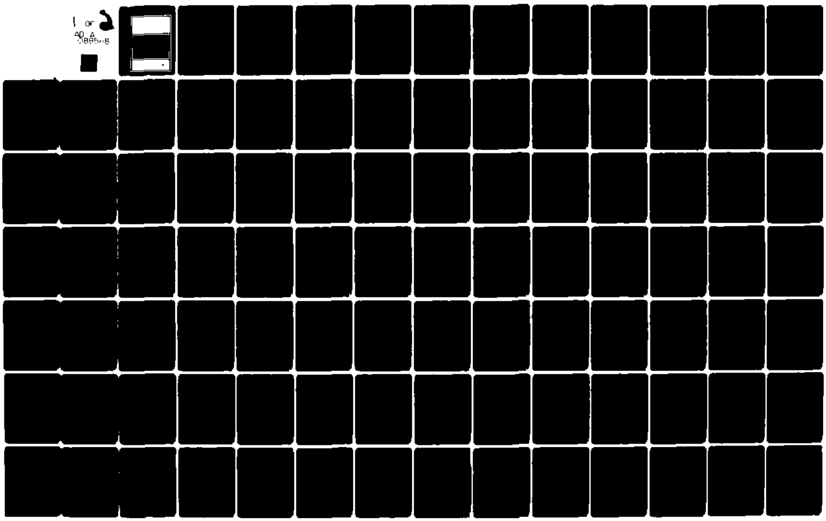
AD-A088 588 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 5/2
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION.(U)
APR 80 K N KOUROGENIS, W A MARTIN
UNCLASSIFIED AGARD-CP-279 NL

1 OF 2
AD A
088588



AGARD-CP-279

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No. 279

International Access to Aerospace Information

DTIC ELECTE 1980 A

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY ON BACK COVER

CP 279

AGARD-CP-279

INTERNATIONAL ACCESS TO AEROSPACE INFORMATION

AGARD-CP-279

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.279

INTERNATIONAL ACCESS TO AEROSPACE INFORMATION

Copies of papers presented at the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the authors.

Published April 1980

ISBN 92-835-0264-7

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London, W1P 1HD

THEME

One of the main elements of the work of the AGARD Technical Information Panel (TIP) is to assist NATO's aerospace research and development activities by improving the effectiveness of scientific and technical information systems which exist in aerospace and defence fields in the various member nations of the Alliance. The choice of theme for this 1979 Specialists' Meeting stems from consideration of this aspect of TIP's work, particularly in relation to the host country, namely Greece. Development goes on constantly, and the Meeting will review recent changes not only in the channels and media of information transfer but also in the nature of the information itself. Arising out of this, it is hoped that firm plans can be laid for a broadening and extension of existing information transfer facilities in the fields of aerospace and defence R and D in particular areas of the NATO Community.

TECHNICAL INFORMATION PANEL

CHAIRMAN: Ir A.S.T.Tan
National Aerospace Laboratory (NLR)
P.O.Box 90502
1006 B.M.Amsterdam
The Netherlands

DEPUTY CHAIRMAN: Mr H.E.Sauter
Administrator, Defense Technical Information Center
Defense Logistics Agency
Cameron Station
Alexandria, VA 22314 USA

PROGRAMME AND MEETING OFFICIALS

PROGRAMME COMMITTEE CHAIRMAN: Ir A.S.T.Tan
(see address above)

PROGRAMME COMMITTEE MEMBERS:

Mr R.Bernhardt
Leiter der Hauptabteilung Datenverarbeitung
GID, Herriotstrasse 5
D-6000 Frankfurt am Main 71,
FRG

Mr D.W.Goode
Chief Librarian
Royal Aircraft Establishment
Procurement Executive, MOD
Farnborough, Hants GU14 6TD
UK

Ing. Gén. C.Guilleminet
Directeur Adjoint du CEDOCAR
26 Bd Victor
75996 Paris Armées,
France

Mr H.K.Krog
Managing Director
Norwegian Centre for Informatics
Forskningsveien 1
Oslo 3
Norway

Mr H.E.Sauter
(see address above)

MEETING COORDINATOR

Captain G.N.Goulios
KETA
Delta P Phalirou
Palaion Phaliron
Athens
Greece

PANEL EXECUTIVE

Mr E.T.Sharp

CONTENTS

TECHNICAL EVALUATION REPORT

by

D.W.Goode
Chief Librarian
Royal Aircraft Establishment
Procurement Executive, MOD
Farnborough, Hants GU14 6TD
UK

SUMMARY

The 1979 AGARD Technical Information Panel (TIP) Specialists' Meeting was held in Athens on 17–18 October at the Hellenic War Museum and was entitled "International Access to Aerospace Information". Choice of the Meeting's theme stemmed from TIP's commitment to improve the effectiveness of scientific and technological information systems in aerospace and defence fields within the NATO member countries particularly in relation to the needs of the host country, Greece.

The programme was divided into four sessions of three papers each, with the exception of Session IV which had only two papers but was followed by a Workshop Session. Sessions I and II were held on 17 October and were entitled "Official Access Facilities for Aerospace Information" and "Requirements and Tools for International Cooperation and Data Exchange". Sessions III and IV, held on 18 October, were entitled "Problems of Utilization of Aerospace Literature" and "Non-Literature Data in Aerospace and Development". Each Session was chaired by a Member of the Technical Information Panel and in addition had two Reviewers (also from TIP) to stimulate discussion following the presentation of each paper. Papers presented included some at an advanced level and also those of a more general nature for the benefit of those member countries where information retrieval systems were still in the early stages of formulation and development.

It was felt that the Meeting successfully attained its objectives and this was also the conclusion of the AGARD Technical Information Panel's Business Meeting held immediately afterwards. The Greek TIP delegates stated that they had obtained many useful ideas for their forthcoming task of setting up their National Documentation Centre, together with its information retrieval system. All attendees expressed their appreciation of the excelient facilities and social activities provided by the Greek authorities.

The complete Proceedings of the Specialists' Meeting for 1979 will shortly appear as AGARD CP 279.

1. INTRODUCTION

The choice for the Meeting's theme, "International Access to Aerospace Information" stemmed from the Panel's commitment to improve the effectiveness of scientific and technological information systems in the aerospace and defence fields within NATO member countries. Recent changes in the channels and media of information transfer were reviewed with the particular needs of Greece, the host country, in mind. It was hoped that all NATO member countries might benefit in their work of broadening and extending their own existing information transfer facilities.

The Meeting was divided into four Sessions, two being held each day. These were as follows:

Session I – Official Access Facilities for Aerospace Information. Session II – Requirements and Tools for International Cooperation and Data Exchange.	Held on Wednesday, 17 October 1979
Session III – Problems of Utilization of Aerospace Literature. Session IV – Non-Literature Data in Aerospace and Development.	Held on Thursday, 18 October 1979

Each Session contained three papers with the exception of Session IV which had only two but was followed by a Workshop Session. A brief summary of the papers together with the main points of the resulting discussion is given below. Each Session was chaired by a Member of the Panel and also included two additional Panel Members acting as

Reviewers to stimulate discussion. The present format for TIP Specialists' Meeting was retained in that the presenter, rather than reading his paper as a whole, simply emphasises his most important points during the first 20 minutes to be followed by comments from the Reviewers and discussion involving attendees in general. The time allocated for each paper was approximately 45 minutes.

2. SESSION I: OFFICIAL ACCESS FACILITIES FOR AEROSPACE INFORMATION

Chairman: Mr R.BERNHARDT, Gesellschaft für Information und Dokumentation, FRG
Reviewers: Mr A.BRUCE (UK), Mr D.W.GOODE (UK)

2.1 Paper 1: Access to Aerospace Information in the Greek Situation

Dr K.KOUROGENIS, YEET, Greece

Dr Kourogenis' paper dealt with the present "state-of-the-art" of information retrieval and exchange in Greece and also the formulation of a policy for a central system based on a National Documentation Centre. During discussions between Dr Kourogenis, the Chairman and the Reviewers, it became clear that the planning was at a very early stage. No decision had yet been made as to whether civil and military requirements would be separated (as in USA and UK) or whether the proposed centre would act as a major loan and bibliographic centre as does the British Library. As Reviewer, I remarked that a recent survey of the "Science Citation Index" between 1974–77 attributed over 2,000 papers to research centres in Greece and entirely agreed with Dr Kourogenis that a centre of the type proposed was very desirable. Dr Kourogenis was unable to comment at this state on the involvement of Demokritos – the Athens research centre for fundamental science, but thought it might well have an important part to play in the future. He also stated that he felt that the present Meeting should be useful to the Greek authorities in providing details of systems and services available in other NATO countries. Although no figures on the present use of computerised on-line information retrieval equipment within Greece were available, he nevertheless hoped that a link with EURONET might be established in the not-too-distant future. It was felt that this paper was most useful to the Meeting attendees in general from the point of view of describing the local scene.

2.2 Paper 2: European Cooperation in the Field of Aerospace Information

W.A.MARTIN, ESRIN, Italy

This was an interesting and very well presented paper which gave details of the origins and establishment of the ESRO/ELDO Space Documentation Service and the present functions of the ESA/IRS system. The presenter elaborated considerably on the framework of his paper and gave many interesting indications as to how future services might develop, including the philosophy behind ESA's decision making processes. The relationship between ESA and the Commission of the European Communities (CEC) was explained in the context of EURONET. Questions answered by Mr Martin during the discussion included advice concerning on-line document ordering and the availability of the ESA/IRS services to Greek centres and establishments.

2.3 Paper 3: The Role of NASA for Aerospace Information

G.P.CHANDLER Jr, NASA, USA

Mr Chandler, a TIP Member, presented a paper similar in format to that of the previous speaker in that he outlined the origins and development of the NASA System and included information on possible future developments. Services being offered at present which were described in detail included the abstract journals *International Aerospace Abstracts* (IAA), *Scientific and Technical Aerospace Reports* (STAR) and the RECON on-line retrieval system. The paper was well presented and included information on many NASA projects over and above that contained in the preprint paper. These included Voyager I and II and the problems facing the Space Shuttle, due to fly in 1980. Questions put by the Reviewers to Mr Chandler included how much of the NASA data base was available to Western Europe, the difficulties in obtaining microfiche copies of non-US documents listed in STAR and the scope and coverage of conferences by IAA. Both the Reviewers and many attendees were concerned about the policy which requires one relevant document input to NASA for each hour of access to NASA on-line files via ESA/IRS. It was felt that this could be a hardship to those establishments whose output is mainly in classified/restricted literature, yet who may assist NASA in other ways, (e.g. as a Committee Member, by supplying research information direct, or participating in joint projects). In his reply Mr Chandler reminded delegates that the requirement had been built into the RECON system since its inception but only now was it being more strictly enforced. Greek delegates also questioned Mr Chandler on the availability of NASA on-line facilities within their country. As already stated in Dr Kourogenis' paper, wide use is already made of the abstract journals IAA and STAR. Mr Chandler gave some information on existing networks and suggested that this should be borne in mind when setting up the Greek National Documentation Centre.

3. SESSION II: REQUIREMENTS AND TOOLS FOR INTERNATIONAL CO-OPERATION AND DATA EXCHANGE

Chairman: J.COYNE, Dept. of Energy, USA
Reviewers: G.P.CHANDLER Jr.(USA), C.GUILLEMINET (FR)

3.1 Paper 4: State of the Art of Standardization and Harmonization of Data Elements

H.DIERICKX, UNIBID, British Library, UK

This paper related to data elements in bibliographic descriptions rather than to current cataloguing codes. It was presented in the light of the increasing trend towards computerisation and bibliographic information exchange which has tended to reduce the problem of choosing the appropriate headings for records. International and national rules and standards were dealt with, also the proposals to establish "Universal Bibliographic Control" (UBC) via the creation of authoritative bibliographic records, including the choice and form of headings by national bibliographic agencies. The majority of the existing standards and guidelines were described in some detail. In conclusion, the presenter stated that standardization of form and presentation of bibliographic data elements is an essential prerequisite to achieving international compatibility of bibliographic records. He stressed the need for close and continued co-operation amongst relevant international (professional) organizations. The paper was at a fairly advanced level and the discussion chiefly involved those NATO member countries who already possessed sophisticated documentation systems. Questions referred to standards already in use and the need to aim for an international system in view of increasing computerisation of records.

3.2 Paper 5: Towards Flexibility of Data Exchange: Recent Developments in Standards and Formats

Dr J.S.MACKENZIE OWEN, Stitchting Nijenrode, The Netherlands

This paper discussed the reasons for the wide gap existing at present between the library community and abstracting and indexing services in the field of compatibility for data exchange. The problems were considered in detail, together with the work now being undertaken by UNESCO to find a solution. Suggestions were put forward regarding possible future developments in this field. The need for flexibility when considering a standardized format was stressed in view of the diverse needs of the many different users. In all probability this paper was most appreciated by those delegates whose organizations or establishments operated large bibliographic systems where this problem was already evident. As a consequence, the majority of questions asked during the discussion were asked by delegates of this type and related chiefly to their own problems and methods of solving them. The paper was well presented and Dr Mackenzie Owen had several constructive suggestions to make in his replies given during the discussion.

3.3 Paper 6: The Role of Multilingual Reference Tools in the Production and Transfer of Technical Information

Dr G.ROSENAU, DFVLR, FRG

In the preparation of this paper Dr Rosenau drew on his own experience as a member of the team responsible for the production of the revision of the AGARD Multilingual Aeronautical Dictionary (MAD). His paper first surveyed the problems of expressing technical information in one language and then went on to discuss the additional problems encountered in producing the same information in more than one (foreign) language. He went on to evaluate the use of the different types of multilingual reference tools necessary to accomplish this and to point out the possible advantages and deficiencies of the tools themselves. As a result of his work on MAD, criteria were put forward for a future multilingual reference system.

As might be expected the questions asked during discussion of this paper came from those delegates whose establishments possessed a translations section or were entirely involved in their preparation (e.g., British Library). Much of the discussion related to the criteria advanced through Dr Rosenau's work on MAD and the presenter was able to give useful advice from his own experience in this field. His basic criteria for any future multilingual reference system were that it should be correct, comprehensive, up-to-date, easy to use and cost effective. He stressed the need for continuous updating to be regarded as an essential factor in the compilation of a dictionary, reference work or system of this type, and the fact that MAD was in machine-readable form should help to achieve this.

4. SESSION III: PROBLEMS OF UTILIZATION OF AEROSPACE LITERATURE

Chairman: A.S.T.TAN, National Aerospace Laboratory, The Netherlands
Reviewers: K.HANSEN (FRG), H.KROG (NOR)

4.1 Paper 7: Kinds of Access to Unclassified Literature

C.P.AUGER, Lucas Research Center, UK

At the request of the Programme Director, this paper was made very general in scope and related to methods now being used to identify, locate and retrieve unclassified aerospace literature, mainly in report form. Types of literature available were discussed together with the necessary reference tools, abstracts, indexes and on-line systems for its identification and retrieval. The needs of a wide variety of users were considered in conjunction with the different routes open to those seeking access, from general institutions to specialist information centres. Suggestions were put forward relating to the future improvement of present access facilities and mention was also made of some current problems and methods by which they may be overcome.

Although much of the information provided by this paper was already familiar to many of the delegates from large systems, it did stimulate much interest amongst those representing smaller establishments lacking the benefit of such tools as on-line information retrieval. Questions were asked by the Greek attendees regarding some of the services and systems mentioned by Mr Auger and the possibilities of their availability in Greece. Difficulties in information retrieval now being experienced by individual establishments were mentioned and the presenter was able to suggest various methods by which they might be overcome.

4.2 Paper 8: Difficulties in the Handling of Limited Distribution Literature

J.H.KLOPP, CEDOCAR, France

M. Klopp's paper concerned all aspects of the security of classified documents. Topics covered included military restrictions, industrial espionage and the use of the various caveats or markings to denote classification within the NATO Countries. Downgrading procedures now in use were also discussed, together with suggested safe methods of storage and transmission for classified literature.

In the absence of M. Klopp, the paper was presented at the Athens Meeting by Ing. Gen. C.Guilleminet, one of his fellow French Members of the Technical Information Panel.

The majority of the delegates had experienced some of the difficulties relating to classified documents dealt with in M. Klopp's paper. As a result, there was a far ranging discussion which included such items as storage and marking of classified microfiche, problems with restricted announcement bulletins and services, and a suggested fixed review period for downgrading. With the aid of the Session Chairman and Reviewers, Gen. Guilleminet was able to deal most effectively with the queries raised by this paper.

4.3 Paper 9: Full Text Handling – A Critical Review

Dr R.P.L.JONES, Computer Sciences Division, Atomic Energy Research Establishment, (AERE), Harwell, UK

In composing this paper the author largely drew on his own experiences in setting up and running the STATUS information retrieval system in use at AERE, Harwell. The free text approach was examined from the point of view of STATUS and certain extended features described to demonstrate its broad application possibilities. The paper began with a description of the full text method and the elements of free or full text systems were defined and described. A comparison was made between retrieval by the full text system and thesaurus-based keyword systems, including the particular strengths and weaknesses of each type. The paper went on to describe the capabilities of a full text system when used as a nucleus for integrated information systems. In conclusion, Dr Jones suggested that with a system of this type it may be possible for information scientists to design their own in-house systems, tailored for their individual needs without recourse to specialists in the fields of system analysis and software development. Mention was also made of the capture of source information and the importance of linking to word processing systems.

The discussion again tended to involve delegates from the larger establishments and institutions already completely familiar with computerised retrieval techniques. Interest centred on the STATUS system itself, particularly by US delegates and further information was provided by Dr Jones in the form of an AERE leaflet describing the system in detail.

5. SESSION IV: NON-LITERATURE DATA IN AEROSPACE AND DEVELOPMENT

Chairman: H.E.SAUTER, Defence Documentation Center, USA
Reviewers: M.DAY (USA), G.TITTLEBACH (FRG)

5.1 Paper 10: The Numeric Aerospace Data: Problems of Evaluation, Handling and Dissemination

Dr G.K.HARTMANN, Max-Planck Institüt, FRG

Dr Hartmann's paper referred to the compilation and use of data banks, that is, numerical or factual information on a subject as opposed to data bases (i.e., bibliographical references). Problems relating to the capture, storage and retrieval of numerical data for aerospace were discussed together with a brief mention of the growth and use of pictorial data. With the rapid growth of information in this format, a data bank of the type envisaged was regarded by the author as being an essential requirement for the researcher in the aerospace field. It is possible for the scientist to accurately retrieve the data directly, without recourse to documentation personnel who may be unable to correctly interpret the information format. In addition to this factor, there is obviously a marked advantage in time saved over the more common data base system.

The discussion mainly involved TIP Members and attendees from organisations closely linked with technical information problems within the aerospace industry. Many firms and establishments already operate a data bank of this type, often relating to information on a particular aircraft type or to electronic components. Questions were asked concerning future developments and the type of information that might be included. The possibilities of scientists and engineers having direct access to numerical data by this method, either with an "in-house" system or on-line to an external computer were considered.

5.2 Paper 11: The Role of World Data Centers and the Lunar and Planetary Institute in the International Exchange of Lunar and Planetary Data

Ms F.B.WARANIUS, Lunar and Planetary Institute, USA

The NASA lunar and planetary investigations have resulted in the accumulation of a mass of data in a wide variety of formats. Ms Waranius explained in her paper how the Lunar and Planetary Institute makes its information available to scientists on a world-wide basis. She went on to discuss the methods by which the Institute satisfied the information requirements of scientists from a wide variety of disciplines. Much of this data is archived by the National Space Science Data Center (NSSDC). The paper also included information on the concept of World Data Centers as envisaged during the organization of the International Geophysical Year and their subsequent growth and development.

With the growing interest in World Data Centers this interesting and informative paper prompted many questions during the subsequent discussions. Many of these referred to information on the Centers in general, and here the author was particularly helpful in providing a considerable amount of information on other establishments of this type in addition to her own. She also provided a series of references and addresses to enable questioners to conduct their own investigations into this current topic.

6. WORKSHOP SESSION

This was held immediately following Session IV and was chaired by Mr H.Sauter (Defense Documentation Center, USA). A panel made up from TIP Members, the TIP Executive, and the conference speakers answered questions from Meeting attendees. A form had previously been circulated to enable attendees to submit questions in advance for consideration at the Workshop which was primarily for the benefit of the Greek delegates. As was intended, most of the questions raised came from Greek attendees and mainly requested additional information on topics raised during the papers presented in the four sessions. Items discussed included:

(a) The administrative structure of information services in Greece.

(b) Present and future availability of the ESA/IRS and the EURONET/DIANE services and the formalities required to obtain membership.

(c) Functions and working methods of Defense Documentation Center (now DTIC), USA and its relations with information analysis/data centres. The possible takeover as a part of DTIC of centres located in the USA.

(d) Availability of British Library Lending Division services to Greece. (Contact via the British Council was recommended).

(e) Language barriers, especially when using on-line systems.

(f) The need for a national bibliographical collection to back up information services. Criteria for a collection of this type and what services might be expected from it.

(g) Problems of access to NASA data base already referred to in discussion following Mr Chandler's paper.

(h) Comments on Professor Cleverdon's (Cranfield, UK) evaluation of a test data-base; according to W.A.Martin (of ESA) results showed that natural language indexing was as good as a controlled vocabulary for retrieval purposes.

It is expected that an edited version of the Workshop Session will appear as part of the complete proceedings of the Meeting (AGARD CP-279).

7. FACILITIES FOR LOCAL VISITS

The facilities provided by the Greek authorities were very good indeed. The Conference Hall of the Hellenic War Museum was comfortable, with good acoustics, and was provided with simultaneous translation facilities and an excellent audio system for questions from the floor.

All Meeting attendees expressed their appreciation of the generous social programme which included a Panel Dinner, a reception for the Conference delegates and two local tours of Athens and district. A visit was also organised to the Hellenic National Meteorological Service located at the Aeronautical Centre at Athens Western Airport. Here delegates were shown information being received via the satellites METEOSAT and NOAA V. Landline connections between Athens and Bracknell and Athens and Washington were also demonstrated. The author noted that equipment necessary for on-line information retrieval appeared to be available at this Centre, and this could be taken into account when making future plans to participate in ESA services.

At the Technical Information Panel Business Meeting, held after the Specialists' Meeting, the Chairman, Axel Tan, on behalf of the Panel Members and all the Meeting attendees, expressed his sincere thanks for all the excellent local arrangements made by General Aktidas and by the Meeting Coordinator, Major Goulios. These arrangements had been a major contribution towards the success of the Meeting.

8. CONCLUSION AND RECOMMENDATIONS

Despite the late starting date fcr its arrangements, the Meeting was very successful and was enjoyed and appreciated by all attendees. The TIP Chairman, Axel Tan, asked the Greek Panel Members if they felt that it had been effective and if the Greek delegates felt that they had learned useful information from it. Major Goulios replied that they had, in fact, learned a great deal about services and systems which they could employ in their proposed national bibliographic system and in his own case, in the Hellenic Air Force information services. Asked about the level of the papers presented, he stated that although the Greek nationals may well have found it rather difficult to understand certain presentations, they nevertheless learned about a variety of services now provided in other NATO Countries and which they felt that they could gainfully use themselves. In view of this, there would seem to be a clear case for retaining in the programme some papers at a general level when the Meeting is hosted by a member country at present lacking some of the facilities available to its larger NATO partners.

The present method whereby the presenter emphasises his most important points during the first 20 minutes of the allotted 45 minutes for each paper is very satisfactory and should be retained. This allows for more participation by the Reviewers and Meeting attendees in general. There are still those who feel it would be a good idea for attendees to receive preprints up to a fortnight before the Meeting itself, in order that these can be read in advance and more significant questions asked in discussion as a result. It is felt that this would be difficult to achieve in practice owing to the tight time schedule for printing and preparation. The Panel Executive does however always ensure that the Reviewers and Panel Chairman receive copies of papers to be presented in their particular Sessions well in advance of the Meeting.

The standard of the Vu-graphs included as a part of the Sessional Papers tended to vary widely in the course of this particular Meeting. Some examples gave obvious signs of hasty preparation and it is thought that for future Meetings the Programme Director should endeavour to vet these items in advance, in order to ensure a high standard of quality. This, of course, means that the Vu-graphs should be submitted for vetting together with the draft copy of a paper.

It is to be regretted that, in view of its success, the Meeting was rather poorly attended. The major factor here I believe was the present financial climate, as I am sure there would have been many more attendees if financial support was forthcoming. The question of finance now represents a considerable problem when attempting to recruit possible speakers for the TIP Specialists' Meeting and many organisations from both the government and private sectors will not allow their employees to accept invitations to deliver papers unless their travel and subsistence expenses are provided.

At the Panel Business Meeting held on 19 October it was suggested and agreed that the Greek authorities might be advised to apply to AGARD for a consultancy to be arranged to assist in their endeavours to set up a National Documentation Centre. It was also suggested that visits by Greek representatives to Documentation Centres in some of the large NATO Countries should be considered. Using the United Kingdom as an example, this might include the British Library plus its Lending Division, Defence Research Information Centre, Technology Reports Centre and the Library of a major research and development establishment such as the Royal Aircraft Establishment at Farnborough. The Greek TIP Members thanked the other Panel Members for their help and advice and stated that they would discuss these proposals with their senior officers.

In view of the difficulties encountered by the author as Programme Director for the Athens Meeting, it is strongly recommended that the proposals put forward by Mr H.Sauter (DTIC, USA) at the Panel Business Meeting on 19 October should be adopted and adhered to. These relate to a strict timetable for the arrangements which should begin three years in advance.

In conclusion, tribute must be paid to the Panel Executive, Trevor Sharp, and to his assistant, Heather Laget (TIP Secretary), for all the extremely hard work done by them towards arranging and running the Meeting. Their helpful assistance did much to make the delegates feel at home and was a marked factor in making the Meeting a very enjoyable one.

ACCESS TO AEROSPACE INFORMATION – THE GREEK SITUATION

by

Dr K.N.Kourogenis
Director of Technology and Applications
Y E E T, Greece

INTRODUCTION

The dynamic developments in the field of Aerospace sciences and their associated technologies during the last two decades, have influenced and are nowadays reshaping many areas of human activities. This evolution, did not only give answers to age old questions about space, but it also gave rise to the development of new applications aiming at a better living for mankind.

The successes of national and international aerospace programs such as those of NASA and the European Space Agency (ESA), did not only satisfy human curiosity and the purely scientific interests of a limited number of gifted specialists, they also became gradually subjects of interest for the large majority of those who were not directly involved by profession or nationality. Non-aerospace scientists and engineers became users of aerospace derived technologies and in some cases they had not only to be consistently informed but even be reeducated in new technologies. Thus, new groups of aerospace information users with differentiated needs were formed.

The quality and quantity of aerospace information transferred from its generator to the professional user may be distinguished to what the average man can absorb. However, since the properly formulated and without noise distortion transferred information to the end user constitutes knowledge, there is a need for high quality aerospace information towards all directions irrespectively of the transfer medium and who the end user may be. Aerospace information sources are dynamically interrelated with their corresponding users in a time dependent relation due to the ever increasing numbers and subject areas of publications and other related forms of such information. The exponential growth of scientific and technical literature is also observed in aerospace literature, whereas the distinction between these two becomes progressively more difficult. The so called "information explosion" phenomenon is extended in this field too, irrespectively of the fate, success or failure, of the large scale national and international aerospace projects and programs.

THE SOURCES

The sources of primary aerospace information are mainly countries or organizations which are producers of aerospace technology. In other countries where no such local technology exists, the need is covered by creating secondary aerospace information sources within the existing general science and technology information centers, or the more sophisticated, computerized national documentation centers. The existing primary aerospace information sources provide documents of various types and formats. Typical examples of this variation are NASA's semimonthly Scientific and Technical Aerospace Reports (STAR), the International Aerospace Abstracts (IAA) published by the American Institute for Aeronautics and Astronautics, NASA's Thesaurus on Aerospace terms, ESA's Scientific-Technical and Contractor reports, the International Astronautical Federation's publications, and the more specialized and timely information provided by the weekly Abstracts of the National Technical Information Service (NTIS) on NASA's Earth Resources Survey Program. In Europe, ESA inherited the Space Documentation Service (SDS) from the European Research Organization (ESRO) and the European Launcher Development Organization (ELDO). The basic operational concept of SDS is the exchange of all unclassified information on Aerospace science and technology between NASA and ESA. Thus a continuous flow of data and systematic information is secured between the two organizations and through them among their member countries. The subject areas of this formal dissemination of information do not only cover aeronautics and spaceflight but they also include earth sciences, biosciences, computers, instruments, etc.

SDS also provides various information services to users such as selective dissemination of information (SDI), standard profiles (SP), reproduction services and answers to specific queries from users. The roles of the European Space Operations Center (ESOC) in Darmstadt and the European Space Technology Center (ESTEC) at Noordwijk together with the headquarters operations in Paris should not be underrated in the successful integration and operation of the European space information network. By participating in this network, countries with no substantial primary aerospace information output, can benefit and share the information it can offer. However, network compatibility reasons advocate that all nodal points should conform to strict rules of script and format, so that Latin script languages may be acceptable by the system but no others.

Countries with non-Latin script languages, in order to participate successfully in the network, have among other obligations to create the infrastructure of translation and abstracting services in addition to the installation of all necessary hardware and software associated with the creation of a compatible nodal point. These extra technical and economic needs may be partially covered by incorporating the functions of an international nodal point to the existing information services of the national documentation centers (NDC). But the NDC's are not that flexible and apply their own priorities and unified methods in treating information no matter if this refers to aerospace technology or basic sciences. Hence the need arises for the parallel creation of a semiautonomous aerospace information center, connected to the international network through the NDC and utilizing its facilities. This center may eventually operate as a focal point at national and international levels and be responsible for the quality, quantity and timeliness of the relevant information it will provide to its national users. Aerospace journals and conference proceedings found in libraries may be considered as sources of such information and knowledge for at least one group of users, the scientists and researchers throughout the world, but they may not satisfy completely their information needs.

THE USERS

The users of systematic aerospace information are not only aeronautics specialists but they may also be telecommunication engineers, Geologists or Meteorologists, to name only a few. Other users may be scholars from all scientific fields, who are usually interested in the new findings of the aerospace programs. Even the general public is a customer of general aerospace information with much of its needs covered by the popularized editions and films of the mass media. The definition of user needs in the aerospace information field, should be the basic prerequisite for the design of a user oriented information system, and these systems seem to be successful nowadays. However, defining the user needs is not that easy to obtain, because the user himself may change attitudes and needs over the time, he may be subjective in defining his current information needs, he may propose expensive ways and means of getting it, and last but not least he may not be properly educated on information processing and the capabilities of the system he is using. The answer to above question may be founded on an analytic and objective evaluation of all factors and parameters which formulate it. Seeking the fruitful cooperation between the user on one hand and the aerospace information source and the information system designer on the other may prove valuable in most cases.

Another aspect of the operational implications between the user and the source is that their interrelation is an unstable equilibrium, the user may equally be unhappy if he receives less or even more information compared to what he actually needs. This is ascertained by the fact that large volumes of unwanted and time consuming data and documents reach the working desks of scientists and engineers every day. One way out of this saturation state is the further development of the selective dissemination of information (SDI) methods addressed to individuals or narrower groups of users, if the cost allows it.

THE GREEK NATIONAL INFORMATION PROGRAMME

The creation and operation of easily accessible aerospace information sources in Greece may not be considered separately from the overall national information program (NIP). Two years ago the decision was taken to establish the National Documentation Center (NDC). The main objectives of the Center are the creation and coordination of the national network of scientific and technical data and information in accordance with international standards and the dissemination of processed information to those interested with a view to the economic development of the country.

The NDC will eventually have the legal and executive power to incorporate or enlarge other existing information centers and units and create new ones. This is an effort to establish a new national service of documentation with as wide a range of subjects as possible by incorporating the existing and new services in a country wide network. The NDC will also be the focus for exchange of information between Greece and the EEC and within EURONET, and in world-wide information services such as AGRIS. In Greece, the NDC will at least play the role of the network coordinator no matter if the national network develops to become of the star-like form or any other form. In order to obtain a clear picture of the size of the national network in its final form, systematic surveys have been carried out for the location and evaluation of the most important documentation centers and libraries in the country. This includes as complete data as possible on the holdings, services, personnel and characteristics of these institutions. The most prominent of these centers per subject area will constitute the sectoral focal points of the subsystem of the network in their specialized fields, e.g., agriculture, medecine, aerospace, engineering, sciences, etc.

Parallel to the above a directory of users of information and profiles of subject interest is currently under way. Scientists and engineers are interviewed in order to establish all the characteristics of the potential users of the information system in preparation for SDI and for the secure design of the future development of the system. The directory of on-going research activities is published intermittently every second or third year by the Scientific Research and Technology Agency, (YEET), this activity will be undertaken by the NDC in the future. Other activities to be performed by the NDC are the acquisitions list of the NDC so that the cooperating libraries in the network will be informed and develop a co-ordinated acquisitions policy in the country.

The union list of periodicals-holdings of important libraries will be published and arrangements will be made to have it updated regularly. Bibliographies, Reprographic services and translation and terminology services are among the goals of the Center. In addition to the above Greek NDC will also set standards and methodologies, provide financial support and technical assistance to co-operating institutions, operate and evaluate pilot projects established for testing new technologies, collect and provide data to government agencies for formulation of the national information policies, act as a national referral center and organize training programmes at the national level. The final stages of the NDC establishment period will be the creation of the computerized information network, the participation in regional information systems and finally the participation in EURONET with a two-way exchange of information during 1981. By that time the infrastructural prerequisites for such networking operation will have been fulfilled in cooperation with the Greek Telecommunications Organization.

THE GREEK NATIONAL PROGRAMME ON RESEARCH AND TECHNOLOGY (EPET)

The effort to develop the NDC is part of a more extensive programme on R and T. In this context it is worth mentioning the objectives of this programme which are the development of scientific R and T, and the application of research and technology in meeting national needs. EPET introduces the concept of project financing on specific objectives as opposed to the old method of institutional funding. In the past, Greek research needs were usually met with the creation of a corresponding institution to undertake it. This institution would soon become a research center, usually with general research to member researchers rather than to the original research need.

This disorientation was further intensified by the lack of any coordination. The only incentive for research became the speedy production of publications, leading to an academic career. These negative aspects are hopefully met by the introduction of the National Research and Technology Programme. EPET is considered the pivot for the development of scientific research in the country, and has the following objectives:

- To fulfil the national needs for natural resources (materials, energy, food, water).
- To improve the country's international economic and technical competitiveness.
- To increase the standard of living and working conditions.
- To protect the natural environment and to protect the national interests generally.

The Scientific Research and Technology Agency (YEET) which is responsible to carry out this whole new programme, has established the procedures for the formulation of National Scientific Policy which includes: the identification of national needs and objectives for scientific research. In the context of this effort, numerous scientists, academics and people from industry and finance participated and assisted in the specification, classification and evaluation of national needs and objectives for scientific research. Then, followed the classification of all proposals into two major groups, namely the infrastructure and the areas of intensive research. To conclude on this matter one may state that EPET is now operational and one of its first infrastructural priorities is the development of integrated information systems in Science and Technology. This is considered to be a basic prerequisite for the development of coordinated scientific research and technology by avoiding costly duplication and delays.

GREEK AEROSPACE INFORMATION SOURCES AND USERS

The efforts of YEET to establish the National Research Programme and the NDC will eventually cover the special needs of aerospace information users in the Country. It is anticipated that the whole picture will change drastically after the international network interconnection of Greece and other European countries. For the time being, many users of Aerospace information exist in Greece: The Academy of Athens created the Greek National Committee on Space Research (GNCSR) in order to achieve a successful cooperation and interconnection between Greece and the Committee on Space Research (COSPAR).

Another coordinating body worth mentioning is the Greek National Committee for Astronomy (GNCA), which connects Greece with the International Astronomical Union (IAU), exchanges information, organizes seminars and other international activities.

The GNCA coordinates the main astronomy, astronautic and space research centers. One is the Research Center for Astronomy and Applied Mathematics (RCAAM), which receives most of NASA's publications, it operates one of the largest special libraries in the Country on astronomy, astrophysics and space sciences. Some 400 institutions throughout the world are its partners on an exchange basis, and 50 subject area journals. This library is accessible by scientists and engineers. The other institutions are namely the Astronomical and Meteorological institutes of the National Observatory of Athens, the departments of Astronomy and Astrophysics of the University of Athens, the departments of Astronomy and Topography of the National Technical University, the center for observation of artificial satellites, the department of Astronomy and Geodetic Astronomy of the University of Thessaloniki, the Astronomic department of the Universities of Patras and Ioannina. Nearly all above centers may be considered as sources of aerospace information being richer in pure Astronomical subjects and poorer in other fields.

Services such as the Greek Telecommunications Organization and the Greek Meteorological Service participate in the international exchange of data and information related to Aerospace applications and services. A number of journal libraries such as that of the National Research Center make available aerospace information to their visitors.

To conclude with a futuristic statement we may say that as the evolution of Aerospace sciences leads to more applications useful to mankind, the information systems will produce high quality, easily digested and timely information, easily accessible by the increasing number of users.

REFERENCES

1. Stolk, H.A. — *Sources of Scientific and Technical Information*, AGARD-LS-44, 1971.
2. Stearns, J.F. — *User Needs*, AGARD-LS-44, 1971.
3. Massil, S.W. — *Greece, Development Information Infrastructures. Establishment of a National Documentation Center*, S. No FMR/BEP/PGI/77/307.
4. Kourogenis, K.N. — Draft Proposal on the *Establishment of the National Documentation Center*, YEET, November 1977.
5. UNESCO — *"Greece" National Science and Technology Policies in Europe and North America*, No.43/1978.

European Cooperation in the Field of Aerospace Information

W.A. Martin
European Space Agency
Information Retrieval Service
ESRIN, Via Galileo Galilei
00044 Frascati, Rome, Italy

SUMMARY

Agreements reached in the early sixties leading to the establishment of the ESRO/ELDO Space Documentation Service are reviewed. European cooperation in both the input of aerospace information and its utilisation are described. Details of ESA-IRS activities in support of the NASA information system are given. The development of the use of aerospace information in Europe as evidenced by the growth of IRS networking is examined. Research into aerospace information handling is outlined. The relationship between ESA and the CEC is explained in the context of EURONET.

1. LAYING THE BASIS FOR EUROPEAN COOPERATION : THE ROLE OF ESRO-SDS

1.1 In May 1964, an agreement for cooperation and exchange in the field of scientific and technical information was reached between NASA and the European Space Research Organisation. Early the following year, in April 1965, an agreement to set up a joint documentation service entered into force between ESRO and its fledgling sister organisation ELDO - the European Organisation for the Development and Construction of Space Vehicle Launchers. The result was the creation of SDS - the Space Documentation Service.

1.2 Under the terms of the agreement reached between ESRO and ELDO the joint documentation service was to cover both space research and space technology, and was to serve not only the staff of the two European space organisations but also their member states. The recipients in the member states were originally defined as "government establishments, universities, and institutes working in international or national space programmes", and "qualified persons or companies residing in the member states of ESRO and ELDO". The possibility of providing services to organisations in non-member states was also foreseen in a paragraph stipulating that such requests must be referred to the ESRO Council.

1.3 A further agreement was reached late in 1966 between ESRO/ELDO and EUROSPACE - the European Industrial Group for Aerospace Studies. This agreement enabled Eurospace member companies to obtain services from SDS, ie. bibliographic searches and/or selective dissemination notifications (SDI). It contained a financial guarantee to SDS for a minimum sum of 75 000 French Francs, and the requirement that "Eurospace shall endeavour to prevail upon its members to increase the number of their subscriptions and to have them renewed, and shall encourage the recipients to take advantage of the (Space Documentation) service beyond the facilities covered by their subscriptions". This preoccupation with financial return was entirely understandable when it is recalled that not only SDS but also its parent organisations were in their very early stages of development. Thus from the beginning SDS has been required to provide services to all sectors of the aerospace community within the member states of ESRO and ELDO.

1.4 A significant aspect of the 1965 ESRO/ELDO agreement concerned the types and sources of material to be provided by ESRO and ELDO to the joint documentation service, and an acknowledgement that these could be made available to NASA. "ESRO and ELDO shall provide scientific and technical documentation, of which the publication is not prohibited, received or produced by them during the performance of the tasks described in their conventions or resulting from these tasks. They will in particular ensure that in research, study, development or other relevant contracts ...... the application of this provision is duly observed.

ESRO and ELDO agree that all documentation which is made available under the terms of this agreement may be channelled to ...... NASA ......; documentation received from NASA ...... shall be brought into the joint documentation service."

The original NASA/ESRO agreement of 1964 included the requirements that "ESRO will provide to NASA abstracts of scientific and technical reports originating from European sources ...... processed in a form suitable for inclusion in NASA's Scientific and Technical Aerospace Reports (STAR)", and "NASA and ESRO will make available to each other single copies of microforms ...... representing ...... abstracts published in STAR." In addition, it was stated that "NASA and ESRO have agreed, in principle, to exchange material for computer searches ......"

