AGARD EXTENDED SUMMARIES 2

AGARD EXTENDED SUMMARIES 2

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92 NEUILLY SUR SEINE FRANCE

Storage and Retrieval of Information

A User-Supplier Dialogue

★

JUNE 1968

NORTH ATLANTIC TREATY ORGANIZATION

Extended Summaries 2

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

STORAGE AND RETRIEVAL OF INFORMATION

A User-Supplier Dialogue

The papers of which these are extended summaries were presented at the TIP-AvP Joint Symposium held at Munich 18-20 June 1968.

The material in this publication has been produced directly from copy supplied by each author.

Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St. London, W.1.

CONTENTS

* Not available for Extended Summary.

COMMUNICATION AND SECRECY IN SCIENCE

Rudolf Schrader

Modern science places an ever-increasing demand on the communication of its results, and one of the most serious problems of today is that of providing the professional worker with speedy access to the specialized knowledge in his field of interest. In fact, no research and development programme can nowadays be initiated and carried out with some hope for success, without having available good documentation. But there is also a great danger that science as such fragments into a number of independent disciplines which bear little relationship to each other, unless we are able to cope with the "information explosion". While advancement in the science and technology of documentation had been slow until relatively recently, with the mechanization and automation of a great many of its processes, progress is now fairly rapid. Against this background, the AGARD Panels for Avionics and Technical Information agreed to jointly organize a symposium devoted to the subject of "Storage and Retrieval of Information".

When the planning for the symposium began, it was noted that several conferences on much the same subject had been held during the last years, and the question was raised as to whether there was indeed a need for another meeting of this kind. In examining the programmes of these conferences, however, it was realized that, in the main, they had focussed attention on those problems which are of primary interest to the documentalist. In fact, none of these conferences had offered much opportunity to the scientist, as the user of scientific knowledge, to discuss his needs in the fields of information processing and dissemination. For this reason, both AGARD Panels eventually reached agreement to propose a programme which should fill this gap and should serve both professional groups in a joint meeting.

As one may see from the programme, the symposium intends to demonstrate how modern storage and retrieval concepts assist the documentalist in the handling of scientific information, and the scientists in audience are expected to express their views on a number of newly proposed designs and methods. But the scientists are also invited to the meeting as active contributers. I know that in a number of countries great efforts are being made to introduce the modern techniques of data handling into the traditional fields of documentation and that promising results are expected in the not too distant future, including automatic reading and language translation, based on the application of modern computers.

Hence, the success of the symposium will depend on the co-operation and joint participation of both professional groups. To this end, invitations for the symposium have gone out to all AGARD Panels and other scientists interested in the subject, and I note with great appreciation how many of them have come, in addition to the great number of documentation experts.

As I have been invited to represent at this meeting the scientist's point of view, I intend to discharge my task in devoting my speech to a discussion of communication and secrecy in science. In so doing, I recognize the important role documentation plays in the general process whereby science advances and our scientific knowledge grows.

Early History of Documentation

Documentation appears to be almost as old as man's interest in science and literature, and among the earliest collections of manuscripts were archives attached to temples and palaces. The great philosophers and scientists of antiquity were active collectors of volumes, and Aristotle was the first man known to possess a collection worthy of the name "library".

For many centuries, the library at Alexandria represented the intellectual centre of the ancient world. It is difficult to give a precise figure, but it can be said that the number of manuscripts collected from all parts of the world and available in the library was very large; however, it should be kept in mind that the papyrus roll of antiquity usually contained less written matter than a modern book. The catalogues and classified lists issued by the library were among the earliest experiments in bibliography.

Although the Romans made excellent use of technology, they did not particularly encourage scientific advancement. As a matter of fact, the Romans were not scientifically imaginative, whereas they demonstrated outstanding capabilities in such fields as jurisprudence, political and military affairs. Science, therefore, remained Greek in nature and spirit under Roman domination, and it was not until the last century of the republic that we hear of libraries in the capital. These libraries ceased to collect Greek writings when in the year 330 the capital was removed to the Bosphorus. Eventually, the aggressions and intrusions of the Germanic tribes swept away the classical learning from Italian soil and, with the fall of Rome in 476, the ancient history of documentation came to an end.

When Christian literature began to grow, libraries became part of the ecclesiastical organizations. The abbey of Monte Cassino founded in Italy about 529 was the first monastry in which a library of religious work was established, and this custom rapidly spread to all parts of the world. Monks began to write, and the use of vellum instead of papyrus resulted in the replacement of the ancient roll by the bound book as we know it today.

The Renaissance brought about a gradual broadening of man's intellectual horizon and an ever-growing desire for the collection of manuscripts and books outside the monastries. To satisfy this desire, large public libraries were created in France and Italy. When in 1453 Constantinople was conquered by the Turks, a full stream of Greek scholars began to flow into Western Europe, and the impact of Greek philosophy on Latin Christianity produced a powerful surge of intellectual activity and a keen interest in science. This development created a high demand for more and more books, and it is one of the miracles of history that the printing press was invented almost at the same time when Constantinople fell into the hands of the Turks. With the enormous increase of the scientific literature, new plans for libraries had to be developed, and the modern history of documentation began.

Communication and Scientific Research

Science and technology can only grow and blossom in an environment which provides for free communication between scientists and for the rapid transmission of scientific knowledge to the engineers who always strive for innovations and continuously wish to use new discoveries for practical purposes. Even in the fields of defence research and development, progress depends on the freedom of exchanging technical information. In fact, communication has always been a necessary part of the scientific process, and there is today everywhere in the world a growing awareness of its increasing importance for the advancement of science: the more the body of scientific knowledge grows, the greater is the need for communication, and the more communication is provided, the better are scientists able to carry out research.

Until the latter half of the 17th century, communication was done by way of direct correspondence whereby scientists kept each other informed about the work they were doing. At about this time, however, scientific journals began to appear and to render a much better method of communication. In addition, they offered an excellent means for publication and, in so doing, provided scientists with opportunities to proclaim their discoveries to the world, in an effort to get in first.

Today, the vast growth of scientific activity makes it difficult for the individual scientist to keep up with the ever-increasing number of publications. So much is nowadays published that he is completely incapable of studying all papers which could be of some help to him in his own field of research. This situation is often referred to as the "crisis in communication", and I share the hope that in the not too distant future a modern computer be designed whereby the steadily growing flood of new scientific information is indexed, processed and assembled, in order to get us over all the difficulties of this crisis.

But whatever computers will be capable of doing in calling a scientist's attention to a piece of scientific information which might be critical to his work, they will never be able to distinguish qualitatively between a good paper and a bad paper, neither will they answer those crucial questions which scientists are used to ask. For this reason and many others, open discussions among scientists working in the same field - scientists who correspond regularly with each other and meet repeatedly at conferences - will continue to play a major part in the communication process and hence in the advancement of science.

When in former times the number of scientists was small, a scientific discovery was usually the product of one single mind and emerged at one particular moment. Today, scientists are counted in hundreds and thousands, and we are likely to find that, at any one moment, a good many of them are engaged in the same piece of research and are about to publish much the same results. As a matter of fact, there are almost always several laboratories in the world which, in the pursuit and exploitation of new scientific ideas, move along the same lines of approach, and it is well remembered how much research was done in parallel during the Second World War, when major parts of science were regulated by the demands of military security and scientific communication was almost non-existing.

Imposing Secrecy on Science

In war, but also in peace, scientific information of direct military impact must be controlled by security regulations in order to prevent its premature leakage to an enemy. Hence this information must be withheld from the traditional channels of scientific communication by some sort of classification, and nobody, of course, would question the need that in the interest of national defence certain kinds of scientific information must be protected. In fact, military secrecy is necessary and often vital to our survival. But in a real sense, it is bad for science.

Classified research is almost always in danger of suffering in quality, as this kind of work is not exposed to scientific criticism to the same extent as research, which is done in an open laboratory. Moreover, secrecy prevents free discussions which are so important for scientific progress. Secrecy furthermore means that large parts of the scientific body are kept in ignorance, at least for some time. It should be added that scientists carrying out research under condition of secrecy are less likely to enjoy the overwhelming wealth of scientific ideas as their colleagues working in a free and open environment. And as it frequently happens that a scientist fails to realize the significance of his own findings, it becomes all the more important that scientific observations be made available to others for further studies.

Indeed, secrecy and great scientific thoughts cannot thrive together. Nevertheless, secrecy is necessary and is often applied to scientific research for reason of national security.

But secrecy also plays an important role in industrial research. It is difficult to think of a firm which devotes substantial funds to research and reveals scientific information gained at great expense to its competitors. However, it seldom happened in the past that the leakage of scientific knowledge caused a real loss to a company. On the contrary, the free release of basic information has in almost all cases been an advantage to all firms working in the same field of manufacture.

Secret scientific information often continues to remain classified, even if it has lost its military values almost entirely and consequently does not become available to scientists as rapidly as that of the open literature. It is therefore necessary that methods be introduced whereby scientific information will not remain classified long after its secrecy has ceased to have military significance. In my view, the proposal that some positive action should be required to maintain classification after a certain length of time merits consideration. Today, action needs to be taken in almost all cases in order to obtain de-classification, since in present security procedures classification is heavily favoured over de-classification. The access to classified information is generally regulated by the criterion of "need-to-know". While this criterion can be easily applied to information of tactical military value, it can hardly be used for information of scientific content, as nobody in this world is able to say precisely in advance which piece of scientific information may or may not be of benefit to a particular research project. Because of a narrow interpretation of the need-to-know criterion, it so happened many times in the past that an information which was not available to a scientist at the right moment has made the difference between the success and the failure of a research project. It is therefore very important to ensure that the criterion of need-to-know be intelligently used and not be allowed to hamper the free flow of scientific information. But it is also necessary to set up some form of special information service whereby those scientists who are entitled to see classified information are aware of its existence in order to avoid costly and wasteful duplication in defence research to as great an extent as possible.

Classification practices adopted by most countries overlook the fact that a scientific discovery made by one scientist can never be kept secret for any length of time, as one day, sooner rather than later, the same scientific discovery will be made by another scientist. Whenever this happened in the past to a piece of scientific information which for security reasons was classified and consequently not published in the usual way, it was in almost all cases the popular belief that the information had been stolen by espionage. In reality, however, the information was rediscovered, as nature and her laws are open to every intelligent mind throughout the entire world, and no nation may claim a monopoly to be the only one with a capability of producing scientific ideas.

What do we gain by imposing secrecy on scientific research? Obviously, the main prize is time; and this is probably all we ever gain in any scientific field, as we may expect to prolong the time it takes our potential enemies or competitors to learn what we already know.

In fact, secrecy does not play a useful part in science, and security regulations should never be applied to pure research and broad fields of basic knowledge. Although the withholding of fundamental scientific information may occasionally provide short-term military advantages, in general it is detrimental to scientific progress and for this reason should always be avoided.

NATO's Interest in Scientific Communication

Within the Atlantic Alliance, the AGARD Technical Information Panel plays quite an exceptional part in the sense that it is the only body of NATO which continuously deals with documentation and its various problems. Established by AGARD in 1953, the Panel has launched a broad programme in the fields of aeronautical publications, but has also served NATO in other fields of documentation, whenever called upon.

At the request of the NATO Science Committee, the Technical Information Panel undertook in late 1958 to study ways and means whereby the exchange of scientific information within the Atlantic Alliance could be improved. As a result of these studies, the Panel recommended in March 1959 that a documentation liaison unit be established, which should not perform functions usually provided by an efficient documentation centre but should keep in touch with research and development activities and supply information to those scientists in the NATO countries who are in need of this information. Initially, this recommendation met with great enthusiasm, but was later-on considered difficult to implement, as a documentation liaison unit of the kind envisaged by the AGARD Panel would be of little use in the case of classified information which by necessity must be exchanged between nations on the basis of bilateral agreements. Consequently, the idea of a documentation liaison unit was dropped, with the understanding however that the subject of improved information exchange among the NATO countries be studied along other lines.

In 1962 the Technical Information Panel submitted to the Science Committee a report which dealt with the exchange of scientific information in the defence field. This report pointed to the need that there should be at least one defence documentation centre in each member country of NATO and that countries without such centres should, as a matter of urgency, initiate their establishment.

All defence documentation centres should be equipped with the most efficient information-handling techniques and should endeavour to arrange the automatic release of all unclassified information to the corresponding centres of the other countries. As far as the exchange of classified scientific information is concerned, the report recommended that countries having mutual interests in a given field make bilateral arrangements. But pending further improvements and as an initial step, lists of unclassified titles of classified information should be given widespread distribution.

The report of the Technical Information Panel was approved by the North Atlantic Council in June of 1963, and the governments of the member countries were subsequently invited to take such action as they deemed necessary for the implementation of the recommendations. Recognizing the significance of the issue, the Council furthermore agreed that work in the field of documentation should continue, as indeed the communication of scientific knowledge within the Atlantic Alliance is a problem area of greatest importance to NATO.

In this discussion, I have referred twice to the NATO Science Committee and its interest in documentation. Now, I should like to call your attention to the Committee's report on "Increasing the Effectiveness of Western Science", issued in the Autumn of 1960. Actually, the report was written by a special study group set up by the Committee, but it reflects the Committee's attitude towards the need for accelerated scientific progress in the Western world by means of enhanced international co-operation. As one may expect, the report deals also with documentation, and it appears to me to be of some significance in this context to mention briefly the main recommendations, as they describe the scientific demands for improved methods in publication and documentation in an appropriate way.

The recommendations, as they are listed in the Committee's report, call for closer co-operation in publication practices between the chief editors of the main scientific journals; the proliferation of scientific journals should be discouraged; the activities of all documentation centres should be co-ordinated, and a single international system of indexing should be introduced; authors of scientific publications should be invited to supply, together with their papers, abstracts which are to be edited according to specific rules and to be classified in accordance with a single and unified system; the abstracts should be published immediately; experienced scientists should be encouraged to periodically review broad fields of scientific research and to summarize their results in an efficient way.

These are the main recommendations made by the NATO Science Committee when dealing with the problems of documentation some years ago. They demonstrate the importance the Committee attaches to these problems. Needless to say, these recommendations are still as valid today as at the time they were written.

Science and Documentation

The critical review of wide research fields periodically done by competent scientists serves a valuable purpose in summerizing large portions of the available scientific literature. In my view, this kind of work should be given more credit than in the past. If wisely done, it will certainly assist us in our desire to better cope with the ever-increasing volume of scientific publications.

As the selection and presentation of scientific information can only be carried out intelligently by those who originate the knowledge, scientists must nowadays devote greater efforts to these activities than heretofore. As a matter of fact, research can no longer be regarded as being completely separated from the communication of its results, and scientists must nowadays join the professional documentalists and accept responsibilities for the transmission of scientific information to the same extent to which they bear responsibility for research.

Scientists often produce a certain amount of redundancy when they publish their work, and it frequently happens that they issue for the same bit of research several reports all of which seem identical to the documentalist. There are reasons for this, one being that this is the way whereby scientists make their work known and consequently stand a better chance of gaining prestige and stature in the world of science. The essential point about publication is that scientists should always be conscious of the right to publish, and, in my view, they should take full advantage of this right, as they have indeed something to say to the world worth consideration, whether it is of pure character in the sense that it does not relate as yet to some known field of exploitation or whether it is of a more practical nature in a given area where the possibilities of material application are already well recognized. Nevertheless, the subjection to a kind of self-control practiced by the authors of scientific reports will certainly help the documentalists to overcome at least some of their difficulties.

There is another point which is critical for a fruitful co-operation between the scientist and the professional documentalist. How should our young scientists be trained to make better use of existing documentation facilities? When I was a student years ago, very little was provided in documentation training, and we first became aware of the existence of and acquainted with the many problems involved when we started to do laboratory work on our own. As by that time we did not know how to handle documentation at all, each of us began to develop his own method, so to speak, and I believe that the lack of proper guidance in this field is often the cause for so many scientists today to prefer the setting up of their own documentation rather than to rely on the official information services. There is certainly a great need for improvement, and I sincerely hope that in the future universities will offer their students better training in documentation.

In conclusion, I should like to say that in my judgement the key to the solutions of the great many documentation problems lies in a kind of co-operative approach, involving the symbiotic activities of scientists and documentalists. Both professional groups must work together in an atmosphere of mutual respect and appreciation, and it seems to be very important to enhance their interplay. I trust, this symposium will contribute to this co-operation and will set an example for many more meetings of this type to be held in the years to come.

The Suppliers Point of View

Introductory Paper by H.F. Vessey, Chairman, T.I.P.

1. General

In considering the need for this symposium, it was pointed out that in the last 18 months there have been more than six International symposia on much the same subject. However, on examining the programmes, it became clear that, in the main, they were of primary interest either to the documentalist or to the computer operators and that the real customer, the Scientist or Engineer has not been allowed much time. It is hoped that this symposium will fill the gap. The papers have been designed to review the present services, to set down what may be possible in the near future and to allow our customers to say what is needed. It is hoped that the Scientists and Engineers present will take their full part in the discussion and that those computer engineers and documentalists speaking will concentrate on the "Service to Customers" rather than lose us in the thickets of computer programming or the forests of Indexing.

The following short notes are intended to form not only a background to the specialist papers which are necessarily limited in number and coverage but also to outline the task, the problems and the equipment available to the supplier, that is the Documentalist or Information Officer.

