Matthews proposed a technique for analyzing a given sentence by synthesis from a transformational grammar, as early as 1961. Work on the recognition problem has now been undertaken, and three different types of grammar are being provided with transformational components designed to recover deep structures automatically. Kuno reports some experiments with the Harvard predictive analyzer, which are designed to produce kernel sentences con-
currently with the analysis of the surface structure. For phrase structure grammars, three methods have been proposed by S. Petrick, by M. Kay, and by the MITRE Language Processing Techniques Subdepartment. An "approximate" formalism to obtain structural descriptions very similar to deep structures is being developed by Lieberman, et al, at IBM Research. Although applied to a phrase structure grammar, the formalism is intended to be applicable to other models as well. Robinson experimented briefly with a paraphrasing routine for a phrase structure grammar, but is currently designing a dependency grammar with a transformational component, in collaboration with Hays and Kay.

Several machine translation groups are also incorporating transformational features into their grammars, designed to uncover deep structures, in accord with Harris' assumption that many languages are more similar in their kernel sentences than in their total structure. Linguistic work in information retrieval is obviously closely related to work in translation, but no attempt is made here to discuss the latter field.

Bibliography


SOME LOGICAL ASPECTS OF THEMATIC REPRESENTATION WITH EXPLICIT TIES.

Linear sequence, an essential feature of speech transcribed in ordinary writing systems, is a valuable conventional feature in some special notation schemes as well. But when expressions given linearly are listened to or read with some degree of understanding, the linear sequence is in effect attacked, and in general altered. Part of the process of grasping what is said or written consists, in effect, in recognizing which of the sequential features are functional and in distinguishing, for those that are functional, significant differences among the functions locally involved.

Noting this, and noting also the utility of diagrams, tables and index lists, one may propose to dispense with the over-all linearity characteristic of speech and ordinary writing through use of schemes which present or re-present message "content" in another fashion -- a fashion such that explicit ties of several kinds are the basic formal means by which functional components are combined and integrated to form components of greater complexity. By conventions for tie direction, the recorded positions of tied components may then be openly nonfunctional. An abstract model for such a scheme is available in the mathematical notion of graphs (especially, directed graphs) composed of points and (directed) point-connecting lines. Such graphs are most plainly represented by diagrams or three-dimensional constructions or by associated matrices. Straightforward development and plain use of graphic schemes in representing messages or message components can result not only in graphs that are trees but also in graphs that contain cycles and graphs that are nonplanar.

Organic chemists exploit a graphic model quite directly when they choose to represent a chemical compound by a graphic formula or a molecular model, instead of a linear formula or name. For chemists, each of these interrelated modes of representation is a tool available, reputable, and particularly useful on certain kinds of occasion. An analogous latitude of choice could conceivably benefit the work of persons concerned with messaging in a broader way -- for example, persons concerned with information processing in documentation systems, or persons concerned with cognitive aspects of language processing in human minds or brains. Conjectures of what might result, scientifically and technically, from the use of graphic modes of representation in functional connection with linguistic modes seem to invite and justify pursuit of a graphic model in developmental efforts more persistent, penetrating and extensive than those now on record. Such efforts would fall largely within a field called semiotic (the science of signs), which on definition encompasses much of what is specialized in logic and in linguistics.

To consider some matters of logic that are pertinent in a graphic approach to message processing, attention is addressed to a particular project pursuing that approach and to the ways in which such matters are at present treated in it. In the project concerned, themes or topics are represented directly, in a relatively neutral way, and themes play a role in the analysis and representation of propositions and statements. The logical aspects considered are discussed by way of such terms as the following: statements (indicative and other), propositions, and themes (topics); predication and predicates; nonsymmetric predicates; syncategorematic predicates; common names and proper names; connectives; individuals (events
and others); parts of things; quantification; normal forms. The current scheme is exemplified in application to topic-indicating texts.

Bibliography

Note. The following bibliography is confined to several introductory works (one each for linguistics, graph theory and logic), one seminal work that is easily overlooked (Kempe), and one paper by the present author containing references to a few of many further works that might be cited.


THE TRANSFORMATION AND ORGANIZATION OF INFORMATION CONTENT: CONTRIBUTION OF PSYCHOLOGY. Philip J. Stone Dept. of Social Relations, Harvard University Cambridge, Massachusetts, U.S.A.

Since psychologists work with information theorists and documentation specialists at a number of junctures, it often becomes somewhat arbitrary to determine what ideas are to be credited to what disciplines. Contributions begun by specialists in one field may be developed by specialists in another. Such organizations as the RAND or System Development Corporations in the United States or the EURATOM or C.N.R.S. projects in Europe have had a particular significance for interdisciplinary work in this area.

Without being too presumptuous in presenting claims for my field, certain viewpoints of a distinct psychological bearing can be presented as perhaps having a significant future role in documentation research:

Organization of information
1) View of man as a limited "list processing" organism.
   - limitations in lengths of lists.
   - limitations in organizations of lists.
   - limitations in handling disjunctive concepts.
   - tendency to ignore negative information, i.e. to ignore considering subsets of items that all share in not having a certain characteristic.

2) Use of tree structures to represent human list processing.
   - types of tree structures:
     - Yngve's "seven plus or minus two" list length hypothesis taken from C.A. Miller
     - Feigenbaum's "EPAT" representation of nonsense syllable learning.
     - Hunt's concept learning model.
     - alternative languages for representing list structures on a computer: IPL-V, LISP, CMUT, SNOBOL, DYSTAL, SLIP, etc.

3) Development of documentation techniques to aid in extending human list processing capabilities; emphasis on "shuffling" information according to presence or absence of characteristics.
   - Hunt's model: emphasis of logical characteristics
   - Sonquist's model: emphasis on statistical characteristics
   - representing shuffling of tree structures on "Project Mac" time sharing system, using typewriter console.
   - example results.

Units of Organization: Defining Basic Elements
1) Humans tend to "chunk" information at a rather high level of abstraction.
   - fashions of jargon often represent different chunking patterns.
   - after chunking takes place, composite elements tend to be forgotten; chunks tend to be taught as unanalyzed entities.

2) Within a field, "building block" elements can be identified.
   - "extension" referents: basic measurement procedures used in research.
   - "intension" properties: basic characteristics considered to be relevant to field.
   - implications of identifying further substructures.
examples of "building block" identification in field of survey research.

3) Importance of identifying "building block" elements within newly developing fields.
   - should be a responsibility of the profession, not left to librarians and later historians.
   - "building block" elements provide a flexible cognitive basis for structuring and comparing different theoretical fabrics; places the biases of an era into an empirical and theoretical perspective.

Bibliography


Miller, George A. "The magical number seven plus or minus two, Psychological Review, 1956, 63, 81-97.


Recent developments in classification research are discussed in terms of philosophic approach and methodology. A series of definitions is given in order to clarify the discussion and to indicate parameters of the study. The relationships of logic, mathematics, scientific and probability methods, and linguistics are explored within the framework of current research in classification. Three major types of research procedures are distinguished, and the various schools of thought involved are tentatively identified. The necessary basic intellectual problems to be solved are singled out for attention and the future of classification research efforts outlined with regard to these problems.

BIBLIOGRAPHY


AREA II. ORGANIZATION OF INFORMATION FOR DOCUMENTATION

Symposium B: Comparison and Evaluation of Transformation Techniques and Organizational Structures in Documentation Systems

THE EFFECT OF A CITATION INDEX ON LITERATURE USE BY PHYSICISTS. Ben-Ami Lipetz. Carlisle, Massachusetts, U. S. A.

An experiment was conducted for the American Institute of Physics Documentation Research Project for the purpose of evaluating the impact of a citation index to physics literature upon the literature-use habits of physicists. A citation index was specially prepared in which it is possible to look up a known reference in either of two very heavily used United States physics journals -- The Physical Review and the Journal of Applied Physics -- to learn whether, and precisely where, that reference was cited in four leading Soviet physics journals which are published in English translation but which have only about one tenth as many subscribers in the United States. The Soviet journals which were used as citation sources in the index were Soviet Physics - Crystallography, Soviet Physics - JETP, Soviet Physics - Solid State, and Soviet Physics - Technical Physics. It was hypothesized that because the citation index would help United States physicists to learn of Soviet papers with subject contents similar to the known United States references they would start with, it would therefore tend to increase the utilization of the Soviet physics translation journals by United States physicists.

The experimental citation index was prepared by punched-card techniques. The Soviet source journals which it covered dated from 1955. About 18,000 citations of the two United States journals were found, mostly citations of The Physical Review. The citations were rearranged for printing into year, volume, and page order for each of the cited journals.

The impact of the experimental citation index was measured principally by means of statistics which were collected monthly on utilization of the Soviet physics translation journals in libraries. (There are almost no individuals subscribing to these translation journals.) More than one hundred libraries throughout the United States contributed data to the study. The statistics were in several categories, reflecting borrowing activity, photocopying activity, and display-issue use. Statistics were reported separately for each of the four Soviet physics translation journals used as sources in the citation index, and for an additional four Soviet physics translation journals which were excluded from the index to serve as controls. Statistics were supplied by the participating libraries for a period of months preceding and following the controlled distribution of the experimental citation index. The index was distributed without advance notice to the 550 subscribers to The Physical Review and/or the Journal of Applied Physics in a restricted, self-contained geographic region which accounts for five percent of the domestic subscriptions to American Institute of Physics publications. Some 85 percent of these subscribers are individual physicists, rather than libraries. Library-use statistics from participating libraries in the test region were compared with statistics from participating libraries in the remainder of the country in order to evaluate the impact of the citation index.

The study of library use was supplemented by a parallel study of changes in the subscriber lists for the test journals as compared to the control journals following distribution of the citation index, and by a study of back number orders for the source journals and control journals received by the publisher before and after
distribution of the index. Each recipient of the citation index was contacted by mail to determine whether he had tried to use the index and to what extent. Personal interviews were conducted with a sampling of the index recipients.

The impact of the citation index on literature use by physicists -- as indicated by changes in the ratio of test-region library statistics to control-region library statistics following distribution of the index -- was anything but dramatic. Although a definite impact was detected in the first six months or so after distribution of the index, it averaged out to only about a 15 to 20 percent increase in borrowing and photocopying and about a 40 percent increase in display-issue browsing; at the end of this period the utilization figures appeared to have returned to normal levels a few percentage points higher than before the index was distributed. Since utilization levels were very low to start with, the impact of the citation index in terms of actual stimulated literature utilization is very small. The statistics gathered in this experiment permit the rough calculation of the number of scientist-hours which were diverted to the source literature by the experimental index.

Results obtained from the other evaluation techniques used in this experiment are in good agreement with the library-use statistics. No indications of a large impact attributable to the citation index were found in the statistics on source journal subscriptions or in the statistics on back-number orders which were supplied by the publisher. A survey questionnaire which was sent to all citation index recipients a few months after distribution of the index confirmed the fact that only a minor fraction (42 out of the 260 who replied) had tried to use the index for information retrieval.

Despite their low usage of the experimental citation index, the comments of physicists who had received the index indicated a generally favorable attitude toward the citation concept in literature retrieval. Those physicists who had actually used the index tended to be extremely enthusiastic in their comments.

It is concluded that although the citation index concept is by no means a cure-all for the literature problems of physicists, it can nevertheless be of definite value as a supplement to the existing "universe" of reference aids. It deserves further study and development, and should be taken seriously. No conclusions can be drawn from this experiment regarding the effectiveness of the citation index as substitute for any specific conventional reference aid or aids.


AUDITING PROCEDURES FOR INFORMATION RETRIEVAL SYSTEMS. Björn V. Tell. AB Atomenergi, Nyköping, Sweden.

The need for effective information retrieval systems has been most strongly felt in science and technology. The researcher in the information field seems to have been influenced to apply models developed in some of these science fields, e.g. information theory, automata theory, switching theory etc. Information systems both of nationwide "macro"-size and down to the "micro"-size of that of a documentation center in an enterprise involve the use of manpower. In many cases we have to deal with man-machine systems. Instead of looking for an applicable model from the science field, it is here suggested that good use can be made of models from the social sciences. Economic models display many features of interest for analysis both of macro- and micro-information systems.

For a long time the pragmatic aspect of documentation has been stressed by raising the question about relevance in information retrieval. But anyone who stresses "relevance" as a criterion for evaluation purpose, must answer the question - relevant to whom? and for what purpose? Thus, the documentalist is trapped in the same situation as the economist when he discusses "utility" or "value" as an important parameter. However, the economists have developed important methods without settling upon this question. This paper concentrates on dealing with a standard costs model for evaluation purposes, leaving the question about the very nature of relevance open.

Like an auditor the documentalist who is going to perform an analysis of an information retrieval system, has to test his results against some standard of which he is aware. Standards are scarce in the information field due to the fact that most systems often are in a state of transition. The "Proposed Standard Description for Reporting Evaluation Tests of Retrieval Systems" represents a consensus about pertinent parameters for describing an IR-system. The quantitative data in this standard form deal with items, such as documents, index terms, time, costs, personnel and equipment. From an auditor's point of view, money is the common denominator in terms of which the exchangeability of these items can be measured.

There is a market price for many of the items in the standard form, but for some of the parameters we have to make use of estimated costs. As in accountancy, if we use estimated costs, the results will be less accurate than based on market prices, but they might well prove useful for the analysis. The items which have no market price can be converted to expenses by adopting some conventions for conversion.

An auditor's evaluation is concentrated on a study of the results. However, these results presuppose a forecast or a plan according to some goal, and his evaluation consists of a comparison between a plan and its execution. The analysis of variations from standard costs is a review technique which operates by means of a comparison between the results expected according to the plan and the results as revealed in the cost accounting. The results expected according to the plan is designated by the term "standard".

The goal of an IR-system includes the maximization or minimization of one or more factors. However, the present performance of a system can also be regarded as
a goal or a plan. If during a stated time period, the observed data are taken as the intended goal for that period, this will fix a starting point or a "standard" for further planning. Suggestions for changing the plan in order to maximize one factor, will result in a new plan, the execution of which will take place in the new time period. The results of the plan on completion can be evaluated by a study of the variations from the standard.

Attention is called to the irrelevant documents which are treated as "scrap", i.e. the cost for their retrieval has to be charged against the value of the retrieved documents. That the "scrap" can be reprocessed for new questions is accounted for.

Examining the standard costs variations opens a possible way to explain differences in performance within a system at different times. These explanations can be used to formulate changes in the initial conditions in such terms that the changed or new model can serve as an explanation of the observations which led to the rejection of the old one. It is suggested, finally, that changes should be undertaken as long as the costs of the changes are less than the increased value of the system.

References.

Bourne, C.P., Review of the criteria and techniques used or suggested for the evaluation of reference retrieval systems. (Working paper for the NSF Conference on the evaluation of reference retrieval systems and procedures. 1964).

EVALUATION PROCEDURES FOR COMPUTER-BASED RETRIEVAL SYSTEMS, Gerard Salton.
Computation Laboratory, Harvard University, Cambridge 38, Massachusetts, U.S.A.

The interest in evaluation procedures for information retrieval systems and techniques stems from two principal causes: first, more and more retrieval systems are being designed, thus raising a question concerning performance and efficacy of these systems; and second, evaluation methods are of interest in themselves, in that they lead to many useful problems in test design and performance, and in the interpretation of test results.

The work reported here differs from other efforts on systems evaluation in that it deals with the evaluation of automatic, rather than conventional information retrieval. The computer-based information systems of the future may be characterized by the following principal properties:

a) the systems will probably operate in a time-sharing environment, in such a way that a multiplicity of users may be given simultaneous access to the files;

b) the search function may be undertaken iteratively, possibly under user control, by performing several partial searches to approach the desired subject area, rather than a single one-shot process;

c) several different analysis and search procedures may be incorporated in such a system, including, in particular, stored intellectual aids for vocabulary normalization;

d) the size of the document file to be searched may be expected to consist of 100,000 or more documents; under the circumstances, a manual assessment of relevance of all documents with respect to each search request will become impractical, and sampling techniques will be needed to measure retrieval effectiveness;

e) the computation of sophisticated correlation coefficients between documents and search requests will make it possible to present ranked document output in answer to the search requests, in decreasing order of the correlations between documents and request.

In the present report, evaluation techniques and procedures are discussed which are specifically based on the foregoing systems organization. Such important systems parameters as cost of retrieval, response time, influence of physical layout, personnel problems, and so on, are disregarded, and the emphasis is instead on the evaluation of retrieval (that is, analysis and search) techniques.

Since the evaluation process is restricted to automatic systems, a number of human problems which complicate matters in a conventional evaluation situation, including, for example, the difficulties due to inconsistency among indexers, or to the presence of search errors, are no longer so crucial. The evaluation process can therefore concentrate on the main systems parameter: its ability to locate relevant information, while rejecting nonrelevant material.

The design of automatic evaluation systems is first discussed in detail. Recall-like and precision-like measures are then introduced which seem appropriate...
for the evaluation of computer-based retrieval systems. The use of these measures in evaluating retrieval performance of automatic, iterative, search systems is discussed, and test results are presented showing results obtained by using the fully automatic SMART retrieval system, now operating experimentally on the IBM 7094. (This system incorporates a large number of automatic search and analysis procedures, and makes it possible to evaluate the effectiveness of each one by processing the same search request against the same document collection in many different ways, while comparing results in each case.)

The prospects and problems of automatic search systems are summarized, with reference in part to existing systems, such as the NASA, MEDLARS, and DDC systems, and tentative conclusions are drawn concerning the design of fully automatic information systems.

REFERENCES


The discussion is focused on two main themes: identification of significant measurable characteristics of retrieval system performance, and development of adequate mathematical models to serve as a basis for making measurements. The performance characteristics to be measured must be clearly relatable to the operationable objectives of a specific retrieval system at hand, and must take into account appropriate boundary conditions such as nature of collection, user population, and expected kinds of search requirements. Quite different mathematical models and measuring techniques might be appropriate, depending on the objectives and boundary conditions at hand. Even within a given set of boundary conditions, the mathematical model should provide means for taking into account the numerous sources of variance encountered in practice—particularly, variance in depth-of-search requirements among different users, variance in relevance judgements among system evaluators, and variance in system behavior from one query to another.
Studies of information gathering habits of scientists and engineers (SIG) are conducted to learn how and why scientists and engineers use information and information systems (libraries, technical information centers, personal indexes, and other sources of information). These studies also reveal scientists' and engineers' opinions of existing and proposed information systems. The results of these studies are intended to provide information specialists (librarians, information scientists, etc.) with a set of guidelines that will help in the evaluation of existing information systems and in the design of improved information systems. Despite the fact that numerous SIG have been conducted in recent years, basic questions, such as, what is a good information system?, still remain unanswered.

A composite picture of scientists' and engineers' information needs, uses and sources is drawn in this paper from reports of SIG. SIG techniques are examined in terms of type of information that each technique can provide, reliability, and other characteristics.

Computer-based information systems, as exemplified by selective dissemination of information systems and the Project MAC Index at M.I.T., open up new techniques for collecting records of information use as a by-product of system use. Records of use of computer-based information systems can be obtained without disturbing the user and without the need to rely on memory or other subjective factors.

Different aspects of the problem of information needs and uses have to be studied with different SIG techniques. Some aspects of the problem, as exemplified by the use of indexes by future generation of scientists, cannot be determined by any technique. While SIG techniques are far from perfect and leave some questions either unanswered or incompletely answered, they are a basic tool in the design and evaluation of information systems.

In a personal interview survey designed to guide future editorial decisions of a scientific journal, an attempt was made to collect information on the user's reading environment, reading behavior, attitudes and opinions of the journal, and respondent's perception of new editorial subjects. In addition, data on respondent's age, educational background, and publishing pattern were collected.

Personal interviewing of users in an effort to determine their information needs requires careful attention to determination of study objectives, selection of sample, collection and tabulation of data, and finally interpretation of the data.

The objectives of a survey are not fulfilled until the results have been analyzed, interpreted, and then applied. Frequently the latter is the most difficult step to perform since the replies are rarely clear-cut guides to action. When the survey is used as a continuing aid to editorial judgment, it serves as a valuable tool to journals seeking to meet the information needs of scientists.
User-needs studies can be put into two major classifications: those of narrow scope for design or improvement of a specific information system and those of broad scope for the planning or improvement of a network of information systems. Most of the user-needs studies conducted to date belong in the first category. They have been confined to either a narrow segment of the technical community, e.g., the users of a particular library, or to a narrow area of investigation, e.g., the media used to acquire information. The study recently completed by DOD is one of the few user-needs studies conducted that fits into the second category, and it certainly is the largest.

The DOD study was extremely broad in scope, both in terms of the interdisciplinary nature of the population surveyed and of the areas of investigation. The principal purpose of the DOD study was to collect and analyze a statistically significant data base on how DOD scientists and engineers presently acquire and utilize technical information in the performance of their tasks. The population surveyed comprise the entire research, development, test, and evaluation (RDT&E) community within the Department of Defense. Its 36,000 members represent practically every technical discipline and are engaged in a wide variety of interdisciplinary tasks of a scientific, engineering, or administrative nature.

The methodologies generally used in the previous user studies of narrow scope were not appropriate to a study of the broad scope and magnitude of the DOD User Need Study. Consequently, it was necessary to utilize some survey techniques which had never been previously applied to information needs studies. For example, a principal feature of the survey methodology was the application of the critical-incident interviewing technique. This involved the identification of a task recently completed by the respondent, and the isolation and description of the actual information utilized in the performance of that task.

The critical-incident technique prevented the study from becoming an opinion poll and thereby eliminated the primary source of bias in most user studies. A semi-structured interview guide and handbook were employed to aid the interview process. The structured portion increased the consistency of the question interpretation and simplified the recording of the responses. The unstructured portion was designed to allow for the identification of unanticipated information patterns.

The methodology best suited to a particular user needs study depends upon the objectives of the study. The methodology of the DOD study is not particularly applicable to system-requirement studies confined to a single discipline or a relatively small organization. In these types of studies, it usually is practical to use a considerably less structured survey technique, which facilitates deeper probing of promising areas as they develop, rather than restricting the interview to a rigid format. It may also be fruitful in such cases to determine the information requirements of specific projects or organization functions rather than a random sampling of recently completed tasks.
On the other hand, the methodology used in the DOD study will be highly applicable to conducting future user studies of large heterogeneous populations. The methodology provides for relating specific information requirements to specific task descriptions. This, therefore, produces reliable data useful in the overall planning of a comprehensive network of information systems before attempting to design or improve various systems and subsystems which only service the needs of individual segments of the population.

AUERBACH Corporation

"DOD user needs study. Phase 1, in two volumes."
AREA III. INFORMATION NEEDS OF SCIENCE AND TECHNOLOGY

Symposium B: Increasing the Efficiency of Information

INCREASING THE EFFICIENCY OF THE USE OF INFORMATION - A BACKGROUND REVIEW.