1.5 In looking back through those early documents dating back some 15 years, I was impressed by the very thorough job which seemed to have been done of setting the stage for a good two-way flow of aerospace information between SDS for Europe and NASA for the United States, and for its proper dissemination throughout the member states of ESRO and ELDO. It is clear that, a decade and a half ago, the manner in which Europe proposed to cooperate in the field of aerospace information had been defined in considerable detail. A tangible demonstration of its resolve to realise this cooperation lay in the creation of SDS, the Space Documentation Service, the role and activities of which had largely been specified as a result of agreements now dating back some fifteen years. The subsequent development of the Service lay in the manner and extent to which it was able to realise the mandate defined in principle in the mid-sixties. I will leave you to decide later the extent to which these early intentions have subsequently been achieved.

1.6 In 1973 the decision was taken to terminate the majority of the activities carried out by ELDO, and to re-orient those carried out by ESRO to include applications satellite projects in addition to the purely scientific missions previously flown as well as the development of a European launcher, ARIANE, and the re-usable Spacelab to be carried by the shuttle developed by NASA. With the fusion of the two original European space organisations a new agency named ESA, the European Space Agency, was created in 1975.

1.7 Later still, in 1978, ESRO-SDS, which had already become known as ESA-SDS, was given a new name. The intention was to reflect the wide range of scientific and technical information and services which were, by then, available from this Department of ESA responsible for STI and other information and data facilities. The new name chosen was IRS - Information Retrieval Service.

2. EUROPEAN COOPERATION TODAY, AS SEEN BY ESA-IRS

Today, IRS activities might be summarised as follows :

2.1 For NASA input, about 4 000 documents originating from European aerospace sources are handled annually of which almost 70 per cent are fully pre-processed to computer readable weekly tapes for STAR announcement. About 100 of these documents are selected for complete translation into English.

2.2 An online interactive search system is operated for ten hours daily on all working days. Some 15 major scientific and technical bibliographic databases plus two specially developed databanks total over 12 million records online at all times. Two databases each now exceed three million bibliographic references (Chemical Abstracts and PASCAL - Bulletin Signaletique), and are believed to be the largest online databases of their type in the world. Each database is updated monthly. Searching may be via a controlled vocabulary, if available, and/or via the natural language of title, title extension, abstract, uncontrolled terms or any other desired data element. A special numeric range search command has been developed for use with the databanks.

2.3 In parallel, an extensive STI network has been built up and provides direct online access to the IRS computer from all ESA member states and from several non-member states. Well over 10 000 kilometers of leased telephone lines stretch from Stockholm in the north, to Rabat, Morocco in the south. This network supports high-speed 2 400 bps (240 char/sec) video terminals equipped with 190 line/minute printers multi-dropped directly, and lower-speed 300 to 1 200 bps dial-up terminals via remote concentrators, provides remote offline printing via medium-speed printers and is used for various purposes by the Agency including facsimile transmission and computer to computer links. Other networks are interconnected to provide more access points and backup redundancy.

2.4 Today, more than 1 000 users of this system, with constant additions throughout currently thirteen European countries, one North American and one North African state, demand continuous support as facilities are improved and new databases added.

Recently an online order facility has been implemented which enables a user to order any original document announced in the NTIS and PASCAL files by means of a quick and simple operation at the terminal. There has been demand from both users and database suppliers alike to extend this feature to other files; as this paper is being prepared online ordering of NASA original documents is being implemented. For those users who require absolute security to surround their online file access a "TOP SECRET" command has been introduced which renders it quite impossible for anyone at the IRS computer centre to know what search the user is executing; no record is accessible other than by the user and this is destroyed instantly at logoff or if a break of any kind occurs. The more sophisticated terminals supported can be switched from retrieval mode to input mode (ODE) and can then be used to create machine readable data in a most convenient and efficient manner.

2.5 Some comments on the file coverage might be of interest. First, the needs of the Agency must be met as far as possible, ie. Agency staff, its advisory groups, its prime contractors, and its scientific collaborators. Because no multi-disciplinary mission oriented service can cover any single category within its scope to the same depth as a discipline-oriented service covering that category exclusively it has proved necessary to augment the NASA aerospace file with a number of discipline-oriented databases, particularly in the fields of physics, chemistry, electronics and engineering. More recently it was decided to add a biological database to cover the life science requirements of the SPACELAB programme; and an agricultural database, plus three small files covering oceanography, environmental and pollution information in anticipation of the information needs resulting from the new earth resources activities. For the latter a special databank of earth resources imagery, available from ESA as either photographic products or computer compatible tapes, has been developed and implemented.

The information needs of users in the member states tend inevitably to be much less well-defined than those of the Agency itself. The most marked demand was for chemical information which confirmed the need to add Chemical Abstracts Condensates. This file has become the largest IRS file and has the highest utilisation factor, over 22 per cent of total connect hours. In fact, chemistry, physics, engineering and general research together account for approximately two thirds of the total access time.

2.6 More recently there has been a growing demand from many member states for IRS to offer a more complete range of the information required. More specifically there has been criticism of the IRS policy of restricting the service to scientific and technical information and thus failing to satisfy the need for business and economic information. This leads to the question "just how can aerospace information be defined?", a question which has been asked many times but remains unanswered. One recent suggestion, for a very pragmatic approach, has been to carry out an analysis of which files are actually used significantly by the aerospace community. If this were done and the result was used as a basis for the definition of "aerospace information" it is clear that business files (which would enable searchers to obtain answers to questions of the type: "what large contracts have been obtained by company X recently?", "which company got contract Y in Saudi Arabia?", or "what are the turnover figures for the last 5 years for bidder Z?") would certainly be included.

2.7 As an indication of Europe's cooperative input to the world-wide aerospace information system operated by NASA, the IRS pre-processing statistics are incomplete due to the fact that the handover to IRS of all member states' document acquisition had not been completed during 1978. As a result, the numbers of documents sent from certain member states, notably the United Kingdom, did not represent the total contribution since some were sent direct to NASA.

As an indication of Europe's cooperative utilisation of aerospace information the IRS statistics of online access are also incomplete since they cannot take account of the searches performed manually by means of NASA STAR (Scientific and Technical Aerospace Reports) and IAA (International Aerospace Abstracts) which may still comprise a significant proportion of the total searches made.

3. EUROPEAN SUPPORT OF NASA SCIENTIFIC AND TECHNICAL INFORMATION ACTIVITIES : A CLOSER LOOK AT INPUT

3.1 IRS acts as a decentralised European extension to the NASA Scientific and Technical Information Facility. It maintains contact with all bilateral NASA partners in Europe and more recently the "tripartites", and arranges information exchange agreements between ESA and European groups working in the aerospace field. All this is intended to ensure the acquisition of as much as possible of the aerospace literature originating in Europe. An attempt is also made to cover conference schedules and programmes. All documents received by IRS are registered, and credited to the forwarding organisation. An Acquisitions Database is updated on a weekly basis and from this are generated half-yearly and yearly detailed and summary reports which are sent to NASA, to all ESA delegations, and are used by IRS itself.

3.2 After registry, all items are submitted to a selection procedure which can result in any of the following treatments :

3.2.1 IRS pre-processing to computer readable tape plus microfiche for announcement in NASA STAR

3.2.2 Despatch to AIAA for announcement in International Aerospace Abstracts

3.2.3 Pre-processing for initial STAR announcement; cover-to-cover translation; pre-processing for second STAR announcement of English version

3.2.4 Discarded as it duplicates an item already in the NASA Information System

3.2.5 Rejected or discarded as out-of-scope or otherwise unsuitable.

3.3 When an item is selected for STAR announcement it is given an English language abstract (the source document may be in any of eight European languages including English); descriptors from the NASA Thesaurus are assigned; it is catalogued; and a mother microfiche of the source document is prepared. Simultaneously, the details coded into the input worksheet (the form 901) are keyboarded to computer readable ATS compatible form using an online input facility. Weekly, the cumulation of data is rolled out onto a mini-tape and sent to the Facility with a back-up copy. Also from this ATS data is stripped the input to the IRS Acquisitions Database.

3.4 At the request of NASA certain series of reports issued by ONERA (Office National d'Etudes et de Recherches Aerospatiales), France and DFVLR (Deutsche Forschungs- und Versuchanstalt für Luft und Raumfahrt e.V.), Germany, are considered automatically to be candidates for full translation by IRS unless rejected for some reason, usually out-of-scope. More recently ONERA has offered to provide the periodical La Recherche Aerospatiale, issued six times a year, in the form of a cover-to-cover translation into English. Following translation the English language version of the report is issued in the ESA TT-series (Technical Translation), pre-processed for STAR, and cross-referenced to the original announcement of the source document in the original language.

3.5 A special report acquisition and STAR pre-processing project was initiated at the end of 1974 as a result of the close-down of ELDO. Due to differences in the Conventions of ESRO and ELDO the latter had not during its lifetime been able to make available a comprehensive collection of technical reports. As part of the liquidation programme almost 15 000 contractor reports and other documents generated during the ten year life of ELDO were examined, of which more than 1 600 documents, containing almost 2 000 discrete papers, were selected for full pre-processing for NASA STAR announcement. These items were all announced in L-STAR but, perhaps more important, now form part of the NASA computer readable database.

3.6 In addition to the largely passive role of acquisition of the European literature, IRS performs an active monitoring and liaison function with NASA's European bilateral partners and, more recently with the "tripartites". The latter organisations are those which wish to have direct online access to the NASA File via the Agency's STI network and online service. When an IRS user requests access to the NASA File he is asked to sign a tripartite arrangement between NASA, ESA and the user organisation. Under the terms of this tripartite arrangement NASA agrees to grant online access to the NASA File in return for the submission of literature within the scope of the NASA Information System. Items published openly, eg. in the periodical literature, are not accepted since these would be processed by AIAA independently of such an agreement. Difficult-to-obtain material such as research reports are sought. From NASA's viewpoint there is little difference between a bilateral and a tripartite. In each case an exchange is involved although in the case of tripartites no material dissemination is required.

3.7 The monitoring function is carried out on the basis of a number of system aids. First, the overall performance of each source organisation is checked annually on a "this year, last year" basis. Second, a regular report is generated from the Acquisitions Database which lists all sources which have sent zero input, one document, two and so on, over the prior period. The sources are listed alphabetically by country and where appropriate, contact is made with the source to draw attention to the absence of, or drop in, input and to establish and, if possible, correct the cause.

In addition to these guides, a special report on all tripartites is produced each half year. This report lists the number and type of documents submitted to IRS by each tripartite organisation and the number of connect-hours in the NASA File.

4. EUROPEAN UTILISATION OF AEROSPACE INFORMATION : IRS'S EXPERIENCE WITH NETWORKING

4.1 Towards the end of 1968, the 400 000 NASA File records needed 20 hours of computer residence time to process a batch of 30 questions; the operation would clearly soar out of hand in the following year. Following an evaluation of available systems and discussions between NASA STIO staff and SDS the decision was taken to implement RECON so as to give an inverted file system to shorten search times and increase capacity, and to have a system compatible with that used by NASA.

4.2 RECON was implemented on ESRO's largest IBM 360 computer installation which was located at ESOC (European Space Operations Centre), Darmstadt, near Frankfurt. In addition to the advantages already mentioned, RECON offered the bonus of being an online system and the original ESRO STI network consisted simply of two leased international telephone lines, one from Darmstadt to Paris where a RECON terminal was installed at ESRO Head Office, and from Darmstadt to Noordwijk near Amsterdam, to support a terminal located in the main library of ESTEC (European Space Technology Centre).

4.3 During that same year great interest was expressed by several member states in introducing direct online access to the SDS database in their national aerospace information centres and, with the approval of NASA, the first such "external" terminal was installed in 1970 at the Technology Reports Centre, near London. The embryo STI network was beginning to grow. The first installation in the UK was quickly followed by one at ZLDI (Zentralstalle für Luft- und Raumfahrt Dokumentation und Information), Munich, and by the end of 1975 some 20 high-speed leased line video terminals had been installed in the member states.

4.4 In 1973 SDS, which had been split between ESRO HO, in Paris, and ESOC, Darmstadt, was transferred to new premises at ESRIN, Frascati, near Rome, which had previously housed fundamental space science research activities. This brought all SDS staff together under one roof and provided the Service with a dedicated IBM 360/50 computer in support of STI activities for the first time.

4.5 To minimise the problems of relocation, it was decided not to disturb the network, then essentially a STAR network centered on Darmstadt and operating at 2400 bps. The intention was to retain the existing network topography and to link ESOC and ESRIN by means of two high-speed channels of 9600 bps using multiplexing modems on geographically independent routings. When the two trunks were handed over by the PTT's however, it did not take long to discover that the line routings were identical; since when one line went down, very often both went out together! This was disastrous for network operation, of course, since any outage on the ESRIN link killed the entire network. The alternative routing was never achieved since apparently it never proved possible to reach the line quality required for data transmission via the only other possible route over the Alps!

4.6 The only solution seemed to be a total revision of the network to bring the "centre" to Rome. There were other factors giving a push to this project such as an increasing user demand for the support of more sophisticated hardware, eg. dial-up concentrators and remote line printers. Following this reconfiguration total network blackouts are now rare.

4.7 In the original concept the network was required to support several tens of 2400 bps terminals multi-dropped from a STAR of lines operated at 2400 bps. As far back as 1973, however, SDS had foreseen the need to introduce much cheaper access to the network. The high-speed leased line terminal was a sophisticated and responsive device but, certainly in European information terms, it was expensive, and could really only be justified in very busy information departments or services. The development of the RTC, or Remote Terminal Concentrator, was disappointingly slow - the first version was not available for testing until 1975. The RTC is based on a PDP-11/10 minicomputer and as originally tested acted as a speed switch from the 2400 bps transmission from the central computer to the PDP-11 on one side, and from four to ten 300 bps dial-up devices on the other. Once again the Technology Reports Centre, near London, was the first member state organisation to install the new device, the penalty for which was a six-month teething period during which would-be dial-up users in the UK were called upon to exercise some patience! The RTC, which behaves in a manner very similar to a 2400 bps terminal, may be multi-dropped from the same 2400 bps line as a group of LLTs in more or less any combination. RTCs have been installed in Copenhagen, Dublin, London, Madrid and Stockholm, whilst concentrators using multiplexers are installed in Brussels, Darmstadt, Noordwijk and Paris. Direct dial-up to Frascati is also possible.

4.8 The introduction of concentrators into ESANET has enabled a new level of service to be provided to users who previously had to request a search to be executed for them by intermediaries in national centres. Though the latter were using an online system, to the user it looked like "batch". Now, through the RTC, he is able to search directly, online, himself. This has led to a change in the role of the intermediary, from that of searcher to that of teacher. It has also led to new demands for yet higher levels of service - RTC centres wanted the printout via their own line printers, driven over the network outside, and even during RECON time, thus making extra demands on the network.

4.9 Dial-up users in France are particularly well-served as a result of the decision to develop a gateway linking ESANET with CYCLADES. The latter is an experimental, but advanced, packet-switched network with nodes in Paris, Rennes, Grenoble, Toulouse, and Lyon. This gateway, developed by the French, permits one way access to the IRS computer. A new network named TRANSPAC is expected to replace CYCLADES in France this year and plans are in hand to connect IRS with TRANSPAC.

4.10 Since 1975, IRS has also been connected to the TYMSHARE network which has nodes located in Europe at Paris, Brussels, The Hague, Zürich and Frankfurt. Particularly useful to IRS users located in the Netherlands and Switzerland this network provides additional capacity to cover peak traffic times, and backup redundancy in the case of network failures.

4.11 More recently IRS has been connected to the Rome node of DARDO, the Italian link with the North American networking services including TYMNET and TELENET in the United States and TELEGLOBE in Canada.

4.12 As this paper is in draft, preparations are well advanced for the connection of IRS to EURONET; this is covered in the next section.

4.13 Overall utilisation of the IRS online service via the network complex described amounted to 22 938 hours in 1978.

5. EUROPEAN RESEARCH INTO AEROSPACE INFORMATION HANDLING : THE IRS INTEGRAL DATABASE CONCEPT

5.1 We have carried out a number of experimental studies using the NASA File in order to try to measure some of the RECON system parameters, particularly in 1972/73 (1) and more recently when we reported some unexpected findings to the 1977 Cranfield Conference (2). Test searches were carried out by experienced users in several member states, also by IRS and other ESA staff. In the most recent study the analysis of the results was carried out by Cleverdon. Broadly speaking the results of the measurements of recall and precision compared with those found by Lancaster and others in studying searching of the National Library of Medicine's MEDLINE system. Our earlier study suggested that in day-to-day use, on average, about 50 per cent recall at 68 per cent precision could be expected for an elapsed terminal time of about half an hour per search. The recent study did not try to measure absolute recall, but only comparative recall. However, correlation of these figures with the earlier study suggests very similar results of 51 per cent recall at 74 per cent precision with elapsed terminal time at 38 minutes.

5.2 Our experience with the NASA File and with several other major databases led us to believe that the NASA File is about the best example of what can be achieved using a well-controlled vocabulary. However, for any database, typical results of 50 or 60 per cent recall should not permit complacency since, for every 5 or 6 documents found which are known to be relevant to the question, a further 4 or 5 equally relevant documents remain "mislaid" in the computer.

5.3 Analysis of controlled retrieval language search failures in the experimental studies readily demonstrates the inconsistency with which valid descriptors from the controlled vocabulary are assigned to similar documents. In other words it is one thing to achieve a well-controlled vocabulary, but quite a different matter to control the indexing based on that vocabulary. We are probably unlikely to improve matters as long as manual procedures are used for checking. In my view, a well-structured thesaurus containing more than a very few thousand terms is far too complex an instrument to be used unaided by the human indexer.

5.4 There is also the need to make it easier for a user to search across several files, preferably simultaneously. However, the studies carried out had served only to emphasise the difficulties and failed to produce any realistic proposals to solve the problem - all approaches were based on the notion of cross-linking the various thesauri. Such a concept is operationally unattractive, smacking of the sledgehammer and nut tactic. Some indication of the complexity of such a task is given in a paper by Niehoff (4) describing the development of an integrated controlled vocabulary in the relatively limited area of energy terminology alone. In my view what was needed to achieve an integral database was a common retrieval vocabulary for all the files to be merged.

5.5 The idea emerged of the possibility of using the natural language of title, title extension, abstract, free-terms, etc, as the primary retrieval vehicle to be used with several files in an integral database and regarding the retrieval keys provided by the tape supplier (eg. classification code, descriptors, specialised codes such as CAS Registry Number, etc) as auxiliary search aids specific to individual files.

Though a great deal of work had been reported on the use of natural language for information retrieval it was not possible to draw any hard and fast conclusions. Most work had been done offline and it was difficult to compare results reported due to the surprising variety of bases for the experiments.

Opinions tended to polarise strongly in one direction or the other, often it seemed, without adequate justification. Perhaps the following quotes will illustrate this situation :

- "Experiments already carried out have demonstrated that natural language, with minimal or no control, is superior to any form of controlled vocabulary." (5)

- "The reasons for the failure of free-text indexing are both theoretical and practical. Many computer experts continue to propose the computer implementation of free-text indexing either ignorant of, or indifferent to, past failure." (3)

5.6 We began to feel that if retrieval performance based on natural language searching could be developed to the point where it was not significantly inferior to that achieved when using a controlled vacabulary, then it could provide the vehicle for an integral database. An experiment was designed by IRS, a test database was created online based on the NASA STAR tapes for 1973 and 1974 with the agreement of NASA, and an evaluation was carried out under the direction of Cyril Cleverdon (6). Experienced searchers from several European countries carried out test searches using either controlled vocabulary or a special natural language index. This experiment was reported and the results discussed at the 1977 Cranfield Conference (2).

5.7 The results surprised us and could be summed up in the following quote from the cautious Cleverdon :

"It appears difficult to reach any other conclusion than that, within the parameters of this test, natural language searching on titles and abstracts proved at least equal to and probably superior to searching on controlled language".

In fact for virtually every parameter used to compare the two techniques of searching, ie. natural language and controlled language, the former was demonstrated to be superior.

5.8 In my own view, these results have shown that retrieval using the developed natural language index was not only "not significantly inferior to" but was actually superior to that obtained on the controlled vocabulary. The techniques needed to create the index were tested and shown to be perfectly practical. We can see that retrieval on titles only is futile though it could be a convenient adjunct to a controlled vocabulary search. However, the amount of processing time needed to create the inversion on the abstract text is not insignificant, whilst the disk demand for the inverted files can more than double, when compared with the basic IF on descriptors, authors and corporate source.

5.9 I hope that further work will be possible on the integral database concept. I believe the results obtained to date, with an index which incorporated only one of three planned approaches and is certainly capable of further improvement, have already demonstrated the feasibility of the approach. I hope also to see the experimental NASA STAR File further developed according to the original plans and to see a second database given identical processing and the inverted files merged. This would constitute the first experimental integral database for evaluation.

5.10 At a different level complete re-design of the online applications software has been in progress for over two years and a very new version is on the way which will include the implementation of the "EURONET Common Command Set" referred to later.

5.11 A policy of database standardisation has been developed such that the different files should appear to be more homogeneous to the user. A series of standard formats has been specified for display, type and print outputs. As these are implemented on all the files, it means that any given data element, be it a document type or a classification code, will always appear at the same position in the record; data elements which are identical but are described by different terminology by the database producer (eg. free-term, identifier, keyword, etc) will be given standard names, eg. controlled term (for descriptor in NASA parlance); uncontrolled term (for identifier on COMPENDEX); category code (section number in Chemical Abstracts). Where possible, abbreviations will be expanded to the full form and a standard table of language names will be introduced. Prefix codes (AU=) and Suffix codes (/CS) in the inverted file will be standardised.

6. EURONET : THE RELATIONSHIP BETWEEN ESA AND THE COMMISSION OF THE EUROPEAN COMMUNITIES (CEC)

6.1 The name "EURONET" means many things to many people and indeed, embraces a very wide spectrum of STI possibilities. A significant amount of money has been allocated to the project - approximately 6 M dollars for the first 3 year period 1975-1977, and a similar amount for the second 3 year period 1978-1980. One CEC spokesman has explained it by saying : "We wish to create a Common Market in scientific and technical information ...... information is a commodity which can be bought and sold like any other".

6.2 EURONET will be a packet-switched STI network serving the member countries of the CEC (France, Italy, Germany, Belgium, Luxembourg, the Netherlands, Ireland, Denmark and the United Kingdom). It will also encompass Community policies on STI, embracing the types of database and databank needed in the CEC countries and including the sponsoring of new databases and databanks where these are seen to be required, and on multi-lingualism ultimately intended to permit access to computer data via all Community languages.

6.3 EURONET as a physical network is being created by a consortium of all nine PTT's from the CEC member countries. The current target date for operation is autumn 1979. The presently foreseen network will be a two-level system with four network switching nodes located at Frankfurt, London, Paris and Rome and five remote concentrators at Amsterdam, Brussels, Copenhagen, Dublin and Luxembourg. Initially, only teletype compatible terminals will be supported. The first host computers scheduled to be connected to EURONET include those of IRS; DIMDI (Deutsches Institut für medizinische Dokumentation und Information) in Cologne; INKA (Fachinformationszentrum Energie, Physik, Mathematik) in Karlsruhe; the British Library Service BLAISE; and a second UK venture named INFOLINE.

6.4 Of the nine member states of the CEC and the eleven of ESA, eight are common to both. It is perhaps hardly surprising that, in view of the comprehensive European STI capability already developed by IRS, and the far-reaching plans of EURONET, ESA and the CEC should be encouraged by the member states to discuss opportunities for cooperation. The result was an exchange of letters between the Director General of ESA and the Director General of the Directorate General for Scientific and Technical Information and Information Management of the CEC.

6.5 Europe needs a reliable trans-European TYMNET type STI network as soon as possible. To help achieve this, following the exchange of letters ESA has collaborated with the CEC to the maximum extent possible. This has included the detachment of two IRS staff to the CEC's DG XIII in addition to a very considerable amount of staff time which has been allocated. A so-called "Mixed Team", comprising three CEC staff and three from ESA, all at high level, is charged with the practical implementation of the cooperation agreed. In addition, IRS staff are inevitably involved in many committee and working group activities associated with EURONET.

6.6 Although EURONET suffered the inevitable delays of any large-scale multinational project, by the end of 1978 there were unmistakeable signs that it would be available in the comparatively near future. A contract was placed with a large marketing agency for the pre-launch phase; a host-network technical unit was created to coordinate host connections; and a user relations unit was set up; a launch team was being established to "help the smooth running of EURONET operations". EURONET also acquired a new name - DIANE ! The new acronym stands for "Direct Information Access Network for Europe".

6.7 By March 1979 the CEC announced that the REX-25 testing facility which had been opened to host information services was being used successfully; this facility (located in Rennes, France) simulated the X-25 standard interface to be used for host-to-network links on EURONET. At the same time it was announced that switching equipment had been delivered to all sites with the exception of Rome! Negotiations with Switzerland and Spain for EURONET access were proceeding well and a dialogue had been opened with Sweden.

6.8 IRS' own preparations for connection with EURONET by mid-1979 were extensive. Delivery of the "black box" device to be used for the connection, a TERPAC H-12A unit made by the French company Sitintel, was scheduled for end-June, to coincide with the installation by SIP (the Italian telephone company) of the line to connect IRS with the Rome node of EURONET. About two months of testing were foreseen, falling in the seasonally quieter months of July and August, leading to an operational EURONET connection by September. Major changes had been under development for the applications software for some considerable time, including the provision of a totally new command language, CCS - the EURONET Common Command Set specified by the CEC for all EURONET hosts. The target for September 1979 was release of version 1 of IRS' implementation of CCS, with the user being offered the choice of IRS' traditional command language or the new CCS by means of a switch.

REFERENCES

1. *Martin, W.A. A comparative study of terminal techniques ...... Paper presented to the First European Congress on Documentation Systems and Networks, Luxembourg, 16-18 May 1973, EUR 5058, pp. 107-164*

2. *Martin, W.A. Publication pending*

3. *Wessel, A.E. Computer-Aided Information Retrieval. Melville Publishing Co., Los Angeles, Calif., 1975, 176 p. ISBN/ 0-471-93376-7*

4. *Niehoff, R.T. Development of an integrated energy vocabulary and the possibilities for online subject switching. J. American Soc. Inf. Science, V27, n 1, Jan - Feb 1976, pp. 3-17*

5. *Bhattacharyya, K. The effectiveness of natural language in science indexing and retrieval. J. Documentation, V30, n 3, Sept., 1974, pp. 235-254*

6. *Cleverdon, C.W. A comparative evaluation of searching by controlled language and natural language in an experimental NASA database*

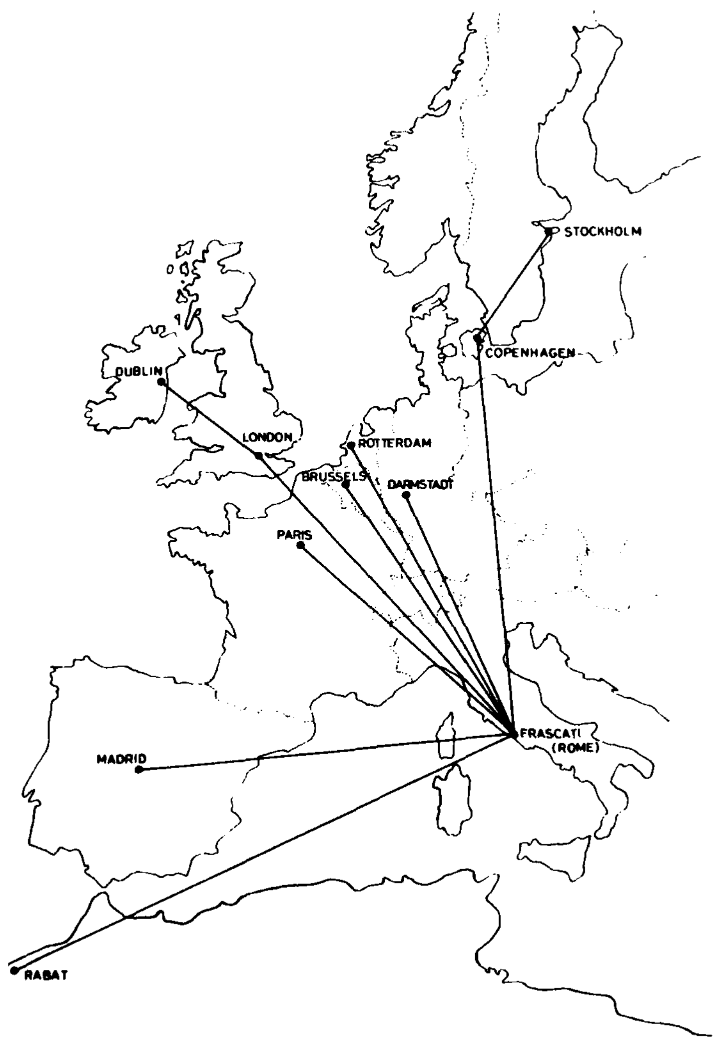
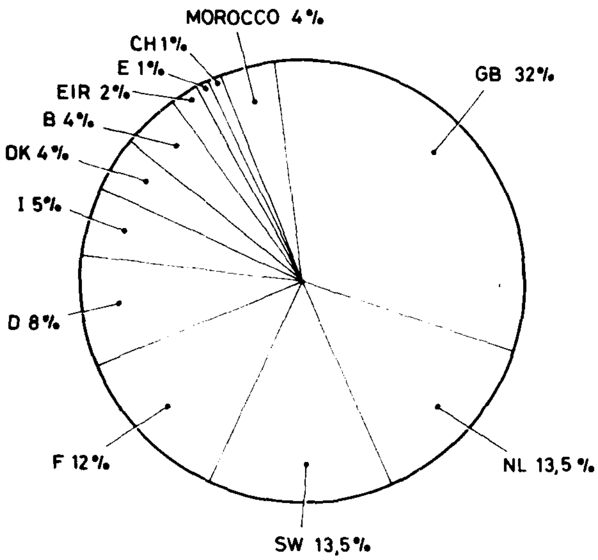
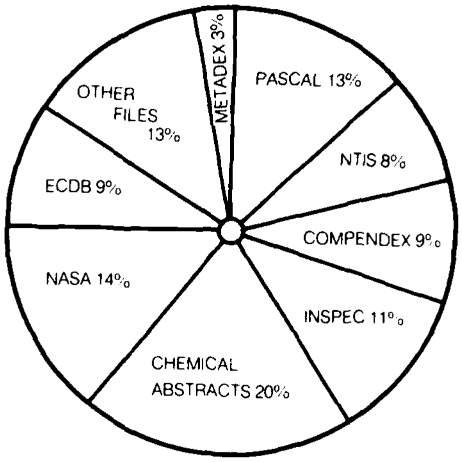


Figure 1



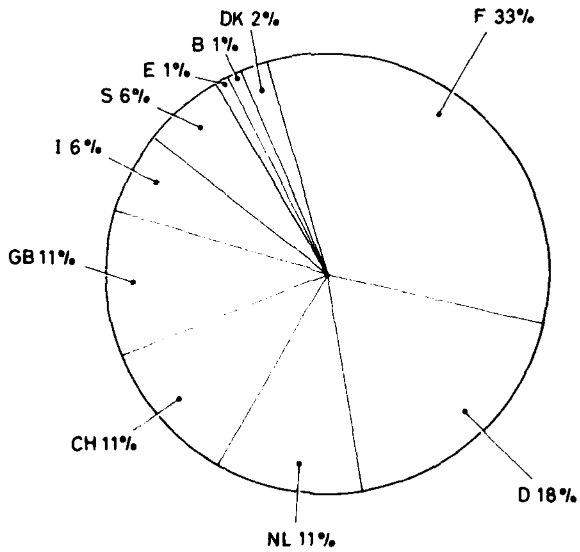
GEOGRAPHICAL DISTRIBUTION OF ONLINE UTILISATION OF AEROSPACE INFORMATION DURING 1978

Figure 2



Total use of files in percentages.

Figure 3



GEOGRAPHICAL DISTRIBUTION OF ITEMS PROCESSED BY ESA IRS FOR NASA STAR 1978

Figure 4

The Role of NASA for Aerospace Information

George P. Chandler, Jr.
Chief, Scientific and Technical Information Branch
National Aeronautics and Space Administration

INTRODUCTION

In 1962 the United States announced its commitment to put a man on the moon by the end of the decade. In the exciting and stimulating environment created by the launching of this commitment NASA began its activities for providing scientific and technical information to the scientists and engineers working to solve the problems of manned space flights. This year we observe the tenth anniversary of man's flight to the moon and safe return to Earth, -by any measure a remarkable feat of applied engineering and of human technical genius.

But from the beginning NASA recognized that its goal of putting a man on the moon was but one part of an aerospace program of research and development that would provide a broad variety of benefits to all mankind. Stated clearly among NASA's goals is the need to disseminate as widely as practicable the scientific and technical knowledge developed as part of the nation's aerospace activities. Today the NASA STI program has become a valuable national and international resource with more than 1.6 million documents recorded in its data base and growing at a rate of 80,000 documents per year. The collection of documents continues to grow in subject scope as well as in size in order to meet the needs of new NASA programs, aerospace programs of other nations and a growing diversity of users. In responding to the needs of this growing universe of users, NASA has developed a highly sophisticated and effective automated system of providing direct access to its data base. It is a dynamic and flexible system that constantly monitors requirements for information by government, industrial and economic groups throughout the world. The needs are translated into new services and products many of which are made possible only through the effective utilization of new techniques and new technologies.

Despite the reduced level of the current U.S. space program, because of budget limitations, many exciting programs are underway: the Space Shuttle, planetary probes, satellite surveys, and others. In addition, many more nations throughout the world now have ongoing programs of research and development related to aerospace activities that create a continuing flow of new information. Certainly an important current need reflected in the incoming flow of documents and requests for information is the design and construction of aircraft to utilize fuel supplies more effectively. Thus, an increasing audience of users continues to require information on space related developments as well as on the utilization of space-developed technology. Today, the system serves more than 11,000 scientists and engineers within NASA, some 100 NASA contractors, subcontractors, and grantee organizations, and more than 600 elements in U.S. Government agencies. Some 60 countries and over 750 governmental organizations, universities, and research institutes are active participants in the NASA International Exchange program. Through the European Space Agency, much of the NASA data base is made available across Western Europe. Increasingly, the public and commercial interests not directly engaged in aerospace activity are making use of the system.

On a typical day the NASA Scientific and Technical Information Facility receives requests for information from a wide spectrum of users. The president of a manufacturing company asks for information relevant to a materials problem. A nine-year old child writes to ask for information on how to build a rocket. In addition, each day hundreds of scientists, engineers, administrators, legislators, and others access online the NASA Scientific and Technical Information data base. They seek information that will help solve problems not only in the fields of aerospace activity, but also in such critical areas as the increase of the world's food supply, control of floods, location of new sources of minerals and energy, and in providing accurate warning of impending natural disasters.

Scope and Coverage of the NASA System

To achieve completeness of coverage and quality of content, NASA has an aggressive policy of seeking sources of documents. NASA utilizes an established worldwide network of cooperative public and private sources, from which it collects each year about 90,000 - 100,000 documents. These incoming reports as well as journal articles, reviews and books from government, industry, research institutions and universities contain the findings of physical scientists, social scientists, medical scientists, engineers and others who contribute to and are concerned with progress in aerospace development. Along with these reports and published articles, the system regularly receives documents and information on NASA-owned inventions covered by U.S. patents and applications for patents, contracts let by NASA, NASA research and technology programs underway or planned, and related research in progress. Such information is also gathered from other U.S. Government agencies and other major sources of aerospace information throughout the world.

The subject matter of these documents ranges widely and includes such areas as aeronautics, astronautics, chemistry and materials, engineering, geosciences, life sciences, mathematical and computer sciences, physics, social sciences, and space sciences.

The NASA STI system is managed by the Scientific and Technical Information Branch (STIB) at NASA Headquarters in Washington, D.C. The actual operations are primarily performed in two contractor-operated facilities. The NASA STI Facility, located near Baltimore, Maryland, employs about 210 people who process report literature, operate the computer complex and provide support for software maintenance and developments. A second contractor, the Technical Information Services of the American Institute of Aeronautics and Astronautics, employs approximately 80 people in New York City and processes the open literature -- journals, magazines, books, etc.

The scope and versatility of the NASA STI program are apparent from the following service highlights:

1. Publication of four announcement journals, two on a biweekly basis, one on a quarterly basis, and one on a monthly basis.

2. Publication of a biweekly current awareness service covering almost 200 separate information categories.

3. Publication of six continuing bibliographies covering specific fields such as "Energy," "Earth Resources," and "Aeronautical Engineering."

4. Distribution of about three million microfiche per year.

5. Distribution of approximately 1.5 million hardcopy documents per year.

6. Provision of interactive data bank and retrieval capability to NASA Field Centers, NASA Industrial Application Centers (IACs), and others having NASA/RECON (REmote CONsole) terminals.

7. Provision of documents and microfiche for sale to the general public.

8. Preparation of indexes to NASA Congressional Hearings, speeches, press releases, and management documents.

9. Provision of magnetic tapes and microfiche to ESA for service to the European aerospace community.

10. Provision to accept system input from ESA through machine readable magnetic tapes.

11. Operation of the NASA library network (NALNET), which provides online search and retrieval access to all books held by NASA libraries.

Online Access Via RECON

The features of the NASA system which are most highly regarded by users are the NASA data base and the RECON system for online interactive access. Terminals link NASA installations and NASA Industrial Applications Centers throughout the United States to the Facility's central computer. Access is also available to other government agencies and to NASA contractors via dial-up capability. In addition to direct online searches via a terminal, data base searches are provided to registered users on request to the Facility in Baltimore, NASA Field Centers, and the NASA Industrial Application Centers. More than 11,600 searches are being made each year (3,600 at STI Facility; 8,000 outside the Facility) and the usage continues to grow. Currently, NASA Centers and Industrial Applications Centers have started using new intelligent terminals as the first step toward a distributed processing network that will provide many added software options at the local level as well as relieve the central computer of many routine tasks. Adjuncts to this retrieval system permit access to the data bases of the U.S. Department of Defense and the U.S. Department of Energy.

As we gain more and more experience in making computer searches with RECON, we become more impressed with the power and versatility of this interactive search and retrieval tool. Not only does it provide expanded capabilities and ease of use in meeting the information needs of scientists and engineers, but it is enabling others to use the scientific and technical literature as a source of economic and marketing data. For example, the results of such searches can produce information on the names of companies doing work in a particular field and thus suggest opportunities for new markets or for mergers or joint ventures. Information from reports of research and technology underway or planned is being used to identify newly developing fields which will foster tomorrow's markets and new technological advances. Increasingly, the NASA STI base is used to produce information of value to workers in the fields of contamination control, medicine, food, technology, and many others.

Since the computer provides only bibliographic citations and abstracts, it is essential that users be provided access to full documents as rapidly as possible. Regular delivery of documents, either in full-size printed form or in microfiche, is made by NASA through its STI Facility to more than 1,900 organizations in the U.S. and to more than 750 abroad. Regular delivery of documents and special publications is supplemented by sale through the U.S. Government Printing Office (GPO) and the U.S. National Technical Information Service (NTIS) to assure the widest possible public availability. Currently, a system of direct ordering of documents via the computer system is being tested.

A significant part of the computer software development program is concentrated on expanding the usefulness of the RECON system. Expanded usage creates demands for new capabilities and for greater ease in the use of the system. Each year, within the limits of our budget resources, we do as much as we can to incorporate new changes and to take advantage of new technology. In the past year, we have made significant progress in improved system reliability and addition of several important new features for users.

Currently being tested as a new capability of the NASA RECON system is a selective dissemination feature whereby individuals may establish and retain in the computer a profile of their specialized interests. This profile is then used to retrieve information covering a given time period, usually that time which includes all document accessions during the previous month and contained in the various announcement journals. Users may update or change their profiles via a direct access terminal online. While intended to supplement existing alerting services, it is not unlikely that this type of computer capability may replace some printed products.

Announcement Services

The basis of NASA information services is a series of announcement journals generated from the computer based information files in the form of comprehensive or specialized bibliographies and abstracts varying in frequency of issue from semimonthly to quarterly. These announcement journals include: Scientific and Technical Aerospace Reports (STAR) which provides bibliographic citations and abstracts on some 1,000 accessions on report literature per issue; Limited Scientific and Technical Aerospace Reports (LSTAR) which provides abstracts and indexes of security-classified and limited distribution documents; International Aerospace Abstracts (IAA) which provides coverage of the published literature of approximately 1,600 accessions per issue. A service of particular value to individuals is Selected Current Aerospace Notices (NASA/SCAN). This biweekly service divides the content of STAR and IAA into some 200 subject profiles, thus providing highly specialized notification on new information entered into the system. STAR, IAA, SCAN, LSTAR, and other NASA announcement publications such as Computer Program Abstracts (CPA) serve to keep users aware of new additions to the document and information collection and to facilitate retrieval of required material. In all, about 40,000 journal articles and books and approximately 24,000 technical reports are announced each year.

To provide for prompt and effective dissemination of these publications, NASA maintains a computer-based mailing list for the generation of mailing labels and control of distribution. The considerable in-depth computerized data on users, products and services allows us to segment, manipulate and cross-match records so as to put any user -- or groups of users -- into categories which will permit us to examine statistically the who-where-what and how many of our audience and its demands. This computerized system assists in establishing resource requirements and allows evaluative analyses to be performed in setting future courses of action.