2. N.A.T.O. Defence Documentation Centres

The task of the National Defence Centre in a N.A.T.O. country is to make known and supply to the country's scientists and engineers the unpublished and sometimes published literature in their field.

This is done in several ways:
(a) By supplying to individual scientists the new literature in their fields.
(b) By issuing accession lists, or,
(c) By preparing Abstract Journals of new material.
(d) By supplying relevant reports in response to an enquiry.

Further, the National Defence Documentation Centre has the responsibility of making known, in other countries, the work of its own scientists and of arranging exchange agreements for foreign literature.

3. Other Centres

In N.A.T.O. countries there are a number of documentation units apart from the National Defence Documentation Centres. Each large research establishment will have its own library or Information Agency and there will be the civil technical libraries attached to Universities, Research Associations, etc. Methods of operation will vary but the basic requirements as previously described apply but with a modified emphasis. It is highly desirable that there shall be collaboration between all these areas of documentation.

4. Type of Agency

The types of Agency may be described under four main headings although there is usually considerable overlap of activities particularly in the first two.

Libraries

Libraries deal with published information in the form of books, periodicals etc. In most cases the records used for retrieval consist of the bibliographic information (Title, Author and Publisher) and a broad statement of the main subject.

Documentation Centres

The main task is the collection, announcement and dissemination of unpublished report literature although published literature may also be processed. The records are usually kept in a more detailed form and most Centres will make subject searches, prepare bibliographies etc.

Information Analysis Centres*

These work in a specialised field, have direct access to working scientists in that field and will provide advice and evaluation in addition to answering detail subject enquiries.

Referral Centres

Referral centres accept enquiries and refer them to the organisation best fitted to reply.

National Defence Documentation Centres fall within the second category but may extend into the other fields.

5. Mechanisation

Mechanisation is increasing rapidly and many Agencies are now using computers in their operation. We shall hear of some of these in the later papers but a few words on the need to mechanise are appropriate here.

Many documentary units are running very efficiently on manual systems and to maintain balance we shall hear of the operations of one or two. There is a very real danger that a manual system that is not operating satisfactorily will be automated to improve its efficiency. The result is unlikely to be satisfactory for a poor manual system will, when put on a computer, be even less efficient and will cost ten times as much.

In an organisation the case for the computer generally rests not on retrieval but on the other operations such as the preparation of Abstract Journals, Indexes, a considerable number of Bibliographies or other such lists and finally on "House-Keeping". House-Keeping will include circulation control, "security" check, stock control and other similar functions. Retrieval will certainly be done, using the information fed in for the other operations but this alone is unlikely to be financially attractive unless the operation is very large indeed or is a "Selective Dissemination of Information" (S.D.I.) operation with a considerable number of customers (S.D.I. of course is a specialised retrieval search).

6. Indexing

Indexing is one of the most difficult problems in documentation for unless a paper is competently indexed it is lost in the system. For some years subject specialists will be required in documentation centres for this work. Good men are scarce and expensive but as they are also required for abstracting and circulation recommendation the indexing is but a small addition to the load. As abstracting becomes less necessary and circulation is taken over by S.D.I. the pressure to develop mechanised indexing will increase.

7. Newer Developments

A number of the newer developments deserve mention for not all will be known by the large range of customers of documentalists.

Information Analysis Centres

Information Analysis Centres have grown up to meet the needs for detail advice in specialist fields. We shall hear of the working of several of these and it is sufficient here to say that they are usually based on a University or Establishment working in a specialised field.

Referral Centres

Referral Centres have rationalised the practice of most documentary units of referring some enquiries to other agencies better able to assist. They accept enquiries, transfer them to the Unit working in the appropriate field and monitor the results.

Selective Dissemination of Information

Selective Dissemination of Information is not new and a large number of Libraries and Information Units keep a "field-of-interest" register which is used to route the

* Sometimes called Specialised Information Centres

new material. However, S.D.I. is the term now applied to the large operations which are now possible by computer of which we shall hear. Basically S.D.I. is a retrieval operation but the questions asked of the computer are "profiles" of each customer defining his interests. The "profiles" have to be built up by allocating index terms so that the search picks out the reports likely to be of interest to the particular scientist. Difficulties arise in establishing and, in particular, keeping profiles up to date and experience seems to indicate that the profile of a group of scientists produces more satisfactory results than individual profiles.

Citation Indexing

Citation Indexing is comparatively new and is a computer operation producing an index of all papers in which reference is made to one older publication. Thus a scientist knowing one authority in the field in which he is interested can turn up the author and title of all later papers which give the original report as references. The process is expensive and time consuming for not all references will be on the main subject but in some cases it will produce papers missed by other searches. It is certainly an additional bibliographic tool which in some cases will be very valuable indeed. However, despite some claims, it cannot replace other methods of search.

Microform

Photography has been used for many years in relation to storage of documents. Records on film, 35 mm, 16 mm or cut film have been made to duplicate records for security or for easy transmission. Some systems too have been developed for storage and retrieval. The important recent development has been the widespread use of microfiche particularly in U.S.A. The standard American fiche is a sheet of cut film 6" x 4" containing 60 images of approximately A4 size documents but the older European systems generally used smaller fiche. Microfiche have the advantages of ease of storage, postage and reading with relatively cheap equipment. At present however, particularly in Europe, printing hard copy from them is difficult and expensive and there appears to be more resistance from scientists to using them in a reader rather than asking for hard copy. The microfiche images are at a scale of 1-18 or 21 but much greater reductions are now being demonstrated. Any documentary centre of reasonable size should have facilities for reading film and microfiche and for taking paper copies of selected pages, that is they should have reader printers. The cost of the printing of hard copy and the preparation of microfiche is still high and these tasks, for some time, will probably be confined to a few centres in each country.

8. The Future

We shall hear of several new proposals which will influence the future and the author will, therefore, merely set down those developments which he considers of major importance.

Preparation of Reports on Tape

Apart from speeding up the issue of reports which require re-drafting the final tape can be fed directly to a computer for documentary processing. At the present time an expensive process of editing and key punching is required to extract the bibliographic data and present it to the computer.

Text Reading

Although more expensive than the above, machine reading of text typed on a standard form in special type face will show advantages over key punching for those reports where tape is not available.

Time Sharing

Time sharing on computers by making use of "dead" computing time occupied by say print out allows a number of unconnected tasks to be done almost simultaneously. As a further extension tasks may be fed to the computer by telephone wires from a number of remote terminals and the coded tasks stored until the computation time is available. We shall hear of one or two of the documentary applications.

Display Techniques

Display techniques are advancing rapidly and it is already possible to envisage the remote display of, say, abstracts in response to a subject enquiry.

Rapid Printing

On a more mundane level there is a very real need for economic printing at a speed compatible with computer speeds. This and display techniques will no doubt progress together.

Finally, it should be noted that for many years there will still be the need for local information centres without much mechanisation, although there will probably be a tendency for them to rely more and more on the large fully mechanised central stores.

9. Education

The need for education of Information Officers or Librarians in the operations and in the customers' needs is now fully recognised and the recommendations are now being implemented in most countries. Less well recognised is the need to educate the customers in the many sources of information that are now available and the best ways of using them. The author being a scientist (and engineer) converted to documentalist is acutely aware of the deficiencies on both sides and hopes that the talks on this subject will do something to remove the misunderstanding which still exist.

The Information Officer is anxious to assist the scientist but his task is made easier and the result will be better if the scientist is aware of the documentary problems. Some specific recommendations are:-

(a) In report writing use an informative title and include an abstract of some 100-200 words which fully describes the report.

(b) In making requests for documents be as specific as possible. Thus where the request arises from a reference, quote originator, Report number and say author. This gives sufficient redundancy to allow for a check against, say, transposition of numbers. If the information available is incomplete, give everything there is and say "that is all available".

(c) In asking for a subject search, be as specific as possible and give an indication of whether the search can be limited by date. If it is to be a detailed search for obscure information give some advice on the type of report which may contain it.

In case it should be felt that these recommendations are too elementary it should be said that in the author's organisation some 15% of requests are incomplete and either require a search on inadequate data or reference back. The delay, nuisance value and increase of costs is quite considerable. Again, T.I.P. was asked to prepare a bibliography on "Brittle Materials" and it was not until the work had started that the author, by approaching the British Panel Member, ascertained that interest was only in the Ceramics.

FOUR "NEW" SCIENCES: AN APPROACH TO COMPLEXITY

Edward B. Montgomery
Dean, School of Library Science
Syracuse University
Syracuse, New York 13210, U.S.A.

An Extended Summary

We are faced with a growing complexity which has been a by-product of our progress, increased knowledge and our vast technological achievements. Like air pollution, we must learn to cope with this by-product before it leads to our destruction. And we must learn to cope with it while we acknowledge the need for still further research and development.

Research and development are being mass produced and automated, and will increase. Increases such as these are of an exponential nature and have the same effect as a chain reaction. A change in one discipline effects changes in other disciplines which, in turn, may affect still others. Some of these chain reactions are short-lived. Others, however, continue branching into many disciplines having large reproduction factors and short reproduction cycles making synthesis of new information very difficult.

This paper suggests a new approach to complexity and it is hoped that it will effect progress rather than confusion.

This approach suggests organization and synthesis of four "new" sciences which will allow a more certain comprehension of interrelationships of knowledge leading to such phenomena as overpopulation and pollution.

The sciences which need development deal with information, communications, systems and applications. None of these are totally new; they are being implemented in varying degrees and in various combinations at present. Much of the application of these fields is, however, on a small scale and is accomplished through societies for cybernetics, general systems, etc. It is recommended that these sciences and their obvious parallel engineering fields be explored, structured and pursued with increased breadth and support.

The interrelationships resulting from study and research are rather unilateral and are generally confined to their parent discipline. Still, there is greater realization of the potential of these interrelationships outside the discipline giving rise to them.

If we postulate the four logical scientific domains of knowledge that provide a continuum from information to its ultimate use, and further require that these domains cross all disciplines at right angles to them, we should add the now missing dimension to our ability to understand the potential interrelationships of knowledge.

Information science is the first science we must consider. To see it in terms of its many contexts, we must consider the broadest possible definition of information. It does not matter if this definition agrees with a dictionary; information exists everywhere. We could define information as the position of all the atoms and molecules in the universe and of all sets and combinations of those atoms and molecules at any time.

Consider next how information operates--as a result of man's knowledge and reflections and as a recurring natural process. Information exists within all of the organisms found in nature. Every cell of every organism contains molecular information which defines and has determined the complete description, make-up and functional specification of that organism.

When organisms reproduce, information about the one or two organisms giving rise to the reproduced organism, is tramsmitted in the reproduction process and a new piece of molecular information is created in the process.

On the highest level of complexity, man is acted upon not only by his environment and his reaction to it, but reacts to his own feelings and thoughts.

As a result of these complex interactions, organisms contain information transmitted frequently about the state of the organism and of the sub-organisms or systems within the organism. For example, one's nervous system is under the constant surveillance of the human body's sensing of the change of state of the organs, the change of state of emotions and the change of state of thinking.

Information is also produced by the activities of organisms. Humans observe, design experiments, create, think, feed and indulge in various activities which increase the amount of information. In a scholarly discipline, this increase in information is sought after in an organized and logical fashion.

Information itself is not sufficiently investigated as a scientific entity. What information is, its roles, its contribution to the disciplines is too frequently a by-product of disciplines with other goals than an understanding of information per se. Information, its dynamics, the tools that elicit it and with which we use it deserve a special science that can lead us, through better understanding, to a far better use.

The consideration of a science of information leads to the need for additional sciences to cover the spectrum of this investigation.

Information in itself, however, is of little use. It must be communicated. We must search scientifically for answers to such questions as: What is communication? How does it take place? Why does it occur? What are its purposes? and What forms of communication are there?

Information, communicated through a system, usually has a purpose or application. Thus, a science of applications following a science of systems is needed to provide full coverage of the information continuum.

Complete understanding then, requires a knowledge of: what information is, how it is communicated, the systems used in transmittal, and ultimately, its application.

External information and communication actions of organisms include communication on a personal basis, on societal, cultural, business, international and many other bases. Among these, too, there are the needs for a better understanding of what communication is, what a system is and why it operates or is applied.

The four sciences should not be limited to the interface with participation in the physical, biological and social sciences. These sciences of information should transcend and tie together all activities of man, including all of the disciplines, from art and architecture to zoological systems.

Because the conferees here are involved in communications and systems work of all kinds, it is unnecessary to comment further about communications and systems as broad scientific disciplines. However, we do need to know what communication is and what a system is from a broad context so that we understand their general properties and dynamics and can anticipate their by-products.

The fourth science dealing with applications should be more thoroughly explained. For example, how does man apply the knowledge or the information that he has? Most of the endeavors in our society pertain to the use of knowledge, yet one of the least perfect areas of scientific knowledge is applications. A science of applications would embrace studies leading to understanding of how man really uses information. This could lead to vastly improved functions of all sorts. For example, education is the application--through a system of communication--of the knowledge and the information man has. The end product of education is the application of knowledge to all phases of learning--further scholarship, a professional career or simply a fuller, better life. Yet, the entire process of application is not very well understood. Do we really know what learning is? The systems called schools, although they are improving, are far from attaining achievement of the best educational objectives because we lack knowledge of application.

Applications knowledge also has to do with the arts, and while the present state of the arts may seem productive and satisfactory, one facet of art is in the ability to apply, but what is applied and how it is applied are not very well understood.

Our list of examples of the need to study applications should also include questions of how man applies his knowledge of political, economic, and social sciences to the relationships among nations. Here again, precision of application knowledge, in fact what knowledge to apply, based on what information, is an enormous problem man has failed to solve over the centuries. Now that some sciences have succeeded in applying their knowledge to the creation of devices with the capability to eliminate man from the face of the earth, our inability to discover, to communicate, to construct fail-safe systems of society lead to the realization that our application of man's knowledge needs even greater depth and surer progress.

A fifth science, which is not a new science at all, might be called the science of science itself. Science can be understood so broadly that its contributions are not used for purposes that could harm or even annihilate mankind. Scientists and technologists should provide fail-safe capabilities for their creations, the ability to anticipate undesirable by-products as well as leadership to eliminate those problems of the world for which man has been responsible.

This may seem naive and impractical to some. However, continuation of the present exploitation of science and technology is truly naive.

Now, why is it that we speak of <u>new</u> sciences and of <u>new</u> engineering areas? We are aware that most of the activities we have discussed are subsumed under existing disciplines such as linguistics, information theory, systems science and other current activities of scientific and technological groups seeking to unite their efforts to promote more general studies of these subjects.

The reasons why we speak about them as new sciences are: Information has dynamic properties. That the pen is mightier than the sword is an old manifestation of the power of information. Yet relatively little effort has been made to really understand information dynamics in terms of our present needs. However, it is quite probable that the problem is now growing at a rate beyond the capabilities of professional societies. The dynamics of information, its communication and its ultimate application in the various systems that exist or that have been created by man are the <u>dynamics of chain reactions</u>.

We must create a synthesis that can include all uses of information. We must plan for the most effective understanding of the good and the bad effects so that we can begin to cope with the complexity that we are creating with the increasing reactivity of chain reactions.

Out of such new organization of disciplines may come new horizons, new approaches, new concepts and new capabilities that are needed for considered growth and implementation in research and development of all fields.

If we consider the field of information science, it is quite probable that a group working to develop information in one discipline may find resources, methods and devices that will be useful in information-seeking in other fields. Approaches and instruments that have been developed in scientific research seldom stay in the narrow field for which they were created.

New principles, points of view, philosophies, definitions and specifications concerning information will be useful in many, if not all, fields. The exploratory work on information in one field should provide insight, impetus and working tools for others fields. This should offer lower costs in terms of manpower and money. Of even more value than lowered cost in times of great need, such as the present, there will be less time elapsed from initiation to broad use.

Similar statements can be made about communication science, systems science and the science of applications. Understanding of systems that resulted from systems analysis, beginning with the aerospace field, has led to new approaches in the many fields including the behavioral and life sciences. On the other hand, as more understanding of the operation and communication of information in the organs of animal organisms is achieved, new insights into social systems will evolve. It is even possible that the fail-safe philosophy developed for the operation of nuclear reactors could be translated into a method of international cooperation.

The repertoire of illustrations of the possibilities of use of these sciences is almost endless and will be limited only by the imagination and creativity of the disciplines involved.

Research and development have reached a stage where the unit cost of knowledge and productivity is increasing exponentially. An even greater exponential applies to the need for increased research. Each new unit of knowledge has an inherent interrelationship with another field. Many of these change those fields so that they, in transmitting their changed states, create new possibilities of interchange in still other fields. So while we are creating more and more knowledge, we are also creating more by-products of that knowledge. This can be considered as an exponential product of the exponential growths.

The ability of mankind to cope with the complexity thus created seems to lag farther and farther behind the creation of new interactions. We cannot slow down, we must increase our research efforts as well as our reflection on its results.

We need to increase the ability to reflect, to think about, as well as the ability to create more research at a lowered unit cost. The mass production of research, in fact, the achievement of its automation, depends on new concepts and new directions.

Hopefully, these new approaches provide a matrix for coping with the problem of rising complexity.

