Concurrent session 5A

Telecommunications

Chair

D. Varloot

# The needs for telecommunication services in the information market

A. Gambaro, S. Poroli - SIP D.G.

The market of "on line" information is rapidly growing. Availability of update information represents the necessary condition for manufacturers, services companies, public and private operators to plan their activities, to enter the national and international markets, to provide more efficient services, conformed to the needs of users and clients.
Telecommunications have a central role in this process; the problem to access host computer and data bases from remote sites is a thing of great importance for the user. Host computers with data bases are numerous, distant and spread in the world; the existance of a telecommunications network with high quality, low costs, interconnected with international networks is an essential condition to transport and to present the information when and where required.

Telecommunications networks may use various switching techniques (circuit, message, packet) and can be characterized by various utilities (speed, protocol, etc.) and various costs.

The ways to connect a terminal to a data base are:

## a. Public switched telephone network

In this case the user dials the host computer number and uses the telephone connection to transmit data through a modem.
The cost, for this kind of connection, depends on distance and time, as the normal telephone communications.
This solution is profitable only if the distance between the computer and the user is small; usually the computer holding the information, called the host computer, can be located anywhere in the world. In this case the telephone call is not convenient for the high costs but also for the high error rate.

## b. Direct connections

In this case a direct and fixed connection exists between a user and a specific host computer. With this kind of connection is possible to transmit, with good quality, at high transmission speed, but the user can't reach different destinations. This solution may be convenient only for long and frequent connections with a fixed host computer.

c. Public switched packet network

The packet switched network uses a store - and - forward technique in which the messages to be transmitted are split up into packets which are then addressed, stored and switched.
Packet-mode data terminal equipment can be connected directly to the network. However, so-called packet assembly/disassembly (PAD) facilities in the network also allow certain non-packet-mode types of data terminal equipment to be attached.
Communicating stations are interconnected by a logical (so-called virtual) connection rather than a direct physical link.
Since the data are stored in the network, connections are possible between stations operating at different speeds.

The diffusion of the public switched packet networks (PSPDN) gives the possibility for the development of "on line" information market.
Through a national PSPDN a user has the possibility to realize connections with external data bases, with the following characteristics:
- cost depending on duration, volume of data transmitted and independent from distance;
- high quality and speed of transmission, because the network is specialized for data transmission;
- complete intercommunication with international Data Bases connected to international packet networks (i.e. an italian user may access Data Bases connected to Transpac, Iberpac, Telenet ..).

The packet switched network is, for the above reasons the more convenient system from the cost/performance point of view for acces to data base application.

We report the example of an Italian user who wants some information from a data base in U.K.. He can use or the packet switched data network (PSDN) or the public switched telephone network (PSTN). If the user, connected to the PSDN through a direct line at the speed of 1200 b/s, exchanges transaction of 14000 characters in 10 minutes the costs are:

PSDN a monthly rental of L.147100
cost for transaction is about L. 3000,
PSTN a monthly rental of L. 122800
cost for transaction is L.14000.

Therefore if the user exchanges more than 5 transaction / month, the packet switched network is cheaper than the telephone switched network.

A problem for a potential user of Data Base is rapresented from the need to be aware about "where and how" access the necessary information.

In order to become user of "on-line" Data Bases is necessary to follow these steps:

a)-to identify Data Bases of interest through:

- -reference books concerning access to Information Provider;
- -collection of information in the preposed organisms;

b)-to select the proper host computer in order to subscribe a contract;

c)-to contact the host computer to obtain password and detailed documentations concerning Data Bases and manuals;

d)-to install equipments (terminal, P.C., modem, printer..) or, if utilizing existing equipments, to check the conformity and if necessary, to choose the communication software;

e)-to request the connection to the telecommunication service (i.e. packet switched network);

f)-to follow training courses for practice in using the inquiry language.

The collaboration among the operator of this sector (information providers, institutional entities and telecommunication companies) is necessary to contribute to the growing of this market and to help the user in his research of information.
SIP, the company managing Itapac, has defined and realised marketing actions in this direction in collaboration with information providers:

1)-presentation of the market of information to the users;
2)-mailing to potential users in different disciplines and promotional campaigns with free access to Data Bases.

Concerning point 1), SIP has organised presentation of the services offered by Itapac, with particular regard to the access to Data Bases application.
SIP has explained the characteristics of:

- -basic network services for transport of information;
- -interface (hardware and software) between terminal equipment and network;
- -tariffs of the network, with particular regard to their convenience for interactive applications.

Some Data Bases producers explained, by means of pratical demonstrations of on-line research, the main characteristics of a research, the contents of their Data Bases, the inquiry languages, the form of a subscription for Data Base access and respective costs.
This activity, promoted by SIP in collaboration with information providers and companies operating in the information market, had great success from a quantitative and qualitative point of view: it rappresented an opportunity to exchange impressions and information among all the partners involved in information market.

The medical sector is particularly interested to this kind of demonstration; there are a lot of medical Data Bases: Medline, Pharmline, Embase, Toxline, etc.

Concerning point 2) the goal of a mailing is to inform the users about applications and services or directly through specific documentation or with indication of existing tools containing information about:

-procedures to be followed in order to become user of Data Base;

-Data Bases of interest for the specific sector of users contacted by mailing.

The mailing gives the possibility to open a dialogue between the company managing the network and the user, giving him the opportunity to find a contact point and to request further informations.

In order to promote the application of Data Base access, SIP has planned mailings to specific activity sectors with promotional offers that give the possibility to the new users to have economical convenience to subscribe a contract in terms of:

-reduction of rental;
-free participation to training courses;
-free connection for limited amount of hours.

Sip may contribute with these initiatives to the development of information market and can help the potential users to find all the information necessary to become "on-line" subscriber.

All the above initiatives concern potential users, of different sectors, without a specific knowledge about the information market, who might be stimulated to the use of telematic services.
There are a lot of sectors interested to "on-line" research; in scientific-technological sector the on-line information give the possibility to spread new knowledge, to verify the

situation for research activity, to increase the applicability of results of basic research; in the industrial and economical sector a rapid and exaustive way to accesss various kinds of information may help in the strategic decisions of industry and company and may constitute a valid tool in the management work.

There is also the need for a telecommunication company to check the impressions of the users of the packet network. SIP has realized some marketing researches in order to check:

-a)which kind of users are connected to the network and which applications are performed.
-b)which are the primary needs for users, the opinion about the service and which aspects of the service may be improved.

It was possible to know that there are several activity sectors interested to the use of the network (industry, commerce, schools, professionals ....) and several applications performed (order entry, store and stock management, message handling, file transfer...).

Moreover it was possible to have indications and suggestions about the modality to give commercial information concerning service,tariffs, interconnection and to improve the quality and rapidity of technical assistence.

# TRANSBORDER DATA FLOW: A PROCESS IN NEED OF ATTENTION FROM MEDICAL LIBRARIANS

BY
JAMES F. COMES
BALL STATE UNIVERSITY
MUNCIE, IN USA

The confluence of computer and communication technologies has resulted in more communication of information at greater speeds and over longer distances, supporting and enhancing global interdependence politically, economically, and culturally. From the marriage of satellites and computers vast amounts of information can be transferred in microseconds by groups as diverse as librarians or national security agencies. Relationships between nation states, once formal and predictable, have been transformed as computerized information systems have reconfigured international trade, financial transactions and information flow. The catalyst for many of the changes now and in the future is the transmission of information across national borders or more commonly known as transborder data flow(TBDF). The purpose of this paper is to demonstrate the importance of TBDF to librarians by identifying the components which most influence the flow of information, discuss the issues and concerns associated with TBDF, explore the future of TBDF and the role of the librarian.

Why should librarians be concerned and knowledgeable on the topic of TBDF? Several years ago in an analysis of issues confronting international librarianship, Russell Shank identified TBDF as a critical issue for librarians since it involved the competing influences of concern for national interests and of a desire to share intellectual wealth.[1] Many of the issues related to TBDF may not affect the flow of medical information directly. However, "it stands to reason that some policies implemented to restrict the flow of information internationally may, by default, affect unfavorably the transfer of medical research information and other scientific and technologic research information related to health."[2]

Since the beginning of this decade a number of developments have occurred which have brought increasing attention to TBDF by transnational corporations, international organizations, developed countries and less-developed countries. The advancement with the greatest impact has been the convergence of computer and telecommunication technologies. As a consequence of the development of the microchip, satellite communication and optical fiber, computers and communication systems have been able to merge voice, image, text, and symbols to render obsolete the customary legal distinctions that governments have used to administer the delivery of information by telegraphy, telephone, television or newsprint.[3] The increasing use of digital data and packet switching is creating faster, cleaner links for computer communications. The distinct features which separated network providers and information providers are becomming less visible. For example, the value-added networks such as EURONET and TELENET have incorporated the convergent technologies for data transmission purposes.

Closely linked to the new technologies is the changing character of information. In a discussion of European information policies, Van Rosendaal stated:

> "Information is one of the few non-scarce raw materials for the economy of Western Europe...It has some qualities of public good...[and possesses]...strategic, social, and cultural value...It will undoubtedly be a source of major investment in the future and information will continue to grow in its significance as an input to all types of goods and services."[4]

Information is not a physical resource which disappears when it is consumed or shared. Sharing through verbal communication or written messages is more likely to add value to the information rather than deplete it. International borders, distance, and volume are irrelevant when using satellites to transmit data.

The convergent technologies and the insatiable need for information have resulted in dramatic increases in TBDF and helped to create an industry. Simpson estimated that in 1986 there were 1500 database producers, over 3,200 databases and 386 companies world wide operating online search services. Sales from these services exceeded $3.5 billion dollars.[5] Corresponding growth in worldwide sales of information based services and telecommunications equipment will be much greater. Together, these developments are adding volume and value to TBDF and have led to three trends identified by Bollinger.[6] First, there will be greater competition for information and knowledge than for manufactured and capital goods. Information stored in scientific and technical databases is viewed as a resource supporting economic development and productivity. The second trend is the development of diverse interests and multiple access points for information. The North American and Western Europe link has been joined by the Far East and many Third World countries who have recognized the benefits of satellites, multinational networks, gateways to support governments, educational programs and private businesses. Third, scientific and technical research has been internationalized. The proportion of scientific and technical literature produced by the U.S. has been decreasing in all fields except biomedicine since 1973. This trend has reflected a growing demand and value of foreign scientific and technical literature. This increased interdependence will magnify the impact of any restrictions or barriers to TBDF and affect research and development for all nations.

## Barriers: Issues and Concerns

The adaptation to social and technological change is never an easy task. It is all the more difficult on the international level where priorities reflected in national policies must be accepted as equally legitimate. Each country desires to maintain its independence, economic strength and cultural identity. However, the means and ends to establish TBDF do not support nationalistic desires in the traditional sense. Concerns are raised when access to scientific and technical research is controlled by another country, when information flow between two countries is interrupted by a power outage, when information acquired from international networks violates the security laws of one country, when political events take priority over normal operating agreements.

In the seventies, the issues associated with the unrestrictive flow of information across national borders were beginning to take shape. Prior to the advent of the computer, limited amounts of information were stored because of the lack of facilities and high costs. Thus, there was little opportunity to misuse stored information. Europeans became aware of the potential for the abuse of personal privacy before Americans, recognizing the ephemeral nature of information and the difficulties of tracking information, finding out what information is being transmitted, to whom, and how it is being used. In addition to the privacy issues, two reports were prepared which had an impact on most countries. The Nora/Mince report, prepared for the French government in 1978, studied the impact of computerization in France and made recommendations for a national information strategy. The Clyne Committee, commissioned by the Canadian government, studied the impact of TBDF's on Canada and recommended a telecommunications plan to avoid exploitation.[7] The issue of dependency helped to stimulate the development of Euronet in 1979 and provide access to a large number of scientific and technical data bases.

As a consequence of these developments and the technological advances mentioned, a substantial number of issues and concerns have surfaced. Baker, from Chemical Abstracts Service, has identified barriers to the international flow of scientific and technical information and grouped them into four categories for purposes of clarity and simplification.[8] They are:

- -Intellectual Property Rights and Legal Mechanisms which include patents, copyrights, contracts and legislation.
- -Economic Issues such as tariffs, taxes, denial of entry, and government subsidized.
- -Technical issues which include standards, new technology, domestic hardware and software.
- -Political, Social and Cultural Concerns such as privacy and data protection, national sovereignty, national security, fair trade, language and stability.