An important element of the NASA mission has become the transfer of aerospace technology for problem-solving and application in the private sector. This includes dissemination of information and online access to the NASA Data Base and also technical assistance in the evaluation of the commercial potential of a given product. The NASA STI Facility also maintains a computerized mailing list of some 20,000 organizations and individuals not directly connected with the U.S. space program who are mailed information about commercially useful spin-off developments from NASA research efforts. NASA stimulates interest in such opportunities through an active program of announcing innovations deemed to be of value to manufacturers for the development of new products. The principal such vehicle is a quarterly publication called NASA Tech Briefs. In further support of this effort, NASA has set up a technology utilization network currently comprised of seven regional Industrial Applications Centers which perform a type of "push or stimulation" dissemination. The IACs interpret the essence of specific available information for practical and useful applications and suggest its usefulness through technical consultation with potential developers. More recently, NASA has established two state technology centers to test the transfer of technology at the state level in the U.S. A major source of information for the IACs and the state centers is the NASA STI system.

International Document Exchange

In the early 1960s, prior to the NASA/ESA arrangement, NASA initiated and implemented an international document exchange program as one of its first moves to develop cooperative programs with European and worldwide organizations. The major objective of the program was to promote international cooperation through an orderly information transfer process. Governments, academia and selected research establishments within the respective countries were made exchange partners. NASA provided selected services, but primarily the semimonthly abstract journal STAR and its semiannual cumulative indexes in exchange for recent technical documents, journal articles and serials. As mentioned earlier, some 60 countries and over 750 governmental organizations, universities and research institutes are active participants in the NASA international exchange program. Receipts from this program and the ESA arrangement account for about 15% of the accessions in STAR and about 5% in IAA.

Recently a new concept was inaugerated in Europe to enhance the acquisition of documents for the NASA data base in exchange for access to the NASA STAR and IAA files via ESA RECON. It is our "Tripartite" agreement program. It stems from an agreement among NASA, ESA and a third party, e.g., an organization in a participating ESA member state. This program has enabled ESRIN-IRS to offer throughout the ESA community, via the ESA RECON network, direct online access to NASA STAR and IAA files. A major provision of the agreement is the requirement to provide one relevant document input to ESA (and thus NASA) for each hour of access to NASA files. At the moment there are about 300 Tripartite agreements throughout Western Europe. Results of the program and status of the Tripartites are reviewed semiannually.

ESA and NASA RECON

The mechanized, interactive online retrieval system at the European Space Agency is a result of a mutually beneficial exchange agreement between NASA and ESA (then ESRO) in 1962. At that time, NASA agreed to make available to ESA its monthly input of STAR and IAA references in machine-readable form in exchange. Today this input by magnetic tape to ESA now totals about 1,000,000 accessions.

In exchange, ESA through its ESRIN-IRS (Information Retrieval Service) at Frascati, Italy, supports NASA by providing the major document acquisition effort in Western Europe for STAR; providing a portion of the document acquisition effort for IAA; processing (cataloging, abstracting and indexing) on tape, in machine-readable form for direct input into the NASA system, about 2,200 STAR accessions and placing these accessioned items on microfiche; forwarding to the NASA STI Facility up to 1,500 other items for input into the NASA data file under one of the NASA unpublished accession series; producing about 100 English language translations of selected technical French and German language reports for input to the NASA STAR file; acquiring and providing to NASA select, difficult to acquire, Russian language aerospace reports through a special document exchange between ESA and the Institute of Space Research, Academy of Sciences, Moscow. The Russian material is provided in exchange for ESA sponsored material. Microfiche masters are forwarded to NASA, along with the documents and their related processing forms, as well as as a machine-readable tape. NASA accession numbers are automatically assigned to the documents after they are processed for entry into the NASA data base. Periodical or open literature items are submitted for review to the Technical Information Service of the American Institute of Aeronautics and Astronautics and those selected for continuous receipt are acquired for inclusion in the semimonthly abstract journal International Aerospace Abstracts.

The online retrieval systems of NASA and ESA have a common ancestry in that both were initially developed by Lockheed in the late 60's and early 70's. Both RECON systems have since been considerably enhanced, and an effective exchange of information on new RECON developments is maintained between NASA and ESA. The ESA main computer facility is located at Frascati, Italy and serves the ESA RECON Network across Western Europe.

In 1975, ESA took on increased responsibility for the European area by collecting and processing the reports involving several hundred organizations. In 1977, NASA in cooperation with ESA produced the camera-ready copy of ESA SP-1006, INDEX OF ELDO PUBLICATIONS. Containing about 1,600 citations and abstracts with six computer-produced indexes, the document covers significant reports generated by and developed for ELDO, the European Launcher Development Organization.

NASA and AGARD

NASA engineers, scientists and technical information specialists have participated in the preparation of AGARD technical reports, books, manuals and journal articles since NASA became an active participant in AGARD. NASA STIB began more direct participation in the AGARD publications program in 1974 when the camera-ready copy of the AGARD INDEX OF PUBLICATIONS, 1971-1973, was prepared at the NASA Scientific and Technical Information Facility. This publication included an abstract section and five computer-generated indexes -- all from the NASA data base. Subsequently, another index for the period 1974-1976 was prepared. In addition, manuscript copies of the bibliographic sections for several AGARD Lecture Series topics, e.g., Strapdown Inertial Systems, Energy Conservation in Aircraft Propulsion, and Methodology for Control of Life Cycle Costs of Avionic Systems have been prepared by the NASA STI Facility. At the moment the AGARD Multilingual Aeronautical Dictionary is in production, and plans are being completed for the production of a third triennial index of AGARD publications for the period 1977-1979. The dictionary is a large and complex publication involving nine languages, including a need for many special characters and accents such as cyrillic and Greek. Scheduled for actual publication in early 1980, the dictionary will be a product of utilizing, at the NASA STI Facility, the latest developments in computer assisted composition and photocomposition technology.

The growing effectiveness and use of the NASA information system reflects the scope and the growth of aerospace programs by the nations of the world. The scope covers not only man's achievements in space, but the wide array of benefits accruing to man as a result of the research and development supporting those achievements. Increasingly, through the work of organizations such as the European Space Agency, those benefits are being made available more widely throughout the world.

We have attempted to use as much new technology, know-how and experience as possible in the building and operation of the NASA Scientific and Technical Information Program. We plan a continued program of keeping the system a highly effective source of information for all concerned with aerospace development and progress. In the near future we see the demands for access to the information contained in the NASA STI system more and more related to the rapid social, economic and political changes taking place throughout the world. It has never been more important that our information systems must be guided by a philosophy of readiness to change rapidly and effectively in anticipation of new needs and new demands. Not only for NASA, but for all information systems, the goals must be an increasing understanding of the ever more complex needs of our users, the provision of fuller access to all kinds of information, and the fulfillment of needs for specialized user requirements.

STATE OF THE ART OF STANDARDIZATION OF BIBLIOGRAPHIC DATA ELEMENTS

by

Harold Dierickx

Director, UNISIST International Centre for Bibliographic Descriptions (UNIBID),
The British Library, Research and Development Department, Sheraton House,
Great Chapel Street, London W1V 4BH, UK.

SUMMARY

The current situation with regard to standardization of data elements in bibliographic records is summarized. The following areas are considered: headings, bibliographic description proper, other areas or particular attributes of the bibliographic record. Mainly international and national rules and standards of international significance are dealt with.

Although the difficulty for standardization efforts to suit the different requirements of various user groups is recognized, the conclusion reached is that standardization of form and presentation of bibliographic data elements is an essential prerequisite to achieve international compatibility of bibliographic records, particularly in an automated environment.

1. INTRODUCTION

The present paper endeavours to summarize the current situation with regard to standardization of data elements in bibliographic records. The areas considered are: (1) headings, (2) bibliographic description in the strict sense, (3) other areas or attributes of the bibliographic record. To avoid any possible misunderstanding, we should like to point out that we understand cataloguing rules to be broader than mere rules for bibliographic description: cataloguing rules relate to the entire bibliographic record, i.e. the description of the item recorded as well as the choice and form of headings and any possible auxiliary information (e.g. codes for languages and names of countries, identification numbers, rules for abbreviations and romanization, etc.); rules for bibliographic description may cover some auxiliary information in addition to formalizing the description of the item referred to in the bibliographic record, but not choice and form of headings (entry points for filing and other forms of processing).

Although we were asked to give a state of the art report we have not attempted to mention every conceivable standard or similar normative document or every initiative towards standardization possibly undertaken. We have in fact limited our attention to mainly international and national rules and standards of international significance. In this way we hope to have succeeded in presenting a reasonably clear overall picture in which it is possible to distinguish the wood from the trees.

2. STANDARDS AND RULES FOR THE ENTIRE BIBLIOGRAPHIC RECORD

Such rules are normally referred to as "cataloguing codes" or "cataloguing rules" and they cover all aspects of the description of a bibliographic item in a record (for the purpose of this paper the terms "bibliographic record" and "record" are used interchangeably). As yet no universally accepted set of cataloguing rules exists, but there are a number of national rules and rules for special applications with varying degrees of international adherence. Several of these have based themselves on a set of general cataloguing principles developed under the auspices of the International Federation of Library Associations and Institutions (IFLA), viz. the "Statement of principles adopted at the International Conference on Cataloguing Principles, Paris, October, 1961" (1). This is particularly true for the AACR and the RAK which are briefly dealt with, among other codes, further on in this paper. These principles, often referred to as the "Paris principles", apply only to the choice and form of headings and entry words, i.e. the principal elements determining the order of entries in catalogues and other alphabetical lists of books and other library materials. Essentially these headings and entry words are author names and titles. The Paris principles have had a profound influence on all new cataloguing codes and revisions of codes that have appeared since (2).

2.1 The "Anglo-American Cataloguing Rules" (AACR)

The AACR are widely used in English speaking countries and more or less adapted translations are in use in many other countries. The first edition comprised a British (1967) and a North American (1968) text which were not completely identical. A case in point was the form of names of organizations as headings. A second revised edition was published in 1978 (3). This time there was only one text and among the stated objectives of this edition may be mentioned: (1) closer conformity to the Paris Principles; (2) particular attention to developments in the machine processing of records; (3) adherence to the principles embodied in the International

---

Any views expressed in this contribution are strictly those of the author and are not necessarily the views of UNIBID or The British Library.

Standard Bibliographic Descriptions (ISBDs) and in particular to ISBD(G) (the ISBDs are dealt with in the next chapter of this paper). The 1978 edition also contains a chapter on the description of "analytics".

2.2 The "Regeln fuer die Alphabetische Katalogisierung" (RAK) (4)

The "Rules for alphabetical cataloguing" (RAK) have replaced the former so-called "Prussian rules" (5) in the German speaking countries and are very comprehensive, covering filing and transliteration rules, lists of abbreviations as well as rules for headings and bibliographic description proper.

2.3 The USSR Cataloguing Code (6)

This code consists of a series of publications covering all types of library materials, published by the USSR Cataloguing Committee since 1960. These rules are used in all USSR Republics, in Mongolia and in Vietnam.

2.4 Specialized Rules for the Entire Bibliographic Record

The three codes referred to above have been conceived in the first instance for libraries. The rules which we shall mention in the present section are different to the extent that they have been drawn up primarily for another broad group of information services, in particular abstracting and indexing (A&I) services, and in some cases are meant for a more narrowly defined category of bibliographic material (e.g. government scientific and technical reports) or of users (e.g. authors of scientific and technical literature). A further distinctive characteristic of this group of instruction manuals is that they are usually little or not at all concerned with specific rules for creating headings because they intend to combine rules for bibliographic description with a record structure in a way which allows for the automatic generation of a variety of headings and other access points by computer processing. It should be realized also that the differences between these two categories of cataloguing codes are not so absolute as one might assume at first sight. Many of the A&I manuals have relied in part on library cataloguing instructions (e.g. the AACR) for the formulation of their rules for bibliographic description. Conversely, at least some of the A&I cataloguing rules could also be used in many libraries.

2.4.1 The UNISIST Reference Manual for Machine-Readable Bibliographic Descriptions (RM) (7)

The RM may be considered as the prototype of an A&I cataloguing manual which contains rules for bibliographic description as well as specifications for record structure, tape formats, character sets and any other information needed to prepare and exchange complete machine-readable bibliographic records. Although intended in the first place as a specifications manual for an exchange format to be used by systems analysts, it can also be used as a manual for bibliographic description and has thus far been used as such more than as a set of rules for an exchange format (8). Form and content of bibliographic data elements are defined at field and subfield level but no firm instructions are given on the creation of headings. Data elements and essential sub-elements are individually identifiable and this permits automatic generation of a variety of headings via computer processing at the discretion of individual user systems. Although specifically designed for automated environments, the rules for bibliographic description in the RM can also be applied in manual systems. The 1974 edition covers monographs, serials, reports, theses and patents. Draft extensions for translations and standards have been prepared. Eventually all bibliographic materials will be covered. A revised loose-leaf edition is planned for 1980. Regular updating and extension of the RM is carried out by the UNISIST International Centre for Bibliographic Descriptions (UNIBID) located in London, UK, in the Research and Development Department of the British Library.

2.4.2 Leitfaden fuer die formale Erfassung von Dokumenten in der Literaturdokumentation (LED) (9)

The "Guidelines for processing of documentary literature" (LED) is a complete cataloguing manual aimed in the first place at A&I services in the Federal Republic of Germany. It has been conceived to suit manual as well as computerized systems. It covers the description of all kinds of bibliographic materials, although the emphasis is on printed material, and contains also instruotions for choice of headings, form of citation and filing. It may be considered to be for A&I services in Germany what AACR is for libraries in English speaking countries. In addition to ISBD(M) (10) and ISBD(S) (11) it has also drawn on the COSATI (12), Chemical Abstract Service (13) and INIS (14) rules for bibliographic description. As most other manuals in this category, it allows for the automatic individual identification of data elements and many of their significant components.

2.4.3 Guidelines for Descriptive Cataloguing of Reports (COSATI rules) (12)

These rules for the bibliographic description of technical report literature were first published in 1966 by the US Committee on Scientific and Technical Information (COSATI) of the Federal Council for Science and Technology. An updated version was issued in 1978. The COSATI rules are addressed to libraries and information centres that are processing/exchanging bibliographic information in the report literature field, especially on magnetic tape or on-line. No instructions for machine processing, tape formats or character sets, nor for creating headings, are included. This is purely a manual for bibliographic description for a very specific type of literature but, as such, very complete although limiting data elements to the minimum necessary for identification and retrieval (excluding subject retrieval). The Guidelines are being used in the USA by - among others - the Defence Documentation Centre (DDC), the National Technical Information Service (NTIS), the former Department of Energy (DOE) (and presumably by its successor), the Aeronautics and Space Administration (NASA). They also have had an influence on the cataloguing manuals of international information systems like INIS (International Nuclear Information System) (14) and AGRIS (International Information System for the Agricultural Sciences and Technology)(15), and apparently on many more information services in the United States and elsewhere (12, p.4).

2.4.4 Rules for bibliographic references

Bibliographic references which appear in bibliographies, in lists and in footnotes in books and in other published and unpublished documents, are usually made by authors and editors and contain only the strict minimum of bibliographic information necessary to identify and retrieve the items cited. They do not contain all the elements of the fuller bibliographic description required by national and some other libraries or A&I services. An international standard is available: "ISO 690-1975: Documentation - Bibliographical references - Essential and supplementary elements". This standard is now being revised on the understanding, among other things, that the bibliographic reference will be based on information taken from the piece referred to and that the order of the ISBDs but not the punctuation will be followed. The present draft of ISO 690 covers only references to monographs, serials and parts in such publications (e.g. chapters in books, journal articles) but an extension to cover other types of material is envisaged.

A national standard equivalent to ISO 690 which is virtually complete in terms of materials covered and detailed guidance for implementation is "ANSI Z39.29-1977: American National Standard for bibliographic references".

In addition to these rules for shorter bibliographic references there exist rules for formatting still shorter references to aid identification and ordering of serials and parts of serials. "ISO R 30 - 1956: Bibliographical strip" has been revised and is now available as a draft international standard: "ISO/DIS 30: Bibliographic identification (Biblid) of serial publications". ISO/DIS 30 makes it possible to identify parts of serials (volumes, issues, individual articles) by printing a very abbreviated reference on one, more or all pages of the serial concerned as considered appropriate by the publisher. This is particularly useful for the identification of photocopies of offprints of journal articles. Compatibility between ISO 690 and ISO 30 is planned.

3. STANDARDS AND RULES FOR PARTS AND PARTICULAR ATTRIBUTES OF THE BIBLIOGRAPHIC RECORD

3.1 Standards for Headings and Entry Words

This is one of the most difficult areas to standardize because national and various local traditions in this respect are strong and divergent. Nevertheless some efforts towards standardization have taken place and we shall mention those which we consider to be of more importance. Before doing this, however, we wish to point out that, with the advent of computerization, the concept of equal value access points is now considered to be a more useful approach than that of main heading or entry (16). However, in manual systems and in single-entry listings (e.g. a bibliography), the choice of a main heading remains important.

3.1.1 Name headings

The principle of accepting internationally the established national form of personal names is gaining ground. IFLA has issued guidelines on national usages for entry in catalogues of names of persons (17) and a supplementary volume covering more countries is in preparation.

Divergent practices exist also with regard to the choice and form of corporate names as headings. The IFLA Working Group on Corporate Headings is preparing basic principles for form and structure and may reach final recommendations by the end of 1979. The Working Group is using as a main source document a fundamental work on corporate headings by Eva Verona (18). Some information services have established their own authority lists of corporate headings which they update regularly and which are often used by other services as well. A leading example of such a list is that of INIS. It is based on a set of general principles for recording corporate names in author affiliations and as headings contained in the COSATI rules (12). To cover the special problems represented by governments and governmental organizations as authors, IFLA has prepared a list of uniform headings for higher legislative and ministerial bodies in European countries (19) which is being updated to cover a period up to 1978-1979. A similar volume for African countries is also being prepared.

3.1.2 Uniform headings/Title headings

When a work does not have recognizable personal or corporate authors or has many authors, it is general practice to enter it under a uniform heading or under title. Typical examples of such works are religious and liturgical works, serials, etc. IFLA has prepared a list of uniform titles for liturgical works of the Latin rites of the Catholic Church (20) and of uniform headings for anonymous classics in medieval European literatures (21).

The classic example of works being entered under title are serials. The identification and description of a serial title, as is well known, is not always straightforward. Some titles are not specific (e.g. Transactions, Annual Report, Bulletin) and an addition problem is that of changes in the title of the same periodical. The International Serials Data System (ISDS) is providing a means of unique identification of all serials by means of a standardized form of title (the ISDS "key title") associated with its International Standard Serial Number (ISSN). ISDS practice in this connection is based on a Unesco manual: Guidelines for ISDS (22).

3.2 Rules for Bibliographic Description

In this section we deal with rules and guidelines which are strictly concerned with the bibliographic description of documents and nothing else. In other words these rules are not concerned, for example, with headings, subject analysis or filing rules.

A set of internationally prepared and agreed (but not universally applied) rules for bibliographic description have been issued by IFLA. The key document of this series of guidelines is the "General International Standard Bibliographic Description" (23), usually referred to as ISBD(G). The ISBD(G) lists all the data elements required to describe and identify bibliographic materials and assigns an order to and prescribes preceding punctuation for those elements. In addition to this general ISBD, ISBDs have also been developed for monographs (M), serials (S), cartographic materials (CM), non-book materials (including audio-visual materials) (NBM). Further ISBDs are currently in preparation for antiquarian books, printed music and "analytics". Since these ISBDs have been developed over a number of years and since the ISBD(G) was not the first one to appear, there has been some difficulty to keep overall consistency between those different manuals. Some slight discrepancies have also been reported between the ISBD(S) and the ISDS Guidelines (22), which specify the data elements to be used by national centres preparing serial records for input into the international serials data base kept by the ISDS International Centre. The ISBD(An) is also bound to have some repercussions on the other existing ISBDs, especially with regard to the treatment of bibliographic levels distinguished in the processing of multi-level records (e.g. a record describing an individual contribution in a volume of conference proceedings which itself is published in a series). The reason is that the bibliographic description of "analytics" (describing the "lowest" component of a multi-level record - in the example given above this would be the individual contribution) and the larger works of which they form a part requires a more refined treatment of the concept of bibliographic level than is the case in the existing ISBDs.

The ISBDs have been adopted by many national and other libraries, but so far they seem to have had little appeal to other information services.

3.3 Standards and Rules for Other Specific Areas and Attributes of the Bibliographic Record

3.3.1 Standard identifying numbers

International standards have been developed for the unique numbering of monographs (ISBN : International Standard Book Number) (24) and serials (ISSN: International Standard Serial Number) (25). CODEN (26) is also used internationally and particularly in the USA for the identification of serial titles. In contrast with the ISSN, this code is composed of characters, not numerals.

The construction of international identification numbers for patent documents is standardized in: "ISO 3388-1977: Patent documents: Bibliographic references: Essential and complementary elements".

Plans are under way in Technical Committee 46 of the International Organization for Standardization (ISO) to develop an international numbering scheme for technical reports, based on a US standard: "ANSI Z39.23 - 1974: American National Standard Technical Report Number (STRN)", American National Standards Institute, New York, 1974. The report numbers of the US National Technical Information Service (NTIS) are widely used also outside the US.

As to numbering systems for sound recordings, two draft international standards exist: "ISO/DIS 3901: International standard recording code" (ISCR), 1975, and "ISO/DP 5956: International standard number" (ISRN).

3.3.2 Codes for the representation of the contents of bibliographic data elements

International codes exist for the abbreviation of names of countries and languages (27) (28) which often appear in different places of the bibliographic record (e.g. country of origin of a record, country of a personal or corporate author, language of the document described, etc.).

Very useful ISO standards are those governing the abbreviation of "typical words" (e.g. illustration, page, volume, edition, etc.) in bibliographic records: "ISO 832-1974: Bibliographical references: abbreviations of typical words" (under revision), and the abbreviation of titles of serials: "ISO 4-1972: International code for the abbreviations of titles of periodicals". ISO 4 (also under revision) deals with general principles of abbreviation and forms the basis of "ISO 833-1974: International list of periodical title word abbreviations", which lists the actual abbreviations of title words. The latter rules will be withdrawn as an ISO standard but will continue under the auspices of the ISDS International Centre that will also regularly update the list.

ISO has also developed a number of transliteration schemes for the conversion of texts in non-Roman scripts into Roman script. So far proposals have been made for the conversion of Slavic and non-Slavic Cyrillic characters, Arabic, Hebrew, Greek, Japanese, Chinese, Korean and Yiddish characters.

Work is also going on within ISO on the standardization of the so-called "implementation codes". These are codes which denote certain attributes which apply to the entire bibliographic record and which are often important for computerized processing. Codes for "type of document", "medium", "record status", "bibliographic level", "target audience" are typical categories of "implementation codes". There is much confusion about the definition, naming and categorization of these codes and a standard, if developed, will be most useful.

4. CONCLUSION

Let us try now to summarize the overall situation.

As far as cataloguing rules for the entire bibliographic record are concerned, the chances to arrive at a unified internationally accepted set of rules seem to be slim. Different national traditions and requirements do not appear to make such an approach possible. On the other hand, the unifying influence of the Paris principles should be acknowledged, especially in the area of choice and form of headings. Nevertheless the creation of headings remains subject to much diversity and national and individual systems' traditions may prove to be strong barriers to further progress. The positive factors to be indicated here are the rules and guidelines developed or being prepared by IFLA for the choice and form of personal and corporate names as headings. There are also proposals for the creation of international authority files for particular data elements which may be considered to be within the framework of IFLA's programme for Universal Bibliographic Control (UBC). Such files for personal and corporate names come first to mind. There is even a proposal for the creation of an international authority file for entire records (29). It is difficult to see, however, how the authority file solution could be made to work satisfactorily in practice. At best this solution could only apply to a limited number of records because - to mention only these - the organizational problems involved with creating an authoritative record or established form of a number of associated personal and corporate name headings for every bibliographic item recorded by the world's library and other information systems are enormous. Even if the job could be done, the authoritative forms of names, records etc. would presumably be available only subject to considerable delays which would make them less useful at the time when they are required for bibliographical lists, card or printed catalogues, or listings in A&I journals. The creation of "provisional" records or entries while waiting for the definitive form would seem to constitute little progress with regard to the present situation.

Serials are the type of document for which UBC seems to be on the way of reaching its objective of "universal" control. The growing data base of the ISDS International Centre in Paris will eventually include standardized key titles and other information, including a unique ISSN, for each periodical publication.

The area where standardization has made the biggest progress is that of bibliographic description proper, especially with the introduction of the ISBDs. It should be noted, however, that, whereas the ISBDs seem to be popular with many national and other libraries, they have been much less so with *other information services, particularly A&I services.* At least two *reasons why* the ISBDs have *proved* to be less than universally attractive are probably that (a) they are a very elaborate display format with a complex system of punctuation which is bound to appeal less to many smaller and specialized libraries and to A&I services which prefer shorter entries, (b) their structure does not fit easily into the "philosophy" of handling records as favoured by most A&I services and which relies on a type of bibliographic description which allows identification of discrete bibliographic data elements and their sub-components to make possible the generation of a variety of headings and display formats via computer processing. The purpose of the Universal MARC Format (UNIMARC) (30) is precisely to provide the ISBDs with this kind of flexibility but it is not yet operational and the first indications are that the A&I services are not very keen on it.

Progress has been most satisfactory for particular areas of bibliographic description: the ISBN and the ISSN are cases in point. Standardized codes for names of countries, languages and abbreviations of typical words in bibliographic references and in serial titles have been developed and are widely used. All except the symbols for names of countries are under revision. We are also optimistic about the likelihood of developing an international standard for implementation codes (cf. section 3.3.2 above). Progress has also been made in the field of devising systems for transliteration of non-Roman into Roman scripts but it may take a while longer before definitive international standards will be completed.

As a final observation we should like to put forward the proposition that, with the growing computerization of the creation and exchange of bibliographic records, the emphasis in standardization efforts should be in the first place on defining discrete bibliographic data elements and their significant sub-elements (e.g. a personal name as such and its different components: surname, forename, prefixes, suffixes, some titles). Assuming an automated environment, this should be an in-built feature of all bibliographic formats and would permit individual systems to create records and headings to their specific requirements by computer programming (31). This would obviate the need for universal cataloguing rules and an elaborate system of international authority files. It is therefore encouraging to note that an international directory of definitions of bibliographic data elements will serve as the basis for the proposed Common Communication Format (CCF), about which the next speaker will undoubtedly have to say more. The entire CCF exercise will stand or fall with the success or failure of this international data element directory. The deliberations of the Second meeting of the Ad Hoc Group on the Establishment of a Common Communication Format, in Paris, on 2-4 October 1979 (32) revealed very clearly this crucial importance, within the context of bibliographic exchange formats, of the definition and presentation of bibliographic data elements. Whereas the Group could agree easily on a number of mandatory data elements which would have to be part of the CCF and ought to be taken from the piece being described, there was considerable argument as to whether or not these elements should be transcribed in the exact form and sequence as shown on the physical item. Whereas library representatives insisted that this should be the case in all circumstances, it was pointed out that this procedure might not be acceptable to most A&I services. The fundamental reason for these two opposing views rests with a difference in format philosophy already referred to and exemplified in respectively the UNIMARC and the RM. Since formats is the next topic on the agenda of this meeting, I do not wish to enter into any substantive detail but, essentially, the difference between the two approaches is that library formats, like UNIMARC, tend to create multiple entries for a number of data elements: formatted differently for display, for search and filing purposes respectively; whereas A&I formats try to enter each data element only once and formatted in such a way as to allow for automatic display as well as for filing or search. It is clear that the second approach cannot always allow for the transcription of data elements from the piece in the exact form and sequence as shown there, which, however, is still different from saying that this approach would not allow unique identification and description of the item catalogued. Moreover, the question may be asked whether formal adherence to the principle of transcribing data elements following their exact form and sequence on the piece necessarily leads to identical description in all cases. One only has to consider, for example, the complexity of corporate responsibility indication on some documents to become aware of this problem.

Libraries and other information systems may have some different requirements but they all need a selection of the same data elements to identify and retrieve bibliographic records. Even when the headings and the form and extension of the records they need are not the same, it makes good (economic) sense if the elements which these records have in common conform to identical definition and form of presentation. Standardization of discrete data elements will avoid trying to standardize which is unnecessary if not unfeasible, i.e. comprehensive cataloguing rules for entire bibliographic records. Would it be too much to say that, at least in an automated environment, this would be the type of standardization to end all standardization in the field of cataloguing/bibliographic description?

REFERENCES

References to ISO standards or draft standards are not repeated here when their full title is given in the preceding text. All are obtainable from the International Organization for Standardization in Geneva and the year of publication can be derived from the ISO numbering which precedes the title of each standard. The following is a very useful compilation of the texts of ISO standards governing information transfer:

International Organization for Standardization (ISO)/United Nations Educational, Scientific and Cultural Organization (Unesco), Information Transfer, ISO Standards Handbook 1, Geneva, ISO, 1977, 516 p.

(1) International Federation of Library Associations, Statement of Principles adopted at the International Conference on Cataloguing Principles, Paris, October, 1961. Annotated Edition, London, IFLA Committee on Cataloguing, 1971. 137 p.

(2) A. H. Chaplin, IFLA Committee on Cataloguing 1954 - 1974. 2. Retrospective View, International Cataloguing, 3(1), January/March 1974, p.718.

(3) M. Gorman and P. W. Winkler (eds), Anglo-American Cataloguing Rules. Second Edition, London, The Library Association, 1978, 638 p.

(4) Verein Deutscher Bibliothekare. Kommission fuer Alphabetische Katalogisierung, Regeln fuer die Alphabetische Katalogisierung. Vorabdruck, Munich, 1969-1976, several volumes.

(5) Instruktionen fuer die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899, 2nd ed. of 10th Aug. 1908. Unchanged reprint,Wiesbaden,Harrassowitz, 1964, 179 p.

(6) Каталогизационной комиссией СССР, Единые правила описании произведений печати для библиотечных каталогов. I-8., 1960-1970.

(7) UNISIST/ICSU-AB Working Group on Bibliographic Descriptions, M. D. Martin (comp.), Reference Manual for machine-readable bibliographic descriptions, Paris, Unesco, 1974, 71 p.

(8) A. Hopkinson, The use made of the UNISIST Reference Manual: a brief history, in H. Dierickx and A. Hopkinson (eds), Towards a Common Bibliographic Exchange Format? International Symposium on Bibliographic Exchange Formats, Taormina, Sicily, 27-29 April 1978. Proceedings, Budapest, UNIBID/OMKDK - TECHNOINFORM, 1978, 214 p.

(9) C. Hitzeroth, D Marek, I Mueller et al., Leitfaden fuer die formale Erfassung von Dokumenten in der Literaturdokumentation, Munich, Verlag Dokumentation, 1976, 505 p.

(10) International Federation of Library Associations and Institutions, ISBD(M) - International Standard Bibliographic Description for Monographic Publications, 1st standard ed. rev., London, IFLA International Office for UBC, 1978, 59 p.

(11) International Federation of Library Associations and Institutions, ISBD(S) - International Standard Bibliographic Description for Serials, 1st standard ed., London, IFLA International Office for UBC, 1977, 69 p.

(12) Committee on Information Hang-ups. Working Group on Updating COSATI, Guidelines for Descriptive Cataloguing of Reports. A Revision of COSATI Standard for Descriptive Cataloguing of Government Scientific and Technical Reports, Washington, D. C., March 1978, 72 p. (available from National Technical Information Service (NTIS), U.S. Dept. of Commerce, Springfield, Va. 22161).

(13) Chemical Abstracts Service, Chemical Abstracts Service specification manual for computer-readable files in Standard Distribution Format, 3 vols., Colombus, Ohio, CAS, 1972- (this manual is being updated continuously).

(14) G. Del Bigio, C. M. Gottschalk, H. W. Groenewegen, E Ruckenbauer, INIS: Descriptive Cataloguing Rules, Vienna, International Atomic Energy Agency (revised editions are published from time to time).

(15) M. T. Martinelli, AGRIS: Guidelines for Bibliographic Description, Rome, Food and Agriculture Organisation of the United Nations. AGRIS Coordinating Centre, 1973 - (these guidelines are being updated regularly).

(16) R. E. Coward, The Impact of Hardware and Software Technology in Catalogues. In: H.Dierickx and A. Hopkinson (eds), op. cit., p. 68-70.

(17) International Federation of Library Associations and Institutions, Names of Persons: National Usages for Entry in Catalogues, 3rd ed., London, IFLA International Office for UBC, 1977, 203 p.

(18) E. Verona, Corporate Headings: their use in Library Catalogues and National Bibliographies. A Comparative and Critical Study, London, IFLA Committee on Cataloguing, 1975, 238 p.

(19) International Federation of Library Associations and Institutions, U.S.S.R. Cataloguing Committee, List of uniform headings for higher legislative and ministerial bodies in European Countries, London, IFLA Committee on Cataloguing, 1975.

(20) International Federation of Library Associations and Institutions, List of uniform headings for liturgical works of the Latin rites of the Catholic Church, London, IFLA Committee on Cataloguing, 1975.

(21) International Federation of Library Associations and Institutions, R. C. Hewett, Anonymous classics: a list of uniform headings for European literatures, London, IFLA International Office for UBC, 1978.

(22) International Centre for the Registration of Serial Publications. Guidelines for ISDS. Paris, Unesco, May 1973 (document SC/WS/538), 58 p., 3 app.(these guidelines are currently under revision).

(23) International Federation of Library Associations and Institutions, ISBD(G): General International Standard Bibliographic Description: Annotated Text, London, IFLA International Office for UBC, 1977, 34 p.

(24) International Organization for Standardization, ISO 2108-1972: Documentation - International standard book numbering (ISBN), Geneva, ISO, 1972,2 p.

(25) International Organization for Standardization, ISO 3297 - 1975: Documentation - International standard serial numbering (ISSN), Geneva, ISO, 1975, 4 p.

(26) The International Coden Directory is published annually on microfiche and is updated with a mid-year microfiche supplement. The Directory as well as any information on CODEN are available from Chemical Abstracts Service, Ohio, USA.

(27) International Organization for Standardization, ISO 3166-1974: Code for the Representation of Names of Countries, Geneva, ISO, 1974, 24 p.

(28) International Organization for Standardization, ISO/R639-1967: Symbols for Languages, Countries and Authorities, Geneva, ISO, 1976, 12 p.

(29) S. M. Malinconico, The coordination of bibliographic control, in H. Dierickx and A.Hopkinson (eds), op. cit., p. 61-67.

(30) IFLA Working Group on Content Designators, UNIMARC - Universal MARC format, London, IFLA International Office for UBC, 1977, 126 p.

(31) An example of such automatic processing of names is described by:
P. Goossens, F. Heymans, R. Gabriel, PAN: Project for automatic name processing in the context of integrated information processing systems. This is the English translation of a slightly revised version of an article originally published in Dutch in Open, 7(7-8), July-August 1975, p. 352-356. The translation is available from the authors at the Koninklijke Bibliotheek Albert I, Brussels, Belgium.

(32) The Second Meeting of the Ad Hoc Group on the Establishment of a Common Communication Format: Conclusions and Recommendations, Paris, Unesco House, October 2-4, 1979, 3 p. This document is available at the discretion of Unesco's Division of the General Information Programme.

STATE-OF-THE-ART OF DATA EXCHANGE: PROBLEMS OF FORMATS AND STANDARDS

Dr. J.S. Mackenzie Owen
Netherlands School of Business
Nijenrode
3621 BG BREUKELEN
The Netherlands

SUMMARY - The vast growth of bibliographic data services over recent years has produced many problems in the field of compatibility for data exchange. It is now beginning to be understood that there exists a major gap between the library community on the one side, and the Abstracting and Indexing services on the other side. It is also becoming clear that any standardized format for bibliographic data exchange should be flexible enough to incorporate the divers needs of many different users. An effort is now being undertaken bu UNESCO to solve these problems. This paper describes the present situation, problems and future developments in this field.

BIBLIOGRAPHIC DATA EXCHANGE IN PERSPECTIVE

In historical times, the universe of documents available to any individual was more or less confined to the library or libraries in the city where that individual happened to live and work. Nowadays, the universe of documents available to an individual consists of almost all existing documents. This is not because libraries are nowadays so large as to contain all those documents, but because librarians have invented bibliographic data exchange and have integrated it with many and advanced modes of communication.
Written and printed bibliographies were the first means of communicating to other people information about documents that were not immediately at hand. Bibliographic data exchange has also become a means of increasing efficiency. Documentation pools and distributed catalog systems have releaved libraries and information services of having to compile the same bibliographic information many times at different locations.

The huge problems created by the vast growth of scientific and technological publication have been solved to a great extent through the use of computers. And librarians were very early at doing so. This means that nowadays bibliographic information has to be exchanged in a form in which it can be understood and processed by computers. Such forms are commonly well defined information structures referred to by the term bibliographic formats. I shall return to bibliographic formats in the next sections of this paper.

The development of computers and related technology during recent years has resulted in on-line accessability of large-scale information systems as well as in advanced modes of communication. Individuals can now communicate with computerized data-bases on the other side of the ocean. Here, exchange of bibliographic data is a trivial matter. As little or no processing is reqired at the receiving end, it comes down to nothing more reading or printing out information in the data-base. This procedure is becoming more and more common, and, through systems such as Teletext and Viewdata may even bring bibliographic information into the private home. Very large bibliographic information systems and sophisticated communication networks may eventually supersede local processing of bibliographic data. In that case, there will be very little need for bibliographic data exchange in the traditional sense. The problem of standardization in man-machine bibliographic systems will move from the input side of the system to the output side, in the fields of access, retrieval and display formats. In my view, this will have detrimental effects on the flexibility of bibliographic information. Local processing means that the system can be adapted to local requirements, to the local document collection and the local type of user. Centralization and standardization are on hostile terms with individuality. The increasing collectivism and centralized control in society as a whole is, unfortunately, sure to be reflected in our future bibliographic information systems.

THE STRUCTURE OF BIBLIOGRAPHIC FORMATS

At the present stage of technological development, bibliographic data is usually exchanged on magnetic tape. Information on the tape has to have a well defined structure in order that the computer may identify individual bibliographic descriptions and their elements. A bibliographic description as a whole is referred to as a record. The exact lay-out of such a record on tape is defined by a well-known standard called ISO 2709, Format for Bibliographic Information Interchange on Magnetic Tape (1). ISO is the International Standards Organization, which has a special technical committee, TC46, for standardization in the field of libraries and documentation. The scope of ISO 2709 is described as follows:

> "This International Standard specifies the reqirements for a generalized exchange format which will hold records describing all forms of material capable of bibliographic description as well as related records such as authority records. It does not define the length or the content of individual records and does not assign any meaning to tags, indicators or identifiers; these specifications are the functions of an implementation format.
> This standard describes a generalized structure, a framework designed specifically for communications between data processing systems and not for use as a processing format within systems. Although this International Standard is designed for magnetic tape, its structure may be used for other data carriers" (2)

ISO 2709 specifies that a bibliographic record on tape should consist of a fixed length record label containing general information on the record, a datafield directory specifying the position of each data element within the record, and any number of variable length datafields, separated by a field separator (i.e. a special character) and terminated by a record separator.

As indicated in its scope, ISO 2709 specifies nothing more about data elements than that there may be any number of them and that they may be of any lenth (although the total length of a record is usually required to be no more than 2048 characters). This standard is used in most bibliographic data exchange applications, except by a number of large commercial data-base systems for organizational, historic and economic reasons. ISO/TC46 is now working on a revision of the 2709 Standard. This is proving to be a difficult and time consuming task, due to differing views on implementation and processing in general and, of course, due to the economic aspects involved with the large investments in software for existing implementations.

ISO 2709 is not what is generally regarded as a bibliographic exchange format. As mentioned, it is no more than a logical structure specifying the lay-out of a record on the physical exchange medium. A bibliographic format in the usual sense is an implementation of the logical structure. Such an implementation has to specify:

- the elements, i.e. the fields and subfields to be distinguished in the bibliographic description (in theoretical terms: the implementation format specifies the set of domains from which the descriptive attributes are to be drawn)
- the identifiers for these elements, i.e. the codes which instruct the computer as to the type of element involved
- structural aspects, e.g. sequential or hierarchic relationships between elements
- processing information, e.g. indicators specifying repetition of data elements, number of non-filing characters, etc.

One of the most important things to understand about formats is that they do not only specify what information you put into the record, but also what information you can get out of it again. Here we come to one problem of format compatibility. If two formats use different identifiers for the same element, a mapping has to be established between the two (or two mappings between both exchange formats and the internal format of the processing system). This is obviously less efficient than having the same identifier in both formats, but it can be done. But if one format identifies an element and the other does not, a mapping cannot be established and the information is lost. Not identifying an element does not mean that the information provided by that element is not in the record. Imagine format A specifying title + subtitle as a whole, and format B identifying title proper and subtitle as two distinct elements (e.g. as subfields within the title field). Translating format B into format A is done by joining title and subtitle into one element. Translating format A into format B just cannot be done, as it is impossible to get title and/or subtitle out of format A separately.

Format compatibility is an economic necessity for efficient processing and optimal data exchange. One seemingly obvious solution, i.e. an extremely detailed general format, is not the optimal solution because it presents other economic problems, especially at the input side (3). Different users require different levels of complexity and definition in their formatted bibliographic records. A truly general exchange format will have to allow for various levels in a uniform way. Incompatibility between different levels can only be solved by standardizing users. I do not believe this would be a justifiable procedure.

