# INFORMATICS TECHNIQUES AND STRATEGIES FOR INFORMATION TRANSFER

John Page

January 1980
PP-81-1



INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

This paper has been designed as a working document for the discussion of available technologies, future potential, and constraints in applying technology to systems interconnection for the Second Meeting of the UNISIST Working Group on Technology of Systems Interconnection.

CONTENTS

# CHAPTER I

# INTRODUCTION AND BACKGROUND

## OBJECTIVES

The primary objective of this paper is to review developments in informatics, both present and future, which are relevant to information transfer and the interconnection of information systems. Because developments in informatics are proceeding at a very rapid pace, while real user needs are difficult to establish in quantitative detail in advance of an informatics application in the form of actual service, it is necessary to consider flexible strategies in which the interaction between new technical possibilities and emerging user needs may be maximized. The need for such an approach is underscored by the fact that there are a large number of influences which together contribute to innovation in informatics, only some of which are directly related to the requirements of information transfer.

It should perhaps be stressed that in the context of this paper by "information transfer" is meant the transfer of scientific and technical information, including economics and the soft sciences generally: the transfer of information in the sense of news, cultural exchanges, commercial transactions, and entertainment is excluded, although some of the techniques with which we shall be concerned have been developed largely as a result of requirements in these fields.

## HISTORICAL BACKGROUND

The application of informatics technology has had a major impact on information transfer and the process seems likely to continue in the foreseeable future. The impact, however, is not simply a matter of a straightforward application of a technological innovation to solve known problems at an acceptable cost,

but appears to involve a complex of factors including cross-fertilization between two previously independent areas of development (computers and telecommunications), the explosive growth of general electronic data processing in the industrialized countries, market forces, and user reaction. An historical prospective may therefore help to establish the dynamics of these continuing interactions between developments in informatics and the concomitant rapidly occurring changes in information transfer possibilities and methods.

## The Marriage of Computers and Telecommunications

Prior to the mid-1960's, changes in the system of information transfer in the shape of primary publications and later in the use of secondary publications to help locate relevant items in the stock of primary literature were relatively gradual: the major changes were in the numbers of items dealt with by the system (the literature explosion), rather than in the methods by which information was made available to the potential user. During the early part of the decade 1960 to 1970, there occurred three developments in technology which, when joined by a fourth at the end of the decade, provided the means for the beginning of a revolutionary change in information transfer systems. The first two of these were the development of relatively cheap (and rapidly becoming cheaper) random access mass memories for computers, and software and hardware systems for using computers in a timesharing mode. These two developments together made possible a much more cost effective use of computing in many applications, including information transfer: the development of timesharing systems made it possible to take economic advantage of the processing power of third generation computers, and the new possibilities offered by the high volume, high speed random access memories led directly to a third development, one which was specific to a major problem of information transfer. This was the development of information retrieval software applicable to a system of inverted and serial files, well suited to the new hardware architectures and offering a highly cost effective alternative to the serial and batch processing retrieval systems which were then current. The new hardware/software combination made it possible to interactively search fairly large databases to provide quickly and cheaply printouts, selected from perhaps half a million items on the database, of the items relevant to a query which had been typed in a more or less natural language on a terminal.

The fourth development did not occur until the latter years of the decade. This was the development of communications networks using telephone circuits in which digital data were substituted for analogue voice signals. The development of TYMNET, which began operations in 1969-1970, was an initiative by the Tymshare Corporation to make available timesharing services for a nationwide clientele in the U.S.A.; the combination of the favorable economics of sharing the resources of major computers and of sharing the costs of the data circuits necessary to access them by a large and dispersed community of users resulted in a

major new market dominated not by the satisfaction of STI user demands, but by the increasing need for cheap and effective computational facilities by commercial and business users. At the same time a major operational experiment in computer science and technology, ARPANET, was being mounted by the U.S. Government to develop a new networking system (packet switching) to optimize the use of expensive communications circuits in a computer network. Both TYMNET and ARPANET employed a marriage of minicomputer and telecommunications technology to control and optimize traffic flow, and this constituted an important breakthrough in making computer networks cheap and effective for data transfer.

## The Growth of STI Transfer Using Computer Networks

The success of these four new developments in the STI field alone may be illustrated by the fact that, starting with the single example of the MEDLARS database, which was made available through TYMNET by the National Library of Medicine in the early 1970's, there are now several hundred STI databases available over TYMNET and TELENET, the commercial network employing ARPANET-developed technology, and probably over 2,000,000 STI searches are carried out annually within the continental U.S.A. using these networks in association with informatics techniques for storage and interactive interrogation of information which are direct descendants of the mid-sixties systems already mentioned.

A new, comparatively recent development is the introduction of fact and numerical databanks; these are of growing importance to users outside the research and development field, since they supply facts and/or data in one search operation, instead of references to journal articles and reports. Databanks of this type range from statistics on international trade and financial questions to performance data on electronic components and thermodynamic properties of compounds. There are already nearly 100 online databanks available over North American networks.

Although both informatics and communications research and development were no less advanced in Europe than in the U.S.A. in 1970, it is noteworthy that the actual realization of network STI operations has been considerably slower in Europe as compared with the level of activity reached by the main U.S. vendors of online services. This is in spite of the fact that the market (on a continental basis) does not appear to differ markedly between the two areas, although the American market is probably about 25% larger, and is homogeneous in terms of language. Probably, however, the most significant difference accounting for the slower paced development is that inherent in the telecommunications structures: the U.S.A. is virtually a single telecommunications market of over 200 million people, while Europe is divided into individual national markets: this has two consequences, first, that comparable economies of scale on an international level are difficult to achieve, and second, that it is much more difficult for regulatory changes which may be necessary in promoting the growth of cheap computer network facilities to be agreed internationally. For example, a

landmark in the growth of networking facilities in the U.S.A. was the concept of Value Added Carriers which was agreed by the Federal Communications Commission in the early 1970's: this allowed the new networks such as TYMNET and TELENET to rent circuits from A T & T, the monopoly telephone carrier, and to use these facilities to provide data services to their clients. This would have been totally impossible on a continental scale in Europe because it would have required that every telecommunications administration was able to agree to give up some of its monopoly rights to a commercial European equivalent of the U.S. TELENET Corporation. In fact, it was not until the mid-1970's that the difficulties were resolved by an agreement between the Commission of the European Communities and the Conference Europeanne des Administrations des Postes et Telecommunications (CEPT) on a means of building up a European packet switched network for STI transfer, funded by the Commission but managed by a PTT consortium.

## Informatics Developments and Information Transfer

The pattern of informatics and teleinformatics developments sketched above would suggest that information dissemination and transfer is not a main driving force, although when it can be associated with other more commercially oriented potential applications, information transfer requirements can significantly influence development. Examples are the development of online information retrieval software following the introduction of time-sharing, and the rapid growth of online STI retrieval services following a commercial decision to invest money in providing a network time-sharing service for business-oriented companies. Nevertheless, an aggressive STI development policy may, as in the case of the EEC's promotion of the EURONET concept, force development to meet what are seen to be solely STI requirements. In this case, the European PTTs were, however, quick to understand the importance of EURONET as an initial stage in the development of a general purpose international data processing facility in Europe: exceptionally, STI was the "front runner" in this instance. Finally, we may note that while the entrepreneurial association of an emerging new technology and its potential market constitutes the main driving mechanism for informatics development, these must take place within the framework of institutions and policies whose rate of adaptation to the new technologies is much slower than the pace of technical development and application. Both driving and braking forces are at work, and it will be necessary to try to evaluate these in the context of current and future informatics development later in this paper.