About Scientific and Technological Information

There is great diversity in the format and content of scientific and technological information. For illustration, a classification of "uses" is offered which covers the range from specific, well defined data to speculative search of a general problem area.

Increased efficiency might be sought by classifying users (bench chemist, research administrator for a pharmaceutical laboratory) and kinds of information and investigating their interactions.

About Combinations of Elements of Information

Some work has been done on frequency of "association" among ideas and on plausible associations in one-step and higher order links. Such approaches can utilize the abilities of computers and might be used to extend the inquiry, to make for more fruitful search, or to stimulate thought. The particular result depends upon system characteristics not identified here.

There is also a case for unlikely combination - for stimulative effect. This possibility has not received much attention. There is some evidence that unlikely combinations and the unconventional application of known techniques or apparatus promote new directions in creative thought.

Increased efficiency might be sought by abandonment of the one-document or one-piece-of-information approach for much of the user spectrum and pursuing research on combinations and linkages. This approach focuses on the information itself rather than the document or index term.

About People

Individuals exhibit wide ranges of creative talent, knowledge, needs. Each person in each situation has his present compromises - unstated assumptions about what it is futile to ask for - and these are certainly far from the ideal satisfaction of all useful needs for access to recorded knowledge. His statements about his needs will almost always reflect his bondage to present habits of obtaining and using technical information.

In practice, his statement of need prior to search will frequently be inconsistent with his success rating after the search - for the later classifications suggested above - because he learned something during the search. Even some evidence that the modifications of statements about need are different for manual and mechanized search.

Increased efficiency of use might follow more studies of real need, as distinct from conventionalized and compromised statements of need.
About Language and its Representation

Scientists communicate in coded languages as well as natural languages; chemical formulas, mathematical formulas, circuit diagrams, mechanical drawings, graphs, block diagrams, maps, photomicrographs, etc.

In such specialized languages, they think, query, make statements, record information, and screen collections of information for their own purposes.

Some of the coded forms of scientific information has been shown to be receivable by computers, thus opening the way to a variety of useful mechanized logical processes. The early experience appears useful but there is a paucity of disciplined experimentation.

Increased efficiency seems likely to follow progress in the technics of computerized manipulation of information, provided that the specific logical operations can be imbedded in appropriate search and retrieval systems.

About Tests and Evaluations

Formal evaluation of actual search experience poses problems of classic difficulty: the activity is largely mental and covert; the subjects are knowledgeable and clever human beings; the object of the activity is subject to change in midstream; and scoring poses difficult and subjective problems. Yet progress depends as much upon valid testing as upon technical achievement.

About Technics and Organization

The above discussion relates to ways of doing things. There is a whole range of additional factors that should affect our efficiency in the use of information - what might be called organizational matters. They relate to the assignments of functions and responsibilities, the provision of financial and technical support, the making of choices of strategy and program.

They are left for one of our distinguished discussants, Dr. Alvin M. Weinberg, whose recent report on these matters has had unprecedented influence. /2/

THE M. I. T. TECHNICAL INFORMATION PROJECT. M. M. Kessler. The Libraries, Massachusetts Institute of Technology, Cambridge, Massachusetts, U. S. A.

The Technical Information Project (TIP) at the Massachusetts Institute of Technology is an experiment in information-system design. It provides a facility to investigate by direct experience what contributions modern technology can make in solving the problems of scientific communication. The system will be described and demonstrated at the FID Meeting by teletype connection between the lecture hall in Washington, D. C. and the installation in Cambridge, Massachusetts.

The system consists of five major components: a sample literature, a computer facility, a library of programs, a population of users, and a test and monitor procedure.

Elements of the incoming literature are key punched, edited, compressed, and transferred to magnetic tape. This tape is kept as a back-up record. For processing purposes the material is transferred from tape to the disc memory where it is always available for use.

The material in store is taken from 25 journals of physics. For each of the articles in each of these journals we record the location of the article (journal, volume, page), the title, authors, the institutional affiliation, the citations (journal, volume, page), and the location of the article in Physics Abstracts. Periodic monitoring programs test the integrity of the data.

The computer facility is itself an experiment (Project MAC). It consists of a central machine with 150 remote consoles. The consoles are standard teletype units distributed largely around the Massachusetts Institute of Technology. Contact with the computer is by means of telephone. The 150 consoles are available to perhaps 500 people who may at any time use the computer on a time sharing basis. Thirty people may use the computer at the same time.

The computer operates on the literature through a set of commands such as "Search Physical Review Volume 128 and print the titles of all papers by John Smith". Searches may be initiated by author, title, citation, index and/or bibliographic coupling. The search range and the output format commands are flexible and under user control. Several logical tools are available to the user such as "and", "or", "but not". The communication facilities of the system may be used to transmit messages from system to users, from users to system and between users. The user may also receive a print out of all available programs. This facility is being elaborated into a teaching program so that new users may be instructed by the computer.

The use of the system as a research tool will be demonstrated at the Meeting.
CASE STUDY: INCREASING THE EFFICIENCY OF THE USE OF INFORMATION, AEROSPACE RESEARCH APPLICATIONS CENTER. Dr. Howard L. Timms. Aerospace Research Applications Center, Indiana University, Bloomington, Indiana, U. S. A.

During the past two and one-half years, the Aerospace Research Applications Center (ARAC) at Indiana University has been providing information retrieval service of various types to the scientists, engineers, and technicians at subscribing companies totaling forty-five at the time this abstract was prepared. The ARAC project is the first university-based information center operating under contract with the National Aeronautics and Space Administration (NASA). The project's objective is to develop methods of disseminating to private companies the scientific and technical information generated in the nation's space programs. It is believed that this information may be useful in the development of new and improved products, processes, and materials by private companies for non-space and non-military markets. The project is one of several being conducted by NASA under a requirement of the 1958 Act of Congress that established NASA.

The major information services provided by ARAC are as follows: (1) retrospective searching of the NASA file now totaling nearly 150,000 documents, (2) selectively searching incoming documents against company interest profiles to provide a current awareness service to company "interest centers." Incoming documents total nearly 4,000 per month. Selective dissemination searches against interest profiles are made twice-monthly. Other services are provided by the Center. They are described in the paper that is abstracted here.

The NASA document file is prepared by Documentation, Inc., NASA's subcontractor that processes raw documents resulting in the file input to ARAC. This file is received monthly by ARAC in two forms: magnetic computer tape and microfiche. Also received are the corresponding abstracts in microcard form.

Increasing the efficiency of the use of this information may be treated in two parts: the first concerns organization on the part of the subscribing companies for efficient ingestion of scientific and technical information; the second part concerns preparation by ARAC of its output to subscribing company people.

Concerning organization on the part of subscribing member companies, this subject is extremely complex and few principles of an operative nature are as yet available to guide companies in this effort. The results of formal efforts by ARAC and company representatives to establish such principles are described in the paper abstracted here. These formal efforts were in the form of a workday on one occasion, and a panel discussion a year later. Over its period of operation to date ARAC has worked informally with its member companies on specific organization for information transfer at each company. The results of these efforts, which for the most part have been successful, are described in the paper abstracted here. They involve specific communication and coordination procedures between ARAC operating personnel, company information users, and company information centers and libraries.

Concerning the preparation of its output for efficient use by company people, ARAC early established several guiding policies that appear to have served very well to date. A few of these policies, which are treated fully in the paper abstracted here, are as follows: (1) design the output in such a manner as to fully
suit the user's needs, not in a manner that one might think the user ought to accept it—in other words the user's needs and desires rule, as in any well-designed consumer-oriented product; (2) never send the user a full-copy document until he has seen its abstract; (3) don't badger the user with questionnaires, forms, and other post-operations control devices—build only the minimum necessary, simply-designed, feedback and control features into the operational paperwork.


The aim of this paper is to relate capabilities and limitations of present information systems to the needs of society for information. The needs have been set forth in previous papers of this section of the Congress. "Present information systems" subsumes both communication systems such as the mass media, mail, telephone, and telegraph and storage-communication-and-processing systems such as command and control systems, management information systems, document systems, and libraries. The latter set, however, constitutes the focus of this discussion. The capabilities of those systems are examined in respect of storage, communication, and processing capacities and of their abilities to meet the needs for pertinence, timeliness, convenience, and economy.

A main conclusion is that, except in one area, the capabilities of the hardware components of present information systems promise more than the over-all systems deliver, that the bottlenecks of information technology lie not in component devices but in the area called "software," in system planning and design, and in the overcoming of organizational and social intransigence and inertia. The exception is the interface between information machines and people, a domain in which device technology has lagged, but in which there is now much interest and good prospect for advance.
POLICY PROBLEMS OF A DATA-RICH CIVILIZATION. Harold D. Lasswell. Yale University, New Haven, Connecticut, U. S. A.

Open access to sources of information and freedom of dissemination will be restricted so long as the arena of world politics is constrained by security considerations. At present world public order provides an incomplete system of national security. Hence every nation state relies on covert as well as overt information.

If information becomes more inclusive and realistic, political decisions may become more focused and rapid. If the United Nations or any organization of states is permitted to improve its sources of information, security estimates may become more reliable for everyone. Hence the possibility of surprise at least can be largely eliminated.

Allowance must be made for the strength of forces that favor monopolies of dissemination held in a few hands. Powers that undergo forced industrialization use information control in order to undermine local sentiment, and to encourage national identity. National not transnational perspectives are fostered. We now see that the initial effect of an expanding civilization of science and technology is not the universalization of a world outlook. Hence it has been suggested that heroic measures be taken in the hope of overcoming the parochializing and divisive consequences of information unbalance along national lines. Drastic proposals include a world communication network that furnishes news to all people everywhere, and a world school board that commands attention of all school children for at least an hour a day.

It is to be predicted that wherever popular government is well established and competitive electoral propaganda is important, research information about public motivation will grow more refined. Unless voters and officials are provided with an up to date stream of information about how they are manipulated, popular government will be deprived of the degree of choice that they might otherwise exercise.

In a data rich civilization the most serious constraints on information are likely to be in processing and use. There is danger that modes of perception will be standardized and that motivations will be turned toward unimaginative channels of expression. As mass society grows the leadership will no doubt become aware of the latent threats to established order. Hence it will be tempting to manage education for purposes of rigid indoctrination.

Two broad lines of policy regulation are conceivable in regard to information: limit gathering and storing; limit access. It is most improbable that initiatives toward the former policy line will be accepted. The latter, however, does come within the scope of likely action. For example, public regulation is likely to define who may obtain access to personal data.

It is generally agreed that great damage is done to children if derogatory information is made public about them individually. Attempts are made to avoid reporting the stigma of illegitimacy, of early delinquency, or of initial failure. Adults are also given the benefit of some non-disclosure policies, and no doubt pressure will continue to be exerted to screen health and other records from miscellaneous hands.
However, I suspect that attempts to protect human dignity by the selective enforcement of privacy will almost certainly become obsolete. Privacy won't work. Coming years may witness a drastic new approach for the purpose of protecting our self esteem. Instead of relying on privacy we will shift to strategies of insight. A comprehensive stock of social information will provide facts that reduce the false pride of an individual, family, or community. If the educational system trains the individual to see his enormous potentialities and to discover the many ways of overcoming specific negatives, the resulting character systems will be able to take or leave privacy without grave inconvenience.

Unless the storage banks of society are made accessible to competing organizations strong enough to utilize and interpret great amounts of data, the monopoly exercised over enlightenment by privileged organizations will throttle the expansion of knowledge. Furthermore, unless provision is made for continuing insight into perceptual limitations and also into restricted motivations, the information potentials of society will remain poorly adapted to the requirements of science and policy.
ISOLATION PROBLEMS IN THE FIELD OF ANTHROPOLOGY. Stephen N.
Shelby, American Anthropological Association, 1530 P St., N.W.,
Washington, D.C. 20036, U.S.A.

The study began with producers of information to determine whether a "backlog" of unpublished research existed. Questionnaires were sent to all professional members of the
Association and 18 percent replied without follow-up. Interviews were conducted with 80. It was discovered that a very
high proportion of the research conducted over the past twenty
years was unpublished and consisted largely of field notes, which
are generally unanalyzed by other workers and unvalued
as to content and significance, and data reports, which
form the organization of data into descriptive categories.

The study was in the pre-conception stage of its work.

Increasingly, the individual scientists, and by implication
anthropology itself, are asking some of the following
questions:

1. What is the current status of the field of anthropology?
2. How can anthropology be improved?
3. What is the role of anthropology in society?

The study began with a survey of the literature, particularly
those works that have been cited by other authors. It was
found that a great deal of the work has been done by
individuals who are not members of the professional
anthropological community. This suggests that there may
be a need for a new approach to the study of anthropology,
one that is more interdisciplinary in nature.

INTERNATIONAL SYMPOSIA

Symposium B: Specific Knowledge Areas

INFORMATION NEEDS IN EDUCATION. John E. Hampf, Educational

The phrase "information needs" implies a situation in
which people have knowledge deficiencies that prevent
them from achieving their goals. Whether the information
needs in this sense can be classified into a number of general
categories is not yet known. Nevertheless, the study of
information needs is important because it can provide
insights into the ways in which people learn and use
information.

The study began with a review of the existing literature
on information needs. It was found that there was a
considerable amount of work in this area, but that
there were few systematic studies. A new approach was
taken in which a sample of users was interviewed to
learn about their information needs.

The study was in the pre-conception stage of its work.

Increasingly, the individual scientists, and by implication
anthropology itself, are asking some of the following
questions:

1. What is the current status of the field of anthropology?
2. How can anthropology be improved?
3. What is the role of anthropology in society?

The study began with a survey of the literature, particularly
those works that have been cited by other authors. It was
found that a great deal of the work has been done by
individuals who are not members of the professional
anthropological community. This suggests that there may
be a need for a new approach to the study of anthropology,
one that is more interdisciplinary in nature.
THE INFORMATION NEEDS OF POLITICAL SCIENCE.
Karl W. Deutsch, Yale University, New Haven, Connecticut, U.S.A.

Political Scientists need three main kinds of information: 1. current data on office holders, legislators, recent votes and other events where speed of information is essential; 2. background data from economics, psychology, sociology and other disciplines, for current or past events, where comprehensiveness and accuracy are essential; and 3. time series and historical data for past crises, decisions, recurrent configurations and long-term trends. Beyond the raw data, Political Scientists need information on their margins of error of these data, and of the permissible errors in using these data in the contexts of particular theories or efforts at prediction. In the unfolding dialog between theory and data—such as at the Yale Political Data Program—the deeper statistical and mathematical analysis of data is becoming as important, or more important, as the more gathering of primary data. Data archives are thus changing into data laboratories, and computing, analytic and retrieval facilities should be designed accordingly.

INFORMATION NEEDS IN THE FIELD OF ECONOMICS
Edgar J. Ducv, Jr., Resources for the Future, Inc. 1755 Massachusetts Avenue, N.E., Washington, D.C., U.S.A.

In recent years research in economics is moving in two directions that require new information resources and place new and unusual demands upon existing information sources. First, the most pressing social problems of the day have to do with structural phenomena at quite a disaggregated level of detail. This contrasts markedly with the earlier preoccupation of the profession with economic stability as measured by aggregate indicators. Second, economists are becoming increasingly satisfied with static state equilibrium models and are seeking new ways to structure their concepts of reality. Stochastic or probability models and systems analysis mark this trend. There is also a trend toward considering a broader set of significant variables extending beyond earlier concepts of the boundaries of the discipline. These trends have been accelerated and reinforced by the advent of the computer which has made available a new order of economy and technical research capacity. These developments have underscored sharply the inadequacy of the systems that generate and make available existing social science and economic information. There are anomalies on both the demand and supply side of the information process. For a number of reasons these problems are qualitatively different from those characteristic of the physical sciences.
SOLUTION TO INFORMATION PROBLEMS OR OBsolescence? Dr. Eugene B. Konecci.
National Aeronautics and Space Council, Executive Office of the President,
Washington, D. C., U. S. A.

Accurate and timely information is the lifeblood of any organized system, be it
government, management of a corporation, scientific knowledge, or intelligence. From the earliest of
times man has been dependent on a means of communication between himself and his immediate
family, his tribesmen, and even communicating with domesticated animals. Since communication
is not a one-way process, even the early cave man had to learn to distinguish not only sounds like a
growl, or later, words which form the first vocabulary, but in fact had to use his eyes in the
same way that we use ours for determining facial expressions, friendly or hostile
acts by other cave men or animals. A great deal of hereditary instinct such as the
need for food, survival, played a large part in the motivation and desire for
communication with other primitive people. The more cultured and advanced
societies of our past history, such as the Chinese, Egyptians, Greeks, Incas,
Romans and so forth, established a communication and information system to a high
degree. Mathematics was founded and based on symbols. People trained in the
field were able to communicate meaningful information. These facts and knowledge
have been winnowed and sifted through generation upon generation for several
thousand years of civilized existence of man.

Today we come to a steppingstone in history in the area of information
processing, storage, retrieval, and so forth. It becomes quite apparent that with
our scientific and technological advances we have become a world that could be
buried in megatons of paper. The written word has become an important symbol of
prestige and earning power. It is the socially accepted manner of communicating
information on a wide scale. The scientist, the systems manager, the executive,
and many others are drowning in this mass of ink and paper. Creativity, free
thought and progress are being held back by our obsolete methods. The technological
solution to this problem is available. However, people in various walks of life
are very reluctant to change the accepted standards of newspaper, magazines,
journals, and textbooks as being their forms of documentation and communication.
This paper will discuss the problem, and alternatives.

Computers, data processing equipment, systems analysis techniques and sophis-
ticated human engineered display and control equipment can help us solve the
mechanical aspects of the information problem. But to maintain a high degree of
scientific, technological and social progress we will have to improve our education
in and out of formal institutions of learning in order to upgrade our people. This
means we will have to very well understand how the human mind works, in other
words foster the new field of Biocybernetics. Information and documentation is not
an end unto itself. It is a requirement in our evolutionary progress. Goods are
plentiful and cheap - where labor, skilled labor is expensive and scarce. We must
have information systems to continuously and dynamically upgrade man's skills to
their fullest capacities and not permit human obsolescence.
PRINCIPLES OF SUBSTANTIVE ANALYSIS OF INFORMATION. Vladimir Slamecka
School of Information Science, Georgia Institute of Technology, Atlanta, Georgia, U.S.A.

In the context of this paper, substantive analysis of information is understood to be
an intellectual or intellect-imitating activity serving the input function in information stor-
age and retrieval systems. A model of this activity is developed permitting to differenti-
ate between the process of information reduction, as in abstracting, and information cate-
gorization, as in indexing. Emphasizing the state of the art of operating systems rather
than of hypothetical or experimental ones, the principles of substantive analysis of infor-
mation presented are of the nature of essential, underlying constituents, and of rules of
action. The principle of relativity of information is introduced as a fundamental con-
straint in the process of substantive analysis of information. Based on this principle,
generalized rules of action, employing predominantly theorems and concepts of mathemati-
cal logic and information theory, are postulated for the activity of information representa-
tion and categorization, and for information structuring (the formalization of intellectual
relations between information classes). Some implications of these principles on the de-
sign of information retrieval systems are discussed.
THE ROLE OF PAPER TAPE AND OPTICAL SCANNING COMPUTER INPUT IN TEXTUAL DATA PROCESSING. Raymond P. Wishner. Documentation Incorporated, Bethesda, Maryland, U.S.A.

For many years to come the search system of a mechanized library will interrogate and manipulate for indexes surrogate records rather than entire documents word by word. Many aspects of the creation and maintenance of a document surrogate may be likened to creating and maintaining a stock record in inventory systems found throughout mass production industry. Unlike the maintenance of an inventory status system is the necessity that many entries of the document surrogate record must be nearly free form text. Examples of this are the document title and abstract. Whereas in a standard file maintenance system the form rigidity of punched cards is welcome discipline to the systems designer, the rigidity of the punched cards is an uneconomic constraint for "inputting" free text. This paper will not fight the ten year running battle between paper tape and punched cards as the basic input medium for a digital computer. Rather it will use the argument of the superiority of paper tape for the inputting of "free text" data into a computer as a bridge to the argument of the potential superiority of a rigid font optical scanning system for the inputting of "free text" data in a mechanized library.

The argument of paper tape versus punched cards as an inputting mechanism has been fought out in terms of rates of key strokes and the character reading rates of computer input devices. In relation to free text inputting, however, these arguments are really secondary to the advantages of textual correction techniques of one medium over the other. In this area, systems designers have often discarded the most important advantage of a paper tape typewriter by using correction procedures well adapted to punched cards but quite senseless for a paper tape oriented system. The two most common paper tape correction procedures are: 1) having an operator blank out a paper tape from the point of error and rekeying the text from this point, 2) a character by character verification of the text by rekeying the data. The object in both methods is to produce nearly perfect input data to the computer in one operational pass.

An alternative and powerful correction technique utilizes the visual copy created in the process of keying the text on a paper tape typewriter. As the data is being keyed, the operator may semantically recognize an error in what he is directly keying or for that matter in a previous paragraph. By giving the operator a base point (i.e., start of last paragraph) and a set of correction codes (i.e., delete, insert, and replace) he can make the correction immediately or delay it to the end of the current paragraph. Regardless of the details of the correction scheme an unambiguous correction can be coded for computer batch processing and a record of the correction made for later visual verification. Further, there is no reason that correction procedures involving the visual copy should stop here. The visual copy can be removed from the typewriter and sent to a proofreader and corrections annotated on the input proof copy. The proof copy can be re-inserted into a tape typewriter and the corrections keyed into an exception paper tape reel. As long as the base record and base point within the base record are clearly coded, an unambiguous correction for computer manipulation again can be created and a record of the correction made for visual verification.
Note must be taken that initial input proofing and correction is often not sufficient in textual data processing. Thus it may be necessary for the computer to print out visual copy for paragraph verification. The same procedure used for initial input proof and correction can be used for computer printed proof copy. . .the advantage is naturally one "well-defined" system for editing.

At this point part of the basic thesis of the paper will be now clearly delineated: The advantages of using the initial proof copy created in keying the paper tape codes as a correction medium are:

1. The ease with which the operator can create corrections for immediate errors and errors created earlier in the keying process.
2. That the proof copy can be removed from the keying machine and later inserted again for the creation of corrections.
3. That the same system used for creating initial textual corrections is nearly the same as the system for correcting computer printed proof copy.

The significant disadvantage to the above described input proofing and correcting scheme is the potential physical separation between the visual proof copy produced by the typewriter and the paper tapes that contain the original copy and later corrections. Systematic handling procedures can negate most of this disadvantage (i.e., batch processing procedures). It is here, however, that the significant advantage of rigid font optical scanning techniques are most apparent.