Another major aspect of bibliographic exchange formats is that they are always implicitly or explicitly based on a set of cataloguing rules (or rather: description rules). Such rules do not only specify the various descriptive elements, but they also define them. Very often, different sets of cataloguing rules specify the same elements, but define them differently. For instance, the element 'title' may be defined as title proper in one format and as title proper + subtitle in another. Moreover, there might not be agreement as to what information within a document is to be regarded as 'title proper'. Another example: one set of rules might allow for more than three authors, while a different set might regard a publication by more than three authors as an anonymous document. There are other categorization problems as well. Most cataloguing rules identify authors, editors etc. But there exist rules which identify primary and secondary intellectual authorship. There is no foolproof way of mapping one into the other.

I believe therefore that standardization of cataloguing rules is the most fundamental problem behind compatibility of bibliographic data exchange formats. The library community has gone a long way in solving this problem, as I shall point out later in this paper. These efforts are reflected in the large degree of format compatibility within the library field. The field of Abstracting and Indexing services is much less uniform in this respect. Compatibility between both fields is virtually non-existing. This leads us to another problem in the field of bibliograpic information exchange. For there are various different types of bibliographic records. Libraries produce catalog type records. Abstracting and Indexing services produce reference type records. Catalog type records can be further differentiated according to various document types (e.g. maps, serials, patents, music etc.). Reference type records can be at the analytical, monographic and collective level. As indicated above, bibliographic records are a function of description rules as applied to documents. Different types of description rules, often necessitated by differences between document types, produce different types of records and this explains to a large extent the variety of formats now in existence. Different organizational contexts (e.g. libraries versus A&I services) also contribute towards variety and hence incompatibility. Integrating different types of records into one general exchange format in order to achieve efficient exchange of information is proving to be one of the most difficult tasks in the field of format standardization. This task can only be accomplished through efforts at all levels, i.e. description rules, formats, and the organization of information exchange as a whole.

THE PRESENT SITUATION IN BIBLIOGRAPHIC EXCHANGE FORMATS

The first format I should like to mention here is the MARC format, or rather the MARC family of formats. MARC stands for MAchine Readable Cataloguing and was initially developed by the Library of Congress (4). MARC-LC has a firm background in library cataloguing practice. It is based on the Anglo-American Cataloguing Rules (AACR) and is a more or less faithfull translation of these description rules for monographs into a machine-readable structure. The fact that MARC-LC has given birth to quite a family of highly related formats demonstrates the underlying problems of format compatibility and standardization. Most members of the MARC family are national or regional alternatives. They have been developed by Great Britain, Canada, France (MONOCLE), Italy, Belgium, Denmark, Austria and a group of French speaking West European countries (INTERMARC) (5).

These formats differ from one another in mostly minor respects due to a number of reasons:

- different cataloguing rules and conventions
- different views on processing
- local and organizational aspects (e.g. the bilingual situation in Canada)
- more recent development, offering more sophisticated techniques (e.g. the linking indicator in INTERMARC)

In general, however, there is a fair degree of compatibility between these MARC formats. They all identify more or less the same elements and they use the same identifiers for most identical elements. A good example of format development based on compatibility is the INTERMARC format, which retains MARC-LC as a valid subset. This means that, apart from having to apply minor software changes in processing, MARC-LC records are compatible with INTERMARC, though not vice versa without loss of information.

Another type of offspring from MARC-LC are the MARC formats for other types of documents than monographs (e.g. serials, maps, etc.).

Within the library community there are, unfortunately, several other formats which are not at all MARC compatible. The most notorious example is the German format, MAB-1 (6), which has a completely different structure. MAB-1 is based on the German cataloguing rules (RAK) which are nothing like the Anglo-American Cataloguing Rules. Also, MAB-1 is not entirely compatible with the present version of ISO 2709 (7)(8).

To summarize, the situation within the library community is as follows. Compatible exchange of bibliographic data is usually 100% within any one country, i.e. there is usually at least a national standard. There is a fair amount of international compatibility within the group of MARC-type format users, especially betwee the United States, the United Kingdom and a number of smaller countries (e.g. the Netherlands accept MARC tapes in their national system and will be MARC compatible when producing their own tapes). This is essentially true for monographs. There is much less compatibility between bibliographic records for different document types, for which you need different formats. This is mainly due to the fact that libraries have a much larger organizational differentiation between document types than the Abstracting and Indexing services. However, MARC users are now developing a general format called UNIMARC which is universal in the sense that it can handle a large number of different document types (9). Countries such as Germany, with its highly individual format, fall more or less outside the mainstream of international bibliographic data exchange. But on the whole, libraries all over the world are now making good use of the principal of Universal Bibliographic Control, recognizing that exchange of bibliographic data makes much cataloguing either unnecessary or at least a great deal more efficient.

The situation within the field of Abstracting and Indexing services is rather different. A number of services fall entirely outside the scope of format standardization by not even using the 2709 tape format, e.g. Excerpta Medica and Chemical Abstracts Services (10). Most formats within this field have been developed for specific types of A&I services, e.g.

- INIS for atomic energy information (also used by AGRIS and ASFIS) (11)
- EUDISED for educational information (12) (28)
- IRRD for road research information (13)(14)

In addition, there are various national or regional exchange formats for documentary information, e.g. MADOK (W. Germany)(15)(16) and MEKOF-2 (member states of the Council for Mutual Economic Assistance - CMEA) (17)(18)(19).

So the situation within the Abstracting and Indexing field is one of diversity. There is a large amount of international data exchange, but very little interdisciplinary compatibility. In other words, inter-system connection is impeded by the lack of a general exchange format, and perhaps also by the reluctance to accept one. The underlying problems in this field are slightly different from those in the library field. There are no really uniform description rules within the A&I community. However, this is a much less problematic issue here than in the library field where cataloguing rules are about the most important thing to argue about. Moreover, an international standard for bibliographic references has recently been developed by ISO (ISO 690), and though it is not specifically meant for A&I publications, it might help to provide more basic uniformity in this field.

Differentiation between document types is virtually impossible in the A&I field, at least at the organizational level. This means that you need one format which is universal to all document types and bibliographic levels. This is an advantage and a problem as well. An extremely elegant solution has been found in the UNISIST Reference Manual (20), which is based on a document type / bibliographic level matrix structure. The Reference Manual is a general exchange format developed by UNISIST in co-operation with the Abstracting Board of the International Council of Scientific Unions (ICSU/AB) for use by the Abstracting and Indexing services (21). Unfortunately, it has not been universally accepted, although there do exist various implementations (22).

It has been necessary to differentiate between the library field and the Abstracting and Indexing field in this discussion of existing formats. I have already pointed out that there is very little compatibility between the two fields. In my opinion, this is a serious matter. We could provide extremely efficient service to information users by integrating documentary and library information systems. This would avoid the frustration of finding a relevant citation in a documentary data-base and then having to spend a lot of time, energy and wit in trying to locate it in a catalog, often in another department, with a different type of description, and all that with a fair chance of not being able to find it at all. From a technological viewpoint, integration of both types of systems is feasible. The high degree of incompatibility between both types of information surely is one of the reasons why such integrated systems still hardly exist. We shall have to look at future developments in order to obtain some hope for a type of service we really should be offering our users now.

NEW DEVELOPMENTS IN FORMAT COMPATIBILITY

As I already have indicated, universal exchange of bibliographic information requires standardization of the information itself and standardization of the structure and content designation of bibliographic records, i.e. of bibliographic exchange formats. A great deal of this work has and still is being done by TC46 of the International Standards Organization. Such standards as for biblographic characger sets, standard book numbers, abbreviations of bibliographic terms, translitteration schemes, country and language codes, etc. all contribute towards more efficient bibliographic data exchange. Another organization working in this field is the International Federation of Library Associations (IFLA) which has developed the International Standard Bibliographic Descriptions (ISBD). These specify the lay-out and interpunction of essential elements for bibliographic descriptions. There are several ISBD's for various types of documents. IFLA has also developed a general ISBD(G), which integrates a large amount of different document types (23). It has now been proposed that ISBD(G) be adopted as an official international standard by ISO.

An interesting and extremely important development is now being undertaken by UNISIST, the Information Program of UNESCO, following a recommendation made by the International Symposium on Bibliographic Exchange Formats held at Taormina, Sicily, in 1978 (24). This major project, in close co-operation with other organizations from the library and documentation fields such as ISO, ICSU/AB and UNIBID, is designed to produce a common bibliographic exchange format which can be implemented at various levels of completeness by various segments of the information community. The format will be based on a so-called Data Element Dictionary, derived from a large scale comparative analysis of all major existing formats. One of the most important aspects of this format will be that it will provide a type of flexibility which I proposed in a paper (3) a number of years ago, i.e. the allowance of various levels of standardization, together with a co-ordinating instrument, viz. the Data Element Dictionary which will be governed by ISO. This will ensure a means of effective compatibility without serving as a uniform straight-jacket for all concerned. Especially, it will encourage information exchange from and to smaller scale systems, which have found the existing formats too complex and expensive for their needs. Hopefully, it will also provide the flexibility needed to adapt bibliographic systems to various types and levels of user needs. For, after all, it is the type of information needed by the user that should determine the contents of a bibliographic record (25).

So where will the future be? A new format will have its own problems. It will have to be accepted and it will have to prove itself. In the meantime work on bibliographic data exchange still goes on. I myself am now involved in a computer experiment with vaguely structured information, i.e. bibliographic information in which elements are not explicitly identified. This invloves having to instruct the computer to interpret bibliographic information along more or less the same lines as human beings do. Perhaps in future the computer will do more things than we can expect at this moment. Other means of establishing links between information services have been proposed, e.g. interfacing mechanisms based on authority control (26). This would be one way of bridging the gap between library (catalog-) systems and A&I (reference-) systems. An example of more specialized developments is the bilateral format for exchange of bibliographic data between the US and the USSR, drafted by the US-USSR Research Group on Development and Testing of a Common Communications Format for Bibliographic Data Exchange on Magnetic Tapes, within the framework of the US-USSR Working Group on Co-operation in the Field of Scientific and Technological Information (27).

In any case, world-wide exchange of bibliographic information is already well on its way. The future of bibliographic exchange formats will probably depend a lot on technological developments. For instance, we may expect exchange on tape to give way to on-line data exchange through communications networks such as EURONET. I believe this will have major consequences for future format design.

REFERENCES

(1) International Standards Organization. ISO 2709 Documentation - Format for bibliographic information interchange on magnetic tape. Geneva : ISO, 1973

(2) International Standards Organization. Draft Proposal ISO 2709-Rev. 1, Documentation - Format for bibliographic information interchange on magnetic tape. ISO/TC46/SC4 doc. N 124 (1978-09-15).

(3) Mackenzie Owen, J.S. Format incompatibility and the exchange of bibliographic information : a comparative study. Paris : UNESCO, 1976.

(4) Avram, H.D., J.F. Knapp & L.J. Rather. The MARC II Format : a communications format for bibliographic data. Washington, D.C. : Library of Congress, 1978.

(5) The interchange of bibliographic information in machine readable form : papers given at the European Seminar held at Banbury, 12-16 May, 1974 / edited by R.E. Coward and M. Yelland. London : The Library Association, 1975 (Library Association Research Publication ; 17).

(6) Deutsche Forschungsgemeinschaft, Bibliotheksausschuss. Maschinelles Austauschformat für Bibliotheken (MAB 1). Arbeitsstelle für Bibliothekstechnik, 1973.

(7) Kohl, E. Current developments in Germany and Austria. In: The interchange of bibliographic information inmachine readable form. London, 1975, pp. 24-32.

(8) Kohl, E. Areas of possible incompatibilities in the international exchange of bibliographic information in machine readable form based on a compatibility study for MAB 1 data. In: Towards a common bibliographic exchange format? Budapest, London : 1978, pp. 52-58.

(9) International Federation of Library Associations (IFLA). UNIMARC : universal MARC format. London : IFLA International Office for UBS, 1977.

(10) Rowlett, R.J., R.L. Eigington & J.L. Wood. Standard distribution format : Chemical Abstracts Service. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 105-111.

(11) Dragulev, V., E. Ruckenbauer & D. Marek. The INIS format : an existing operational exchange format. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 122-131.

(12) Linford, J. The EUDISED format : a structure which will contain any type of formalized document record. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 115-118.

(13) Organisation for Economic Cooperation and Development (OECD). Specification for magnetic tape transfer of bibliographic information in the IRRD scheme. 3rd edition. Paris, OECD, 1977.

(14) Mongar, P. The International Road Research Documentation System's experience and views in bibliographic exchange formats. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 112-114.

(15) Zentralstelle für maschinelle Dokumentation (ZMD). MADOK : Magnetband-Austauschformat für Dokumentationszwecke. Munich, Verlag Dokumentation, 1977.

(16) Hitzeroth, Ch. MADOK - the German magnetic tape exchange format in the field of documentation. In: Towards a common bibliographic exchange format? Budapest, London, 1977.

(17) Mezdunarodnyj kommunikativnyj format - MEKOF 2. Soderzanie i sposob predstavlenija dannyh v zapisi dlja obmena informaciej na magninoj lente (NTP MCNTI 19-77): Specifikacija elementov dannyh. Moskva, MCNTI, 1977.

(18) Skripkin, V. & V. Kodola. The communication format in the international system of scientific and technical information of the CMEA countries. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 162-171.

(19) Vajda, E. Some ideas about the compatibility problems of bibliographic data interchange formats, based on a comparative anlysis of the MEKOF-2 format. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 172-192.

(20) Reference manual for machine-readable bibliographic descriptions / prepared by the UNISIST/ICSU-AB Working Group on Bibliographic descriptions with the assistance of ICSU and ICSU-AB member services ; compiled by M.D. Martin. Paris, UNESCO, 1974.

(21) Dierickx, H. The UNISIST Reference Manual for Machine-Readable Bibliographic Descriptions : current and anticipated developments. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 147-161.

(22) Hopkinson, A. The use made of the UNISIST Reference Manual : a brief history. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 49-51.

(23) International Federation of Libraray Associations (IFLA). ISBD(G) : General International Standard Bibliographic Description : annotated text. London, IFLA International Office for UBC, 1977.

(24) Towards a common bibliographic exchange format? : International Symposium on Bibliographic Exchange Formats, Taormina, Sicily, 27-29 April, 1978, organized by the Unisist International Centre for Bibliographic Descriptions (UNIBID) in cooperation with ICSU-AB, IFLA and ISO and sposored by UNESCO : Proceedings / edited by H. Dierickx and A. Hopkinson. Budapest : OMKDK-TECHNOINFORM, London : UNIBID, 1978.

(25) Mackenzie Owen, J.S. The fundamentals behind information exchange : an approach to optimal solutions. In: Towards a common exchange format? Budapest, London, 1978, pp. 9-12.

(26) Malinconico, S.M. The co-ordination of bibliographic control. In: Towards a common bibliographic exchange format? Budapest, London, 1978, pp. 61-67.

(27) Gove, N.B., T.E. Hughes, G.D. Matiushin & N.V. Turtanov. A common communication format for exchange of bibliographic data between the USA and USSR. In: Towards a common bibliographic exchange format? Budapest, London, 1978.

(28) Council of Europe. EUDISED : standards, formats, character representation. Strasbourg, Documentation Centre for Education in Europe, 1973.

THE USE OF MULTILINGUAL REFERENCE TOOLS IN THE PRODUCTION AND TRANSFER OF TECHNICAL INFORMATION

by

Gerhard Rosenau

Institut fuer Flugmechanik
Deutsche Forschungs- und Versuchsanstalt fuer Luft- und Raumfahrt (DFVLR)
Postfach 3267
D-3300 BRAUNSCHWEIG
(Germany)

Summary

Technical information in most cases is at least partly expressed in written or spoken language. A person desiring to understand such information needs either to know the language used or to have the information translated into his or her own language. In both cases it will be necessary to use technical dictionaries or other multilingual reference tools.

The paper first gives a survey of the problems of expressing technical information in one language. The discussion of the additional problems encountered in producing the information in more than one or in a foreign language leads to the description of the different types of multilingual reference tools, of their advantages and of their deficiencies.

Based on the experience gained during the revision of the AGARD Multilingual Aeronautical Dictionary, criteria are developed for a future Multilingual Reference System. The criteria are presented for discussion.

1. Introduction

Success in almost all technical projects depends on three human capabilities:

- technical skill and knowledge,
- organizational talent,
- communication between individuals.

The first capability alone would be sufficient for a single technical achievement. The second comes into play as soon as the process becomes more complicated, and the third is necessary as soon as more than one individual takes part in the technical process.

The normal means of communication between individuals is language in its spoken or written form. Technical knowledge and organizational talent are not specifically national, but language is. Thus we are confronted with the phenomenon that technical projects would be restricted to a national scope, if it were not possible to overcome the language barrier.

The aim of this lecture is

- to review the tools we use to this purpose,
- to discuss some of the aspects of their use, their preparation and their future development.

2. Characteristics and problems of technical language

2.1. Characteristics of technical language

In the transfer of information by means of written or spoken language three characteristics play a decisive role:

- the set of words available in the language (vocabulary),
- the rules for the combination of words into phrases and for the transformation of words necessary in this process (grammar),
- the representation of words and phrases in the spoken or written language (pronunciation and spelling).

According to the different applications of these characteristics, classes may be distinguished within a language. For instance a certain dialect is characterized by local pronunciation. The use or avoidance of expressions is distinctive for the language of a social group, a sociolect.

How may the limits of technical language be determined within the language as a whole? It is easy to see that technical language has some essential characteristics in common with the general language. There is no reason why there should be a different pronunciation. Also the grammatical rules should be the same for technical as for non-technical texts. As a distinguishing characteristic then there remains only the vocabulary.

For a technical language it is required to designate the many technical concepts by as far as possible unequivocal terms. That leads automatically to a rich technical vocabulary and to differences between technical and general language. On the other side remain narrow links between these two language classes. The majority of the terms appearing in a technical text belong to the general language, even for very special subjects.

The technical terms in fact are symbols invented by technicians or scientists or adopted from other special vocabularies. These symbols shall facilitate the exact expression of information. Basically every technical term is a symbol "ad hoc" and its meaning cannot safely be guessed from the meaning which an equally pronounced or spelt term would have in a non-technical context. In order to understand a technical term it is, therefore, necessary to have a sufficient background of knowledge. This knowledge, differing according to depth and scope in the various general and technical domains, leads to a classification of terminologies, as it is shown in figure 1. This classification is hierarchical as the more specialized domains in the lower parts of the figure are imbedded into the less specialized domains higher up. Because of this fact, the number of persons acquainted with the vocabularies of the different classes depends on the degree of specialization of the corresponding technical domain. The majority of speakers of a language - if not all of them - will be able to understand the general language. Only a minority of persons, however, will know the terminology of a special technical domain.

Available Reference Tools	Terminology classes	Examples of terms	
very many	General terminology	*wind*	*air*
many	General technical terminology	*wind speed*	*air density*
few	Aeronautical terminology	*wind sack*	*airspeed*
	Terminology of a narrow subject field	*down-wind leg*	*air scoop*

Figure 1: Hierarchical terminology classification

The examples in figure 1, derived from the words "air" and "wind", illustrate the different classes. The two basic terms belong to the general language but the terms in the last line are used in special contexts only: "down-wind leg" is a term of piloting and "air-scoop" belongs to the now obsolescent domain of lighter-than-air air vehicles.

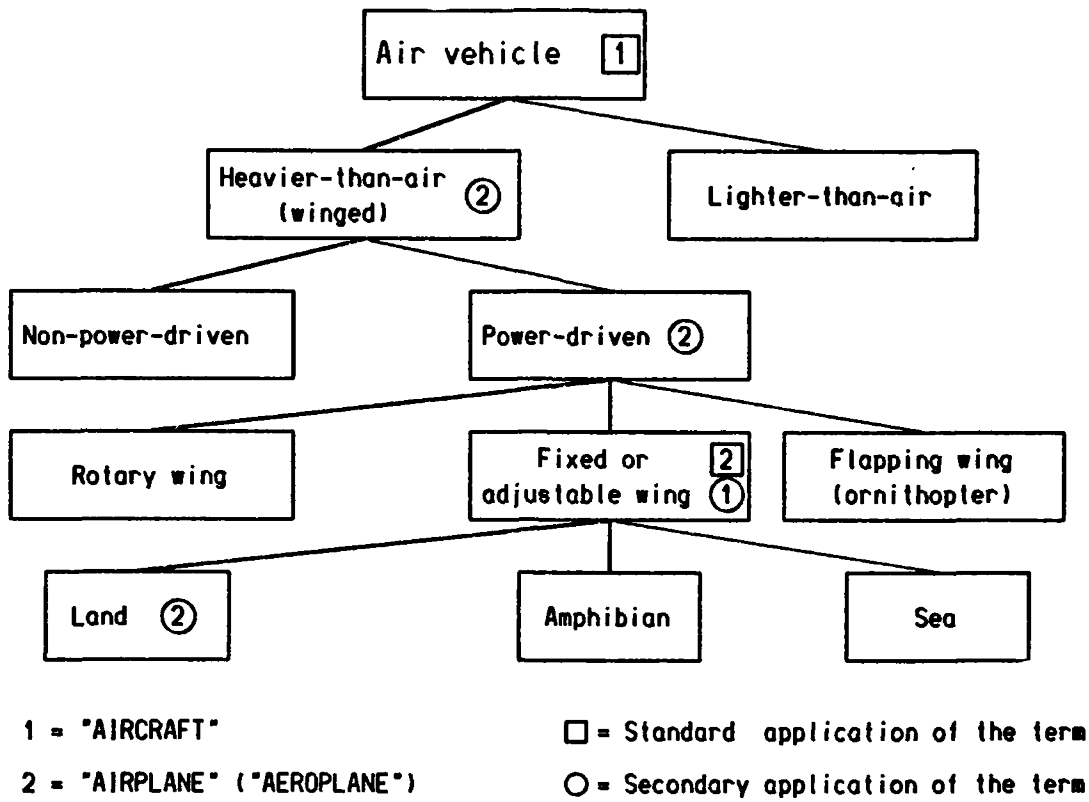


Figure 2: Application of the terms "Aircraft" and "Aeroplane" to different classes of air vehicles

2.2. Problems of technical language

The notion of an obsolescent technical domain leads to another important problem: the variation of the vocabulary with time. Technical progress always introduces new concepts and creates the need for new corresponding terms. On the other side, the same progress makes older concepts and terms obsolete. Sometimes such obsolete terms gain a new meaning and come back to the vocabulary, perhaps in a different technical field. This constant change of the vocabulary is the cause of the never-ending demand for up-dating of the reference tools.

No general rules can be given for the formation of new technical terms and, consequently, also not for the composition of a technical vocabulary at a given time. Certainly many of the special terms originally came from the general language, many are adopted or adapted from neighbouring technical or scientific domains. Let us consider terms used in the description of an aeroplane: "wing" is the a term originally used for a similar part of the body of flying animals, "rudder" designates a part of a ship. The term "elevon", however, is a special artificial word, formed by contraction of "elevator" and "aileron", names of other components of an aeroplane.

For most of the important languages there are monolingual reference tools, which may be used to obtain information on the vocabularies of the general or technical language. Number and scope of these reference tools depend highly on the degree of technical specialization involved. While one can find very many reference tools covering the general language, there exist only a few for relatively large technical subject fields as e.g. aeronautics. For the narrow special fields generally no reference tools at all are available.

Using the example of monolingual reference tools, some problems may be discussed, that exist also and to a higher degree for multilingual reference tools. As has been mentioned before, the basic requirement is that the information transferred by means of the language must be corrrectly understandable. To achieve this; the information must be unequivocal, and this should be true for all the technical terms used. The examples of Figure 2 show that this is not the case, even for very important and basic terms, as e.g. those used for the different classes of air vehicles. The classification scheme for the concepts is taken from the "ICAO Lexicon" (/6/). From the same source also come the standard uses of the terms "aircraft" for an air vehicle in general and "aeroplane (airplane)" for a power-driven, heavier-than-air air vehicle with fixed or adjustable wings. Besides these standard uses exist secondary uses listed in the "United States Air Force Dictionary" (/7/). Behind the externally identical terms "aircraft" hide at least four homonyms with different meanings. For the two terms considered in figure 2 the differences between the respective homonyms are only generic ones, i.e. one of the homonyms covers the meaning of all the other ones represented by the same term.

This is no longer the case for the examples of Figure 3. Here we have to distinguish between "homophones", terms pronounced identically, and "homographs",terms spelt identically. The different meanings of the homographs ("distance or time ahead" / "wire" / "metal") are no longer related to each other.

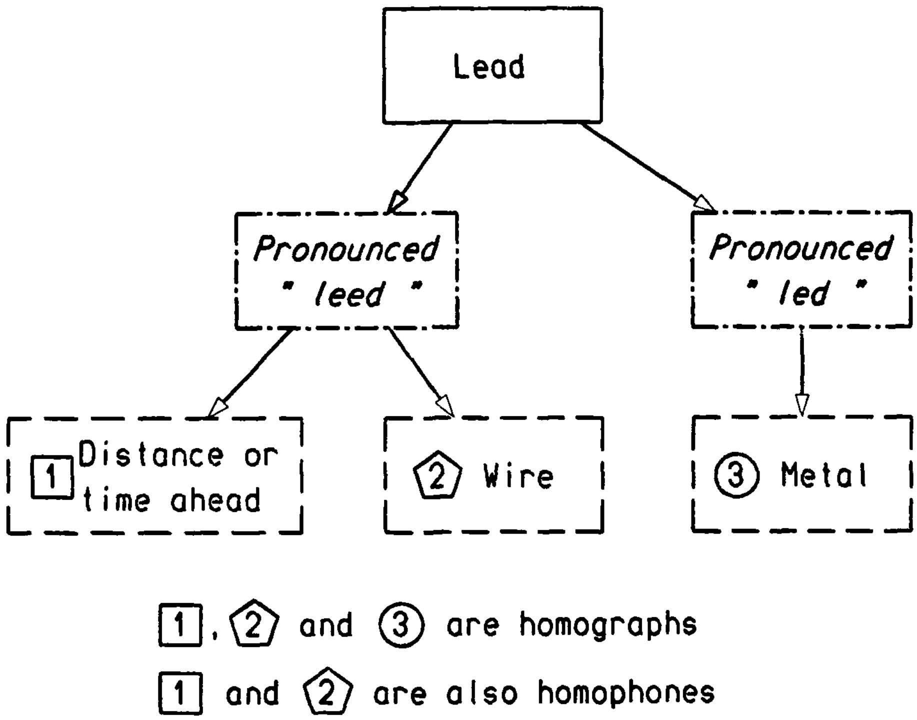


Figure 3: Homographs and homophones

Different meanings of the noun "lead" in technical contexts

Special care is required even in monolingual reference tools to distinguish between homonyms. Most of the reference tools list the entries alphabetically and, therefore, often do not make a difference between homographs. If as usual in a technical reference tool the pronunciation is not indicated, this characteristic cannot be used for distinguishing purposes. Thus confusion is created very easily which than propagates into the multilingual reference tools (cfr. chapter 3.2).

Another important point for reference tools is the synonym problem. By "synonyms" we understand terms different in pronunciation and spelling, but having an identical or almost identical meaning. One example for synonyms is the use of "aeroplane" as well as of "aircraft" for the powered heavier-than-air air vehicle with fixed or adjustable wings. At first sight the quasi-identity of meanings of synonyms is predominant, and under this viewpoint they often are listed in the reference tools. In reality, however, the special importance of synonyms lies in the different shades of their meanings. The terms "large", "big", and "great" are often interchangeable in meaning of more than usual size, extent, etc. (a large, big or great tree), but in other cases they imply or exclude quite different characteristics. So it is correct to say that Professor von Karman was a great man, though he certainly was not a big man. The differences in meaning of synonyms must be indicated in the reference tools, if possible by comparing these differences.

Closely connected to the homonym and synonym problems are the facts represented in Figure 4. The three terms "dihedral", "anhedral", and "cathedral" have to do with the angular position of the two half-wings of an aeroplane with respect to its longitudinal plane of symmetry. It is quite generally accepted to call a dihedral positive, if the half-wings are inclined upwards, but there may be a "negative dihedral" for downwards pointing wing-tips. The use of the two other terms is based on two different distinctions:

1. The triplet "anhedral - dihedral - cathedral", where "anhedral" means upward and "cathedral" means downward inclination.
   This is probably the correct use of the Greek prefixes "ana-" and "kata-".

2. The pair "anhedral - dihedral", where "anhedral" designates exactly the opposite characteristic as before, viz. the downward inclination.
   The prefix "an" in the sense "no" or "without" does not make any sense here.

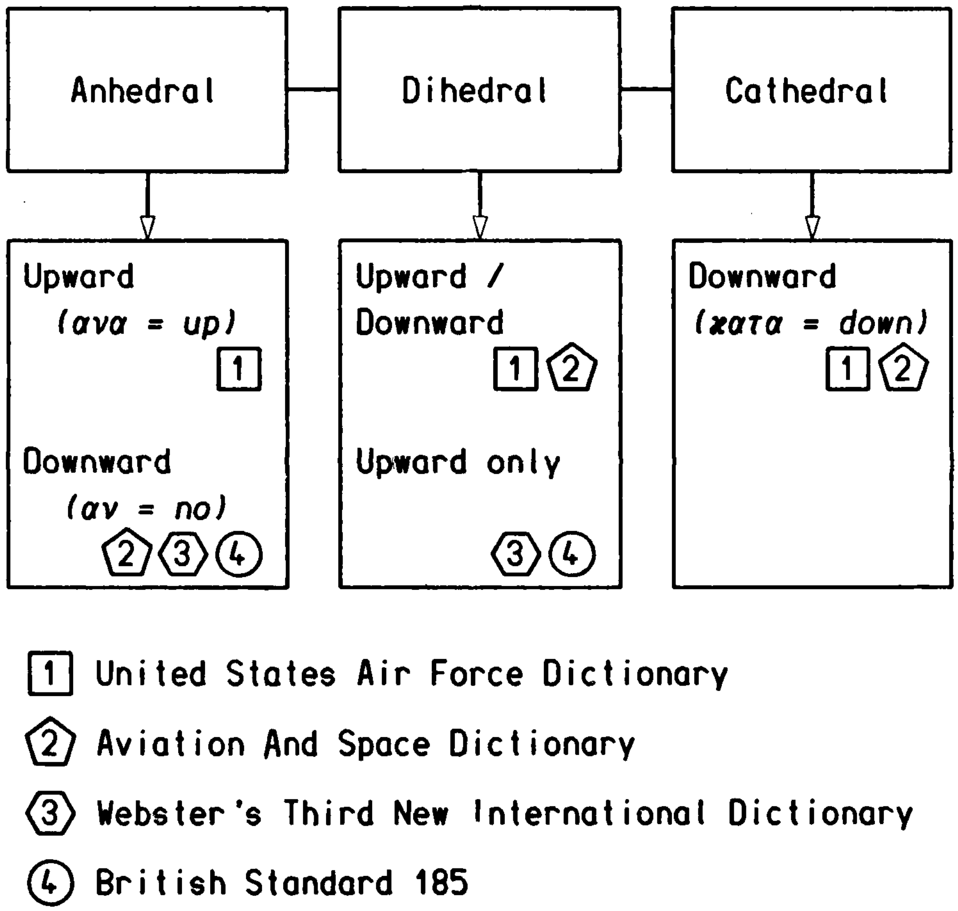


Figure 4: Different uses of "dihedral", "anhedral", and "cathedral" in dictionaries

From etymological and notational errors result here uses of terms, which should be classified as wrong or incorrect, but must be accepted as correct, as they are widespread and found in otherwise correct texts. The reference tools in such cases cannot avoid listing all the different and contradicting meanings of the terms.

Only a standardized terminology could eliminate these difficulties. In most cases, however, the accepted use of the terms is very deply rooted and would resist to any attempts of a contradicting standardization.

The last observations show that it is necessary to characterize the various uses in the reference tool. Examples of such characterizations would be "standard use", "non-standard use" or even "wrong use". Other indications could be added to indicate the language levels, e.g. if a term belongs to normal language ("aeroplane") or to some jargon ("bird").

3. Production and transfer of technical information in more than one language

3.1. Overview over the tasks

Not only there are distinct domains within one language, but there exist many different languages, separated from each other by language barriers. Certainly there are persons mastering more than one language, but for most persons and most cases the language barriers represent considerable obstacles. These obstacles must be surmounted if an information is to be produced in more than one language or to be transferred from one language to another. This may be done in two different ways:

1. actively, if the information is newly formulated,
2. passively, if the information is only received and understood.

In both cases the understanding of the information is the final goal. While by the active process the information is made available in the desired language, in the second case the person receiving the information has to overcome the language barrier himself or herself.

Figure 5 shows the different possible processes in the production of information in a foreign language or more than one language. Such production processes naturally are active ones. There is, first, the production of spoken information. Examples are talks or discussions but also the issuing of military commands or the transmission of flight safety information by flight controllers. The requirements on the precision of the information representation in spoken language may vary widely according to the situation. In informal talks and discussions high precision is generally not required, as the speaker may correct his deficiencies, if necessary with the help of the listeners. Military commands, however, and flight information must be correctly understood immediately. These, therefore, must be formulated very precisely and contain normally but few information and appear in only a few standard forms.

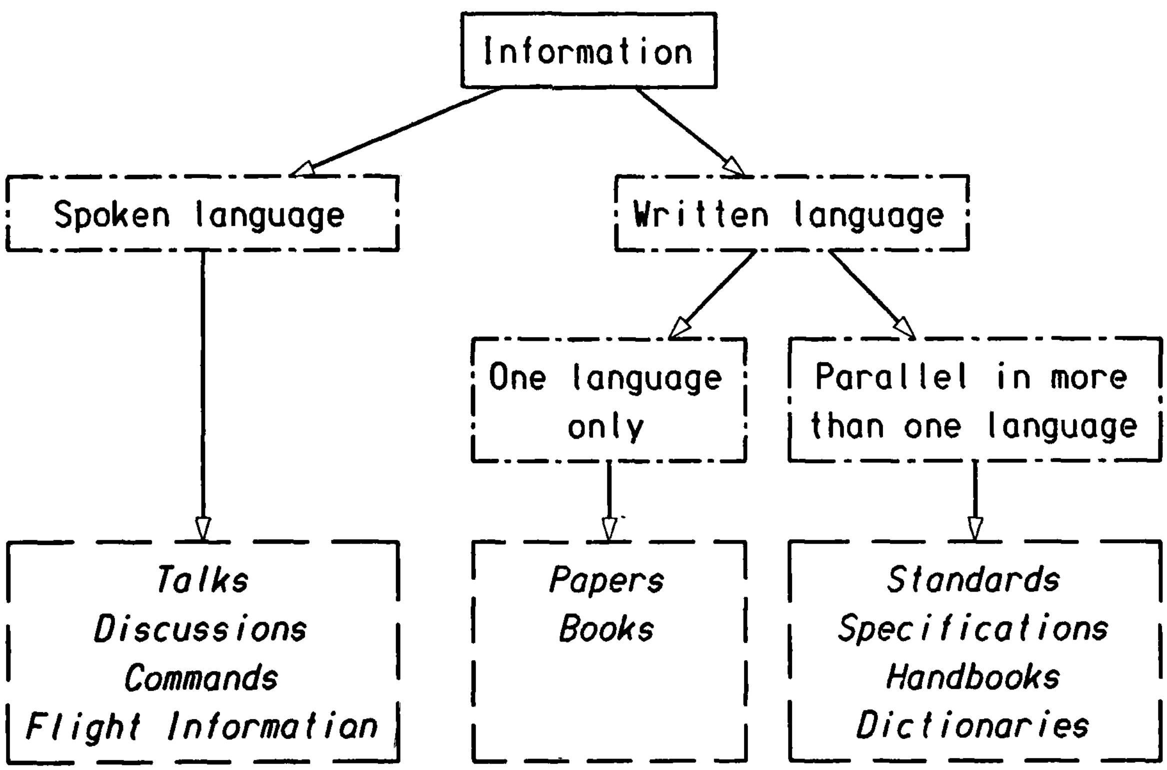


Figure 5: Production of information in a foreign or in more than one language

Whereas the production of spoken information is an ad-hoc-process and may be repeated if necessary, the written language supplies the means for a permanent representation of the information. In this case there is normally a distance in space and time between the producer and the receiver of the information and no feed-back between them is possible, which could help to avoid or to correct errors.

The most common process here is the production in only one but foreign language. Often e.g. it is not efficient to write a paper or a book in a less known language. At all times there were dominating languages, which an author would have to use, if he wanted to be read: Greek in the antiquity, Latin during medieval times, French in the eighteenth century. To-date English is the dominating language in our technical litterature, and a book on an aeronautical subject must be written in this language, if the majority of the aeronautical specialists shall take notice of it.

The difficulties increase if a parallel text with identical content is to be produced in more than one language. Examples are the elaboration of standards, specifications and multilingual handbooks. The standards of the International Organization for Standardization (ISO) are published in the three languages English, French, and Russian. The work on such standards in international committees requires considerable linguistic competence and the greatest care. Even than the difficulties sometimes are insurmountable. The German standardization institution (DIN) renounced to a description of programming languages (e.g. FORTRAN and ALGOL) in German as it could not be guaranteed that a German text would exactly and everywhere have the same meaning as the English text.

As compared to the translation of texts there is often some advantage in the parallel production of texts in different languages, as the versions of the texts in all languages can be adapted if difficulties arise for some language.

A special form of parallel production of information in more than one language exists for dictionaries. This appears explicitly if definitions are given in several languages, as in the first edition of the AGARD Aeronautical Multilingual Dictionary (/1/). If no definitions or definitions only in one language are given, the task of furnishing correct equivalents in different languages remains a problem of information production as it is necessary to detect the concepts behind the terms in the original language and then to search for the correct equivalents in the other languages.

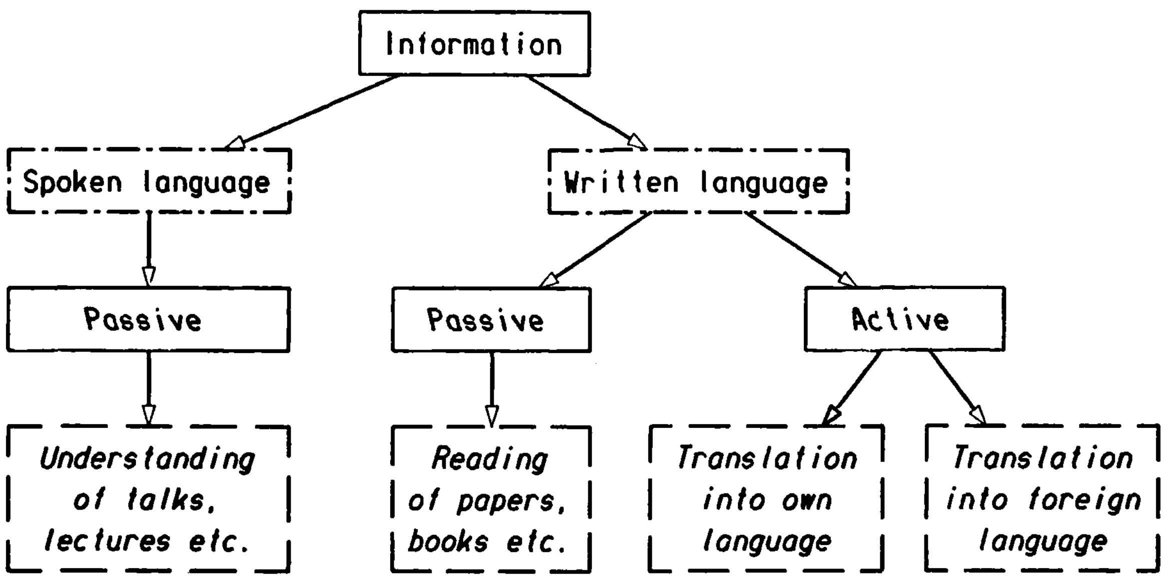


Figure 6: Transfer of information from one language to another

An information not produced immediately in the native language of the information receiver needs to be transferred into that language. This can be achieved by a passive process if the receiving person understands the original language well enough. This is the normal case for spoken language and frequent for written language (Figure 6). The receiving person generally translates internally and only parts of the information or nothing at all. In the active transfer, the whole information is translated externally from one language to another. The translated text should represent the original as exactly as possible. In the technical domain only the content and not the style will be important, but the danger of substantial errors is always present.

3.2. Types of Multilingual Reference Tools

For all work, where more than one language is involved, a certain personal linguistic competence is necessary.

This competence may be assisted by a number of multilingual aids for surmounting the language barrier. We shall call such aids "Multilingual Reference Tools". The best-known of these tools are dictionaries, but there are also others which are important for the technical language (Figure 7).