## NON-INFORMATICS CONSIDERATIONS IN INFORMATION TRANSFER

While the scope of this paper is clear from its title, it is obvious that there are other considerations which must affect the choice of strategies for application of technology to information transfer. Some, such as the (usually) paramount role of

other electronic data processing applications in determining the direction of technological development, and the dominance of the market for other applications of remote computing services in determining the shape and extent of penetration of the physical facilities themselves have already been noted. Brief mention must, however, be made of two further factors, which, it is understood, will be fully discussed in other working documents. These are user needs in information transfer and telecommunications technology for information transfer. Without anticipating the discussions and conclusions relating to these topics in other documents, it may be useful here to briefly note the interface between these topics and that of informatics with which this paper is chiefly concerned, and to state some broad assumptions on them which underlie the discussion in subsequent sections.

## User Needs

STI users are primarily research and development scientists and engineers, but the existence of the many newer databases in the fields of economics, statistics, and social science, together with the inclusion in service operators' offerings of fact and numerical databanks, as distinct from bibliographic databases, has widened the potential audience to planners, analysts, and policy advisers generally. The needs of research scientists and engineers in developing countries for information probably do not differ significantly from those of their colleagues in industrialized countries: they need access to bibliographic databases in their particular field to identify relevant literature, and thence, easy access to copies of the original documents. Users outside the R & D field will need specific information drawn from outside the formal discipline-based or mission-oriented material covered by STI secondary publications and their corresponding bibliographic databases. In developing countries these needs beyond formal STI may be satisfied from imported know-how information and data (e.g., from technical cooperation or assistance programs), locally produced information of specific relevance to the country, usually in the form of reports not part of the formal resource, and indigenous information, or knowledge gained and transmitted informally. Woodward [1] suggests that locally produced information is perhaps of the most importance for developing countries; however, it seems to be generally held that in many regions, use of the formal STI resource through the medium of online interrogation of databases in Europe and the U.S.A. is becoming a highly important method by which these countries satisfy their information needs. In any case, all users of bibliographic databases, whether in industrialized or developing countries, require efficient and speedy access to primary documents. The importance of locally produced information would also suggest that developing countries have a special need to develop methods for identifying, collecting, and handling information of this type. Thus, in subsequent sections, two types of interface between informatics applications and information needs in the developing countries will be discussed: an

interface with existing services and systems in industrialized countries, and informatics techniques for the ordering and retrieval of locally produced and imported information.

## Telecommunications Techniques for Information Transfer

It is assumed that existing networking technology, typified by TYMNET and TELENET in the United States, and EURONET and national data networks in Europe, will be applicable to developing regions. The pace at which existing networks will be extended to provide services in developing countries, will, however, be conditioned by the potential growth in traffic foreseen as a result of a particular extension, expected revenue, and costs. It seems reasonable, however, to assume that such commercial decisions will not be mainly determined by potential STI traffic, which will probably be a relatively minor factor. Another possible scenario which might be considered is that local telecommunications authorities in developing regions obtain funds to subsidize a network extension, controlling the price local users pay for the service: there are precedents for arrangements of this kind between commercial network operators and telecommunications administrations. A different type of interface between informatics technology and telecommunications may also be considered: this involves the use of small earth stations, not necessarily connected to the ground based telecommunications networks; such communications satellites are already state of the art. (See for example the plans and proposals for multipurpose satellite networks made by the Satellite Business Systems Corporation and the Xerox Corporation in the U.S.A. described briefly in Reference 2, and the work of COSADOC, a working group of the European Space Agency, in specifying experiments in information transfer which could be carried out using that agency's Orbital Test Satellite or OTS [3]. The conditions under which small ground station systems might be used for information transfer, and the interfacing of this technology with the informatics elements will be mentioned in subsequent sections of this paper.

## STRUCTURE OF THE REPORT

In the next three chapters we examine the state of the art in each of the main functional areas in the transfer of scientific and technical information, followed by a brief analysis of the problems in each area, including those which particularly affect information transfer to developing countries. Possible solutions to these problems in terms of application of informatics technology are then discussed, including other elements of an infrastructural and legal or regulatory nature. In the final chapter, strategies assisting the solution of these problems are then derived, with recommendations for action at the national, regional, or international level.

This approach means that some interesting new informatics technologies are not considered in any detail, because so far as

can be foreseen at the moment, they do not contribute to the solution of an actual STI problem. For example the use of Videotex systems has already made a considerable impact on the dissemination of certain types of information of interest to the general public in industrialized countries; the use of such systems has even been suggested as a means of technology transfer, for example, in agriculture.

# CHAPTER II

# INFORMATION CAPTURE AND ORDERING

## THE PRESENT SYSTEM

### Information Capture

All of the formal STI literature, and a large proportion of the gray literature (reports, etc.), is generated in the traditional way; i.e., the author prepares a manuscript, it is typed, submitted for refereeing or approval, edited, typeset, (or an equivalent process to produce a reproduceable master), published, and distributed. In the case of the formal serial literature, articles are associated together and published as an issue of a journal. In the case of report literature, individual reports may be one of a numbered series (findable through periodic lists of titles) or not. While there have been important improvements in methods of preparing master texts and in printing technology, the system in general has not altered greatly in the last few decades when seen from either the author's or the user's point of view. New developments have mainly served to keep pace with the growing volume of publication and to keep cost increases within reasonable bounds.

### Ordering of Information

By "ordering" is meant the process by which the stock of information is systematically arranged so that individual items can either be identified or located for retrieval. The formal literature (and reports in certain circumstances) is scanned, abstracted, catalogued, and indexed, the result being a secondary publication most often prepared from computer readable records, which can then be readily reprocessed to provide a bibliographic

database arranged for batch processing or online retrieval. Essentially, the system still retains many of the features of traditional library cataloguing and indexing but with new informatics technology added to input to the computer: examples are optical character recognition (OCR) and direct online input using a video display terminal for editing an intermediate record, followed by merging with the master database. Various possibilities in economizing the intellectual labor associated with the ordering process as a whole have been tried; tables of contents of journal issues, rather than the articles themselves, may be processed to record only titles and authors' names: free indexing based on titles or abstracts may be employed in place of controlled terms in a thesaurus.

## GENERAL PROBLEMS IN INFORMATION CAPTURING AND ORDERING

In this section, an attempt will be made to outline some general problems in this area of information transfer and to suggest how informatics technology might help to solve them. The problems dealt with are essentially general in character; that is, they are faced by all users from industrialized and developing countries alike.