With an optical scanning system for inputting textual data, the proofing medium and the inputting medium are one and the same thing. Thus a typist using a typewriter equipped with optical scanning font and subject to no more format discipline than with a tape typewriter can create initial input for a computer. This copy can be removed from the typewriter and blue penciled by a proofreader for corrections. The copy then can be re-inserted into the typewriter and corrections typed at the bottom of the form. These corrections are coded to refer to some base point in the original input text or in an earlier correction line.

The final thesis of this paper is that rigid optical scanning systems reduce the complexity of the textual data inputting systems by creating an absolute one to one correspondence between the proof copy and the input medium.

Although the price of rigid optical scanners has decreased significantly in the last few years their cost still does not justify their use on small volume library applications. However, as it has been foreshadowed in the argument of this paper, the basic system for using paper tape equipped typewriters and optical font equipment are nearly the same. Because of this close relationship between the two inputting systems, we are experimenting with a data inputting system compatible with both methods.
The purpose of this paper is to examine the characteristics of file structure and search techniques in information retrieval systems. The main concentration will be upon the problems and possible solutions available for file structures and searching procedures on general purpose computers. The paper is given in terms of specifications that would be encountered if one were faced with the problem of designing a file and its associated search program for an application in information retrieval.

The first section concerns itself with the actual content of the file. That is, what specific pieces of information must be maintained in the file for each item. Two criteria are employed which jointly provide necessary and sufficient conditions for successful utilization of the file. The first of these is to determine the minimum information required for each item so that it might be matched to any question likely to be posed. Even though the particular pieces of information will vary from system to system, experience over the last few years with very large systems, as well as the folklore of traditional librarianship allows one to generate a sufficient minimum basis for all systems. The second criterion to be applied is that the items in the file must contain sufficient information in order to allow a given searcher or user to determine that a selected item is or is not relevant to his interests.

The second section of the paper concerns the relationship between the physical and logical organization of the file items themselves. The limitations and capabilities of existing general purpose computer systems effect the physical organizations that are possible. Alternative logical organizations are considered in the light of their effect on the economics of file manipulation. All known physical and logical organizations are described and a tentative selection is made of each for the purpose of the subsequent discussion.

The third section of the paper concerns an analysis of the various alternative procedures that may be used to match a given question to a given file item. Standard Boolean, associative, and probabilistic matching criteria are described and one new alternative is suggested. It is important to note that the matching criterion is independent of both organization and search strategy.

The fourth section of the paper concerns itself with the alternative search strategies that may be employed, given the selection of the file content, file organization, and matching procedure. It is seen that the selection of these prior alternatives, especially with respect to organization, severely constrains the available choice of search strategy. In the light of this constraint the problem of file organization is reconsidered and alternatives previously rejected are examined in more detail.

The final section of the paper deals with the options available for output of the search system. In particular, systems which include so-called on-line capabilities are discussed in detail. In this section it is argued that satisfactory
selection of output format and content is critically dependent on vague and relatively unknown psychological characteristics of the searcher and the user. Some suggestions are made for increasing our knowledge of users.

Selected References


An information system is no better than its final products. The characteristics of good output from a computer system include: simplicity, readability, compactness, completeness, reproducibility, economy, timeliness, and proper sequence. Output products may be characterized by form (printed, other visual, machine-readable), by size of audience (individuals, large groups), and by response time (immediate, rapid, delayed).

**Equipment Review.** A wide variety of devices are available for producing output from computer systems. Printers are the most common. Electric typewriters, the slowest kind of output printers, operate at speeds of approximately 10 characters per second. Electro-mechanical line printers (which operate at speeds of about 150 to 1500 lines per minute) fall into four categories: type-bar, type-wheel, wire-matrix, and rotating-drum. Chain printers are more flexible, since they permit expansion of the character set at some sacrifice of speed. Electrostatic printers are much faster, operating at speeds greater than 5,000 lines per minute. Photocomposition devices produce film copy of graphic arts quality which can be used directly for plate-making and subsequent publication. Other types of visual output include graph plotters and visible displays.

Some retrieval systems produce machine-readable output for use at other establishments on local equipment. Such output may take the form of punched cards, paper tape, magnetic tape, disk packs, or removable magnetic card packs. Other systems produce remote output by means of data communication links to console units. Finally, the actuator (a digital to analog converter) is an output device which enables a computer to control the operations of machinery and processes. So far, actuators have generally had industrial applications, but they may become important in the information field for linking digital and graphic retrieval systems.

**Systems Considerations.** Designers of an information retrieval system must have the answers to several basic questions in determining output requirements. Who needs the information? How quickly does he need it, and in what form? Why does he need the information? What does he do with it?

Output requirements directly affect the input to a system. The system designer must ensure that all needed data are entered into the computer. This always includes the substantive data to be printed in reports, and may also include control data, such as typographic symbols for photocomposition and special sorting elements for publications. Output requirements should also be considered in optimizing internal file construction for a system.

Special attention must be paid to designing an adequate character set for a published index with due consideration to problems of special characters and dia-critical marks used in various languages. Other important design considerations are the highlighting of "scan" elements; margin justification; and the use of multiple columns, running heads, and page numbers. The system designer must also weigh the advantages of direct access or index access to a bibliography and give careful thought to the frequency of cumulation.
In developing individualized output products from a retrieval system, flexibility of both format and sequence is desirable. One particularly useful technique is to list citation references in sequence according to a weighted scale based upon the numbers of "hits" of search terms and index terms.

There are special design considerations when providing machine-readable output for others. External users impose limitations on the parent system--i.e., a change to the system has multiple effects. Complete program documentation is essential, and training and technical liaison should be provided. Effective procedures must be developed in advance for maintenance of files released to field users.

Computer manufacturers provide "software," or generalized programming routines, which can be useful in developing output systems. Sort and merge routines linked with report generators can simplify the programming task considerably. Data processing compiler languages are particularly useful in facilitating changes to output records.

This paper presents a detailed review of output equipment for automated retrieval systems, and describes some of the more important elements of system design to be considered.

Bibliography


CONTRIBUTED ABSTRACTS
A SYSTEMS APPROACH TO DATA BANKS INCORPORATING BOTH MANAGERIAL AND TECHNICAL DATA. K. B. Andread, A. K. Dunlap, and M. R. St grotes, TRW/Space Technology Laboratories, One Space Park, Redondo Beach, California, USA.

An approach to the interface problems of managerial and technical data in future international data banks is presented. The impact of the scientific and medical data banks in purely managerial areas such as law, economics, and marketing is clearly evident today. The timely flow of information between the two fields is of the utmost importance. At the same time private proprietary rights and governmental interests must be preserved. A model of an international oceanographic data network containing managerial information is discussed in some detail. System-type flow diagrams are presented showing the interfacing of managerial and technical data as well as charts to show the hierarchy of levels of information in the network at international, national, and regional levels. The relationship between the oceanographic data bank and other data banks which are linked is discussed, as is the system such as weather information, food and agriculture, etc., are pointed out. A built-in statistical screening procedure for keeping data in the system current and relating other data to historical files as various projects phase out is discussed. Oceanographic research is on the verge of opening up vast new areas in mining, ocean-agriculture, law, marketing, etc., all of which add to the complexity of management problems.

A PHARMACEUTICAL DATA HANDLING SYSTEM FOR SCIENTIFIC AND MANAGEMENT REPORTING FUNCTIONS. Frances H. Arendell, Carol R. Englebor, Cynthia L. Godara, Dr. Lee K. Stenaker, W. Lambert Research Institute, Morris Plains, New Jersey 07950, U.S.A.

Large numbers of synthetic organic compounds must be screened by preliminary categorizing tests, and a few subjected to intensive animal study to produce one clinical candidate. This process results in large volumes of data, much of it negative, but most of it in consistent formats. Unfortunately, this data never dies, but must be re-searched for compounds having certain patterns of activity newly found to be a meaningful lead for useful clinical properties. The biological data can fall fairly readily into three types: 1. Animal dose-response data from fixed testing procedures, amenable to standardized reporting formats. 2. Scores of subjective observations for increased or decreased behavior, or blocked or potentiated challenge drug responses, for which score sheets can be used. 3. Evaluation data from non-routine testing procedures devised or modified to bring out the particular properties of a drug in advanced development, reported in technical reports. Basic tab card formats were designed to handle these three types of data in an integrated system using an IBM 039 sorter and 070 document writer to start. All data may be listed by compound, or each test reviewed for the most active compounds. Simple fixed programs may be used for fairly effective tabular displays of the data. VM or without Markush variants of chemical types. Quick search cards for screening status reviews or pattern searches can be generated from the master cards, using one column per test. A series of management reports were devised covering active compounds (permitting a pre-edit of the report by the scientific staff), quantities of tests run per month, project surveys, etc. We have proven the feasibility of our concept of a unified reporting system from the single input to the data store on a card handling basis over the past two years. The present size of the operation and the data store warrants conversion to magnetic tape.
ANALYSIS OF INFORMATION NEEDS OF INDIVIDUAL SCIENTISTS WORKING IN THE RESEARCH ORGANIZATIONS OF THE RUMANIAN ACADEMY. Dan Atanasiu and Dorin C. Lollancu, Scientific Documentation Centre of the Rumanian Academy, Gutenberg St., 3 bis, Bucharest, Rumania.

An analysis of methods utilized in investigating the users' needs and a review of worldwide results obtained up-to-date are made. Using the inquiry method by means of a questionnaire distributed to 500 Romanian academic research scientists, the authors establish the preferences, uses, and individual information gathering. A critical comparative study of similar data obtained abroad is done. The study intends to draw up conclusions concerning both the efficiency of the investigation methods for establishing the users' needs and some local specific aspects, the latter determining differences in information gathering behaviour.

CURRENT INFORMATION PUBLICATIONS. CONSIDERATIONS ON THE EXPERIENCE OF THE SCIENTIFIC DOCUMENTATION CENTRE OF THE RUMANIAN ACADEMY. Dr. Aurel Avramcu, member of the Rumanian Academy, Scientific Documentation Centre of the Rumanian Academy, Gutenberg St., 3 bis, Bucharest, Rumania.

The usual types of current information publications and their efficacy are compared. The paper points out the necessity of providing bibliographic references both with UDC numbers and descriptors, for unspecialized journals. In author's opinion, although classification and indexing require more time and delay of publication, they yield better informational value and wider possibilities for retrieval. Solutions adopted by the Rumanian Documentation Centre and measures taken for improving initial type of its current information publications are discussed. An increase of the descriptors number assigned for each paper and the utilization of automatic indexing methods are contemplated.
AN INTEGRATED EXPERIENTIAL TRAINING MODEL IN DOCUMENTATION, G. O. Batty, College of Librarianship, Aberystwyth, Wales.

To facilitate training in documentation the College of Librarianship, Wales is establishing an integrated set of manual and mechanized indexing systems all handling the same body of prepared information from a limited field. This constitutes a model to be used for training by students (including a simple approach, but not a complex one). The model's scope will include and modify subsequent stages in documentation activity. The material used as a basis for the model is the literature on information retrieval, because of the familiarity of the concepts and terminology, the limited size of the field and therefore also of the index languages, and the existence of Library Science Abstracts. Systems already in existence include a conventional classified catalogue using the CNG scheme for library science, a rotated classified file in visible index form, a thirteenth system, a manual co-ordinate system using random coding on edge-notched cards and a mechanized co-ordinate system using positional coding on body-notched cards. Additional or substitute systems proposed for a computer now under consideration include a version of the classified catalogue, a serial file and an inverted file based on the co-ordinate systems above, an IBM index and a "lattice" index.

Supplementary to these systems, which are dynamic in the sense that they will develop through use, is a static set of variant systems for demonstration only, e.g. Peek-a-boo as a variant on Thumter. The model is "primed" with 500 documents coded in each system for initial demonstration and use; indexing practices by the students will extend as the model develops. It is essential for all systems to be kept at the same stage of development for purposes of demonstration and comparison, though not all systems will be developed by the students. The model also offers experiential training in the preparation of documents (e.g. abstracting) and dissemination of information. It will also offer some training in SOI systems, and on later stages of its development may set itself as material for research.

ACCELERATING INFORMATION TRANSFER IN SCIENCE AND TECHNOLOGY, Charles L. Bernier, The National Library of Medicine, 8600 Rockville Pike, Bethesda, Maryland, U. S. A.

This is an analysis of: what can be done with existing technology; where resources are required; and what new research and development (R&D) in information science is needed to speed application of results to R&D. Between information generation and use there are many blockages and delays. Resistance to novelty the heart of successful R&D is a block. Another block is indicated by delays between reading and application. Delays between results and their communication; between writing and publication; and between publication and reading, extracting, abstracting, indexing, reviewing, teaching and learning. Delays in printing of R&D results invite comparison with newspaper accomplishments. Delays in reading relate to competing diversions of a complex civilization including the increased complexity of R&D itself. New techniques for selecting, condensing, and presenting information (including data) are needed to make salient facts interesting. The potential user needs more attractive leading into reading. New selectors for documents and data are required as well as more extensive use of existing selectors. Measurement of reading and scanning times for different formats are urged. Better abstracts are considered to enable the scanner to select more accurately and rapidly. Inadequate support of information services and systems causes delays. Support includes decision to act, selling necessary ideas to resource sources, and resources (dollars, position allocations, and personnel) to carry out the decision. Backlogs of papers need resources for publication. New periodicals, evaluative reviews, journals of indexed abstracts, and data compilation programs are needed. Techniques for producing these are discussed. Full application of existing techniques through improved resources can contribute to promptness of communication.
A discussion of some indexing problems. Massa Bloomfield and 
Dell Schaefer, Hughes Aircraft Company, Culver City, California 
U.S.A.

A brief study was made of the indexing approach of several different indexing-abstracting publications. In the sample taken, five different articles were selected. Each article is indexed in Chemical Abstracts, Nuclear Science Abstracts, Engineering Index, ASTHA Technical Abstract Bulletin, Physics Abstracts, and in a KWIC (Keyword-in-context) system, manually prepared. The derived results are discussed and graphically displayed. Widely different patterns of indexing were observed, indicating either a lack of effective indexing rules, or that each indexing journal not only serves a different public, but that each public has little in common with the other. One of the tables developed indicates in detail the semantic apparatus associated with each set of indexing terms. From this, the individual patterns established by each journal may be ascertained. It is concluded that different systems will not only arrive at different basic vocabularies, but that these vocabularies will express different meanings in each system. The effect of the indexing system on the vocabulary chosen to describe the information content should be investigated further.

CUE INDEXING SYSTEM (CIS). Eric H. Boehm, American Bibliographical Center, 800 East Micheltorena Street, Santa Barbara, California, U.S.A.

The Cue Indexing System (CIS) has a dual function: 1) the establishment of a new knowledge classification system which is entirely alphabetical and has mnemonic capabilities greater than that of Dewey System, the Universal Decimal Classification or the Library of Congress System, and 2) creation of a new type of index which applies computer capabilities to print specific subject facets -"the cues" - with each index entry. The main cues are 3-letter abbreviations, such as FOR for "foreign relations," or acronyms such as CMP for "communications mass media or publishing." The index carries the cues with each number entry. It is conventional in other respects: the headings consist of a complete text and they have subheadings and third-order headings when required. A computer is used for permutation of the cues for printout and later cumulation, information retrieval, and SDI (selective dissemination of information). The first use of CIS will be for the Index Number of Volume I of AMERICA: HISTORY AND LIFE (1965). Some specimen lines:

<table>
<thead>
<tr>
<th>Abstract Cues Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAW</td>
</tr>
<tr>
<td>501 CMP: Press SLA REL 1710</td>
</tr>
<tr>
<td>218 ECO:Ind. 1952-45</td>
</tr>
<tr>
<td>109 ECO:Ind.: (steel) LAB 1933</td>
</tr>
<tr>
<td>17 SLA REL, POL 1800</td>
</tr>
</tbody>
</table>

CIS has general applicability to the special needs of the social sciences and humanities, just as the KWIC index has had general applicability to the current awareness needs of scientific research. CIS has these characteristics or advantages: 1) the cues are selected by indexes on the basis of an authority list, 2) the printed text is much less than KWIC, 3) CIS can be scanned more quickly than KWIC, and 4) CIS is not so bulky as to make multi-volume cumulations prohibitive expensive.

DEVELOPMENT OF AN ACADEMIC CURSE ON THE ENVIRONMENT OF TECHNICAL DOCUMENTATION SYSTEMS. Jon W. Bohnert, The American University, Washington, D.C., U.S.A.

The course is based on the assumption that technical documentation systems are best understood as a part of the total communication process as technically trained personnel. Documentation is only one of the regular means of communication in organized research and development work. If the function of documentation systems is to unite members of a specific group of technical users with appropriate documents at suitable times and places, then indexes and evaluations of documentary services and techniques can best be made in the context of the particular work environment. Study of the environment, both of the technical users to be served and of the documents appropriate and available, will mainly determine what documentary services are practicable, and what few documentary techniques are applicable. An American University course, "Technical Information," was first given from this viewpoint in the Spring Session, 1965. Three main kinds of users of technical documentation were distinguished: en, management, operating, and research and development personnel. The considerable documentation experience of students in the nation's capital benefited class discussions and resulted in individual student projects evaluating the effects of a particular work environment on certain documentary activities. One of the main difficulties encountered was in securing adequate accounts, verbal or written, of the various types of work situations that condition existing documentation systems.

DEVELOPING AN INFORMATION SYSTEM FOR MENTAL HEALTH: A PROGRESS REPORT. Lorraine Bouthilet and Julius From. National Institute of Mental Health, Bethesda, Maryland.

The National Clearinghouse for Mental Health Information (NCMHI), an information evaluation center, is developing a mechanized information storage and retrieval system as part of its total system. At present, an interim system has been developed for the IBM 1401 type computer. Index-code sheets as well as thensauri have been prepared by techniques which call for a combination of extraction and coding. Other documents were themselves as well as obtaining the assistance of experts in the field. Terminology has been developed in such areas as psychopharmacology, crime and delinquency, mental retardation, drug abuse, and occupational mental health. The indexing system is a coordinate indexing system with provision made for the elimination of false drops inherent in such a system. The retrieval system is based on a combination of inverted and serial files, using the best features of each. The computing system is modular in that it can search both specific areas as well as a complete master file. Over a thousand computer searches have been made on a file of over 25,000 documents, which is growing at a rate of 2000 documents a month. The computer search print out contains both the citation of the document as well as a complete abstract. The computer has generated reports in the field of mental health as well as other reports concerning a comprehensive survey of the field of mental health information. Upon analysis of ten months of operation of our indexing system, the interrelationships of language, automatic encoding and formatting, the use of a random access computer, suggests themselves as improvements.
A COMPUTER-GENERATED INDEX PUBLISHING SYSTEM.  

Bruce Carter, Barbara Shaffer and Diana DeWitt, American Society for Metals, Metals Park, Ohio, U. S. A. 44073

The computer-generated published index system, a joint design and developmental effort of the American Society for Metals (ASM), Engineering Index (EII) and the International Business Machine Corporation (IBM) is a generalized and multipurpose program with a wide range of applications and subjects. Designed as a component of the IBM 1401 Combined File Search System (CFSS), it will be included as a module of the System/360 version of the 1401 Information Storage and Retrieval System. The publishing system program arranges all the entries-subject headings, titles, cross-references, authors-in an alphabetical sequence, automatically inserts cross-references and prints the index. Abstracts are also handled in this system, and an entire volume of indexes and abstracts, such as ASM's REVIEW of METAL LITERATURE or EI's PLASTICS OR ELECTRONICS sections, can be produced as a part of the computer operation. There are five programs comprising the publishing system, the first stores the abstracts, the second does initial formatting, the third generates cross-references and sorts the index, the fourth sets up page and line format, and the final program produces copy ready for photoreduction. Both the basic format of the CFSS package and the design of the ASM-produced Publishing System allow for great flexibility of index type and content. In addition to the present subject heading and author indexes, many variations are envisioned, including bibliographic citation indexes, listing of notations on content rather than titles and classification by journal. The development, function, utilization, flexibility and description of the publishing system programs are discussed. Detailed program descriptions and system flowcharts are included.
OPTIMIZING RETRIEVAL RESULTS WITH MAN/MACHINE INTERACTION. Robert M. Cartton, Victor Rosenberg, Lehigh University, Bethlehem, Pa., U.S.A.

It has been repeatedly shown that document retrieval is as much an intellectual process as a mechanical one. By combining the intellect of the user with the files searching and display capabilities of the computer, an associative retrieval strategy has been designed which enables the user personally to navigate the search within the parameters of his own interests and the contents of the collection. The system eliminates the need for a thesaurus and may be used with most types of indexing.

The user approaches the system with a request containing at least one term used as an index term in the collection. The true or terms are coordinated and the file searched. A list of associated terms (terms occurring together in documents) is then displayed to the user along with the number of documents initially retrieved. The term associations are determined statistically. On the basis of the number of documents initially retrieved, the user may either expand or narrow his search. From the list of terms the user selects those which seem to suit his needs and which will yield a reasonable number of documents, based on the predetermined maximum number of resulting documents. The additional selected terms are then coordinated in a manner which depends on the choice to expand or narrow the search. The user may repeat the above process with the additional terms as many times as necessary. When the user elects to conclude the search, an expanded list of documents is produced. The above type of retrieval strategy is particularly applicable to existing collections and could be optionally implemented on time-sharing computer configurations. Results obtained to date on a 1,000 document collection have been particularly encouraging from the point of view of user satisfaction as well as objectivity within systems criteria such as computer running time and number of documents resulting from a given search.

LA DIFFUSION DES PUBLICATIONS AMERICAINES EN EUROPE AU XIXE SIECLE PAR LES EXCHANGES. Julie D. Brant, Service des publications internationales, 16 rue des Deux-Eglises, Brux., Belgique.


La documentation échangée à travers le monde dépasse les 300,000 volumes, enrichissement extraordinaire pour l'Europe.
Le traitement des documents contenant des informations textuelles est un domaine très actif de recherche. Les avancées technologiques ont permis de développer des outils qui permettent de traiter de grandes quantités de données de manière plus efficace. Cependant, les défis restent nombreux, notamment en ce qui concerne la reconnaissance et l'interprétation des textes manuscrits ou manuscrits.

Pour être efficace, le traitement du texte doit être capable de gérer une grande variété de formats et de préserver la structure du document source. Il est également important de garantir la fiabilité et la précision des résultats, en particulier dans les domaines où la qualité du traitement du texte peut avoir un impact significatif sur le processus de communication, comme dans le domaine juridique ou médical.