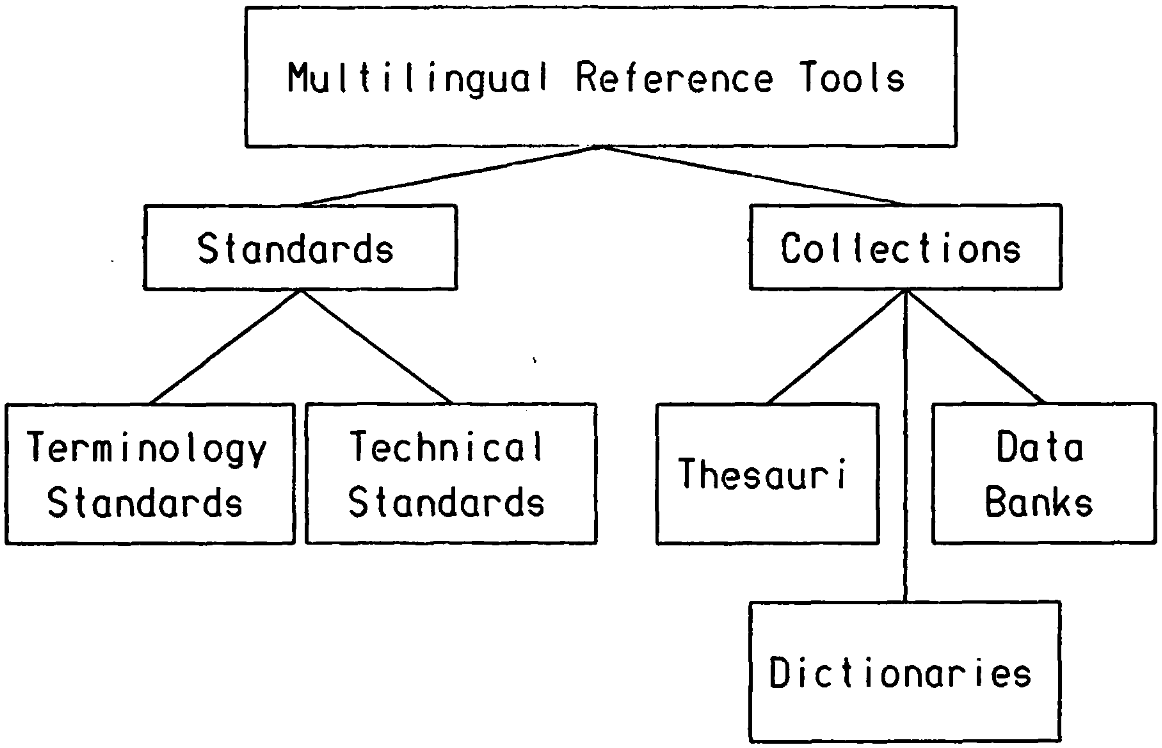


Figure 7: Types of Multilingual Reference Tools

The reference Tools may pursue different basic intentions:

- definition of a standardized terminology in different languages
- collection and presentation of existing terminologies.

The consequences of the basic intentions are important. The standardization of terminologies supposes the existence of a corresponding system of concepts. This system must be free of contradictions and sufficiently complete. Only such a system can guarantee that a certain term can be attributed to a defined concept, and will then have a well-defined meaning. In a standardized terminology the definition leaves in general no room for the many shades of meanings that the normal use of language requires. Because of this, standardized terminologies are restricted to very limited fields and difficult to produce. Standardized terminologies are not as useful as they may seem to be, as their use can only be enforced by additional external conditions, as e.g. legislation.

The reference tools which in figure 7 are called "collections" cover a much greater scope of uses. They accept the termionologies as these exist at a given time. If definitions are contained in this type of reference rools, they describe the underlying concept a posteriori in order to assure the correct correspondence between concept and term or in order to facilitate the understanding of the meaning.

The first intention mentioned above, standardization of concepts and terms, is mainly pursued by standards. Terminology standards serve this purpose only. Examples for this in the field of aeronautics are the ISO flight dynamics standards ISO 1151 I/VI (/8/) and the definition part of the ICAO Lexicon (/6/). Efforts within the ISO to produce a multilingual aeronautical dictionary, the contents of which would be of normative character, seem to be successful only for those subjects for which there are already internationally accepted standardized definitions and the corresponding national standards.

Technical standards intend to fix a restricted set of characteristics for certain technical objects. They may be used as reference tools in our sense if they exist in different languages and if the terms in question fall into the scope of the standard.

Under the heading of collections the first group is that of thesauri. They are similar to terminology standards, as they intend to cover the set of concepts relevant to a certain subject field as completely as possible by a set of conveniently chosen terms. Thesauri generally serve documentation and retrieval puposes, but are also valuable reference tools if they are multilingual. This is the case for the international editions of Universal Decimal Classification (UDC, cfr. /9/), which are published parallelly in a great number of languages. In the UDC the subject fields are arranged in a hierarchical form starting with the most general and going to the special. Definitions of the concepts corresponding to the classes are not given, the meaning of the class names may, however, been deduced from the position of the concept within the hierarchical class system.

Until to-date dictionaries are the most important reference tools. They exist in a great variety, and more and more new ones are added to their number. As shown in figure 1, the choice decreases, if one comes to more special fields. Mr. Stork in his rationale for the revision of the AGARD Aeronautical Multilingual Dictionary (/1/) listed 11 different aeronautical or astronautical dictionaries published between 1965 and 1971. All of them were at least bilingual, one contained terms in seven languages. Since than some more multilingual aeronautical dictionaries have been published.

The existing dictionaries differ considerably according to their purpose but also according to their scope and internal quality. Some of them contain definitions, others present only equivalence lists of terms. In some dictionaries very many equivalent terms are listed in an average entry, but they are not systematically arranged within the entry, homonyms are not marked and no indication concerning the differences in the meanings of symonyms is given. Such dictionaries reseable more a storehouse than a true reference tool.

The last group of reference tools in figure 7 are the data banks. To-date they are not yet available for public use. They consist of information stored in and accessible by computers. The access methods they offer exceed by far those for dictionaries, as e.g. searching of terms in the context of definitions or sorting according to subject categories. In principle they can be kept continiously up-to-date. Their content, however, as far as I know, does not go far beyond that of dictionaries. Some of their other characteristics do not compare favorably with those of dictionaries. They generally dispose of a very restricted set of symbols (in some cases only upper case Latin letters and numbers), which creates restrictions in the representation of non-English words. The special conventions for computer storage and output ("linear" organization not permitting the usual sub- and superscripts etc.) do not allow the normal representation of mathematical formulae. Due to the very short time of their existence, data banks are certainly still in their first stages of development and use as reference tools.

3.3. Use of Multilingual Reference Tools

The use of dictionaries should be treated in every teaching of languages. There most often the general language is concerned and the use of dictionaries generally does not present great difficulties. This may be completely different for technical language. Some aspects of the use of dictionaries and other reference tools need therefore discussion.

Probably everybody with a sufficient linguistic and technical competence who has used dictionaries or other reference tools will be acquainted with a dilemma: one gets the impression that they contain only what one knows already or does not need, but the subjects not known but needed cannot been found. Certainly this dilemma is only partly true. If it is true it is often caused by the rapid changes in the technical vocabulary and the unavoidable lag of the work on dictionaries and of their publication with respect to those changes.

Often the discontent with the reference tools is related to the linguistic and technical competence of the user. Figures 8 and 9 illustrate this relationship. They compare the benefits from dictionary use with its safety. In both cases linguistic and technical competence have a combined effect. If the two are great the benefits (figure 8) are good. Already if only one of the two competences decreases, the benefits diminish rapidly until, eventually, it is useless to consult the reference tool, as either the adequate technical or linguistic competence is missing. The safety of reference tool use varies accordingly. It ranges from "safe" for the highly competent user to "disastrous" (i.e. untolerable technical or linguistic errors would result) in the case of insufficient competence of the user.

The two diagrams have been drawn almost identically in order to put emphasis on an important point: for the majority of users reference tools are useful or necessary. At the same time their use may be precaricus or even dangerous.

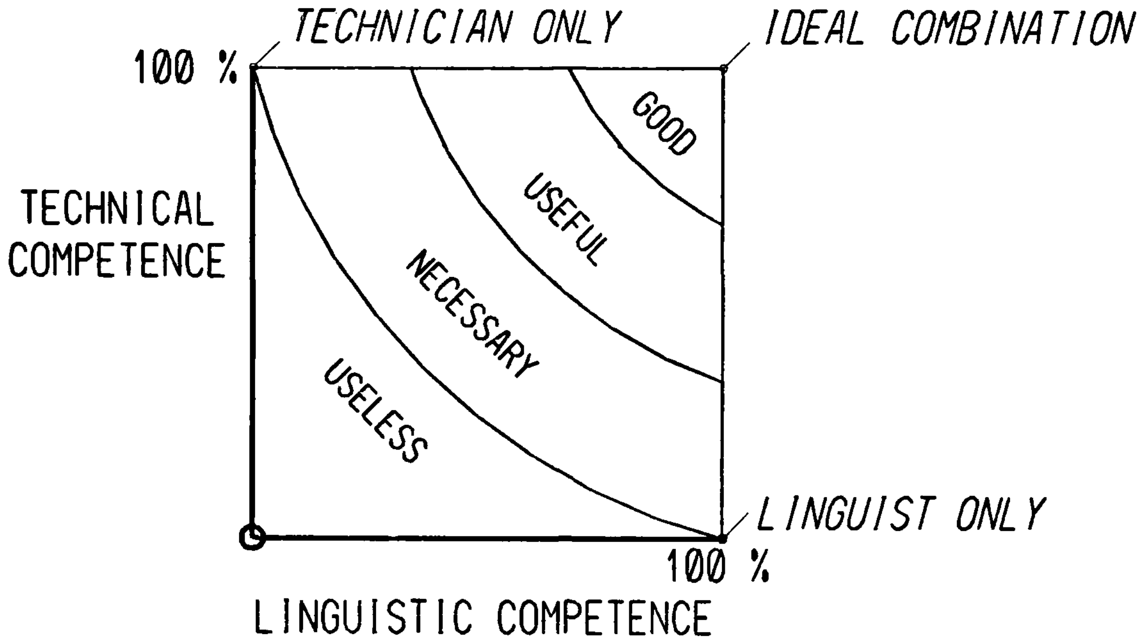


Figure 8: Benefits from dictionary use

What is the reason for this apparent discrepancy? In the discussion of technical language we encountered the homonym and the synonym problems. These problems are aggravated if more than one language plays a role. In every language homonyms relate to different concepts and synonyms have different shades of meanings.

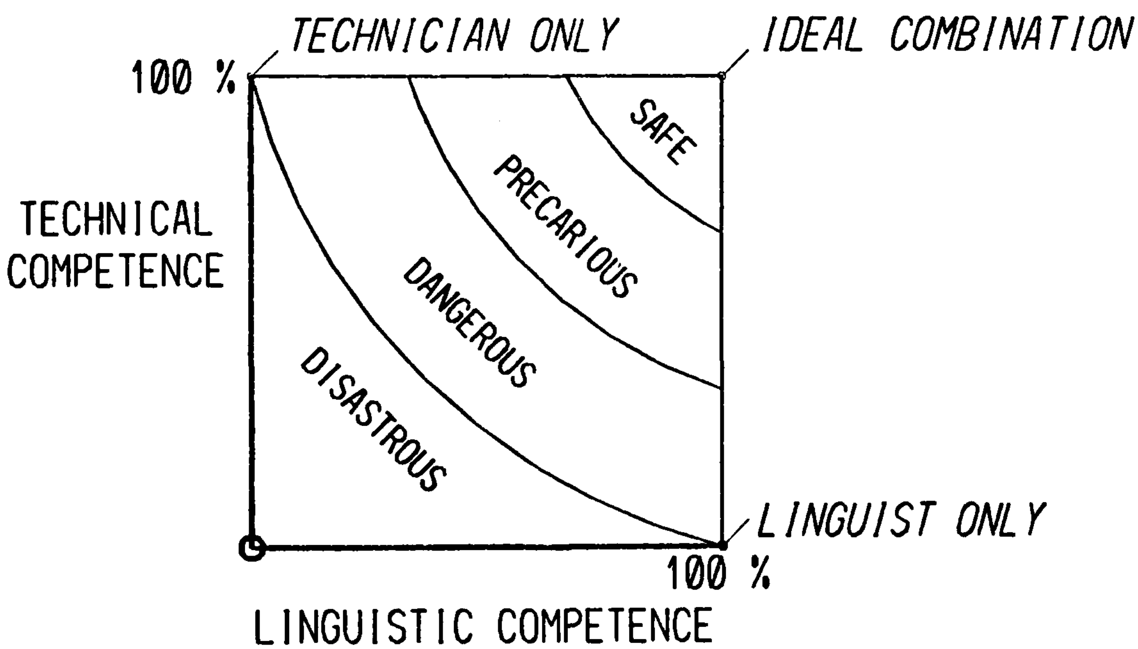


Figure 9: Safety of dictionary use

In figure 10 this situation is compared for the pair of English terms "altitude" / "height" and their French and German equivalents. To the terms correspond four different concepts wich share the basic idea of "extension in a vertical direction". In English "altitude" and "heigt" are synonyms for concept 3, the same is the case for the two French terms. Concept 4 usually is designated by "altitude" in English and by "hauteur" in French. There are parallel relations between the concepts 1 and 2 and the corresponding English and French terms. In German all the four terms may be designated by "Höhe". If therefore this German word with the meaning 4 shall be parallelly translated into English and French, the basically different terms "altitude" (English) and "hauteur" (French) must been used. In the sense 2 the term "altitude" shall be used which is an English-French homograph. Only few reference tools distinguish exactly between the four different meanings, such that the incompetent user is in the danger of using the incorrect terms.

The above discussion shows another problem. Languages which have a long history in common posess in their vocabularies many words that are homographs or almost homographs of terms of the other language. This is typical for the relationship between French and English but it exists also e.g. for Dutch and German. These bilingual homographs are the dreaded "false friends", which easily mislead if no sufficient care is taken. This care should be extended to the use of reference tools, as even these may not be precise enough.

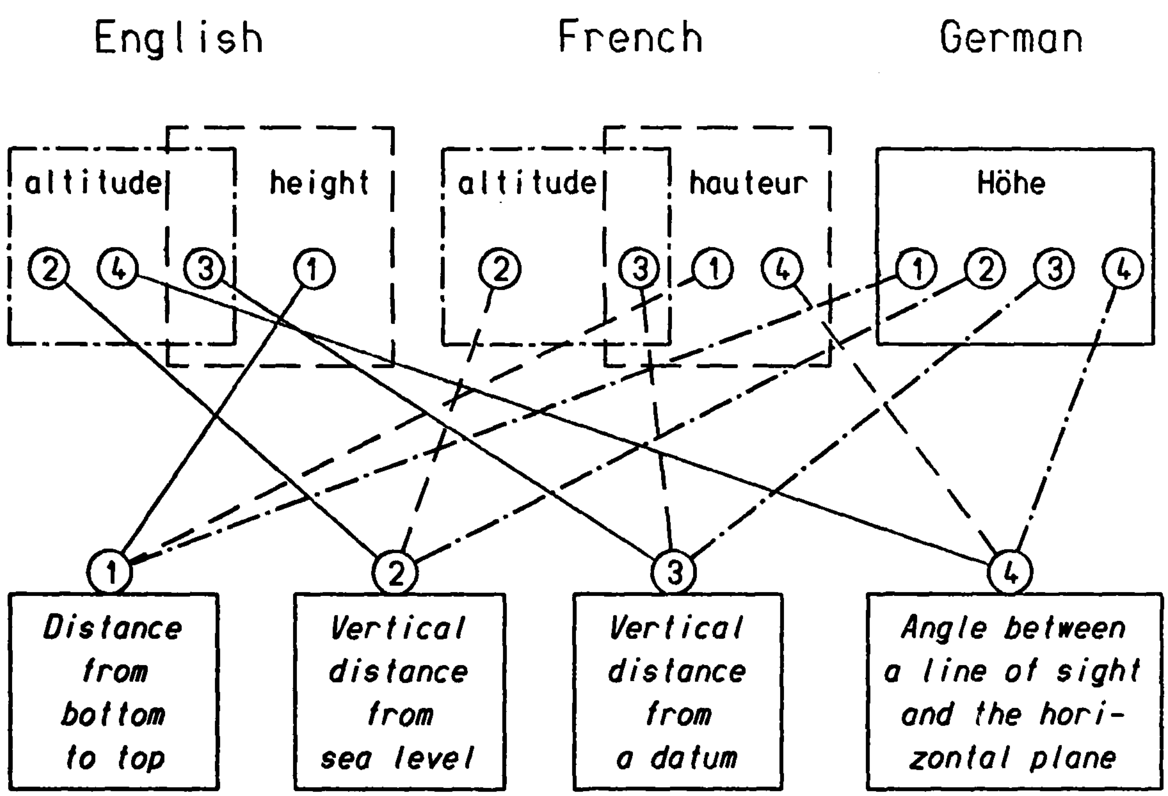


Figure 10: Synonyms and homonyms in different languages

A similar situation exists for the term "lead" (figure 3). Two of the most important English-German technical dictionaries do not distinguish between neither the homographs nor the homophones, but arrange the equivalent German terms indiscriminately, which will cause confusion even with a competent user.

In the example of figure 2 ("aeroplane" / "aircraft") it can happen that under the influence of insufficient reference tools the general concept "air vehicle heavier-than-air" becomes the restricted concept "landplane".

For the angular positions considered in figure 4 the signs of the quantities may be inverted by wrong translations from one language to another and thus conduct to completely false consequences for flight dynamics.

In using multilingual reference tools the greatest care, therefore, is necessary. In practice their use depends on the situation (figures 5 and 6). If spoken language is used, normally no reference tools are used, except if the information first was prepared in a written form. Passive processes often do not require the use of reference tools, as the meaning of unknown terms can often been deduced from their context. If in a passive process reference tools are used, the list of the various equivalent terms in an entry will permit to find the correct native language term or at least to understand the foreign term correctly. The situation in the case of translation into the native language is similar. Again reference tools are at the same time most necessary and most precarious to use for information representation in a foreign language, as the reduced linguistic and technical competences impair the judgement in the necessary selection of terms.

4. The AGARD Multilingual Aeronautical Dictionary and a future Multilingual Reference System

4.1. The AGARD Multilingual Aeronautical Dictionary

AGARD is called to participate in the elaboration of multilingual aeronautical reference tools as it is a organization with two official languages and a considerably greater number of national languages, and its tasks of research and development require clear concepts and terms.

A first Aeronautical Multilingual Dictionary was published by AGARD with a supplement in the years 1960 and 1963. Its first part (/1/) was based on the than existing version of British Standard 185 (/4/). The supplement used the NASA Dictionary as a source.

In the beginning of the seventies it became evident that a revision of the Dictionary was required. The revised edition should retain the multilingual properties of the first edition, as there scarcely were or are other suitable multilingual dictionaries. As before, definitions should guarantee the correct understanding of the concepts and the corresponding terms. With respect to its predecessor the revised Multilingual Aeronautical Dictionary (MAD, /2/) should be expanded by the introduction of new subject categories and new terms.

Work on the revision stated in 1973 and is now almost completed. A Working Group (now a sub-committee) of the AGARD Technical Information Panel supervised and coordinated the work but has also actively participated in many of the individual tasks.

The working programme consisted of the following steps:

1 Adoption of a basic English language dictionary with terms and definitions.
2 Incorporation of additional subject categories and introduction of further terms and definitions.
3 Establishment of a first draft to be submitted to the different AGARD Panels and to specialists for revision and for making proposals for additions and deletions.
4 Incorporation of amendments into the final English language version.
5 Production of national equivalents of the English terms.
6 Processing of the non-English inputs.
7 Proof-reading of the complete dictionary.
8 Establishment of three-language lists of equivalent terms with the different non-English languages as first language and English and French as second and third languages.

The programme corresponds roughly to the fundamental diagram of the development cycle for a multilingual dictionary. (Figure 11).

This programme was followed essentially, but it became evident in the later stages, that the collatoration of technical/linguistic editors was required. The editors had to revise and unify the results elaborated by a multitude of persons, mostly technical specialists with little or no experience in the dictionary field and disposing of quite different technical and linguistic competence.

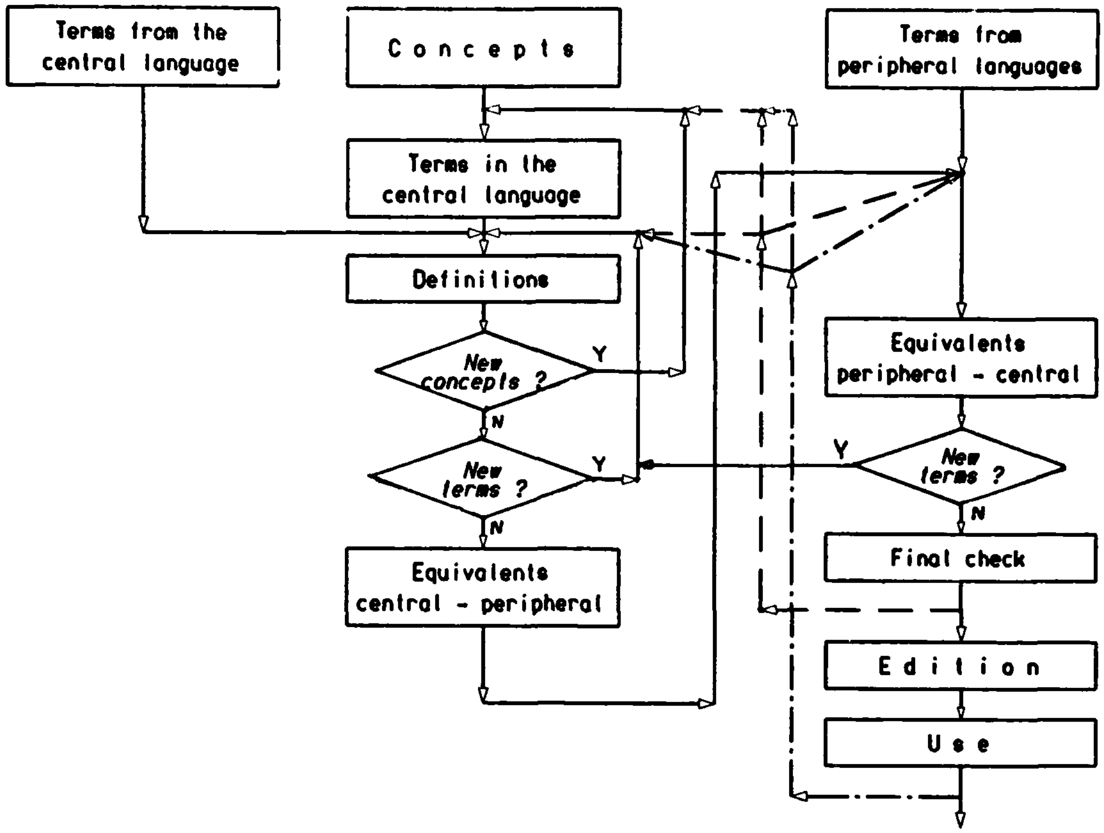


Figure 11: Multilingual dictionary development cycle

In the first step of the work some important decisions were taken. The MAD would have a "central organization", i.e. one language only would be used for the definitions. As for the supplement to the first edition, English was chosen as this "central language". Equivalent terms would be given in nine other languages, mainly languages of NATO countries plus Russian and Spanish (Figure 12). An alternative to this type of organization would have been the original polygonal organization of the first edition (Figure 13). In this case it would have been necessary to give rigourously equivalent definitions in all languages. Than all languages would have been of equivalent weight, but the size of the dictionary and the amount and cost of work would have become excessive.

As a source for the basic English dictionary a revised edition of British Standard 185 (/5/) was adopted. Additional sources for this basic dictionary were the ICAO Lexicon (/6/) and the US Air Force terminology.

Proposals for the content of the new subject fields (e.g. avionics, reliability) came mostly from technical specialists. Special military subject fields were not included.

In step 3 of the working programme the AGARD Panels were required to review the subject categories of the basic English dictionary related to their activities. The results of this review were not always satisfying. Such reviewing requires the assistance of a technical editor.

In step 5 national specialists were asked to provide terms in their languages corresponding to the concepts defined. Here again the competence of the collaborators was rather different. It proved difficult to avoid a mere translation of the respective English terms into the other language, and so the central organization of the dictionary may introduce a direction dependence in the equivalences. The national terms than would be correct translations of the English term but with an extended or restricted meaning as compared to the corresponding definition. Therefore it will be necessary in cases of difficulties with homonyms and synonyms to recur to the definitions.

The equivalence lists of step 7 will permit to get from one non-English language to another. It must, however, always be remembered, that the equivalences are valid only for the terms used in the sense of the definitions. So when going from an Italian to a Greek term, one will pass through the definition and perhaps implicitly through the English term.

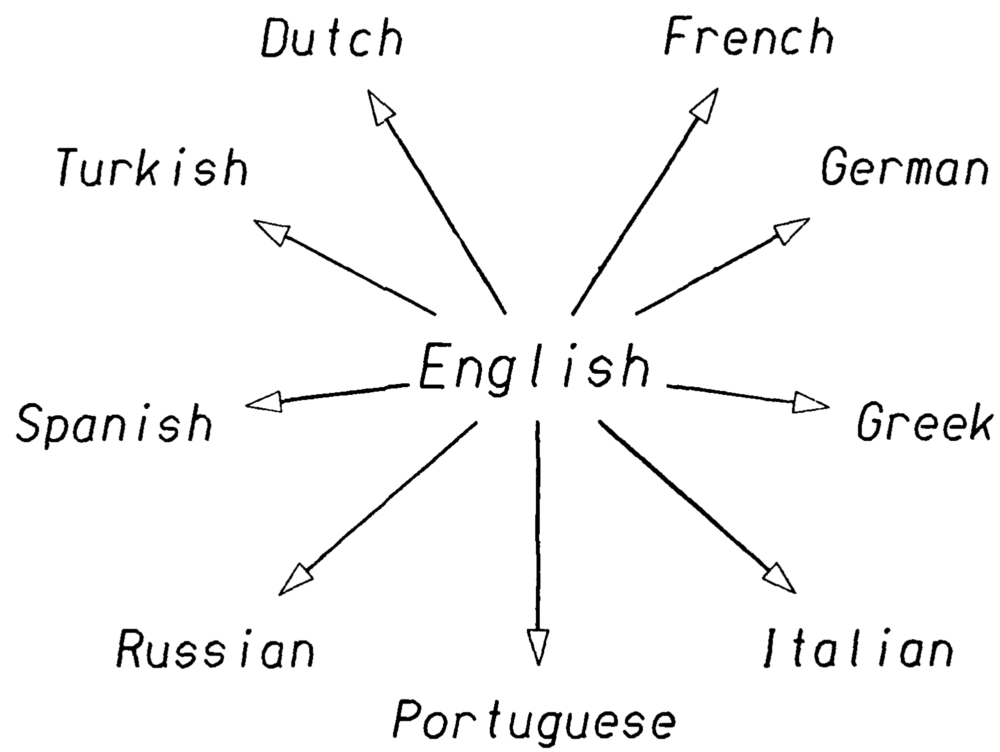


Figure 12: Central dictionary organization
with English as the central language

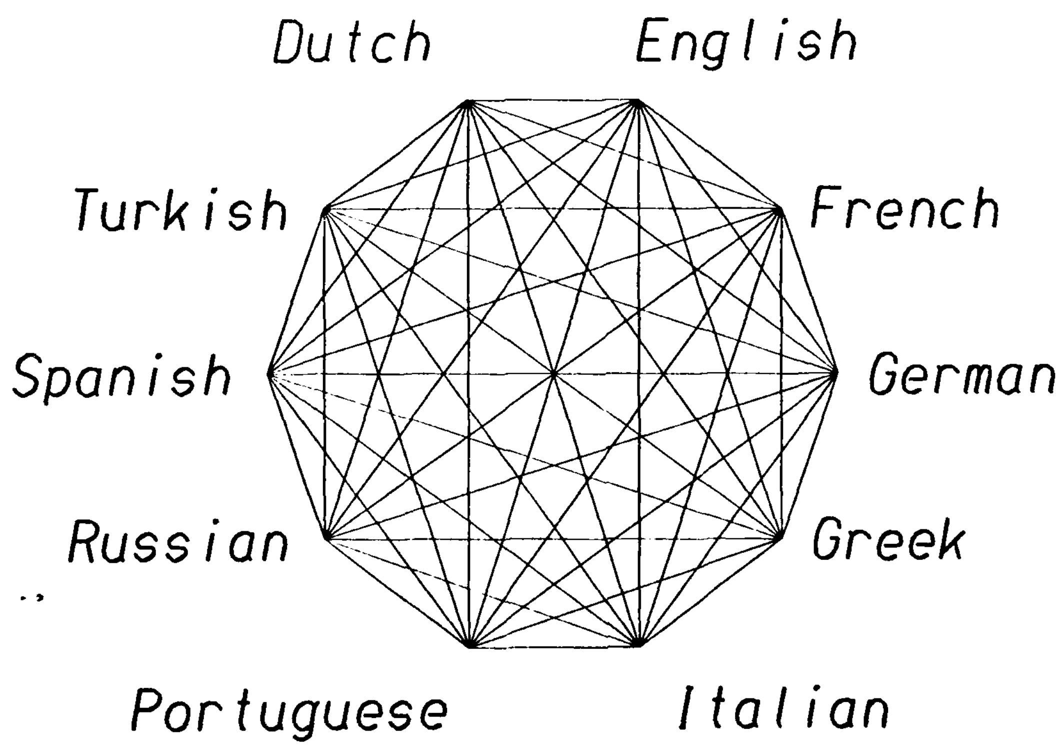


Figure 13: Polygonal dictionary organization
All languages are equivalent

4.2. Requirements for a future aeronautical Multilingual Reference System

It may seem premature to consider a revision of a dictionary which has not even been completed and published. But the experience gained during the work on the still incomplete edition, the rapid technical development and the possibilities provided by the available computers suggest such a revision. One decisive reason to start the revision as soon as possible is the long time required for such work.

The revision should at least be a step towards a Mutilingual Reference System (MRS), combining different types of reference tools. In order to define the direction of that step, we will try to outline the basic characteristics of such an MRS.

The revised MAD will contain about 8000 entries. Not all of them contain definitions as some of the entries are only cross-references to defined concepts (though they are not necessarily synonyms of the corresponding terms).

Even with the numerous newly added subject categories, the MAD will only partly cover the domain of aeronautics. Consequently aditional subject categories should be included, e.g. flight safety, flight testing, aircraft and ground support equipment, and air-defence.

The content of the existing subject categories is not yet sufficient. This is due to the fact that the primary source was a terminology standard with its unavoidable limitations. Two of the German specialists who took part in the work of providing German terms estimated independently, that their special field of interest was only covered to about one third by the terms and definitions now available.

It will be necessary to replace most of the cross-references now existing by definitions indicating the different meanings of synonyms. This may introduce definitions covering the different meanings of non-English synonyms.

The new subject categories and definitions would add considerably to the number of concepts in the MRS. The total number of these concepts may reach or even exceed 20000. It should be possible to understand and translate any aeronautical text using only the MRS and a dictionary of general language. This requirement implies, that all terms used in the definitions that cannot be found in a good dictionary of general language, must be contained in the MRS itself.

	PART 1	PART 2
Set of concepts	[A]	[A] + (B)
Definitions	Yes	No
Content	Up to 20 000 definitions	≈ 50 000 terms / language
Organization (Central language)	Central (English)	Polygonal

Sets of concepts:

[A] = Basic concepts requiring definition

(B) = Concepts to be derived from the corresponding terms or from the terms corresponding to set [A]

Figure 14: Characteristics of a future aeronautical Multilingual Reference System (MRS)

In many cases the texts for the understanding of which the MRS would be a tool could be used as sources for additional definitions and terms. Such texts are e.g. AGARD publications, military and other standards and specifications, and handbooks.

The definition of the concepts together with the corresponding terms in all languages covered by the reference system would form a "Part 1" of the system (Figure 14). As in the new MAD the organization of this part could be central, sorting the entries alphabetically according to the English terms and presenting the definitions in English only.

The number of more or less special terms used in aeronautics for any single language, however, is still larger than 20000. Many of these terms are derived from terms corresponding to the defined concepts of Part 1. In this case no additional definitions are required but one needs the equivalent terms in other languages. These terms should be included in the equivalence lists, which would than form a second part of the MRS (Figure 14). Generally the meaning of these terms should be clear. If not, there might be a short attribute indicating the relevant subject category or some short description, as it is used in most monolingual dictionaries for the general language.

In contrast to Part 1 all equivalence list should be strictly bilingual, the entries consisting of pairs of terms from the two languages concerned. Thus each equivalence would represent one direction on the connections of the polygonal scheme in figure 13. All equivalence lists would include the terms contained in part 1. The total number of entries per equivalence list probably would finally exceed 50000 in a more advanced stage of the MRS.

The supplement to the old MAD made some use of drawings. In preparing the new MAD it was thought, however, that drawings and mathematical formulae would add too much to the production costs. Moreover it was not possible to handle this type of information adequately with the available computer software. These restrictions will probably not exist for a future MRS. It would, therefore, be useful to include drawings as well as formulae, as these two express in many cases information much more clearly than a verbal definition and they are language independent.

On the linguistic side the future MRS should indicate pronunciations of the terms - at least if they are distinguishing characteristics as in figure 3 - and the necesary grammatical information as accents and other diacritical marks. It would even be desirable to find irregular grammatical forms of conjugation and declination again at least for those terms, which may not easily found in a normal dictionary of the general language. In addition to synonyms there should also be cross-references to antonyms and to related terms.

The new AGARD MAD will be published in the form of a printed book. This traditional form was chosen though, at some time, there was a suggestion to produce it on microfiche. This suggestion did not seem feasable due to the lack of very small and cheap microfiche readers. A future complete MRS could hardly be produced in the form of printed books if it would be expanded as proposed.

As in other fields the solution would be to create a computer-based system. This computer-based MRS would consist of a data-base and the necessary software. It could be used immediately on the computers in a manner similar to the retrieval systems of NASA. In addition to the normal dictionary-like use of the MRS it might be possible to obtain graphical information on terminals or on other output equipment and this graphical output then could handle mathematical formulae in the traditional manner. Finally it would not be to costly to produce dictionary-type computer-output-microfiches (COMs) covering any desired level of information (e.g. equivalence lists, contents of single subject categories, graphical information). The data-base as well as the software and the COMs could be updated regularly in any desirable time intervals. The updated COMs would simply replace the earlier ones. If this method were used supplements would not be needed and the use of the MRS would be greatly simplified.

The future computer-based MRS may seem difficult and expensive to produce. Certainly it will require man-power and funds. It would, however, be cost-effective as in every revision the results of previous revisions would be completely available and easy to use, whereas for a printed book at least the high printing costs are lost as soon as a new edition is printed.

The aeronautical MRS proposed would not be an isolated system but could be linked to or included in other technical reference systems, which may already exist or be developed in the near future. Furthermore it is important that the total cost of a reference system does not only include production costs but also the costs of its use and possible losses of time and money, which occur if no adequate reference system is available. Finally the data base from which the new MAD is produced would be a very good starting capital for the MRS.

- Correct
- Comprehensive
- Up-to-date
- Easy-to-use
- Cost effective

Figure 15: Summary of criteria for a future aeronautical Multilingual Reference System

Figure 15 summarizes the basic criteria for a future aeronautical Multilingual Reference System. The final goal will be reached in successive steps. The first step should be made in the very near future. It should concentrate on two main topics:

- correction of remaining errors and unification of the existing definitions,
- revision of some existing and addition of some new subject categories (e.g. flight safety, flight dynamics).

This would be a direct continuation of the now almost complete revision. With the available and the new material the necessary software for the production of the dictionary in a new form should be prepared and tested. The whole process should take less time than the current revision.

5. Conclusions

The lecture attempted to give a survey of problems connected with multilingual information production and transfer and of the reference tools used in these processes. It was necessary to point out various difficulties encountered in the use of reference tools. Based on the experience gained during the revision of the AGARD Multilingual Aeronautical Dictionary, criteria for a computer-based Multilingual Reference System were developed and discussed.

6. References

/1/ AGARD Aeronautical Multilingual Dictionary
G.H.Frenot and A.H.Hollcway, ed.
Pergamon Press, 1960

/2/ First Supplement to AGARD Aeronautical Multilingual Dictionary
A.H.Holloway, ed.
Pergamon Press, 1963

/3/ AGARD Multilingual Aeronautical Dictionary
Second revised and expanded edition
(in preparation)

/4/ British Standard 185
Glossary of Aeronautical and Astronautical Terms
Revised edition 1950-1951

/5/ British Standard 185
Glossary of Aeronautical and Astronautical Terms
Revised edition 1969-1973

/6/ ICAO Lexicon
Volume I: Vocabulary
Volume II: Definitions
Fourth edition
International Civil Aviation Organization, 1974

/7/ The United States Air Force Dictionary
Woodford Agee Heflin, ed.
Air University Press, 1956

/8/ International Standard ISO 1151
Terms and Symbols for Flight Dynamics
Parts I to VI
International Organization for Standardization
1972 sqq.

/9/ Universal Decimal Classification, International Edition
BS 1000 (English)
British Standards Institution

KINDS OF ACCESS TO UNCLASSIFIED LITERATURE

by

C P Auger
Lucas Group Services Limited
Lucas Research Centre
Shirley Solihull
West Midlands B90 4JJ

SUMMARY

The paper reviews the nature of unclassified aerospace literature, and examines what categories are available and how they are organised. The preponderance of reports literature is noted, and current awareness and on-line services receive particular attention. Consideration is given to the great variety of users seeking access for one reason or another, and the different routes open to them, from general institutions such as public libraries to specialist information centres such as that at the Royal Aircraft Establishment. Finally attention is turned to some of the problems which need to be overcome if access is to be improved.

DEFINITIONS

The term classification has several meanings, but here refers to the classifying of documentary information into groupings according to the stringency of the measures to be taken to prevent its falling into the hands of an enemy or potential enemy. Typical designations are Top Secret, Secret, and Confidential. By inference therefore, and for the purposes of this paper, Unclassified means, in the context of documents, not subject to a security classification.

The other definition which needs stating at this point is that for aerospace itself, and the term may be taken to mean an industry whose products world-wide include civil and military aircraft, helicopters, aero-engines, guided weapons, hovercraft, space vehicles, and a whole range of aircraft and airfield systems.

WHAT SORT OF LITERATURE IS AVAILABLE?

Unclassified aerospace literature can be divided into a number of categories depending on the depth of understanding and degree of specialised knowledge necessary to make the fullest use of it. AGARD for example is concerned primarily with the exchange of scientific and technical information; other bodies are more interested in commercial and marketing information; others again in the historical, educational and strategic aspects; and of course aerospace literature is of considerable appeal to the general reader, whether as a schoolboy enthusiast or a regular airline traveller.

The wealth of aerospace literature and the great interest in it reflects the often mentioned point that a progressive industrial society needs a pioneering spearhead technology to stimulate growth, promote innovation and set standards of excellence for a country's entire range of manufacturing, production and trading activities. Factors which contribute directly to the volume of aerospace literature produced and disseminated, especially reports literature, are the immenseness of the projects involved; the employment of large numbers of highly qualified, highly skilled personnel; the importance of aerospace in terms of national security; and not least the ruling that contractors working on government projects are required to report on progress at regular intervals. To this list can be added public concern over safety, noise and pollution.

Secrecy has always been a major feature of aerospace information, and many documents stay inaccessible to those who cannot demonstrate a need to know. Nevertheless there still remains a large body of aerospace publications which is actively promoted to benefit the aerospace industry in general by making available a common core of knowledge, and secondly to try and justify and recoup some of the costs involved by promoting product spin-off and technology transfer.

HOW IS THE LITERATURE ORGANISED?

The forms of publication in which aerospace information appears are no different from those of any other major branch of knowledge, namely books, journals, translations, trade literature, standards and specifications, technical manuals, patents, and reports. What is different is the concerted effort which has been made to ensure such publications are readily accessible, and as a consequence the abstracting and indexing services associated with aerospace literature can be held up as a model to other areas of science and technology.

Another major difference is the preponderance of reports literature. Indeed it can be argued that the history of reports literature coincides almost entirely with the development of aeronautics and the aircraft industry. The series of reports usually accorded the honour of being first in Great Britain is the R&M (Reports & Memoranda) series of the Advisory Committee for Aeronautics, now the Aeronautical Research Council, which began appearing in 1909. In the United States the aircraft industry has been represented continuously by the National Advisory Committee for Aeronautics (NACA), now of course the National Aeronautics and Space Administration (NASA), which issued its first report, On the behaviour of aeroplanes in gusts, in 1915.

Reports have since spread to other areas, notably nuclear energy, but still form a large and characteristic part of the aerospace literature.

As befits the country with the largest and most comprehensive aerospace industry in the Western world, the United States had led the way with two major abstracting publications. The first of these is Scientific and Technical Aerospace Reports (STAR) which is published twice a month as a guide to current technical reports issued by organisations around the world. STAR began publication in 1962, and since then some 325,000 citations have been listed. New citations are being published at a rate of 24,000 per year. In detail, STAR announces technical reports issued by NASA and its contractors, by other US government agencies and their contractors, and by both domestic and foreign companies, universities, and research organisations.

STAR reflects the interests of the NASA Scientific and Technical Information Branch (STIB) and its coverage embraces the basic and applied sciences relating to aeronautics and space research.

Despite the careful specification of NASA's subject interests by means of annotated scope notes, it has often been remarked that STAR is full of surprises. Thus for example it is possible to find details on establishing habitability factors for the design of office environments (N79-10744).

STIB is also concerned with the announcement of published documents other than reports, and accordingly sponsors the publication of International Aerospace Abstracts (IAA) by the American Institute of Aeronautics and Astronautics (AIAA).