### Delay in Information Capture and the Ordering Process

In has been noted that the processes by which STI information is recorded, captured, and ordered are still based on traditional methods of primary and secondary publication, and these carry a heavy burden of intellectual work. These factors make for relatively long delays between the completion of a piece of research and development and its availability first as a primary document and second in retrievable form in a secondary publication or a bibliographic database. Publication delays of six months to a year are typical, and most secondary services find a mean delay of six months between publication and appearance on a database the best they can achieve.

The degree to which such delays can be tolerated will vary greatly between different groups of users. In some areas of pure science, especially those of a somewhat esoteric nature with a rather small population of researchers, there is probably little need for really rapid exchange of results through the formal scientific primary and secondary literature; this is relatively easily achieved by personal contact. However, in a rapidly advancing area of applied science or technology, in which the results of research and development could be of immediate economic or commercial consequence, there will be a greater need for fast exchange of results, and in this case, probably the present processes are too slow and cumbersome. Solutions to the delay problem are not simple, and we know little in quantitative terms about the real need for speeding up the process. One group of solutions involves changes of emphasis in the structure of the present system for capturing and ordering of information. For example, since the exchange of research and development results does not necessarily involve the prior production of a journal

article or research report, merely knowing that a particular research project is in progress, its scope, and the stage reached, is probably sufficient to promote the necessary contacts between different groups of researchers. Emphasis should therefore be placed on collecting and organizing material for databases on current research projects as a natural supplement to bibliographic databases which provide a historical picture.

Major increases in speed of publication cannot, however, be expected within the formal publication process; such devices as on-demand publishing, tailor-made journals, and synopsis journals may effect some improvement in publication delays, as well as provide other convenience features for users, but informatics technology has no ready-made solutions to offer in the short term. In the longer term, electronic publishing and the paperless exchange of scientific information (see Chapter IV) could drastically change the present situation. It is worth repeating, however, that up to the present, we do not know enough about the need for speed in the recording and capturing of STI.

Undoubtedly, improvements are possible in delay in the production of data bases for information retrieval. At one extreme, short cuts such as input limited to journal contents pages have already been noted and such products already command a considerable market. At another, the process of collection of items of primary literature from dispersed sources, their centralized processing into machine-readable records, and the final creation and dispatch of copies of the database updates could well be streamlined. This might involve the use of electronic media for transmitting bibliographic information including abstracts, editing it at a central location with text processing equipment, and broadcasting it, again using electronic methods, to host machines carrying the database. In this way, updates could be made daily, using terrestrial high speed data communications, or (preferably) satellite systems (see Chapter IV) with their combined broadband and broadcast properties. Here again, however, we are uncertain, in the present state of knowledge, about the degree to which users need to have greatly improved timeliness in the appearance of items on bibliographic databases and the costs of making such radical changes in the database production and distribution system. As with other radical changes involving text transfer, the practicabiltiy of the system suggested depends on cheap telecommunications bandwidth, such as could theoretically occur with the new generation of communication satellites.

## The Growth of Bibliographic Databases and Databanks

Prior to 1970, the number of databases available for public access in Europe and the U.S.A. was very small. By 1975, 386 databases and databanks had been identified, while in 1979 EUSIDIC [4] had identified a total of 1,280, of which 565 were bibliographic databases and 715 were fact or numerical databanks. The fastest growth had occurred in the databank category (from 51 in 1975 to 715 in 1979). There are available for online

access through the major public networks such as TYMNET, TELENET, and EURONET (including ESANET) a current total of about 400 databases and databanks of all kinds. The major U.S. online vendor, Lockheed Information Systems, adopts a supermarket approach to information, offering access to some 100 databases, ranging in subject field from the natural sciences to economics, from business and commercial data to art history.

A highly important development in recent years has been the online availability of statistical data from various fields, including international trade, energy production, currency and share values, industrial statistics, and related fields of particular concern to planners, managers, policy advisors, and other classes of potential users entirely outside the R & D field. The very large numbers of databases and databanks now available online, and the wide but probably highly overlapping subject coverage, makes it extremely difficult for a potential user to select the database or combination of databases most likely to answer his query. Even in the more familiar fields of pure and applied science, there may be ten or twenty databases having some potential coverage of a particular topic and it could be an an expensive and time-consuming task to search all of them in order to ensure that valuable references were not missed. The large discipline-oriented databases such as CHEMICAL ABSTRACTS and BIOSIS contain a wealth of material important to applied disciplines such as agriculture and environmental science and technology. While in a majority of cases, these references will have been picked up also by the specialized databases, for example the branch of agriculture concerned, this cannot always be taken for granted. The situation becomes even more difficult once one moves into the more economic- or commerce-oriented fields, where the vital piece of information may be in any one of a number of special databanks and databases or even in a major newspaper information bank such as the New York Times System.

Solutions to these types of where-to-look problems have been sought in a number of directions. The major database vendors in the United States are beginning to produce a species of online database dictionary to their offerings; in these, key words are listed alphabetically with an indication of the frequency of their use in the databases offered. Other solutions applicable perhaps to information centers in major government or commercial enterprises undertaking a large volume of online searching in particular fields lie in specialist training of intermediaries to improve and maintain their expert knowledge of information sources in their own specialization. From an organizational point of view, the institution of highly specialized information centers on the pattern of those in some of the U.K. Industrial Research Associations or the Fachinformationzentren in the Federal Republic of Germany may be appropriate: such solutions, involving much highly skilled manpower and a sophisticated infrastructure, may not be very helpful in some developing countries. Another possibility is the improvement of referral systems, including online dictionaries of database coverage, comprising not only indexing term statistics, but also primary sources of

information scanned for each database, and an index of primary publications by subject coverage. An ideal tool of this kind would be continuously updated, be highly user-friendly, and facilitate entry into the information by a number of routes.

## SPECIFIC DEVELOPING COUNTRY PROBLEMS

### Potential for Information Management

It seems reasonable to assume that, for the immediate future, developing countries will not be major producers of primary material for the formal stock of STI. They will thus be faced with the inadequacies of the present system of information capturing and ordering primarily as users. In data bases such as INIS and AGRIS, produced by international cooperative programs, all participants agree to provide input to certain uniform standards, and in this respect, developing countries may find it difficult to organize collection and processing of the material material, particularly if they are faced with a shortage of skilled manpower. It could, however, be argued that providing input to an international information system helps a developing country to develop an infrastructure and local skills in information management which it badly needs, particularly if the sponsoring organization can use technical assistance programs for pump-priming purposes in the countries concerned. Should there be sufficient volume of information to justify it, more advanced informatics technologies might be introduced as a second stage, for example, key to tape machines, text editors, etc., once again as a means of developing new skills and facilities which can be used for other purposes.

### Processing of Locally Produced Information

This type of information, which Woodward [1] stresses as a highly important information resource in developing countries, is essentially non-conventional in character; it may consist of resource inventories, consultancy and project reports, studies, annual reports, and administrative documents, some of which may be of limited circulation. At all events, there would appear to be a major problem in tracing, collecting, and organizing this material so that it is available locally where it is needed, and when appropriate, can be exchanged with other countries, either bilaterally or through the medium of regional centers or international programs. Part of the problem is therefore organizational, requiring a national focus (or foci, depending on local structures and responsibilities in relation to the subject and quantity of the material). The linguistic problems discussed in Chapter 3 as they affect information retrieval, may also be important in local database creation and exchange systems based on such material. To construct a database (or to provide input to an international or regional cooperative system) may possibly imply a requirement to translate material into a common language.