La recherche dans ce domaine est active et progresse rapidement. De nouveaux algorithmes et techniques sont continuellement développés pour améliorer la performance des systèmes de traitement du texte. Ces développements permettent de traiter des documents de plus en plus complexes et de plus grande taille, ce qui ouvre de nouvelles perspectives pour l'automatisation de processus de traitement du texte.

Les avancées technologiques dans le traitement du texte sont également importantes pour l'amélioration de l'efficacité et de la productivité dans des domaines tels que l'éducation et le travail administratif. Les systèmes de traitement du texte permettent de gérer de manière automatisée les tâches répétitives et de libérer du temps pour des tâches plus complexes ou créatives.

En conclusion, le traitement du texte est un domaine en constante évolution, où les chercheurs continuent de dénouer de nouveaux défis et de développer de nouvelles techniques pour améliorer la qualité et l'efficacité des systèmes de traitement du texte.
CONTRIBUTED ABSTRACTS


This is a study of the development of librarianship, focusing on the role of libraries in the education of scientists. The study examines the evolution of the relationship between libraries and the scientific community, highlighting the changing nature of information needs and the development of new technologies and methods for managing and disseminating information.


This paper discusses the development of information and data retrieval systems specifically for analytical chemists. The author describes the needs of the analytical chemist and reviews various systems designed to meet these needs, including those developed by the I-13 Subcommittee of American Society for Testing and Materials on Absorption Spectroscopy, and Proton Technic Abstract Services for gas chromatographic data and nuclear magnetic resonance data.


This paper explores the theory of notification, focusing on the role of libraries in disseminating information. The author discusses the development of notification theory, emphasizing the importance of library services in supporting the needs of scientists and researchers. The paper also addresses the evolution of library services and the impact of new technologies on notification practices.

For a full understanding of the development and impact of library services in the field of information and data retrieval, you should consult the original sources provided.
INDEXER REQUIREMENTS FOR THE RECOGNITION OF SCIENTIFIC CONTENT AND CONTEXT.

W. B. Foster, Associate Director, Life Sciences Division and David F. Hervey, Deputy Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N.W., Washington, D.C., U.S.A.

It is essential for the sake of economy that an information system use the least expensive personnel consistent with effective retrieval. The Science Information Exchange has for 15 years provided information concerning current research to both scientist administrators and individual investigators, with emphasis on recognition and preservation of scientific context and current research concepts. An experiment was designed to determine the extent to which subprofessional personnel with various levels of educational background can handle different types of subject content in the indexing process, with respect to efficiency as well as quality. Using a standard of correct index points determined by senior professionals (Branch Chiefs) in the light of established working definitions of topics, subprofessional personnel were able to index an average of 30% of the possible number of correct index points, as compared with 80% for the earlier professional analysis of the same projects. Subprofessional personnel took longer per topic indexed, making the estimated cost about the same for both groups. The correct index points which were missed by the subprofessional group were much more likely to be major or critical ones than in the case of the original analysts. There is a strong positive correlation by rank order between educational level and performance. It should not be concluded, however, that subprofessional personnel should not index scientific research content, all of the test group did relatively well on "Test" types of indexes, such as chemical substances and biological organisms. Many other types of categories were successfully handled by the test group, depending primarily upon level of educational background and experience.

THE USE OF CURRENT SCIENTIFIC RESEARCH INFORMATION FOR ADMINISTRATIVE PURPOSES. Morris E. Freeman, Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N.W., Washington, D.C., U.S.A.

Information concerning who is doing what in current research is useful to agencies administering grants programs, and the Science Information Exchange of the Smithsonian Institution is primarily oriented to this service. Other directly related purposes of the Exchange include the provision of research information to laboratories sponsored by Federal grants and contracts, and to individual recipients of these grants.

It is of continuing interest to the Exchange to evaluate the types of scientific subject matter information most often called for by these three groups of users. Data collected to date suggest that the first class of users requires the full spectrum of S.I.E. facilities, both general and detailed information on subject questions, usually associated with some administrative parameters as well. The need for broad surveys of subject areas is greater for this group, as might be expected; there was a surprising frequency of need for detailed breakdowns of these broad areas on the part of grant administrators. The second and third classes of users require primarily the more specific types of subject matter coverage, although many programs in the second class do require rather generic level searches.

More recent and complete data will be presented on the points just discussed, and the relationship of services to the individual scientist to that of a more administrative orientation will be considered.

EMERGING PATTERNS OF NATIONAL INFORMATION SERVICES.

Bernard M. Fry, Clearinghouse for Federal Scientific and Technical Information, 3255 Port Royal Road, Springfield, Virginia, 22151, U.S.A.

National technical information requirements and the principal problems and issues involved in fulfilling them are discussed and recommendations made for steps toward their solution. The following central problems are examined in relation to their impact upon the structure and effectiveness of national information services in science and technology: 1. The disparate though inter-related roles of data, information analysis, and information dissemination centers--together with research libraries--considered in terms of their distinct functions and their contributions to advanced information systems. 2. Achievement of network capability utilizing special information sources both government and non-Government. 3. Development of comprehensive and non-duplicative information centers performing indexing and abstracting services for the major disciplinary fields. 4. Establishment of criteria for centralized and decentralized services where appropriate, based on effectiveness, efficiency and economy. 5. Provision of information services in various forms adapted to the needs and interests of particular scientific, technical and industrial groups. 6. Effective application of scientific and technical information to economic growth, with particular reference to the transfer process. 7. Ways and means to obtain user feedback and to measure the effectiveness of information services.

A NEW TYPE OF CONCERN IN TERMS OF DOCUMENTATION:


The Documentation Unit of Ispra's Euratom Technical Services takes care of a very peculiar kind of documents which consists of catalogues, pamphlets, leaflets and research studies sent by research or production firms for the use of Scientists. This scientific and Technico-Commercial Documentation became so essential for the use of all technicians within Ispra's European Atomic Energy Commission Research Center, that a special classification has been worked out and data is stored by Flexowriter for the mechanization of the system through IBM 1401.
CONTRIBUTED ABSTRACTS

FONCTIONS ET ACTIVITES DU GROUPE D'ETUDE SUR L'INFORMATION SCIENTIFIQUE : "OCOCUMILATON DE LA DOCUMENTATION". Jacques Gérard, Section d'Automatique Documentaire, C.N.R.S., Marseille, France.

Le Groupe d'Etude sur l'Information Scientifique a été créé en 1963 par la délégation générale à la Recherche Scientifique et Technique. Ce Groupe est chargé d'établir un "Inventaire des études et des applications en cours intéressant le traitement automatique de l'information scientifique, principalement la documentation et la traduction automatique". Cet inventaire se matérialise sous la forme d'un fichier analytique établi à partir de la documentation technique non publique et des revues spécialisées dans les domaines en question. A ce jour, 5000 documents ont été dépouillés (3 dater de 1963). Le fichier comprend d'une part un ensemble de rubriques correspondant aux différents aspects du domaine, et d'autre part une série de subdivisions dites "annexes", concernant les chercheurs, les Centres de Documentation, les Congrès, les Équipements, etc. Un "Groupe de Scecteurs" permet l'analyse de ces textes et des différentes chapitres donnent le plan du fichier. Le lexique, bilingue pour l'anglais et français, sera quadrilingue dans sa version définitive (russe et allemand). Augurant, l'on espère parvenir, avec la collaboration d'organismes étrangers homologues à une uniformisation des principales notions permettant d'incliner la "Documentation de la Documentation". Enfin le Groupe a mis au point une enquête sur les 'Moyens Nouveaux de Documentation Scientifique et Technique en France, enquête équivalente aux "Systèmes non-conventionnels d'information technique actuellement en usage" publiée par l'Office de Service Information Service de la F.I.D. L'ensemble des travaux du groupe permet ainsi la fabrication d'ouvrages de références, de bibliographies spécialisées à partir du fichier analytique, en même temps qu'il va dans le sens des recommandations de la F.I.D. en matière de coopération internationale en Information Scientifique.

DESIGN DECISIONS FOR AN INTEGRATED LIBRARY SYSTEM. Marjorie Greif, International Business Machines Corporation, Advanced Systems Development Division, Los Gatos Laboratory, Los Gatos, California, U.S.A.

The library's responsibility in documentation begins with the entry of the bibliographic information which is the source of all subsequent library records for each document acquired. The steps from acquisition to circulation are now mechanized in many libraries, but seldom as part of an integrated system. Such a system has been developed in an experimental technical library ideally situated: it is a working library for an IBM laboratory and serves as a testing ground for advanced data processing equipment and programs. To keep up with the increasing information supply and demand, this library was moved from piecemeal mechanization to an integrated system which now provides mechanized bibliographic control and will eventually include mechanized information searching and retrieval. Design decisions for this system were based on an analysis of the effectiveness of existing procedures as evaluated by both the library staff and library users. The basic question was not "How can we mechanize this operation?" but "Why do we do this...and why do it this way?" The success of the system now implemented depends directly on the thoroughness of that questioning. As the answers disclosed duplicate effort, avoidable sources of error, and unverified needs, the system design was modified as permitted by the equipment to be employed. In operation, a single machine-readable entry establishes bibliographic control when a document is ordered, and on-line processing maintains this control, automatically accommodating corrections and revisions as needed. The same input, printed out selectively, provides purchase orders and subject catalogs. With the flexibility demonstrated and the unexpected value of unplanned by-products, this integrated system seems to successfully anticipate both the future needs of the library and the possibilities of present data processing technology as it applies to the phases of documentation which are concentrated in the library.

A RETRIEVAL ALGORITHM PERMITTING RANKING OF DOCUMENTS RETRIEVED BY DEGREE OF PERTINENCE TO THE USER. Mary A. Hawes, UNIVAC Div. of Sperry Rand Corp., 650 M. Sepulveda Blvd., El Segundo, Calif., U.S.A.

A Retrieval Algorithm has been developed to meet the requirements for an on-line document retrieval system that can be expanded to an information retrieval system. These requirements are: (1) relatively complex requests can be formulated in a straightforward manner without necessitating assistance from a retrieval specialist, (2) the language and the rules used to specify the criteria permit the user to state both minimal requirements and maximum pertinence, (3) documents retrieved can be ranked by degree of pertinence to the user, (4) the request statement is suitable for direct entry into a computer system, (5) the request statement can be verified as being processable while the requestor is on-line to the computer system, (6) procedures and machine coding for processing the request can be generated dynamically, (7) procedures allow processing of multiple requests in a parallel operation, (8) techniques used for demand requests are suitable for information dissemination, (9) the retrieval algorithm can be expanded for information retrieval.

In order to meet the above requirements, it was necessary to develop some new techniques. Among these is Information Grouping Logic (IGL), an extension of Boolean logic incorporating the minimal requirement concept, a problem oriented language for stating the retrieval request, and a diagrammatic representation of the request. Decision Tables are used to define the algorithm which generates the procedures for processing the request and ranking the documents retrieved by degree of pertinence to the user. A number of examples are included.
CATALOGING AND RETRIEVAL STATUS OF STATE EXPERIMENT STATION TECHNICAL BULLETINS AT INSTITUTIONAL LIBRARIES, John E. Hogue

The purpose of this study was to determine the extent to which technical bulletins are cataloged and retrievable from institutional libraries and to compare the retrieval status of bulletin collections held in governmental and non-governmental repositories. The study involved an exploratory survey of six government and seven non-governmental libraries.

The six government libraries were:
1. Iowa State University
2. University of California, Davis
3. University of California, Berkeley
4. University of California, Los Angeles
5. University of California, San Diego
6. University of California, Riverside

The seven non-governmental libraries were:
1. University of Washington, Seattle
2. Stanford University, Stanford
3. University of Oregon, Eugene
4. Oregon State University, Corvallis
5. University of Oregon, Portland
6. University of Southern California, Los Angeles
7. University of California, Irvine

The study found that:
- The majority of technical bulletins are not cataloged in government libraries but are more likely to be cataloged in non-governmental libraries.
- Retrieval status varies significantly between government and non-governmental libraries.
- The use of automated systems for cataloging and retrieval is more prevalent in non-governmental libraries.
- The adoption of consistent indexing practices is more widespread in non-governmental libraries.

CONTRIBUTED ABSTRACTS

TOWARDS AN INTERNATIONAL SOCIAL WELFARE DOCUMENTS SYSTEM, Joe B. Wofford

Documenting and exchanging knowledge is an inseparable part of planning, research and development in any field. Social welfare is no exception. The problem of creating improved methods for searching and retrieving information is difficult because of the nature and complexity of social welfare. Social welfare is defined as "field"...
INDEXER CONSISTENCY TESTS - ORIGIN, MEASUREMENTS, RESULTS AND UTILIZATION. Robert S. Hooper, International Business Machines, Federal Systems Division, 7220 Wisconsin Ave., Bethesda, Maryland, U.S.A.

The efficiency of a document reference retrieval system may be measured in terms of the success in retrieving documents. This success is often measured in terms of "recall" and "precision" ratios. Efforts are extended to make these ratios as large as practical to meet the needs of a real system. Recall ratios below the 100% mark are often attributable to indexing failures. The nature of indexing failures can be determined by studying why documents are not recalled, if it can be assumed that the collection is sufficiently large and sufficiently well-known to enable the measurement of recall, and if enough search requests are available for study. Another method of determining these failures, is to perform a controlled indexer consistency test where indexers will, each in turn, index a significant number of "live" documents. By analyzing the reports of indexer consistency tests, it has been found that there is no standard measure of consistency; hence, a set of three measures has been derived and is used to analyze the data in these reports. The analysis indicates that a range of consistency values between 10% and 80% may be expected, depending upon the indexing parameters. Additional analysis shows that uniquely structured classification schedules and indexer training programs are the indexing parameters most often associated with the higher consistency results. Further, the analysis suggests that consistency tests have an important role as a tool in the design of indexing systems and in the training of indexers and quality control of the indexing.

SUBJECT INDEXING USING A COMPUTER-MANIPULATED THESAURUS AS VOCABULARY CONTROL. Marjorie R. Hystop, American Society for Metals, Vnetels Park, Ohio, 44073, U. S. A.

In redesigning the information retrieval system of the American Society for Metals, it was decided that the capability should be included to automatically generate a subject index to the Review of Metal literature (abstract journal). It should be generated as a computer product of the over-all information retrieval system, and based on a thesaurus vocabulary. Previously the index was manually prepared utilizing a modified classification scheme as authority. The first step was to compare the two types of vocabularies. For test purposes the "EIC Thesaurus of Engineering Terms" was used, recognizing that it is more restricted than will be the eventual ASM Thesaurus. Nevertheless, comparative tests on 100 abstracts showed that 83% of the indexing concepts could be intellectually duplicated with the thesaurus vocabulary. Tests for indexing consistency were performed by comparing entries submitted by 9 indexers for the same body of 25 abstracts. Indexing rules were devised based on these tests and also on the thesaurus structure and the computer processing capabilities. One of the important features of the new method is the merging of two intellectual indexing functions previously performed separately on the same document--deep indexing for machine retrieval and shallower indexing for publication. Other features are (1) the provision of two levels of indexing, (2) the automatic generation of cross references, (3) the provision of monthly indexes in addition to a cumulated annual index, and (4) the inclusion of titles with index entries.
EXCEPtS to MEASURING TOOL FoR INFORMATION SYSTEMS. Jay


CONTRIBUTED ABSTRACTS


WADEX (86th ed.) is an IBM (IBM) prepared mechanical index, extension of KWIC Concept. This paper describes greatly improved third version where intelligent effort ranging from zero to considerable can be employed depending on titles and purpose. WADEX system usable for browsing, searching, express information, also for preparation of specialized deep indexes of specialized holdings. First experimental WADEX (1) and second experimental WADEX (2) are based on 8,000 titles of APPLIED MECHANICS REVIEWS (AMR), an international critical review magazine. In WADEX titles printed fully with author names, and as many lines used as necessary (50 characters per line). WADEX entry: word in title (except forbidden word) or author’s name.

Alphabetically sequenced entries printed out of context. Words are single words or hyphenated word pairs or “Tagged Words” (TW). Letter consists of significant “word compounds” in which one or more words are forbidden words. TW selected from subject index authority list of AMR and hyphenated words from frequency list of Words of previous WADEX. For more than one identical word entries, titles alphabetized according to first author and for more than one identical author entries according to reference number, which printed in line with entry. Multiple entries listed only once per column. WADEX arranged in 2 columns, tabulation, columns printed simultaneously with dictionary entries at top and pagination at bottom by computer. After addition of 26 alphabets “Compuserch” is the photoready copy with 6.5 average entry per paper including 1.5 for authors. Published WADEXES based on available extended system for titles with intellectually added keywords, all individually in brackets. Documentation of machine program now available. APDR, NSF, & ONR are sponsors of WADEX and/or AMR.


OPERATION OF THE AIR FORCE MACHINABILITY DATA CENTER. John F. Kahles, Air Force Machinability Data Center (AFMDC), 3980 Rosslyn Drive, Cincinnati, Ohio, USA.

This paper outlines the basic design concepts leading to the development of AFMDC, reflects procedural changes made since its design was completed, and discusses actual operation of the Center since it was started on 1 Oct. 1964. Strong economic justification for establishment of AFMDC is indicated in that more than $34 billion are spent annually for labor and overhead in the metalworking industries alone in the United States. Other significant economic data include expenditures for machine tools amounting to more than $1 billion annually plus $787 million for machine tool accessories, including small expendable cutting tools. Machining information is stored on a set of 10 punch cards, including general information as well as specific numerical machining data such as speeds, feeds, depth of cut, tool material and geometry, cutting fluids, and other significant variables.

The entire system is based on the “machining situation” concept. A machining situation is defined when a particular material, with definite chemical, physical, and mechanical properties, is being processed by a specific material removal operation, either a conventional or nonconventional process, and it is the focal concept for acquisition, interrogation, or presentation of information. This paper shows how forward-looking design concepts can anticipate updating of a system, such as conversion from punch card storage to computer. Computer programs have been developed and demonstrated for converting coded machining data to printout in machinist’s language. The paper summarizes the inquiry experience with more than 300 specific inquirers and discusses procedures used in development and maintenance of a User File.

PROGRAM OF THE CENTER FOR INTERNATIONAL BIOMEDICAL COMMUNICATIONS RESEARCH. Allan Kent, Director, Knowledge Availability Systems Center, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, U.S.A.; Pierre J. Vinken, M.D., Executive Chief Editor, Excerpta Medica Foundation, Amsterdam, Holland.

Research aimed at studying the most effective means of transmitting new medical information from research to the practicing physician will be conducted by the University of Pittsburgh-Excerpta Medica Foundation Center for International Biomedical Communications Research. The Center will investigate techniques of information retrieval, publishing and libraries, especially, as well as computers and other hardware elements in an attempt to develop an integrated communications system which will provide the physician with medical knowledge while it is new and vital. The joint University of Pittsburgh-Excerpta Medica Foundation project will utilize, as the chief source of research material, the Foundation’s twenty-year collection of about 1,300,000 abstracts, indexes to the Excerpta Medica specialized abstracting journals, and the microfilm library in which original articles from over 3,000 biomedical serials are stored. The initial research will be concentrated in three areas: a selective dissemination program for physicians; a study of the problems relating to medical terminology; and an investigation of the feasibility of applying automatic indexing techniques to medical literature.
CITATIONS IN SOVIET METALLURGICAL LITERATURE: Ladislav Rehacek, Institute for Technical and Economic Information, Naka 137, Prague 1, Czechoslovakia.

The present study examines the citations given 20 years ago in the three Soviet journals "Stal", "Metallovedeniye i keramika", and "Fizika metallov i metallocv-dereni" which deal with the problems of metallurgy, heat treatment, and inertial physics. The paper studies the cited documents from the viewpoint of type of document, language, country of origin, and age. The results were tabulated and reviewed. The most cited journals were tabulated both for all three source journals together and for each source journal separately; the results descnribed in the present paper were compared with the Burton's results. The cited documents were divided into two groups, up to ten years old and older than ten years. Most cited were the documents in Russian, further in English, German, and French. The documents in Russian, English, and German were cited in 97.0% of source journals. The results described in the present paper were compared with the Burton's results. The cited documents were divided into two groups; up to ten years old and older than ten years. Most cited were the documents in Russian, further in English, German, and French. The documents in Russian, English, and German were cited in 97.0% of source journals. In journal citations not older than ten years Russian participants with 59.65% English about 25%, German at the cost 8.124 and French at the most - 2.5%. In comparison with older journal citations Russian and English increased their share whereas the percentage of American and French journal citations dropped. 50% of cited papers from the last ten years appeared in 3.6-4.4% of the total number of journals cited. The first 30 journals most cited in all the source journals (not older than ten years) include 16 most cited journals in all source journals separately. From 19 most cited Western journals 16 titles can be found among the 20 most cited journals according to Burton's survey.


The CAS computer-based substructure search system retrieves information about compounds on the basis of structural characteristics. This paper describes the computer techniques used in the present system. These techniques range from the use of simple screens to iterative queries. One important phase under development employs set reduction to decrease the amount of unproductive searching.

The more emphasis of the paper is on searches by means of which inappropriate compounds are rejected, promptly and without iterative search. Criteria for evaluating searches are based on a predicted question pattern and include development costs, run times, and per cent success. Other factors discussed are the effect of "true" file organization, the value of nested machine records, and the relative merits of special purpose searches versus blanket searches.

Pilot results are presented and are projected to a full scale system. Present and proposed output formats are described, and the importance of interaction with users and potential users is stressed. This work is partially supported by funds from the Department of Defense, the National Institutes of Health, and the National Science Foundation through a contract with the National Science Foundation.

P 1 C 8: THE PHARMACEUTICAL INFORMATION CONTROL SYSTEM OF MERCK SHARP & DODGE RESEARCH LABORATORIES, Margaret C. Koeb, Jerome F. Maddock, and Barbara E. Weaver, Merck Sharp & Dodge Research Laboratories, Rahway, N.J., U.S.A.

The Pharmaceutical Information Control System (P 1 C 8) developed at Merck Sharp & Dodge Research Laboratories provides central control and methodology for a series of decentralized information areas in the Division. It is compatible with and instrumental in total data processing and analysis of research information. Serving as a Core Index to all information resources of the Research Laboratories, it also processes, stores, and retrieves research project information for the staff members for planning and retrospective searches. A register of all domestic and international clinical research information on experimental and in-line products of Merck & Co Inc., is provided by this system.

An eight digit dual-faced classification code was developed based on a company-wide program identification scheme. This code of two mutually exclusive facets enables us to identify a product with a field of research. This code has been adopted for use in administrative planning, cost accounting, time allocation, and internal reporting. The uniform use of this information code within the Research Division minimizes the vocabulary barrier between the use and information system and provides the system with a self-indexing device for internal reporting.