IAA lists all journal articles, conference papers, books, and other forms of published literature, and the object is to complement rather than duplicate STAR, a task which with rare exceptions it manages to accomplish. IAA began publication in 1961 and by 1977 had presented around 500,000 citations of aerospace literature. Currently about 35,000 citations are announced each year.

Both STAR and IAA have been developed into a combined on-line data base which is available in Europe, but only to accredited users who enter into a tripartite agreement with the European Space Agency (ESA) and NASA, via the ESA-RECON installation at Frascati.

The NASA file is unique in that no royalties are levied - instead NASA requests users of its file to contribute on a regular basis suitable documents for inclusion.

In passing, it should be mentioned that the term RECON refers to remote console. In the United States, the data bases are available through NASA-RECON, Washington.

WHO WANTS ACCESS?

The range of users seeking access to unclassified aerospace information is as diverse as the forms of literature itself. The categories include research and development workers, performance engineers, company executives, university lecturers, students, and members of the general public. Each person will have his own reasons for wishing to consult the literature, for example the advancement of an experimental project, the establishment of design criteria, the assessment of a market, the submission of a thesis, or the study of a particular type of aircraft.

Detailed studies of specific categories of users have been presented at an earlier AGARD meeting (1), and include the contributions of J R Sutton (2) on the information requirements of engineering designers; Margaret O Sheppard (3) on users at the Aeronautical Research Laboratories, Australia; and Harold E Pryor (4) on methods used by NASA to evaluate user satisfaction.

HOW ACCESS IS ACHIEVED

Ease of access will depend to a large extent on the user's background. People working in research and development establishments and in large industrial companies will have the problems of identifying and obtaining literature considerably eased through the services of technical libraries and information departments. Lecturers and students in universities and colleges will be able to make use of main library and departmental collections, whilst the general reader's first resort will be the local public libraries, in particular the larger reference collections.

By using the appropriate channels, access can proceed from the local or regional level to national institutions, in the case of the United Kingdom the British Library Lending Division (BLLD) and the Science Reference Library (SRL).

The practical business of gaining access to aerospace information can be broken down into two components - identifying the publications required for the purpose in hand, and acquiring copies on loan or for retention.

Identification may be accomplished by consulting library and information specialists, by studying guides to the literature, and by making use of current awareness services such as STAR and IAA mentioned above. The widespread introduction of on-line searching facilities has considerably reduced the period taken to process enquiries, whilst at the same time it has greatly increased the comprehensiveness of the searches which can be undertaken.

In the particular case of aerospace literature, access to the ESA-RECON installation in the United Kingdom can be arranged through DIALTECH, the technical information service provided by the Department of Industry's Technology Reports Centre, (TRC) Orpington. Users in other ESA member states have similar access through the national centres nominated by ESRIN.

Details of experience with ESA can be found in many recent papers, including S Olmer - Court (5) on the French connection; L Maat (6) on the system at the Technical University, Delft; R Gülich (7) on experience at Sulzer AG; and S K Kumar and V A Kamath (8) on access to data bases in India.

Experience has shown that it is essential to impose an intermediary between the literature user and the on-line service. Usually the intermediary is an information specialist who can on the one hand exploit the interactive nature of the facility and make full use of the great number of access points, and on the other hand overcome the disadvantages of systems difficulties and data base inconsistencies. A further reason for using a trained intermediary is that an untrained user may inadvertently run up a large bill in his attempts to find references which match the requirements of a badly framed question.

Once references have been identified as relevant, the question of document provision crops up. References to published literature usually mean originals can be obtained on loan through whatever library channels are more convenient, or purchased through the book trade or official agencies. Because data bases provide full and detailed bibliographic information, problems of verification and checking are not normally encountered.

Items identified as reports however can present difficulties which are best resolved by applying to institutions which are known to specialise in such documents and to possess large reports collections. The TRC is one such source, another BLLD and a third SRL. Confirmation of holdings in the case of BLLD can easily be obtained for major reports series by consulting the latest edition of its publication "Current Serials Received", wherein are quoted the shelf marks for AD Reports, NASA publications of all types, PB Reports, and many more.

Again similar collections are available in ESA member states, although many European organisations prefer to apply direct to BLLD for a wide range of literature.

In the United Kingdom, a key collection of reports devoted specifically to aerospace is that maintained by the Royal Aircraft Establishment (RAE), Farnborough. Firstly it contains the research work of RAE itself, as recorded in Technical Reports and Technical Memoranda. In addition the research reports of other organisations all over the world are of great interest and use to RAE staff, and so are collected as well.

RAE make the point that the procedure for authorising the release of its own reports outside the Ministry of Defence (of which RAE forms part) is necessarily complex, even though the documents in question may be unclassified. The arbiter in such situations is the Defence Research Information Centre (DRIC), neighbour at Orpington of the TRC. RAE will advise on the availability of other reports in its collection.

PROBLEMS AND REMEDIES

A great deal of effort has gone into making access to unclassified aerospace information as easy and as efficient as possible. Providing the user for his part takes the trouble to think out and present questions in as precise a manner as possible, the system is available to provide speedy and relevant answers.

Three major areas of difficulty face the providers of information. Firstly, the need to educate the user so that he presents his requests logically and concisely, providing for example details of sources already known or tried, limitations on the period of interest, back-ground information likely to be helpful, and terminology characteristic of the subject under review. Only when the user knows clearly in his own mind what he wants and can be persuaded to take the information specialist fully into his confidence, will the best results be achieved. Admittedly many seekers after information are genuinely not able to refine their ideas since they are not quite certain what it is they are looking for until they actually see it. Nevertheless they can be encouraged to try.

Secondly people seeking access to information need to be aware of the importance of accurate data. Documents misquoted and authors' names misspelled are but two of the hazards which can prevent questions reaching successful conclusions.

Finally the user needs to realise that his and his alone is the task of understanding and digesting the information he has succeeded in gaining access to. Many readers are daunted by the sheer volume of the material they have unearthed and far from appreciating the efficient and rapid systems which have made it possible, are frequently moved to comment that they wished they hadn't asked!

APPENDIX

One of the first steps in the preparation of this paper was to conduct a search of the appropriate RECON data bases to see whether the subject of classified aerospace information had been examined, by other reviewers. It was felt that classified information ought to be easier to define than unclassified material, and that any discussion of such documents would inevitably provide an indication of the freedom of access to material not subject to security restrictions.

Accordingly a search was run in the NASA data base covering the period 1961 - 1979, using the following strategy : (1) REPORTS or (2) LITERATURE or (3) DOCUMENTATION or (4) INFORMATION RETRIEVAL, making set (5), which was combined with (6) CLASSIFYING and (7) SECURITY. No hits were registered with (5) and (6) and (7). Combining (5) and (6) only gave one reference, a report on the application of LANDSAT imagery. When (5) and (7) were combined, a total of 33 hits was indicated, but the papers themselves proved unhelpful, and the first couple of titles were concerned with the security of the national energy plan and a signature scheme for computer security.

Attention was then turned to the NTIS data base over the years 1969 - 1979, using the terms (1) DOCUMENT or (2) REPORT or (3) INFORMATION RETRIEVAL to make set (4), which was then combined with (5) CLASSIF and (6) SECURITY to give set (7), with an indicated total of seven hits. However when an extra term (8) AEROSPACE was added, no hits were indicated. Thus it was decided to see what set (7) contained, and the documents retrieved covered the following topics : solid waste composition and emission factors; military message experiments; the delivery of social services to elderly persons; Multics security enhancements (the Honeywell Multiplexed Information and Computing Service); a security compliance study of the Air Force Data Service Center; papers from the Classification Management Journal; and the Special Work Project for the Unemployed.

The most relevant of the above items was the collection of papers from the Classification Management Journal (AD-A025 341), the controlled and uncontrolled indexing terms for which at least confirmed that the strategy used was the correct one. However the search as formulated did not achieve the desired objective, namely the identification of papers on classified aerospace literature. No doubt other search strategies might have produced a different result.

REFERENCES

(1) The problem of optimization of user benefits in scientific and technological information transfer. NATO Advisory Group for Aerospace Research and Development. Neuilly sur Seine, AGARD 1976, 126p illus. tables. (AGARD Conference Proceedings No 179).

(2) J R Sutton. Information requirements of engineering designers. (AGARD Conference Proceedings No 179, 12-1 - 12-8).

(3) M D Sheppard. User response to the SDI service developed at Aeronautical Research Laboratories, Australia (AGARD Conference Proceedings No 179, 10-1 - 10-9).

(4) H E Pryor Listening to the user : a case study. (AGARD Conference Proceedings No 179, 11-1 - 11-7).

(5) S Olmer-Court ANRT, the national connection to the European Space Agency. (Information et Documentation, (3) Sep 76, 29-34).

(6) L Maat. The ESA/SDS information system at the Technical University, Delft, (Open 8(6) Jun 76, 304-308).

(7) R Gülich. Experiences with on-line searching (Nachr./Nouv./Notiz, 53(3) 1977, 97-107).

(8) S K Kumar, V A Kamath Access to world data bases and ESRIN/RECON on-line demonstration in India. (Ann. Libr. Sci. Docum., 23(3) Sep 76, 205-212).

LA GESTION DES DOCUMENTS AYANT UN CARACTERE DE RESTRICTION

par

Mr J H Klopp
Chef de la Division Information
CEDOCAR
26 Bld Victor
75996 Paris Armees
France

RESUME

La classification des documents : le secret militaire; tableau des marquages de restriction dans différents pays; le secret industriel. Importance de la déclassification des documents quand elle devient possible. Le signalement, le stockage et la diffusion des documents à caractère de restriction.

L'information étant le bien le plus précieux d'un pays il importe de la traiter et de la diffuser à bon escient.

Une organisation susceptible de la gérer d'une façon efficace doit être mise sur pied.

On distingue bien entendu deux types d'information, l'information ouverte, accessible à tous et celle à caractère de restriction réservée aux besoins de l'Etat.

On pourrait ajouter une troisième catégorie, appelée "littérature grise", c'est-à-dire l'information n'ayant aucune restriction, mais non diffusée pour des raisons diverses (oubli, manque de moyens etc...).

L'objet de notre présentation ne couvre que la documentation non publiée, c'est-à-dire celle qui pour une raison ou pour une autre ne se trouve pas sur la place publique. Le problème des archives n'est pas traité non plus.*

Dans le domaine ouvert je me permets de vous rappeler la Lecture Series N° 69 de l'AGARD, "How to obtain information in differenct fields of Science and Technology. A user's guide", ainsi que le rapport établi en 1978 à la demande du Comité for Information and Documentation on Science and Technology, de la Commission des Communautés Européennes "Transfert de l'Information pour l'Industrie".

Ces documents font le tour du problème de la Gestion de l'Information ouverte.

Les documents à caractère de restriction se distinguent en général de la littérature ouverte ou grise par un signe particulier permettant de voir si la communication des documents au grand public ne risque pas de porter préjudice, sous une forme ou sous une autre à quelqu'un ou à un organisme privé ou officiel.

On remarque également que d'une façon générale les rapports techniques ou scientifiques sont souvent diffusés de façon trop exclusive surtout lorsque ces documents contiennent des informations confidentielles.

Un juste milieu doit donc être trouvé pour améliorer la circulation de l'information tout en assurant la protection des sources des informations.

1. CLASSIFICATION DES DOCUMENTS

La gestion des documents à caractère de restriction, doit tenir compte des limitations imposées par:

- La protection du secret de la Défense Nationale (donc par les règles de sécurité qui en découlent).
- La protection des intérêts légitimes de l'organisme émetteur du document.

---

* A quelle stade de la vie d'un document peut-on parler d'archives?

Nous nous trouvons de plus en présence de deux aspects de la question, à savoir la protection du secret militaire et celle du secret industriel.

1.1 **Le secret militaire** comporte un certain nombre de dégrès, allant de la Diffusion Restreinte (Restricted) au Top Secret (Très Secret). L'emploi des timbres correspondants est strictement réglementé et l'application d'un de ces timbres sur un document ne doit être effectuée qu'aprés mûre réflexion; il est nécessaire d'etudier et d'évaluer si la divulgation d'un document ne risque pas d'entraîner un dommage exceptionnellement grave ou un dommage sérieux, ou seulement un préjudice plus ou moins important etc . . . et c'est alors seulement que l'on choisira le timbre de protection approprié. Ces timbres sont complétés par d'autres timbres, (cela dépend des pays), spécialement adaptés à des situations particulières; c'est ainsi que l'on trouve en complément des timbres précités, des timbres précisant des limitations complémentaires ou au contraire permettant un élargissement de la diffusion de ces documents à un cercle différent de celui initialement prévu. Par exemple: timbre précisant que le document ne peut être diffusé que dans tel ou tel pays, à l'exclusion des autres ou bien limitant la diffusion aux seuls pays membres de l'OTAN, etc. . .

Le tableau ci-dessous montre les timbres de restriction utilisés par les différents pays members de l'OTAN.

OTAN	COSMIC TOP SECRET	NATO SECRET	NATO CONFIDENTIAL	NATO RESTRICTED
Belgium	TRES SECRET	SECRET	CONFIDENTIEL	DIFFUSION RESTREINTE
Canada	TOP SECRET	SECRET	CONFIDENTIAL	RESTRICTED
Denmark	YDERST HEMME-LIGT	HEMMELIGT	FORTROLIGT	TIL TJENESTEBRUG
France	TRES SECRET	SECRET-DEFENSE	CONFIDENTIEL-DEFENSE	DIFFUSION RESTREINTE
Germany	STRENG GEHEIM	GEHEIM	VS-VERTRAULICH	VS-NUR FÜR DEN DIENSTGEBRAUCH
Greece	AKROS APORRI-TON	APORRITON	EMPISTEFTIKON	PERIORISMENIC CHRIS-SEOS
Italy	SEGRETISSIMO	SEGRETO	RISERVATISSIMO	RISERVATO
Luxembourg	COSMIC TRES SECRET	SECRET	CONFIDENTIEL	DIFFUSION RESTREINTE
Netherlands	ZEER GEHEIM	GEHEIM	CONFIDENTIEEL or VERTROUWELIJK	DIENSTGEHEIM
Norway	STRENGT HEMMELIG	HEMMELIG	KONFIDENSIELT	BEGRENSET
Portugal	MUITO SECRETO	SECRETO	CONFICENCIAL	RESERVADO
Turkey	ÇOK GIZLI	GIZLI	OZEL	HIZMETE OZEL
United Kingdom	TOP SECRET	SECRET	CONFIDENTIAL	RESTRICTED
United States	TOP SECRET	SECRET	CONFIDENTIAL	–

Vous remarquerez que les Etats Unis n'ont pas de classification Diffusion Restreinte (Restricted). Les documents entrant dans cette catégorie sont reclassés à un degré supérieur mais avec un marquage spécial, permettant des procédures de circulation identiques à celles couramment utilisées pour le "Restricted".

Outre ces marquages de restriction il existe de nombreuses annotations qui orientent différemment la diffusion des documents, c'est ainsi que l'on peut trouver sur un document en provenance d'un pays ami, membre de l'OTAN:

DIFFUSION RESTREINTE

DOCUMENT RESERVE A LA DEFENSE

et un marquage particulier:

DIFFUSION INTERDITE A L'INDUSTRIE

1.2 **Le Secret Industriel** et la protection des intérêts industriels, ont fait l'objet de nombreuse études et un groupe de travail de l'OTAN a même recommandé, outre l'utilisation des timbres de sécurité normaux, l'emploi de timbres particuliers. Ces timbres ne sont pas utilisés pour des raisons de sécurité, mais pour sauvegarder les intérêts du propriétaire de l'information.

Les accords OTAN du 19.10.1970 sur la communication des Informations Techniques pour les besoins de la Défense ont même recommandé l'emploi de ces timbres.

La nécessité de protéger les intérêts des industriels de chaque pays, entre autre la prise de brevets, est bien entendu évidente et l'échange d'information entre pays ne peut s'effectuer que si chacun respecte et accepte les différents timbres ou marquages figurant sur chaque document.

Mais il ne faudrait quand même pas tomber dans l'absurde,à savoir, trouver sur un document non classifié à diffusion illimitée un timbre précisant que la divulgation de ce document est interdite aux industriels.

Ceci nous amène à préciser également qu'un document non classifié, c'est-à-dire sans restrictions de sécurité, n'est pas forcément à diffusion illimitée. Des raisons valables "limitent" leur diffusion (par exemple, droit de propriété privée ou d'Etat). Et cette diffusion ne peut être effectuée qu'avec l'accord de l'organisme émetteur de document.

En résumé les documents échangés entre pays comportent presque toujours des timbres dits de "Sécurité"* et des timbres ou marquages dits de "protection industrielles". Précisons que les informations contenues dans les documents peuvent perdre de leur importance avec le temps et par conséquent une politique de "déclassement" des documents, est indispensable, afin d'obtenir une meilleure circulation des informations au sein des pays membres de l'OTAN.

2. DECLASSIFICATION DES DOCUMENTS

Comme pour le classification des documents, la déclassification des documents n'est pas du ressort des Centres de Documentation de la Défense, si ce n'est de fournir une liste de documents classifiés aux autorités compétentes afin qu'elles puissent prendre des décisions idoines.

En général, la décision qui modifie ou supprime la protection d'une information appartient toujours à l'autorité d'origine.

Les modifications de la protection sont effectuées de façon automatique quand elles peuvent être prescrites dès l'émission de l'information par rapport à un critère déterminé.

C'est ainsi que plusieurs pays ont adopté un système voisin. A chaque émission d'un document, un formulaire doit être rempli par l'émetteur, précisant le pourquoi du timbre de restriction; par exemple: informations commerciales, droit de propriété, informations étrangères, inventions etc . . . . ainsi que les limitations de diffusion. Une précision complémentaire, quant à la déclassification est également demandée (déclassification automatique après . . . n . . . années, ou bien déclassification automatique refusée, mais a revoir dans . . . n . . . années).

Par ailleurs des indications bien visibles sur les documents classifiés permettent aux destinataires de ces documents de savoir quand ils peuvent de leur propre autorité déclassifier ces informations.

* Les timbres Unclassified – Limited ou Unclassified – Unlimited n'existent pas dans la législation française.

Il est à noter que le destinataire d'un document à caractère de restriction peut demander à l'autorité émettrice de modifier le degré de restriction d'un document, (en fournissant des justifications), soit en demandant son déclassement, soit en exigeant un relèvement du timbre de sécurité.

Le rôle des Centres de Documentation de la Défense consiste en général à éditer régulièrement pour les besoins de la Défense (puisqu'ils sont destinataires d'au moins un exemplaire des Documents Scientifiques et Techniques de la Défense), une liste de documents déclassifiés afin d'alerter le plus grand nombre de personnes de la disponibilité de nouvelles informations.

Pour terminer cette question, je voudrais insister sur l'importance du problème de déclassification. En effet dans ce domaine plusieurs points importants sont à noter:

- il n'y pas de normes OTAN sur les procédures de déclassification des documents
- Le côut du "surclassement" ou du maintient des documents à un niveau de restriction trop élevé, est très important (stockage dans des locaux spéciaux, traitement particulier des documents, surveillance renforcée etc . . .).
- Les auteurs émetteurs de documents classifiés changent souvent de domaines techniques ou d'organismes et par conséquent sont difficiles à joindre pour obtenir de leur part une autorisation de déclassification de leurs documents.
- Dans le domaine de la déclassification il serait nécessaire qu'il puisse y avoir un point de vue unique au sein des pays membres de l'OTAN, la circulation de l'information y gagnerait.

3. GESTION DES DOCUMENTS A CARACTERE DE RESTRICTION

En règle générale, la gestion des documents à caractère de restriction est effectuée de façon analogue à celle de la documentation ouverte, mais avec des règles de sécurité importantes et toujours contraignantes pour un Centre de Documentation.

Il est préférable pour un Centre documentaire de la Défense de créer une chaîne d'exploitation parallèle à la chaîne d'exploitation classique de la documentation ouverte, plutôt que de les mélanger.

C'est ainsi que plusieurs pays ont créé des Centres Spéciaux pour le traitement des documents classifiés afin d'éviter des interférences entre ces deux chaînes.

L'avantage le plus important d'une telle décision, est incontestablement un contrôle plus aisé du signalement, du stockage et de la diffusion de ces documents.

3.1 Signalement

Les Centres éditent régulièrement un bulletin signalétique des documents classifiés. Les signalements de ces documents sont présentés de la même façon que ceux des documents ouverts, mais avec indication du niveau de restriction.

Certains organismes fournissent des signalements non restrictifs de documents classifiés, d'autres, par contre, considèrent que les signalements ont le même caractère de restriction que le document lui-même*. Dans le premier cas le bulletin signalétique ne comporte aucun caractère de restriction, dans le deuxième cas le bulletin possède en général un degré de restriction correspondant au degré de restriction du document le plus restreint.

A noter que des recherches bibliographiques ainsi que des D. S. I. en batch peuvent également couvrir des documents à caractére de restriction.

Par contre l'interrogation à distance, d'un fichier de documents classifiés est bien entendu soumise à des règles de sécurité très strictes (sécurité des lignes, cryptographie etc . . .).

3.2 Stockage des Documents Classifiés

Le stockage des documents classifiés est soumis à des règles de sécurité bien précises, par exemple accès contrôlé aux bibliothèques, stockage des documents dans des armoires fortes, etc . . . contrôle journalier des documents existants, bilan des entrées et des sorties etc . . .

---

* *Les documents ayant un haut caractère de restriction, ne sont en principe jamais signalés, car ils sont réservés à un cercle très restreint de destinataires.*

3.3 Diffusion des Documents

Si la diffusion inopportune de signalements de documents à caractère de restriction peut déjà présenter un certain danger, la communication des documents eux-mêmes est encore plus grave. Le seul contrôle de l'habilitation du demandeur est insuffisant et risque de mettre le Centre de Documentation gestionnaire dans une situation très délicate. C'est pourquoi le Centre exige, de plus, la justification du "besoin de connaître" (need to know) du demandeur déjà habilité au degré de restriction du document.

Dans le domaine de la Défense la justification du "besoin de connaître" est délivrée par l'organisme de la Défense responsable de la recherche, ou de l'étude ou de la réalisation d'un matérial à laquelle participe le demandeur (dans un établissement de la Défense ou sous contrat dans l'industrie). L'habilitation au degré de restriction est placée sous la responsabilité du responsable de la "sécurité défense" de l'établissement.

4. CONCLUSION

Le problème le plus délicat de la gestion des documents à caractère de restriction est de transmettre l'information nécessaire aux personnes habilitées tout en protégeant les documents, les organismes émetteurs ou transmetteurs de ces documents et le Centre gestionnaire lui-même.

Le responsable d'un Centre de documentation classifiée doit arriver à concilier deux impératifs contradictoires: restreindre pour protéger et diffuser pour informer qui de droit.

Pour terminer, je me permets de rappeler l'excellent ouvrage de nos amis

A.H.Holloway, Elizabeth H.Ridler, Domenic A.Fuccillo, Marvin E.Wilson et B.Yates,

intitulé "Information work with Unpublished Report", publié par l' Institute of Information Scientists" (monograph series), London UK en 1976 et par Westview Press, Boulder, Colorado, USA en 1977.

FULL TEXT HANDLING - A CRITICAL REVIEW

by
Dr. R. P. L. Jones,
Computer Science Division,
Atomic Energy Research Establishment,
Harwell, Oxon OX11 ORA.
G.B.

SUMMARY

The paper sets out to demonstrate the wide capabilities of full text information retrieval systems. In the first part, the basic capabilities of such systems are described including the strengths and weaknesses that seem to be inherent in the techniques that are in use. In the second part, the wider application of such systems is described. Emphasis is placed on their capability to form a nucleus for integrated information management systems which are flexible in use and do not require programming skills to exploit them.

1. INTRODUCTION

Full text handling by computer is an approach to information retrieval which has a long, if rather chequered history. The techniques actually go back to the manually produced concordances used in biblical research to identify the sources of quotations by indexing the occurrences of the words in the text. The first computerised concordance based systems were applied to the field of legal research, and this is the area which many people still associate with full text retrieval. It would appear that this occurred because lawyers were suspicious of having legal text classified and indexed for them, and insisted on having the original source text fully indexed and available to them. The other reason for the expansion of legal applications was that lawyers seem to have had sufficient financial resources to be able to afford the cost of retyping and storing the voluminous amount of text involved.

The acceptance of full text techniques in other areas such as bibliographic applications has been more recent.

The author believes that this is due at least in part to confusion as to what is meant by full text retrieval. It is, therefore, appropriate to begin with the basic definitions of the terms used in this subject.

2. DEFINITIONS

The author believes that the phrase free text handling is a more accurate description of the techniques in use, rather than full text handling. The justification for this change will follow later.

<u>A free text information retrieval system</u> is fundamentally a body of software that stores and indexes a stream of characters, called the <u>text</u>. The text is made up of groups of two types of characters, called <u>concordable</u> and <u>delimiting</u> respectively. Concordable characters are usually the letters forming the alphabet and often the digits. A typical delimiting character is a blank. Groups of concordable characters, terminated by one or more delimiting characters, form a <u>word</u>. A word forms the basic token for indexing purposes. The words will be grouped into <u>records</u> for informational reasons in that the records will form the basic unit of retrieval.

The significant feature that distinguishes a free text system from other information retrieval techniques is that, in general, all the words in the text are indexed, so that the occurrence and position of every word in the text is recorded. The index thus formed is the <u>concordance</u>. A concordance is made up conceptually of two parts. The first is a <u>dictionary</u>, a structured list of all the unique words forming the <u>vocabulary</u> of the text. The other part of the concordance is the <u>reference lists</u> giving all the occurrences of each word in the dictionary. These define not only the records which contain that word, but often the word position in the record.

The most typical example of the text in a free text system is natural language. The words are those of the natural language, delimited by spaces or punctuation. The records are documents or articles, or whatever is appropriate for the uses to be made of the text.

The author postulates that this natural language example forms the definition of a typical <u>full text retrieval system.</u> The more general definition was given to underline the fact that in a <u>free text</u> system the text does not have to be natural language. The other feature of free text systems that are being used in real environments is that further structures are added to the system to aid in the retrieval and display of information. This will be elaborated at a later stage in the paper.

Examples of free text systems in commercial use are the IBM system STAIRS, the BASIS system from Battelle Laboratory, and STATUS developed by Harwell Laboratory in the U.K. These may be contrasted in their scope and breadth of use with LEXIS, developed by Mead Corporation, and STATUS I, an earlier product from Harwell, which in our definition are full text systems, both developed for legal research.

3. BASIC RETRIEVAL FUNCTIONS

Every practical free text system must have two fundamental functions. It must be able to search the references in the concordance, and display the records containing the references satisfying the search criteria. Of course, nearly every system will have many other capabilities in addition to these.

These are discussed in a later section of this paper since they are fundamental to the second theme of this paper, the application of free text retrieval techniques to provide integrated information management systems.

The simplest form of search is to retrieve all records that contain a given word. All free systems in practical use, use a set of Boolean operators to search for combinations of words in the same record. The typical operators are AND, OR and NOT. AND demands that the two words linked by the operator both occur in the same record. OR requires that either or both words occur in the same record, while NOT selects records that contain one word but not the other.

It may be noticed that the above operators use only a single level of reference, the record number, for the occurrence list of a word. However, other operators in common use, the collocation operators, demand that the relative position of the word in the record be retained.

The simplest collocation operator is the phrase. The two words specified must be adjacent in the same record if the search criterion is to be satisfied. More generalised collocation operators allow the user to specify that two words must be separated by a specific number of words or a range of intervening words.

Words and operators may be strung together according to a given set of syntax rules to form complex queries. It is interesting to note that, in spite of much academic research in this area, all implemented systems use these fundamental operators to retrieve records from the data base.

A commonly supplied feature is a mechanism whereby a word stem may be supplied in a question, rather than in a word. A word in the vocabulary satisfies the search criterion if it begins with the specified stem. This feature allows for grammatical variations of a word to be included in a search. It has the disadvantage that it may include all sorts of miscellaneous words which the questioner had not intended to specify. Further, very deep truncations, down to a few characters, pull in a vast number of words, which can soak up a massive amount of computer time.

The records which satisfy the criteria specified in the search form a retrieved list. This list is always retained at least until a further search is made, and in many systems it may be retained and referred to until the user session ends. The text specified by a retrieved list can be displayed at the terminal, so that the relevance of the answers may be assessed by the user.

4. THE PHILOSOPHY OF THE FREE TEXT APPROACH

The free text approach to information retrieval is underpinned by a number of basic principles. The most basic is that such systems seek to put as much skill as possible into the hands of the person retrieving the information, rather than to put the onus exclusively on the person who is preparing the information. The retriever must have a feel for the language in which the text is couched. He may then convert his informational need into a query involving the words that express the concept for which he is searching.

Such systems are, therefore, most effective if the end customer actually interrogates the system rather than use an intermediary. It follows that free text systems must be easy to use. Further, since each person is likely to want to follow his lines of questioning in his own way, the system must be command driven.

Since it is likely that a series of iterations will be necessary to obtain a satisfactory answer to a query, it is essential that such systems be interactive. The human terminal operator is a vital link in the retrieval process, since he can examine the answers retrieved to indicate whether he couched his query sensibly. The system must be human oriented, and one aspect of this is that the source text must be interfered with as little as possible by the system. It must be possible to preserve the original layout of the text, and, as far as possible, not force a fixed layout on the source text.

Nearly all these points are diametrically opposed to the other type of information system in common use, the thesaurus based, keyword approach. These systems rely on an indexer selecting out of a fixed set of words to describe the contents of a record. The retriever of information is much less involved in the creative part of information retrieval, the couching of the query, since the range of words is much more limited.

Such systems tend to be batch oriented, or to have developed from batch systems. This is possible, as a more concise question can be asked initially. The systems are very often not easy to use and are computer oriented in their interface with the user.

The reason why such systems are better known than the free text systems is simply that they have been around for longer. Batch processing with over-night turnround has been a practical possibility for a lot longer than the interactive computing facilities needed by the free text systems. It is significant that a number of keyword systems are beginning to offer free text capabilities in addition to the keywords. Conversely free text systems are beginning to offer thesaurus and keyword facilities.

5. THE PERFORMANCE OF FREE TEXT SYSTEMS

Although free text systems look very different at the user interface level, most systems will share common internal structures. Hence they will tend to have similar performance characteristics, especially in the area of searching the data base.

The search overheads will be dominated by the number of references to be combined, if more than one search term is to be used. It is likely that the processing involved in finding the start of a set of references will be quite modest, no matter what the size of the data base. Thus the major impact of a large data base will be that there will tend to be a larger number of references for a given word, rather than that the vocabulary is larger.

The major retrieval performance problems in free text systems arise out of their generality. Frequently used searches will have to be reprocessed since, although many systems allow the storage of references, for later use, such lists tend to become obsolete as the data base changes. Systems do not yet possess good mechanisms to analyse the pattern of queries and maintain frequently used reference lists of popular searches for general use.

The author is strongly of the opinion that effective free text systems must have the ability to carry out truncated searches and phrase searches. However, these facilities allow the user to ask expensive questions without intending to do so. Truncations may well bring in many more search terms than the user intended. Queries linking together words to form phrases may be expensive, since the two words forming the phrase may occur many times in many different contexts, whereas the phrase may have only a small number of references. In both cases it is impossible for the system to judge whether such requests are reasonable or not. The best that can be done is to warn the user that he is consuming significant amounts of resources.

Such performance points can be overstressed, however, since the user will spend a major part of his time reading the results of his searches. The requirement from the user point of view is that text display must be virtually instantaneous, while the queries may take longer, as the user intuitively feels that this is reasonable. However, the interactive nature of free text systems requires a response time which does not make the user lose his train of thought completely.

The other performance area of vital importance to free text systems is the time taken to update the data base. The author's view is that it must be possible to alter small portions of the text, and update only those parts of the concordance that are affected by these changes. This capability is normally referred to as dynamic update. Otherwise it becomes impossible to maintain the currency and accuracy of a data base, particularly when the text is changing very rapidly.

The requirement for dynamic update almost certainly adds an extra degree of complexity to the concordance structure. This will reflect itself in higher unit costs to concord small quantities of text. However, when large volumes of text are to be processed at one time, the overheads of concording can be substantially reduced by various optimising techniques, so that the requirement for a dynamic update capability does not impose a significant extra cost on concording.

There is therefore a price to be paid for the generality and flexibility of a free text system. The system has indexed everything at the level of a word, so any more complex searches have to be made by combining reference lists together, at the time when the query is made. A further point is that free text systems can consume a large amount of disk storage. An expansion factor of 2 or 3 to 1 of raw text to final data base size is typical. There is not much scope for data compression, again because of the generality of the system.

It must be borne in mind that the decreasing costs of computing make all the above physical overheads less and less significant. This is especially true on dedicated mini-computers, on which a number of free text systems will run. Also the cost of human resources is increasing, as is the volume of information, so that free text systems, which can require minimal text preparation, seem to be destined to come more and more into their own.

6. ENHANCED FEATURES OF FREE TEXT SYSTEMS

The features common to most free text systems have been described in an earlier section of the paper. Every system in practical use will offer extra features on top of these basic ones. These will have been included for one of five reasons; to increase the power of searches, to provide more flexible display of text, to allow greater integration with other systems, to provide security, and to enable the user interface to be tailored to a particular need. Each are discussed in turn.

One of the problems for free text systems is how to help the user think of the alternative ways in which the concept he is searching for may be expressed. This is particularly the case when the vocabulary of the text is not constrained, but may be equally true when there is a controlled vocabulary. The provision of a synonym capability to provide terms to be OR'd in a query, and less commonly a thesaurus to link synonyms by various relations such as 'broader term' will help considerably. A thesaurus in a free text system will have a rather different function from a thesaurus in a keyword system. In particular, not all words in the vocabulary need be in the thesaurus, and not all queries need involve the use of the thesaurus.

Most records will contain numeric information as well as text. There is a widespread need to be able to search for records that contain values (such as dates) lying within certain ranges, as well as possessing other characteristics.

In many cases, there is a need to present the text satisfying a search in a different format from its original form. This may be done by presenting parts of records, or by presenting the text to fill different width terminals. The facility to highlight search terms as they appear in the text is a very useful way to assess the relevance of the record.

The integration of free text systems with other systems is a subject discussed in more depth in the next section. The most important areas are to integrate information capture and correction with the retrieval system. Thus it should be possible for a user to enhance the information in a data base by adding to the text of a record, while he is interacting with the system. At the other end of the process it is essential to be able to print text and also to route the output to a disk file rather than to the terminal.

The security offered by any system is a major factor in deciding its usability. It must be possible to protect a complete data base or individual records or parts of records against unauthorised access. In addition, the integration of updating with retrieval demands control over the alteration of text, and a vetting process which must take place to examine the modifications that have been made before they are integrated into the data base proper.

The final area where free text information retrieval systems can offer enhancements, is in the user interface. It should be possible to alter any command name or message, either into another language, or into a set of terms that are more familiar to the user. To go beyond this is perfectly possible however. For example, the STATUS system allows the user to define a macro which has a name and a body. The body is substituted into the input command whenever the name occurs. Thus complete commands or partial commands can be replaced by simpler constructions. An enhancement under development allows groups of commands to be combined into a single macro, allowing the construction of a purpose built language.

7. APPLICATION AREAS FOR FREE TEXT SYSTEMS

In this section, the author draws on his experience with the STATUS information retrieval system to indicate the breadth of application that is possible with free text systems. The areas in current use include:

Legal text, both statute and case law.
Bibliographic retrieval.
Library management.
Experimental results.
Chemical and pharmaceutical information.
Personnel information.
Accident reports.
Police records.
Production records.

In the majority of these areas, the information retrieval system provides the framework, not only for ad hoc enquiries, but also to supply information on a regular basis for other processes. For example, a legal publishing company is providing an enquiry service for lawyers, but is also using the same information in the data base to produce typesetting tapes for their normal publishing business. One technical library has integrated bibliographic retrieval with library management, with STATUS controlling the issue of overdue notices, and all ordering information. In another application, the reduced results of experiments are stored with the values of various physical parameters in force at the time, together with textual comments about the experiment. A set of experiments which satisfy a particular range of values can be retrieved from the data base and the results presented in graphical form, or as reports.

In all these applications, the flexibility of the system allows the retrieval of records satisfying essentially any criteria that are required. Communication with related special purpose programs is achieved simply through data files of retrieved records. Thus there is a natural split between the retrieval system which can retrieve on virtually any criterion, but does not understand the details of the information and the special program which cannot retrieve the information but understands its structure more clearly and also knows what specific action is needed.

It is interesting to note that many of the application areas mentioned above are traditionally handled by special purpose programs. The differences between keyword based systems and free text systems have already been mentioned. The numeric capability available in some free text systems makes them contenders in areas often thought of as data base management (DBMS) applications. The DBMS approach is to index on a small number of fields and to build each record as a set of pigeon holes, with each box containing a pointer to another record, or a piece of data. The free text approach indexes virtually everything, so that cross references are made by referring to another record, perhaps by a reference number in the text. There is no limitation on field lengths or record lengths. Of course, in many applications the DBMS approach gives more rapid responses, and will occupy less disk storage. However, if records vary in size, and a wide number of ways of retrieving information are required, then the free text approach is appropriate. It is vital to remember that no programming is needed to retrieve information from a free text system. Police applications have these requirements, and it is significant that there is a growing interest in police forces in the uses of free text systems.

Since the typing of text is a very expensive process, it is important to minimise the number of times it has to be retyped. The advent of word processing terminals has produced a supply of source text in machine readable form, which may often be automatically converted into an appropriate input format for free text systems. The same is true of computer typesetting tapes. In both cases this is possible because such systems require a minimum of information apart from the text itself. Hence data bases may be built up as a by product of other processes, thus completely altering the cost justification for them.

It is interesting to note, however, that an organisation often begins by regarding an information retrieval system as a convenient sink for information produced elsewhere. However, with realisation of the true potential of the system, the information retrieval system becomes much more central, and provides the information that may be disseminated by other processes such as typesetting, report generation, bulletin production, selective dissemination of information (SDI) or graphics. All these processes can be carried out using a system that is basically interactive in nature, and does not need any programming skills to exploit.

8. CONCLUSIONS

This paper has set out to examine the possibilites of free text information retrieval. It defined the concepts behind the free text approach, and has looked at the limitations to the method, and the relationship with other technqiues. While not pretending that the approach is universally applicable, the paper attempts to illustrate how very flexible information management systems can be built up, using a free text information retrieval system as a nucleus. The important point to remember is that whereas computer costs are going down, staff costs are going up, and western society is in the middle of an information explosion. All these factors point to free text systems.

9. ACKNOWLEDGEMENTS

It is a pleasure to have an opportunity to thank the people involved in the STATUS project at Harwell for many helpful discussions. I would particularly like to mention I. F. Croall and Dr. R. D. Mount.