The level of sophistication of the informatics technology to be applied in the creation of a local database will depend on several factors: volume, availability of resources (both human and material), and decisions on how the material is to be announced and made available. Tell [5] points out that where responsibilities for information management are reasonably centralized in an executive ministry, and this ministry possesses a satisfactory computer, locally produced and other information may be aggregated into what amounts to powerful local databases, tailored to the needs of the decisionmaker concerned. Such solutions, however, presuppose a certain level of infrastructure and technical expertise which may not be available in a very large number of cases.

# CHAPTER III

# INFORMATION RETRIEVAL - INFORMATICS AND THE USER

## THE PRESENT SYSTEM

The historical background to the very rapid development of information retrieval services using extensive computer networks to provide terminal access to databases resident on remote computers has already been discussed in Chapter I, and it is not proposed to enlarge on this here. Online retrieval is no longer a cottage industry in industrialized countries: Cuadra [6] has identified over 60 organizations offering online retrieval services internationally, and some authorities believe that the current U.S. market is about $40,000,000 per year in size, with Japan and Europe each at about the $10,000,000 per annum level. While a number of online searches are being made by countries outside these regions, these are "imported", mainly from the U.S.A.: there appears to be little activity as yet in supplying online services from other regions.

In terms of equipment, i.e., terminals, recent trends are essentially evolutionary, in that the equipment is now cheaper, more efficient, and easier to service than hitherto, rather than embodying revolutionary new features. Accoustic couplers are now accepted by most telecommunications administrations, and this makes it possible to use portable terminals, such as the Texas Silent 700, where the local telephone network is reasonably good and free from noise. Intelligent terminals working in block mode are also becoming much cheaper, but their use in information retrieval is not yet widespread, perhaps because interfacing them with computer networks (and the public telephone system) can be difficult or expensive, and in any case, they are probably much more useful for online input or file creation, rather than simply for retrieval purposes. With some possible exceptions, the specific needs of online information retrieval have not been

a particularly dominant factor in terminal development: instead, the STI user has been able to take advantage of the results of the growth of the market for general purpose terminal equipment.

The trend in software development and retrieval system architecture has also been essentially evolutionary; the initial systems, such as early versions of ORBIT and DIALOG, have been considerably improved and extended, permitting such features as free text searching with left or right truncation, storing of search strategies for automatic running against file updates to provide a remotely controlled SDI service, automatic ordering of document copies, etc., etc. The number of retrieval systems in common use via existing networks is now well into double figures, and some are expressly designed to retrieve numerical data from databanks: others have computational modules to permit manipulation of numerical data. Still other systems, such as the QUEST software of the European Space Agency's IRS, are designed to search both bibliographical records and numerical data, having a range-of-values search possibility. Most of the current software systems include commands which provide dictionary displays, numbers of references indexed (or occurrences of words or combinations of words), Boolian logical combinations between selected groups, displays of individual items, and local or remote printing of any item. Searching by author or corporate source is also usually possible. A recent development in Europe has been the creation of a common command set to avoid the problems encountered by users in having to remember the different commands and features of many different systems if they wished to get the best out of a multi-host network such as EURONET-DIANE. When implemented on a host computer, the common command set permits a single set of commands to be used, these commands being automatically translated by the host into its own retrieval system, which is thus transparent to the user.

For bibliographic searching, all systems use an inverted file systems architecture, in which numbers identifying a particular reference corresponding to a particular search term (e.g., a keyword or an author's name) are stored in central memory for rapid manipulation, while the text of the references themselves, including abstracts where provided, are stored on serial files in peripheral mass memories. These serial files are only searched when a display is called for. The technique is efficient in the use of computer resources during searching, but involves considerable extra file maintenance in the shape of file inversion programs in updating them. A new development, applied on an experimental basis to automatic telephone directory look-up, is that of ICL's Content Addressable File Store (CAFS), in which a combination of parallel processing and associative matching, and selection of data, enables very rapid multiple searching of a single serially-arranged file. Such a system would have obvious advantages for STI information retrieval and it is understood that this possibility is under investigation.

## PROBLEM AREAS IN INFORMATION RETRIEVAL

A main difficulty which affects users in developing countries to an even greater extent that users in industrialized countries

is the fact that a search of an online bibliographic database will provide information about information, i.e., addresses, and not the information itself. A second process to obtain the actual information sought is therefore required, and it is here that severe difficulties may be encountered, particularly if there is no convenient depository library or document supply center. This problem will be discussed in Chapter IV. Retrieval of information from a fact or numerical databank is a different matter, since the operation provides the user with the information he wants in one step, if the subject matter of his query is such as could be included in this type of databank. Although some of the material sought for in the context of an R & D user's enquiry of a database may also be of this type, such users will also need more general information. He will want to scan the original document as a necessary part of his search: reading the abstract online will not give him what he needs. Nevertheless, he may often not want the entire text of an article: possibly a graph, a table, or an equation will constitute the desired end result of the search. However, the user cannot know this in advance, and therefore it has been suggested that, ideally, an R & D type literature search with a browsing and text selection possibility should be carried out on the full texts of documents. An experimental system of this kind has recently been announced by Bibliographic Retrieval Services in the U.S.A.; Such a database would be expensive to store, and therefore to interrogate, but possibly, since supply of the full text may not be required, the user may find the single-search operation more cost effective than the normal bibliographic search followed by document ordering and supply.

In the normal case of information retrieval from bibliographic databases, the main problem areas may be grouped into two classes: the imperfections of the user/system interface, and the economic and operational problems associated with trade-offs between use of intercontinental computer networks and the use of local facilities. These are discussed in the next sections.

## Imperfections of the User/System Interface

The interrogation of an online bibliographic database, however well managed, is from the end user's point of view at best a substitute for interaction with an unknown number of authors of relevant papers. While he will be able to interact with the material resulting from his first formulation, and will thus be able to adjust and refine his question, there is bound to be a degree of imperfection in the process as a whole. At the one end of the chain, the database as the user sees it in the retrieval operation is largely the creation of the production team (document scanners, analysts, abstracters, and indexers): relevant items may be missed, the level of understanding of their content may not have been complete enough to prevent some loss of information, and in any case, the user will not necessarily have a total appreciation of the indexing philosophy employed. Finally, the user will have to interrogate the database using a retrieval language which is not quite as flexible as natural language, and which will have to be learnt, if only

to avoid simple pitfalls such as wrong spacing, inappropriate spelling, misuse of punctuation, or special symbols interpreted by the computer as commands, etc., and which, unless thoroughly understood, will result in a series of frustrating "invalid argument" replies from the machine. These types of difficulties are usually the reason for employing an intermediary or information analyst to act for the end user on the terminal, the intermediary having a professional knowledge of the structure of the relevant databases and information retrieval systems. However, efficient intermediaries require expensive training, and in the context in which some information centers in developing countries have to operate, this training may be difficult to organize on the spot; in addition, the skills required of a potential intermediary are likely to be in short supply. Even if good intermediary services can be organized, there is a further potential loss of information because, however knowledgeable the intermediary is, he cannot react precisely as the end user would in modifying his query following the first responses of the system. Thus, where possible, end users and intermediaries should sit together at the terminal for the best results. Even in industrialized countries with good communications and highly developed information infrastructures, this is often very difficult. In developing countries, where the population of users may be widely dispersed, served only by few and distant information centers, joint work on the terminal may be virtually impossible.