AGARD SYMPOSIUM ON STORAGE AND RETRIEVAL OF INFORMATION

Münich 18/20 juin 1968

Les problèmes posés par le vocabulaire documentaire et l'organisation des dictionnaires et thésaurus

F. LEVERY - Ingénieur Conseil à la Cie IBM FRANCE

Quel que soit le système documentaire utilisé (méthode de classification traditionnelle hiérarchisée, indexage à l'aide de mots-clés ou de descripteurs), on suppose que le même langage documentaire permet, à la fois de signifier correctement des documents et des demandes de documentation.

Le principe général consiste à considérer qu'une demande peut être rédigée en décrivant le document espéré.

Cette hypothèse suppose, en fait, que le demandeur possède au départ une bonne connaissance du sujet sur lequel porte son interrogation. Dans le cas contraire, il est obligé de faire appel à une personne spécialisée capable de "traduire" sa question d'une manière correcte.

Il est à noter que cette difficulté était pratiquement surmontée dans tous les systèmes traditionnels utilisant une classification hiérarchisée.

En effet, le système classificatoire servait de guide au demandeur et lui apportait les informations nécessaires pour poser correctement sa question. L'organisation hiérarchique permettait par ailleurs de fixer facilement le niveau de généralité désiré dans les réponses.

Il est certain que ces facilités n'existent plus dans les systèmes documentaires qui utilisent le langage naturel pour constituer les fichiers de recherche.

L'expérience montre que le langage utilisé par les demandeurs (langage du non-informé) est très sensiblement différent du langage utilisé par les auteurs (langage de l'informé).

La comparaison brutale entre la liste des termes utilisés par un demandeur et la liste des termes utilisés par un auteur risque d'être insuffisante et un grand nombre de documents pertinents peuvent échapper à la sélection.

C'est la raison pour laquelle il convient d'établir des dictionnaires (ou thésaurus) permettant d'assurer les échanges entre le vocabulaire du non-informé et celui de l'informé. Ces dictionnaires peuvent en fait être utilisés de deux manières différentes :

- ou bien au moment de la constitution du fichier ; il s'agit alors de rechercher, grâce au dictionnaire, l'ensemble des notions susceptibles d'apparaître dans une demande pour laquelle le document considéré serait pertinent
- ou bien au moment de la sélection documentaire ; il s'agit alors de rechercher, grâce au dictionnaire, l'ensemble des notions qui ont pu être utilisées par des auteurs et qui se cachent implicitement derrière les termes utilisés par le demandeur (termes synonymes, hiérarchiques c'est à dire plus spécifiques ou plus génériques et termes sémantiquement voisins)..

Que ces dictionnaires ou thésaurus soient utilisés à l'entrée ou à la sortie, leur forme ne varie guère. Le choix entre ces deux modes d'utilisation dépend en fait de facteurs pratiques .

Si le nombre de documents est grand vis à vis du nombre de demandes, il y aura intérêt à utiliser le thésaurus à la sortie (fichier rétrospectif peu interrogé), si au contraire, le nombre des demandes est important pour un faible fichier, il y aura intérêt à utiliser le thésaurus dès la mise en mémoire.

La constitution des thésaurus peut être envisagée de plusieurs manières différentes :

- soit à priori et ceci est obligatoire lorsque le thésaurus intervient au cours de l'indexage

- soit à postériori lorsque le fichier documentaire générateur du lexique est constitué.

Cette dernière solution n'est envisageable que lorsque le thésaurus est utilisé à la sortie.

Plusieurs expériences ont permis de dégager des méthodes pratiques de constitution de thésaurus :

- récupération des associations créées pour l'interrogation des fichiers
- comparaison des définitions
- exploitation des structures documentaires hiérarchisées etc..

Ces différentes solutions seront commentées au cours de l'exposé.

MECHANICAL READING

TRENDS AND DEVELOPMENTS IN CHARACTER AND PATTERN RECOGNITION

Lawrence A. Feidelman

1. Introduction

This talk will describe the up-to-date developments in the fields of pattern and character recognition, and indicate technical and economic trends in the near future. Character and pattern recognition represent automatic techniques, employing mechanical and electrical devices, for reading of patterns which have been placed on paper forms or films. Pattern recognition is the general term denoting automatic identification of all types of patterns. Character recognition specifically relates to the automatic identification of alphanumeric characters and symbol patterns and, therefore, forms a subdivision of pattern recognition.

2. Definition of Techniques

Pattern recognition can be defined as a technique for automatic identification of a given figure or arrangement which is known to belong to one of a finite set of pattern classes. This figure may relate to a missile launch site on an aerial photograph, a tumor identification on an X-ray, or an identification of resistor and capacitor symbols on an electric circuit diagram. The automatic reading of patterns replaces the present method of visually inspecting each film frame, which is a physically exhausting task prone to errors.

Character recognition is a technique for automatic identification of alphanumeric characters or symbols. These characters or symbols have relatively clearly defined property characteristics as compared with the general class of patterns.

The character recognition technique has been implemented in a device called a character reader, which is primarily a replacement for the keypunching and card reading operation. The character reader permits printed, typewritten, or handwritten data to be entered directly into the data processing system from the source document. In practical operation, this direct conversion is not always possible due to uncontrolled data preparation conditions, and a retranscription of data via typing is necessary. However, the typing operation has proven to be faster, more reliable, and more efficient than keypunching, and requires fewer hours of training.

3. Principles

Both character and pattern recognition may be implemented into equipment in a similar manner; the essential difference is the means for recognition of patterns. Figure 1 illustrates a general configuration for a pattern or character reader.

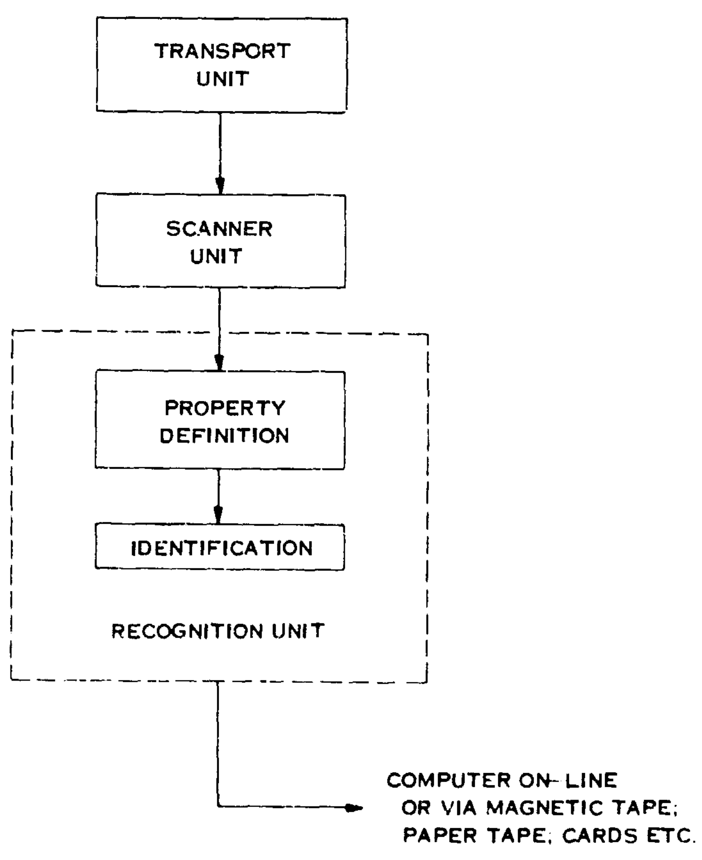


Figure 1. Block Diagram of Pattern/Character Recognizer Reader

The transport unit is a mechanical device for positioning the paper form or film by the reading station. The scanner converts the pattern appearing on the form into some analog or digital representation using optical or magnetic techniques.

The recognition unit, which is the heart of the system, extracts the significant properties from the pattern, thus identifying it as belonging to a given class. The problems of property definition and identification of general patterns are obviously more complex than just identifying a character belonging to a given character set.

4. Application Areas

The application areas for character and pattern recognition are quite diverse. Character recognition is business-oriented, and is directed toward supplying data to the computer more efficiently. Pattern recognition is science-oriented, and is directed toward analyzing visual information which has little semantic value.

The following are examples of formats that can be read by character recognition devices:

(1) Embossed customer credit card information
(2) Invoice billing
(3) Customer statements
(4) Transportation tickets
(5) Application forms
(6) Questionnaires
(7) Mail
(8) Book titles
(9) Phone directories

Pattern recognition work has been most pronounced in aerial photograph interpretation, medical film analysis, voice print identification, and graphics interpretation.

5. Trends

Awareness of both character and pattern recognition has increased significantly in the past year basically due to the commercial acceptance of the character reader. Once considered a research device, this reader is now taking its place in the data processing system. Therefore, pattern recognition, which has taken secondary consideration, can now be given more technical attention. The future challenges to the scientist and engineer definitely lie in the area of pattern recognition.

While pattern recognition work is involved in property definition and determination of proper identification algorithms, character recognition work is based on increasing the reader flexibility, reliability, and speed while reducing the costs to cover a wider area.

In this respect there are eight basic trends in character readers.

(1) Software

There is an upward trend to use a programmable unit for recognition rather than to rely on special purpose hardware. Programmable units increase flexibility; e.g., they can read forms having different character fonts and field formats. In addition, the programmable unit can be used for data extraction, sequencing, and manipulation to reduce the computer load further.

(2) Recognition of Handwriting

The work being done on the recognition of handwritten characters can be divided into two classes: handprinted and script. The ability to read numeric handwritten characters already exists in the readers manufactured by IBM and Optical Scanning Corporation. Recognition Equipment has recently announced the capability to read handprinted alphanumerics. The IBM 1287 Optical Reader, for example, can read handprinted numeric digits and five alphabetic control symbols, but the rigid set of rules emphasizes that the concept is still quite restricted in practice. However, reading of script characters is only in the developmental stage, the Post Office being the primary customer.

(3) Context Recognition

Context recognition, a long-range effort to reduce reject and error rates, is an attempt to simulate the human ability to apply contextual significance to characters or elements which might otherwise be devoid of meaning. When a person reads, the legibility of individual letters, or even of individual words, is not usually critical. Human beings "read" or perceive letters within the context of the entire word, and words within the context of an entire sentence. Consequently, the reader easily identifies "Sxrxxt" (where the "x's" represent garble), in the phrase "2231 South 12th Sxrxxt," as "Street" even though only 50 percent of the letters are readable. This identification is possible because of the conditional dependency of letters and words in human communication.

Although context recognition is not yet sophisticated enough to become a major factor in a recognition scheme, it can be used as a back-up method for identifying illegible characters. The most obvious advantage of this technique lies in its potential to identify a complete word in which one or more characters may present serious recognition difficulties.

(4) Off-Line Versus On-Line

One of the benefits of OCR is placing the input peripheral device off-line (in most cases). The on-line card reader has slowed down the computer and has, in some cases, required a separate computer for preprocessing. The trend of character readers is towards independent off-line units producing magnetic tapes for computer processing.

(5) Speed

Another, but less critical, area of developmental emphasis in character readers is speed. Reading speed is presently limited by the amount of time required to mechanically move the document past the reading station. The overlapping of the reading and transport operations is accomplished by using storage tubes (i.e., vidicon scanners) or by reading "on the fly." Speed can also be increased by using form controls which perform selective field reading, and skip blank spaces. The reliability of the character reader not only affects its accuracy, but also has a significant impact on document-reading capability; actual reading speed is obviously affected if a document must be read more than once.

(6) Improvements in Reliability

Naturally, reliability in the form of low error and reject rates is a prime consideration in all the development work being done on character readers. One approach to reduce these rates is to improve the resolution of the scanning units and thereby increase the number of sample points from which the equipment can make an identification. Philco-Ford Corporation is using a cathode-ray tube that has a resolution of 2000 optical lines, and even better resolution can be expected in the near future.

The reading reliability of the character readers, in terms of reject and error rates, has improved substantially due more to the source document preparation control, typist training, proofreading, and special checks within the character reader than to the recognition logic itself. The reject rate for in-house form preparation is presently two to three percent. Based on improvements in preparation, this reject rate will drop significantly to below 1 percent in the next five years, and error rates will fall to below 0.5 percent. Keypunching error rates are presently 1.5 percent.

(7) Cost

Present commercially available character readers are designed for large-scale operations (more than 10,000 documents per day), in which cost can be justified. There is, however, a definite need for a low-cost single-font character reader (approximately $20,000) which could read fixed-format single-document types. In view of the recent character set standardization, it would appear that a trend toward such a device is now likely.

(8) Remote Scanners

The use of remote scanners connected in a time-sharing configuration with a centralized recognition unit is within the state of the art and can be expected soon.

Non-Numerical Mathematics and Data Processing

F. Krückeberg

1. Non-Numerical Mathematics

Many problems of modern mathematics are non-numerical in structure. It is a simple matter to calculate the path of a rocket numerically, based on the theory of differential equations. This subject properly belongs to Analysis. On the contrary, the topological structure of cyclic satellite orbits is non-numerical. It is possible to classify types of orbits topologically.

There are mathematical areas which contain no numerical components as, for example, graph and network theory. With the help of these theories, complicated problems of strategy can be analysed. In the theory of games, graphs allow an intuitive grasp of the problem to be readily achieved. There is a close correspondence between graph theory and logic. Graphs can also be described in terms of Boolean matrices. With this, a link to algebra emerges. A special topic in algebra is Group Theory which can be used for the investigation of graphs. The importance of Group Theory is, however, much greater than this and more general. For example, one can with the help of groups, describe the symmetric properties of elementary particles.

Many problems in geometry are non-numerical in nature. Hilbert's research on the foundations of geometry are especially worthy of note here. A further important subject in mathematics is logic. This topic is, at present, being very actively pursued. It is possible to prove that large classes of problems can be solved without direct consideration of the individual problems through the utilisation of very broad logical generalities. An extention of this idea leads to the new topic of model theory. Modern mathematical theory is becoming ever more generalised and distant from the classical world of numerical analysis.

2. Non-Numerical Data Processing

The field of non-numerical data processing is so large, if one regards it with full generality, that no list of subjects can be exhaustive. We can merely cite several important new efforts without prejudice to the large number of others not mentioned.

Group Theory

It is possible to store finite groups in the core store and all group operations can be described in subroutines. In this way it is possible to manipulate groups in computers. For example, all sub-groups can be determined automatically and even more complicated problems of group theory can be solved very conveniently. In this domain, it is certain that very interesting new results will be obtained.

Graph Theory

Very complicated graphs can be stored in the computer, and the structure of the graph can then be investigated. It is possible, for example, to determine the shortest connection between two nodes of the graph. Further, cyclic sub-graphs can be discovered. Such questions are of the greatest practical interest. Tecnniques, including signal-flow graphs, flow graphs and k-trees allow one to obtain a clear intuitive picture of the functioning of a linear electrical network, after which analysis is much easier. Knowledge of the graph in topological network analysis eliminates many time consuming mesh and node calculations. In the chemical industry, the flow of material can be described by these methods. This is of considerable importance for the solution of management problems in such large factories. A large German chemical factory is, at present, actively using this technique in its daily operation.

Games Theory

It is well known that it is possible to programme a computer to play a complex game like chess. Since the theory of games is of the greatest importance for scientific management and logistics, game playing in the computer becomes a very serious occupation indeed. This application of the computer will be, in future, one of the most important. It is planned that German government practice will be strongly influenced by these methods in the near future.

Translation

Machine translation is of considerable importance due to the continued growth of international scientific exchange. It would seem that the pessimism expressed by a number of authors concerning the possibilities of automatic translation is unwarranted. It is merely a matter of time until this method has reached a usable degree of perfections. Successful attempts at Russian-English translation have delivered satisfactory results for scientific and technical texts. A Russian-German translation project is being carried out in collaboration with the Deutsche Forschungsgemeinschaft (German Research Association) and the University of Saarbrücken.

Pattern Recognition and Enhancement

The recognition of shapes is a very difficult and interesting problem whose solution has many applications. The problem can be divided into two parts. One is the decision as to which class out of a large number of possibilities a given well defined pattern belongs (character recognition, for example). The other is concerned with improving the definition in patterns which are greatly disturbed by extraneous influences and, given a limited number of classes, deciding into which such patterns most probably fall. For example, in the later category, in collaboration with the Rheinisches Landesmuseum, Labor für Feldarchäologie, a project in the enhancement of buried archaeological monuments seen in the results of surface geophysical measurement is in progress. Although the method requires much numerical manipulation of the data, the end result must be presented in a form which enhances the ability of the human eye to distinguish faint shapes in a noisy field.

Form, pattern, structure, logic trees, graphs, language, algebraic manipulation, all of these are but a few of the non-numerical problems which are yielding to the attack of non-numerical mathematics and data processing, giving new results in areas where hitherto insurmountable difficulty prevailed.

EFFICIENT TRANSFER OF TEXTUAL INFORMATION

James W. Altman
American Institutes for Research
Pittsburgh, Pennsylvania

PROBLEM

The need for efficient transfer of information from authors to potential users hardly requires belaboring. Yet, three central problems remain:

1. Text-sensitive tasks of scientists and engineers have not been delineated and analyzed in such a way as to define clearly their requirements for textual information support.