Bakers list has covered the spectrum of issues, concerns and barriers associated with TBDF. A European perspective was offered by Van Rosendaal when he defined the barriers inhibiting the exchange of information services as: the respective roles of the public and private sectors in information supply and dissemination, the physical infrastructure, differing national rules and regulations, and the impact of technology.[9]

Although the terminology and classification of barriers are described differently, there are three underlying issues common to both viewpoints which illustrate how national concerns have affected the development of international information policies. A number of issues arising from TBDF originate (stem from) because of the varying legal systems and laws in each country. One example is the Norwegian social research worker found guilty of espionage after publishing an article on NATO installations. He had obtained the documents from an information service in the U.S. In another example, a German company, Siemen, was not permitted to transfer Swedish employee records to Germany for storage. Part of the problem with the issue of privacy resulted from the differences between the European and American legal systems. The Europeans have used civil law whereas the United States depends on the common law tradition. The issue of privacy, however, has been handled reasonably effectively at the international level by the OECD and the Council of Europe. Additional efforts must be made to improve the cooperation between national legal systems in order to achieve the full benefits of TBDF.

A second underlying issue is trade in data services and telecommunications. The volume of information carried over the international communication system is now doubling every five years. If the present trend continues, it will double every two years. This is a big business producing billions of dollars in revenue. Trade problems have resulted because some countries have erected nontariff barriers to protect their economies, local industries and state monopolies. A related issue is the treatment of information as a commodity and TBDF as trading in goods. When information is considered as a resource, the wider question of social allocation and social control is raised. It appears that existing differences in legal structure and attitudes are likely to make the development of an agreement in this area difficult to achieve.

Third, access to data and information refers to the need to ensure that governments do not take unilateral action that would curtail the use of international services. Interruptions to the flow of information could occur because of labor unrest, unauthorized use of service and a lack of guidelines or criteria to access data and information normally protected. In the fields of technology and science strict controls will

retard advances. Jacob noted: "since information grows primarily through cooperative research, analysis, combination and synthesis, free flow is the strongest means of ensuring new indigenous information sources."[10]

## Future

Approximately ten years ago U.S. Senator George McGovern commented:

> "One way to attack a nation such as the United States which depends heavily on information and communications is to restrain the flow of information---cutting off contact between the headquarters and the overseas branches of a multinational firm, taxing telecommunications crossing borders, building information walls around a nation."[11]

The benefits of information to all nations are real and measurable. They impact on all phases of human interaction and a range of activities such as science, culture, education, finance and libraries. In recognition of the opportunities and advantages offered by TBDF a number of international organizations such as the OECD, the Council of Europe, the International Telecommunications Union and the Intergovernmental Bureau of Informatics are working to understand the new complexities with service and use issues. The effort by nations and international organizations to regulate TBDF will be difficult for several reasons.

- The rate of technological change has continued unabated. Given the pace of change in computers and telecommunication technologies, a successful intervention by a national government or international organization will be more difficult. For example, the recent announcement by the MacDonald's corporation that it is using an Integrated Services Digital Network(ISDN) to link its restaurants to send voice, data and video over ordinary telephone lines simultaneously will challenge regulators.

- At the international level there is a lack of consensus on an approach to meet the communication and information needs of all countries. For example the United States does not have a national policy for communication and information science. Even though our influence is significant worldwide, the quality of our input is less. The European Community faces an uphill battle to convince member states to change telecommunication and equipment policies to ensure the development of new products and services.

- Competition has entered the global telecommunications environment and calls for deregulation have challenged government telecommunication services. The traditional practices and international relationships which were based on premises assuming slowly changing technologies, homogenous services and regulatory regimes are no longer characteristics of the global telecommunications realm.

Although our efforts may be meeting the needs of our patrons in-spite of these difficulties, change will affect us directly or indirectly. A number of individuals and groups have called for a New Information Order to take advantage of the new technologies, support equal access to all scientific and technical resources, and encourage cooperation between governments and international institutions. The New Information Order would be characterized by the demand for timely and high quality information services, increased efficiencies and effectiveness in the flow of information, and support to share resources and the exchange of information.

In the information age the librarian can play: The role of broker, facilitating the flow of information and services between the public and private sector; the role of provider,

ensuring that the archival nature of libraries is maintained; the role of guardian, promoting equal access for all people; the role of educator, teaching and encouraging citizens to learn information handling skills.[12]

TBDF offers librarians alternatives for the future. It may also be used to widen the gap between those who have access to information and the means of using it and those who do not have access to this information. More dialog, more cooperation and understanding is needed by all parties involved in this endeavor.

## REFERENCES


1. Shank, Russell. "IFLA, ALA and Issues in International Librarianship." Library Journal 107(July 1982):1300.

2. Smith, Kent A. "Transborder Data Flows: The Transfer of Medical and Other Scientific Information by the United States." The Information Society 5(1987):68.

3. Branscomb, Anne W. "Global Governance of Global Networks: A Survey of Transborder Data Flow in Transistion." Vanderbilt Law Review 36(May 1983):981.

4. Van Rosendaal, C. Jansen. "European Information Policy Situation." Aslib Proceedings 36(January 1984):15-23.

5. U.S. Congress, House of Representatives, Testimony of J.W. Simpson Before the Committee on Government Operations Regarding H.R. 145, Washington, D.C. February 26, 1987, p. 3.

6. Bollinger, W.A., and D.C. Ellingen. "Evolving United States' Information Policy and Its Effect on International Access to Online Technical Databases." In 11th International Online Information Meeting, December 8-10, 1987, London. Learned Information, Oxford, 519-528.

7. Miller, Anthony Paul. "Teleinformatics, Transborder Data Flows and the Emerging Struggle for Information: An Introduction to the Arrival of the New Information Age." Columbia Journal of Law and Social Problems 20(1986):89-144.

8. Baker, Dale B. "Chemical Information Flow across International Borders: Problems and Solutions." Journal of Chemical Information and Computer Sciences 27(1987):55-59.

9. Van Rosendaal, C. Jansen. "European Information Policy Situation." ASLIB Proceedings 36(January 1984):15-23.

10. Jacob, M.E.L. and D.L. Ring. "National and International Information Policies." Library Trends 35(Summer 1986): 142.

11. McGovern, George. "The New Information Age." New York Times 9 June 1977 A21, col. 2.

12. Edwards, Lee. "Information Protectionism: A Clear and Present Danger?" The Information Society 5(1988): 161-168.

# Biomedical Libraries in the Age of Telecommunication

D. Scherf, Product Manager
Radio-Suisse Ltd. Telecommunications + Databases

## Abstract

Due to important technical improvements, in the last five years, in the field of telecommunications, hardware as well as software, new needs and of course new tasks have arised for modern biomedical libraries. The paper will give an overview of all the elements which belong to telematic as well as an outlook for the next five to ten years. Further, a specific telematic service, the interlibrary communication and ordering system, will be discussed.

## Introduction

Before the introduction of printing in Europe all copies of books had to be written by hand. Costs and slowness of copying limited the size and number of libraries that were usually located in monasteries or in cathedrals.

The advent of printing in the 15th century brought down the price of books by speeding up the process of copying and of distributing.

Private libraries became more common and by the 17th century even a few public libraries had come into being. However, users remained few until literacy became more widespread during the 18th century.

At the end of the 20th century the reality and problems have changed completely. Literacy is widespread and we may talk today of an information explosion as well as a democratisation of knowledge. Policy of today's modern libraries is clearly towards the ideal of making all information available, without delay and this to an increasing number of readers.

For the first time in the human history this goal can be reached with the help of modern technologies.

## Telecommunications and EDP Systems

The conjunction of telecommunications and EDP systems, so called telematic, have given a new dimension and of course new tasks to medical libraries. There is no doubt that telecommunications will play in the near future an important role in our daily life and will affect sooner or later our way of working, living and thinking.

In the last 10 years this sector of the economy has made great improvements. Long distances are now overpassed in an ellegant way and of course in a cost effective relation by satellites. New based computer technologies have given new tools to telecommunications technicians to optimize trunks capacities, with an improvement on data transfer quality and last but not least by upgrading transfer speeds.

The utilization of those technologies e.g. packed switched data networks, have also permitted to drop transmission costs on a level which may be defined as accessible to anyone.

With the implementation of the optical fiber, by the end of the 80's beginning of the 90's, speed will again be upgraded and costs will dramatically drop, due to the fact that networks will have more capacity.

On the other hand a modern network without any value added services is like a harbour without a vessels. As stated before, telematic is a conjunction of telecommunications and EDP systems. Thanks to powerful computer systems, a library can develop applications in order to support and optimize internal but also external libraries tasks (e.g. catalogue, indexing facilities, loan control system, billing, etc.).

Another important factor in this context is the appearance on the market of data terminals e.g. PC's, home computers, videotex/minitel terminals, which became available already in the early 80's and may nowadays be purchased with relatively little money.

All those elements have helped to create new needs from the medical libraries users point of view and to give new objectives to those organizations.

Telematic offers the possibility to consult databases on-line as well as library catalogues with respective additional services as e.g. on-line ordering. All this independently, from the user's geographical location. In other words telecommunications help to disseminate in a fast, cost effective way, all the world's knowledge in this area. No matter if the user is living just beside a library or in remote areas.

## Major Telematic Services

The major medical libraries telematic services are as known the followings:

- Library catalogue on-line
  accessible with ASCII oriented terminals such as personal computers, large main frames, home computers or via videotex/minitel terminals.

- Ordering system enabling any of the above mentioned customers to order on-line books or genuine articles.

- Delivery service of genuine articles by using mail box facilities.

## Interlibraries Communication and Ordering System

Another aspect which can be considered as vital, especially in the context of 1992, is the tele-ordering between libraries within a specific country or throughout Europe. The realisation of such an application needs again a telematic system, known as electronic mail. Such a system is e.g. the BELGIAN TELEDOC SERVICE of the University of Louvain's faculty of medicine library, which uses the Radio-Suisse e-mail service called DATA-MAIL as an interlibrary communication and ordering support.

## TELEDOC the Belgian Interlibraries Communication System

As soon as an article reference is encoded by a user, Teledoc first indicates whether the requested article is available at the library; then it may accept the orders of the document, irrespective of whether it is available or not. As soon as the orders appear on the librarian's PC, Teledoc consults the BIOMED union catalogue (permanent inventory of biomedical periodical subscriptions in Belgium) and automatically selects from the 40 libraries listed therein one of the libraries holding the document, thanks to a preestablished order of priority.

At the end of the day (or more often where necessary), TELEDOC forwards the requests by electronic mail to the different selected libraries. In the case of those libraries which are not connected to the electronic mail, but which have a telex, the order will be sent automatically from DATA-MAIL to the telex subscriber. Very soon the system will also be able to send messages or ordering forms via faxsimile.

In the case of articles from periodicals not included in BIOMED's, TELEDOC makes it possible to send out the order by electronic mail to another library freely selected from traditional union catalogues or others.

The TELEDOC software indicates the awaited documents and, in the case of an exaggerated delay, makes it possible to send out reminders to the relevant libraries or new orders to the library which appears next in the priority list.

Upon receipt of the ordered document, the library records the document's date of arrival and forwards it to the final user.

This tele-ordering system is fully automated, from the formulation of the demand by the user, to invoicing. The system is still flexible however as the librarian may intervene at any moment; for example during the formulation of the request, in the choice of the corresponding or in the means of communications.

The automated selection and automated dispatch of orders by electronic mail will considerably reduce the time taken to receive documents. Finally, telefax will make it possible to provide instantaneously the copying of urgent documents between one library and another or to a final user.

## Conclusions

This example shows how modern libraries could solve a specific problem by introducing a telematic service for the interlibraries communication. Such systems could help an organization to speed up document deliveries on one hand but also to cut down administration costs on the other.

Such facilities would also correspond to an increasing number of users, with new needs and follow the main policy of making all information available to a larger audience without delay.

# Concurrent session 5B

# Microcomputers in libraries I

Chair

P. Morgan

**Mme Anne PARRICAL, Centre International de l'Enfance**
**Château de Longchamp - Bois de Boulogne - 75016 PARIS**

**MINISIS: logiciel multi-utilisateurs: un an d'expérience de gestion de base de données et de bibliothèque par l'équipe de la bibliothèque du Centre International de l'Enfance - Paris.**

## LA BIBLIOTHEQUE DU CIE ET SES COLLECTIONS

La création du Centre International de l'Enfance, fondation d'utilité publique, remonte à 1949, alors que pour compléter l'action des organisations internationales, orientée vers les besoins d'assistance immédiate, la nécessité s'était fait sentir d'une institution se consacrant davantage à la formation dans les domaines de la santé de la mère et de l'enfant.

La bibliothèque du CIE, ouverte en 1950, a été, dès le départ, créée pour répondre aux besoins de l'institution avec également une vocation de service public.

Son fonds couvre la santé maternelle et infantile, les vaccinations, l'alimentation et la nutrition, la santé communautaire, tant dans les pays en développement, que dans les pays industrialisés. 15.000 ouvrages, 600 collections de périodiques y sont actuellement stockés.

En 1981, le volume de cette documentation couramment tenue à jour, a semblé d'importance suffisante pour envisager de la gérer sous la forme d'une base de données, plutôt que de continuer une gestion sur catalogues manuels de fiches.