Some years ago, it was often held that the role of the intermediary would tend to disappear: informatics development would lead to more intelligent and adaptive database and retrieval system structures, including dynamic prompting for optimizing search strategies allowing the end user to carry out his own search as a matter of routine. This has not happened to an appreciable extent for a number of reasons. The bridge between computer science research and informatics application has so far not been crossed, and costs may be prohibitive. Meanwhile, new databases have been produced for specialist fields by repackaging existing material and adding to it from non-STI sources, thus (in theory) making it simpler to make searches in the new fields of applied science and multidisciplinary activities. On balance, therefore, it seems that developing countries wishing to exploit online information sources in industrialized countries as a first step in establishing their own information systems must reckon with the need to provide intermediary services on an expanding scale.

There are three areas in which a directed and organized informatics development could be of great help in improving the user/system interface, short of the fundamental breakthroughs in intelligent databases and the like. These are, respectively, improved user friendliness, multilingual systems, and software for mini- or microcomputer retrieval systems. These will be discussed briefly in the following paragraphs.

## Improvements in User Friendliness

While this is an essential part of the user/system interface, the need for attention from informatics specialists here

is particularly necessary if the developing countries are to make the best use of online systems. User friendliness has, in fact, considerably improved since the early days of online, and this improvement has been affected by relatively simple steps, such as the production of user manuals designed for those who are not computer specialists, the use of online tutorial systems which can be called up by typing a simple command, such as "help" or "?" followed by the topic on which an explanation is required. While the position now seems relatively satisfactory, this may be because intermediaries in industrialized countries are now in general sufficiently experienced to require reminders, rather than initial explanation, and in any case, user seminars and the like are frequent and easy to attend. None of these conditions is likely to be obtained in the early exposure of information centers in developing countries to online, and it may be necessary to study specific problems encountered in order to find appropriate remedies. It should not be forgotten that the user sees the entire system of connection to the network--the network itself, the host computer, and the database--as one, and therefore error messages or failures to obtain correct responses should not only indicate that something is wrong, but where it is wrong and what, if anything, the user might do about it. A European user interrogating a database on a European host finds it quite easy to call up the host operator, explain his problem, and get advice on what to do, but this would be difficult if he were located on another continent. A means is required for obtaining help and advice when necessary over the network itself: the help-type commands in present systems go to the retrieval software level, not to a level at which messages can be received and acted upon at the host operator or network level. This, plus the capability for monitoring a user at his request, would be a relatively simple development.

## Multilingual Systems

The vast majority of online databases and retrieval systems currently available are in the English language: the use of English as the *lingua franca* for exchange of scientific and technical information is probably totally acceptable in the pure sciences, and in a number of applied disciplines, but is certainly a constraint which may hinder understanding in particular regions. Information in such areas as planning, law, social questions, business, and commerce will be handled by users with less knowledge of English than their scientific colleagues. Dubois [7] points out that an an international database on football, for example, might well justify on economic grounds alone "output in a dozen languages as diverse as Portuguese and Polish". Conversely, users who only speak English will require translations of important material in other languages, such as French, German, and Russian. Language ability in one of the main world languages will depend, in many developing countries, on their past colonial history: users in Latin America or in French-speaking Africa may not normally have English as a working language. The need for a multilingual approach is not necessarily, however, a problem mainly of developing countries: the Commission of the European Communities has organized a large multilingual program in order to provide better access to scientific and technical information within the nine member states.

Multilingual tools range from machine translation of whole texts to multilingual thesauri and switching language enabling indexing in several languages. The state of the art in each is reviewed by Dubois in Reference 7; he concludes that fully automatic systems are not yet sufficiently developed to be regarded as operational tools, while systems such as TITUS, which require specially written abstracts, but can then be searched and the results output in several languages, are practical but expensive. He favors use of switching language systems as a cost effective means of providing fully automatic translation of indexing terms effused in conjunction with a minicomputer, without degradation of the parallelism of meanings as additional languages are added.

It may be observed that we are rather far from economic solutions to language problems in information retrieval, since there is no practical way of converting whole databases into another language, although it may be economic to provide handmade tranlations into English of material in lesser known languages, since the market for translations into English is probably large enough to bear the cost. It may be that the TITUS principle could be applied to other databases, but again, the economics of doing so would depend on estimates of the real market in for example Spanish, which could only be reached by this means. A multilingual approach limited to parallel meanings of descriptor/indexing terms provides a limited but practical solution to input and retrieval opeations in a language other than that of the database itself.

Another type of tool of use in a multilingual environment is what might be described as bilingual hardware. There are particular problems in languages such as Arabic, in which different scripts may convey different meanings, and other features may contribute to ambiguity, making computer input operations very difficult. A computer terminal has been designed by ESA/IRS which can employ either the normal Latin alphabet or a standard Arabic with a character coding agreed by 21 Arabic speaking countries. The terminal is able to operate in a normal networking environment; i.e., its signals and signaling requirements are no different from those of a normal Latin alphabet terminal: it does not therefore have to operate in a closed environment. This type of device is specially suited for the computer production of terminology dictionaries, for example, listing the Araric equivalents to French and English technical terminology.

## Software for Retrieval Using Mini- and Microcomputers

There are a number of pertinent arguments suggesting that developing countries, especially, might be able to satisfy needs in the area of information retrieval and information transfer generally by the use of small, relatively portable minicomputers. See, for example, Griffiths [8]. The economic aspects (to be discussed in the next section) make the possibility look attractive, and in the information retrieval area two main technical problems would have to be solved if this possibility is to be

realized. Although the cost/performance ratios of mini- and microcomputers have dramatically improved over the past few years, there is probably an upper limit to the amount of storage which can be handled. Also, large applications programs for information retrieval such as STAIRS, DIALOG, and QUEST, etc. require large amounts of core, impracticable in these machines. Data compression techniques, and the removal of data from each record primarily of concern to librarians but which does not assist the retrieval process, could cut file size by 50%. This would make it possible to store two or three years of an averagely large bibliographic database on the sort of disk which could be handled by a small machine. In parallel with this, simplification of information retrieval systems to make them usable on minicomputers without loss of vital functions seems quite practical: it is understood that a streamlined version (mini-ISIS) of that well-known retrieval system has been developed; Bibliographic Retrieval Services are believed to be working on a streamlined, STAIRS-like retrieval system suitable for very small minicomputer systems. If the software development could also include file creation and file maintenance modules, the essentials for a small but independent stand-alone system, with the possibility of adding local data would be provided: such a system might find ready acceptance in a number of developing countries.