2. Methods have not yet been established which permit characterization of text in terms which support generation of a technology for efficient transfer of text to users.

3. Little concentrated effort has gone into attempts to establish lawful relationships between text-sensitive tasks and textual characteristics.

It is the purpose of this paper to review some of the findings already available concerning these issues and to suggest ways in which they might be resolved. Three principal phases of textual use are discussed: (1) screening, (2) gaining and maintaining awareness of a scientific or technical field, and (3) application of text-mediated information. These phases have been somewhat arbitrarily chosen; a reflection of a well-mapped domain of human behavior is not implied, for such mapping has not yet taken place.

SCREENING

Document screening can be conceived as occurring in three principal waves. The first wave involves selection of a set of documents from the total body of scientific and technical text, the phase perhaps currently containing the most dysfunctions between textual information systems and scientist-engineer needs. This paper, however, does not attempt to review the current status of document retrieval schemes. Rather, a distinction is drawn between recall systems and non-recall systems and some of the salient features of each discussed. In addition, one approach is explored which seems to have promise for clarifying the underlying logic of document retrieval systems, as well as practical implications for improvement in retrieval, particularly automatic, of documents. The second wave involves selection from the set of initially considered documents those that are to receive serious scrutiny or, in other words, the compilation of a bibliographic listing of some sort. The merits of attenuated representations of documents for this phase are recognized. The third wave includes the reading of documents and selection of those parts that will be remembered or applied. This wave is inextricably entwined with maintaining awareness and application.

GAINING AND MAINTAINING AWARENESS

A comparison is made between the textual requirements for gaining awareness of text for an immediate and specific purpose and those for maintaining current awareness. Concerning the latter, several studies are cited which demonstrate the utility of attenuated text (abstracts and extracts) in comprehending the essential meaning of scientific and technical articles. In addition, it is noted that differences between comprehension from full text versus comprehension from attenuated text tend to be maximal for text of intermediate complexity. That is, comprehension levels from attenuated and full text tend to be most similar where comprehension from full text is either very poor or very good. Mention is also made of the relevance of the comments concerning the screening of documents for a specific purpose to maintaining general awareness. Just as specific requests can, at least theoretically, be scaled according to some set of underlying dimensions which structure the field, so also might an individual's general interests be similarly scaled. Documentation having single multidimensional scaling might then be brought to his attention on a regular basis.

Gaining specific awareness usually requires greater detail than is provided by abstracts or extracts. According to the findings of one study referenced in this paper, an average loss of 30 percent (across a number of scientific and technical fields) resulted from efforts to use extensive descriptive abstracts as a source of specific facts. Interestingly, neither the extensive abstracts nor briefer abstracts brought about a time saving over the use of full text for fact retrieval. Although this inadequacy of abstracts and extracts for specific applications makes it tempting to conclude that the original document is required in most cases for such purposes, some notions which suggest that such a conclusion may be premature are presented later in the paper.

APPLICATION

Some considerations pertaining to textual analysis are presented, followed by the relationship of text characteristics to levels of application. Included in the discussion of textual analysis is a description of the operations involved in the procedure of breaking scientific and technical text into elements called "textual units." A description is also provided of the criteria involved in the testing and improvement of the uniformity of textual units resulting from this procedure, the categorization of these units, and their arrangement in convenient spatial configurations which reflect the apparent temporal progression of the author's writing. Major types of textual units are defined, in addition to idealized configurations of units for deductive, empirical research, and development studies. This characterization of text is then used as a basis for relating it to levels of application, arbitrarily limiting levels to these four: (1) reflecting simple awareness, (2) reflecting sophisticated awareness and/or collation, (3) reflecting analysis and synthesis, and (4) reflecting evaluation. Following a description of each of these levels, it is noted that any given application is likely to have relevance to only a selected sub-set of the entire set of more or less independent logical strings of textual units in a given document. The final point made is that given highly accessible and facile information systems in the future, it is not inconceivable that a prospective user would be selectively exposed to only those strings of textual units having likelihood of being relevant to a particular application and in an order most consistent with his scientific and technical task--with subsequent units in the string being presented on demand.

CONCLUSIONS

Based on the various findings and arguments advanced in this paper, the following conclusions are suggested:

1. In the preparation of original text:

   a. Some of the extensive attenuations which have been accomplished in text without measurable loss of information suggest that there may be considerable reduction yet possible in the volume of text generated through more rigorous attention to the preparation and monitoring of publication standards. This prospect seems particularly rich when it is recalled that the attenuations were accomplished on articles in published journals, which tend to be among the most terse of scientific and technical text.

   b. The traditional ordering of text to parallel the logical sequence followed by the investigator may be less than optimum for the individual who uses text. More optimal orders of presentation for use should be sought.

   c. The delineation of textual units in the process of preparing text for storage and retrieval may serve as a powerful aid to efficient textual transfer. Its potentials should be investigated.

2. The structuring of phenomenological fields by the methods of multivariate analysis and related rational processes should be explored as an aid to:

   a. Retrieval of documents on the basis of specific queries.

   b. Selective dissemination of document lists to match individual interests.

   c. Characterization and retrieval of individual textual units as well as whole documents.

3. Abstracts and extracts show promise as efficient aids to maintaining current awareness of a field. Only a rudimentary technology now exists

concerning this issue. That technology should be strengthened by greatly expanded empirical studies of the characteristics of attenuated text which make it a suitable substitute for full text.

4. Information systems to date have concentrated almost exclusively on the retrieval of whole documents. Work should be initiated on determining the feasibility and value of retrieving individual textual units and strings of related units.

ON-LINE INFORMATION STORAGE AND RETRIEVAL

Prepared for the

AGARD Symposium on
"Storage and Retrieval of Information"
For Session 3 - The Evolution of Current Methodology

18-21 June 1968
Munich, Germany

Noah S. Prywes
The Moore School of Electrical Engineering
University of Pennsylvania

This paper reviews briefly the components of a total storage and retrieval system while referencing relevant developments.

The storage process described in this paper includes all the functions which take place in libraries and information centers from acquisition to the placing of the documents in the repository. This process, which includes indexing, cataloging and vocabulary maintenance, demands a great deal of time and expertise. In any one of the large libraries or information centers there are thousands of monographs and serials that are waiting to be catalogued and indexed. These often lay unused because of the dearth of competent cataloguers and indexers, especially those expert in particular subjects and languages. The increased amount of material which is being circulated soon may require substantial increase in staff. Staff with this competence is extremely scarce; low salaries discourage young people from library work. For these reasons the storage process tends to constitute a serious bottleneck.

On the retrieval side, evaluation tests indicate that libraries and information centers operate at a low, almost inacceptable retrieval effectiveness. The library user requiring specific information is overwhelmed with information, much of which is irrelevant.

The mechanizing of procedures in an information center or a library does not need any more justification than the notion of mechanizing any other industrial, commercial, or service function. The premise of this paper is that automatic storage processing and on-line retrieval are competitive in effectiveness with manual procedures. The automatic procedures are not especially complex and they can be readily applied.

The automated storage processing discussed here includes the following steps. Citations and sometimes abstracts of incoming documents are first transcribed into machine readable form. Natural language processing of title and abstract results first in a concordance of stem words. The concordance may also provide information about the frequency of stem words. In a semi-automatic process, words may be omitted, added, or various relationships established between words to form an open-ended thesaurus. Then, based on this thesaurus, the incoming documents are automatically indexed. Finally, an automatic process may be applied which generates a library classification system for the collection. Such a classification then represents a scheme for placing documents on shelves, in microforms, or in the computer, as appropriate.

The interactive man-computer retrieval process, which follows the storage process, offers the best potential for improving retrieval effectiveness to the point where information storage and retrieval systems become really useful. This interactive process has a number of aspects. An individual can communicate with a central computer through a remote terminal "on-line", i.e., where the terminal is continuously monitored by a central computer. The computer deletes, changes and analyzes the queries and retrieves information in "real-time", compatible with the normal working speed of a human. To fulfill these functions, the computer must have a storage capacity of billions of characters with fractional second access to any information.

In the interactive retrieval process the user may search thesauri, classification schedules, catalogues or the documents in a manner very similar to that employed traditionally in libraries; however, with far less effort and much greater speed. He may for instance reference documents by title, author, publisher, citation, subject, or browse through citations or abstracts of documents on a common subject, placed together in the memory of the computer.

Methods and procedures like those described in this paper, such as content analysis, concordance and thesaurus preparation and indexing, which require merely clerical procedures, have been proposed for centuries. They have been opposed by those who believe that manual processing of the document has a "quality" superior to algorithmic processing based on selection of words from the abstract or even from the title. The manual approach has a number of ancillary positions that are contested here. For instance, the manual approach also conveys the notion that the subject term vocabulary needs to be controlled, and that only highly competent persons in specific areas should exercise judgment in regard to adding terms; these positions are contradictory to the approach in this paper. The objective of the procedures described here is to do away with much of the vocabulary maintenance work currently prevalent, especially the notes and instructions directed to indexers and cataloguers which would not be required in an automated system.

M A N U A L S Y S T E M S

TDCK CIRCULAR THESAURUS SYSTEM

by J.A. Schüller - TDCK.Neth.

Paper presented during the Symposium on Storage and Retrieval of Information.
Joint NATO AGARD Avionics and Technical Information Panels.

Extended Summary

As an example of a manually operated documentation centre a description is presented of

1. the organization
2. the functions
3. the systems

1. For the organization I refer to fig.1.
The centre falls directly under the Minister of defence. An advisory council consisting of four members advises the minister, at his request or on their own initiative on TDCK-policy matters. Three of the members are high-ranking university trained officers of the Navy, the Army and the Airforce and the fourth member is the director of TDCK.
The personnel strength is 63 people.
Following the scheme of fig 1. we see at the lefthandside the technical divisions and at the righthandside the special-library department and the administration division.
The technical division, subdivided in sections, is manned by scientists and technical engineers, in total 25.

2. The primary functions of TDCK are:

   2.1 To collect, evaluate and store new scientific information from all over the world, which may be useful for military purposes in general.

   2.2 To be well informed of highly specialized information-sources in and outside the country.

   2.3 For use available information for giving assistance to those scientists technical investigators and officers, who are involved in solving problem in research, technology, education, management, military sciences and other fields of military interest.

A documentation centre is a model for applied efficiency. It's charge is to use available information to inhibit costly duplication on the one side and to organize the transport of information in a broad sence, on the other side.

On fig. 2.a table to sources of information is shown, a list which is intended to remind our information officers of all sources available at TDCK which should be consulted when, for instance, a selective bibliography has to be compiled. If, for example a bibliography on "Inertial Navigation Systems" has been requested 18 different sources will have to be consulted according this schedule, all of them packed with information and all of them using their own indexing system. This last remark suggests one of the reasons why TDCK does not feel like using a computer for its own system; some 35 other systems would still have to be scanned in a conventional manner.

Requests put to TDCK can be answered in four different ways:

a. by making available a selection of reports or articles;
b. by making available a selective bibliography consisting of titles and abstracts - all well indexed and crossreferenced;
c. through the production of a literature search report or a state-of-the-art report, in which also verbally obtained information is included.

3. The systems in use at TDCK.

For defining the contents of the literature and for specialized retrieval search TDCK is using two different systems: Namely the UDC (Universal Decimal Classification) and the TDCK-Compact System.
Apart from these systems, a source index is kept for retrieving reports, papers etc. according to their issuing organization, an institute, a laboratory etc. Part of this activity is mechanized by a lectriever, which is an electrically driven push-button file selector.

When designing the TDCK thesaurus, we conciously sought to obtain a combination of:

1. a systematic subject set up;
2. alphabetical arrangement of descriptors;
3. coordinate indexing principle;
4. mutual relations and facets; and finally
5. visible directions display.

Figure 3. is showing an example of a very simple circle scheme. Notions which are related to each other and which can be placed in the familiar family-tree pyramid, are placed on cencentric circles; that is to say we have simply made a plan view of our tree, because on a circle we have in practice, about five times more room than in a pyramid structure.
Arrows, fanning out in all directions, are used to display relationships between concepts.

The word at the centre of each circle scheme is the descriptor of a special descriptor-field.

The following rules are in force:

1. The thesaurus consists of descriptors;
2. Each framed descriptor appears only once in the system;
3. At the input, coding will start from the centre of the circle, following an arrow until the wanted descriptor has been reached. In one circle more than one radius may be followed.
4. All descriptors encountered will be noted down. When a descriptor from another circle scheme has to be used, we can "borrow" such a notion and add it to our circle-scheme outside the ultimate circle, in which case we do not use a frame.
5. New descriptors have to be defined by the subject specialist concerned.
6. On the retrieval side the circle-schemes will always be used. The user (a specialist in the field) will be led automatically to the pertinent descriptors, once the appropriate scheme has been selected.
7. The thesaurus, as a one-language technical index, can be translated.
   The use of homonyms or synonyms is avoided. The word "measurement", for example, will be used in several descriptors, which could read:
   measurment of time, measurement of distance, ballistic measurement etc.

   In all 376 circle-schemes have been designed to date.

.-.-.-.-.-.

MECHANISED SYSTEMS

by

N. E. C. Isotta
European Space Research Organisation, Paris

INTRODUCTION

Until not very long ago, perhaps only about four or five years, the authors of most papers on mechanised systems of information retrieval, would be mainly concerned with one of two things, and these were in very general terms, either the necessity for the justification of the commencement of a machine system, or an attempt to prove that having started such a system, the results were worthwhile or at least as good as expected. Nowadays, the atmosphere in these matters is rather different since most people, and by this I mean both the customer and the supplier, have realised and accepted the need for mechanical methods of handling large files.

THE MACHINE VERSUS THE PROFESSIONAL

There is certainly complete acceptance of the fact now, that documentary processes are particularly amenable to mechanisation in what are often known as "business activity" areas. Such areas may be fairly easily defined; they include operations which can theoretically be performed without the direct intervention of professional labour, even though such labour may have been necessary initially to establish the operational procedures. These will include such matters as stock control, catalogue or index printing, preparation of announcement journals or accession lists, the establishment of "field of interest" registers etc. The most important activities remaining which still require professional attention are therefore document analysis for input, question analysis for retrieval, and retrieval result evaluation. Speaking rather heretically, primarily as a documentalist, and not as a user, it seems to me however inevitable, that even these areas must eventually succumb to machine treatment, simply because of the sheer weight of material involved, in conjunction with the increased effectiveness of the machine systems available.

MACHINE INPUT

The main problem area for many years has been the one of getting the documentary material into the machine. Keyboarding in one form or another has remained essential. Technically speaking it has been possible for some time to arrange for input to be made directly to a computer without the necessity for such keyboarding operations. Economically however, such systems, which are normally accompanied by very large workload capacities, have not been justifiable in circumstances where the capacity would never be fully taken up. This is therefore still one of the most expensive, time consuming, and error ridden parts of an integrated machine system. Optical scanning could be a solution if standardised print formats were used in document production, in order to avoid an intermediate keyboarding into such a standard type script.

COOPERATIVE SYSTEMS

The ESRO/ELDO Space Documentation Service based on an exchange arrangement with NASA, is one of the first ventures of the kind where the agency receiving the machine system i.e. ESRO/ELDO, is also responsible for the provision of machine input to the system operated by the supplying agency i.e. NASA. This has underlined the problems mentioned above and has certainly indicated the enormous advantages which could be gained if greater standardisation in this area could be achieved.

MACHINE OUTPUT AND USER REACTION

It is clear that in spite of the advanced computer age in which we live, there has been a general diminution of standards of production resulting from the use of computers, and many of the users of machine document systems are accepting this with reluctance. The computer manufacturer's philosophy until quite recently has been that the advantages inherent in machine processing in respect of time saving, and capacity, have outweighed any disadvantages apparent in the final machine product. In my view, they have been totally wrong. The manufacturers of such things as detergents can teach the computer manufacturer a great deal concerning "eye appeal" and "packaging". There are known cases where the cost of the package is greater than the cost of the contents; one specific example outside the detergent field, is the can of water supplied on certain European flights. Even now upper and lower case computer output is a rarity and is often associated with some other extremely expensive offline printing machine. However, by now, the user too should have become somewhat more sophisticated in his reaction to the current standards of computer print-out. He should make the best of what is available since it is a retrograde step to interpose between the computer output and the user, some intermediate manual stage, be it editing or the improvement of appearance of the output by some other printing or reproduction process. I feel sure that what must be aimed at, is a completely satisfactory direct computer output; but certainly, in the meantime, the user must overcome his prejudices, although at the same time he should be sufficiently vocal to indicate that the result is not really pretty enough to encourage him to make the greatest use of it.

DIRECT USER ACCESS TO THE MACHINE

This must of course come. A habit which is, I think, engrained in most of us after centuries of the existence of libraries, is that of browsing. This is something that the machine has been tending to deprive us of, since somehow, wading through a computer listing is not quite the same thing as browsing through a shelf of books. It is now possible however, to approach a similar situation by means of direct interrogation of the machine file using a remote visual display console. The ESRO/ELDO Space Documentation Service hopes to have such a capacity available initially for its own analytical staff, early in 1969, closely following a NASA lead. Ultimately such consoles would be installed at strategic points throughout the European network of ESRO establishments, thus enabling the user to go direct to the machine as and when he feels like it. For some time it has, I think, been apparent that future development would be in this direction. It is essential that the future of machine information retrieval is not designed around the capabilities of the first and second generation computers with which the technique was born. Joint effort on the part of the supplier, i.e. the documentalist, and on the part of the user should soon achieve the desired result.