APPENDIX

FIGURES USED IN TALK

Fig.1.Free Text Information Retrieval
Text Structure

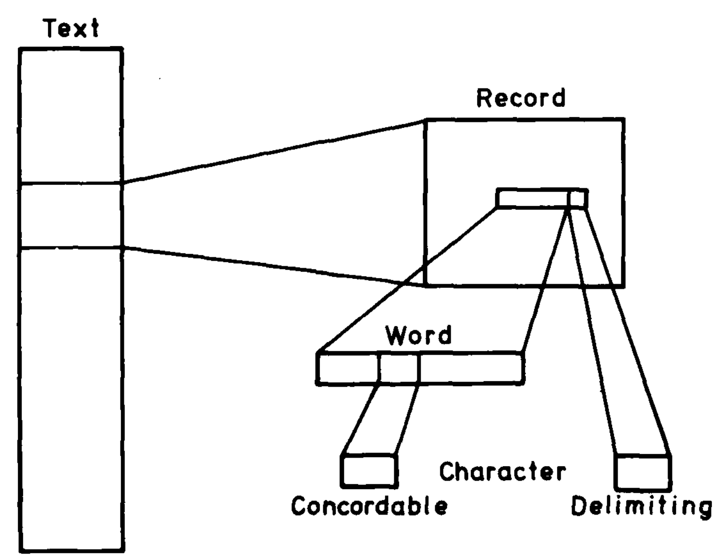


Fig.2.Free Text Information Retrieval
Concordance Structure

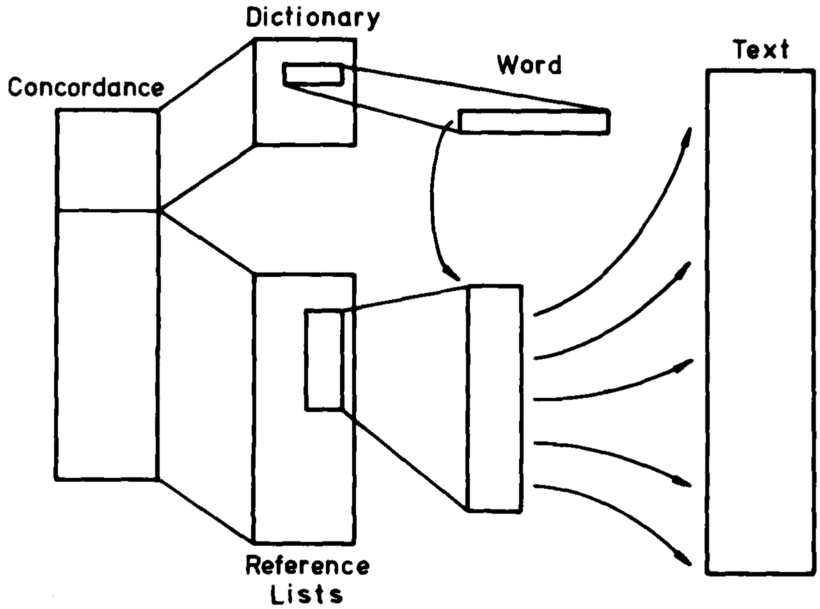


Fig.3.Free Text Information Retrieval
Basic File Structure

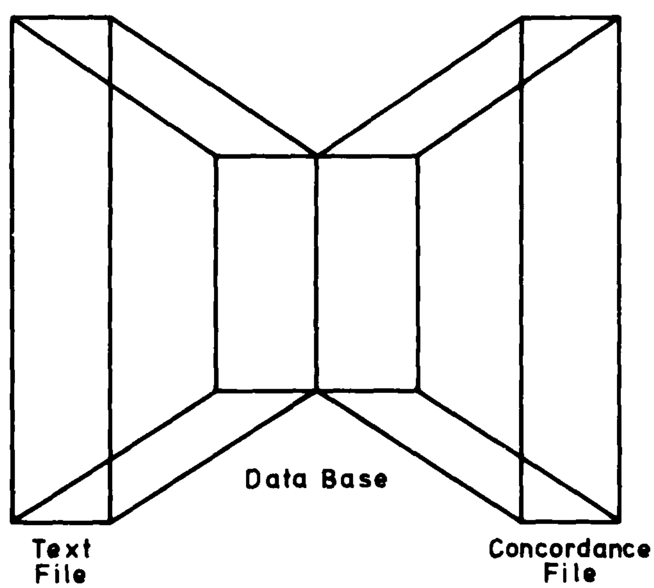


Fig. 4. Full Text Example

Text

Record 1

Thermionic Vacuum Tubes by Appleton

Record 2

Physical Constants by Childs

Record 3

Wireless Receivers by Oatley

Record 4

Physical Principles of Wireless by Ratcliffe

Concordance

Dictionary	Reference Lists
Appleton	1,5
By	1,4 2,3 3,3 4,5
Childs	2,4
Constants	2,2
Oatley	3,4
Of	4,3
Principles	4,2
Physical	2,1 4,1
Ratcliffe	4,6
Receivers	3,2
Tubes	1,3
Thermionic	1,1
Vacuum	1,2
Wireless	3,1 4,4

Fig. 5. Advanced Features of Free Text Systems

- Powerful Search Aids
  - Field Structure in Records
  - Synonyms
  - Thesaurus
- Numeric Capability
  - Range Searching
  - Sorting on Value
- Security
  - By Data Base
  - By Record
  - By Part of Record
- Integrated Systems
  - Tailored Interface
  - Information Capture
  - Post Processing
  - Spool File

Fig.6. Current Application Areas for Free Text Retrieval from Status

- Legal Text
- Bibliographic Retrieval
- Library Management
- Experimental Results
- Chemical and Pharmaceutical Information
- Patents
- Personnel Information
- Accident Reports
- Police Records
- Production Records

Fig.7. Integrated Free Text Management Systems

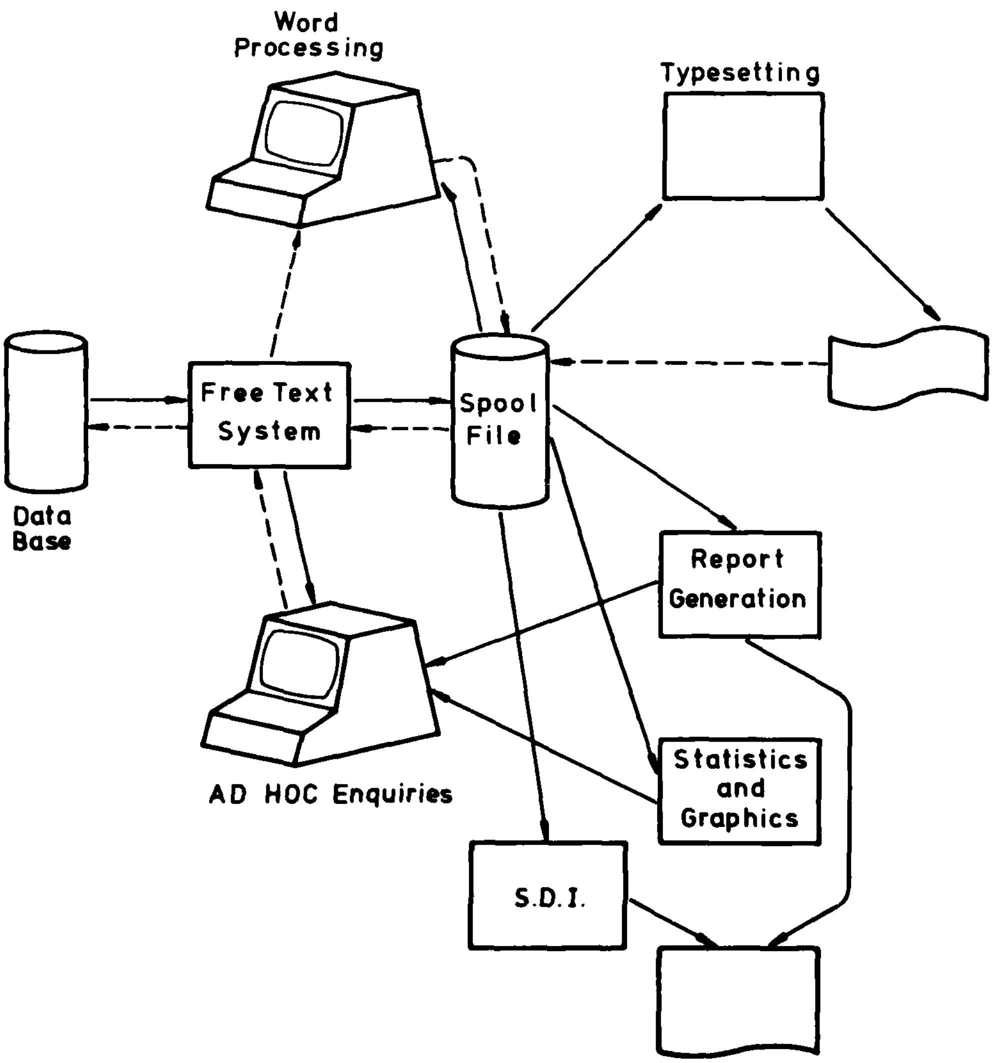
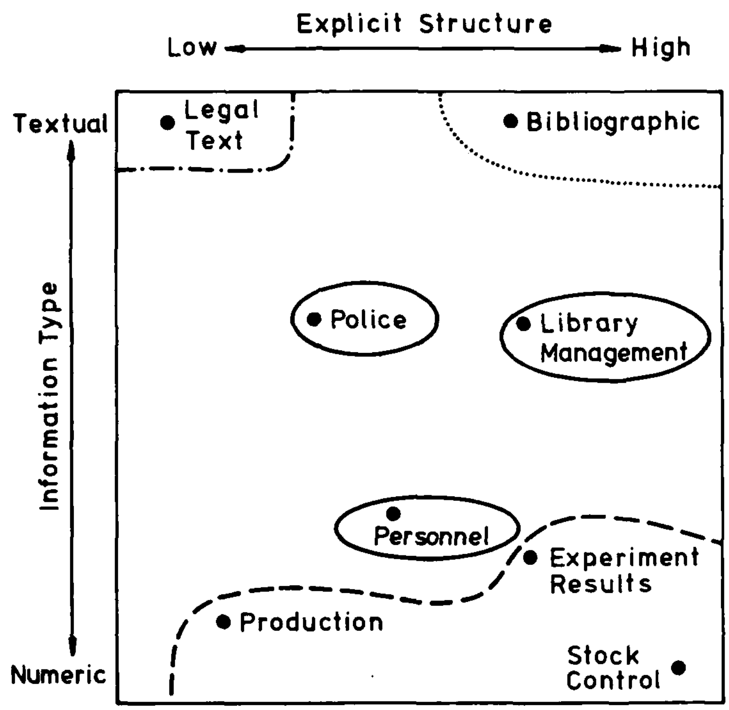


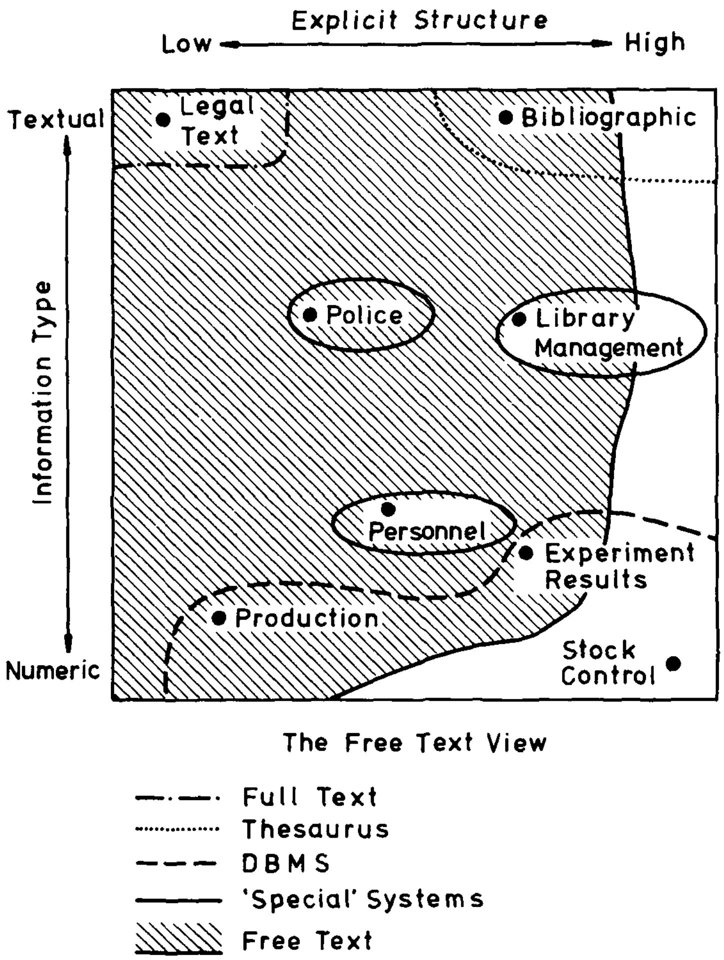
Fig.8.Information Classification



The Traditional View of Application Areas

—·—· Full Text
········ Thesaurus
– – – – DBMS
——— 'Special' Systems

Fig.9.Information Classification



The Free Text View

THE NUMERIC AEROSPACE DATA:
Problems of Evaluation, Handling and Dissemination

G. K. Hartmann
Max-Planck-Institut für Aeronomie
3411 Katlenburg-Lindau 3, P.O.Box 20
Germany

SUMMARY

The main (numeric) data sources QI in science are till now scientific journals, books, tables, and reports ("Data from Literature" - in contrast to data about literature: "Bibliographic Data"). The advent of the new digital measuring and data processing techniques produced a second source QII of fast growing importance (data from measurements, from model calculation by computers, and from statistical inquires - briefly: "Measured Data"). The large data growth rates due to the very recent technical progress give rise to a slow increase of a third source QIII ("Picture Data"). Since the evaluation-, dissemination-, and handling-problems of Aerospace data from QI are very small - and principally solved - compared to those considering QII and QIII they will be disregarded here. Since the problems with QIII are even more complex and more recent than those of QII, they have to be disregarded, too, due to the short time available here. The organic specialized information growth led not only to the beginning of the second data source QII but also to the necessity to redefine or subdivide the terms data acquisition and data processing. For example it became necessary to introduce the new terms "raw data acquisition" and "data preediting" as well as to subdivide the term data processing, in context with data evaluation procedures. In context with dissemination and handling procedures we are now faced with the problems to distinguish between small, medium, and large information systems.Due to the fact that most of the numeric Aerospace Data ("Direct Aerospace Data" from source QII) are non reproducible time series data - often also space dependent - with tremenduous growth rates it will be extremely difficult and time consuming for any scientist ("secondary user") and information system (documentation personnel) not actively working in the specific field, to evaluate, disseminate and handle those data with a high degree of efficiency and reliability. Thus it is very likely that also in the future numeric Aerospace data exchange activities will by far dominate the pure handling activities despite any technical progress.

I. INTRODUCTION

The fact that we have to talk about such general problems in context with Aerospace data is due to the **tremenduous data growth rates (information explosion) which we observe in the industrialized** nations during the last two decades.

There are two principal information problems [3, 4, 5, 6, 7, 8]:

1. Information can be accumulated, time (e.g. human life time) cannot be accumulated.

2. Further information growth rates will increase the tension between structural differentiation (i.e. branching which is a consequence of so called organic growth) and integration (i.e. striving for better overall views, which are a need of our present societies and which are a consequence of new scientific results and new technical tools).

Branching now often leads to an organisational and institutional separation between Data Acquisition, Data Documentation, and Data Processing Activities as well as between Research and Service Activities. This often leads to various unexpected, negative consequences. Documentation activities [1, 2, 8] will play a key role in avoiding and/or reducing the negative separation effects. (Documentation means: the systematic collection of literature in a broader sense, the analysis of the facts and the storage of its principal elements with the intention of providing a fast and complete retrieval of the information upon request.)

The main (numeric) data sources QI in science are till now scientific journals, books, tables, and reports ("Data from Literature" - in contrast to data about literature: "Bibliographic Data"). The advent of the new digital measuring and data processing techniques produced a second source QII of fast growing importance (data from measurements, from model calculation by computers, and from statistical inquires - briefly: "Measured Data"). The large data growth rates due to the very recent technical progress give rise to a slow increase of a third source QIII ("Picture Data"). Since the evaluation-, dissemination-, and handling-problems of Aerospace data from QI are very small - and principally solved - compared to those considering QII and QIII they will be disregarded here. Since the problems with QIII are even more complex and more recent than those of QII, they have to be disregarded, too, due to the short time available here.

Talking about Aerospace data in general and numeric Aerospace data in special we have to distinguish in both cases between the following two categories:

A. Direct Aerospace data: Data that stem from the investigation of the Aerospace.

B. Indirect Aerospace data:
   a) Data about any device for the investigation of the Aerospace
   b) Data about any type of Aerospace transportation and/or communication/navigation system

Here we will only deal with numeric data from category A, respectively QII, and denote them for the

sake of brevity "numeric Aerospace data". Already now they represent the major portion of all the available Aerospace Data.

II. EVALUATION PROBLEMS OF AEROSPACE DATA

1. Problems due to different properties of data

Generally speaking, any information that can be processed by a computer can be regarded as "data". However this general meaning produces many misunderstandings in almost all practical cases when we are faced with more detailed data activities. Thus we have to specify this term [4, 7, 8].

Tables 1a and 1b show the main areas of data documentation and some data filing aspects.

More than 90 % of the numeric Aerospace data belong into the category of time and space dependent non reproducible data, which introduces special problems for any data evaluation and interpretation activity.

Fig. 1 which shows the elements of empirical science will clarify this aspect from another point of view. It becomes clear that all data that belong into the category "reproducible" belong here into the category "conditions controlled by man"; "non reproducible data" belong into the category "uncontrolled conditions". This presentation shows simultaneously that most of the interesting relations deduced from numeric Aerospace data have to be obtained by statistical methods. There are two principal statistical methods of treating large amounts of data a) Frequency distributions, b) Descriptive statistical parameters. An additional problem is introduced by the fact that very many of the just mentioned numeric Aerospace data are time series data. (A time series is a set of observations at different times. Each observation or measurements consists of a quantitative number that describes the measurement as well as of a second number that describes the time at which the measurement was carried out.) If we deal with time series (data), then there are till now no well developed and tested methods of inference (conclusion), applicable to large areas, like those which one knows from analytical statistics. This is mainly due to the following fact: Concerning time series data we find that adjacent observations are mostly correlated, e.g., they are not independent from each other. This so called "series correlation" does not allow the generally used estimations of standard deviations and troubles the "range-estimations" or check of hypotheses. This implies that the evaluation or interpretation of those (non reproducible) time series data requires a much larger and longer practical experience than the evaluation of reproducible data. This is very often not enough considered in context with the planning of those information systems separated from the relevant research institutes and dealing with (non reproducible) time series data. The problems increase further if we have to consider simultaneously stronger spatial variability of the data.

2. Data Acquisition and Data processing problems

Despite the fact that in general numerical data are the most objective information sources and therefore are the most important basis for rational decision finding processes, one must not forget that very likely the majority of decisions and actions are based on emotions, i.e. they depend on subjective information [4, 7, 8]. Together with the tremendous information growth rates, especially for QII and QIII, and the still increasing software-hardware - gap this leads to the increasing use of subjective information. This ultimately implies that decisions reached on a golf course or by drinking a beer with just a second person, for example, might dominate by far those based upon more objective data that stem from specific information systems. In the long run this will strongly affect all kinds of official information channels, especially the large Information Analysis Centers (IACs) [1, 2], the WDCs (World Data Centers) [10], as well as all larger data acquisition systems, e.g., SELDADS [9].

The most measured data from category QII are stored digitally these days and mostly without any explanatory remarks. This "minimum documentation" makes the data only usefull for an extremely small user-community, in general only the principal investigator (PI), who just matched the procedure to his own special needs. These days attempts are made that part of these data - at least those that stem from very expensive experiments or those that are of increasing interdisciplinary interest - is made available by additional documentation procedures for a broader user-community, e.g., with the help of various IACs. This would be a large step beyond the till now predominant data-exchange or cooperation schemes within one specific field of science. Already now there become visible quite a number of problems e.g., the unconscious and/or conscious data misuse by secondary users, data protection problems as well as many administrative and economic problems.

Due to the large data growth rates we have to subdivide not only the term Data Acquisition - see fig. 2 - but also the term Data Processing - see fig. 3. Fig. 2 shows also the meaning of the newly defined term Data Preediting and shows further a new term, defined by the author namely: Datography, which is used in analogy to the term Bibliography. (Datographies give that minimum of information - e.g. type of data, accuracy, calibration, sampling rates, formats etc. - which is required that "secondary users" can use numeric data - generated by principal investigators - with minimum chances of misinterpretations.)

In the context with all Data preediting activities Video-Graphic-Communication and Documentation Systems (VIGRODOS) will play a very important role [4, 7, 8].

Special data documentation steps or datographies have to be performed on each step from the data acquisition procedure to the final data processing (data evaluation and interpretation). The broader the "secondary" user-community will be and the less "background" information these user have, the more detailed the datographies have to be written. Very detailed datographies however lead to two problems on the one hand it is very time consuming to write those detailed descriptions and on the other hand it also very time consuming to study very detailed descriptions. The less detailed a datography the smaller the number of additional users!! Thus we have to envisage a compromise between the maximum wanted number of "secondary" users and a necessary minimum text for the datographies. This will lead in the documentation area to a number of principal problems. If information-systems deal with large amounts of data and they want to supply various related services, then in addition to datographies they have to introduce another "activity", which is denoted as "Data-Catalogue" by several information handling agencies, e.g., by WDC-A in Boulder, Colorado, USA (fig. 3, Table 5).

3. Some examples for the production of large amounts of data

a) Space Transportation System (STS)
The STS-program plans about 500 flights between 1979 through 1991 for the Space Shuttle and the Spacelab. A maximum data flow of 5×10^7 bits per second is envisaged. This corresponds the "production" of a "DIN A4" file with 500 pages densly printed with text or tables. Thus for a 7 days flight period with 100 % duty cycle we could get 600 000 of such files with information from several scientific domains. After 12 years - 500 flights - this would be 30 000 000 of such files with 500 pages each, almost twice as much as the largest library of the US Congress (Library of Congress) stores today concerning the number of volumes.

b) Other programs
The just mentioned STS-program is only out of a larger number of data intensive scientific programs which are already in operation or which are under preparation, e.g. satellite-meteorology, Earth resource satellite programs, Aerospace programs etc.
In the 1976 report "Geophysical Data Centers: Impact of Data-Intensive Programs", the Geophysical Research Board of the U.S. National Academy of Sciences' National Research Council evaluated the impact of large scale geophysical programs on the U.S. and WDC-A. They found that the National Climatic Center had about 77.000 reels of digital magnetic tape. The National Space Science Data Center had about 41.000 reels, and the National Geophysical and Solar Terrestrial Data Center had about 600 tapes from only a few years of data acquisition of this type. These Centers, together with the National Oceanographic Data Center, had millions of feet of film records and cubic feet of paper documents. Further, a sampling of some 14 national and international data-collection programs then in progress indicated that these data loads would increase by some 10^{14} "bits" (equivalent to 2.5 million digital magnetic tapes). This is in addition to the continuing important Center roles of archiving data in photographic, graphical, and paper tabular forms - truly a formidable task [4].

c) Directly application oriented programs
For example the oil-companies need large amounts of geophysical data for their resource-technologies. Some of the very large companies have thus several millions of tapes with data. The progress of measuring techniques as well as the progress in evaluation techniques as well as the specific characteristics of time series data imply that these tapes will be processed not just once but several times spaced by several years. This leads to serious documentation problems, which have to be solved and which are not only due to the large amount of data tapes but also to their technical maintenance.
However there is nowadays in other fields than physics or geophysics a "production" of large amounts of data visible, due to the very recent developments in electronics etc. and due to very many model-calculations which are performed with the help of modern computers.

III. DISSEMINATION AND/OR HANDLING PROBLEMS

1. Data Exchange (exchange of measured data, source QII)

One of the most important activities in the Aerospace domain was and is the international cooperation of institutes, the performance of joint experiments and the exchange of data between institutes that performed the same or similar measurements. For a very large amount of these data international cooperation is one of the main postulates for an effective data interpretation. This mutual exchange dominated till now the whole data documentation scene in geophysics (Aerospace). This exchange took place directly between research institutes or via the WDCs etc. This was in contrast to most other areas in physics where the data, for example from laboratory physics, were available not only for the specialists but also for a wider user-community, that was often only interdisciplinary related to the subject. Just now there are some additional attempts initiated to try to make available parts of the geophysical data - at least those that stem from very expensive experiments or those that are of increasing interdisciplinary interest - to a broader user-community, e.g., with the help of newly established IACs. However, this produces not only economic and administrative as well as organizing problems but also a principal one as already mentioned. This implies also that one of the major tasks of the new information systems is "Advice" for secondary users who request non-reproducible data. This advice is getting the more difficult the more we deal with time series data.

2. Other problems

One major problem arises due to the fact that we mostly have to deal with various generations of data as is shown in table 2 presenting various steps of empirical, experimental sciences.

A further problem is due to the fact that large amounts of data can stem from principally different data sources and generations - table 3 - and can be used for different purposes, for example Research or Service Activities - table 4.

Another very important problem arises these days due to the fact that we have to deal with three types of information systems: small, medium, large - see table 5.

It seems important to mention that we have not only to distinguish between the various services that are supplied by the information-systems but also to keep in mind that we also generally deal with different "generations" which lead to different "main tasks" for the systems.

In the first generation when a new field of research starts the main task of a system is the collection of data and to supply the information who is doing what and where. The data should in general be preedited but not compressed. We deal only with fairly small amounts of data. In the second generation when the field of science is well established we deal with a lot more information. Then the main task for the information-systems is the documentation. The data are mostly distributed in a slightly compressed form. In the third generation we deal with very large amounts of information (literature and data) and a wider distribution of "raw data" and/or preedited data by information centers is impossible. The data have to be either compressed or small portions have to be selected before they are mailed. Thus the compression of data and the compilation of data catalogues become now the main task besides the older already "classical" tasks.

IV. CONCLUSIONS

Due to the fact that most of the numeric Aerospace Data (Direct Aerospace Data from source QII) are non reproducible time series data - often also space dependent - with tremenduous growth rates it will be extremely difficult and time consuming for any scientist ("secondary user") and information system (documentation personnel) not actively working in the specific field, to evaluate, disseminate, and handle those data with a high degree of efficiency and reliability.

Thus it is very likely that also in the future numeric Aerospace data exchange activities will by far dominate the pure handling activities despite any technical progress.

V. REFERENCES

[1] Abbel, R., 1978
Dateninformation und -dokumentation -Einführung und Überblick-
In: Deutscher Dokumentartag 1977 (03.10.-07.10.) in Saarbrücken, S. 133 ff.
Bearb.: Mathilde von der Laake, Peter Port (DGD-ATS), Verlag Dokumentation Saur KG, München, New York

[2] CODATA, 1977
CODATA (Committee On Data For Science and Technology)
CODATA Secretariat, 51 Boulevard de Montmorency, 75016 Paris, France

[3] Hartmann, G. K., 1977
Why do we need data, what principal data evaluation problems have to be considered?
Herausgeber: ZAED, Kernforschungszentrum, D-7514 Eggenstein-Leopoldshafen 2

[4] Hartmann, G. K., 1978
The information explosion and its consequences for data acquisition, documentation, and processing
World Data Center A. Solid Earth Geophysics, Report SE-11; Solar Terrestrial Physics Report UAF-65, U.S. Department of Commerce, NOAA EDS, Boulder, Colorado 80302, USA

[5] Hartmann, G. K., 1978
Betrachtungen über Energie und Information. ntz, Bd 31, Heft 8, S. 535-536

[6] Hartmann, G. K., 1978
Informationsexplosion - Grenzen, Folgen, Handlungsmöglichkeiten
Teil 1: Energie und Information. ntz, Bd 31, Heft 8, S. 568-571
Teil 2: Probleme des Informationswachstums. ntz, Bd 31, Heft 9, S. 654-657

[7] Hartmann, G. K., 1978
Numerical Data Growth Rate Problems
6th International CODATA Conference, Santa Flavia, Italy, May 22-25, 1978 (Proceedings in press)

[8] Hartmann, G.K., 1979
Documentation problems of large amounts of measured data
Edited by FIZ 4; Kernforschungszentrum, D-7514 Eggenstein-Leopoldshafen 2, Germany (In press)

[9] Williams, D., 1976
SELDADS: An operational real time solar terrestrial environment monitoring system
NOAA Techn. Report ERL-SEL 37, U.S. Department of Commerce, Boulder, Colorado 80302, USA

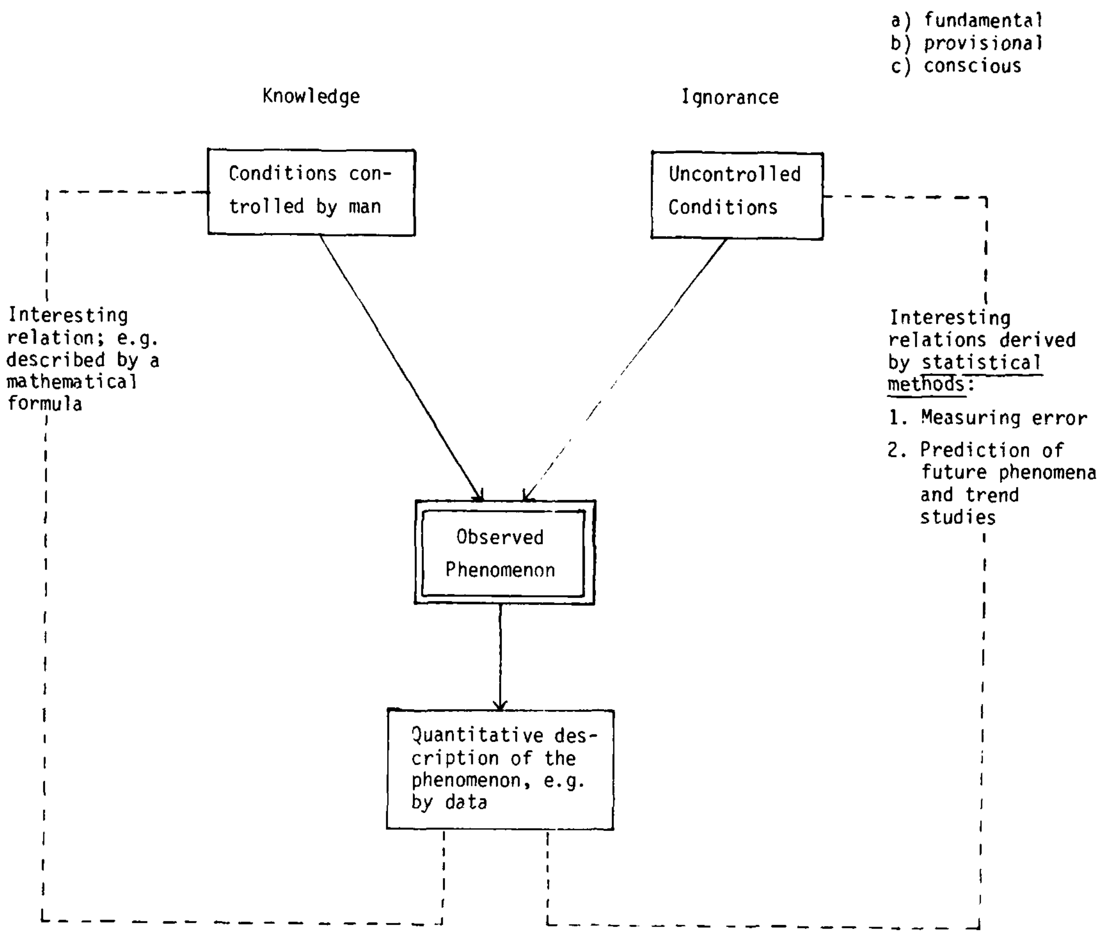
[illegible] NOAA EDS, 1978
[illegible]eedings of the NGSDC Pilot Workshop on the information explosion and its consequences for [illegible] acquisition, documentation, and processing
[illegible] NGSDC, Boulder, Colorado 80302, USA

VI. ACKNOWLEDGEMENT

The author thanks the Max-Planck-Institut für Aeronomie, represented by the director, Prof. Dr. W. I. Axford, for the support. Many colleagues he thanks for critical discussions, especially at the WDC-A-STP at Boulder, Colorado, USA and at FIZ 4 at Karlsruhe, Germany.

VII. FIGURES AND TABLES

Fig. 1: Elements of empirical science



Remarks:

A. The ignorance about the uncontrolled conditions can be of three different types:
   a) fundamental type - if for instance the influence of a parameter cannot be determined in principle (time of nuclear solitting)
   b) provisional type - if we can hope that by the use of new information we can expand the range of controllable conditions
   c) conscious type - if we know of the existence of some uncontrolled conditions and disregard them, since regulating them is either of no interest or technically not feasible or too expensive.

B. There are two principal statistical methods of treating large amounts of data.
   1. Frequency distributions
   2. Descriptive statistical parameters

Fig. 2: From data acquisition to processing

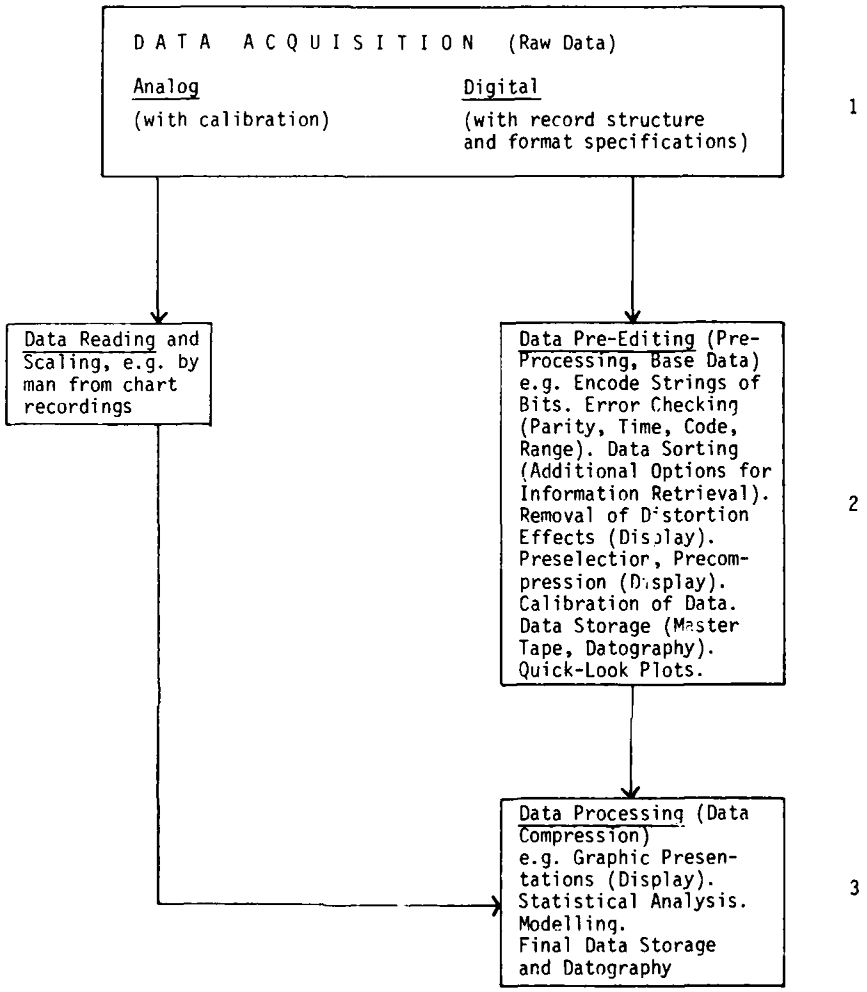


Fig. 3: Various steps of data processing

1. Data collection (only preedited data are collected)

2. Data collection documentation (compilation of datographies and preparation of data catalogues)

→ Archives and Data Bases

3. Data compression
   a. Graphical representation
   b. Statistical analysis
   c. Models

→ Publications (literature)

The third step, the data compression is again subdivided with an increasing degree of complexity.

Table 1a Main areas of data documentation

I. Alpha-numerical data documentation

Systematic filing of facts
Examples: Bibliographic files, address files, merchandise lists, geographical distribution of insects, etc.

Remark: Hartmann, 1977, used the term non-numerical insteaJ of alpha-numerical, which produced misunderstandings in context with text- and picture documentation. [3]

II. Numerical data documentation

1) Physical-chemical data, i.e. properties of matter and substances;
   Physical constants
   Data are independent of time and space
   Data are reproducible since the conditions of the first measurement can be reproduced for any subsequent measurements

2) Raw material- or Engineering data
   Data are dependent on the production technology of the material, its preliminary treatments and its "geometry"

3) Environmental data
   a) Data from natural matter and from nature, i.e. from the three Geospheres (1. Lithosphere, 2. Hydrosphere, 3. Atmosphere), the Biosphere, and the Interplanetary Space, e.g.: Geosciences, Biosciences, Medicine, Astrosciences, numerical Aerospace Data. Data are practically all non reproducible - conditions of the first measurement cannot be reproduced for any subsequent measurement - and are time and/or space dependent.
   b) Data from Economics, Social and Political Sciences, and from Arts and Humanities. Example: The statistical almanac issued by the Federal Republic of Germany give data which are mostly time and space dependent and which are non-reproducible.

Table 1b Some data filing aspects

I.	Properties; characteristics e.g. time and/or space dependent	(1. Step)
II.	Origin; domain	(2. Step)
III.	Measuring methods	(3. Step)
IV.	Utilization, application	(4. Step)

Table 2 Various steps of empirical, experimental sciences

First generation

(Simple Data Technologies, SIDAT)

Data Acquisition (DA), Data Preediting (DP), Data Documentation (DD), Data Processing (DPr), Data Exchange (DE), and Data Handling (DH) are done at the same location and by one institution. Here we deal with a relativ small amount of data and mainly with pure research activities.

Second generation

(Intermediate Data Technologies, INDAT)

The specialisation and "branching" processes due to the organic information growth imply that the steps DA, DP, DD, DPr, DE, and DH, are getting complexer and are slowly institutionally and organizingly separated, i.e., they occur generally at several different locations, being accomplished by several different organisations. Here we deal in general with medium amounts of data. We still have more research than service activities. (Service means in this context the foundation of so called information systems for data handling.)

Third generation

(Recent Data Technologies, REDAT)

The above mentioned steps DA, DP, etc. are mostly completely separated, spatially, temporally, and organizingly. Here we deal with a large amount of data. Generally there are more service activities required than research activities. (Service for Research and Application.) Foundation of Information Analysis Centers (IACs), Data-networks, and Data Base Management Systems (DBMS).
Here we reach at present the limits of the human information processing capabilities and capacities.

Fourth generation ????

There might later occur a leap across the just mentioned limits, possibly initiated by new ways of information experience and information understanding which amongst others will be supported by VIGRODOS-Systems (Video-Graphic-Communication and Documentation System; term defined by the author).

Table 3 Various "technical generations"

By using large amount of data one has to take into consideration various "generations" and sources.

A. Very many different, short-time sources

↓

Generally three technical generations have to be considered:

a) SIDAT (see Table 2)

b) INDAT (see Table 2)

c) REDAT (see Table 2)

↓

Are they comparable and how?

↓

Compilation of a data catalogue (datography etc.)

↓

Edition of an "electronic book" with the data (e.g. cassette for a mini-computer with graphic display)

↓

Multidimensional "still picture presentation e.g. on a television screen

B. Very few, long-time sources which produce time series data

↓

Generally three "time generations" (old, medium, and young) have to be considered. They are represented by:

a) Very coarse models

b) Improved models

c) Fine scale models

↓

Where lies the "data noise" threshold?

↓

Compilation of data for selected short time periods, with special datography

↓

Graphical representation of selected data (e.g. with cassette on a mini-computer with graphic display)

↓

Multidimensional "motion-pictures" (film, holograms, etc.) on television screen using also video mixing techniques

Table 4 Data Usage

Large amounts of data might be used for the following two purposes:

A. RESEARCH, (Science)

a) by the principal investigator (PI)

b) by the secondary user

↓

about 60 % of the data can be directly processed using computers and filter techniques according to "apriori knowledge" (redundant). Here we deal with the application of technical intelligence.

↓

about 40 % (rest) requires the direct interaction of human beings and their decisions, i.e. organic intelligence.
Only this part leads to real new knowledge.

B. SERVICE, (operational activities)

a) by the data producer

b) by the secondary user

↓

about 100 % of the data can be directly processed using computers due to "apriori knowledge"

↓

The data are used for prediction purposes and for improving the prediction schemes (the longer the time series the more we are moving from "frequency distributions" to "probability distribution). Only in this context we obtain really new information.

Table 5 The three most important types of information-systems and some of their characteristics

		Small information-systems	Medium information-systems	Large information-systems
I.	Main-Service	preedited, uncompressed data	slightly compressed and selected data	data-catalogues (preedited and compressed data only on special request. Store period > 10 years).
II.	Main-Technique	SIDAT + INDAT** (VIGRODOS)	INDAT + REDAT** (Data-Networks)	REDAT** (DBMS and Data-Networks)
III.	Main-User	Many individuals	Research groups	Large organizations
IV.	Number of users	Very large	Medium	Small
V.	Budget	Public and/or private	Public and/or industry	Public, industry, military
VI.	Store period	< 10 years	≈ 10 years	> 10 years *
VII.	Amount of data	Small	Medium	Large
VIII.	Flexibility	Large	Medium	Small
IX.	System time constant	Small	Medium	Large
X.	Expenses	Small	Medium	Large
XI.	Computer	Small	Medium	Large
XII.	Accumulationproblem	Large	Large	Large
XIII.	Location	Decentralized	Decentr. and centralized	Centralized
XIV.	Ratio of formal to informal information	Small	Medium	Large
XV.	Ratio of research to service	Large	Medium	Small

** see also table 2

* the uncompressed, preedited data are only stored longer (>) than 10 years, if they were requested during this period. Otherwise only the data-catalogues and the compressed data will be further stored.

The sign ≈ means "approximately", < means "smaller", > means "larger"

THE ROLE OF WORLD DATA CENTERS AND THE LUNAR AND PLANETARY INSTITUTE IN THE INTERNATIONAL EXCHANGE OF LUNAR AND PLANETARY DATA

Frances B. Waranius
Lunar and Planetary Institute
3303 NASA Road One
Houston, TX 77058

SUMMARY

The success of many lunar and planetary investigations has resulted in the accumulation of a mass of data in a myriad of formats and medias. The application of these data to comparative planetology, origin of the solar system, and potential industrial applications of space, has made it necessary for scientists from many disciplines to access the data.

Although much of these data are archived at the National Space Science Data Center (NSSDC) and made available to scientists world-wide through the World Data Center-A for Rockets and Satellites (WDC-A-R&S), the collections of these data are so diverse that it is often difficult to select what is needed based on catalog information alone. In addition, there are other valuable data sources which are not at NSSDC.

The Lunar and Planetary Institute (LPI) has one essential goal...to bring the user and the data together. As a research support organization operated by the Universities Space Research Association (USRA), the Institute through its active Visiting Scientist Program, a balanced program of study workshops and topical conferences, and organized and supervised data collections, has assisted scientists, educators, and students to review, study, and obtain the data necessary to the pursuit of their research.

The concept of World Data Centers for the international exchange of data was initiated during the organization of the International Geophysical Year (IGY). Prompted by the problems which occurred in obtaining data from the observations of the Second International Polar Year (1932-33) the Comite Special pour l'Annee Geophysique Internationale (CSAGI), in September 1955, resolved that "observational data to be exchanged in accordance with the IGY programme shall be available to scientific institutions in all countries" and accordingly authorized "the establishment of at least three IGY World Data Centres, of which one will consist of a number of parts. Each Centre will be international in the sense that it will be at the service of all countries and scientific bodies." As a result of offers by the United States and USSR to establish centers for all IGY disciplines, it was decided to establish WDC-A in the United States, and WDC-B in the USSR; a third center, WDC-C is comprised of a number of discipline centers in various nations.