## ACCESS TO ONLINE SYSTEMS: THE OPTIONS

Access to online systems in North America is simple and cheap, since TYMNET and TELENET tariffs are inexpensive, and a node of one or the other network will exist in most major cities. While the telecommunications costs involved in using EURONET or ESANET in Western Europe may be between twice and four times the comparable network charges in the U.S.A., and the telecommunications charges to interrogate U.S. information services are a minimum of $30 per hour, these charging levels seem to be acceptable to a good proportion of potential users in Western Europe. While network nodes are not as widespread as in the North American continent, the use of the public switched telephone network to reach national EURONET, ESANET, or TYMNET, etc. nodes is perfectly practicable and, in most cases, not prohibitively expensive. In Eastern Europe, data networks and links are being established, and regulatory and legal problems inhibiting cross-border data flow in scientific and technical information are now being solved.

In other parts of thw world, however, there are relatively few possibilities for access to online systems: it is very expensive to lease intercontinental circuits, thus inhibiting the growth of commercial networks such as TYMNET to reach developing regions. As already noted, the volume of STI traffic is, of itself, unlikely to be sufficient to justify investment in network extension, and thus STI access in non-industrialized regions through the commercial networks will normally only be possible as a side effect of intensive commercial activity justifying the network extension (e.g., Hong Kong). In general, therefore, developing regions will only be able to use online services in North America and Western Europe by leasing expensive

intercontinental circuits or by using the international public switched telephone system, also a very expensive alternative, and one which may be in any case impracticable for other reasons, for example, the absence of direct dial facilites and the overloading of circuits. In the next section we examine some technical possibilities which might help to solve these problems.

## Telex as a Data Terminal

Several online systems have a telex interface which allows the host computer to be called up on the telex network from a remote station, which can then use its telex machine to send commands and receive responses. The use of the telex network over intercontinental distances is not cheap, but such a system may be more cost effective than use of the telephone network, particularly since it will be unnecessary to rent or buy separate data terminals and modems; telex is moreover fairly widespread throughout the world; maintenance facilites for telex will be part of the telecommunications infrastructure in developing regions before parallel facilities exist for data terminals and modems.

On the negative side, although high speed telex machines operating at similar speeds to data terminals are becoming increasingly popular, the generality of systems is five or six times slower than a data terminal operating over the telephone network; slow speed will inhibit the interaction between the searcher and the data base, as display of a large number of references will take a considerable time and therefore will be expensive. Nevertheless, for a relatively large number of countries, telex offers the only immediate possibility of obtaining some direct benefit from the online information resource; references to major operational experiments in the use of telex in this way do not, however, seem to be very common in the literature. It may be noted that, as part of its Direct Access Project, the International Atopmic Energy Agency has for the past year offered telex access to the INIS and AGRIS databases resident on its computer in Vienna as a supplement to access by TYMNET for regions outside Europe. Use of this facility has not yet built up to a point at which general conclusions on its acceptability and feasibility can be drawn.

## Local Database Operations

It is in theory possible to lease copies of STI databases from the producers, mount them on a suitable computer, and provide services without incurring the high costs of circuits to Europe and the U.S.A. While hardware is becoming cheaper, a major computer installation capable of handling several STI databases simultaneously will cost over a million dollars a year, will require special facilities such as air conditioning and constant voltage power supplies, and most importantly, will need trained personnel for operation, and for file maintenance and update. A local online service will imply an adequate telecommunications structure and a sufficient concentration of users within the

area served to make this a viable option; almost certainly such a system, utilizing a dedicated main computer as the host machine for a local network, may be ruled out for most developing areas. The use of an existing computer installation for STI purposes may, however, be less problematic, particularly if spare capacity exists; however, the provision of network services may still be impracticable either from a technical infrastructure or economic point of view, or both. Nevertheless, the gradual introduction of locally operated STI database services using an existing computer center as a growing point is one of the longer term strategies to be considered. Reference has already been made to the benefits to be gained by exploiting locally produced information by assembling this into a database for planning and decision making in developing countries and regions: the two requirements for developing a potential for database operation and local database creation reinforce each other. Another important point is that while spare capacity on an existing main frame computer is a reasonable starting point for local stand-alone systems, the new possibilities offered by minicomputers are an attractive alternative. Either possibility would lead to decentralized systems not requiring the use of intercontinental computer networks.

## Dynamic Database Distribution

On the basis that minicomputer systems employing streamlined versions of STI databases and retrieval systems were likely to prove a feasible and economic alternative to intercontinental interrogation of online services in the U.S.A. and Europe, COSADOC included in its proposals for communications satellite experiments a dynamic version of minicomputer storage and retrieval [3]. The dynamic database distribution concept is based on the fact that bandwidth in satellite telecommunications systems is not limited by the constraints present in ground based networks. Further, the characteristics of the system as a whole, including its omnidirectional properties, render it distance and location independent within the area covered by a satellite in geostationary orbit. Using a Time Division Multiple Access (TDMA) system, the "space segment" costs (i.e. the costs involved in the satellite channel itself) are likely to be such as to make data transfer at at a megabit rate quite feasible and economic. At certain frequency ranges, notably 11-14 GHz, transmission and reception requirements are such that large central earth stations of the type used in the INTELSAT system are not obligatory. Both transmitter and receiver require a dish antenna of no more than three meters, and are easily transportable. Receive-only stations of this kind in series production will cost in the region of some thousands of dollars: transmit and receive stations are considerably more expensive.

From the informatics point of view, the transfer of a two year portion (inverted and serial files) of a major STI database from a central computer to a remote minicomputer via the satellite would be a straightforward high speed file transfer at megabit rates, which has proved feasible in the European Space Agency's Project STELLA [3]. One version of the experiment

proposed, therefore, was the transfer of a major portion of a database to a "secondary host" (minicomputer) for local searching by a family of terminals attached to this secondary host, the database being deleted from memory on completion of a series of searches in order to free this storage for the next database requested. Other variations of the experiment using both satellite and terrestrial networks in combination were also suggested, but the primary version as described could be relevant to the information transfer needs of developing countries. Should transmission costs for streamlined, data-compressed versions of databases prove to be as expected, it becomes economic to consider searchable portions of them as throwaway articles in order to avoid the storage costs and file maintenance costs associated with maintaining permanent copies of a large number of databases.

The difficulties of implementing such a system is probably less technical than regulatory and economic. The existing INTELSAT series of communications satellites has global coverage, but depends on large and expensive earth stations requiring normal ground based telecommunications networks to feed them. Small earth stations economics do not therefore apply, and in fact, data transmission using INTELSAT, including the ground segments, costs the same to the user as a conventional terrestrial link. It is not known what view the national telecommunications administrations and international carriers will take of small earth station possibilities in setting tariffs for systems favoring this method of operation, although in Europe, the ECS (European Communications Satellite System), due to go into operation within the next two to three years, will be based on experience gained using the currently operating Orbital Test Satellites (OTS), which are designed to use three meter dish antennae as ground stations.