An Introduction to the Study of Cost Effectiveness in Information Systems

Professor J.N. Wolfe
University of Edinburgh

The following is a brief summary of the programme to be covered at the AGARD Conference in Munich.

Large sums are spent each year on information services in each of the NATO countries. As an example, the United Kingdom alone spends about 50 million pounds each year on library services only. We have now no reliable and consistent statistics for the amount spent in other NATO countries, and in particular, we lack information on the amount spent on information services other than libraries. The OECD is in the process of attempting to collect this information and the study is underway under the direction of the Studiengruppe of Heidelberg, Germany.

We know however that the total sum being spent is sufficiently large and growing with sufficient rapidity to present a serious economic problem. This economic problem has several aspects. First, there is the question of how much ought to be spent on information services in general. Secondly, there is the question of how rapidly this sum should grow. Thirdly, there is the question of the most appropriate division of expenditure among the competing types of information service which might be offered, and fourthly there is the question of the most appropriate organisation of information services both within a single country and between countries.

These sorts of questions may have seemed to be of only academic interest during the last decade or so, for there has been general agreement that the volume of funds available for information services has hitherto been too low, and funds have been expanded with considerable rapidity. During this period too there has been rapid technological change in the information industry. There are now many more technically developed candidates for absorption of information funds than was the case even a decade ago. As new techniques pass from the laboratory and pilot stage into the world of practical possibility the question of economic viability and value for money becomes a very real and pressing one.

It was in this context that the Office for Scientific and Technical Information in the United Kingdom, acting in collaboration with the OECD, decided to undertake a study of the economic aspects of information systems. The study was commissioned from the Department of Economics in the University of Edinburgh, and involves a large team of workers including five full-time economists and a full-time information officer, two accountants, two statisticians, and five part-time economists. The work has been underway for approximately four months but will not be completed until the end of the calendar year 1969. One aspect of this work which is already rather far advanced is an economic study of the library system and particularly the public library system in the United Kingdom. It is proposed to publish very shortly a volume of essays on this topic. Most of the papers involved are quantitative and econometric in character and it would be difficult to summarise any of them briefly. I would however like to mention here two papers in particular which seem to me to offer considerable interest. One of these is a paper by Mr. Ralph Young on the Forecasting of the Demand for Library Services in the Public Library sector by econometric means. This paper provides what I think is the first attempt to offer a quantitative forecasting technique for library demand which is not simply an extrapolation of past trends. Mr. Young shows that even at this early stage of analysis it is possible to forecast the appropriate level of library provision in a general way at least with considerably improved accuracy. This technique has been applied, as I say, to the public library system but I think that it offers considerable possibilities of extension to library systems within private firms or government agencies. Another paper of some interest is that prepared by Dr. Jacob Moreh which examines in a statistical and econometric way the problem of economies of scale in library services. Dr. Moreh attempts, and I think for the first time, to go below the level of simply comparing large groups of dissimilar libraries with one another on the basis of an average cost figure. Such a procedure, while common enough in practice, is of course statistically exceedingly unreliable. Dr. Moreh on the other hand utilises techniques made familiar in production function studies to examine the cost functions of operating within the public library system on the basis of a variety of independent variables including number of branches in each library system, the number of employees, the number of volumes, and the volume of ancillary services such as gramophone record issues. While his results are not yet completely analysed, they do seem to indicate that the popular belief in economies of scale in the library world may be somewhat over-simplified.

Before moving to some account of the larger economic study now underway, it may be useful to provide some introductory observations on the nature of cost effectiveness studies in the context of information and library services. It will be recalled that cost effectiveness techniques were given substantial development by work undertaken on behalf of the United States Department of Defence largely in the Rand Corporation of Santa Monica, California. Put in the simplest way, the notion of a cost effectiveness study is an attempt to discover the relative magnitude of costs and benefits accruing from alternative forms of expenditure. More concretely, the early studies involved assessment of the relative cost per ton of bomb delivery for example. The essence of a cost effectiveness study is the reduction of the benefits of alternative task systems to some kind of commensurable unit. Once this is done the problem becomes merely one of comparing the alternative task outputs with their costs.

Looking at the matter in another way, we may see the cost effectiveness study as simply an improvement on the more normal cost study. The traditional cost procedure involves an examination of the costs of two alternative tasks. But clearly costs are not a sufficient determination of which task provides the best outcome. We must consider as well the benefits achieved in each outcome.

Let us take an extremely simple example drawn from everyday life. Supposing we wished to determine which was the wiser purchase, an orange or a lemon. We could easily determine the cost of the orange and the cost of the lemon. The question of which of the two fruits provides the better buy for money depends however upon what we wish the fruits for. If we are anxious to obtain a given quantity of Vitamin C, for example, it may well be that the lemon provides the better bargain. If our object is to provide a refreshing morning drink, and we wish therefore to maximise the sugar content of the citric juices, then a different answer may be obtained. We cannot therefore tell which fruit it would be worth our while to purchase until we determine the objectives for which we are purchasing them.

With this introduction in mind there should be little difficulty in understanding the procedure which is being adopted with respect to cost effectiveness in information services. The first part of our job is to determine cost for alternative types of service. This presents certain features of difficulty because of the fact that information services, like most public services, do not normally keep accounts upon what is called a functional basis. That is to say, the accounts of most information services take the form of a list of expenditure by name item of expenditure. That is to say labour, materials, rent, heat, etc. They do not normally assign these expenditures to the manifold functions which an information service in fact attempts to achieve. It is therefore necessary to recast the accounts of the information services in functional form before any serious further work can be done.

One of the basic difficulties here is of course the assignment of overhead costs to the various alternative functions. We have to ask for example what proportion of the time of a head librarian ought to be attributed to his work as head of an information service as well as of a library service, in a unit which offers both library and information services. Similarly we may ask what proportion of the cost of heating an information centre is to be attributed, let us say, to the preparation of abstracts on the one hand or to the preparation of translations on the other. It will be clear that, however much care is taken, there will be a certain measure of arbitrariness in such calculations. It is our object not to eliminate arbitrariness entirely, but rather to reduce it to manageable proportions.

One important issue is the extent to which information services may be added to existing library activities at lower costs than information services can be provided, in a purpose-built organisation. On the one hand we might expect that the sharing of certain overheads with a library would produce lower costs in the integrated operation. On the other hand the greater expertise which can be developed in a specialised and purpose-built organisation may conceivably offer economies of substantial importance. This balance between economies of scale and economies of specialisation is, as everywhere else in industry, an important question deserving the most careful examination.

The central core of our method consists of evaluating two particular types of situation. The first is a situation in which an old type of information service is to be superseded by a new type. This situation provides alternative information on costs and also provides information on the change in value of the service received by changing over between the two systems. An alternative approach consists of examining situations in which two alternative types of information service exist side by side. For example, we may have certain organisations which utilise an advanced information service while other organisations utilise an older style of information service. Here costs and effectiveness may be compared on a cross-section basis. It will be understood, however, that in this case there may be expected to be a substantial amount of extraneous information introduced because of the possibility of underlying quality difference between the units using the technically advanced information service and those using the technically less advanced information service. The final part of our work consists in evaluating the services provided by alternative information systems. This is clearly the most difficult part of our job. It is difficult partly because previous attempts to deal with user requirements and user needs have not been directed specifically to economic investigations. There is a fundamental difference between technological criteria of efficiency in this context and criteria of economic efficiency. Ideally, one would like to obtain estimates of the impact of the information service on the productivity of the workers receiving the information service. In practice this level of productivity is likely to be very much influenced by extraneous factors. This is a particularly damaging point if we are dealing with cross-section studies of a particular industry which has different information services in different firms. We are likely, I think, to find that good information services are in fact characteristic of technologically advanced firms, and if this is the case any attempt to correlate efficiency with information services is likely to give us too optimistic a result. When we deal with changes in information services affecting all the units in an industry, we have, I think, a rather more practical proposition, although here we will, I am afraid, be hampered for some time yet by a shortage of instances. I would expect, however, that as the number of information services examined increases, a statistically reliable result may eventually be approximated.

There are alternative methods of obtaining effectiveness measurements from information services. Some of these consist of sampling opinion about efficiency. Others consist of obtaining objective characteristics of the functioning of the information service. I would rather, however, leave a further consideration of this particular problem to discussion at our forthcoming meeting.

PAPER XII

USER NEEDS

by

W.C.Christensen

Department of Defense, USA

Not available for Extended Summary

SELECTIVE DISSEMINATION OF INFORMATION

Melvin S. Day

Deputy Assistant Administrator for Technology Utilization

National Aeronautics and Space Administration

Washington, D. C., U. S. A.

Today's scientist and engineer who attempts to keep aware of current developments in even a limited area of interest encounters a real problem in the growing number and size of the abstract journals that he must review. One successful approach to reducing literature-review efforts is Selective Dissemination of Information (SDI). SDI applies computer technology to the task of providing a personalized current awareness service.

In the past, libraries have provided patrons with a current awareness service by circulating copies of current documents of particular interest, by distributing library accession lists, and by routing current issues of abstract journals. The usefulness of such methods is limited by inconsistent selection or by excessive volume of material announced. For the great majority of scientists and engineers, a more formal and efficient service is necessary to alert them to current documents of specific significance.

In describing SDI, it is essential to refer to the mechanism of the SDI operation and especially to the relative costs and efforts involved in the many different SDI systems that can be designed. Potential users should participate actively in the design of information systems and be able to demonstrate the cost effectiveness of the service received.

SDI services have been initiated by many organizations, both in the United States and Europe. NASA has operated several types of SDI program since 1963. These have been distinguished by volume of input and by size of user population. During 1967, 875 users received, four times each month, announcements selected from the current issues of Scientific and Technical Aerospace Reports and International Aerospace Abstracts. Input to the system during 1967 thus consisted of a total of 63,700 aerospace reports, journal articles, and conference papers. The users received a total of 800,000 announcements, selected according to their specific interests, an average of 19 per user per issue of the abstract journals.

Principles of SDI

SDI is a current-awareness tool and results in the selection and announcement of current documents having a high probability of interest to the individual user. The fundamental element in selection is the computer comparison of two data files. One is the file of bibliographic data covering newly received reports. These data include subject index terms, authors, corporate sources, supporting agency, contract or grant numbers, etc. The other file contains the users' interest profiles, which in essence are equivalent to bibliographic search strategies. The interest profiles consist of selected bibliographic data elements, such as subject index terms, related by the common AND, OR, or NOT Boolean logic expressions. Other methods of relating terms are possible; e.g., by assigning relative weights to each term, a certain minimum total weight of matching terms then being required before a document is chosen for announcement.

The interest profile is thus a rational set of terms in the same technical language used by the document indexers. Structuring an interest profile may require considerable skill. As an example of profile complexity, a potential user interested in supersonic transports would not be satisfied if only the term Supersonic Transport were placed in his profile. Documents specifically on the Concorde might be indexed to the term Concorde Aircraft and not to the general term Supersonic Transport. Other concepts of possible interest would be clear air turbulence, sonic boom, international law affecting civil aviation, or basic engineering problems involving supersonic heat transfer, supersonic flutter, or supersonic wind tunnels. In preparing the interest profile the user should also consider whether he desires to limit the total number of announcements to him. How can this be done? It can be done by limiting the number of terms in the profile, or by restricting selection through the use of Boolean AND or NOT relations, or by arbitrarily cutting off the printout at a prescribed number.

Experience has shown that in structuring their SDI interest profiles most scientists and engineers require considerable help from professional reference analysts who are familiar with the authorized vocabulary terms and the indexing practices and patterns. I should point out that the success of an SDI program is directly related to the quality of the user profiles.

Beside the interest profile, additional features essential to an SDI program are:

1. A standard form for announcing the selected documents.
2. A method for requesting a copy of the announced document.
3. Routine feedback by the user to provide a quantitative measure of the performance of the selection process.

Types of announcement form

Numerous forms have been designed for announcing selected documents to users. A card type of form announces each report or journal article as a unit record, a listing type prints the announcements continuously on sheets.

Cards are supplied to the user in the majority of SDI systems. Typically the card is the standard electronic data processing (EDP) card. The card may present a full abstract of the selected document or merely a bibliographic reference. If only a document reference is presented, the index terms may also be printed out to add to the informativeness of the title.

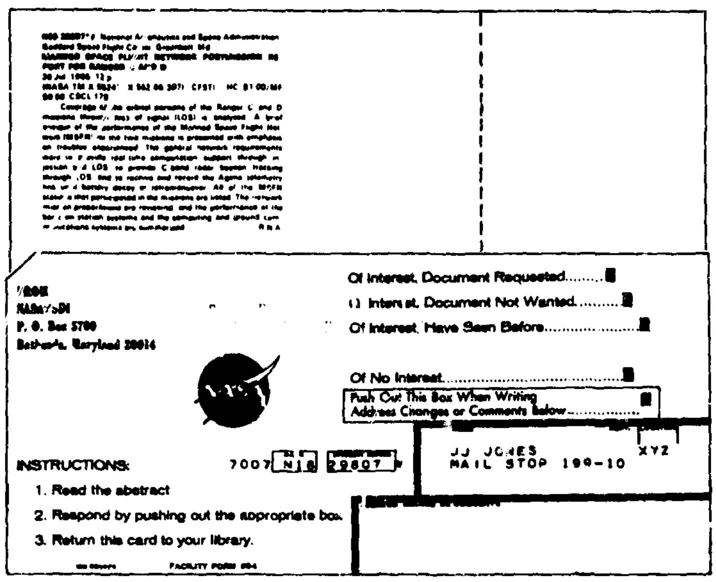


Fig. 1. NASA two-card SDI announcement.

The SDI document announcement appearing on the card may be computer printed, or it may be duplicated, as by offset reproduction. Offset reproduction permits full abstracts to be reduced in size and presented on a single card. A disadvantage is the need for two operational procedures. The computer first punches a card with the user's name and address and a document identification number. After sorting the cards by document number, abstracts are reproduced on the appropriate cards by offset.

Notification cards are commonly designed so that a stub may be detached and returned by the user to request a copy of the document or to express his evaluation of the announcement as feedback to the system.

The user's address and the abstract need not be printed on the same card. NASA's first SDI program, operational from December 1963 to January 1966, provided the user with two cards for each announcement (Fig. 1). One was an EDP card punched with the user's name and address and the document number. Prescored blocks could be punched out to express the user's evaluation of the announcement. The second card was not computer manipulated, but presented the full offset-printed abstract of the selected document. The two cards for an announcement were inserted into a single window envelope for delivery to the user. A vertical perforation allowed the abstract card to be converted to a 3 x 5 inch file card.

Cards are very popular with the SDI users, as he may file those of particular interest in a personal desk-drawer file. Costs of a card-type system are dependent on many factors, but might fall in the range of $100 to $150 per user per annum for a large volume of input references; e.g., the total references (approximately 65,000) in Scientific and Technical Aerospace Reports and

International Aerospace Abstracts.

Less expensive than cards are computer-printed listings of selected bibliographic references. The abstract may be printed out if it is on machine-readable files, at the expense of added computer use and increased bulkiness of the announcement package.

NASA's present system, in effect since February 1966, is an adaptation of the simple listing. A three-copy, no-carbon-required form is used (Fig. 2). The computer-printed references of course appear on all three sheets, together with the user's name and address. When the user marks his evaluation or document request he simultaneously marks all copies. The original may be retained, one of the copies is used to fill document requests, and the other copy is returned to the system operator, who tabulates the responses as a measure of system effectiveness.

N A S A / S D I   N O T I F I C A T I O N

Please check the appropriate box. USER keep the top copy, and send the second and third copies to your library.