La Banque d'Information Robert Debré (Base BIRD) a donc été installée en 1982 sur un centre serveur extérieur, le G.CAM, tandis que la sélection et la saisie des références étaient effectuées au CIE et les bandes transmises et retraitées par les ordinateurs du centre serveur. La recherche était effectuée en se connectant au serveur, à partir d'un terminal informatique au CIE.

Ceci a duré de 1981 à 1987 et a paru au CIE assez coûteux. En 1987, BIRD(1) comptait 85.000 références. Les interrogations de l'extérieur sur le serveur n'ont pas atteint un niveau probant pendant cette période, et ceci n'a pas encouragé le CIE à poursuivre ce système.

En 1987, le CIE a pris la décision de changer le mode de gestion de BIRD, et de commencer une gestion interne entièrement sur un système matériel et logiciel qui seraient présents à la bibliothèque.

## CHOIX DU LOGICIEL

Il est apparu à l'équipe, qu'il n'y avait pas assez d'expérience interne pour faire un choix "matériel-logiciel" sans se faire aider. Le CIE a fait appel à un consultant externe.

La bibliothèque a alors fourni une liste de ses besoins pour la base de données et pour la gestion informatisée de la bibliothèque. Ceci incluait la gestion de la base, la gestion des périodiques à un niveau national ou international.

Après avoir pris connaissance des articles et ouvrages techniques spécialisés sur les logiciels, l'équipe a identifié 10 logiciels possibles. Ceux-ci comprenaient MINISIS, BASIS, BRS, et d'autres encore. Chacun des fabricants des 10 logiciels a été visité et interrogé sur les possibilités, la structure, l'architecture du logiciel et son prix. A chaque visite, assistaient le consultant, moi-même et un membre du personnel de la bibliothèque. Une liste de 200 questions était passée en revue avec le vendeur qui devait spécifier, également, quel type et marque de matériel pouvait accueillir le logiciel.

A cette époque, des visites d'utilisateurs eurent également lieu, et l'équipe de la bibliothèque était particulièrement désireuse de savoir si les sites visités fonctionnaient sans l'aide d'un informaticien, puisque les spécifications internes n'avaient pas prévu de créer un poste supplémentaire pour un informaticien dans le service.

Pour finir, le rapport du consultant a recommandé le choix du logiciel MINISIS et deux sites MINISIS tournant sans informaticien avaient été vus à Paris. MINISIS tourne actuellement sur un miniordinateur multipostes HEWLETT PACKARD de la série 3000 et a donc été choisi par l'équipe.

Le logiciel MINISIS a été développé par le CRDI (Centre de Recherches pour le Développement International à Ottawa) en vue de répondre à la demande d'un outil de recherches bibliographiques et de gestion de bibliothèques, solide et convivial, particulièrement adapté aux centres de pays en développement et installé également dans des pays industrialisés, puisque les institutions comme le Bureau International du Travail à Genève, le Sénat Français, la Fondation Aga Khan, l'utilisent.

Le CRDI est un organisme canadien à but non lucratif, financé par le Gouvernement du Canada pour entreprendre des projets dans les pays en développement. Il est très actif dans le domaine de la santé, et des sciences de l'information. En 1975, le CRDI voyant qu'il n'y avait pas de système pour gérer les bases de données miniordinateur, a produit MINISIS qui, 12 ans plus tard, est installé dans 300 sociétés à travers le monde et est disponible en version française, anglaise, espagnole, arabe.... et même chinoise.

Le CRDI a également développé une méthodologie de construction de base de données appelée DEV SYS, dont le but est d'arriver à une structuration facilitant les échanges internationaux.

Le point fort de ce logiciel est le module "recherche", très efficace. La structure des données est relationnelle et en même temps le logiciel autorise la création de fichiers d'accès rapide - dits fichiers inversés ou indexés - qui autorisent les recherches sur certains éléments particuliers préalablement inversés (auteurs-titres-mots-clés).

La saisie se sert de l'éditeur de HEWLETT PACKARD. On peut également entrer les saisies en fabriquant des masques de saisie pleine page.

Une autre commodité importante de MINISIS est que l'on peut à tout moment ajouter un élément à la structure de la base de données. En clair, il faudrait normalement décider de tous les éléments de la base avant de la construire. Mais si un champ supplémentaire doit être ajouté après coup, cela est tout à fait faisable.

Au début de l'installation de MINISIS au CIE, certains organismes français se sont préoccupés de savoir si le choix d'un système canadien n'allait pas empêcher une coopération bibliographique en réseau, le jour où cela serait souhaité.

En fait, cet aspect là de la question avait été envisagé dès le développement de MINISIS. Au départ, en effet, MINISIS a prévu de pouvoir importer et exporter des données bibliographiques dans le format ISO 2709,

format d'échange international pour données bibliographiques sur support magnétique.

On choisit souvent un logiciel en fonction d'une ou plusieurs caractéristiques que l'on trouve importantes. Dans le cas de MINISIS, il s'agit de la fonction "recherche", de la possibilité de gérer un thésaurus et d'interfacer avec d'autres systèmes ainsi que du fait que l'ensemble peut être géré sans informaticien. La gestion des périodiques et des prêts par codes barres ainsi que la gestion des acquisitions peut se faire de même que la DSI.

## DEFINITION DE LA BASE DE DONNEES

Le Centre International de l'Enfance commença par tester MINISIS en Novembre 1987.

BIRD(2), la nouvelle base, a démarré en février 1988 et compte, à ce jour (octobre 1988), 3.000 références. Le service a mis 2 mois à travailler sur la structure de la base de données, en examinant de près :

- les champs requis (auteurs, titres, références bibliographiques, descripteurs, etc..)
- lesquels doivent être obligatoires, répétés, inversés.

L'expérience antérieure de base de données avec BIRD(1) a aidé beaucoup l'équipe dans cette tâche.

Pour construire l'ensemble, le CIE s'est aidé de la méthodologie canadienne DEV SYS et de l'exemple concret de Bamako, Resadoc, base de données de l'Institut du Sahel.

Une liste complète de descripteurs a été refaite par l'équipe, en français, en anglais et en espagnol, et parallèlement un thésaurus géographique a été structuré.

Le service de documentation, qui à la fois indexe les documents et fait les recherches dans la base, discuta beaucoup de la terminologie. Les thèmes de la base de données ont été regroupés en 7 domaines.

Tout ce travail a été fait avant l'implantation du logiciel et après avoir comparé le champ de la base avec d'autres bases de données dans le domaine "santé de la mère et de l'enfant". Le contenu a donc été recentré. Le logiciel permet de faire des statistiques annuelles des nombres de documents rentrés dans chaque thème

La base de données recense des articles, des livres, des thèses et de la littérature grise. Tous les documents référencés sont physiquement présents au CIE, et on trouve des documents en français, en anglais, en espagnol et quelques uns en italien.

Bien que l'équipe soit consciente de l'utilité de fournir des résumés aux usagers, et aurait souhaité le faire, en regardant nos limites de temps et de personnel, nous avons décidé de ne pas inclure de résumés. Alors pour permettre un accès maximum en recherche, nous utilisons entre 8 et 20 mots-clés à l'indexation des références de BIRD(2).

## PROJETS POUR LE FUTUR

Alors que, à présent, la base de données tourne, le service pense à automatiser sous MINISIS la gestion des 600 périodiques de la bibliothèque, et est en train d'installer le cardex informatique.

La nouvelle version "G" de MINISIS permettra un transfert de données sous micro ordinateur PC par le logiciel CDS ISIS.

Une autre priorité sera d'importer sous MINISIS notre base de données BIRD(1): 85.000 références bibliographiques sont actuellement toujours sur

serveur G.CAM. Egalement, la confection d'un CD ROM est prévu pour 1989-1990.

Nous souhaitons également offrir des diffusions sélectives d'informations aux usagers voulant suivre un thème.

## CONCLUSION

Le travail passé et présent a été rendu possible grâce à une attention importante portée à l'organisation du travail du service. Un groupe de pilotage de 4 personnes, que je coordonne, suit les problèmes de logiciel et de matériel, les nouvelles applications, les créations de bases de données et le planning de travail du service.

Cinq bibliothécaires et 1 médecin-adjoint ont pris en charge le travail sur le vocabulaire, font la sélection courante, la recherche et signalent les améliorations souhaitées. Les secrétaires et le personnel administratif (4 personnes) prennent en charge la saisie.

Chaque personne du service a dû s'adapter, en l'espace d'un an, à accomplir une tâche différente de celle faite auparavant, et à faire face à des contraintes technologiques nouvelles.

Heureusement, ces changements se sont fait assez bien et la nouvelle base BIRD(2) a pu effectivement démarrer en février 1988.

Il a fallu préparer l'installation du site et l'installation du mini-ordinateur, la gestion des sauvegardes et leur stockage.

Cela a pris un peu de temps, de voir également les questions d'isolation, la température de la pièce (pas de climatiseur) et de résoudre les problèmes de stabilisation de l'électricité. On peut dire que l'équipe a su faire face en 1987/1988 à ce qui avait été considéré comme un changement important, et néanmoins un passage obligé, pour rendre l'information bibliographique du Centre International de l'Enfance sur la santé plus opérationnelle, à un meilleur rapport coût-efficacité, et en même temps gérée par des bibliothécaires désireux de maîtriser également l'outil informatique tout en ne perdant pas de vue le contenu scientifique de l'information.

# A STEP BY STEP INTRODUCTION OF AN AUTOMATIC SYSTEM IN A MEDICAL LIBRARY

*Donatella UGOLINI,**Giovanna ALLORO

* Istituto di Oncologia Clinica e Sperimentale - University of Genoa
** Istituto Nazionale per la ricerca sul cancro - Genoa

## INTRODUCTION

The National Institute for Cancer Research and the Clinical and Experimental Oncology Institute of the University of Genoa share rooms and staff in the same organization having, however, a different holding inventory. The resulting Library is part of the Scientific Services and operates in strict cooperation with the Scientific Information and Documentation Service for on-line bibliographical research. Both services are managed by the same head.

Aims: to satisfy the information requirements of the users and to make cultural progress possible; to cooperate with general projects of the Scientific Direction.

Users: research staff of the Institute (involved in research on cancer) and scientific personnel from outside the Institute; usually highly specialized and in relatively small numbers (on the average twenty people per day).

Collection: 287 subscriptions to scientific journals of which 215 currently active; about 2600 books, with an annual increase of about 250 (150 increase the Library collection and approximately 100 are purchased by specialized funds and located in the Services which assume the responsibility for them).

The Library, initiated at the beginning of 1980, had no catalogs until late in 1984; only a short description of the books in a chronological register was available. By the end of 1984, a total stock revision had taken place and up to now about 1600 books have been catalogued. The organizational planning which was carried out concerns classification, indexing and automation.

## CLASSIFICATION

The bibliographical standard cards were prepared on the basis of the ISBD rules arranged to satisfy the specific requirements of the entire system. The National Library of Medicine Classification (NLMC) was adopted, since our Library is a biomedical Library which deals in particular with tumor research.

Furthermore, an expansion of the QZ-200-NEOPLASMS schedule was carried out, in order better to define the various aspects of Oncology, which is not well-handled by any available classification. For non-medical subjects class Q of the Library of Congress classification was used.

## INDEXING

Key-words extrapolated from MeSH were assigned to the books. A thesaurus was thus compiled, simplifying the hierarchical structure and the logical correlations, in order to create a simple vocabulary for both the indexer and the user. Rules were established so as to standardize the description of the textbooks, sometimes sacrificing the scientific rigour in favour of simplicity, comprehension and the homogenity of the subjects. For complex subjects the cooperation of staff experts of the Institute was requested.

## AUTOMATION

Aims: the automation of the Library must provide a simple, rapid and complete retrieval of bibliographical information for Library users and new management strategies which improve the Library system for its staff.

Hardware: our equipment currently consists of:

- 2 IBM PC AT3 with 512 kb memory
- 1 IBM Proprinter 4202 XL for standard outputs
- 1 HP printer Laser Jet II for quality outputs
- 1 3M streamer with 40 Mb memory to back-up the information files.

Software: at present there are two types of software on the market: a) software addressed to library management which are too inflexible, so that the Library requirements have to meet the software possibilities; b) software used for many applications other than bibliographical information and designed to meet most requirements.

Therefore, it was decided to conduct a market analysis, in order to find a software suitable for our library's needs, examining some packages directly, conversing with other Library managers who use these types of software, or reading literature on the subject.

Three packages in particular were tested: TINlib, INMAGIC and TEXTO. The first is addressed essentially to Library management; the second and third are used for collection and retrieval of general information (documents, employees, patents, etc.).

Information retrieval software was preferred, in order to manage other files in addition to the bibliographical ones, along with the possibility of creating an information retrieval data base which provides simple and rapid search methods.

Our ultimate decision resulted from a compromise between our requirements concerning archiving, input, search and output (as well as other different applications) and the package possibilities themselves, considering that the software must satisfy three main requirements: a flexible file, a simple and rapid search method, free and flexible output formats.