Incoming mail is copied and registered by the information center prior to transmittal to the addressers. Copies of all outgoing mail and internal correspondence are directed to the center. An information scientist analyzes each document and selects the project code and document descriptors. This information is punched into 80-column cards. The documents are filed by code and the cards are filed alphabetically by name, term and by date. Output forms include information, documents, printouts of document citations, and reports to drug regulatory agencies. Continuous system evaluation results in reduced 1/0 time, better utilization of personnel, and improved user feedback and contact.

A RETRIEVAL PROFILE FOR CURRENT RESEARCH INFORMATION, Frank J. Kley, Associate Director, Physical Sciences Division and W. R. Foster, Associate Director, Life Sciences Division, Science Information Exchange - Smithsonian Institution, 1730 M Street, N.W., Washington, D.C., U.S.A.

The Science Information Exchange (S.I.E.) provides information concerning details of ongoing research in the sciences to Federal and private agencies engaged in or supporting such research, as well as to individual investigators working on specific problems. Emphasis is placed on professional level retrieval of current research concepts. To assess the adequacy of such retrieval, it was desirable to have an evaluation of the retrieval profile in routine use. In recognition of the subjective nature of the determination of "relevance" of an answer, an experiment was designed in which agreement was obtained on a specific subject question with a technically trained representative of a Federal agency whose program is included. In this way the recall of the S.I.E. system with respect to that agency's projects was assessed along with the relevance determination. Results to date indicate that under the conditions of the S.I.E. operation, with the same professional scientists doing both indexing and retrieval, reviewing all outgoing subject material, it is possible to have a high level of both recall and relevance. Index structure is an important contributory factor to this result, in that relevant concepts are not scattered alphabetically but are grouped by and for the use of scientist-indexers.
CONTRIBUTED ABSTRACTS

WALTER PERSONAL INFORMATION STORAGE AND RETRIEVAL SYSTEMS. 1100 1st Street, N.W., Washington, D.C.

By. Vancouver Bush compiled a amazing set of predictions in his article, "As We May Think" in the July 1945 Atlantic Monthly. He predicted a day where computer-driven highly individualized, personalized, organized, and retrievable systems an extension of one's memory; therefore, he dubbed it "Memos." All of his predictions are essentially within the state of the art today. Yet we have no Memos. However, more and more people are requiring better and better retrieval systems for personal and professional use. This is happening while simultaneously the amount of information to be stored is mounting at an ever increasing rate.

Large institutionalized document storage and retrieval systems are growing rapidly. Many are computer-driven. Humans seem to be considered as peripheral to these systems. Consideration should be given to increasing the efficiency as well as reducing the cost of personal oriented information storage and retrieval systems. The recent development of multiple access time sharing systems like Project MAC and implicit programming show promise in this area.

It is proposed that an all out effort be instituted in terms of equipment, program languages and classification schemes to make available to a large number of individuals in the next decade an updated version of Memos. The problem is not one of gigantic strides to be made in the state of the art as much as it is a reorientation of our human efforts to accomplish this goal.

ORGANISATION ET CONSULTATION D'UN THESRAUUS.

Lorsque l'on interroge un fichier d'indices organisé à l'aide d'un système de mots-clés ou de termes, il est en général nécessaire de consulter une mise en réseau de thésaurus permettant de déterminer les termes ou mots-sujets satisfaisants pour définir la demande de documents. La recherche de ces termes revient à déterminer la validité du mot ou du terme qui constitue l'environnement déterminant d'une notion. Cet environnement peut être automatiquement calculé en tenant compte des rubriques qui forment une notion ou des sections dont nos notions sont d'autant plus voisines que leurs dénominations sont plus semblables. Il est donc possible de consulter un thésaurus et d'obtenir un fichier de documentations lexico-grammaticaux concernant les mots-clés. Chaque mot-clé est indexé à l'aide de dictionnaires qui permettent la sélection automatique de docum-ventation de l'environnement. L'avantage essentiel de cette méthode est de permettre une construction facile du thésaurus, puisque les relations sémantiques sont déterminées automatiquement. Il est d'autre part possible de calculer des environnements plus ou moins grands, en imposant un nombre de dictionnaires communs plus ou moins grand entre deux notions pour les considérer comme reliées. Les moyens automatiques et les programmes utilisés pour la sélection automatique de documentations peuvent servir à l'exploitation d'un tel thésaurus. Une application expérimentale a été réalisée dans le domaine médical. Un thésaurus a été constitué pour des noms de maladies, en utilisant un vocabulaire fondamental de définitions. L'expérience a montré que l'environnement déterminé automatiquement pour chaque notion était excellent et permettait de ne plus avoir de documents manquants dans les réponses.

ECONOMIC ANALYSIS OF A TECHNICAL INFORMATION DISSEMINATION SYSTEM. Nathan P. Levy and Ross M. Zigmon, Western Electric Company, Inc., P.O. Box 600, Princeton, N.J., U.S.A.

A crucial question posed by management when developing a technical information dissemination system is: will the participating engineers gain sufficient benefit from its service to pay for the system's operating costs? An economic study, conducted on an experimental basis in 1964 at Western Electric, indicated that a technical information system can pay for itself if it improves the service in the operating costs. An economic evaluation was based on a concept which is described as the "Cost of Information Transfer." This paper describes this concept in detail and relates the concept to a real technical information dissemination system.

"La Propriété Industrielle Monétaire - Théorie des Patents"
CONTRIBUTED ABSTRACTS

THREE EXAMPLES OF COOPERATION. A. Leynaire, High Authority of the European Coal and Steel Community (E. C. S. C.), Luxembourg.

1. Medical, Mining and Iron and Steel Documentation Pool.

2. The Metallic and Iron and Steel sector, the research institute has been grouped in a "Association européenne pour l'échange de la littérature technique dans le domaine de la sidérurgie" (A. E. S. C. L. T.). An agreement linking the High Authority to the specific institute specifies, on the one hand, the terms and conditions for authorizing subsidies to encourage trans-border work, and on the other, the obligations of those who benefit by it. III. Project for an automatic documentation pool in the iron and steel sector. In order to improve their exchange of information, the specialized institutes, who are members of A. E. S. C. L. T. aim, with the aid of the High Authority, to make increased use of electronic machines. A working group is preparing a thesaurus of keywords in the metallurgical and iron and steel sectors, taking into consideration all the terms of the existing classifications as well as the lists of key-words known at the present time. The thesaurus will be established in five languages, which will all be registered on an IBM 360 and linked together, after which the machine will be able to provide, for a given subject in particular, all the references to articles in periodicals analyzed in the various languages.

THE USE OF SECOND ORDER DESCRIPTORS FOR DOCUMENT RETRIEVAL. M. J. Libbey, The MITRE Corporation, Bedford, Massachusetts, U.S.A.

It is proposed that one way to increase the efficacy of document retrieval is to define for the computer the descriptors used to index the file. A computer program written in COMIT to implement the proposal and to facilitate testing its capabilities is described. Definitions are given to the computer as a string of terms called a "definitor." These terms, which act as "second order descriptors," are not normally those used as file descriptors. Their introduction provides a more versatile and powerful search capability to the user. The thesaurus is not exhaustive; it is designed to cover only a limited number of topics. The thesaurus is used to index documents and to retrieve corresponding documents. A similar study of libraries of difficult languages, destined to Library of Congress Classification, is described. The systems were associated with the work of the pool. The research was made with the aid of the High Authority.

A THREE-DIMENSIONAL ARRAY USEFUL FOR MAPPING IN FOUR-DIMENSIONAL, NON-COMMUTATIVE FUNCTION-SPACES. K. J. Lisson, Patrolite Corporation, 369 Marshall Avenue, St. Louis, Missouri 63119, U.S.A.

In a recent article, (1) a detailed description was given of the use of non-commutative composition spaces for the plotting of properties of polymeric materials. Several alternative geometric arrays were illustrated for plotting in spaces of up to three dimensions. Since that time, a method has been developed and used in our laboratories for depicting four-dimensional composition-spaces. The method is general and should be useful for the study of any four-dimensional function-space. The model to be described in this paper, although three-dimensional, has many of the properties of its analogous four-dimensional space, including a place to "put" the fourth dimension. It can be considered a three-dimensional prospective drawing of a four-dimensional space. It has proved very useful in discussions of four-dimensional problems with less mathematically inclined persons.


COMPARATIVE COSTS OF DOCUMENT INDEXING AND BOOK CATALOGING. L. H. Lindley, Aeromotive, Division of Philco Corporation, Ford Road, Newport Beach, California, U.S.A.

Costs of storing and retrieving information are not well understood, and little information is available in the literature to provide this understanding. This study was undertaken in an effort to determine and compare true costs of indexing reports and cataloging books. A prior study the author reported an average input cost of $2.99 per report indexed into a machine document address storage system. Indexing was to a depth of 12.6 access points per document. A similar study of traditional book cataloging costs in the same organization is reported here. Traditional book cataloging techniques using Library of Congress Cards, the Library of Congress Subject Heading List and standard card catalog cabinets were applied to a collection of 10,800 books and monographs. Labor costs ($36,683), supplies ($2,312) and equipment amortization ($641) totalled $39,624, or an average of $3.67 per item, an amount similar to that encountered for indexing reports. However, since the systems contained a significant number of duplicate copies, the cost of cataloging a book for the first time averaged $5.68 while added copies were accomplished for an average cost of $1.42 each. A further difference is noted when access points are compared. For the report indexing system the average cost per access point was less than 24¢ while in the book cataloging one it averaged nearer 95¢. Since the two systems deal with significantly different sized units of knowledge, and because the associated subject authority control established for the book catalog was a more time-consuming task than the comparable vocabulary control used for report indexing, it is concluded that book cataloging in situations as described above is inherently a more expensive procedure than is coordinate indexing. The final evaluation of the two systems would, of course, depend significantly upon retrieval effectiveness and retrieval costs of the two methods.

THE USE OF SECOND ORDER DESCRIPTORS FOR DOCUMENT RETRIEVAL. M. J. Libbey, The MITRE Corporation, Bedford, Massachusetts, U.S.A.

It is proposed that one way to increase the efficacy of document retrieval is to define for the computer the descriptors used to index the file. A computer program written in COMIT to implement the proposal and to facilitate testing its capabilities is described. Definitions are given to the computer as a string of terms called a "definitor." These terms, which act as "second order descriptors," are not normally those used as file descriptors. Their introduction provides a more versatile and powerful search capability to the user. The thesaurus is not exhaustive; it is designed to cover only a limited number of topics. The thesaurus is used to index documents and to retrieve corresponding documents. A similar study of libraries of difficult languages, destined to Library of Congress Classification, is described. The systems were associated with the work of the pool. The research was made with the aid of the High Authority.

A THREE-DIMENSIONAL ARRAY USEFUL FOR MAPPING IN FOUR-DIMENSIONAL, NON-COMMUTATIVE FUNCTION-SPACES. K. J. Lisson, Patrolite Corporation, 369 Marshall Avenue, St. Louis, Missouri 63119, U.S.A.

In a recent article, (1) a detailed description was given of the use of non-commutative composition spaces for the plotting of properties of polymeric materials. Several alternative geometric arrays were illustrated for plotting in spaces of up to three dimensions. Since that time, a method has been developed and used in our laboratories for depicting four-dimensional composition-spaces. The method is general and should be useful for the study of any four-dimensional function-space. The model to be described in this paper, although three-dimensional, has many of the properties of its analogous four-dimensional space, including a place to "put" the fourth dimension. It can be considered a three-dimensional prospective drawing of a four-dimensional space. It has proved very useful in discussions of four-dimensional problems with less mathematically inclined persons.

CURRENT FID ACTIVITY AND FUTURE PROSPECTS FOR THE UDC

International responsibility for maintaining and redacting the UDC rests with the FID's secretariat (The Hague), whose small Classification Department handles all routine work. Responsibility for the production of schedules on more modern lines is carried out by the Classification Committee for the UDC (CCC). Although the FID secretariat normally does not itself publish UDC editions, it controls all proposals for updating and revising the UDC, issuing P-notes to subscribers in many countries, with a 4-month period for comments. P-notes, unopposed or after approved amendment, are authorized by the FID and included in the half-yearly cumulative "Extensions & Corrections to the UDC". Specific subject areas are covered by a network of main UDC committees with working groups, such as FID/C/3 Social sciences (Dutch-German secretariat), FID/C/55 Earth sciences (U.S. secretariat) and FID/C/69 Building & architecture (JDC, with joint FID and CIB participation). It is expected that more than 30 such main committees will eventually be required to fulfill the CCC's long-term program, which aims to redevelop unsatisfactory schedules on more modern lines and introduce new 'common categories' in the auxiliary tables. Vacations of the class number 4 (by merging Linguistics with Literature in 8) has opened the way, after the prescribed 5-10 year 'starvation' period, for a planned new class spanning the pure and applied sciences in classes 5 and 6 and perhaps bridging the gap between social sciences 3 and sciences 5. In applying more facet analysis to categories and transforming them into index symbols, his means are twofold: (1) words and phrases (subject headings, descriptors, etc.); numbers and letters (classification schemes). In the first case difficulties arise when the symbolic function of the index word or phrase is overlooked because of the absence of the index symbols in the textual words. In the second case the notational significance of the numbers or letters is obvious, but their meanings are not always clear. Having surveyed the general statements intended to explain them, Interpretation difficulties in indexing and searching have in recent years led to the compilation of athesauri, some of which try to standardize terminology. The recognition of the complexities inherent in language expression should not induce us to demand that every author use the same words for similar topics, nor that all indexes be confined to the same symbols. We might instead try to create instruments that would help to convert textual or inquiry languages into index-symbols (of various kinds) and vice versa, and that would also act as 'translators' between different indexing languages. This translation function might be carried out by a thesaurus, consisting of clearly defined terms, grouped in systematically classified and alphabetically sequential tables, similar to Koyt's Thesaurus. The structure of this thesaurus would incorporate two different, but complementary search methods: the 'topic-word' and the "synoptic survey or topic-in-context" approach, whose combination might help to deal effectively with different user situations.
TEXTUAL INFORMATION RETRIEVAL AND THE IBM SUGGESTION PLAN. Joseph J. Magnino Jr., Thomas J. Watson Research Center, CHQ YRRC, P. O. Box 218, Yorktown Heights, N. Y. U. S. A.

The IBM Technical Information Retrieval Center provides a computerized in-house information retrieval and dissemination system for the scientific and technological community of IBM. It is in operation using a textual information retrieval system for retrospective searches, current awareness, monthly bulletins, and special library tools, of IBM technical documents and selected non-IBM documents. Over 120,000 technical documents are on magnetic tape for this decentralized service. A specialized application of the normal text information retrieval system is using this system to search IBM Suggestions, to provide the necessary capability to determine whether or not a particular suggestion is new or is one that was previously submitted. The system, tested on searches of 40,000 suggestions, uses computer and teleprocessing techniques to assure that 200 searches (new suggestions) are accomplished daily within a 24-hour turn-around time. A description of the normal text IR search system and the operational procedure, using the IBM Suggestion plan as a special application, is detailed in this report.

PRACTICAL COMPUTER PREPARATION OF BOOK INDEXES. Clifford J. Maloney and Marvin Katz, National Institutes of Health, Bethesda, Maryland, U. S. A.

The preparation of an index to appear in the back of a published book (an internal index) is to be sharply distinguished from the preparation of an index accompanying an entry in a bibliographical index (possibly) unknown at the time of search (an external index). The development of an efficient method for internal index preparation by computer could have considerable interest in general but seems particularly advisable under three definite conditions. The employment of computers in the publication of newspapers, books, and other material is now performed on a routine basis. (Proceedings of the International Computer Typesetting Conference, The Institute of Printing Limited, July 1964) Any additional service obtained as a by-product would, of course, add to the attractiveness of computer typesetting. The publication of reports and congresses such as this one, in which earliest possible appearance of the published proceedings adds greatly to its value, but where author indexing is not practical, or, because of the large number of authors of various backgrounds, is, at the best, very difficult, has more to gain than would be the case in more conventional book publishing. A third field of usefulness exists in rendering indexing possible for material not now indexed. While a few major newspapers have in the past been able to provide indexes to their contents, it is not usual to prepare them and it is not likely to become so until some method can be found to make indexing a by-product (not necessarily entirely automatic) of original publication. The features which distinguish internal from external indexing and which make it easier to achieve have been examined in a previous study. (Indexing Reports by Computer, Clifford J. Maloney, James Dupee, and Sterling Green, TECHNICAL PRECONDITIONS FOR RETRIEVAL CENTER OPERATION edited by Benjamin Chayeleu) One major difference lies in the desirability of very detailed indexing, yet little space can be allowed to print the results. Study of this problem has led to the development of what we believe to be a practical scheme, which has been tried on sample material. We are now seeking to apply it to an actual book.

ARRANGEMENT OF PRIMARY PUBLICATIONS. D. J. Malony, Director of the Center for Agricultural Publications and Documentation, Wageningen, Holland.

Primary publications constitute the basis of documentation of the sciences. The greater the uniformity of these publications in the arrangement and development of topics, the easier is the task of recording documentary data from them. Various attempts have been made to increase this uniformity, but these efforts have met with resistance from the authors because they feel that their personal freedom to publish is hampered by the observance of rigid rules. It would, in fact, reduce the level of scientific output if the author's ideas were to be confined in a straitjacket. But it is a generally accepted requirement that in a primary publication the researcher should follow the shortest possible path between the exposition of his problem and the conclusions, and that he should pursue a logical sequence of thought. This means that the same arrangement can be recognized in many primary publications and that a number of fixed points may be indicated in this arrangement. It would be extremely useful for the processing of documentary data from primary publications if these points could be readily identified. It is therefore proposed to mark these points by means of symbols in front of the line or in the margin. The requirements that such symbols should comply with are being discussed.
ON THE ECONOMICS OF COMPUTER STORAGE AND RETRIEVAL

H. Marron, Assistant Director, and M. Snyderman, Deputy Director, Science Information Exchange - Smithsonian Institution, 1730 M Street, N.W., Washington, D.C. 20036

Despite the widespread interest in the economics of computer storage and retrieval activities (specifically the cost of performing searches and preparing reports) there is little definitive discussion of this subject in the open literature. A few operating installations have released data on unit search costs obtained by dividing total computer processing time for a batch of jobs by the number of search questions in the batch. By so doing, however, the costs for separate single question searches or for building, modifying, maintaining and updating the file are neglected. These costs may be large and can hardly be disregarded.

Alternative techniques have been studied for costing unit searches which include these factors: (1) Computer times for all searches batched or single (2) Proportion of the cost of a batched search to each of its component parts and (3) All computer times for file modification or maintenance tasks. Theoretical results have been compared with actual operating data from computer searches performed at the Science Information Exchange. Further, some general relationships have been developed which can be used by managers, analysts and designers for qualitative and quantitative assessments of the impact of file modification and maintenance tasks on search costs.

This paper confines its attention to direct computer costs only. Later papers will deal with the overall costs which include programming, analysis and overhead.

"WHERE CAN WE GET COMPUTER HELP?"

John Martinson, Inst. for Advancement of Medical Communication 5550 Wisconsin Ave., Bethesda, Maryland, U.S.A.

"All chiefs and no Indians" is a common complaint. Yet, the proper recognition and determination of professional from technical or semi-professional tasks in the information sciences is only beginning to emerge. If professionalisation of this field follows the pattern of others, sharper delineation of responsibilities will occur. (Many analogous examples can be drawn from medical fields.) -- The growing demand for library personnel has brought such terms as LIBRARY TECHNICIAN or LIBRARY TECHNICAL ASSISTANT into wider use. This development has been formalised in the past decade (and largely within the last 5 years) through establishment of special curricula to prepare students for jobs as Library Technicians, Apparaturists, etc. in schools in the U.S. offer such courses. These are mostly four year programs leading to an Associate of Arts degree from a Community or Junior College. -- Through site visits, phone interviews and correspondence these training programs have been surveyed to answer a number of questions. These include:

- Who is in these programs and why?
- What training do they receive?
- What is the design of the curriculum?
- What are the chief problems in the administration of these programs?
- What are the chief problems in the administration of these programs?
- Where are the library technicians placed and what is the nature of their employment experience?
- What future problems are anticipated by those responsible for the development of these programs?
- What lines of research are indicated to help solve present and anticipated problems?

These programs are often characterized by all the difficulties of rapid growth in an effort to meet expanding problems with rather meager resources. The growth rate of these programs appears to be increasing, however, much work on curriculum development needs to be done. Many examples of successful employment can be cited. -- Copies of the Final Report to the U.S. Office of Education will be available. Oral presentations will focus on discussion of selected examples of Recruitment, Training and Employment experiences.
CONTRIBUTED ABSTRACTS

A PROPOSAL FOR REDUCING THE TIME LAG IN SCIENTIFIC PAPER
PUBLICATION SCHEDULES. Hunter P. McCartney, IBM Federal Sys-
tems Division, 7220 Wisconsin Avenue, Bethesda, Maryland, U.S.A.
The lag between receiving and publishing papers by scientific jour-
nals acts as a cross-purpose with these journals' objectives. Authors
hesitate to risk being "skipped" during this long period and resist the
obsolescence a year can add to subjects on the other (a year is a
bout the average time lag under our current schedules). Those con-
cerned with a paper's publication - the author, the reviewers, the edi-
tor, the publisher - are well aware of this excessive lag, but not much
is being done about it. The seriousness of the problem is shown by
the proliferation of "letter"-type publications (Applied Physics Letters)
prime example) to assure authors earlier exposure of timely
subjects. But the advantage of faster publication by these "letters"
and by the "correspondence" sections of journals could be offset by
a lower threshold of quality, for to meet these accelerated publication
schedules, editors are forced to settle for less than the preferred rigor-
ous review they arrange for normal papers. Much of the publication
lag stems from a paper's review, for most reviewers sandwich this time-
consuming, and often onerous, work into crowded schedules and fre-
quently delay these gratuitous efforts because of their job demands.
Another delaying factor is the shortage of the specialized talents and
facilities needed to read articles in type. This paper is built on the
author's observations, on conversation with editors, authors, reviewers,
and others concerned with the timely dissemination of scientific infor-
mation, and on selected articles. No simple remedy is proposed, but
improved reviewer reward, an increase in special facilities and repro-
duction talents, and more attention to origination and presentation, as
well as to dissemination, would provide short-term relief. The ultimate
solution to these problems is to depart from our present methods of pub-
lication and to use the computer more efficiently in publishing, class-
ifying, and disseminating information.