## THE PROBLEM OF FUNDING DIRECT ACCESS

The fact that STI traffic would not of itself be sufficient to provide economic justification for the extension of commercial networks to many developing regions has already been noted. It has also been argued that satellite systems employing small earth stations provide a means of reducing the real costs of direct access in a totally distance-independent manner, and without the need to postulate a highly developed local telecommunications infrastructure at the receiving end. Nevertheless, the basic problem of financing the means to provide a direct access facility still remains; communications satellites are not cheap to design, build, and launch, and although a single satellite has a very large potential area of coverage, justification for the capital investment required would be impossible on the basis of STI traffic alone.

It would be out of place here to discuss in detail a possible requirement for a special telecommunications policy for developing regions, but it seems clear that without some kind of external financing, or regional arrangement, or both, means to provide direct STI access, either by terrestrial or satellite

networks, will only be developed very slowly. Satellite systems at least offer the possibility of serving more users, including those in areas beyond major centers of population and thus having only a rudimentary telecommunications infrastructure. Moreover, satellite systems are broadband and omnidirectional: they can therefore carry a mixture of locally and internationally originated traffic, and can distribute television or video signals in addition to data and telephony. From a financial point of view, therefore, there are a number of possible additional trade-offs which might promote regional or international financing.

# CHAPTER IV

# INFORMATION DISTRIBUTION - DOCUMENT DELIVERY

## THE PRESENT SYSTEM

### The Traditional Role of Libraries

The supply of information in the shape of books, monographs, journals, and reports in science and technology is traditionally the responsibility of libraries. Today's system for collecting and storing material and its provision to the end user is not significantly different, in terms of systems components and functions, from that of the earlier decades of this century. A possible exception is the development of more complex linkages between components to deal with the greatly increased flow of information into the library system as a whole and the consequent heavier demands placed upon it by end users.

The information explosion in science and technology has tended to make it impossible for individual libraries, at the local, university, or even regional level, to be able to support its users solely from its own stocks, and this has had two effects. First, extensive interlibrary lending systems have grown up in order to permit load sharing, and second, side by side with the establishment of a local-regional-national hierarchy of libraries linked by interlibrary loan systems, libraries specializing in particular scientific topics or areas of applied technology have become indispensible in catering to the needs of specific classes of users. As has been noted in Chapter II, user needs in many areas are no longer entirely satisfied by access to the formal literature supplemented by report series: other sources such as trade papers, market surveys, and other varied types of materials are also necessary.

## Informatics Applications - Library Automation

The routine library functions of acquisition, cataloguing, and locating individual items have been greatly assisted by informatics technology designed to streamline these processes generally and to economize in manpower. The MARC database system provides information on national reference/deposit library holdings, originally for the U.S. Library of Congress, but now extended to several other national collections. Systems such as the Ohio State College Library Catalogue (OCLC) and many others enable cataloguing data to be exchanged between libraries, thus saving labor, and also provide convenient locating systems for specific books, journals, etc. Location, that is the identification of libraries holding a particular item not in stock at the requester's own library, is of primary importance in the interlibrary lending system: computer produced union catalogues, sometimes online, are useful tools in this respect, but the compilation of up-to-date and regularly updated union catalogues for journals, report series, and monographs on a national level is an extremely difficult undertaking, if it is to be totally comprehensive.

A further area in which informatics technology is applied is that of routine housekeeping and management of information within a library. Such systems, providing online input and output can be of great value in saving clerical and other labor in even small libraries. They provide a check on orders, books received, invoices received and paid, journal subscriptions due, loans, etc., etc. Cataloguing information can also be added and indexes to the contents of library holdings can be generated if desired.

The use of reprographic techniques in easing the problems of library operation and management may also be briefly noted. Microfiche and microform document stores are now universal for all kinds of documents: computer output on microfiche avoids an unnecessary photography-from-print-out stage in storing analog versions of texts, etc. produced on a computer.

## Document Delivery Mechanisms

An increasingly large number of requests for document copies or loans which cannot be met from local library stocks is being generated. In the case of some industrialized countries such as the U.K., the total figure of requests reaching the main supplier at the national level (British Library Lending Division) is known, but what percentage this is of the total demand can only be estimated. BLLD fulfills about 3 million requests per year, some of them from overseas: probably an equal number of document requests are fulfilled at local or regional levels. All the major document supply centers in Europe reported growth rates of around 10% per year during the mid-1970's with some tendency to reduce toward the end of the decade. There is also a growing international trade in document requests, indicating that even major national centers are unable to meet a small but significant proportion of requests,

and also perhaps reflecting a growing requirement for assistance in this respect from developing countries who may be faced with difficult problems in this area.

Document supply systems at major supply centers involve mechanisms for handling orders, locating documents for loan or copy, transmission of documents, and billing and invoicing.

Ordering is usually by post, possibly by a prepaid coupon system which avoids the need for invoicing and accounting procedures. BLLD accepts telex orders, automatically reformatted by minicomputer on reception to conform to the order form layout. (A multiple copy order form is used as the basis for process and stock control.) Online ordering procedures have now been introduced by the major online services for certain classes of documents where a particular center has the unique responsibility for document supply, for example, the National Technical Information Service (NTIS) or its overseas agents, for research reports resulting from U.S. Government contracts. Current online ordering systems involve marking of individual items displayed at the user's terminal by symbols recognized by the host computer as indicators that they should be transferred to an order file, which is then interrogated by the supply center on the mailbox principle. A more elaborate system, which provided for a delay to check local library stocks, the possibility of using several supply centers, and for user preference, was suggested by the present author in Reference 9.

Location at the supply center will involve the use of stock lists, union catalogues, and the like, but an interesting variant to deal with items not found in this way is the creation of an open-order file on a computer network which can be interrogated by other supply centers who may have the item in question: TYMNET is used by NLM, NTIS, and BLLD for this purpose.

Transmission of documents is normally by post, and it is this fact that causes much of the delay in fulfilling document requests. The Franklin Report [9] analyzed this problem, and showed that the best that could be expected within an industrialized country was one week to ten days, even though 60 to 70% of the requests were dispatched within 48 hours of their reception. If the international mail service had to be used, the total delay was much longer.

## PROBLEM AREAS

The problems in document supply range from technical and operational to legal and political. Although elaborate supply systems have been introduced to keep pace with increasing demand,

delays involved are thrown into sharp contrast with the speed at which relevant literature references from a bibliographic database may be found in an online search. The increasing demand for copies of documents is one factor which has brought the copyright question into prominence: the growing international traffic in copies of documents has further complicated this issue, since copyright practices are not uniform throughout the world. Political problems will be encountered as soon as documents are delivered electronically, since this will involve complicated questions of cross-border data flows, possibly including legislative aspects. These do not arise when the flow is by mail. All these problems affect both industrialized and developing countries, although such difficulties encountered by the developing nations may be compounded by the consequent drain on their convertible currency resources. Original copies of journals will normally require subscriptions in hard currency, and photocopies of articles of journals not available in the country will normally have to be obtained by the same means. Many developing countries will have to rely to a great extent on foreign sources for a large fraction of the material demanded by their users: central and university libraries specializing in science and technology may be of comparatively recent foundation, and therefore, will only be able to supply comparatively recent items from stock. Some notes follow on the copyright problem, particularly as it is likely to affect developing countries.