DOCUMENT REQUESTED | OF INTEREST, NOT REQUESTED | NO INTEREST

STAR ISSUE #06, 23 MARCH 1968
A E ANDERSON AME0101
N-227-1

N68-15019 #INSTITUT FRANCO-ALLEMAND DE RECHERCHES, ▪ ▪ ▪
CAT. 02 ST. LOUIS /FRANCE/.
BOOMS PRODUCED BY A MIRAGE 3 B IN
ACCELERATED FLIGHT - OPERATION JERICHO-
FOCALIZATION IN ISTRES FROM 8-16 DECEMBER
1966
FLUGZEUGKNALLE EINER BESCHLEUNIGT FLIEGENDEN
MIRAGE III B, OPERATION JERICHO-FOCALISATION
IN ISTRES VOM 8.-10. DEZEMBER 1966
FROBOESE, M. DATE- 10 AUG. 1967 COLL- 61 P REFS
LANG- IN GERMAN
ISL-T-30/67
PEAK PRESSURE OF SONIC BOOMS PRODUCED BY MIRAGE 3
AIRCRAFT MEASURED OVER FIXED GROUND RANGE
*ACOUSTIC MEASUREMENTS, ELASTIC WAVES, FLIGHT PATHS, JET
AIRCRAFT NOISE, MIRAGE 3 AIRCRAFT, *PRESSURE DISTRIBUTION,
PRESSURE RECORDERS, *SONIC BOOMS, SUPERSONIC FLIGHT

N68-15022 #CORNELL AERONAUTICAL LAB., INC., BUFFALO, ▪ ▪ ▪
CAT. 11 N. Y.
THE MULTI-RECOMPRESSION HEATER, A NEW
CONCEPT FOR LARGE SCALE HYPERSONIC TESTING
WEATHERSTON, R. C. DATE- DEC. 1967 COLL- 61 P
REFS
CAL-AD-2390-Z-1
THERMODYNAMICS, HEAT TRANSFER, AND MECHANICAL
DESIGN OF MULTI-RECOMPRESSION HEATER FOR
SIMULATION TESTING OF HYPERSONIC VEHICLES
*ATMOSPHERIC ENTRY SIMULATION, *CONVECTIVE HEAT TRANSFER,
*EXPERIMENTAL DESIGN, *HEATING EQUIPMENT, HYPERSONIC
VEHICLES, HYPERVELOCITY WIND TUNNELS, THERMAL ENERGY,
THERMODYNAMICS

N68-15099 SUD-AVIATION, PARIS /FRANCE/. ▪ ▪ ▪
CAT. 02 CONCEPTION OF THE AIRFRAME AND
AERODYNAMIC PROBLEMS OF THE AEROBUS
CONCEPTION DE LA CELLULE ET PROBLEMS
AERODYNAMIQUES DE L'AIRBUS
ROCHE, C. DATE- ▫1967▫ COLL- 23 P LANG- IN
FRENCH CONF- PRESENTED AT A.F.I.T.A.E. 4TH
COLLOQ. ON APPL. AERODYN., 8-10 NOV. 1967
EUROPEAN AEROBUS CONFIGURATIONS, PASSENGER
TRAFFIC ECONOMICS, DEVELOPMENT COSTS, AND
AIRPORT PLANNING
AIR CARGO, *AIRCRAFT CONFIGURATIONS, *AIRPLANE PRODUCTION
COSTS, AIRPORT PLANNING, ECONOMICS, EUROPE, OPERATIONAL
PROBLEMS, PASSENGERS, *TRANSPORT AIRCRAFT

PAGE 01

FF No. 779a June 66. USER COPY

Fig. 2. NASA list-type, three-part SDI announcement form.

Costs of a list-type system depend on the information presented and the computer and design factors, but might fall in the range of 60 to 70 per cent of the card-type system.

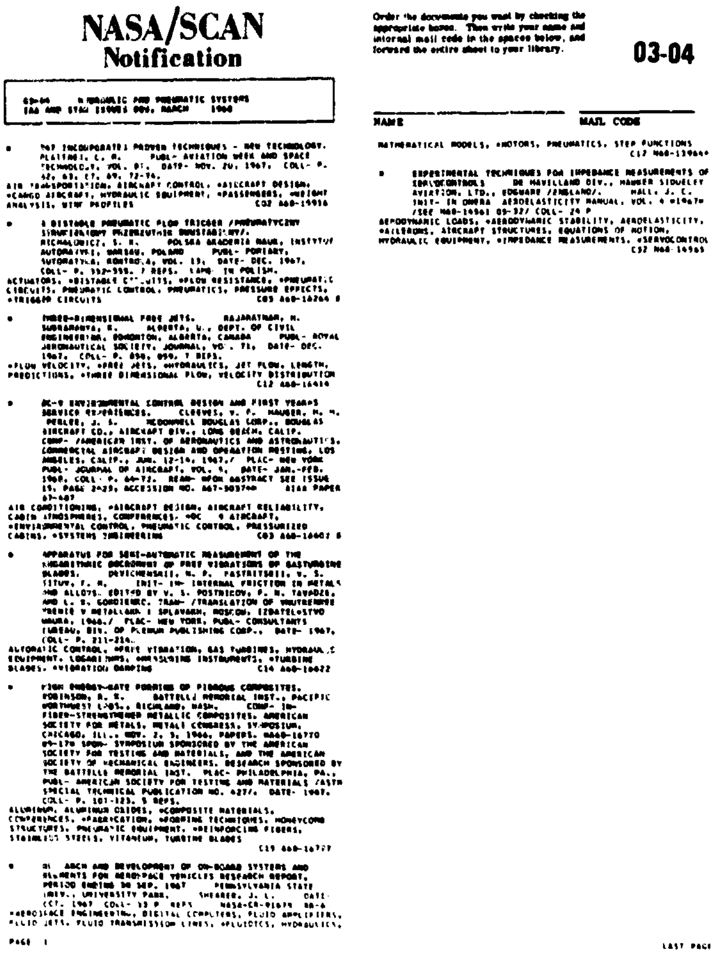
User feedback

Optimum SDI service consists in assuring that as high a percentage as possible of the announcements he receives are of interest to the user. Furthermore, these relevant announcements should be as high a percentage as possible of the potentially useful documents in the file. By tabulating the user's interest evaluation, the operator can identify those interest profiles that need improvement and thus optimize the system. In NASA's SDI experience, an average of 75 percent of the announcements are of interest to the user.

Trend toward standard profiles

SDI systems tailored to individual users are very costly for providing information to large numbers of users. Each new user adds to the required computer time. User turnover and changes in interest can be high; thus, updating the user profiles can be a constant activity. Professional assistance in structuring interest profiles also increases with the number of users.

A solution with great promise is the standard interest profile. Instead of profiles tailored to a particular individual's interests, a series of profiles selects announcements according to subject topics. The user receives copies of the computer-printed topic notification listings that, in combination, best cover his specific interests. As a consequence of transferring much of the over-all effort from computer operations to the relatively inexpensive operations of traditional printing and sorting, this type of service is much less expensive per user than is SDI.


NASA/SCAN Notification

Order the documents you want by checking the appropriate boxes. Then write your name and internal mail code in the spaces below, and forward the entire sheet to your library.

03-04

NAME

MAIL CODE

MATHEMATICAL MODELS, *MOTORS, PNEUMATICS, STEP FUNCTIONS

EXPERIMENTAL TECHNIQUES FOR IMPEDANCE MEASUREMENTS OF ... DE HAVILLAND DIV., HAWKER SIDDELEY AVIATION, LTD., EDGWARE /ENGLAND/. HALL, J. C.

AERODYNAMIC LOADS, *AERODYNAMIC STABILITY, AEROELASTICITY, *AILERONS, AIRCRAFT STRUCTURES, EQUATIONS OF MOTION, HYDRAULIC EQUIPMENT, *IMPEDANCE MEASUREMENTS, *SERVOCONTROL

Fig. 3. NASA/SCAN notification listing.

Typical of the standard profile type of current awareness service is the NASA/SCAN (Selected Current Aerospace Notices) Program. A sample notification listing is illustrated in Fig. 3. SCAN is a developmental program with a limited participation at present, but it offers the possibility of providing a selective current awareness service to possibly tens of thousands of aerospace scientists and engineers. Its coverage is the same as the NASA/SDI program; i.e., the current files of Scientific and Technical Aerospace Reports and International Aerospace Abstracts.

SCAN offers great flexibility in adding or deleting topics and in modifying the scope of topics through profile changes. For example, SCAN topics include Supersonic Transports, Clear Air Turbulence, and Aircraft Noise and Sonic Boom. The first provides broad coverage, while the others meet the needs of users with specific interests. SCAN topics can overlap in coverage and announce the same document under a number of appropriate headings, permitting the user to match either specific or broad interests by a minimum of notification listings. SCAN optimization is accomplished by studies of topic output and by brief questionnaires requesting user comments on their needs for new topics or modification of existing topics.

SCAN appears to be the path that current awareness service will take to provide service to large numbers of users. Several U. S. Government agencies are testing programs much like SCAN. Projected costs for large-scale SCAN programs appear to lie in the range of $10 to $20 per user per annum, again depending on the details of the service provided.

Computer programs for SDI service can be written by an organization that has the requisite staff and computer facilities. Programs may also be available from computer manufacturers or associations of users of particular computers. For input, bibliographic data on computer tapes are increasingly available from professional societies and commercial firms. SDI service itself may be purchased from societies and firms that will match files of reference data against SDI profiles supplied by the customer.

Future developments

The coming generation of SDI users may have direct access to a computer through the console of a remote interrogation station, with his SDI announcements displayed on a screen rather than being printed out as at present. NASA has underway a continuing study of remote interrogation of large document data files known as RECON (for remote console), in which it is planned to incorporate an SDI direct display capability. Even further in the future, the increasing capability of computers to organize raw data, together with advances in content analysis of documents, may permit the user to be alerted to new information rather than to merely a listing of relevant documents, and to display information in its order of significance. The ultimate goal is the enhancement by these means of all the SDI user's capabilities to contribute to the advance of science and technology.

Interactive Information Processing, Retrieval, and Transfer

J.C.R. Licklider
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

ACKNOWLEDGEMENT

This paper was prepared under the aegis of Project MAC, supported by the Advanced Research Projects Agency through the Office of Naval Research under Contract Nonr-4102 (01). In addition to some of the work of Project MAC, the paper describes some of the work of Project TIP and Project INTREX. Project TIP is supported by the National Science Foundation under Research Grant GN-589. Project INTREX is supported in part by a grant from the Carnegie Corporation, in part by Contract NSF-C472 from the National Science Foundation, and in part by the Council on Library Resources, Inc.

INTERACTIVE INFORMATION PROCESSING

By far the greatest part of the experience in interacting directly with computers comes not from experiments in information retrieval or transfer but from use of computers in preparing computer programs and in solving scientific and engineering problems. The earliest digital computers were programmed on line, and there has always been a bit of on-line-programming and "debugging" (elimination of program errors), but not until the advent of multi-access computing, based on the technique of "time sharing", was it possible for large numbers of people to work as a matter of course, day after day, at computer consoles. Now there are several hundred experienced console users at M.I.T. and many times that number elsewhere.

The most widely used interactive computer systems are two almost identical systems using IBM 7094 computers and a supervisory program called the "Compatible Time-Sharing System". Connected through a telephone switchboard to the 7094 computers are about 200 consoles. Most of the consoles are merely typewriters, I shall assume that your console is a typewriter. What you do first depends upon the kind of work you have before you. The most trivial thing to do --- and therefore a good introductory task --- is to write a memorandum. The computer will make it easy for you to correct typing mistakes and to effect editorial corrections, and then type out a "clean copy" of your memo.

You call the writing and editing program <u>typset</u> by typing "typset", and then you type the name you want to give the file you are going to create --- for example, "jsmith".

The computer then types "W1720.2" (which means "Wait a moment." and "It is now 2/10 of a second past 17:20 o'clock.") and then "R000.2+000.1" (which means "Ready" and "You have thus far used 2/10 second of processor time and 1/10 second of drum-core transfer time.") and then "Input" (which means "I am ready for input from you.") You start to type the memorandum:

```
To:  J. R. Smitt#h
     .space
     Subjeck:
     @Subject:  Plans for Improvement of planning Strategy
     .space
       Next Tuesday is the last day for ###to submit your ideas
```

The character (#) ordinarily means to erase a character. The character (@) ordinarily means to erase a line. The control word ".space" means to skip a line. Your memorandum therefor stands as:

```
To:  J. R. Smith

     Subject:  Plans for Improvement of planning Strategy

       Next Tuesday is the last day to submit your ideas
```

You notice that you forgot to give the date and that you need to capitalize the "p" in "planning". To go into "edit mode", you press the carriage return key twice. The computer thereupon types "Edit". You are supposed, of course, to know the control words (or characters), only a few of which appear in this example. You type: "i Date: June 15, 1968" (the "i" meaning "insert"), press the carriage-return key, and then, for format control, type ".space". Then, realizing that you want to center the date, you type "t" (for "go to the top of the file") and ".center" (for "center the next line"). In order to capitalize the "p" in "planning", you type "l plan" (for "locate the character string 'plan'") and "c /plan/Plan" (which changes "planning" to "Planning") and press the carriage-return key twice to go back to the input mode and complete your memo.

And when you have completed the typing --- and have corrected all your mistakes --- you file the memo by typing "file jsmith". The computer types "W1735.1" and then "R011.3 + 008.1" You type "runoff jsmith". The computer waits for you to put a fresh sheet of paper into the typewriter. You press the carriage-return key. The computer types the memo perfectly at 15 characters per second.

Such are the mechanics. They are very convenient if you want to prepare a long paper and don't type well.

Programs

One of the main aims of Project MAC has been to see what kind of a community would grow up around multi-access computers with rich and growing software resources. The system now has over a million words of public programs and over 6 million words of private programs. Let me illustrate with a few examples: A Mathematical Assistance Program makes transformations, and solves equations for you. It handles algebra, trigonometry, differential equations, Fourier and Laplace transforms, and so on. It plots graphs. You do not have to know all about it to use it: it asks you questions until it "understands" your problem.

Another program solves even quite complex problems in symbolic integration.

ADMINS facilitates the preparation, maintenance, and use of data bases.

TEACH teaches computer programming.

OPS is a large system of programs for interactive, incremental simulation and modelling. It provides a language in which you can define objects or entities and specify their properties and the relations. It lets you set into motion the situation thus described and make its behaviour unfold. It records the history and prepares summaries. It lets you intervene at any time to modify anything you like.

The On-Line Community

The foregoing paragraphs described a few of the hundreds of programs of the accumulated software resource. By using their programs, each user of CTSS can take advantage in his own work of pertinent efforts of his predecessors and his colleagues. An accumulative process is beginning to operate also in the domain of data.

Through computer-facilitated human interactions, a new kind of research community is arising at M.I.T. It is of course only in an early, formative stage, but there is little doubt that something significant is happening.

As the local on-line community has emerged from concept into actuality, the idea of a broader, geographically distributed community has taken form in the minds of several people. This idea involves inter-connecting several multi-access computer systems and combining their communities of users into a super-community. When geographically distributed computers and information networks come into being, their impact upon the process of information transfer may be great.

Some Conclusions Based on Project MAC's Experience

During its five years of operation, Project MAC has explored more fields than I can summarize, but let me, nevertheless, attempt to state some conclusions pertinent to information storage and retrieval:

(1) The computer turns information into a dynamic, living thing.

(2) Everything one does in an active informational environment is "complexity limited".

(3) Man-computer interaction is the most hopeful approach to the mastery of informational complexity.

(4) Even with the help of on-line interaction, it will take cooperative on-line teamwork to achieve significant solutions to the "big" problems of science, technology, industries, cities, nations, and alliances.

(5) The basic thing in the user's concept of an interactive information system is the "name space" of the filing (i.e., memory or storage) system.

(6) Although the term "time sharing" has achieved wide currency, the sharing of processor time is not fundamentally important. Much more important are memory sharing and communication. The aim of multi-access design should not be to make each user think he has a computer all to himself; it should be to immerse each user in a cooperative, interactive, computer-based community.

(7) The importance of controlled access to files can hardly be overstated.

(8) The importance of fast interaction makes itself felt when a problem gets complex.

(9) In a community with many interests, the "general-purposeness" of the general-purpose multi-access computer system has real meaning and significance. The system must lend itself to a great variety of applications and serve as ready host to diverse subsystems. Generality and open-endedness cost something, of course, but Project MAC's experience indicates they are well worth it.

(10) Reliable operation is vital and --- since the reliability will not be perfect --- effective "back-up" arrangements and recovery procedures are vital, also. Before they will invest their main intellectual capital in, or entrust it to, a multi-access computer system, people have to be confident that the system will be available when they want it and that it will not lose their valuable programs and data.

Interactive Information Retrieval

Project TIP uses the facilities of the 7094-CTSS multi-access systems. The main TIP data base is a growing collection of bibliographic data, presently from almost 100,000 journal papers in the field of physics. The TIP programs are programs for processing the data in ways formulated by the user during his interaction with the system. Using TIP, you type in lower case and the computer types back to you in all capitals. To explain what some of TIP's abbreviations mean, I shall insert comments in parentheses.

tip (You type "tip" to evoke the TIP program)
W1019.5 (Wait. It is 10:19 and a half.)
TYPE YOUR REQUESTS.
search annals of physics v.26 to v.28
find title pion not author boyling j.b.(Find all articles with titles containing "pion" except those by J.B. Boyling, whose work you already know.)
output print title a i and l (One-letter abbreviations are adequate for the TIP words such as "author", "identification", and "location". You could just as well have typed " o p t a i l" to instruct TIP to type as output the specified information about the items found.)
go (Go to work, TIP.)

ANNALS OF PHYSICS
VOLUME 26
VOLUME 27
J384 V027 P0079
DEUTERON PHOTODISINTEGRATION AND N-P
CAPTURE BELOW PION PRODUCTION THRESHOLD
PARTOVI F.
CAMBRIDGE, MASSACHUSETTS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LABORATORY FOR NUCLEAR SCIENCE AND
PHYSICS DEPARTMENT

VOLUME 28
J384 V028 P0034
ANALYSIS OF THE PHOTOPRODUCTION OF POSITIVE PIONS
HOHLER G.
SCHMIDT W.
GERMANY
TECHNISCHE HOCHSCHULE KARLSRUHE
INSTITUT THEORETISCHE KERNPHYSIK

(No article meeting the specification was found in Volume 26. One was found in volume 27. One was found in volume 28. "J384" stands for "ANNALS OF PHYSICS". If you ask for "pion you will find "pions", also --- but not if you ask for "pion*".)