The first and second of these requirements took priority and they led to the selection of the TEXTO software, which presents the following characteristics: it runs with more than 25 operating systems; it has variable field and record lengths, which are almost unlimited (each record can contain up to 99 fields with a maximum of 4000 characters). Each field can be divided into directly retrievable items; its memory is equivalent to that of the hard disk on which it runs; it is compatible with other software (word processing); the parameterization is done directly by the user.

TEXTO

Methodology: Table 1 shows some Library functions which will be activated with TEXTO: some applications have already been tested, others have not. The principal points will be considered in detail, carrying out some examples.

Tab. 1

- CATALOG
- SEARCH
- INVENTORY
- LOAN
- ADMINISTRATION
- OUTPUTS (controlled vocabulary, Library news bulletin, catalog cards, purchasing orders, etc.)

General structure: each TEXTO file must be associated to a Parameter Document (PD), freely structured by the program user and reporting the details of the file and the names of the fields contained in each record. The file and the related Parameter Document are indispensable to TEXTO to satisfy user requirements, and thus they must be assessed with careful attention. However, it is also possible to rearrange them in a second phase.

Table 2 shows the activation procedures for beginning the work and the PD associated to our Library bibliographical file. A short explanation of each field content follows.

Tab. 2

```
Texto Version 4.1.2  20/11/87  Copyright(C) CHEMDATA (87)
Distribution  : CFM - Milano
Licence d'exploitation: Universita' di Genova - Ist. di Oncologia

*/PBIBLIO

*L /PBIBLIO

name     .pbiblio
format   .8 1 : -
fields   .REF INV AU,(X+)INDAU ED TI TLO CICEA AN SEV,(X+)INDSEV
         .VOL CON DCON ISBN NLM (?)PC,(X+)INDPC DF COL DING PR
         .FOR,(4)FORNIT<NOME INDIRIZ> RIS ($)PZ DR NBO NBL NFA
         .SO PS DP DSP DRP NP
```

REF: reference number of records
INV: inventory number
AU,(X+)INDAU: author or editor names, updating index
ED: edited by
TI: book title
TLO: original title and book translator's name
CICEA: city, publishers and publishing year
AN: publishing year
SEV,(X+)INDSEV: serie and its number, updating index
VOL: serie number
CON:information about the congress to which the book refers
DCON: congress date
ISBN: International Standard Book Number
NLM: National Library of Medicine classification
(?)PC,(X+)INDPC: key-word field, it is default during the query, updating index
COL: location number
DING: entry date in the Library
PR: supplier
FOR,(4)FORNIT<NOME INDIRIZ>: file chaining,information about supplier
RIS: who requested the book and his institute department
($)PZ:book price,numerical field
DR: book request date
NBO: order number
NBL: delivering number
NFA: invoice number
SO: progress of the book order
PS:laste person that has borrowed the book and his department
DP: date of he last loan
DSP: date of the loan maturity
DRP: date of book return
NP: total numbero loans

Input: the program allows for a choice between a "line editor" mode (command A) and a "full screen" mode (AW command) for inputting documents. The document numbering may be a manual one (by the user), or else an automatic numbering (by TEXTO). During the adding procedures access to other files is possible (file chaining) which contain information about periodical files, in order to avoid input of identical information repeated in different records in the same file.

Search features: TEXTO search is rapid and may be addressed to a single item or to a character string contained in any field. There are two possible search processes:

- Simple query (Q command): this can be used for an immediate retrieval of relevant documents which may be displayed or printed in edition formats which are specified each time. The Q command allows the user various formats to display the results obtained. TEXTO can list the complete contents of the retrieved documents (Q command alone), or only number of documents which answer the question (Q1 command) or the reference numbers of the documents (Q2 command) or else the reference numbers and the contents of the fields searched (Q3).
- Multi-step query (QUERY mode): each question creates a numbered dollar ($) set which is stored in a memorization file (up to 25 questions) and which can be recalled and combined with other dollar sets to build more complex questions. The output of each question result can be decided at any time during the session, in any print or display format, sorting between the documents retrieved. When it enters in QUERY mode TEXTO abandons the standard prompt (*) and activates another one (?). During the session it is possible to use the main standard commands to list, to create editions, to print or to query in simple mode, as with standard prompt.

In QUERY mode it could be very useful to create subfiles from the source file or to merge several files into one. These files become TEXTO files having characteristics which are defined by one or more questions.

Table 3 shows an example of creation of a subfile called IARC. This file will contain only documents pertaining to books published by the International Agency for Research on Cancer in the years 1986 and 1987. The F command (to create or to merge files) can be used with the standard prompt in the same way.

TEXTO allows for the use of Boolean operators AND, OR, NOT, in both simple mode and QUERY mode and it also permits left, right and internal truncations and masks.

Tab. 3

```
?CICEA=*IARC*
$1          25 Member(s) for  :CICEA=*IARC*
?$1 AND AN=(1986 OR 1987)
$2          11 Member(s) for  :$1 AND AN=(1986 OR 1987)
?L1F
File                          :IARC
Resequence          REF (y/n) :Y
First REF                     :=00000001
Step                          :1
number(s)     :*
```

Index: searching becomes easier and more rapid by creating query indexes (or inverted files), especially when there is a considerable number of documents in a file, due to the costs of interrogation and the response time. These query indexes can be created each time TEXTO has to carry out research in a given field, or can be declared in the PD as updating indexes and thus, every time new documents are added to the file, these indexes are updated.

Table 4 explains the commands to create the query index regarding the AU field with the itemizing specifier (*) before the field name in order to extract every item within the AU field. TEXTO lists in an inverted file all items of this field in alphabetical order and adds the reference numbers of the documents corresponding to each of them.

The query index of PC field allows for building of the thesaurus, i.e. the controlled vocabulary of key words assigned for book indexing. The thesaurus output provides an important aid for searching and it may be matched with the frequency of usage of every key-word (Tab. 5). The vocabulary will be arranged with logical correlations between the different key words by a word processing package that may be used to edit and to rearrange outputs.

Tab. 4

```
*INDEX
Index            (To)                    :INDAU
Do you want a Standard Index (y/n)?      :Y
Field            (From)                  :AU
Maxi lenght for Items          ( < 47 )  :

Answer(s)     72

98  item(s) in your Index
```

Tab. 5

```
*X3
Index            (From)                  :INDPC
1st,last word:

 1 ANTIBIOTICS/adverse effects
 1 ANTIGENS, NEOPLASM
 2 ANUS NEOPLASMS
 2 ATLASES
 1 BILE DUCTS NEOPLASMS
 1 BIOPSY
 1 BONE MARROW/transplantation
 1 BRAIN NEOPLASMS
 1 COLONIC DISEASES/surgery
 4 COLONIC NEOPLASMS
 1 COLONIC NEOPLASMS/diagnosis
 1 COLONIC NEOPLASMS/occurrence
 1 COLONIC NEOPLASMS/prevention & control
 1 COLONIC NEOPLASMS/therapy
 1 COLONOSCOPY
 1 COLOSTOMY
28 CONGRESSES
 1 CONTRACEPTIVES, ORAL/adverse effects
 ..........................................
```

Output: TEXTO allows for various formats for printing, editions and tabulations, which can be temporary (to be used for the output of the search results) or catalogued (these are quite complex and can be used at any time to print a thesaurus, library news, catalog cards, etc.).

The Library provides its users with a quarterly "Library News Bulletin" which lists newly acquired books according to subject matter. TEXTO extracts from the source file, searching with combined commands, for the most recently purchased books which have already been indexed and catalogued and subsequently, the Library news output can be printed with a suitable edition; as with the vocabulary, it is possible to arrange this by means of the word processing package (Tab. 6).

Tab. 6

CYTOLOGY

PACKER L-DOUCE R
PLANT CELL MEMBRANES
SAN DIEGO,LONDRA:ACADEMIC,1987
METHODS IN ENZYMOLOGY,VOL.0148
COLL

DNA, RECOMBINANT

WU R-GROSSMAN L
RECOMBINANT DNA.PART E
SAN DIEGO,LONDRA:ACADEMIC,1987
METHODS IN ENZYMOLOY,VOL.0154
COLL

ENVIRONMENTAL POLLUTION

HUTZINGER O
AIR POLLUTION
BERLINO,NEW YORK:SPRINGER VERLAG,1986
HANDBOOK OF ENVIRONMENTAL CHEMISTRY(THE),VOL.0004 PART A
WA I 029

The online output of the catalog cards is carried out combining several files and catalogued editions. Table 7 shows an example of a catologued edition and a catalog card according to the author.

Tab. 7

```
*L/ESCHEDA

name    .escheda
fields  .'' (=2d)COL '' TI '' ED '' AU '' CICEA '' SEV '' DF ''
        .TLO '' CON ISBN INV (=2)NLM '' PC ''
line    .70
margin  .0
sheet   .18
page    .18
skip    .Y

Answer(s)     1

PIPER DW                                                    WI 024

STOMACH CANCER.A SERIES OF WORKSHOPS ON THE BIOLOGY OF HUMAN
CANCER.REPORT N.6
edited by
PIPER DW
GINEVRA:UICC INTERNATIONAL UNION AGAINST CANCER,1978
UICC TECHNICAL REPORT SERIES,VOL.0034
viii,138pp:ill,cm24
Workshop on STOMACH CANCER,GINEVRA 12,16/04/1977
ISBN    :92 9018 034 X
INV     :0682DV
NLM     :WI 320

STOMACH NEOPLASMS-CONGRESSES-NEOPLASMS, EXPERIMENTAL
```

Administrative applications: the automation of administrative Library procedures, using the same TEXTO commands, should allow for:
- new acquisition management, in order to immediately verify if a new book requested for buying already exists in the Library or if there are other similar books; input into the Library file of all the information regarding purchase of the book (the person who requested it, the supplier, its price, the order number, etc.); checking of the delivery time and the completion of the book data when it arrives (for indexing and library news);
- loan management, in order to input information about those that have borrowed the book and the pertaining date; to verify the loan maturity date, the fact that the book has been returned and if the loan has been changed or renewed. It is possible immediately to verify which books are most requested for borrowing;
- inventory management, which makes the book inventory redundant and allows for a rapid checking of the Library holdings, with the possibility of obtaining the sum of numerical fields, declared in the PD, like our PZ field ((\$)PZ).

Output formats have been made, or are in the process of being made, for each Library service, in order to improve and complete the automated management (letters for book orders, loan cards, letters to press users for loan return, etc.). These editions will be brought about by an interaction between TEXTO and other packages, or other TEXTO applications (addressed file managed by TEXTO).

EMPLOYED STAFF

Staff with degrees: indexing, choices, start and step control of the automation program are carried out by two biologists of the Scientific Information and Documentation Service. The work began in September 1984; each person has spent about 20% of his yearly working time on this (at present n. 1600 volumes have been treated).

Administrative staff: data was input by administrative staff. The work started in February 1988. About three hours per day are spent for this work (at present n. 1200 bibliographical records have been input).

CONCLUSIONS

On the basis of our acquired experience, the following considerations may be made:
- A detailed analysis of the environment and of the users' needs has to be carried out.
- The attitudes of the staff using the system must taken into consideration.
- An attentive examination of Library materials and a careful indexing of the texts must be carried out.
- The system must constitute a basic and familiar working instrument, both for the users and the administrators.
- The automated file has to be used regularly.

The last two points are more difficult to understand when the new system is introduced in a traditional Library. It must be made clear that automated management represents an improvement of Library activity, both for the quality of information and for the user.

At present we find ourselves in a transition period and we are checking the system. A catalog in card form does still exist but we think it will be eliminated later on.

Regarding the file structure the initial choices have been revised and modified more than once. The output options have been studied with particular attention and some fields have been duplicated in order to build useful search indexes.

TEXTO has proved to be of great flexibility for new requirements which were not foreseen in advance, but in the outputs option it does present some limitations, however. The program has been studied in an overall fashion. Some functions have been examined closely, others will be well studied later on, but on the whole TEXTO does seem adequate for meeting our requirements and resolving our problems.

# TESTING THE MICROSAILS SERIALS CONTROL SYSTEM IN THE CLINICAL RESEARCH CENTRE LIBRARY

Jenny Wade and Philip Emery
MRC Clinical Research Centre
Watford Road
Harrow HA1 3UJ, Middlesex, Great Britain

## Background and test objectives

The Clinical Research Centre (CRC) at Harrow, in north-west London, is one of the major research institutes of the British Medical Research Council, and its library serves the needs of research scientists and the whole range of staff of the large district general hospital on the same site. The library has eleven full-time staff; its serials collection is moderately large, with over 650 current titles, and in a research environment it is naturally a most important component of the information services provided. A collection of this size presents particular management problems, being small enough to make a standard Kardex record system perfectly convenient in most respects, but large enough to make regular checking through the records, for example to claim missing issues, tedious and time consuming.