TEXT90. V.S. Mercer and F. E. Franklin, Automated
Documentation, Department D78, Systems Development
Division, International Business Machines Corporation,
Poughkeepsie, New York.
TEXT90 represents a major breakthrough in automated
text preparation. An IBM 7090/94 program, it accepts free-
form text data and formats this data into "camera ready"
print that is suitable for reproduction. Through easy-to-use
controls, the user can specify formatting of a line, page or
the entire document. TEXT90 has both update and format
capabilities. The updating capabilities range from the de-
letion, insertion or replacement of a character, word or
line(s), to the moving of entire blocks of text from one area of
the document to another. This gives the user complete
control of the document during its various stages of pre-
paration. Formatting capabilities provide for normal
functions (e.g., skipping, indenting) and complex functions
(e.g., table and figure generation). These capabilities, when
combined with double or single column format, automatic
hyphenation and justification, assure the user of a properly
formatted document. TEXT90's freeform input can be pre-
pared on punched cards, or the data stored in the Adminis-
trative Terminal System can be converted to TEXT90 using
the ATS-TEXT90 conversion program. TEXT90 output is
printed on an IBM 1401-1403 equipped with a 120-character
upper/lower case print chain. TEXT90 has been proven.
Thousands of pages...hundreds of thousands of lines...have
been published using TEXT90. It has made possible the
rapid dissemination of up-to-date documentation so vital to
the success of the IBM System/360 programming effort.

INCOORDINATION AND SELECTIVE PERMUTATION. Nan P. McCandless,
Pacific Aerospace Library, 7150 Beverly Blvd., Los Angeles 36,
California.
Incoordination of multiple associations printed out in order to
provide immediate access to the entire store of an information system will provide a high degree of spec-
ficity. This may increase the volume of the printout and also, when the permutation printout requires several pages
for a single term, make it difficult to relate the associated descriptors together. There are various methods to correct
this situation, one of which is selective permutation. By this method all descriptors which are linked together by
coincidence of the accession number are no longer permitted indiscriminately, and various criteria for selection can be
adopted.
1. A selective permutation based on type of descriptors or
a specific ranking. In this case two or several descriptor
variants are accepted, they may be called major or minor
descriptors if more than two categories are considered,
a numerical value of descriptor values may be adopted.
By this method it will be possible to select the type of
descriptors which will be permitted. It will also be possible
to promote a major look-up descriptor or index entry any
of the types of descriptors, or include in the associa-
tive permutation any of these categories, without hav-
ing them as a major index entry.
2. Selective permutation may also be based on a combina-
tion of descriptors and classification or group identification.
By this method the alphabetical sequence of the associated
descriptors may be related to one or several major
groupings. By this method it is more likely that descrip-
tors will maintain similarity in concept or function will
appear in the same printout area.
The coordination of the Pacific Aerospace Index pro-
vided statistical observations for the alphabetical pre-
coordination, and use for selected areas, practical
shows for the incoordination and permutation of the
following method considered above.

SHOT, OVERLAP, INCLUSION AND OVERLAP OF ABSTRACTS. 65.
John Karton and Irvingatt Elder, 01116A, 3 Belgrave Square,
London S W 1, England.
Studies of the coverage and subject indexes of abstracting and
indexing services have been made using the technique re-
bibliography on specific subjects are taken, and the appro-
piate abstracts and journals examined to determine coverage, and
overlap between abstracts journals, of the references in each
bibliography. Then the subject indexes of the abstracts
journals are examined to find the terms under which the
covered references were indexed. Figures are presented for
the coverage and indexing of the following subjects: Hydro-
meteorology, Drought control, Determination of tritium
radioactivity, Fallout monitoring, Motivation research. The red
mail, Magnetic and electric suspensions and Bionsic con-
trol.
Abstracts services covered include "Technical Abstracts, Science
Abstracts (A & B), Biological Abstracts, in A
Engineering Index, Nuclear Science Abstracts, and
others.
CONTRIBUTED ABSTRACTS

FUTURE ROLE OF THE REACTIVE TYPewriter IN DOCUMENTATION.

Cathin Koore, Rockwell Research Institute Incorporated, 401 Mount Auburn Street, Cambridge, Mass., 02139, U.S.A.

Within three years the reactive typewriter will come into use in documentation in USA and overseas. The reactive typewriter is the successor to the tape typewriter, with a telegraph line in place of the paper tape. Specifically, the reactive typewriter is a teletypewriter which acquires powerful new capabilities through on-line connection to a nearby multi-access time-shared computer having large-scale storage capabilities. (C.S. Koore, "The Reactive Typewriter Program", Comp. J. U.K., 4/6, p. 46, Jan. 1962.) Between twenty and several hundred reactive typewriters will be connected to the same computer, sharing the cost. The cost is expected to be less than $200 per month, complete. They will be widely used, replacing many ordinary typewriters in laboratories and libraries. We estimate that at least 500 are now in use. Applications in documentation are beginning. Typical uses include manuscript preparation, business letters, library cataloging, union list preparation, transmission of bibliographic information, as well as computation. Users of the reactive typewriter will set up private indexing and retrieval systems, with retrieval being performed from their typewriter. Using communication facilities (e.g., Telex and Telex) similar access will be had to various documentation files in national and international bibliographic networks (cf. de Groller in "L'Organisation de la Documentation Scientifique" ed. Polonon, pp. 93-194, Gauthier-Willars, Paris, 1964.) We are developing a machine-independent computer programming language called TRAC for controlling the interaction between the user and the reactive typewriter. TRAC has been running under test for 1-1/2 years. With TRAC, from the reactive typewriter keyboard, one can: (1) store, make any changes in, and type out any text; (2) store and index collections of any (3) store procedures for doing things; and (4) access external communication channels. (Cf. Koore, "TRAC, A Text Handling Language", Conference Proceedings, ACM National Conference August 1965, to be published.)

INFLUENCE OF THE NEW GENERATION OF COMPUTER SYSTEMS ON LIBRARY AUTOMATION. 

Ormanation Systems, Czechoslovak Academy of Sciences, 907 Pinkovy 7, Prague, Czechoslovakia.

The purpose of research: creation of efficient, coordinated network of SI centers. Method: statistical questionaries. Discussions with SI workers, managers & senior staff scientists. Attention to the correlation of the personal & institutional documentation systems. Results obtained: efficiency of managing, rationalization & economization of scientific work is in direct proportion with informational level of managers & scientists. Heuristic & methodological differences amongst scientific disciplines do not allow to create unique all-sciences SI system. Factors limiting individual SI projects: extent of problems to be solved, ratio of experimental, theoretical & informational work, necessity & possibility of using external sources of SI, application of scientific results in the sphere of technology, existence & attainability of relevant abstract journals, indices & other secondary resources, useful life of actual SI. Scientists cannot be excluded from the information process. He should indicate what information has to be put into memory of SI system & cooperate in creation of effective indexing plan. It is necessary to raise theoretical & technical level of SI work, as well as qualification of information scientists & other SI workers. The problem is that centralization of SI workplaces needs more investigation. Having introduced abstracting & indexing machines synchronized with effective telecommunication means, it seems possible to build up central SI institutions.

SYSTEM ANALYSIS OF SCIENTIFIC INFORMATION IN THE FIELD OF BASIC RESEARCH. 

Augustin Marte, Czechoslovak Academy of Sciences, 907 Pinkovy 7, Prague, Czechoslovakia.

Object of research: theoretical, methodological & organizational problems of SI (scientific informa-

tion) in the institutes of natural, technical & social sciences of the Czechoslovak Academy of Sciences. The purpose of research: creation of effective, coordinated network of SI centers. Method: statistical questionaries, discussions with SI workers, managers & senior staff scientists. Attention to the correlation of the personal & institutional documentation systems. Results obtained: efficiency of managing, rationalization & economization of scientific work is in direct proportion with informational level of managers & scientists. Heuristic & methodological differences amongst scientific disciplines do not allow to create unique all-sciences SI system. Factors limiting individual SI projects: extent of problems to be solved, ratio of experimental, theoretical & informational work, necessity & possibility of using external sources of SI, application of scientific results in the sphere of technology, existence & attainability of relevant abstract journals, indices & other secondary resources, useful life of actual SI. Scientists cannot be excluded from the information process. He should indicate what information has to be put into memory of SI system & cooperate in creation of effective indexing plan. It is necessary to raise theoretical & technical level of SI work, as well as qualification of information scientists & other SI workers. The problem is that centralization of SI workplaces needs more investigation. Having introduced abstracting & indexing machines synchronized with effective telecommunication means, it seems possible to build up central SI institutions.

INFLUENCE OF THE NEW GENERATION OF COMPUTER SYSTEMS ON LIBRARY AUTOMATION. 

Ormanation Systems, Czechoslovak Academy of Sciences, 907 Pinkovy 7, Prague, Czechoslovakia.

The purpose of research: creation of efficient, coordinated network of SI centers. Method: statistical questionaries. Discussions with SI workers, managers & senior staff scientists. Attention to the correlation of the personal & institutional documentation systems. Results obtained: efficiency of managing, rationalization & economization of scientific work is in direct proportion with informational level of managers & scientists. Heuristic & methodological differences amongst scientific disciplines do not allow to create unique all-sciences SI system. Factors limiting individual SI projects: extent of problems to be solved, ratio of experimental, theoretical & informational work, necessity & possibility of using external sources of SI, application of scientific results in the sphere of technology, existence & attainability of relevant abstract journals, indices & other secondary resources, useful life of actual SI. Scientists cannot be excluded from the information process. He should indicate what information has to be put into memory of SI system & cooperate in creation of effective indexing plan. It is necessary to raise theoretical & technical level of SI work, as well as qualification of information scientists & other SI workers. The problem is that centralization of SI workplaces needs more investigation. Having introduced abstracting & indexing machines synchronized with effective telecommunication means, it seems possible to build up central SI institutions.

DISSEMINATION OF JAPANESE NUCLEAR LITERATURE IN THE WORLD.

Mieko Matsumura, Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken, Japan.

The dissemination of Japanese nuclear literature involves many problems, because of the language barrier and the way of dissemination. An investigation has been made of the subject matter, using the papers whose abstracts have appeared in the "Nuclear Science Abstracts of Japan." Approximately 2000 papers, covered in the above abstract journal in the past two years, were analyzed; as a result, the Bradford's chart was prepared. Thirty journals were selected as the core and critical ones, and their dissemination in the world has been analyzed in detail. Forty percent of all the papers have been written in English; of these, 65% were carried in English-language journals and 35% in Japanese-language journals. The percentage of papers written in English has differed from one field of research to another. 70% for physics, 50% for chemistry, etc. Abstracts in English were attached to 20% of the papers. In the case of English-language journals, there is a problem of correcting the English used in these papers. These journals are intended for overseas distribution, and 20 to 30% of the numbers of copies are sent to overseas countries. The distribution differs considerably from one country to another, with the United States receiving nearly 30% of the copies sent. For Japanese-language journals, on the other hand, only 1 to 2% are sent. Reprints are playing a role in the dissemination of papers. Thus, the state of dissemination of papers in the world is considered unsatisfactory. International abstract and index journals are made use of in this connection. Papers written in English are easily accessible. For the papers written in Japanese, however, there is a problem of translation.
CONTRIBUTED ABSTRACTS

ECONOMIC POLITICS IN THE IDEA PLANES IN THE DESIGN OF A SCHEME FOR DEPTH CLASSIFICATION. A. G. Setham and J. A. Gunenth. Documentation Research and Training Centre, Bangalore, India.

Three concurrent developments in India classify thought area: investigation on the methodology of design, making the design work assemble to scientific method, and making the work of the classificationists and of the classifier more productive. Productivity in design is increased by the separation of the work in the idea, verbal, and notational planes, bypassing the confusing phenomenon of level by giving deeper to the near-semantic levels to sense the essence of the bonds, assigning a scale of priority to the problems to be dealt with, formulating guiding principles, and reducing the strain on the memory by the use of mnemonics. Discusses the residual problems in the application of the Wall-Picture Principle in determining a helpful sequence among the First Characteristics in the schedule for the personality facets in subjects going with a basic class in commodity production engineering. Observes a pattern in the formation of groups with internal affinity among the First Characteristics. In step 2 of the design suggests such grouping of the First Characteristics, the application of the Wall-Picture Principle first to the groups to determine a helpful sequence among them, and then to the First Characteristics within each group to determine a helpful sequence among them. This nearly eliminates the need necessary to select pairs of First Characteristics one after another for the direct application of the Wall-Picture Principle to one pair at a time, and also provides the time taken to determine a helpful sequence among the totality of the First Characteristics. Further, the groups of First Characteristics with internal affinity are recognized even by a beginner without much difficulty. The groups of First Characteristics with internal affinity lie at a level nearer to the seminal levels than those of the First Characteristics themselves. The Wall-Picture Principle gives a sequence of the First Characteristics which nearly parallels the sequence of thinking by the majority of the production engineers.

THE HYPertext. Thomas H. Nelson, Johnstone College, New Mexico, U.S.A.

We are at an historic divide like that created by movable type. In supplement movable type was a further invention as the Offprint, text, magazine, bookshelf and dictionary. The technology of automatic display, stored-program computers and bulk information storage now for a line of inventions and conventional display modes, organizing arrangements, and anecdotal and professional roles and practices that may be the most from these new devices.

This paper proposes a new medium, the hypertext. This is a generic term for texts and combinations of texts with other materials which, because of their structure, require automatic handling and dictionary devices. The hypertext will typically be non-linear, branching, and large, with various returns to the idea.

The idea of a hypertext is distinct: for Information Retrieval; Specialist Information and query search, Information Display, Information Storage, Query Answering, Information Retrieval from large sequence organization and retrieval. Some of these files the general case. The hyper text is defined not as a specialty material (which is an oxymoron), but to give general reference and instruction. It may contain a core of general and specific materials, with their many interconnections indexed. Arbitrary nodes of subject matter can be interconnected, offering a choice of sequential generation. Like an anthology, the hypertext contains pre-populated text and information from sources of original text. It may also contain memories and abstracts unique to the user, assuming a memory foundry to store different backgrounds, competencies and interests, allowing the user to add such details to his own file. A hypertext system will have the form of linked indexes, hierarchies, or explorations. It could have new types of graphic handling (for example, moving a mouse to show operations (volute, zooming,芜cal), areas of broaching or scrolling in, correspondence, conceptual and general information. It will permit self-assessment activity to interest and motivation, even existing natural facility and intellectual function, which are more hidden conventional instruction.

STATUS AND CONTROL OF UDC FULL EDITIONS


A survey is made of existing UDC full editions, i.e., in French, German, English, Spanish, Portuguese and Polish, and of projected full editions, e.g., in Russian, and a tabulated comparison of the classes, divisions and sections available in all languages is provided. The international and national organs for the control, revision and extension of the UDC are described giving an indication of the purpose, structure, membership and activities of such organs, i.e., the Central Classification Committee and its special sub-committees and working groups, the various national UDC committees and their sub-committees and working groups. Finally, an outline is given of the existing Rules for UDC Revision Procedure: general principles, classification and notation principles, preparation and submission of proposals, treatment and circulation of draft proposals, publication of proposals as P-notes, conversion of approved P-notes into "Extensions and Corrections to the UDC."
CONTRIBUTED ABSTRACTS

PEER-GROUP JUDGMENTS ON SCIENTIFIC MERIT: EDITORIAL REFEREEING.
Richard H. Orr and Jane L. Kasab, Institute for Advancement of Medical Communication, 4540 Locust Street, Philadelphia, Pennsylvania 19104, U.S.A.

In making decisions on awards, honors, promotions, and publication, the scientific community relies on peer-group judgments of a scientist's work, i.e., on the consensus of colleagues qualified to assess his accomplishments. Despite the importance of this method of evaluating scientific merit, neither it nor alternative methods has been examined objectively. Because the practice of using editorial referees to judge manuscripts submitted to journals provides a model of the general method that lends itself to systematic analysis, and because editorial refereeing is considered to be the primary means of quality control in the documentation chain from generator to user, it was selected for the first of a series of studies on evaluation of scientific merit. The aims of this study are: 1) to assess the editorial refereeing process quantitatively, 2) to identify factors that affect the process, 3) to generate hypotheses on how its reliability and efficiency might be improved, and 4) to compare the process with other means of evaluating scientific merit. Editorial records on all manuscripts (almost 5,000) submitted to two biomedical research journals from 1957 through 1961 were the chief data source. These journals used a form for referees' reports that required a numerical rating of scientific merit (1 = "poor", 2 = "good", 3 = "low but acceptable", and 4 = "acceptable"). In addition, a sample of the manuscripts submitted in 1957 were returned to the referees who originally reviewed them to be reevaluated in retrospect. The findings to date include: 1) On the critical decision of whether a paper was acceptable for publication, referees agreed in about 70% of all cases where only two opinions were obtained. 2) There was closer agreement on "poor" than on "good" or "acceptable" papers (ratings 1 and 2). 3) Referee's retrospective judgment about the acceptability of a paper agreed with his original opinion shown as often as two different referees agreed. 4) In the retrospective judgments, referees tended to find fewer papers unacceptable.


The standardisation of documentation is carried out mainly by the International Organisation for Standardisation Technical Committee 46 (ISO/TC 46), which approved as far ten of ISO recommendations on documentation. Unfortunately Japan has no national standards of documentation. Though the editors of scientific journals in Japan had their own rules of editing according to the modern society's needs, they were never conscious of the ISO recommendations. Eighteen leading chemical journals in Japan were examined for the possibility of accepting the ISO recommendations: ISO R215, "Presentation on Contribution to Periodicals," is generally received except recording classification marks; the use of ISO R214, "Abstracts and Synopses," is not sufficient; and most of the journals come to introduce gradually synopses in each articles, but resist to have abstract column; for ISO R77, "Bibliographical References," the bibliographical descriptions of books or other separately published works as references are not standardised, and if contributions to collective works or synopses are used, their descriptions are desperate; ISO R8, "Layout of Periodicals," is generally received, but in some journals there is no existence of title pages, incomplete bibliographical marks on each page, or doubling of pagination (pagination both for issue and for volume are recorded on each page); ISO R8, "Short Contents List of Periodicals or Other Documents," is received; many journals; ISO R30, "Bibliographical Strip," is generally applicable in Japanese journals; and ISO R4, "International Code for Abbreviation of Titles of Periodicals," is never applicable in chemical journals. It is very necessary for Japanese chemists to read "Chemical Abstracts" code for abbreviation.
TESTING THE EFFECTIVENESS OF A THESAURUS-CONTROLLED SUBJECT INDEX. Paul E. Palatt, American Society for Metals, Metals Park, Ohio, 44073, U. S. A.

The objective of this study is to determine the content and format of the Index. The objective of this study is to determine the content and format of the Index so that the user will be provided with an effective retrieval tool. The measure of effectiveness was used in the computer output from the 'telegraphic' abstract system employed by ASM over the past eight years. The same file of documents (approximately 3000) was both hand searched through the subject index and computer searched. The search questions were selected from those submitted by subscribers to the ASM Information Retrieval Service and represent a cross-section of all the types of questions received. The study analyzes the reasons a document was retrieved by the computer search of the deep index and not retrieved by the subject index search. Quantitative results concerning the effect of indexing depth on retrieval efficiency are presented. The results of this study have led to changes in subject indexing practice and these changes coupled with subsequent testing are expected to lead to a truly effective subject index.

CATEGORIES AND RELATORS: A NEW SCHEMA. Jean M. Perrasault, Florida Atlantic University, Boca Raton, Fla., U.S.A.

In a mechanized system of bibliographical storage and retrieval, the need for a more adequate system of relational terms with which to eliminate the chance of false coordination (and to make possible new complex classifications) between extent indexing/classing terms (whether verbal or notational) led originally to a comparative tabulation of various available relational and categorical schemata. Their omissions seem traceable to their pragmatic (non-systematic) origins but even when taken as one collated system some of these omissions remain. (Several available schemata are shown and discussed.)

An alternative route to an adequate and universal schema is described. Just as there are signs of interplay between 'general categories' in the setting up of the extent schema, such an interplay systematically carried through, is made the basis for a new non-pragmatic schema, the general categories used are a triad of triads: CANONIC (e.g. lateral, axial, vertical), POSITIVE--INDETERMINATE--NEGATIVE (e.g. active, interactive, passive), and TOTALITY--TOTALITY/ELEMENTS--ELEMENTS (e.g. submissive, determinative, ordinal). A notation of letters (a--j) or of numbers (1--9) embodies the concepts in a hierarchical arrangement, offering relations from the most general to the most particular.

(A chart abstracting the conceptual types from the various available lists is given as introduction to (1) a fully systematic schema, (2) a classified-index tabulation, and (3) an alphabetical chain-indexed tabulation.)

Examples are given of use of the schema in combination with U D C numbers in indexing narrowly- or broadly-focused documents, and projections of results of the combined system when used as a strategization for mechanized searching are given in detail.

AN INFORMATIVE ABSTRACTING TECHNIQUE: DEVELOPMENT AND VERIFICATION. Dan Payne, John Hale, and Sara Munger, American Institute for Research, 410 Amherst Avenue, Pittsburgh, Pennsylvania, U. S. A.

The goal of the research was to develop and verify guidelines for preparing highly informative abstracts. In terms of objectives, the study was intended to develop guidelines that result in abstracts which (1) provide maximal support to abstract-users, and (2) are consistent, or reliable, abstracts of scientific/technical material. An Abstract form and instructions were developed by integrating features from 122 different sets of abstracting instructions received from requests for such instructions sent to 276 organizations. The form consists of a number of sections and sub-sections, generally related to the organization of journal articles, with specific instructions for each sub-section. Two studies were conducted to verify the reliability of the abstracts produced by the guidelines. In the first study, a test of consistency of the abstracts was conducted. This test was based upon expert judgment of similarity of the amount of information contained in each of the sub-sections of an abstract. Six papers were abstracted by three different abstractors. The 18 abstracts were judged to be 88 percent consistent in terms of information content. Tests of performance support involved comparing performance using abstracts with performance using full text, under two different conditions of time, on tasks known to be dependent upon textual material. The first study, employing 20 senior electrical engineering students, indicated that there was virtually no difference in quality of performance, but that significantly less time was required by students using abstracts. The second study involved four different general tasks encountered by scientists in industry, and included two types of abstracts: the general-purpose, or full, abstract, and an abstract "tailored" specifically to each of the tasks. Subjects were 228 students and 40 professional scientists. Results indicated that both abstracts served about as well as full text, but with significant decreases in time required over the two abstract types.

ORGANIZATION OF NUCLEAR SCIENCE CONFERENCE LITERATURE. Margaret L. Pfueger, Division of Technical Information, U. S. Atomic Energy Commission, Oak Ridge, Tennessee, U. S. A.