The examples exercised only a few of the TIP commands, but it may have conveyed a notion of how one works with TIP. The TIP commands may be combined in many different patterns. Users develop ingenious strategies for filtering out irrelevant articles without losing the ones they want.

While some TIP programs make searches and print lists, other TIP programs take notes on how the system is used. The data thus collected are periodically analyzed, and modifications and adjustments are continually made. TIP has developed through a process of guided evolution.

A bibliography, two review articles, and a catalog of the books in the Student Center Library have been prepared with the aid of TIP, and it is now being used to prepare a catalog of the journal and periodical holdings of the M.I.T. libraries.

Most of the work done thus far with TIP has been hampered by the slow pace of typewriter output. We are looking forward eagerly to cathode-ray displays.

Interactive Information Transfer

The purpose of Project INTREX is to conduct experiments that will clarify design objectives, methods, and techniques for information-transfer systems of about 1975. Emphasis is placed on the word "experiments".

Experiments have been planned in four main areas:

1. bibliographic access,
2. physical access,
3. fact retrieval, and
4. network integration.

Thus far the project has concentrated on the first and second.

Bibliographic Access

The purpose of "bibliographic access", of course, is to take the user from a nebulous idea of what he wants to the accession numbers (or equivalent identifiers) of the documents that will satisfy his requirements. Most of the INTREX effort towards that end is centered upon a computer-based "augmented catalog" that will approximately contain 50 "fields" of information about each of approximately 10,000 documents in materials science and engineering: journal articles, theses, and reports as well as books. With so many kinds of information it should be possible to determine which kinds are helpful enough to warrant inclusion in an operational system. Much of the work of selection of the 10,000 documents is being done by the research people and will serve in the planned experiments. Consoles will be located in the Materials Science and Engineering Center and the Engineering Library, and the experiments will be conducted within the context of actual use.

Physical Access

Bibliographic access must of course lead directly to physical access. Limitations of the present technology make digital storage and processing of a library-sized corpus uneconomic. The course being followed by Project INTREX is therefore to hold the substantive documents themselves in a non-digital microform storage system associated with the computer, and to use the computer to execute their delivery to the user. Images of the pages of the 10,000 documents are being made in microfiche, and a computer-controlled subsystem for picking out and scanning selected pages is being constructed. Plans call for experimental investigation of such interrelated factors as the speed, the form, the resolution, and the cost of physical access. By varying the parameters and making measurements of preference and performance under conditions of actual use, optimal engineering compromises will be approached and design objectives for operational systems will be formulated.

Advanced Experiments in a Library Context

Looking beyond the experiments with the 10,000-item collection in materials science and engineering, Project INTREX is conducting design studies that postulate a corpus of a million documents. At the same time, the M.I.T. Engineering Library is being reconstructed in such a way as to provide for simultaneous operations in conventional and computer-based modes. Card catalogs, book stacks, reading tables, microform equipment, and computer consoles will be brought together in an arrangement designed for advanced experiments in an operational library setting.

Synthesis and Prospect

Throughout MAC, TIP, and INTREX, and indeed throughout M.I.T., there is a feeling that a great and fundamental change is taking place in the way men relate to information. The force behind the change is the computer, of course, but it is not the same computer we have known these last 20 years. It is the computer cast in the new role of the moldable and retentive yet dynamic medium --- the medium within which one can create and preserve the most complex and subtle patterns and through which he can make those patterns operate (as programs) upon other patterns (data) derived from nature or the works of other men. In that role, the computer will change the very nature of libraries and information systems.

Man-Machine Interface

W. Händler

The problem of man-machine interface can be traced back to the development of combustion- and steam-engines in the last century. The rather clumsy push-buttons and levers of the initial period were largely replaced by switches and relays with the introduction of electric power. Human interaction with the machine took place mainly during the primary power-push. Subsequently only minor corrections in the amount of steam supplied or the number of resistors in operation required human intervention. In general man-machine interaction was thus confined to checking power supply and motion of the machine.

Apart from the development of nuclear energy data processing techniques may well prove to be the most essential invention of our century to affect man's environment. We may therefore legitimately inquire as to the present state of man-machine interface and its possible future trends, not stressing however, any more or less accidental technical realizations.

The first computers designed more than 20 years ago were very much characterized by the notion of entirely predetermined computational processes set off by a single starting signal. Some years later the computer was equipped with additional switches, among them so-called selector-switches for guiding the course of the programm along alternative pathways.

At first the inventors themselves who were intimately tied to their creation operated the system. Soon other people came to use the machine and accomplished their work on the basis of a start-button philosophy.

Frequently there resulted a feeling of utter resignation in the face of the microsecond. The user could hope to trace computational progress only in a very crude, overall manner, perhaps merely with regard to the beginning and end of his job. However, noteworthy efforts were made repeatedly in the attempt of re-establishing a somewhat closer contact to the ever faster growing computer. Suitably installed loudspeakers, for example, provided a means of global monitoring.

With further increase of speed it became necessary to exclude direct customer operation in favor of batch-processing thus saving costly computer time previously wasted on tedious input-output procedures. Trained people undertook to devote their endeavour to the task of empirical scheduling. In this manner the problem of man-machine interface found a somewhat fictitious solution; the real user rarely came to see the computer or press any of its buttons. Man-machine interface was restricted to a rather small group of trained personnel.

Consequently customer and computer were alienated from each other, a small group of technicians - the operators - providing their sole common tie and most flexible but entirely human interface because man is still the most adjustable and adaptive instrument.

The disadvantage of the method described is due to missing the rather vital experience of observing the machine do your assignment. If you get an error-message - say on sheet 2 of your programm while debugging at the console you still have a fair chance of achieving your objectives quickly. If on the other hand you are kept at a distance from the computer you may well take weeks to accomplish the same end.

This kind of disproportion can be corrected in our days by allotting a substantial number of consoles to various permanent customers. Certainly the computer-frames still remain outside the user's field of vision. Yet he will be in touch with the machine's essential responses to the task he posed. Certain disturbing effects, for example sharing the same central unit with numerous other customers, should pass unnoticed by him. To be sure, typewritten or teletyped communication is only one of the present possibilities and - having originated in a pre-computer-age not necessarily the most efficient one.

Presently we are investigating new forms of contact with the computer. Prime importance must be attached to the development of CRT-display methods which permit almost instantaneous transmission of mixed-mode text or picture information eliminating the need of tedious waiting for carriage return or line feed motions of printer or teletypewriter. Via display-screen the computer is able to offer alternatives to be chosen effortlessly and swiftly with the aid of a lightpen, for instance.

In the past most transitions to advanced techniques have been marked by their initial tendency to duplicate already existing achievements using the new means. But when we employ the display consoles mentioned we will have to do some fundamental rethinking. Evidently the flow of information out of the computer can be speeded up well beyond the limits hither-to known. We must ask for the maximal amount of information per unit of time and the optimal manner of presentation best suited for processing by the human user.

There are several distinct categories of representation:

alphanumeric text
black and white shading
scaled shading
colored pictures
moving pictures.

They may occur on the screen separately or in different combinations with each other.

Determination of suitable quantities or qualities of information is exceedingly difficult. Many partially subjective factors are involved. Our judgement will also depend on the particular topic under consideration. Graphic representations may be of insignificant value when applied to codes of law or administrative regulations. Alphanumeric text on the other hand is inadequate for describing processes of industrial production or technical design. If you interpret the list above as ordered according to quality it may be valid perhaps in a majority of cases but certainlay not in general. We are still short of experiences in many specific areas of application. This gap should be filled soon by means of well-planned experiments and numerous measurements separating objective from subjective components wherever possible.

Another important aspect must not be disregarded. Up to now most display applications have been of a rather static nature in spite of some rudimentary attempts at making computer-controlled movies. Usually considerations have centered around the effects of single pictures rather than taking account of the complete sequence of textual and graphic information as a totality meant to establish dynamic contact between user and machine.

Lately a great deal has been said about lengthy waiting times leading to disturbed contact and distracting the user from his task. Conversely an excessive reaction speed of the computer may "overfeed" the user and cause in him what the psychologists call a "mental block".

Situations of this type are quite conceivable and have indeed been observed already. Some people prefer punching their program on tape or cards before feeding its presumably final version into the computer. They tacitly assume the computer expects reactions without delay and seem to feel pressed for time. Reasonable programming for time sharing systems should avoid creating this feeling.

So far I have mentioned a number of mainly psychological factors influencing man-machine interface.

In contrast a rigid and formal access to the subject is opened by the mathematical theory of automata. Here the user is considered to be a finite automaton in the sense of a working hypothesis. Finite automata are capable of occupying a finite number of states. On reception of an input symbol they pass in a well-defined manner from one state to another emitting an output symbol. Present-day theories however are still inadequate to cope successfully with the formidably large number of states actually existing in both the computer and the human organism.

We may expediently conceive the notion of a telescoping sequence of successively refined structural levels one inside the other in order to facilitate the analysis of our subject matter. For example take Runge-Kutta integration of differential equations. The method may be interpreted as representing a certain state either of man or of the machine. A close-up view of the algorithmic procedure would of course analyse this state into a manifold meshwork of simpler ones. "Runge Kutta" in turn is a refined state relative to the coarser level containing the state of - say - "Numerical Analysis". Transitions from refined to coarser structures are reflected for instance algebraically by chains of homomorphisms.

Within the framework of present-day theory the automaton is called upon to establish equivalence classes of patterns (for example the class of all syntactically correct sentences of a language) or - less ambitiously - perhaps to make decisions regarding the membership of a given pattern to one of those classes. Another part of the theory is devoted to the design of a suitable series of experiments for determining the initial state of an automaton. These investigations can and should be extended to the reciprocal effects of two automata on each other - their dialogue - and to the mutual interaction of many automata - their communication.

The axioms and conclusions of mathematical theory will depend on the fundamental concepts underlying our comprehension of man-machine interface.

One view is dualistic. It treats the machine as man's conversational partner with special abilities for solving certain problems. The other view can be called monolithic and teleologic in a way.

Man and machine are thought of as a single integrated system organized for the purpose of solving a problem or mastering the environment.

The latter notion is unfortunately subject to severe practical limitations because we still possess too little knowledge of how information processing takes place within the human nervous system.

Once deeper insights into these mechanisms have been gained we stand a good chance of constructing very effective man-machine interfaces on the basis of the concept last mentioned.

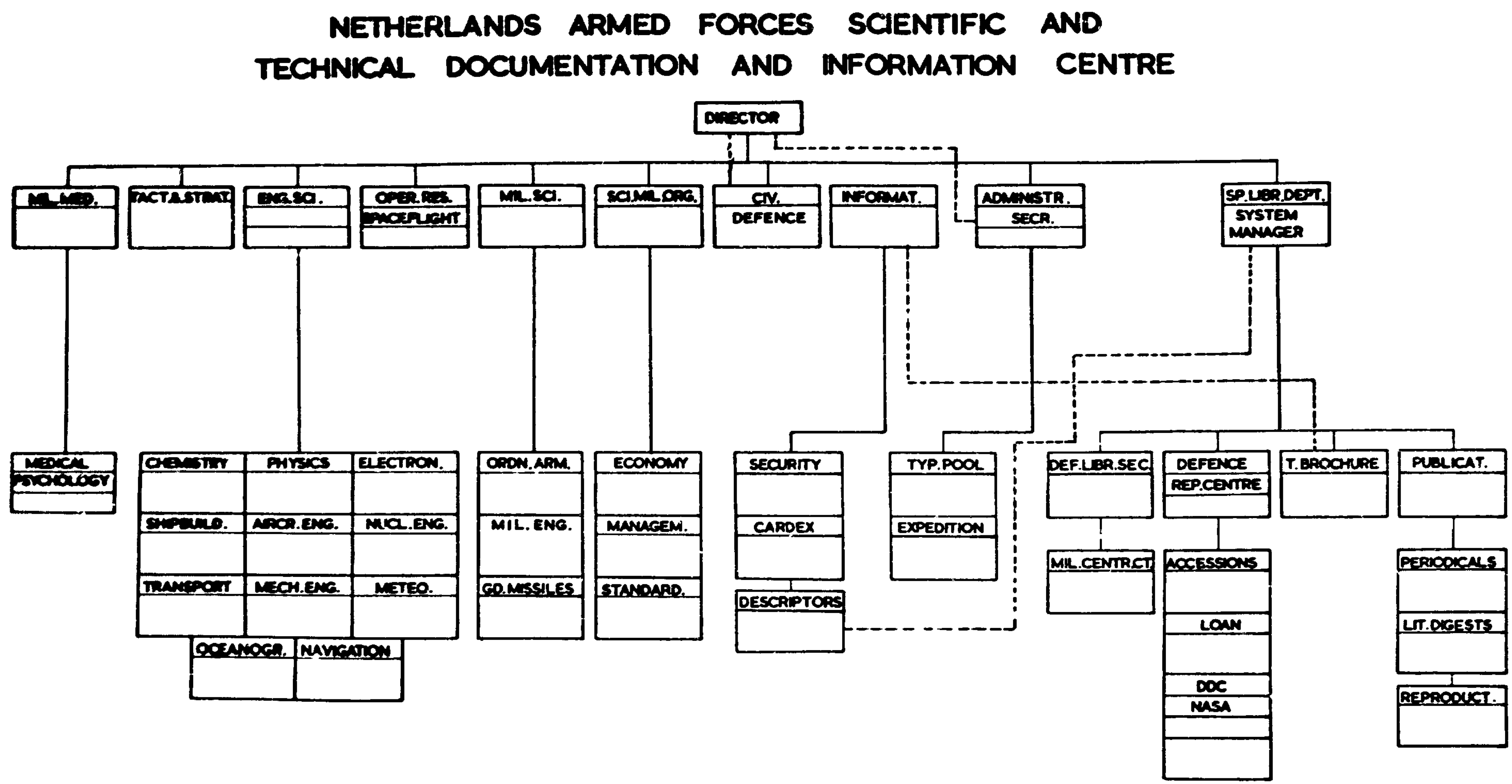
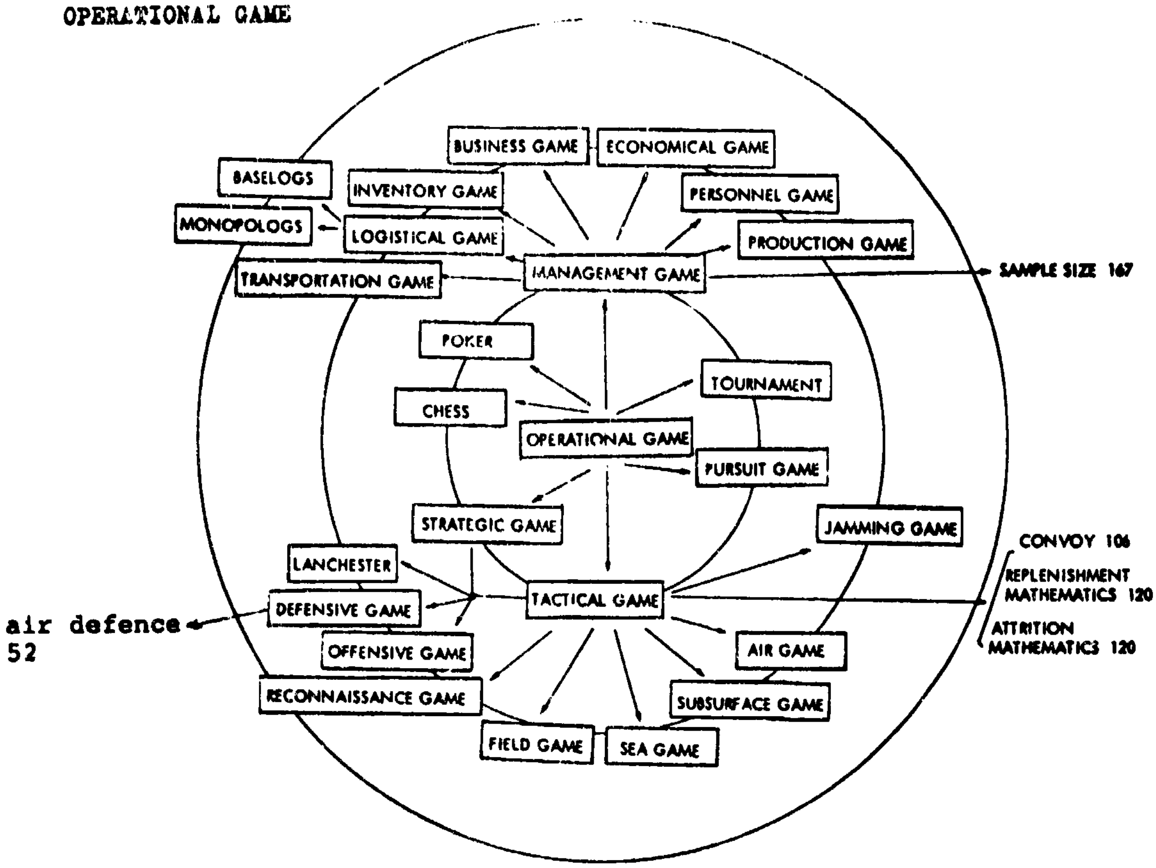

NETHERLANDS ARMED FORCES SCIENTIFIC AND
TECHNICAL DOCUMENTATION AND INFORMATION CENTRE
DIRECTOR
MIL.MED.
TACT.& STRAT.
ENG.SCI.
OPER.RES.
SPACEFLIGHT
MIL.SCI.
SCI.MIL.ORG.
CIV.
DEFENCE
INFORMAT.
ADMINISTR.
SECR.
SP.LIBR.DEPT.
SYSTEM
MANAGER
MEDICAL
PSYCHOLOGY
CHEMISTRY
PHYSICS
ELECTRON.
SHIPBUILD.
AIRCR.ENG.
NUCL.ENG.
TRANSPORT
MECH.ENG.
METEO.
OCEANOGR.
NAVIGATION
ORDN.ARM.
MIL.ENG.
GD.MISSILES
ECONOMY
MANAGEM.
STANDARD.
SECURITY
CARDEX
DESCRIPTORS
TYP.POOL
EXPEDITION
DEF.LIBR.SEC.
MIL.CENTR.CT.
DEFENCE
REP.CENTRE
ACCESSIONS
LOAN
DDC
NASA
T.BROCHURE
PUBLICAT.
PERIODICALS
LIT.DIGESTS
REPRODUCT.