Several areas of the library had already been automated, including various housekeeping functions and, of course, information retrieval. We were interested in the possibilities of the various microcomputer-based serials control systems available, but none of those that we had seen had persuaded us that the benefits of introducing automation in this particularly complex area would outweigh the resulting problems.

For some years the majority of the CRC's journals subscriptions had been handled by Swets (UK), Ltd., based at Abingdon, near Oxford. In May 1987 we were aproached by staff of this company to assess our willingness to become the British test centre for their new microcomputer-based serials control system, known as MicroSAILS, since it had been developed from their mainframe/mini system, SAILS. This was an interesting proposition for a number of reasons, not least because what was envisaged was practical development testing before marketing, with the aim of evaluating and refining the system in a real working environment and thus avoiding many of the problems that had occurred with other systems. In addition, the opportunity to test such a system and input ideas for modifications and

development based on real requirements had great professional interest for us. These aspects of the proposal had of course to be weighed against the inevitable disruption to work patterns in the busy serials department, and the unknown commitment of staff time that would be required.

The decision to take part in the development test was made in August, and the initial time schedule was that the test should run from October 1987 until March 1988. It was important for the company and the library to complete the test as soon as was compatible with good results, but in the event we have had to extend this period for various reasons.

The library staff were to test the basic modules of the system already written, and new modules as received. We knew that initially the system would not be complete, and accepted that testing was very much experimental at this stage. Swets provided intensive induction training for the staff involved, as well as system documentation. From the first day a detailed log was kept noting faults and problems in the system as well as ideas for possible improvements, and there were regular meetings and discussions with company staff throughout the trial. The distance between the CRC and Swets offices did not result in any great difficulty, at least for the library staff.

**The system and experience of its use**

The MicroSAILS system was installed in the CRC library at the end of October 1987, on an Amstrad PC 1640 microcomputer. It is in fact designed to run on any IBM or compatible machine, as a single user system or part of a local area network. The system is menu-driven, with numerous sub-menus and functions. We shall not provide a detailed technical description here, but the main features of MicroSAILS available at the time of testing were:

Ordering: The system includes a facility for raising orders for new journal titles, single issues, etc., and will print order forms for sending to suppliers. Order records can be tagged in various ways, for example with requestor's name.

Pregeneration: For journals published on a regular basis there is the facility of automatic pregeneration, whereby once frequency details for the title have been entered future issues expected can be pregenerated automatically.

Checkin: Checking in journals received is simply a matter of confirming the pregenerated issues expected information. Screens of data for issues

checked in and expected can be retrieved for each journal entered on the system.

Claiming: The system will generate lists of claims for expected journal issues that have not been received, arranged by order of supplier.

Bibliographic and supplier details: Such data can be entered using the bibliographic updates menu and changed as necessary.

Library patron details: For each title circulation lists can be produced and details of the initial requestor entered.

Financial administration: It is possible to allocate journal invoices to specific account funds. A more complex fund accounting facility will be available later.

In the first two weeks of operation in the library a few test titles were entered and most of the functioning menus as outlined above were tested. After this initial trial, which had also facilitated staff training, journal titles, bibliographic details and issues were entered on a daily basis as issues arrived in the library. A major problem soon emerged, in that the systems memory management protocols and database management utilities were not as sophisticated as was required for long periods of use on a day-to-day basis. This had the effect of producing an unacceptably slow response time as well as all too frequent software 'crashes'. Furthermore, serious data corruption occurred owing to a software bug, which had hitherto remained undetected. As a result it was necessary for Swets to remove the system from the library for about four weeks to make modifications.

**Advantages of the system**

After this rather discouraging start we were able to make considerable progress in our practical use of the system and found many attractive features, notably the following:

Pregeneration: The use of automatic pregeneration for future issues of regular journals is one of the most effective aspects of MicroSAILS, and permits very fast checking in of new issues.

User friendliness: The system proved particularly easy to learn and use after the most basic training. Routing, help and error screens are available within each menu, in case the user has any problems, and these actually removed the need to consult the printed documentation except in one or two rare instances. Also, the provision of a test database enabled staff to be trained without fear of corrupting real data.

Modification: In respect of modification the system is very flexible, and can be tailored to the individual library and user's needs. Modifications can be made for the layout and wording in all menu, routing, error and help screens, and to system and function messages. In addition the screen colours, text colours and level of access for each user can be pre-set and changed as required.

Keyword searching: Three different methods of searching for data on journal titles, suppliers and library patrons were provided. The availability of searching on truncated versions of words proved an effective time saving device.

**Problems**

Because of the nature of this development trial problems arising in use were to be expected, since the objective of the project was to refine the system before marketing it. Response from the company to problems with the system and to practical improvements suggested by the library staff was excellent, and in this respect the difficulties experienced produced positive results. The major problem areas were as follows:

Response time: One of the main difficulties was the slow response time of the system. Checking in the daily intake of new journals, for example, at first took twice as long as entry in the manual Kardex system, and entering bibliographic details for a journal new to the database took about twelve minutes. Improvements have since been made to the response time, but it is worth mentioning that one reason for this problem seems to have been the hardware used. Response was faster on the Olivetti microcomputer used at Swets office than on the Amstrad PC 1640 used at the CRC.

Irregular journals: A surprising number of journals appear at irregular intervals or with idiosyncratic publication schedules. A well known example is Biochimica et Biophysica Acta, with 127 issues a year and higher volume numbers appearing before lower. In these cases it is not possible to use automatic pregeneration for expected issues and time-consuming manual pregeneration must be used. Here one may compare the ease with which publication freaks of this kind may be handled by a Kardex system.

Deletion of mistakes: Although it was easy to modify layout and wording of screens, it was not so easy to alter input data if an error had been made; for example, if a journal title had been ordered twice by mistake it was impossible to delete one of the orders. Mistakes in the database could also be caused by software problems; for example future issues may be

pregenerated with the wrong date or with no copies expected, so that the issue is checked in as an extra copy. Again there were problems in rectifying such errors.

Stability: As already described, we experienced problems with the stability of the system, and failures in the memory and COBOL programming led in some cases to the corruption or loss of input data. This problem was later rectified.

Ergonomics: Staff using the system experienced some problems with the colour graphics, with particular complaints about brightness, despite the availability of brightness and contrast controls. Headaches and eyestrain resulted when the system had been in use for some time, and this led us to doubt the value of the initially attractive colour graphics. However some of this problem may again have been hardware-related.

**Management aspects**

Much has already been written about the impact of the introduction of automated systems into libraries, and certain clear guidelines are available for the manager to follow in choosing, installing and operating a commercially available system. However, taking part in a product development test is a somewhat different proposition. The library concerned has to maintain its existing systems and service level while running the test system, without any definite prospect of concrete benefits at the end of the test period. There is a considerable extra workload, which will be difficult to maintain when the inevitable staffing crises occur. Balanced against this, the experience of sharing in the development of such a system is of very considerable professional interest, as is the opportunity to work closely with the commercial sector and to foster mutual understanding of the constraints and demands of the working methods of each.

The main area of impact of MicroSAILS on the CRC library has been on the three staff members directly involved with the trial. Training has not so far been extended to all staff, largely because of the regular modifications to the system resulting from the trial. Further staff training will probably be required in the near future, and will certainly be necessary if it is decided to adopt the system permanently. The serials staff have had some difficulty in keeping up-to-date with their regular routines for maintaining the collection in addition to their work on the trial. In retrospect it would have been more satisfactory if an additional

part-time library staff member had been employed during the test period, but the uncertainty of the time commitment required made it virtually impossible to take this decision in advance.

The danger of a too heavy workload for the library staff was that service to readers would suffer, and there was certainly a reduction in staff time spent on certain journal maintenance procedures. Otherwise the impact on library users has been minimal, since the Kardex records have been maintained and there has been no public access to MicroSAILS. If this were introduced at some later date a whole new range of problems would need to be addressed, but these do not concern us at present.

**The future**

Although we have not yet made a decision to adopt MicroSAILS or any other automated serials control system on a permanent basis, our experience with it has done much to allay our doubts as to the desirability of using such a system in a medium-sized library of our type. Future prospects for the system are attractive, and we hope to continue working on it with Swets for some time to come. Major developments which have been introduced since the initial installation include the implementation of more sophisticated searching facilities to allow exact match and keyword searching as well as browsing. There has been a considerable improvement in response time, and new menu options have been fed through regularly. In the pipeline for inclusion in the first release are a more sophisticated financial administration module and hopefully better management information. It is also planned to link MicroSAILS with the DataSwets system, providing access to the company's main database of bibliographic, publisher and subscription information, as well as a mailing service.

We should like to close by emphasising the positive professional benefits to be gained from involvement in this type of test, which is demanding but should lead to an end-product free of many of the problems that beset systems designed and tested only in a more restricted or artificial environment. The element of risk and inconvenience for the testing library should be offset in the short term by the inherent interest of the project and in the long-term by improved automated systems, more effectively designed to meet the real needs of library activities. The basic requirements for all involved are enthusiasm, imagination, commitment and patience.

# Concurrent session 5C

# End-user searching

Chair

S. Adamic

# ARTIFICIAL INTELLIGENCE IN VALUE ESTIMATION OF DATABASES

Rodes Trautman and Sara von Flittner

The authors conducted this work at the Department of Library Science and Information Science, Abo Academy, 20500 Turku, Finland, under a grant from the Academy of Finland.

The use of online databases has grown tremendously during the last decade. Today there are more than 4000 publicly available databases, and at least 500 vendors (Cuadra, 1988). The number of available databases grows with about 10% per year. These databases represent an enormous amount of information, and the individual user needs effective ways of choosing the right database to find an answer for a question. Reference librarians had no problem with database selection a decade ago when only a small number of online databases existed but today, even information specialists are interested in help. Even though Evans (1980) found that, in each multiple database search, one file, termed a principal online database, generally gives a sufficient number of references, we believe that other databases may offer more pertinent information for a particular end-user. Furthermore, we believe that a wider selection of online databases would be used, by both professionals and end-users, if expert systems that give guidance were available. Currently, database selection aids range from printed directories to online term frequency indexes. Problems with online aids are that (1) they are expensive to use, (2) some cover only one vendor, and (3) none offers advice. Our artificial intelligence system is intended to address these problems, being an offline selection aid that extends across all vendors and all subjects.

The prototype system described in this paper is to be used offline to select a small group of likely databases for answering the actual research question online. There are a large number of journal articles relating to database selection. These journal articles reveal similarities and differences within a narrow group of related databases. In contrast, directories cover most databases, but with less discriminating detail. They have subject indexes of 500 to 2000 terms, varying from specific words like "wine," with just four databases assigned, to broad terms like "economics," with hundreds. All these sources represent "experts" who have published about database selection. Our "knowledge acquisition" consisted of extracting from their printed record only the comments or parts pertaining to database selection itself, capturing all the different methods used by these experts.

The 15 online vendors included handle about a fifth of all the available databases. "IBM PC and MS-DOS compatible" microcomputers were used and a computer with a 10mb fixed disk and full 640k primary memory was essential. GURU (Micro Data Base Systems, Inc., P.O. Box 248, Lafayette, IN 47902) was selected as the artificial intelligence developmental "shell." Our program represents an application that can also be executed with a less expensive "run-time" version. The program consists of command files in the GURU procedural language, which is a very comprehensive and powerful computer language. The same command mode language is used across all of the individual modules and permits sharing of data between them. The current version of the program uses word processing, database management, and expert system

modules. The "scheduler" is a script performance file that has the main menu and that calls the various performance files for each option (Croft, 1987). Each of these has (1) a menu display, (2) optional HELP display, and (3) optional ADVICE consultation. The consultations are rule-based expert systems in their own right that provide a "doctor- patient" interaction to help the user select an appropriate option. These internal experts are:

1. The "user modeler," which determines the depth of interest, the time, language, and geographical coverage desired, and the user's intended application of the assorted kinds of information available online.
2. The "question clarifier," which narrows the user's need to a few of its 240 viewpoints.
3. The "searcher," which makes a Boolean search of an internal database that has nine coded characteristics for each of the online databases known to the program and then makes appropriate data available to the user and the "ranker" expert.
4. The "ranker," which presents the unique or "special features" information about each likely online database from a separate internal text database and asks for pertinence evaluation.
5. The "evaluator," which sums up the session with a ranked list of recommended databases.
6. The "browser," which provides the user with an alternative method of access, such as scanning of the internal files or getting statistics about them or the system.