Scientific information presented at conferences is becoming a permanent part of the recorded scientific literature. A study of review articles on nuclear subjects shows frequent references to conference papers with no indication of publication. While many conference papers are eventually published in journals, they are often quoted and referenced as papers before they become part of the journal literature. Many conference papers are never published in more permanent form, and the author is the only source of availability. The U. S. Atomic Energy Commission has attempted to cope with this type of nuclear science literature in a number of ways; by acquiring and announcing it in Nuclear Science Abstracts, by making it publicly available, and by using special indexing techniques to aid in the easy identification and location. A unique KWIC index provides access to information on past conferences by date, city, and important words in the title. A discreet numbering system brings together preprints of papers presented at a conference and links them to the published proceedings.
CONTRIBUTED ABSTRACTS

LATERAL AND MULTILATERAL INTERNATIONAL COOPERATION IN SCIENTIFIC AND TECHNICAL INFORMATION. Wosleieh Pirg, Central Institute for Technical Information and Documentation, Warsaw, Poland.

A concise outline of current trends in and development of international cooperation, bilateral and multi-lateral, in the field of scientific and technical information, with examples cited from several countries and regions. The cooperation of international organizations concerned with problems of scientific and technical information, the necessity of further development of cooperation in the field mentioned, in order to match the violent progress in science and technology. Suitable services rendered by the national information centres, the character of these services, the exchange of the knowledge, the problems of scientific and technical information, the selection and distribution of tasks in documentation. Optimal conditions of cooperation. The basic concepts concerning the principles and norms of bilateral and multilateral cooperation & bilateral, the need for development, and multilateral agreements; jointly organized and sponsored international agencies and organs. The problem of the international service of scientific and technical information. The role of FID in initiating and coordinating the international cooperation in the field of information and the tasks assigned for the coming years.

SCIENTIFIC AND TECHNICAL INFORMATION FILE FOR THE NATIONAL CANCER INSTITUTE. A. E. Pratt and W. C. White, National Institutes of Health, Bethesda, Md.

A computer-based information processing system is being evaluated as an aid in the management and analysis of the Grants Program of the National Cancer Institute, NIH. The individual record which forms the data file describes one grant and contains administrative, fiscal and scientific data. Each record is composed of an ordered set of nineteen data-line entries: five entries are of fixed-length format, the remainder have a variable-length format allowing a multiplicity of data items to be entered per line. The scientific data are subject-descriptive in terms of phrases abstracted from the grant document by human indexers. Management activities require identification and retrieval of any subset of records relative to any specified logical combination of administrative, fiscal and scientific data. The analysis activities require in addition, the potential for the exhaustive examination by scientific subject content of any identified subset of the file records. A paper tape-oriented, IBM 1620 with 80 digits per tape and random access capability plus a printer comprise the machine system. Computer programs provide for: 1) the creation and maintenance of the machine-stored, natural language information file, 2) the retrieval of records by search on specific data items using combined logic operations, i.e., conjunction (AND), and (OR) and disjunction (XOR), 3) the systematic organization of the set of any subset of the file records by an in-depth matching of a record's subject content against a machine-selected list of subject content automatically abstracted from the entire record set, 4) the routine sorting, alphabetizing of data lists and a variable output format capability.

ANALYSIS OF CITATIONS IN SOURCE PAPERS IN THREE RETROSPECTIVE BIBLIOGRAPHIES ON (I) HYDROFOIL CRAFT, (2) CONTINENTAL DRIFT, AND (3) TECTONICSEISMIC SEA WAVES. Sharlene Rafter, U.S. Coast & Geodetic Survey, Library Br., 13800 Old Georgetown Rd., Rockville, Md., 20855, U.S.A.

Analysis of citation patterns in source papers in three retrospective bibliographies, each covering a specific field, with controlled time span, 1960-1964, was made to establish use factors and citation patterns in scientific and technical literature to guide acquisition, retention, and indexing policies in technical libraries. Hydrofoil Craft, Design and Development, 1960-1964, Continental Drift, a theory, 1960-1964, compiled by the author, and Annotated Bibliography on Tsunamis, compiled by U.S. Coast & Geodetic Survey (1960 Monograph 7) were used for the study. Source cards were made for each citation in the bibliographies. References cited in source papers were checked, items contained in bibliographies were noted, "cited ... on source cards. Items cited, not in bibliographers, were added with reference as first citation. ZATCR edge-notched cards were used as source cards. Analytic of citations shows that: (1) Principle factors in subsequent citation of papers are professional standing of author and journal of origin, publication and language of original. (2) "Classic papers" are cited repeatedly, while major portion of published literature is occasionally cited. (3) Papers are cited more frequently during first five years after publication, less frequently during next five years, and sporadically after first ten years. Most frequently cited papers are included in bibliographies, specialized monographs and reviews of the literature, which are then cited, rather than original papers. (4) Major portion of source papers is originally published in a comparatively few leading journals. (5) Citations are the key to retrospective bibliography preparation, serving as multiple access points to specific areas of literature. Importance of these five factors to technical library acquisition, retention, and documentation policies is discussed.

HOW BIOMEDICAL INVESTIGATORS USE LIBRARY BOOKS. L. Miles Raisig, Meredith Smith, Read McCuff, Frederick G. Kilgour, Yale Medical Library, 333 Cedar Street, New Haven, Connecticut, U.S.A.

Only a few studies have been concerned with the use of biomedical books. This paper reports an investigation into use made of library books by biomedical investigators. Cancelled charge slips were collected at the Yale Medical Library circulation desk each day. On the following day, those slips for books which had been returned by research investigators were segregated. Next, an appointment was made with an investigator for an interview. The interviewer obtained answers from the investigator to questions on a questionnaire designed to elicit information as to how the investigator had learned of the existence of the book, whether or not the book had been useful to him, and if it had, to what use he put it.

During the six-month period of the study, researchers returned 2,735 volumes of which 831 or 30.4 per cent were monographs. Nearly four-fifths of the books withdrawn supplied information wanted and about four-fifths of the books used were printed in the previous decade. Nine-tenths of the use of books by 130 academic investigators was research-related; the other one-tenth was used for lecture preparation. Over a quarter of book usage is to obtain general information and about 15 per cent is associated with the intellectual aspects of scientific activity.
CONTRIBUTED ABSTRACTS

EDUCATION FOR DOCUMENTATION S. R. Ranganathan National Research Professor in Library Science, Documentation Research and Training Centre, Bangalore, India.

The umbra region of documentation consists of documentation, work and documentation service. The penumbra region consists of computerised work, translation, and use of machines for retrieval, translation, abstracting, and indexing. The work of the documentationist lies only in the umbra region with a dash of interest in the top management in the work of the penumbra region, which belongs to different kinds of techniques. Education for documentalists should cover depth classification, subject heading cataloguing, facet analysis of queries, reader, administrative subjects such as selection of documents, acquisition, circulation work, and intimate familiarity with documentation, periodicals and other reference books, layout and fitting, and furniture of documentation centre and the national and international organisation for documentation. Recent fundamental research has opened the way for continuous developmental research in the umbra region. The teaching method should be largely discussion and practical work by tutorial work, practical work and study by students, colloquium work, and project work. The training should inspire the imagination of the students and stimulate them about the unwinding challenge of documentation as a career and as a social necessity. The current emphasis on automation caused by the use of the term 'documentationist' both to denote the work of the documentationist working in the umbra region and also the different kinds of techniques used in the penumbra region should be removed.

A. B. REASER, MULTI-LANGUAGE DOCUMENTATION SYSTEM IN USE AS AN ACTIVITY POINT, Malcolm Price, 1 N Weather Bank, American Meteorological Society, Washington, D.C.

Limited-use documentation systems may be built around single-access or possibly double-access concepts, but systems designed for many types of users and work are a long time away. They must not be confused with a limited or dead-end concept, but must have the flexibility to be used as a database, a library, a catalog, a historical, sequential, serial titles, language, form of work, and several other kinds of access points to each document with or without abstracts. The system may be built for listing, or indexing, or at card or bibliography preparation, either in categories or as a selective retrieval and printed system. The process is completely automatic in that the page size, line length, page numbers, running heads, standing subheadings, suppression of or inclusion of any capital material, bibliographic citation, index or abstracts of a desired degree of abridgment or expansion, and in the form of printed documentation, may be used for highly specialized section of text display by press, hierarchical subdivision, automatic technical and bibliographic limitations, and these, if used, in a limited sample of 1000 documents already stored and indexed by this system, aid in predicting or solving problems which one might encounter in processing a much larger sample. The system can be used equally well with IC and D or a combination of the two, with only slight adaptation.

EURATOM'S NUCLEAR DOCUMENTATION PROJECT, Loli Balbuono, European Atomic Energy Community, Str. rue Belair, Brussels, Belgium.

The paper describes the information needs of Euratom and its member countries and the aims and purposes of the Documentation Project. The fundamental decisions underlying the Project are explained and the development of the system is described and an outline is given of the present activities, including the user service, now in its starting period. Particular emphasis is given to the various problems to the computer processing of index terms, to retrieval strategy and performance, and evaluation of results.
CONTRIBUTED ABSTRACTS

85

DOCUMENTATION TRAINING IN INDIA. J. C. Jadhav, Indian statistical Institute, Calcutta 70, India.
There is no general agreement on the definition of "documenta-
tion" and the professional image of the documentalist or the
information scientist is still not defined. A few years, however,
has come up to integrate scientific activity into information
with different stages of activity, e.g. scientific research. There
is subsequently in the pattern of ICML (International
Conference) in India, the documentation scientist and the
information specialist is defined as a person who studies
information and the professional image of the information
scientist is defined as a person who studies information.

AUTOMATIC INDEXING OF A SCIENTIFIC ABSTRACT
JOURNAL BY HIP UDC WITH UNI DEX. Martin Russell,
American Geological Institute, 1444 14 St., N.W., Washinton, D.C., 20005, U.S.A., and Robert R. Freeman, American
Meteorological Society, F.O. Box 3, 50, Washington, D.C.,
20013, U.S.A.
The development of Meteorological and Geophysical Titles of UNIDEX (currently suggesting "DC" mixed), a
computer arranged title index based on the Universal Decimal Classification (UDC), suggested the possibility of high-speed,
low-cost computer indexing. The present paper describes the
programmes which can be used to incorporate computerized
searching into the existing unification of the University of
Regina, Saskatchewan, Canada, and the University of
North Dakota, Grand Forks, U.S.A.

A PROGRAMME FOR BETTER ORGANIZATION OF DOCUMENTATION. Dr. J. Lamens, 79, rue des Saints-Pères, Paris, France.
The Symposium organized by A.I.D. (Association Internationale des Documentalistes) on the "Adaptation of the
document to its function" has reached the conclusion that
collaboration is necessary between documentalists and pub-
lishers or producers of documents. New techniques such as
photography, dissemination, retrieval, and regrouping of
documents are powerless to satisfy the increasing
demand for information if the documents are not designed
and presented from the very beginning, as a function of
these requirements and techniques. This cooperation could
take place within a large federation grouping documentali-
sts, publishers and institutions, in which the problems raised by standardisation, distribution, publicity,
copyright and international language would be solved on a practical basis. This cooperation could lead to
a flexible and efficient organisation, which is described and
which is based on the following principles:

- all documents could be published with a complete abstract
in English on the first page with reference, author,
address, title, complete summary, general line, key words
and key ideas.

- each abstract can then be photographically recorded on
to microfiche (Filmores type) so that it can be filed,
retrieved and regrouped electronically or visually.

- the abstracts filed or regrouped on microfiches can be
disseminated by means of photograpnic process (Biblioraphia,
photocopying).

- a general or specific indexes could be formed by means of
computers (IBM, Pull, ...) from photographically
grouped abstracts and from the recorded keywords.

DEVELOPMENT OF A MEDICAL SPECIALTY RECORDING BIBLIOGRAPHY
The Index of Rheumotology is a newly-developed recurring
bibliography produced by the Medical Literature Analysis and
Retrieval System (MEDLARS) of the National Library of Medicine,
and published semi-monthly by the American Rheumatism
Association. The Index is a product of a year-long literature
analysis project (1) "Rheumotology" was defined and delin-
eated according to the results of a survey among rheumatologists.
A Discussion of Rheumotology was prepared to improve the index-
ing of publications and to verify the vocabulary. The Index of
Rheumotology was developed in cooperation with MEDLARS.
This report describes preliminary testing of MEDLARS' output,
migration of MEDLARS' Medical Subject Headings (MSH), and
the computer formulation for retrieval of the bibliographic
solutions comprising the Index. a) Rheumotology MSH terms
were created by MEDLARS. The retrieved citations were
examined or correlated for refinement of the retrieved
citations. b) A manual search thru specific issues of journals known to
be stored in MEDLARS indicated additional activity, categories, and library
modifications. c) Subject headings of peripheral interest were
listed in retrieval manual suggested searches from MEDLARS,
until the most distinct correlations were discovered.

- New rheumotology terms were introduced into MEDLARS.

- A computer formulation was modified and introduced into
MEDLARS' 1st version of a biomechanics project.

- The Index, which has been in publication in January 1965, has a present subscription
of 3,000. It is expected to provide over 8000 bibliographic
citations annually from MEDLARS' list of periodical
literature.

A. Z. Kohlhoff, W. A., and Babolat, L. The literature - a special
project of the American Rheumatism Association.
Arthur Blume, 1. 674, 1965
This project was supported by Grant No. AM-0705-410,
National Institutes of Health, to the American Rheumatism
Association.
IDENTIFICATION AND CONTROL OF VARIABLES IN INFORMATION RETRIEVAL

I. INTRODUCTION

Terko Saracevic and Alan M. West, Center for Documentation Research, School of Library Science, Western Reserve University, Cleveland, Ohio 44106. 

Experimentation in information science has suffered from lack of definition and deficiencies in experimental design and control. With the absence of such control it is often difficult to determine what has in fact been tested and under what conditions. The reproducibility of experiments, if desirable, is consequently open to doubt. Inadequate attention has been given to experimental design with even less care devoted to the provision of procedures for the successful execution of experiments. This paper identifies the essential elements necessary for valid and reliable experiments in testing retrieval systems. The variables (components and sub-components) operating within the context of retrieval systems are enumerated, defined and discussed. Models, experimental design and methods of controlling variables in actual experimental situations are illustrated. The meaning of control in retrieval experiments is discussed. Attention is given to sources of experimental bias which contaminate research findings. Specific points are discussed with reference to experimentation currently being undertaken in the Comparative Systems Laboratory at Western Reserve University. 

II. DEFINITION AND PROVISION OF PROCEDURES FOR THE SUCCESSFUL CONDITIONS

For valid retrieval systems it assists the program in developing a criterion group to characterize the same documents. In the present study, author indexing and indexing derived from document titles ("Comparative Indexing: Terms Supplied by Bibliographic Authors and by Document Titles," American Documentation, in press) were assessed relative to the term-choices of the criterion group, and of the document corpus, needed for any program with insufficient control. The findings to date indicate 1) that assessments of relative merit based on this method are reproducible (a) with different criterion groups drawn from the same scientist population and (b) with different document samples from the same universe; 2) that the size of the criterion group, and of the document corpus, needed for reliable assessment is small enough to make the method practical; and 3) that various methods of weighting and scoring term sets gave much the same results. Tests of other variables are in progress. In addition to comparing the quality of indexing achieved by different means, possible applications of the method include 1) measuring the indexing ability or performance of individual indexes, 2) the output of indexing services, and 3) improvement of authority lists.

III. CRITICISM

Mairtini Schuchmann, Deutsche Normausgabe, Deutscher Normenausschuss, Division IV, Düsseldorf 175.

The Dewey-Classification (1) was introduced in the 19th century and has been continuously developed and extended since. It was designed to meet the needs of library users and to facilitate the search for information. The Dewey Classification is based on the decimal system, where each subject is assigned a number ranging from 000 to 999. This system is used in many libraries around the world and is still widely used today.

IV. ABSTRACTS


The second edition of the Dewey Classification was published in 1927 by the Library of Congress. It includes 100 new numbers and a number of modifications to make it more suitable for modern library needs.


The third edition of the Dewey Classification was published in 1980 and includes changes to the classification system to reflect changes in library practice and the information needs of library users.


The fourth edition of the Dewey Classification was published in 1995 and includes updates to the classification system to reflect changes in library practice and the information needs of library users.


The fifth edition of the Dewey Classification was published in 2010 and includes updates to the classification system to reflect changes in library practice and the information needs of library users.


The sixth edition of the Dewey Classification was published in 2015 and includes updates to the classification system to reflect changes in library practice and the information needs of library users.


The seventh edition of the Dewey Classification was published in 2020 and includes updates to the classification system to reflect changes in library practice and the information needs of library users.
The "LITERATURE EXPLOSION" and its effect on documentation theory. T.V. Scrivenor, Secretary, Commonwealth Agricultural Bureaux, Farnham House, Farnham Royal, Buckinghamshire, England.

The "literature explosion" is an article of faith with documentationists. It is often asserted that the number of scientific publications is increasing exponentially. Scientific literature is undoubtedly growing, and in certain fields it may even be increasing exponentially; but this is not true of all fields, and in some the rate of growth seems relatively slow. Growth can be measured only in relation to the number of publications in circulation, and estimates of this vary from 30,000 to 100,000. It is important to establish and maintain an agreed figure so that the rate of growth in the various fields may be measured accurately. Because it has been assumed that the rate of increase is explosive, elaborate and expensive systems are devised to deal with a situation that in some areas does not exist. These systems emphasize "current awareness" at the expense of depth of coverage. But informative abstracts are still needed by working scientists, especially in developing countries, and the abstract journal is still an effective vehicle of scientific information. It would not be easy to compile a universally acceptable "world list" which would involve agreement on what constitutes a scientific publication and when its death may be presumed, and this would demand a degree of practical international co-operation rarely achieved except in war. It is, however, a task that could be tackled jointly by UNESCO and FID.

ASCIA - AN AUTOMATIC CURRENT ALERTING SYSTEM MADE OF CITATION INDEXING. Irving H. Sher and Eugene Garfield, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pa. 19106, U.S.A.

Automatic Subject Citation Alert (ASCIA) is a system for weekly dissemination of data, which selects relevant items from the literature, based on any of several criteria including citation indexing. ASCIA bypasses certain problems involved in formulating questions for language-oriented retrieval systems. Through "SPECIFIC CITATION" citations (any known published work), ASCIA will alert a client of the current articles or patents citing any items in his interest profile. The questions in a user's profile are compared with every reference cited in the current literature. In ASCIA, the references in the bibliographies of current works serve as unambiguous indexing terms, identifying appropriate current papers. The specificity of citation indexing enables ASCIA to select appropriate items from a wide variety of publications including "peripheral" journals which only publish occasional items of interest to a particular user. Other types of questions which may be entered in an ASCIA profile include: the REFERENCE AUTHOR (or CITED AUTHOR) QUESTION - which retrieves all current journal items or patents citing any work by a given author; SOURCE AUTHOR QUESTION - retrieves all current journal items or patents by a given author, regardless of whether he is a primary or secondary author; ORGANIZATION QUESTION - retrieves all current journal items attributed to a given corporation, institution, or other organization; PATENT ASSIGNEE QUESTION - retrieves any current patent assigned to a given firm; PATENT CLASSIFICATION QUESTION - retrieves any current patent by the specified class or subclass number. ASCIA covers comprehensively, on a current basis, over 1,070 journals and all U.S. Patents. ASCIA allows the user to add new questions at any time, as additional pertinent works are discovered. The weekly ASCIA reports also provide a measure of pertinence since the reports show which and how many of the questions in a profile are cited by each current paper to which the client is alerted.

RECURRENT BIBLIOGRAPHIES AS CURRENT AWARENESS TOOLS FOR PROBLEM-ORIENTED RESEARCH: EVALUATION OF AN EXPERIMENTAL PRODUCT OF MEDLARS, Andrew M. Sherrington and Richard H. Orz, Institute for Advancement of Medical Communication, 4040 Locust Street, Philadelphia, Pennsylvania 19104, U.S.A.

Advanced documentation systems, such as the Medical Literature Analysis and Retrieval System (MEDLARS) of the National Library of Medicine, can periodically provide bibliographies of current literature tailored to the special interests of relatively small groups of scientists working on a common problem. Seventy scientists engaged in research related to cerebrovascular disease constituted a test group for a prototype service, the Cerebrovascular Bibliography. Issues of this bibliography contained all the citations that had appeared during a 4-month period under 101 selected headings in Index Medicus. The utility of this publication for searching and for current awareness was assessed by questionnaires and use diaries completed by the test group. This report deals only with data on the potential value of this publication as a current awareness tool. These data include 1) the headings under which a scientist expected to find citations relevant to his work before he became familiar with the publication; 2) in one issue, every citation that, on the basis of the information provided, appeared to be relevant to his own work; and 3) of the citations he identified as relevant, those that were "new" in that he had not previously known of their existence. The data were analyzed 1) to assess the efficiency of the publication in serving the group and each of its members, 2) to develop group and individual relevance patterns, and 3) to characterize reading habits. Among the more interesting findings are the following: The test issue contained 3052 citations of different articles in about 800 different journals; 45% of the articles were in English. The papers relevant to more than 1/3 of the test group could all be found under 16 subject headings, which accounted for 1/3 of the total pages. On an average, each scientist found that about 7% of all the citations were relevant; whereas, under the headings he had a priori expected to include relevant citations, about 20% proved to be relevant. Of the relevant citations, some 20% were new to him.
A METODOLOGY FOR THE ANALYSIS OF INFORMATION SYSTEMS,
David R. Sparks, Mark W. Chodrow, and Dell M. Welsh
Information Dynamics Corporation, 80 Main St., Redwood, Mass. 01867
The management of scientific and technical information is of
fundamental importance as an element of the administration of
research and development. Questions of information system
design relating to user characteristics, costs, manpower, flow
patterns, etc., are, from a management standpoint, at least
equal in importance to the technical aspects of information
handling and its detailed organization for any specific pur-
pose. In the past, these problems of systems design have
largely been left unexplored by those active in documentation
and information sciences, or have received only general
empirical treatment. Modern scientific research, however,
is recognized as far too important an endeavor, and its
information needs far too complex to allow the development
of large-scale supporting information systems to proceed
without adequate and scientific planning and testing of alter-
atives.
The work reported in this paper is an initial effort to
assemble the elements of information system design as a prob-
lem in scientific management, and to propose a methodology
for its solution. The application of this methodology in-
volves the analysis of the features of science information
systems and the identification of the interrelationships
among them. Techniques for a structured mathematical rep-
resentation and quantification of these features and relation-
ships were developed and reduced to computer practice. Sim-
lated data were used to test the methodology. Developed from
current statistical sources and mathematical models of a
range of alternative information systems, data were con-
structed. The mathematical calculations in the models
were performed by computer, and the results analyzed, for
the purpose of effecting meaningful comparisons of system
alternatives.
This work was sponsored by the National Science Founda-
tion under contract NSF-C-263 and NSF-C-370.