Fig. 1

TDCK SOURCES INDEX	LOCATION	AIRCRAFT ENGINEERING	ARMAMENT	CHEMISTRY	CIVIL DEF EMERG PLAN	ECON.-MANAGEMENT	ELEC ENGINEERING	MECHANICAL ENGINEERING	METEOROLOGY	MIL. ENGINEERING	MIL. MEDICINE	NAVIGATION	OCEANOGRAPHY	OPERATIONS RESEARCH	PHYSICS	PSYCHOLOGY	SHIPBUILDING	SPACE FLIGHT	TACT - STRATEGY	TRANSPORT		
CARD SYSTEMS																						
TDCK COMPACT SYSTEM	K	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
UDC	K	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
CODE CLAS LUCHTVAART CCL	K	•								•		•										
INST SCHEEPVAART EN LUCHTVAART ISL	K							•									•					
TECHNICAL BROCHURES	BROC	•		•			•	•	•	•	•	•			•		•			•		
COOPERATIVE AUTHORS INDEX	RC	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
BOOKS OF REFERENCE																						
CODEX MEDICUS	G										•											
DECHEMA WERKSTOFF TABELLE	C			•																		
ENCYCLOPEDIA CHEMICAL TECHNOLOGY KIRK - OTHMER	C			•																		
MEDICAL DICTIONARY DORLAND'S	G										•											
ABSTRACT JOURNALS																						
AED-AB INFO ZUR KERNFORSCHUNG UND -TECHNIK	N														•							
ARMS CONTROL AND DISARMAMENT	PUBL					•													•			
ASME INDEX 1880-1956	W	•		•				•									•			•		
AUSTRALIAN SCIENCE INDEX	PUBL												•									
BATTELLE TECHNICAL REVIEW ABSTRACTS	PUBL	•		•			•	•							•							
BSRA ABSTRACTS	S			•				•				•			•		•					
BULLETIN SIGNALETIQUE SDIT	L	•	•	•			•				•	•		•	•	•						
CORROSION ABSTRACTS	C			•			•	•									•			•		
ELECTRONICS ABSTRACTS	E						•															
ENGINEERING INDEX 1950	W	•	•	•			•	•	•	•	•	•	•		•		•			•		
INDEX AERONAUTICUS	L	•								•	•	•										
INDEX MEDICUS	G										•					•						
INDEX US GOV RESEARCH REPORTS	INFO	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
INTERNATIONAL ABSTRACTS ON OPERATIONS RESEARCH	OR													•								
INTERNATIONAL AEROSPACE ABSTRACTS	L	•		•			•		•		•	•			•							
METEOROLOGICAL AND GEOASTROPHYSICAL ABSTRACTS	M								•													
MONTHLY CATALOG USA	INFO	•	•	•			•	•	•	•	•	•			•		•		•	•		
NLL TRANSLATIONS BULLETIN	PUBL																					
RES AND DEV ABSTRACTS. MIN OF TECHNOLOGY	INFO	•	•	•			•	•	•	•	•	•		•	•					•		
SCIENCE ABSTRACTS. A PHYSICS ABSTRACTS	N														•							
SCIENCE ABSTRACTS. B ELECTRICAL ABSTRACTS	E						•					•										
SCIENTIFIC AND TECHN AEROSPACE REPORTS INDEX NASA	RC	•		•			•		•		•	•			•							
SOVIET ABSTRACTS MECHANICS. MIN OF TECHNOLOGY	PUBL	•	•	•				•	•	•					•		•					
STATISTICAL THEORY AND METHODS ABSTRACTS	OR													•								
TECHNICAL ABSTRACTS BULLETIN DDC	RC	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
TRANSATOM BULLETIN EURATOM	PUBL			•							•				•							
US AIR FORCE RESEARCH RESUMES	RC			•							•				•							
US GOV-WIDE INDEX TO FED RES AND DEV REPORTS	RC																					
US NUCLEAR SCIENCE ABSTRACTS	N														•							
VDE SCHNELLBERICHTE	E						•															
WORLD INDEX OF SCIENTIFIC TRANSLATIONS	PUBL																					
SELECTED RAND ABSTRACTS	RC	•	•		•	•	•		•		•	•		•	•				•			
ENVIRONMENTAL EFFECTS ON MAT AND EQUIPM ABSTRACTS	C		•	•			•															

Fig. 2

OPERATIONAL GAME



OPERATIONAL GAME

Operational game	E	4	R
Tactical game	E	14	R
Defensive game	E	16	Q
Reconnaissance game	E	9	R
Air game	E	13	Q
Air defence		52	

Air Game	E	13	Q
Baselogs	O	6	D
Business Game	E	14	Q
Chess	E	15	Q
Defensive Game	E	16	Q
Economical Game	E	17	Q
Field Game	E	18	Q
Inventory Game	E	19	Q
Jamming Game	E	20	Q
Lanchester	O	4	D
Logistical Game	E	1	R
Management Game	E	2	R
Monopologs	O	5	D
Offensive Game	E	3	R
Operational Game	E	4	R
Personnel Game	E	5	R
Poker	E	6	R
Production Game	E	7	R
Pursuit Game	E	8	R
Reconnaissance Game	E	9	R
Sea Game	E	10	R
Strategic Game	E	11	R
Subsurface Game	E	12	R
Tactical Game	E	14	R
Tournament	E	13	R
Transportation Game	E	15	R

fig.3

STORAGE AND RETRIEVAL OF INFORMATION

The user - supplier dialogue

E D U C A T I O N

by Felix Liebesny
B.Sc., F.I.Inf.Sc., F.I.L., A.L.A.
Head of Group Technical Intelligence Service
The British Aluminium Co. Ltd.

ABSTRACT. The author discusses the educational requirements for users and suppliers of scientific and technical information and the steps taken to provide professional education for various levels of attainment.

The very magnitude of the problems connected with the information explosion is such that several approaches have to be made in the attempts towards overcoming them. Whether we use such almost meaningless figures as

1. that the world's output of scientific and technical articles is of the order of 1,000,000 per annum;
2. that the number of periodical titles in the disciplines of science and technology increases by two every day;
3. that about half of the world's literature in those self same disciplines is written in languages other than English (which for the purposes of this argument includes American English); or
4. that the periodical literature increases at a compound rate of approximately 6-7% per year,

we are still left with the almost irrefutable fact that all of us, users and suppliers alike, are gradually becoming submerged by this flood of paper. Although the argument of quality of the literature is being ignored in such discussions, it is obviously becoming increasingly difficult to cope with the output of the world's presses. This tendency - which is certainly not new - has led over the years towards more and more specialization, both in the user and supplier of the specialist literature. This has led on the one hand to the emergence of the information scientist, and on the other hand to, albeit, isolated endeavours to train the user in some of the more elementary forms of the techniques employed in coping with the literature.

The term 'information science' - though of comparatively recent origin - has acquired in this short space of time several shades of meaning. Therefore in order to define the way in which the term will be used in this paper, the activities embraced by the definition as given in the Articles of the Institute of Information Scientists should provide some guidance. These are:

(1) abstracting, reviewing progress and other similar technical writing;
(2) translating scientific and technical writings;
(3) editing such writings as emerge from (1) and (2) above;
(4) indexing, subject classification and retrieval of scientific and technical information;
(5) searching scientific and technical literature, preparing bibliographies, reports, etc.;
(6) providing scientific and technical information and tendering advice thereon;
(7) dissemination of information and liaison and field work for that purpose;
(8) research on problems in information work.

It is obvious from this recital that a professional information scientist should combine a considerable knowledge and skill for the proper execution of his work; normally this knowledge is obtained from courses of study while the skill should be derived from appropriate practical experience such as work in an information department. In order to attain a corporate membership in the Institute of Information Scientists it is thus necessary to provide evidence of both the required knowledge and skill; thus for Membership a candidate would normally be expected to possess a science degree and to have worked at least five years in an information department.

In training the user to enable him to deal competently with the documentation of his special subject field it would be unwise to aim at such a high degree of professionalism; firstly, a detailed training programme would turn the user into an information scientist and thus divert him from his own special field; secondly, it would be wasteful to impart knowledge and expertise of which a considerable amount would never be required by a specialist user of the literature since the full training of a documentalist involves matters relating to several disciplines; and thirdly, the time generally available for such training of the user is not sufficient for more than a somewhat superficial approach to the many problems of information storage and retrieval.

Therefore the few attempts that have been made to familiarise the scientist and technologist with proper means of using his subject literature require careful study to elicit their useful and successful features. The results of these courses are, however, not quite so easy to establish as most of them have only been conducted for a few years and it is thus still too early to quantitatively assess this criterion.

Before entering into a more detailed discussion of the courses available it must be stated - perhaps somewhat shamefacedly - that information on such educational ventures is not very easy to come by; it appears that the discipline of information science is not too well provided with means of keeping its practitioners informed on what is going on elsewhere. Although there are many periodicals and even abstract journals in the fields of librarianship, documentation, information science etc. their coverage on the educational front is not too extensive. It was for that reason that the Federation Internationale de Documentation (FID), the (British) Office for Scientific and Technical Information, Aslib and the Institute of Information Scientists organised in 1967 an International Conference on Education for Scientific Information Work. The proceedings of this conference were published in September 1967 by FID and its 33 papers tried to survey the international scene by focusing on the activities in the most important countries. That this attempt was not 100% successful can be deduced from the fact that there were no contributions from the USSR and that the Indian delegates did not attend the conference and thus take part in the discussions. However, these Conference Proceedings seem to constitute the most comprehensive review of the activities in educating users and suppliers of scientific information.

Training of user

The user's training can be initiated at two levels: either before he graduates or afterwards.

Training at the undergraduate stage is frequently very difficult because most of the syllabi are so crowded that it requires considerable persuasion of the university authorities to devote any of that precious time to such peripheral subjects as documentation. Furthermore, there is still a great deal of that old belief that every scientist knows - or at least should know - the literature of his own subject. This fallacious attitude ignores completely any subsequent developments in the subject field and its literature or the use of modern techniques in dealing with it.

Training at the post-graduate stage, on the other hand, is likely to ensure that the recipient is in the proper frame of mind to accept the training as by then he is more mature and more aware of his needs with respect to documentation.

1. Undergraduate training: in the United Kingdom organized training of students is conducted at the Universities of Liverpool and Bradford [1] where the university authorities took the initiative, and at a series of six universities [2] (Edinburgh, London (University College, and Chelsea College of Science and Technology), Oxford, Warwick, and York) where the Office for Scientific and Technical Information is arranging for some 500 students to receive an information service and some instruction in its use.

2. Postgraduate training: the main centre for education in the use of literature at that level is undoubtedly the National Lending Library for Science and Technology, at Boston Spa, Yorks. where courses lasting about 10 working days have been run since 1965. A specimen timetable for such a course on the use of scientific literature is given in Appendix I.

   A more recent development is now being supported by OSTI whereby retrospective searching of the medical literature by means of the MEDLARS technique (MEDical Literature Analysis and Retrieval System) will be made easier by providing facilities for consultation with specially trained liaison officers. Five such officers - at Newcastle, London, Edinburgh and two yet to be appointed - will help users in the formulation of search profiles and will advise on availability and capacity of this computer-tape index. Courses are also being held at the Hatfield College of Technology.

Training of supplier

The education of the librarian, information scientist, documentalist, etc. is perhaps somewhat better organized than that of the user. Nevertheless, there are today so many different avenues of training towards producing a qualified information scientist that this plethora of facilities may create an impression of diffusion and even confusion. In order to reduce this seemingly unmanageable pile of information into some state of order it may be advisable to classify these data according to the type of supplier to be produced:

1. Librarian:

   a. Chartered: organised full-time courses of two years duration are provided at several library schools which are housed within colleges of further education, polytechnics or universities.
   Recent developments towards creating a degree in librarianship (as opposed to a post-graduate qualification as is provided at University College, London) under the aegis of the Council for National Academic Awards have

led to the setting up of such a course at Newcastle to commence in the autumn of 1968. Other such courses are likely to commence with the following academic year.
Another scheme now under consideration by the (British) Library Association is to permit graduates to amass a number of credit points (say 30) so that they may become registered as chartered librarians. These credit points can be obtained by attending specified courses and by possessing relevant professional qualifications (e.g. linguistic, legal, computer technology).

b. assistant: it was felt some time ago that a qualification of a lower level of competence than that of the chartered librarian should be provided for those people having to interrupt their professional life - especially married women - or those not wishing to attain the higher professional status. Towards this end, a Library Assistants Certificate has been proposed; this scheme will be administered by the City and Guilds Institute of London.

2. Information scientist: since 1961 a post-graduate course has been run at the City University in London (formerly the Northampton College of Advanced Technology). This two-year course is run twice weekly (two hours each) to enable students to attain an entrance qualification for the corporate membership of the Institute of Information Scientists. A post-graduate one-year full-time course was started in 1963 which, since 1967, will lead to a M.Sc. degree.
Similar courses have been running at the University of Sheffield Postgraduate School of Librarianship and Information Science since 1964. At present this one-year course leads to a Diploma, but from October 1968 will lead either to a M.Sc. degree in Information Studies or in Librarianship.

Owing to the previously mentioned fact that about half of the world's literary output by scientists and technologists is written in languages other than English and that many of the readers of such writings are only too rarely qualified to comprehend those foreign languages, it is obviously important that the suppliers of information should possess some proficiency in handling foreign language material. Therefore many of these courses lay considerable stress on such ability and even include some training in foreign languages in their syllabus. In the M.Sc. course at the City University one examination paper is devoted to testing the required level of proficiency.

While the foregoing survey has been directed largely to activities and developments in the United Kingdom it should not be thought that the rest of the world is standing still. Indeed, as R. T. Bottle[3] has shown, similar schemes for training the user are in operation or being planned in many countries of Europe (to which his survey was confined) and in Australia[4].

In the user-supplier dialogue which forms the theme of this symposium it is essential that each party should understand and appreciate the basic requirements of the other; this can be achieved, inter alia, by proper teaching and training in the elements of storage and retrieval of information for which usually too little time is available in the over-crowded timetables during the university courses. The few endeavours described above are hopeful omens that the need for a deeper understanding is being more widely realized.

References:

1. BOTTLE, R. T., Proc. intern. Conf. Educ. sci. Inf. Work, 1966, 59-69.
2. SOMMERFIELD, G.A., Chem. in Britain, 4(2), 71-73, Feb., 1968.
3. BOTTLE, R. T., J. chem. Educ., 6(1), 3-6, Feb., 1966.
4. WARD, J. L. (ed.) Library services for chemists. Royal Melbourne Technical College Press, 1960.

APPENDIX I

The use of scientific literature - A course for research students

TIMETABLE

Day	Time	
1	2.00 p.m.	Tour/description/services of the National Lending Library for Science and Technology (in particular Reading Room and Staff Library).
2	9.15 a.m.	Guides to published information
		1. Serials: Current awareness tools Abstracting journals Indexing journals Annual reviews Review serials
	10.30 a.m.	2. Tools for student's specified interest
3	9.15 a.m.	3. Reports: Indexes
	10.30 a.m.	4. Books: Theses Annual Reports Yearbooks Monographs Technical dictionaries Language dictionaries Encyclopaedias Bibliographies
4	11.15 a.m.	Language problems
5	11.15 a.m.	Record keeping
6	11.15 a.m.	Information bureaux
7	11.15 a.m.	Keeping up with current literature
8	11.15 a.m.	Library resources in the U.K.
9	11.15 a.m.	Films: The National Lending Library for Science and Technology National Library of Medicine, U.S.A.
10	11.15 a.m.	Criticism and discussion of the course

Time not devoted to lectures will be spent on literature searching.

(N.B. As can be seen from this schedule use is made of the few films available on this topic).

Source: J. Doc., 22(1), 22-32, Mar., 1966.

Information regarding the availability of further copies of AGARD publications may be obtained from

The Scientific Publications Officer,
Advisory Group for Aerospace Research and Development,
7, rue Ancelle,
92 Neuilly-sur-Seine,
France.