The system has a scheme for classifying databases via nine attributes. The first eight define the database kind, its time and geographic coverage, etc. The ninth is a subject classification defined in the prototype as a three-level hierarchical structure for all knowledge. The category selection codes are controlled to make retrieval reliable and to reduce the size of files. The subject classification scheme was developed by analyzing subject indexes of standard online database directories. The editors of the directories, experts in database selection, have assigned subject terms (descriptors) to the online databases. Each of these terms represents the name of a family for which its members are already listed. We call examples of all of these by the generic term "viewpoints." As an aid to the user, we limited our categories in all levels to a number that could fit, together with brief explanations or examples, on a single 25-line screen. Ten first-level categories of viewpoints were adopted and have these brief names: TIME, FACTS, QUALITY OF LIFE, COMMUNICATIONS, COMMERCE, TECHNOLOGY, ADMINISTRATION, LIFE, ENVIRONMENT, and REALITY. Further explanation of the categories, shown in italics, can be obtained through the HELP option and expert consultation obtained through the ADVICE option.

Each of the first-level categories has a variable number of second-level categories that range roughly from specific to general in order to encourage the user to pick a narrower category first. The subject indexing of the directories is often very narrow; consequently, the few databases listed are those to which the editor wishes to call attention. A serious concern is that there are major, comprehensive databases that also have a considerable amount of information on the same narrow topic. Such databases we call "principal databases". We used our own judgment in assigning our viewpoint categories to these, introducing a special truncation code to indicate their broad coverage in either the second or third levels, or both. The program retrieves four sets of online databases: (1) those that share the user-selected generic viewpoint (the ninth

attribute), but are restricted to those that meet the user's requirements specified by the first eight attributes; (2) those that have the same viewpoint, but without the user's restrictions; (3) principal databases with the same viewpoint; and (4) principal databases with the next broader viewpoint. The postings are displayed first. The user can decide which set to inspect, whether to restart, or whether to ask the "evaluator" expert for advice. For whatever set is selected, the program displays specific, uncontrolled text that pinpoints differences or special features that distinguish one member from another. In each screen display, the user may ask for more data in order to assign the online database into one of three categories: (1) pertinent, (2) not right, or (3) not now. The last category provides a list of those online databases that didn't meet the user's immediate information need, but awakens some other interest. The second category represents "false drops." These might be attributed to data entry coding errors. Alternatively, they could be "relevant" to the search statement, but not "pertinent" to the user's possibly subconscious information need (Fugman, 1985). The final display is a ranked list of the pertinent online databases with a few of the attribute mnemonics listed for each.

Our "database" is actually several relational database files about our subject: online databases. This internal database is the heart of our system and the rule-based internal experts provide user-friendly access to it. Consequently, we can take advantage of sophisticated database management techniques that have been developed over the years and can incorporate new artificial intelligence concepts as they emerge. Often, such a clear distinction between conventional database and rule-based schemes is not made. In fact, many expert systems have the underlying data completely embedded in the rule structure, with the term "knowledgebase" used to denote the aggregate. Dividing the overall problem into parts and assigning each to an internal expert, besides clarifying our thinking, facilitates independent programming of the parts in as much detail as desired and independent updating as the project progresses. Because the database underlies reasoning in the program, it was desirable to include fields and manipulations that have overtones of current artificial intelligence thinking, such as:

1. A value judgment (rank) for each attribute field and an overall ranking for each database formed by dynamically combining individual attribute ranks into a single value at the time of display. The user can modify this overall ranking algorithm by adjusting the weights assigned to the attribute ranks.
2. Coded records for speed and quality control, and free-text, variable-length records for uncontrolled keyword descriptions that might spark the user's imagination.
3. An ADVICE function, provided by consultation with one of the several internal expert systems, as well as a HELP function provided in a conventional, context-sensitive manner. The user can opt to select HELP or ADVICE, or neither, at several points during the session.
4. Browsing of indexes and internal text files, as well as alternative methods of subject classification.

These options all enable the user to exercise a modicum of common sense. Common sense has so far eluded programming and so is lacking in the current generation of artificial intelligence programs. We treat users as "subject specialists" who know when they are not getting the information they need and will pursue other options offered.

The complete viewpoint classification schedule is a composite index, by generic categories of user information needs, to those directories included. It shows

in a glance how various editors differ in their assignment of subject terms. But this is the essence of the artificial intelligence approach. We wanted to capture the various ways in which different human experts view the problem. Our procedure captures the concepts deemed important by those already familiar with online databases. We want to facilitate new conceptual linkages in subject categories. and interdisciplinary linkages in subject categories. The menu-driven program is intended to avoid getting too enmeshed in details of the classification (contact either author for information concerning a discette with our current classificationschedules). Our ultimate goal is to get more online databases to be used -- or at least to be considered. The figures 1-3 show the main menu, the viewpoints in the first level, and the given information about one database.

---

WHAT DO YOU WANT ME, AS A COMPUTER, TO DO FOR YOU?

Step
1. CONTINUE from last session
   UPDATE your coverage restrictions -- New users pick this one
2. ASK you about your current information need -- YOUR INQUIRY
   SHOW you what I know about specific databases -- FREE-TEXT BROWSE
3. EXPLAIN how my program works
   LEARN from your comments or from supplementary data
4. QUIT with optional saving of current session

---

Fig.1 The main menu. Cursor movement ("arrow" keys move a highlight bar line-by-line through the options. You begin the search by picking one of the two options in Step 1. In Step 2 you learn something about databases from the computer, normally by allowing the computer to probe about your information need, but alternatively, by allowing you to look up a specific database or to browse. Step 3 provides you an opportunity to find out how the computer reached decisions, especially useful if you disagree; even if you don't care to know, the computer (programmers) want your feedback, and the second option in Step 3 puts you into the word processor for your comments or permits you to add your own records to the system. Step 4 ends the session. Any option may be selected at any time the main menu is displayed.

---

CAN YOU CHARACTERIZE YOUR INFORMATION NEED BY ONE OF THESE VIEWPOINTS?

| | | |
|---|---|---|
| A-TIME | past (history), present (news) & future events (schedules, forecasts, investments, citations) | |
| B-FACTS | organisations, people, places, products, terminology (library & museum holdings, funding research) | |
| C-QUALITY OF LIFE | culture, literature & leisure time activities (entertainment, build & run a home | --- ARTS & HUMANITIES |
| D-COMMUNICATIONS | linguistics, education, publishing, computing (natural & machine languages, library | ---INFORMATION SCIENCE |

| | |
|---|---|
| E-COMMERCE | domestic & international business & trade (banking, investment, consumer affairs) ---ECONOMICS |
| F-TECHNOLOGY | manufacturing, producing, transporting products (construction, natural & synthetic materials) ---ENGINEERING |
| G-ADMINISTRATION | corporate & governmental management & human rights (labor, national & international law, patents) ---LAW |
| H-LIFE | plants, animals & humans (sickness & health) (agriculture, human & veterinary science) ---BIOSCIENCES |
| I-ENVIRONMENT | global view of our surroundings & survival (ecology, pollution, energy, space) ---EARTH SCIENCES |
| J-REALITY | physical & philosophical nature of matter & thought (pure science, logic, philosophy) ---INTERDISCIPLINARY FIELDS |

HELP ADVICE

---

Fig.2 Screen display of the ten first-level categories of the Viewpoint attributes. (In parenthesis optional help examples).

---

2 out of 9 AGELINE (aging) Rank 38

Contains references to and abstracts of materials on aging and the aged. Covers psychological, medical, economic and political concerns.

(AC 05780:AARP vendor
NA AGELINE
PE From January 1978; UPDATED every two months
SO journal articles, books, dissertations, symposia,
SO papers produced in research organisations)

Pertinent Not right Not now More data Debug New view Main menu

---

Fig. 3 The result of the selection of the right databases is a ranked list of databases. One of them is shown in the figure. In parenthesis More data.

References.

Croft, W.B. & Thomson, R.H. I R: A new approach to the design of document retrieval systems. Journal of the American Society for Information Science 38 (3): 389-404, 1987.
Cuadra. Directory of online databases. New York:Cuadra/Elsevier.
Evans, John Edward. Database selection in an academic library: Are those big multi-file searches really necessary? Online 4 (2): 35-43, 1980.
Fugman, Robert. The five-axiom theory of indexing and information supply. Journal of the American Society for Information Science 36 (2): 116-129, 1985

# MEDLINE A LA PORTEE DU PLUS GRAND NOMBRE GRACE A UN SYSTEME D'INTERROGATION SIMPLE ET GUIDE EN MODE VIDEOTEX

Jacques HALPERN, Paulette DOSTATNI
IMA, Centre de Documentation de l'INSERM
Hôpital de Bicêtre, 94270 Le Kremlin-Bicêtre, France

*
* *

ABSTRACT : MEDLINE AVAILABLE TO A GREATER NUMBER OF USERS THROUGH A SIMPLE USER-FRIENDLY SEARCHING SYSTEM ON VIDEOTEX

Today, biomedical scientific information takes place among services provided to practitioners through the Minitel, a cheap and easy to use videotex terminal which is more and more widely distributed. A user-friendly interface, developed by INSERM and the host Telesystemes/Questel, allows users who have no training or experience of information retrieval to conduct their own bibliographic searches on MEDLINE, a very essential information tool, almost at any time of the day or night.
Using a menu approach and guide screens, the user enters specific keywords (either French or English), retrieves in a few minutes the current or main citations on the subject of his choice and may display the corresponding abstracts and order directly the copy of primary documents. While remaining simple, the interface system strives for the quality of searching, taking into account the best features of MEDLINE with matching of keywords to Mesh terms, use of exploding procedure, application of subheadings and distinction of most relevant citations.
MEDLINE Minitel is suitable to simple searches limited to one or two concepts. For those who want to deal with complex questions or reach a high level of relevance and exhaustivity, it is recommended to require traditional, sophisticated searches from a well-trained specialist of MEDLINE/Questel Plus language. MEDLINE/Minitel and MEDLINE/Questel Plus appear as two options proposed by INSERM and Telesystemes/Questel, useful to answer the information need of the biomedical community. MEDLINE/Minitel gives a new dimension by its practical approach suitable for immediate or current retrieval. It is a widening of the access to MEDLINE.

## INTRODUCTION

Le professionnel de la santé, chercheur ou praticien, a un besoin croissant d'une information scientifique disponible rapidement et directement, sans qu'il ait à se déplacer ou à passer par des intermédiaires (bibliothèques, centres de documentation). Il dispose de plus en plus fréquemment, à son domicile ou sur son lieu de travail, d'un Minitel ou d'un micro-ordinateur.
Conscient de cette situation nouvelle, l'INSERM vient de mettre au point, avec le serveur Télésystèmes-Questel, une interface conviviale facilitant l'interrogation par l'utilisateur final de la grande base de données bibliographiques internationale MEDLINE. Il s'agit d'un accès simplifié, en français ou en anglais, utilisable à partir d'un Minitel, terminal videotex français ou d'un micro-ordinateur avec émulation Minitel.

Après avoir décrit les caractéristiques de cette interface, nous examinerons les modalités d'accès de ce nouveau service destiné tant au médecin individuel qu'au service hospitalier, à la clinique, au laboratoire ...

## 1. POURQUOI UN ACCES SIMPLIFIE ?

L'interrogation des bases bibliographiques est jusqu'à présent restée l'apanage des documentalistes, spécialistes de l'information. Elle implique, en effet, l'utilisation de logiciels de recherche documentaire relativement complexes. Chaque logiciel, spécifique à un serveur documentaire (QUESTEL, DIALOG, ESA, NLM...), comprend un ensemble de commandes indispensables à la consultation; les éléments de recherche sont regroupés en équation en suivant les principes de la logique booléenne : il faut manipuler les opérateurs booléens dans des conditions particulières au logiciel. Les index et champs sont organisés, libellés et présentés différemment selon la base et le serveur.
Pour maîtriser ces différents éléments, il est nécessaire de suivre des séminaires de formation, d'avoir des manuels techniques... c'est une affaire de spécialistes !
La base de données MEDLINE comporte en elle-même un ensemble de particularités qui rendent son interrogation difficile ; elle implique notamment l'utilisation d'un thésaurus en anglais, le MeSH (Medical Subject Headings) de 15.000 mots-clés, complétés par de nombreux synonymes et classés thématiquement dans une hiérarchie à 7 niveaux. Les mots-clés se combinent, selon des règles déterminées, avec des qualificatifs (au nombre de 77), ce qui permet d'effectuer des recherches précises.
L'interrogation directe de MEDLINE par l'utilisateur final impliquait donc la mise au point d'un système supprimant les procédures complexes des logiciels documentaires, permettant une interrogation en français et facilitant l'utilisation du thésaurus. Telles sont les spécifications principales de l'interface que l'INSERM vient de réaliser en coopération avec le serveur Télésystèmes-Questel.

## 2. CARACTERISTIQUES DE L'INTERFACE MEDLINE MINITEL

### 2.1. Définition

Il s'agit d'une interface installée au niveau de l'ordinateur serveur (host computer-based software), qui traduit en langage Questel Plus les demandes que l'utilisateur exprime à travers des écrans-menus qui lui sont proposés, et vice-versa du langage Questel Plus en langage utilisateur. L'interface est conçue pour le Minitel comme terminal d'interrogation.