This new form of international cooperation--international exchange of data through World Data Centers organized for the IGY--was found to be very effective. Instead of having to address themselves to many national organizations, scientists could receive data necessary for scientific work directly from the WDCs. Thanks to the collection and exchange of data through WDCs, it became possible to investigate phenomena on a planetary scale and to study the interdisciplinary relationships among various phenomena.

The World Data Centers collect data and publications in the following programs:

- Solar-Terrestrial Physics
- Rockets and Satellites
- Meteorology
- Oceanography
- Glaciology
- Solid Earth Geophysics

World Data Center A (WDC-A) consists of the WDC-A Coordination Office within the National Academy of Sciences and seven subcenters at scientific institutions in various parts of the United States. Four of the WDC-A subcenters are operated by the Department of Commerce's National Oceanic and Atmospheric Administration, one by the National Aeronautics and Space Administration (NASA), one by the University of Colorado, and one by the U.S. Naval Observatory. WDC-A follows the guidelines set down in the International Council of Scientific Unions (ICSU) Fourth Consolidated Guide to International Data Exchange through the World Data Centres (1979). The Guide is available from any of the World Data Centers, their subcenters, or the National Academy of Sciences. (1)

The Center with the prime responsibility for the exchange of lunar and planetary data is WDC-A for Rockets & Satellites (WDC-A-R&S) which is located at the Goddard Space Flight Center in Greenbelt, Maryland. Because of the contiguous location of WDC-A-R&S and the National Space Science Data Center (NSSDC) which was established by NASA in 1965, a close cooperation exists in obtaining data for scientific requests. The same policies apply to requesters both within and outside the United States. NSSDC was commissioned with the responsibility to collect the scientific data obtained from space probes, satellites, sounding rockets, stratospheric balloons, and high altitude aircraft. It is responsible for the acquisition, organization, storage, retrieval, announcement, and dissemination of this information. The types of data archived in the WDC-A-R&S subcenter include:

- Reports of sounding rocket launchings
- Reports of satellite and space probe launchings
- Detailed descriptive information on spacecraft experiments
- Scientific reports of limited distribution on results of experiments
- Orbital elements and ephemerides which are of scientific interest and value

NSSDC has reduced and analyzed data from almost 800 satellite experiments covering such disciplines as lunar and planetary photography; x-ray; gamma-ray, RF, UV, and visible astronomy; cosmic rays; magnetospheric energetic particles, plasma, and electromagnetic fields; interplanetary energetic particles; ionospheres; and meteorology and IR data. (2)

The number of experiments for which data were available at NSSDC grew from a total of 105 in 1967 to 225 in 1969. With the advent of the Apollo program, the total number of experiments climbed until in 1974 there were 548 experiments. Recent tabulation shows that this total rose to 793 in 1978. (3) Of this total, 116 experiments are from 18 planetary exploration spacecraft and 175 experiments are from 29 lunar exploration spacecraft.

Planetary data available to researchers include more than 150,000 frames of photographic data, 600 reels of magnetic tape, 110 rolls of microfilm, and 1,400 sheets of microfiche. Lunar data available includes more than 152,000 frames of photographic data, 3,100 reels of magnetic tape, 1,500 rolls of microfilm, and 140 sheets of microfiche. (4)

Thus, even though the mechanism for the international exchange of space science data did exist through the WDCs, it was recognized that the body of knowledge resulting from the massive technological and scientific effort known as the Apollo Program would create special demands for interaction among the world's scientific community.

In a speech at the Manned Spacecraft Center (now the Johnson Space Center), on March 1, 1968, President Lyndon B. Johnson proclaimed:

> "As a further step toward joining hands with the world's scientific community, I want to announce that we will build facilities here in Houston to help the world's scientists work together more effectively on the problems of space. We are going to have a new Lunar Science Institute alongside this great Center. ... This new Lunar Science Institute will provide new means of communication and research for the world's scientific community. It will help unite the nations for the great challenge of space."

And so the Lunar Science Institute (LSI) began. It was formally established on October 1, 1968 by a NASA contract with the National Academy of Sciences (NAS). Renovation of a vacant old mansion which had been the home of a Texas millionaire in the 1930's was begun and on October 27, 1969, the LSI moved into its permanent quarters.

Since it was necessary to have an organization to permanently manage the LSI and possibly additional space-related facilities, the NAS convened two meetings of university representatives to discuss a new university consortium. Thus, the Universities Space Research Association (USRA) was incorporated in the District of Columbia on March 12, 1969; the first meeting of USRA institutional representatives was held in Washington, D.C. on June 9, 1969, and USRA assumed the management of LSI under contract to NASA on December 11, 1969. To accomplish the goals of international cooperation outlined by President Johnson, the LSI initiated several programs:

Visiting Scientist Program
Conferences, Symposia, Workshops
Publications
Library

The Visiting Scientist Program is an effective means of bringing together scientists from both domestic and foreign universities and industries who wish to begin or continue work on the results of the lunar and planetary missions. The Visiting Scientist Program provides for experimental and theoretical research and educational program development in the lunar and lunar-related sciences and in comparative planetology. Experimental research programs proposed by appropriately qualified individuals may be accommodated through utilization of laboratory facilities at the Johnson Space Center. Appointments vary in length from a week to as long as a year, with option for renewal at the discretion of the Institute's Director. Appointments range from the Associate Professor level through Post-Doctoral appointments, Graduate Fellows and Undergraduate Interns.

To promote the dissemination of information and research results, an active schedule of conferences, topical symposia, study workshops, and educational short courses, are conducted by the Institute's Symposia Office. The results of these meetings are edited and published by the Publication Office. Chief among these, of course, is the annual Lunar Science Conference (now the Lunar and Planetary Science Conference). This major conference which regularly brings together scientists from all over the world is held under the joint sponsorship of the Institute and the Johnson Space Center. The average attendance at this week-long meeting is usually about 650. Two major publications which represent the greatest single summary of the results of the lunar and planetary investigations are the abstracts of papers presented to the conference and the proceedings which each year have been a three-volume work averaging about 3,000 pages. The first of these conferences was held in 1970; the most recent, the Tenth, was held in 1979.

Recent topical conferences have included a Workshop on the Thermodynamics and Kinetics of Dust Formation in the Space Medium, Conference on Origins of Planetary Magnetism, Workshop on Ancient Crusts of the Terrestrial Planets, Workshop on Remote Sensing of Volcanic Gases, and a Workshop on Glass and Ceramics Industry in Space Based on Lunar Materials. Each of these conferences results in a variety of publications: an abstract volume prepared before the meeting, and a proceedings volume which may be published as an independent monograph or as a special volume in a journal, or in a workshop summary which is distributed in the Institute's Contributions series.

A library to provide literature support to the visiting and staff scientists at the Institute and JSC was part of the original organization of the Institute. It contained a basic collection of geology and astronomy texts and journals and provided a full-range of library services to the community. In 1973, the collection of Apollo data consisting of the photography, the cartography, and supporting documents which was housed at the Johnson Space Center was transferred to the Institute and a Lunar Data Center including the Library, the Photo/Map Library, the Lunar Sample Information Library, and the Geophysical Data Center was established.

Although NSSDC remained the major distribution center for lunar data, it became apparent that the specialized and in-depth knowledge of the collections and workings of the data resulting from the missions, which the experienced staff of the Institute had acquired, would be useful to the scientist, educator, or student attempting to access the data.

For example, a knowledge of the history of the organization of the lunar sample collection is highly valuable to properly access the information available at the Johnson Space Center and at NSSDC.

The Lunar Samples are a data set which presented some unique and exciting challenges. The Curatorial Facility at the Johnson Space Center is responsible for the use and preservation of the samples.

Each sample has been carefully documented and a record of the allocation of the samples, the experiments performed, and the results of those experiments are filed in a data pack at this Facility. As the samples of each mission were returned to Earth, a special team called the Preliminary Examination Team (PET) carefully opened each sample bag collected by the astronauts and did a visual and physical description of the sample. A series of photographs was taken in the laboratory to show all six faces of each sample. These macro descriptions done by members of the PET and at least one of the photographs were published in a series of Lunar Sample Catalogs.

A method of cataloging the samples was devised. In general, a five digit generic number was assigned to each sample. The system of numbering evolved through the Program and became more significant in the later missions. The general scheme is:

Apollo 11	10abc
Apollo 12	12abc
Apollo 14	14abc
Apollo 15	15abc
Apollo 16	6xabc
Apollo 17	7xabc
Luna 16	21abc and/or 11016
Luna 20	22abc and/or 20abc
Luna 24	24 abc

For Apollo 16 and 17 the second digit, "x" is indicative of the station at which the sample was procured. For Apollo 15, 16, and 17, the last digit "c" is indicative of sample types such as 0=Unsieved Fines; 1=less than 1 mm fines; 2=1-2 mm fines; 3=2-4 mm fragments; 4= 4-10 mm fragments; and 5-9=rocks and greater than 1 cm fragments. A full description of the numbering system for the samples is contained in the Lunar Sample Information Catalogs for each mission. (5)

As each sample was divided and allocated to investigators for various kinds of analyses, a daughter sample was designated by a comma and a sequential number following the generic number. Thus 15455,54 is a section of Apollo 15 sample 15455. As the samples were divided, cutting diagrams were drawn so that the history of the sample would be documented.

The Lunar Sample Information Library at the Institute contains these catalogs, a mini data pack consisting of the sample photographs, description, cutting diagrams, and in some instances, photomicrographs, and many other documents which have been written on specific sample types by the various investigators working at the Curatorial Facility.

To maintain a record of the actual analyses which were obtained from the experiments conducted by the researchers, the Curator's Office maintains a "Lunar Sample Analysis Data Base." This data base contains the analytical data for each sample as reported in a selected set of published literature. The base is arranged by sample number and in the first section records each complete analysis plus the authors and bibliographic reference in which the analysis is published. The second part also arranged by sample number, lists the elemental analyses recorded for each sample along with the accession number of the bibliographic reference in which the analysis is reported. The third section is the bibliography arranged by accession number.

To obtain information about a particular sample, a researcher might consult the mini-data pack at the Institute, run a search for a particular element on the Sample Analysis Data Base, conduct a literature search through the Lunar and Planetary Bibliography maintained at the Institute, and then prepare a proposal for submission to the Lunar and Planetary Review Panel. Based on the results of his investigations of the data available, the researcher could then query NSSDC for the specific information he would need to support his research effort.

Another significant set of data which resulted from the scientific packages left on the lunar surface is called the Geophysical Data. Although these data are archived at the NSSDC, its usefulness to the researcher is enhanced by a review of what the data set contains.

The Apollo Lunar Surface Experiments Package was a completely self-contained science station deployed and activated by the Apollo astronauts and left on the lunar surface. The ALSEP collected scientific data on the lunar surface and transmitted the data to earth where the information was collected as part of the ALSEP support operations conducted at the Johnson Space Center. A forerunner of ALSEP, known as the Early Apollo Scientific Experiment Package (EASEP) was deployed by the Apollo 11 crew.

Experiments included in the package on various missions were: Passive Seismic Experiment, Active Seismic Experiment, Lunar Surface Magnetometer, Solar Wind Spectrometer, Suprathermal Ion Detector, Heat Flow, Charged Particle, Cold Cathode Gage, Lunar Ejecta and Meteorites, Lunar Seismic Profiling, Lunar Mass Spectrometer, Lunar Surface Gravimeter, and Dust Detector.

Scientific analysis of ALSEP data was accomplished by NASA contracts with specific investigators, and these contracts stipulated the archiving of analyzed data. To ensure proper data archiving, JSC management created the Geophysical Data Evaluation Working Group. The Group was asked to study the data processing and make recommendations on the archiving and distribution most appropriate for the present and future needs of the community. The data resulting from these experiments are collected and archived at NSSDC.

To make the data archived as accessible to the scientist as possible, the Institute was designated a "Geophysical Data Subcenter" by NASA. A copy of the geophysical data on microfilm and microfiche stored at NSSDC was incorporated, along with supporting documentation, into the Geophysical Data Center at the Institute. A researcher can review the data at the Institute, select what he needs and request just those subsets of information from the NSSDC.

Since much of the data archived is in the form of magnetic tapes, the Geophysical Data Curator at the Johnson Space Center undertook the project to prepare computer programs and associated user's guides which would make it possible to read these magnetic tapes. These programs and guides are available from the Geophysical Data Curator at the Johnson Space Center. (6,7)

The largest data suite resulting from the various lunar exploration programs is the photography. This data suite actually began with the Ranger series on July 31, 1964 and ended with Apollo 17 on December 19, 1972. The photography exists in many formats and media ... 16 mm, 35 mm, 70 mm, negative and positive transparencies, in color or black and white, stereoscopic close-ups and orbital panoramas. This data suite is one of the more frequently requested from NSSDC. However, to prepare a request for data, it is often most helpful to the user to visually review the photography in the various formats available.

The Photo/Map Library at the Institute has a complete collection of the Ranger, Orbiter, Surveyor, and Apollo photography of the Moon and various photographic indexes which provide access to the collection. It also has a selection of Zond and Luna photography. The collection is particularly user-oriented with the photographs arranged in notebooks, kardex files, or flip-files allowing the user to browse through the collection easily. Some selected sets of the photography have been assembled ranging from an approximate 100-item set useful to the educator or student, to the full-range of camera clicks taken on a particular mission. Personnel well-versed in the collection can assist the researcher to locate the items needed and then request them from NSSDC. If a researcher is unable to come to the Institute, the staff of the Photo/Map Library can respond to letter and telephone requests and make recommendations of those photos which would best assist the off-site researcher in his studies.

Another significant data set, which has been part of man's knowledge of the Moon since the days of Galileo, is the collection of lunar cartography. From the early hand-drawn efforts of Galileo, Hevelius, and Riccioli through the 1:1,000,000 scale set of lunar astronautical charts (LAC) prepared by the U.S. Air Force Aeronautical Chart and Information Center (ACIC), man had been dependent on telescopic observation and photography for the details of his lunar maps. With the return of over 1,654 high-quality lunar photographs from the Lunar Orbiter flights in 1966-1967, man's 357-year dependence on the telescope for lunar mapping ended.

Several significant maps on various scales were prepared during the period between 1966 and 1972. The small scale maps (1:2,000,000 and smaller) have served for reference and planning purposes and have been used as the basis for many specialized maps prepared in support of the Apollo missions. Among these are:

- Lunar Planning Chart (LOC) series. 1:2,750,000. 1969-1971
- Lunar Equatorial Zone Mosaics. 1:2,500,000. 1968-1969
- Lunar Earthside, Farside and Polar Charts (LMP). 1:5,000,000. 1970 (new edition issued 1979)
- USAF Lunar Reference Mosaic (LEM-1 and LEM-1A). 1:5,000,000 and 1:10,000,000. 1966-1967
- Lunar Earthside Maps, Topographic. Scale varies. 1962-1964

In the medium scale (1:250,000 - 1:1,999,999) the following maps are representative:
Lunar Astronautical Chart (LAC) series. 1:1,000,000. 1965-1967. (42 sheets)
Apollo Intermediate Chart (AIC). 1:500,000. 1965-1967. (20 sheets)

Production of large scale lunar maps (1:250,000 and larger) became possible with the availability of spacecraft photography at scales allowing detailed description of lunar features. Series of maps were prepared from Orbiter, Surveyor, and Ranger photography. These maps were used in the studies to select the Apollo landing sites.

In 1973, NASA in conjunction with the Defense Mapping Agency (DMA) commenced a mapping program that would lead to the production of a series of 1:250,000 maps based on the photography from the Apollo 15, 16, and 17 missions. This program has resulted in the production of orthophotomaps and topographic orthophotomaps. A number of special scale maps have been prepared including 1:1,000,000 and 1:5,000,000. Detailed maps of the Apollo landing sites, called Traverse Charts, show the routes taken by the astronauts on the lunar surface. A program to revise some of the LAC charts is currently underway. (8,9,10)

In addition to the lunar maps and charts based on telescopic and spacecraft photography, the U.S. Geological Survey (USGS) has prepared a series of maps at the scale of 1:1,000,000 showing the geology of most of the visible side of the Moon. Other geological maps on the Moon's hemispheres have been published in the 1:5,000,000 scale. These maps in the Survey's Miscellaneous Investigations Map Series (I-No.) are available from the USGS Publications Office.

A full complement of the lunar maps and charts including many of the earlier maps and atlases, and some of the Russian maps and charts, are available for study and in some cases, for loan, in the Photo/Map Library at the Institute. In many instances, staff personnel can assist the researcher to obtain copies of those maps needed. Individual copies of maps in print are available through the NSSDC. A guide to the lunar maps at NSSDC is being prepared and will be distributed in early 1980.

As the emphasis of the space program shifted from the Moon to the other planets, the scope of the Institute also changed. In 1978, the Lunar Science Institute became the Lunar and Planetary Institute (LPI). The Lunar Data Center was reorganized to include:
- The Library/Information Center
- The Laboratory for the Analysis of Planetary Surfaces (LAPS) which includes:
  - The Photo/Map Library
  - The Geophysical Data Facility
  - The Lunar Sample Information Library
  - Photogeologic Interpretation Facility

And just as the sets of data resulting from Mariner, Viking, Voyager, and other planetary missions are being archived at NSSDC, the LPI is expanding its collections and information services to assist the lunar and planetary community.

The scope of the Bibliography of Lunar Literature maintained by the Library/Information Center was expanded in 1978 to include the Moon, planets, asteroids, comets, and meteorites. Indexing using a controlled thesaurus was also begun with the 1978 literature to increase the access points from just author and title key-words to a full range of subject-oriented index terms. Currently the on-line bibliography contains about 17,000 references. Searches are run based on telephone, letter, or in-person requests.

The Lunar and Planetary Information Bulletin (formerly the Lunar Science Information Bulletin) is compiled and published quarterly by the staff of the Information Center. It is circulated free to approximately 3200 scientists, educators, and students. The Bulletin contains news about personnel changes, conferences, publications, reports and reviews of space missions and mission opportunities, a calendar of events, and the listing of the current bibliography.

The collections of the Photo/Map Library have expanded to include Mariner, Viking, and Voyager imagery and maps. The present planetary imagery collection consists of:

Mariner 6	8x10 prints
Mariner 7	8x10 prints
Mariner 9	microfiche, 8x10 prints
Mariner 10	70mm negatives, positives, contact prints

Viking Landers-Orbiters	5" prints and negatives 8x10 prints and negatives, mosaics 20x24 prints, mosaics
Viking Orbiter 1	8x10 color mosaics slides, color mosaics
Voyager I	
Jupiter	Selected frames; films of mission overview, Jupiter rotation and motion of the Great Red Spot
Satellites	5" prints, negatives and positives 10" prints

A lunar and planetary slide collection is accessed by a computerized subject index. Slides may be borrowed from this collection to support lectures, publications, or other educational and research efforts.

The map collection is also expanding to include the planetary maps for Mercury and Mars prepared by USGS using Mariner imagery. These maps are available as shaded relief, topographic, and geologic representations. Additional planetary maps prepared from Viking and Voyager photography are being considered by the Mapping Program of the Planetary Programs Office at NASA. These maps will be included in the collection as they become available.

Through the development of LAPS, equipment for data analysis at the Institute includes a plane-table digitizer, Zeiss TGZ-3 Particle-Size Analyzer, Traid Viewer, Itek Variable Magnifying Viewer, Zoom Transferscope, and stereo viewers. Proposals for research projects using the LAPS facility and the collections of the Photo/Map Library may be submitted to the Director of the LPI.

To assist users to identify the lunar and planetary data which is available, the NSSDC and the LPI prepare catalogs and information resource guides. These are available from the issuing agency. (11,12,13,14)

Thus although the NSSDC/WDC-A-R&S is the major international distribution center of lunar and planetary data, the LPI and its staff serves as a liaison between the scientific and lay community and the data sets themselves.

REFERENCES

(1) International Council of Scientific Unions. Panel on World Data Centres, Fourth Consolidated Guide to International Data Exchange through the World Data Centres, Washington, ICSU Panel, 1979, 113 pp.

(2) WDC-A; World Data Center-A for Glaciology, Meteorology, Oceanography, Rockets and Satellites, Rotation of the Earth, Solar-Terrestrial Physics, Solid-Earth Geophysics, Washington, Environmental Data Service, 1978, 16 pp.

(3) Report on Rocket and Satellite Information and Data Exchange, Greenbelt, Md., World Data Center A for Rockets and Satellites, April 20, 1979, 14 pp.

(4) Vostreys, R.W., Planetary Data at the National Space Science Data Center, IN: Reports of Planetary Geology Program, 1978-1979, edited by J. Boyce and P.S. Collins, NASA TM-80339, 1979, p. 444.

(5) Warner, J. and C. Dardano, User Guide to the Lunar Sample Data Base, Houston, TX, Johnson Space Center, Curator's Office, July 3, 1978, 11 pp.

(6) Lauderdale, W.W. and W.F. Eichelman, The Apollo Scientific Data Experiments Handbook, Houston, TX, Johnson Space Center, NASA TM-X-58131, 1974, 898 pp.

AD-A088 588 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 5/2
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION.(U)
APR 80 K N KOUROGENIS, W A MARTIN
UNCLASSIFIED AGARD-CP-279 NL

2 OF 2
AD A
088588

END
DATE
FILMED
10-80
DTIC

(7) Bates, J.R., W.W. Lauderdale and H. Kernaghan, ALSEP Termination Report, Washington, D.C., NASA, NASA RP-1036, April 1979, 162 pp.

(8) Schimerman, L.A., Lunar Cartographic Dossier, v.1, St. Louis, MO, Defense Mapping Agency, Aerospace Center, 1975, looseleaf.

(9) Kinsler, D.C., User Guide to 1:250,000 Scale Lunar Maps, Houston, TX, LSI, 1975, 24 pp. (LSI Contribution No. 206)

(10) Kinsler, D.C., New Lunar Cartographic Products, Proc. Lunar Science Conf. 7th v.3, New York, Pergamon Press, 1976, p. i-x.

(11) Cameron, W.S., E.J. Mantell and E.R. Miller, Catalog of Lunar Mission Data, NSSDC 77-02, 1977, 204 pp.

(12) Vostreys, R.W., Catalog of Viking Mission Data, NSSDC 78-01, 1978, 36 pp.

(13) Waranius, F.B. and M.A. Hager, Lunar and Planetary Science Information Resources, Houston, LPI, 1976, 24 pp.

(14) NSSDC Data Listing, NSSDC/WDC-A-R&S 79-05. (to be published)

ACKNOWLEDGMENT. The author wishes to thank R.W. Vostreys, NSSDC, and Ron Weber, Dan Kinsler, and Kin Leung, LPI for their help in preparing this paper. The Lunar and Planetary Institute is operated by the Universities Space Research Association under Contract No. NSR-09-051-001 with the National Aeronautics and Space Administration. This paper is Lunar and Planetary Institute Contribution No. 382.

USEFUL ADDRESSES

World Data Center A
Coordination Office
National Academy of Sciences
2101 Constitution Avenue
Washington, DC 20418

National Space Science Data Center
Code 601.4
Goddard Space Flight Center
Greenbelt, Md 20771

World Data Center A
Rockets and Satellites
Code 601
Goddard Space Flight Center
Greenbelt, MD 20771

Curator's Office
NASA Johnson Space Center
Code: SN2
Houston, TX 77058

Geophysical Data Curator
NASA Johnson Space Center
Code: SN6
Houston, TX 77058

Planetary Programs Office
NASA Headquarters
Code: SL-4
Washington, DC 20546

WORKSHOP SESSION

Summary edited by
D.W.Goode
Programme Director

INTRODUCTION

The Workshop Session was held on the afternoon of Thursday, 18th October 1979. Its primary function was to act as a discussion on technical information topics or problems between the Meeting Delegates and Members of the Technical Information Panel (TIP). Priority was given to items of particular interest to the host country, Greece. It also allowed for extra questions relating to the Sessions papers and for which there may have been insufficient time available during the discussion periods which followed each individual paper. With this requirement in mind a panel consisting of Members of the AGARD TIP supplemented by some of the Meeting's speakers was formed for the occasion to answer questions received. An opportunity was also afforded for Delegates to submit questions in writing in advance of the Session. The Workshop Session was chaired by Mr H.Sauter (USA) the Deputy Chairman of AGARD TIP.

Questions submitted together with answers received from Workshop Session Panel Members and other participating delegates are listed below.

Submitted by the Joint Committee of the Hellenic Air Force

"What implications or difficulty do you see in a terminal between KETA, (The Technical Research Centre of the Hellenic Air Force) and Frascati?"

In the absence of Mr Martin (ESA) who was unable to be present at this Session, this question was answered by the Chairman, Mr Sauter (USA). He stated that he did not foresee any particular problem providing that there was already in existence a good telephone connection between Italy and Greece. With the aid of a portable terminal Mr Tan (the Panel Chairman) had been able to demonstrate the on-line system to Portuguese Panel Members in Lisbon using the nearest available computer centre which was, in this case, Paris. He did not foresee any hardware difficulties but did not like to make any estimation regarding costs which he felt should best be discussed with the European Space Agency. The primary need of user education for on-line systems was also stressed by Mr Sauter who again suggested that the advice of ESA should be sought in this matter.

Mr Tan (Netherlands – TIP Chairman) stated that there was a great deal of literature available relating to on-line systems now in use. He also recommended that ESA should be contacted and that literature and advice should be obtained from them.

Mr Tittlbach (FRG – TIP Member) stated that, as he understood that Greece was not yet a member state of the European Space Agency, this might raise certain difficulties.

Mr Coyne (USA – TIP Member) agreed that this was the present situation but believed that this problem could be overcome by Greece making an application to become an observer state. Once this status had been achieved it would then be advisable to consult ESA about the mechanical aspects of establishing a terminal at the Greek end. Much would depend on the particular needs (e.g. a multiplexer system or minicomputer system) and the number of terminals required. Mr Coyne reminded the questioners that they would need to specify which data files they would wish to consult. Those most commonly used in the aerospace industry included NASA, COMPENDEX, NTIS and the French PASCAL. Mr Martin (ESA) had informed Mr Coyne that he would be looking into the Greek situation and letting Dr Kourogenis know of the present status and whether in fact an application for observer status had already been received from Greece.

Submitted by the Joint Committee of the Hellenic Air Force

"We are going to develop a centre for aerospace information and would welcome your suggestions on the minimum requirement for its implementation."

Mr Sauter (US) felt that this would be an extremely difficult question to answer effectively in the short time then available. Detailed information on the type of service envisaged for the centre would obviously need to be provided.

He was strongly in favour of a consultancy service to advise on the setting up and development of the projected establishment, and suggested that AGARD itself may be able to provide this. Mr Sharp (AGARD TIP Panel Executive) gave some brief details of AGARD education and consultancy programme under which member states may ask for advice. He felt that this matter could be raised by the Greek delegates at the Technical Information Panel's Business Meeting to be held on the following day. This advice was accepted and the request formally raised by Major Goulios at the Panel Business Meeting on 19 October. The Panel was able to formulate a recommendation which was passed to AGARD Headquarters in Paris for consideration and possible implementation.

Submitted by Mr R.W.Slaney, Librarian, Defence Operational Analysis Establishment, Ministry of Defence – UK

"Attention has been drawn to the unhappiness of users who receive more information than they need, and to the problems caused by unwanted and time-consuming data. This seems to be a growth area in information service criticism which reflects both the information explosion and the increased effectiveness with which we do the job we aim at."

Drawing upon his own experiences at the US Defense Documentation Center, Mr Sauter (US) felt that with modern search techniques it was inevitable that the requester would receive more information than he really wanted. This he felt was largely due to the fact that large information centres, such as DDC, carry out searches on a broad and general basis. However, in the case of his own establishment (soon to be retitled the Defense Technical Information Center) he would in the future be given the responsibility for a number of information analysis centres and, with this, an opportunity to refine search techniques and to supply more specific information. Mr Sauter suggested that although general searches and the bulk processing of information could still be handled most effectively by DDC the Centres could best concentrate on searches in their own clearly defined subject fields. This, he felt, should result in an improved situation but would never completely eliminate the comment that was made.

Dr G.Rosenau (W.Germany) also expressed his concern about the large release of information now being produced by computerised search techniques particularly at a time when stringent economies in such items as paper are being demanded. He felt this was due in some measure to what he referred to as "the unstable equilibrium of the relations between user and sources". Many retrieval systems in use were not flexible enough to meet specific needs of users. Should there be standards set for search effectiveness and if so who should set them? It is the requester who must be satisfied but he does not always set standards himself.

M.Deschamps (France) observed that the use of subject specialists to index material in their specific fields for input to a data file may help improve the situation. He further suggested that the same specialists should handle search requests again in their own subject fields. Mr Goode (UK – TIP Member) stated that with general adoption of the computer-aided search this problem of "overkill" seemed inevitable. However, a method in use at his own establishment (Royal Aircraft Establishment, Farnborough, UK) relied on the requester being physically present whilst the search was taking place. Thus the search strategy could incorporate his expertise and suggestions and if necessary varied or expanded to accommodate them. He was also able to select his own "lists" from the VDU and only these need be printed out thus economising on paper. Mr Goode realised that this method would, of course, be impossible to adopt at the major general information centres such as DDC and DRIC or where the requester was not a member of the establishment concerned. A problem now being faced by Librarians and Information Officers is that when faced with a mass of computer produced references a requester often expects them to select the key items for him which could have unsatisfactory consequences.

The questioner himself (Mr Slaney) suggested that to achieve more satisfactory results from the point of view of selectivity, tools more suited to the job than the present secondary sources needed to be developed. A possibility might be a move from abstracts to synopses similar to those being tried by chemists which were described in a paper given at the 1977 TIP Specialists' Meeting held in Oslo.

Submitted by Mr J.Chillag, British Library, Lending Division, Boston Spa, UK

"What provisions are there for document back-up supply for conventional and non-conventional literature which can be supplied to users abroad with special reference to Greece?"

Mr Coyne (USA) said he was not acquainted with any existing or proposed arrangements to enable Greece to obtain literature from foreign sources but he felt sure that this would be closely looked into in connection with the setting up of the National Documentation Centre. It was also a topic for consideration by any team acting in a consultancy capacity on behalf of AGARD or similar organisations. He felt that information should be sought from large national documentation sources on services provided for overseas requesters. Amongst the establishments he had in mind were the National Technical Information Service, NASA, and the Library of Congress in the USA, the University of Hanover in West Germany and of course Mr Chillag's own establishment, the British Library in the UK.

Mr Chandler (USA – TIP Member) reminded the Meeting that in some instances a bi-lateral agreement might well be required. In the case of NASA for instance they would probably be willing to provide Greece with **International Aerospace Abstracts** and **Scientific and Technical Aerospace Reports** but would require copies of Greek reports in return for NASA documents supplied.

The questioner, Mr Chillag, stated that in certain less developed countries the services of the British Library are made available through the British Council. He suggested that the Greek authorities might like to take advantage of this service if they had not already done so.

Submitted by Dr A.Cockx (TIP Member) Belgium

"In his introductory speech Dr Kourogenis gave us a fair description of the Greek situation especially on the programme for the projected National Documentation Centre. I would like to know what part the needs of Greek industry plays in the formulation of a national science policy and what contacts have been made with industry to assess its requirements?"

Dr Kourogenis (Greece) stated that the fullest cooperation was being sought between those responsible for a national information programme and scientists, engineers, industrialists and representatives from all walks of life. At present Greece is an importer of technology involving large sums of money being spent in such areas as patent fees and similar royalties. This too must be borne in mind when setting up a national technical information service. The design concept for the National Documentation Centre is a very broad one and will almost certainly include the setting up of an Industrial Information Centre in cooperation with the Ministry of Industry and the Union of Industrialists. This then is the projected target which may, of course, differ somewhat from the end product.

Submitted by Dr A.Cockx (TIP Member) Belgium

"Optimal introduction of international information networks such as ESA or EURONET in a small country can only be successful if a number of minimum requirements are fulfilled: existing infrastructure in library collections, the education of training of intermediaries/staff etc. Are there any studies now taking place in Greece to evaluate the number of prospective users and to obtain other marketing information in view of possible high installation and running costs for these services?"

Dr Kourogenis (Greece) replied that although some work had taken place he did not believe that any projected figures for user numbers together with their subject requirements were available for Greece as a whole and he felt that this information may be difficult to compile. However, the picture was much clearer in relation to specific industries, for example, aerospace. In this particular field, there are only between 12 and 14 establishments so that it would not be at all difficult to get some feedback as to what their requirements might be. Those requirements could include the need for information on such subjects as geophysics and geology, meteorology, and remote sensing, in addition to normal aerospace literature. From this it would seem that an assessment of projected users in clearly defined subject areas might well be attempted. It was probably too early in the planning stage to attempt any assessment at this time but clearly advice could be sought from the consultants if and when they were appointed, together with their suggestions for a user education and training scheme.

Submitted by Mr R.Slaney (UK) Defence Operational Analysis Establishment, Ministry of Defence – UK

"The results achieved by the Cleverdon Tests are less surprising if the concepts implied by the types of indexing language are studied. The intellectual processes required for the control of natural language are very similar to those used in conventional subject indexing, the main variable being the "location" of the thesaurus."

This particular question was primarily addressed to Mr Martin of the European Space Agency but as he had already left the Meeting Mr Sauter (USA) agreed to reply. In doing so he drew upon his own experiences in the use of national language indexing at the US Defense Documentation Center. He agreed that the required intellectual processes are very similar and was not convinced that any savings could be made. In his work at DDC it had cost as much to control the vocabulary for the natural language as it did for the initial indexing effort.

In connection with natural language DDC uses a machine-aided indexing process which he personally believed had some advantages in this system or concept although opinions were somewhat divided amongst his indexing staff. He believed the advantage to be that this system does in fact take the natural language contained in the title of the abstract and in the description set supplied by the author and puts it into a standardised form. In doing so it gets away from what has been called the over-professionalism of the indexers themselves. So often indexers are prone to add their own thoughts about what was really contained in a document and whilst this may be all very good, we must bear in mind the fact that the requester who receives a bibliography from a documentation centre does not have the advantage of processing the full text of that report or document. What, in fact, he has to do is to look at the title or abstract and make a decision as to whether it is relevant to his particular needs. At the same time he must decide whether he should order the document and in most cases pay for it. The danger then in adding additional terms or concepts that are not reflected in the national language of the title or abstract is that the requester may think that the information supplied is irrelevant to his request.

The questioner, Mr Slaney, (UK) thanked Mr Sauter for his reply with which he was entirely in agreement. He had studied a contract report on the operations of his own library carried out by Professor Cleverdon and had come to almost exactly the same conclusions as those drawn by Mr Sauter. The real benefits he felt came from a more responsive indexing system when using a natural language, but strict control must be maintained by using the conventional intellectual processes.

Mr Sharp (TIP Panel Executive) read out a comment from Mr C.Schuler, (UK – Consultant and former Chairman of AGARD TIP) relating to Dr Kourogenis's review paper on the plans for a Greek National Documentation Centre and for other specialist information centres to serve the various scientific and technical communities. In this comment Mr Schuler drew the attention of Dr Kourogenis and his Greek colleagues to the AGARD Technical Information Panel sponsored "*Manual of Documentation Practices for Aerospace Information*" which is in the course of being published. The Manual will be in four volumes, two of which have already been published and are available as AGARD Agardograph AG-235 Vol I and II. It has been prepared as a practical "how to do it" for those users wishing to start a new technical centre or perhaps to improve an existing service. Topics covered by the first two volumes include acquisition, indexing, abstracting, mechanisation using minicomputers and the creation of a selective dissemination of information service. Volume III covering such subjects as information retrieval, microfilm, reprography and dissemination is scheduled for printing early in 1980. Both Mr Sharp and Dr Kourogenis expressed their thanks to Mr Schuler for his helpful comment.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-279	ISBN 92-835-0264-7	UNCLASSIFIED

5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France
6. Title	INTERNATIONAL ACCESS TO AEROSPACE INFORMATION
7. Presented at	the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

8. Author(s)/Editor(s)	9. Date
Various	April 1980

10. Author's/Editor's Address	11. Pages
Various	106

12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Aeronautics	Information
Cooperation	Technical reports
Europe	Research projects

14. Abstract

These Proceedings concern an AGARD Technical Information Panel meeting held in Athens ard are prefaced by a review of the status of Greek organization, facilities, etc. with respect to access to aerospace information. The review is then widened to consider the state of European co-operation in the field and the nature of NASA's important contribution. Later papers deal with requirements and tools for international co-operation and data exchange; problems in the utilization of aerospace literature; and non-literature data in aerospace R&D. Transcripts of the discussions of the various papers and of a closing round-table discussion are included.

AGARD Conference Proceedings No.279
Advisory Group for Aerospace Research and Development, NATO
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION
Published April 1980
106 pages

These Proceedings concern an AGARD Technical Information Panel meeting held in Athens and are prefaced by a review of the status of Greek organization, facilities, etc. with respect to access to aerospace information. The review is then widened to consider the state of European co-operation in the field and the nature of NASA's important contribution. Later papers deal with requirements and tools for international co-operation

P.T.O.

AGARD-CP-279

Aeronautics
Cooperation
Europe
Information
Technical reports
Research projects

AGARD Conference Proceedings No.279
Advisory Group for Aerospace Research and Development, NATO
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION
Published April 1980
106 pages

These Proceedings concern an AGARD Technical Information Panel meeting held in Athens and are prefaced by a review of the status of Greek organization, facilities, etc. with respect to access to aerospace information. The review is then widened to consider the state of European co-operation in the field and the nature of NASA's important contribution. Later papers deal with requirements and tools for international co-operation

P.T.O.

AGARD-CP-279

Aeronautics
Cooperation
Europe
Information
Technical reports
Research projects

AGARD Conference Proceedings No.279
Advisory Group for Aerospace Research and Development, NATO
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION
Published April 1980
106 pages

These Proceedings concern an AGARD Technical Information Panel meeting held in Athens and are prefaced by a review of the status of Greek organization, facilities, etc. with respect to access to aerospace information. The review is then widened to consider the state of European co-operation in the field and the nature of NASA's important contribution. Later papers deal with requirements and tools for international co-operation

P.T.O.

AGARD-CP-279

Aeronautics
Cooperation
Europe
Information
Technical reports
Research projects

AGARD Conference Proceedings No.279
Advisory Group for Aerospace Research and Development, NATO
INTERNATIONAL ACCESS TO AEROSPACE INFORMATION
Published April 1980
106 pages

These Proceedings concern an AGARD Technical Information Panel meeting held in Athens and are prefaced by a review of the status of Greek organization, facilities, etc. with respect to access to aerospace information. The review is then widened to consider the state of European co-operation in the field and the nature of NASA's important contribution. Later papers deal with requirements and tools for international co-operation

P.T.O.

AGARD-CP-279

Aeronautics
Cooperation
Europe
Information
Technical reports
Research projects

and data exchange; problems in the utilization of aerospace literature; and non-literature data in aerospace R&D. Transcripts of the discussions of the various papers and of a closing round-table discussion are included.

Copies of papers presented at the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

ISBN 92-835-0264-7

and data exchange; problems in the utilization of aerospace literature; and non-literature data in aerospace R&D. Transcripts of the discussions of the various papers and of a closing round-table discussion are included.

Copies of papers presented at the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

ISBN 92-835-0264-7

and data exchange; problems in the utilization of aerospace literature; and non-literature data in aerospace R&D. Transcripts of the discussions of the various papers and of a closing round-table discussion are included.

Copies of papers presented at the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

ISBN 92-835-0264-7

and data exchange; problems in the utilization of aerospace literature; and non-literature data in aerospace R&D. Transcripts of the discussions of the various papers and of a closing round-table discussion are included.

Copies of papers presented at the Technical Information Panel's Specialists' Meeting held in Athens, Greece 17–18 October 1979.

ISBN 92-835-0264-7

8129/9

AGARD

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Telephone 745.08.10 · Telex 610176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Defence Science Information Services
Department of National Defence
Ottawa, Ontario K1A OK2

DENMARK
Danish Defence Research Board
Østerbrogades Kaserne
Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92320 Châtillon sous Bagneux

GERMANY
Zentralstelle für Luft- und Raumfahrt-dokumentation und -information
c/o Fachinformationszentrum Energie, Physik, Mathematik GmbH
Kernforschungszentrum
7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force General Staff
Research and Development Directorate
Holargos, Athens

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3, Piazzale Adenauer
Roma/EUR

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
2600 A.C. Delft

NORWAY
Norwegian Defence Research Establishment
Main Library
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Direcção do Serviço de Material
da Forca Aerea
Rua da Escola Politécnica 42
Lisboa
Attn: AGARD National Delegate

TURKEY
Department of Research and Development (ARGE)
Ministry of National Defence, Ankara

UNITED KINGDOM
Defence Research Information Centre
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RE

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Field, Virginia 23365
Attn: Report Distribution and Storage Unit

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

Microfiche or Photocopy
National Technical
Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

Microfiche
Space Documentation Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

Microfiche
Technology Reports
Centre (DTI)
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RF
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Information Facility
Post Office Box 8757
Baltimore/Washington International Airport
Maryland 21240, USA

Government Reports Announcements (GRA)
published by the National Technical
Information Services, Springfield
Virginia 22161, USA

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London W1P 1HD

DATE
FILMED
-8