COORDINATION OF INFORMATION SOURCES AND DOCUMENTS BY
MEANS OF RELATIONAL DATABASES BY
James W. Siegel, University of Wisconsin-Madison
Relational databases are used to store, retrieve, and
coordinate data from a variety of sources. The storage and
management of data in a relational database is simple and
efficient. However, when the data is to be used by a
variety of applications, the need arises to coordinate the
data with the applications. This can be done by the use
of a relational database management system. A relational
database management system is a software system that
provides a means for managing and coordinating data.

MATHMATICICAL ANALYSIS OF DOCUMENTATION
SYSTEMS: Dagobert Soergel, 78 Freiburg, Rotebueck 68,
Western Germany
As an attempt to a general structural theory of informa-
tion retrieval a document system (DS) is defined as a normal
system consisting of: a) a set (a) of objects (documents); b) a set
(A) of elementary attributes (key-words), from which a
further attributes may be constructed (A); c) a set (N) of names
connecting attributes with objects; from the names further
theorems may be constructed. The theorems may be con-
structed. Using the theorems, different mappings (A) + (a)
(set of documents retrieved) are defined.
The type of a DS depends on two basic decisions: 1) choice
of (M) M may consist of the logical constants "true" and
"false", or of some positive integers etc.; 2) choice of
the rules for the construction of attributes and theorems; e.g.
logical product in coordinate indexing (CI); links. Further
practical decisions: (A) hierarchical or not; kind of map-
ing, introduction of rules (i.e., further attributes). The most
simple case - ordinary two-valued CI - is discussed in de-
tail: (A) is a tree distributive (but not Boolean) lattice,
the homomorphic image a c of subsets of (a); a useful retrieval
operation "practernagation" is introduced. Furthermore
are discussed: a generalized definition of superimposed
coding; some functions for the distance of objects or attri-
butes, optimization and automatic creation of classifications.
The model may be extended to take into account semi-term-
and document-document-relations. It may serve as a struc-
tural frame in terms of which the functional probabilistic
properties of retrieval theory may be expressed more clearly.
Acknowledgments to Prof. Dr. J. Nitsche, Institute of
Applied Mathematics, University of Freiburg.
FACTORS AFFECTING INFORMATION DISSEMINATION IN SCIENCE AND INDUSTRY. LUI Sappi, and Vladimir Bojdaszewski, Institute for Technical and Economic Information, Bratislava, Slovakia.

When choosing the right organization, methods and means for information dissemination in science and industry, one has to decide first of what this importance this information is to the manager, scientist, .

The value of this information to any user should be determined by a formula containing all factors concerning the user's profession, branch of information activity, type of document, form of dissemination and the point of view of the information. This formula serves both for disseminating current information and individual research. Each factor is defined by a combination of circumstances existing and its value equals the value of C. The extremes of circumstances, the value of the exponent and on the basis of all factors concerning the user's profession, branch of information activity, type of document, form of dissemination and the point of view of the user, the value of the extent is determined. Each factor is defined by a combination of circumstances, the value of the exponent and on the basis of all factors concerning the user's profession, branch of information activity, type of document, form of dissemination and the point of view of the user, the value of the extent is determined.

Received a survey of various forms of disseminating scientific information and practical examples of how such information in management and creative processes.

FLEXIBILITY PLUS-
THE IBM 870 DOCUMENT WRITING SYSTEM
Karen G. Takele and Burton E. Lamien
Systems Development Division Library
IBM Corporation
San Jose, California

Some of the significant methods by which the IBM 870 Document Writing System can speed acquisition, processor distribution, reference and educational services in libraries are illustrated in this paper. The IBM 870 feature third the many advantages of IBM Equipment for a library. Each 870 system consists of operating units which can be interconnected for ease of use and for added advantages. Data entered into the system is transferred to other units and is available to the user's files in the form of documents either printed or in electronic form. The IBM 870 has been designed to meet the needs of the library and the documents it produces. It is not a complete library system but a component of the IBM System.
CONTRIBUTED ABSTRACTS

A computer-based system has been established for storing and retrieving the structures of chemical compounds along with information to describe the compounds. The system, called the Registry System, includes a set of molecular files, a molecular language description of the structural diagram, the molecular formula, accepted name/literature, and bibliographic references. Registration entries into the system include assigning a unique machine address (Registry Number) to each compound. The system enables the previously assigned Registry Number from the files. It is intended that the system ultimately includes a record of every chemical substance reported in the literature and identification of all useful published literature bearing on each substance.

Following information is necessary to define the Registry Number. (1) a unique machine address associated with each compound; (2) a unique machine address associated with each compound. It is intended that the system ultimately includes a record of every chemical substance reported in the literature and identification of all useful published literature bearing on each substance.

In addition to being assigned a Registry Number, a compound may also be assigned a name or its equivalent, which is treated as a subject to which the compound is related. This relationship is defined as the relation of a word to the concept it labels, following Hauser (Proc. of Amer. Soc. for Information Science, 1981, p. 21). Meaning is a tetradic relation of a concept to a degree, within a corpus. Indexers, in choosing or assigning all words strongly associated with concepts of a document, assert that the document means the word; in this reup, consistency of indexing measures the precision of which meaning is understood by the indexer. From an indexing experiment, designed to avoid all judgments of truth or correctness of indexing, the following results were obtained: 1) when descriptors are freely chosen, the more descriptors assigned to a document the more difficult is the retrieval of it. 2) By restricting the descriptors used in searching, the disadvantage noted in (1) can be overcome. 3) Possibly, older words are used more frequently and less precisely than newer words. 4) A graph of the total number of applications of a descriptor to the ranked document serial number usually illustrates the precision of meaning of the descriptor; a precisely understood word gives a rectangular curve, a less precisely one an S-shape. Many words have imprecise meanings, even to specialists in scientific fields.

COMPARISON OF CLASSIFICATION SYSTEMS IN THE FIELD OF CHEMICAL ENGINEERING, João Fernando Camargo Tavares, Núcleo de Documentação Técnica, Rio de Janeiro, Rio de Janeiro, Brazil.

After stressing that the problem of classification is basic in every information system, the author briefly mentions the most used traditional systems and the modern information retrieval methods. The classical UDC system and the automatic methods based on "key-words" operated by optical means or punched cards are described. Applying both systems to an instance, the author compares their efficiency, concluding that the advantages of UDC, vanish in fields requiring great detail. Finally, some conclusions are drawn which can influence the choice of classification system, in function of the following parameters: field of knowledge covered; number of references to be classified; extend of detail required; type and frequency of information demands.

IMPRESSION IN MEANING MEASURED BY INCONSISTENCY OF INDEXING, Dr. John F. Tingley, Exxon Research & Engineering Co., New York, N.Y.

Meaning is defined as the relevance of a word to the concept it labels, following Hauser (Proc. of Amer. Soc. for Information Science, 1981, p. 21). Meaning is a tetradic relation of a concept to a degree, within a corpus. Indexers, in choosing or assigning all words strongly associated with concepts of a document, assert that the document means the word; in this reup, consistency of indexing measures the precision of which meaning is understood by the indexer. From an indexing experiment, designed to avoid all judgments of truth or correctness of indexing, the following results were obtained: 1) when descriptors are freely chosen, the more descriptors assigned to a document the more difficult is the retrieval of it. 2) By restricting the descriptors used in searching, the disadvantage noted in (1) can be overcome. 3) Possibly, older words are used more frequently and less precisely than newer words. 4) A graph of the total number of applications of a descriptor to the ranked document serial number usually illustrates the precision of meaning of the descriptor; a precisely understood word gives a rectangular curve, a less precisely one an S-shape. Many words have imprecise meanings, even to specialists in scientific fields.

In Hungary Russian and English language abstracts are often the base of information work. The Scientific Institute for Technical Information is working on a project in order to develop a method combining the mechanization of several steps of information work with machine translation. The proposed system works as follows: The important words in the Russian/English abstract are punched into tapes and read into the computer. A Russian-Hungarian and an English-Hungarian dictionary of a special subject field are stored on magnetic disks and the significant words of the abstracts are matched against the dictionary. The computer assigns the Hungarian equivalent to the word found at the same time the synonym of the Hungarian word which is compared with the profile of the users in order to find the person waiting for information on a special subject. The user gets the bibliographic reference with all the significant words of the abstract translated into Hungarian. The data remains stored in the computer to form a document file that may be searched for documents containing specific information.


The bookseller has an important role to play in the organisation of information. The information must be at hand, ready to be consulted or understood; and for the procurement of material the librarian must rely upon the professional skill, experience and perseverance of the bookseller. Anybody can buy the latest item on a well-known publisher's list, but such work is required if the documentalist is to obtain for his use the cataloguer complete, the newspaper periodical, the unadvertised report. Bookseller standards vary from country to country and firm to firm, but documentalist everywhere require the highest standards.

These can only be attained in certain basic conditions are fulfilled. There must be fundamental willingness on the part of the bookseller to attempt to obtain any published item for which he received an order, and to tell and he must know and record the names, customs and idiosyncrasies of publishers of many lands. He must be prepared to train, and pay, staff to be professional booksellers, not just counter assistants; and the training of a good bookseller is an arduous and demanding task of any librarian. He must be prepared to disseminate information on publications, new and old, by traditional methods such as catalogues down the latest machine-recorded profile of subject interests of his customers. Booksellers should also be efficient business men; they need to be if they are to stay in business. And for his part the customer must accept that his bookseller is an expert working towards the same ends as himself, and not expect a first class service at a second class price. There is an increasing need for greater co-operation and understanding between librarians and their booksellers.

SYSTEMS CONCEPTS FOR ESTABLISHING AN AUTOMATED INFORMATION CENTER. W. J. Weglein, IBM Corporation, 555 Madison Avenue, New York, New York, U.S.A.

The introduction of automated information services at engineering indexes is used as a case history for discussing the systems concepts for establishing an automated information centers. The installation represents a joint effort by Engineering Index and the American Society for Metals in conjunction with Pathfinder Memorial Institute and the IBM Corporation. Recognition of pre-operational areas, and objectives are considered to be primary factors for a successful transition from manual to machine operations. The conversion of people and efforts to machine readable forms, and the aspects of program, program, and computer operations are presented. A generalized combined file uses a computer programming system for the IBM 1401 Data Processing System has been used to create a dictionary, a master file, and an inverted file. To perform retrospective searches. In addition a series of publishing programs utilizing the master file produce subject heading and author indexes. The dictionary is used for descriptor control, whereby input to the system is checked for validity and correctness. The master file contains the information that is pertinent to the documents which have been abstracted and indexed by a professional staff. The inverted file lists each descriptor followed by the numbers of the documents to which that descriptor has been assigned.
A STUDY OF SUBJECT ANALYSIS. Lib. Ranky, Institute for Technical and Economic Information, Konvikta 5, Praha 1, Czechoslovakia. A methodology of subject analysis has been elaborated for use in the Czechoslovak information services. General criteria have been suggested for acceptable abstracts, machine-made, and for other products of document analysis (selection of descriptors, subject indices). Abstracts are to be distinguished as to their purpose (a) for current awareness use and (b) for storage and retrieval (IR), rather than to their form (informative and indicative). Tools for GA abstracts the generally accepted informative form is recommended, special recommendations are set up for SR abstracts, which differ from GA abstracts in the basic approach, style, selection of data, type of documents analyzed, life, date of first use, classification requirements, processing, application attitude to formalization. In a conventional record file, SR abstracts may be considered as classification extended data into area the degree of non-linearity should be given for any point so as to enable the user to boost the abstract to the original size of the document. The SR abstract may be made with regard to the content of all preceding documents in the file and even documents expected to come. Any incoming information may be conditioned with information stored already, all redundant data should be rejected in order to save storage medium. Machine-made abstracts fail in these points. Certain types of information (e.g., statistical data, parameters, transactional matter) shall be extracted along special rules and stored separately for direct access. In formalizing abstract syntax and standardizing terminology we approach from an upper set of rules, a) to follow the language of the subject; b) to select different fields for different types of abstracts; c) to select different fields for different types of information; e) to select different fields for different types of users. The kind of abstracts is set by the purpose of the document. The subject analysis is performed on the basis of a catalogue of subjects, a) for current awareness use and (b) for storage and retrieval (IR).
CONTRIBUTED ABSTRACTS

INFORMATION STORAGE AND RETRIEVAL BY PRE-COORDINATION.

Eugene P. Willging, Catholic University of America Library, Washington, D.C., U.S.A.

A method by which all information stored in a system is permitted in anticipation of possible or organized associations of descriptors provides a printout ready for reproduction, with the answers to possible inquiries. As a result, one has at all times a complete inventory of the stored information, accessible from the various points of view from which this information may be considered.

Aside from the advantage of permitting, an inventory look-up for the subsequent search, the print-out opens an excellent browsing field guiding the inquirer to synonyms and associations of other useful terms which were not anticipated, without look-up in a thesaurus.

The printout is modular in regard to the various items of the bibliographic citation which can be included. The display of the punctation of descriptors and document numbers may vary according to the type of collection or the printout method selected. This printout may include a display of associated descriptors with all the document numbers common to those descriptors; the occurrence of pairs or multiple associations can be traced.

The specific descriptors assigned to each document or the abstract of this document may be included in the printout.

The bibliographical information may be included in the printout under each heading or may be printed out in a serial manner.

Conclusions: The method consists in a synoptic grouping of indexing and bibliographic information available on demand. Further refinement in association terms is possible; the technique used is flexible and can be adapted to the changing needs or the size of a document collection.

This work was performed while this author was associated with the Prevention and Treatment Center, National Academy of Sciences-National Research Council.

PROJECT CADER, CURRENT AWARENESS DOCUMENT RETRIEVAL FOR ENGINEERS: A STUDY IN PROGRESS AT ENGINEERING INDEX. Fred R. Cher, Battelle Memorial Institute, and Carolyn Flacker.

Engineering Index, 345 E. 47th Street, New York, N.Y., U.S.A.

Engineering Index has embarked on a program aimed to provide broader, deeper and faster abstracting and indexing services in all the engineering disciplines. As an initial step toward this goal, a pilot study is under way in the electrical electronics field and the plastics field. A monthly abstract bulletin has been launched in each of these fields as segments of Engineering Index. Eventually a collated index of all the segments will be available. Cumulative index will also be compiled. In addition, a deep index, listing terms and their roles in being constructed, will be the basis of a central mechanized document retrieval service. The present current awareness services of Engineering Index will be extended and refined so as to give selective dissemination of information to subscribers. These combined services constitute an important step toward the long-range goal of establishing an Engineering Information Center at the United Engineering Center in New York. The system is called CADER, for Current Awareness and Document Retrieval for Engineers. Journals are indexed and pertinent articles analyzed by professional personnel in accordance with the EIC system of abstracting and indexing. Their work sheets are processed using a computer for both the index print-out for the bulletin and for the machine search and I/D capability. The heart of the system is a their, which serves not only as a visual aid to indexers and inquirers, but also directs the following computer functions: (1) validate all index terms and subject headings; (2) substitute preferred terms and (3) generate in the published bulletin "see" or "see also" references for terms previously designated for this purpose in thesaurus. The basic computer programs were developed by IBM, but many special features have been added for this project. The EI thesaurus is based on the Thesaurus of Engineering Terms (EIT), but contains about 30% new terms. The American Society for Metals has adopted the same overall system, and a cooperative program between EI and ASM is in effect.

INTERNATIONAL BIBLIOGRAPHY OF THE SACRED SCIENCES.

Eugene P. Willging, Catholic University of America Library, Washington, D.C., U.S.A. 20001

Of more than 800 scholarly titles of serials pertaining to theology, church history, canon law, Scripture, and related disciplines, more than 90% are indexed fully or partially in duplicate (up to ten times) in 20 different bibliographic tools. To show extent of coverage and amount of duplication, a comprehensive list is scheduled for Sept. 1965 publication. As a by-product a plan is being developed to coordinate the indexing of books and serials in the sacred sciences in a new tool tentatively titled INTERNATIONAL BIBLIOGRAPHY OF THE SACRED SCIENCES, which would appear monthly under interdenominational auspices. Through inclusion of subject headings, classification numbers, multiple entries for editors, translators, etc. with the main author entry for books, it would attempt to solve a major library problem of cataloging backlists as well as provide theological scholars with quicker indexing as well as more intensive coverage and more frequent cumulations.

The United States has no established forestry classification. The Oxford system enjoys some use (Forestry). In the Forest Service began to evaluate Oxford's potential for literature control in research and operating units. A number of 20,000 microfiche citation cards was built. This took several years to complete. The validity of the system's claim that it permits the use of a special degree of preselection, flexibility, and capacity for growth was investigated.

A detailed study was initiated to determine the validity of the system's claim. A special degree of preselection, flexibility, and capacity for growth was investigated. The validity of preselection, flexibility, and capacity for growth was investigated.

The validity of the system's claim was also investigated when Oxford's scope is better understood. The validity of the system's claim was also investigated when Oxford's scope is better understood.

The ten thimble issues for July 1984 carried a simple questionnaire asking whether readers had found one or more articles of interest. The geographic distribution of responses corresponded closely to that of FASEB membership. Distribution by discipline showed that members of one of the six FASEB societies found less of interest in the other five (90%) than did members of the other five (90%).

A possible explanation may be the criteria for journal selection. Serials could not be scanned if they were being translated by others. Important biochemical and biological journals were thus excluded. Translated articles were screened by referee editors for novelty of material. Statistical validity, concepts, conclusions and citation of current literature were excluded. Translated articles were screened by referee editors for novelty of material. Statistical validity, concepts, conclusions and citation of current literature were excluded.

The availability of streamlined and definitive medical information on magnetic tape derived from the American Medical Association staff has prompted us to manipulate this tape in such fashion as to generate teaching sequences by direct similarity analysis of the work and phrase content of the information. Generation of branching tree of differential diagnostic sequences is followed by the insertion of general statement and question and answer sequences at predetermined iteration rates and with logical and work similarity cues. By this sequence, any body of work-related information can be processed to yield at least a rough draft of programmed text material, in any language. The saving in man hours using the IBM 7090 is of three orders of magnitude.
<table>
<thead>
<tr>
<th>Name</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sparks, D. L.</td>
<td>88</td>
</tr>
<tr>
<td>Spur, J.</td>
<td>88, 90</td>
</tr>
<tr>
<td>Storker, L. N.</td>
<td>59</td>
</tr>
<tr>
<td>Stobaugh, R. E.</td>
<td>90</td>
</tr>
<tr>
<td>Stone, P.</td>
<td>23</td>
</tr>
<tr>
<td>Suarez, R. J.</td>
<td>89</td>
</tr>
<tr>
<td>Summers, R. W.</td>
<td>89</td>
</tr>
<tr>
<td>Swanson, D. R.</td>
<td>15</td>
</tr>
<tr>
<td>Takic, K. G.</td>
<td>89</td>
</tr>
<tr>
<td>Lai, F. A.</td>
<td>90</td>
</tr>
<tr>
<td>Tauber, S. J.</td>
<td>7</td>
</tr>
<tr>
<td>Tavares, J.</td>
<td>90</td>
</tr>
<tr>
<td>Tchernow, A. B.</td>
<td>88</td>
</tr>
<tr>
<td>Thunnus, V.</td>
<td>88, 89</td>
</tr>
<tr>
<td>Tinker, J. T.</td>
<td>90</td>
</tr>
<tr>
<td>Teeman, J.</td>
<td>47</td>
</tr>
<tr>
<td>Tobiakiewicz A.</td>
<td>90</td>
</tr>
<tr>
<td>Tazi, B. V.</td>
<td>58</td>
</tr>
<tr>
<td>Vasecky, P.</td>
<td>1</td>
</tr>
<tr>
<td>Vass, W.</td>
<td>74</td>
</tr>
<tr>
<td>Vrankin, P. J.</td>
<td>74</td>
</tr>
<tr>
<td>Viste, G.</td>
<td>1</td>
</tr>
<tr>
<td>Washner, R. P.</td>
<td>51</td>
</tr>
<tr>
<td>Weisgerber, H.</td>
<td>91</td>
</tr>
<tr>
<td>Wsucki, A.</td>
<td>94</td>
</tr>
<tr>
<td>Yarke, T. B.</td>
<td>94</td>
</tr>
<tr>
<td>Young, G. A.</td>
<td>93</td>
</tr>
<tr>
<td>Zahn, H.</td>
<td>93</td>
</tr>
<tr>
<td>Zweitner, R. L.</td>
<td>91</td>
</tr>
<tr>
<td>T</td>
<td>V</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Tabor, K. C., 83</td>
<td>Vernier, P., 150</td>
</tr>
<tr>
<td>Eule, T. A., 90</td>
<td>Vesper, W., 151</td>
</tr>
<tr>
<td>Tredter, S. I.</td>
<td>Vinkel, P. J., 121</td>
</tr>
<tr>
<td>Tymmettes, J. E., 180</td>
<td>Vincens, V., 181</td>
</tr>
<tr>
<td>Traylor, R. S., 58</td>
<td>Tomich, A. B., 91</td>
</tr>
<tr>
<td>Trumpower, A. B., 88</td>
<td>Thumann, A. S., 95</td>
</tr>
<tr>
<td>Upaker, J., 100</td>
<td>Iune, H. E., 90</td>
</tr>
<tr>
<td>Toman, J. E.</td>
<td>Tekavec, A., 90</td>
</tr>
<tr>
<td>Turnar, P. A., 150</td>
<td>Tekavec, A., 90</td>
</tr>
<tr>
<td>Tuxen, F. C., 150</td>
<td>Tuxen, F. C., 150</td>
</tr>
<tr>
<td>Tymmettes, J. E., 180</td>
<td>Vincens, V., 181</td>
</tr>
<tr>
<td>Turnar, P. A., 150</td>
<td>Tuxen, F. C., 150</td>
</tr>
<tr>
<td>Tymmettes, J. E., 180</td>
<td>Vincens, V., 181</td>
</tr>
<tr>
<td>Turnar, P. A., 150</td>
<td>Tuxen, F. C., 150</td>
</tr>
<tr>
<td>Tymmettes, J. E., 180</td>
<td>Vincens, V., 181</td>
</tr>
<tr>
<td>Turnar, P. A., 150</td>
<td>Tuxen, F. C., 150</td>
</tr>
<tr>
<td>Tymmettes, J. E., 180</td>
<td>Vincens, V., 181</td>
</tr>
</tbody>
</table>

Y
Yerke, T. B., 94
Young, G. A., 93

Z
Zinsner, H. H., 94
Zwemer, R. L., 94