## 2.2. Accès à partir d'un Minitel ou d'un micro-ordinateur

Le Minitel est l'outil privilégié d'interrogation en raison de sa grande diffusion (près de 30.000 médecins en sont équipés en France), de sa facilité d'utilisation et des possibilités de gestion qu'il offre (système kiosque télématique).
L'interface utilise largement les touches de fonction du clavier du Minitel ; par contre, les possibilités graphiques de l'écran, qui ralentissent la consultation, sont employées modérément.
Un micro-ordinateur, doté d'un modem et d'un logiciel de télécommunications avec émulation Minitel, peut également être utilisé, avec l'avantage de pouvoir enregistrer les informations sur support magnétique.
Les références bibliographiques sélectionnées peuvent être expédiées par courrier du serveur à l'utilisateur qui peut donc consulter MEDLINE sans être équipé d'imprimante.

## 2.3. Système par menus

L'utilisateur est guidé dans l'interrogation par des écrans-menus qui lui proposent les différentes possibilités de recherche.

Trois menus principaux sont utilisés :
- le menu "période consultée" : la recherche peut être limitée aux trois dernières années, aux cinq dernières années ou effectuée sur toute la base (depuis 1971)
- le menu "critères principaux" (fig. 1) présente les cinq éléments fondamentaux de recherche :
  1. mot-clé français
  2. mot-clé anglais
  3. mots libres anglais
  4. auteur
  5. périodique
- le menu "critères secondaires" ajoute trois éléments complémentaires pour limiter une recherche :
  6. langue française
  7. espèce humaine
  8. références prépondérantes

La sélection d'un critère fait apparaître un écran de saisie spécifique qui comporte un guide facilitant l'introduction de l'élément de recherche. L'environnement alphabétique du terme saisi est affiché sur l'écran suivant, ce qui permet à l'utilisateur de choisir le terme autorisé ; s'il sélectionne plusieurs termes, ceux-ci sont automatiquement reliés par l'opérateur logique "ou". La recherche est alors effectuée et le nombre de références sélectionnées est indiqué.
Le choix d'un second critère limite la recherche puisqu'il sera combiné au précédent avec l'opérateur logique "et". La recherche peut ainsi se poursuivre par étapes successives, aboutissant à une sélection de plus en plus fine.
La logique booléenne est prise en compte de façon "transparente", sans manipulation des opérateurs.
A chaque étape, les références sélectionnées peuvent être visualisées.

## 2.4. Apprentissage en ligne

Un manuel comprenant une trentaine de pages-écrans réparties en plusieurs chapitres, peut être consulté soit séquentiellement avant toute recherche, soit ponctuellement à une étape de la recherche, l'accès se faisant directement au chapitre pertinent.
Des guides de saisie et des messages signalant des erreurs d'interrogation complètent l'apprentissage du système.

## 2.5. Utilisation privilégiée des mots-clés du thésaurus

Le MeSH, thésaurus de MEDLINE, a été entièrement traduit en français. L'interrogation peut donc se faire indifféremment avec les mots-clés français ou anglais.

La liste alphabétique des mots-clés est systématiquement affichée dès la saisie d'un mot-clé ; c'est une liste permutée, ce qui facilite l'accès aux nombreux mots-clés composés. La synonymie est automatiquement gérée par l'interface qui affiche l'environnement alphabétique du mot-clé correspondant au terme entré. L'utilisateur est ainsi placé dans les meilleures conditions pour trouver le mot-clé exprimant son sujet.
La hiérarchie est traitée selon deux principes :
- l'utilisateur a la faculté de visualiser les mots-clés spécifiques du terme choisi et de sélectionner l'un ou l'autre de ces mots-clés (fig.2)
- si l'utilisateur ne visualise pas la hiérarchie, tous les mots-clés spécifiques du terme choisi sont retenus.

Les qualificatifs, qui permettent de se limiter à un aspect précis d'un mot-clé, sont affichés systématiquement, afin d'en faciliter le choix. Le système ne propose que les qualificatifs autorisés pour un mot-clé donné.
Trois autres possibilités d'affiner la sélection des références sont proposées après l'obtention des premiers résultats de la recherche : limitation aux articles de langue française, aux articles cliniques (espèce humaine), références prépondérantes (pondération d'un mot-clé).

## 2.6. Information fournie

La recherche sur Minitel permet :
- de visualiser des références avec résumé sur le Minitel
- de recevoir par courrier une liste complète de références
- de commander les photocopies des articles

Une procédure très simple est utilisée pour la commande des photocopies des articles : chaque écran présentant une référence comporte le message "commande du document : C ENVOI" ; il suffit de taper C (ENVOI) pour que la commande de photocopies correspondant à la référence visualisée soit enregistrée.

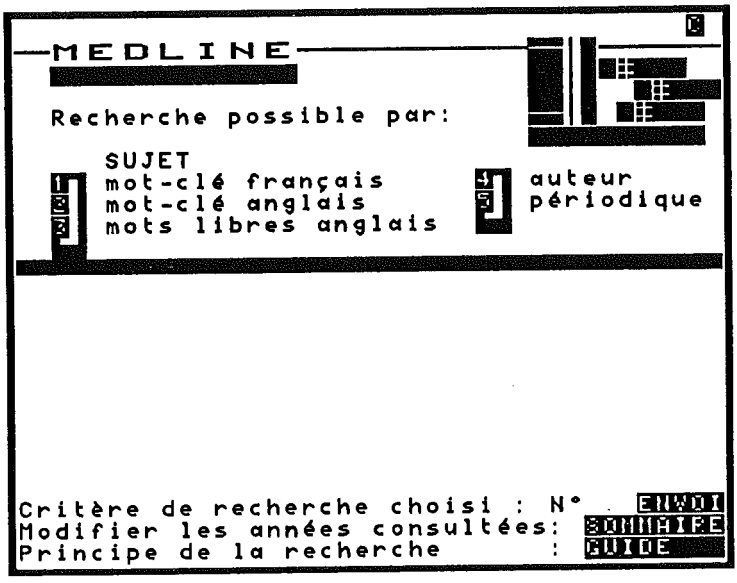


Figure 1 : Ecran menu critères principaux

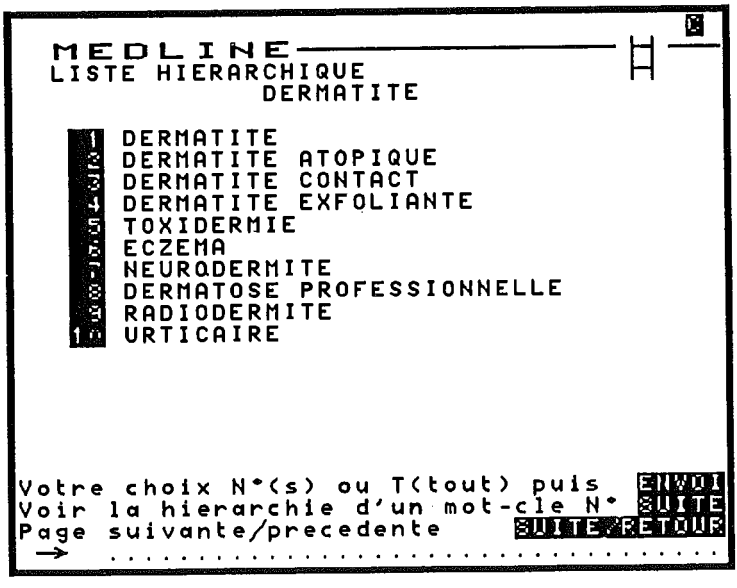


Figure 2 : Visualisation de la hiérarchie du MeSH

## 3. CONDITIONS D'ACCES ET TARIFICATION

Deux modes d'accès sont proposés :

- 36.29.00.36 : kiosque professionnel. C'est un accès direct, sans abonnement, destiné à l'utilisateur individuel et occasionnel. La facturation est faite par les PTT avec la facture téléphonique. Le tarif est de 7,64 HT F par minute; il comprend la visualisation des références sur Minitel et l'envoi par courrier de 15 références par recherche.
- 36.13 code INSERM : sur abonnement, un mot de passe est nécessaire. Cet accès est destiné à l'utilisateur régulier ou institutionnel. L'abonnement minimal est de 2.000 F HT ; il permet de bénéficier d'un tarif réduit, d'une assistance technique de l'IMA et de recevoir gratuitement le vocabulaire de poche et le guide d'utilisation. Le système comprend un compteur à décrémentation qui indique le solde de l'abonnement à chaque connexion ; le compteur prend en compte trois paramètres de tarification :
  . temps de connexion
  . écrans-références visualisés
  . références expédiées par courrier (sans limitation)

## CONCLUSION

L'interface conviviale MEDLINE Minitel permet à l'utilisateur non initié aux particularités des langages documentaires, d'interroger MEDLINE seul, où il le souhaite et quand il le souhaite, en français ou en anglais, et d'obtenir rapidement une information à jour.

Tout en devenant simple, le système interfacé s'est efforcé de préserver la qualité de la recherche en prenant en compte les principales possibilités du thésaurus.

MEDLINE Minitel convient à de nombreuses recherches bibliographiques relativement simples, limitées à quelques concepts. Le recours à un spécialiste interrogeant en mode expert avec le logiciel Questel Plus et connaissant toutes les finesses de MEDLINE, est conseillé pour traiter rapidement des questions complexes et atteindre un haut degré de pertinence et d'exhaustivité.

Ces deux accès complémentaires à MEDLINE, interface conviviale et mode expert, répondent aux besoins d'un plus grand nombre d'utilisateurs et contribuent ainsi à améliorer l'information scientifique des professionnels de la santé.

# Concurrent session 5D

## Data bases

Chair

V. Pistotti

# Health for All, but also Health Care Online Information for All? What HECLINET can do

Rüdiger Schneemann

Technische Universität Berlin
Universitätsbibliothek und Institut für Krankenhausbau

## 1. The WHO Targets

The theme of our session is "databases". For a lot of people it has the meaning of a magic word; for them it is like a key for the solution of most problems. On the other hand there are more the sceptical ones; for them it became a synonym with the vision of an unacceptable future.
It is not my intention to decide if there is any right of these poles. The contrasts should only demonstrate the enormous breadth of our field of discussion and how difficult it can be to come to an all convincing estimation of the value of databases.

The general targets of the development of the different aspects of health are formulated by the WHO. They should be compulsary for the policy of all member countries; the motto of our congress "Health Information for All - a Common Goal" is a good example for this effort. So too I try to orientate my paper to the WHO-Strategy "Health for ALL".

Of special interest for us in the European Region is the target 35 "Health information system", which was endorsed by the Regional Commitee in September 1984. Therin is requested, that "Before 1990, member states should have health information systems capable of supporting their national strategies for health for all" (1).

Now the intention of my paper is first to discuss who the users of such system are (a more general view), second to show which place in an information system databases should or could have, and third to demonstrate what an existing database like HECLINET can offer to the health care community.

## 2. Health Information for All?

The typical "user" of health care is every human being, but is he also meant in case of health information? I think there are some differences and we should have in the following a deeper look to this reflection.

### 2.1. Who is the User of Health Information?

One of the standard problems of documentation departments and libraries is the definition of the "typical user".

If we pass through the program of the WHO with the hope to see some answers, we naturally read above all theoretical and very common formulations. That is not surprising, because in such large programs you can't expect too detailed answers. So we find in the comment to the target 35 the first statement relating to our problem : "Steps should be taken to make health information easily accessible to the public". Now we have a tautology: all user means the public.

But fortunately other points in the paper give us more help for interpretation, especially if we go through the part 'suggested solutions': "Such systems should ... meet the requirements for planning, management and evaluation, ... .". Although later on in the same sentence it is stated, that " ... systems ... should be a source of open information easy accessible for ... the public at large" there is in my opinion no doubt, that

the first statement is much more important than the second one. It reveals the decision-makers, the managers as the main target group of the program. So we come to the conclusion that health information can be for all, but mainly it is defined as a tool for professional health people.

## 2.2. Which Kind of Information should be Expected?

A normal citizen for example expects from a health information system answers of problems like: " What is the best therapy for my illness?" or "Is it good for my health to go once a week into the sauna?". A health manager is much more interested in replies to "Can hospitals with less than 50 beds be economical?" or " How many ambulance cars do we need at the weekend in our region?".

So different are these examples, common to them is the wish of the people to have a clear and easy understandable reply. It illustrates the dilemma of information intermediaries: they are collecting millions of information, but they have no answers to some of the most simple questions. It is indeed also my daily problem: I explain to the user that he can have an outprint of say 124 titles dealing with his special problem, but -please- reading and evaluating those documents is his job!

Most people do correlate with the term "information" dates, facts, etc. If they are lucky they will get some, but the very normal situation is that such "hard dates" are not available, at least not immediately.
But it is not only our task to bring isolated facts to the user, we are also obliged to provide him with background information, which means longer textes and statements. The answer of the headline of this chapter is accordingly twofold: the most actual and best possible facts and the most concerning literature should be expected - with a big restriction: only from the just available and searchable system(s), not from all sources of the world.

## 2.3. Who has access?

The principal demand is that everybody should have free access to information. I want to discuss it neither in a legal sense nor a financial one. Of course both are worth debating; but here I want to stress other points.

Most reasons for the demand of information are search or better research work; its quality depends very much on the grade of education which the asking as well the seeking people have. I think the following statement is correct: that to get well(or bad) information depends at the first level on how well(or bad) the user resp. the searcher is educated.

Another point has more to do with "to know where the knowledge is". It is very important to be in contact with information specialists or centres, but also to know and look through specialised directories, and to be aquainted with experts of the field of interest.These points we can call 'flexibility'or 'cleverness'.

The next argument has to do with the place of living: in a big city with many scientific units it is easier to come and stay in contact within a reasonable amount of time and money than in a remote situated village.

In addition to financial and legal aspects there are obviously some relevant barriers in the persons of user and searcher.

Now back to 'databases': have'nt they the chance to bring everybody the information wanted - especially if we consider the disadvantage of a remote living place - or do they create additional restrictions?

## 3. The Importance of Online-Databases

Because of the technical possibilities of today's communication system, data storage, knowledge transfer etc. databases are normally online. First some words about EDP (Electronic Data Processing).

### 3.1. EDP - the Solution of the Information Problem?

Each information, which is stored in a EDP-system, has lost its original identity. On the screen you can no more recognize, if the source was a valuable, rare document from the last century or only the headline of a poor masspaper. The underlying technology, the digitalization of every kind of information is transferring every thing in the collective bit-pattern, the trivial yes-no system. This brings an enormous effect of standardization; if once a document is stored it can be printed in thousands different versions, transformed etc..

A second very important principle of our industrialized society is the work-sharing. Some few people construct the computer, some others run it and write the software, but the mass of the applicators looks to the screen and acts as the others have prescribed. This alienation on one side, the enormous variety of possibilities on the other hand offer the computer a dominating role, 'he' becomes autonomous. It is evident, that many working processes are designed for the optimal use of this machine, and not - as it normally should be - for the man.

So impressive is the performance of a computer, so much it can help in innumerable cases of information problems, one thing is not possible: it can not make the information better as it is. As a pure machine, without intelligence or own judgment, it can only warm-up what men have put in its storage. This is the main reason, that EDP by itself is unable to solve the 'Information problem' as a whole.

### 3.2. Databases: Information as a Market

The rules of the present situation can be formulated as follows: the better databases are, the more amount of capital and manpower is to be invested. And: the more databases and hosts exist the more dominate economical principles. The whole system became in the last decade a marketplace, with all its advantages and disadvantages. Few big vendors (as Dialog) are the leaders, the others have more or less to follow, imitate or interprete them. Also a remarkable event is the strength a monopolist like Chemical Abstracts conducts its market.

But back to health information: the undisputed market leader is again an institution in the USA, the NLM. One proof is the reversion of EMBASE as one of the biggest competitors from its own thesaurus to the MeSH, anotherone that some countries (Sweden, Yugoslavia) start national medical databases as supplement to MEDLINE.

The fact, that databases are belonging to one of the most international markets, is in some respect a big advantage for the user. He only needs to go through few systems to get a satisfying overview of the literature of many countries and languages. But there is no guarantee of completeness. Even "super-databases" like MEDLINE have no intention to put each published literature from every country in the system. If those countries do not want to lose their national literature they have to start bases of their own. This brings us to the next point.

### 3.3. Databases for ALL? The New Barriers

To the poorer and less developed countries the "market law" brings of course disadvantages; but the real big problem is, that the contents of the leading information systems is orientated to the requirements of the rich countries. I think that an information about the organization of an health center in Tanzania is very different from one in the USA, and very seldom stored in a base like Medline.

Other new barriers have to do with the use of EDP: The technical access (for example by phone and modem) costs a lot of money, the database producers itself want to earn money (and more than by the old printed versions), the hosts charge for their services. There are of course many new capabilities to assist the user, which are only possible by the use of EDP: online-ordering; fast switch-over to other databases; Keep-files and so on. But all these offers are not cheap, and even in industrialized countries exist a tendency

to a "two-class information society". An unpleasant side aspect is that more and more printed services reduce or finish their products because of the competition of databases (also if they are from the same producer).

I try to find a résume: "health information for all" is very, very difficult to gain, may be even more than "health for all". But in order to bring at least some parts of the widespread, divergent information to the user, the implementation of EDP was the logical consequence and the only chance to keep up with the incredible mass of literature.

In this situation a lot depends on database producers and information imtermediaries to lower the hurdles. So we come back to the WHO-targets because I think we have to support them; in the following i try to show what HECLINET as one system has done resp. can offer.

## 4. The example of HECLINET

The acronym HECLINET is derived from the term "Health Care Literature Information Network" which was the theme of a 1975 meeting of the "Study Commitee on Documentation and Information handling" of the International Hospital Federation. Today HECLINET is a synonym for an international cooperation in tackling with the information problem in the field of hospital care, and for a database offered by the host DIMDI. Some figures: HECLINET has about 80,000 entries since 1969 up to present in all languages (50 % German titles, 30 % English).

### 4.1. The "Niche"-Function

The term "Hospital Care" refers to the fact that it is not just the object "hospital" which is the issue, but rather beyond construction and administration of a hospital it is the legal and economic frame work, the entire health care system, the staff, the health economics. In addition, all facilities of health care are included as well, i.e. outpatient depatments, specialized clinics, university medical centres, nursing centres etc..

There is no exact definition of the interdisciplinary area of "hospital care", but the best understandable description is "the nonclinical aspects of ...". So we come to our niche: HECLINET on the one hand documents that literature which is not or not sufficient enough in Medline, Embase, ISI or other special databases. The other aim is to index literature published anywhere in the world (in fact mostly in Europe) in order to fill up regional gaps of the large US-oriented databases.

Another ambition of HECLINET is to give support to interdisciplinary questions, which can not be satisfied by an only subject oriented pool. In addition to this every kind of publishing is included, which means scientific books, journal articles as well as trivial notes from newspapers. Our target groups are among scientists, doctors etc. mainly practical working professions like nurses, administrators, but also patients, in one word - to pick up the WHO target - the public.

### 4.2. The Cooperative Way

A further difference to most other databases is the multinational disposition of HECLINET. Following institutes are partners:

Dansk Sygehus Institut (DSI), Kopenhagen
Österreichisches Bundesinstitut für Gesundheitswesen (ÖBIG), Wien
Schweizerisches Institut für Gesundheits - und Krankenhauswesen (SKI), Aarau
Sjukvaardens och socialvaardens planerings- och rationaliseringsinstitut (Spri), Stockholm
Deutsches Krankenhausinstitut (DKI), Düsseldorf
Institut für Krankenhausbau der TU Berlin (IFK), Berlin

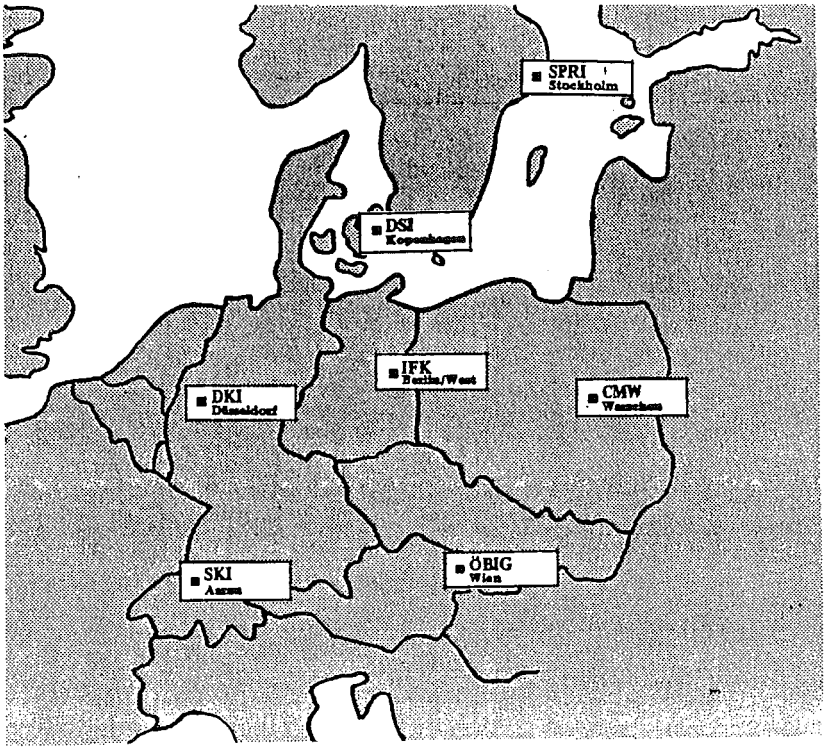


Figure 1: Partners of HECLINET

Very close contacts exist to

Nationaal Ziekenuisinstituut (NZI), Utrecht
Kings Fund Center (KFC), London

The cooperation is based on the following consensus:

- Each institute analyzes its national literature according to jointly agreed upon rules and makes it available to the partners.
- The keywords are based on the "Thesaurus Krankenhauswesen" (2). Titles or abstracts may be written in German or English, other languages have to be translated.
- Data processing up to printing of the "Informationsdienst Krankenhauswesen/Health Care Information Service" (3) as well as the generation of magnetic tapes is handled centrally in Berlin by the IFK.

The basic rules of HECLINET allow on the one side interested partners to participate, on the other side each institute has the chance to index the literature most interesting for his own country. This again is a positive aspect to come closer to what different user at different countries want to know from one system.

### 4.3. User Orientation as Philosophy

Some remarkable points of user-friendliness are mentioned. A new aspect has to do with language-barriers. The thesaurus of HECLINET is translated in English, so even the German documents are searchable for English speaking people. Connected with our host DIMDI in Cologne, the user can switch over from any (English based) database to HECLINET, and supported by the software of DIMDI he can automatically start a search with the terms of the previous base.

Great attention is laid by HECLINET on the problem of making available the original literature to the user, especially because there is a lot of grey literature indexed. Documents with less than 30 pages are generally microfilmed. Each partner has a complete microfiche library with the result, that more than 90 % of the documents are nearly at once available. The rest of the documents (books etc.) are at least provided with notices to clarify the place of location; the user can then get the document by lending.

At last it is to add that the use of EDP brings of course important improvements. The most impressive one is the outprint: the laser technology provides us with clear and good-readable letters, the searcher can define by himself the wanted layout, and the speedness of the printers brings the information very fast to the reader. For HECLINET these aspects were the most convincing ones to change from an university computer to a professional host like DIMDI.

## 5. The Pilot-Role of HECLINET

To come to a conclusion I hope my paper could show that there are chances to come to "health information for all", despite all those problems discussed in the first chapters.

I think the time up to 1990 is to small to meet the ambitions of the WHO regional commitee concerning target 35. But it is to recognize that there are a lot of activities going in the right direction and may be a system like HECLINET can set a good example how to succeed with organizing a sufficient health information system. The more realistic motto can be "Health information For Many"

REFERENCES

(1) WHO, Regional Office for Europe
Regional Targets in Support of the Regional Strategy for Health For All.
Document EUR/RC34/7 Rev.1, Copenhagen, 21 December 1984

(2) Thesaurus Krankenhauswesen.
Berlin, Düsseldorf, 5.Ausg. 1982

(3) Informationsdienst Krankenhauswesen/Health Care Information Service.
Berlin, Düsseldorf, 1(1969) ff, ISSN 0341-0595

# A NEW DATABASE AND A NEW SOFTWARE FOR MEDICAL LITERATURE

Robert Kimberley
Institute for Scientific Information

This paper discussed the development at the Institute for Scientific Information of a new database and its associated access software.

Current Contents Search is the online version of the Current Contents series, providing cover-to-cover indexing of over 4,500 medical and scientific journals. The content and special features of the file are described, as well as its use in different environments, both as a "stand-alone" online service and in conjunction with the printed version. The range of search keys is discussed, together with the various output methods. The paper considers the use of Current Contents Search - by individual researchers or practitioners, and in the setting of a medical information department - for browsing tables of contents, for verification, for SDI profiles, for the creation of current awareness bulletins, and for online document ordering.

The paper also describes Current Contents Connection, the state-of-the-art user interface which ISI has developed. This is a low cost, high performance package offering the user of medical information a variety of functions. The paper will discuss the user-programmable script language which enables automated logon and logoff via any international packet-switching network, and the package's capabilities for assisted searching, display and downloading. The system's speed, power and ease of use is examined in relation to such features as the pull-down menus, pop-up windows and its "point-and-click" action for indicating individual records.

The paper concludes by considering this software in the broader context of information retrieval.