## The Copyright Problem

The Franklin Report [9] gives some detail on the way in which copyright is handled in various European countries: in the U.K., for example, photocopying is allowed without payment to the copyright owners if the user certifies that the copy is made for his own use for private study and research. In other countries, a photocopy made under these conditions is, theoretically at least, a violation of the copyright law. Multiple copies of the same document are regarded as a direct infringement, and there has been at least one successful law case by the publishers in this respect.

Scientific publishers, particularly those of journals with small circulation and those publishing in languages other than English, fear that a combination of rising costs, limited library budgets, and extensive photocopying will destroy their business. In the U.S.A., a new copyright law concerning photocopies is now in operation: royalties are paid to a central clearing house, and these amount in some cases to several dollars per copy of an article. In Europe, the Commission of the EEC has suggested that, while harmonization of copyright practices may be the long term goal, a system of voluntary license arrangements between libraries and publishers should be negotiated; so far, response has not been very enthusiastic. Users, who already have to pay the document supply center for the costs of the photocopying service, would be required to pay a further royalty charge. While this seems to be accepted in the U.S.A., this

would add still further to the problems of developing-country library and information services already short of funds for overseas purchases, and for that matter, would increase the financial difficulties for many university libraries elsewhere. A suggestion has been made that a two tier royalty system might be acceptable, zero royalties being charged for certain classes of user: obviously there are practical difficulties in such an arrangement. For the present at least, a universal ban on royalty-free photocopying for private study and research does not exist.

## A TECHNICAL SOLUTION: ELECTRONIC DOCUMENT DELIVERY

A few years ago it would not have been possible to conceive of a system in which whole texts were stored on a computer and delivered electronically to the user, but technical developments in computer memories, in electronic data capture, and in high speed data transfer have totally changed the picture. The larger scientific and technical publishers are already showing a lively interest in electronic document delivery in the context of electronic journal publishing. Lerner and Mick [10], in a report on an experiment on electronic document transmission carried out under the leadership of the American Institute of Physics, point out that the implementation of electronic publishing and delivery systems by publishers would give them back the control over the circulation of their copyrighted material which they lost on the introduction of widespread photocopying.

The present author [9] has pointed out that a marriage between text processing systems communicating with each other by computer networks and satellite telecommunications for distribution is now technically feasible and promises to be economically attractive to producers and users where speed of information exchange, rather than the recording of material in archival form, is the main requirement. The European Space Agency's COSADOC group proposed two experiments on electronic document delivery, as reported in Reference 3: the first of these concerned the transfer of pictorial data in the form of earth resources satellite pictures, while the second was a text transfer system for the experimental delivery of the texts of articles and reports resulting from a normal online search.

### The Artemis Project

The Commission of the EEC is holding a demonstration/workshop in December 1980, to be attended by information policy makers, industry, online service organizations and user representatives, publishers, and librarians. It is proposed to demonstrate individual components and total systems for the various elements of electronic document capture and delivery, to assess technological and economic feasibility, and to discuss how actual systems based on the physical elements demonstrated might be implemented. This follows an in-depth study of document digitalization and transmission carried out by

Arthur D. Little for the Commission in 1979-80. This study resulted in the publication of the ARTEMIS Report [11]; it concluded that all the processes for capture, storage, retrieval, and teletransmission of STI material are already state of the art, and that their application could be economic provided that certain conditions were met; the authors were, however, of the opinion that market forces alone were not sufficient to bring the system into being. There is thus considerable movement on both sides of the Atlantic towards practical electronic document delivery systems: the present status of the required informatics technology is summarized in the following paragraphs.

## Document Capture: Digitalization

An existing document in printed form may be converted into digital form by several methods. First, high speed digital facsimile, corresponding to Group III of the CCITT standard, is now commercially available and will convert an A-4 page into data-compressed digital signals, temporarily stored in the buffer memory of an intelligent terminal into which is incorporated the document scanner system. Speed of scanning is fast: a second or two is sufficient to capture a page of text, and there is a possibility of adjusting scanning rate so that for difficult text, high resolution can be applied at some slight sacrifice in scanning speed. Line drawings may be reproduced and for other drawings, shades of gray between black and white are also possible. These systems thus produce an electronic image of each page: the number of bits required to capture a page of text is, however, rather high.

Second, optical character recognition (OCR) may be employed to "read" existing texts directly into digital form on a computer. These systems employ pattern recognition algorithms in which characters read by the scanner are compared with standard images of characters; the advantage of OCR-type methods is that they require fewer bits per page of text than digital facsimile, since once identified, each character can be handled in standard coded form, while the disadvantage is that although present systems are able to read several different type fonts, input must be monitored to deal with the occasional unrecognizable character.

Slow-scan video systems may also be employed to read pages of text or microfiche by what is essentially a television camera: because the video system does not have to reproduce movement, scanning can be so arranged that storage of scanned material on videotape or cassettes at reasonable densities is quite practicable. The disadvantage of such video scanning systems is that the derived electronic signal is broadband and is therefore difficult to accommodate on normal data circuits.

## Teletex

If texts can be directly input from a normal computer terminal keyboard, employing one of the standard systems for coding characters into bits, a digital record is created which can then

be stored, transmitted, and reproduced by classical data processing methods. The advantage over facsimile systems is that many fewer bits are required to store a page of text: the disadvantage is that while a graphics terminal can be used to input line drawings, tables, etc., digitalization of more complicated illustrations is difficult and costly and the resulting quality may not be acceptable to a reader used to high quality printed products. Nevertheless, many manufacturers are developing whole systems of text composition, electronic manipulation and reproduction, partly as a result of the new markets opened up in commerce and industry by word and text processing systems, and partly because of the extension of such systems into electronic mail and the fully automated "electronic office". CCITT is now considering a draft standard which could be applied to text processing and electronic mail systems so that all could communicate over the public switched network, in much the same way as telex subscribers communicate with each other today. One major manufacturer is already marketing text processing systems which can accommodate many different types of "virtual keyboards", controlled by software, so that all kinds of type fonts and symbols may be used, without changing the physical keyboard itself. Another major manufacturer [12] has pointed out the extremely rapid growth of sales of word processing systems adapted for remote communication over data or telephone circuits: from a few percent of their total sales in 1978, the proportion of word processing systems for remote use employing telecommunications rose to 30% in 1979.

The implications of teletex systems for document delivery are probably less immediate than digital facsimile, since manual re-keyboarding of existing texts is expensive and may involve difficult problems in certain areas (complicated mathematical equations, chemical formulae, etc.). Nevertheless, a system based on re-keyboarding, but co-opting digital facsimile to solve such difficulties has been suggested, and it is claimed to be economic for document delivery.

The real interest in document capture by coded character methods is, of course, that if used at the outset of the publication process, they can be used both to create a master text for photocomposition or other form of computer-based printing operation, and at the same time provide an electronic master text for distribution in any way that may be required (single articles, personalized journals, and any form of repackaged collections). The electronic master for printing will differ slightly from that for teletransmission, in that the former will incorporate special composition codes for the computer controlling this operation, but these may be stripped off to provide the second master. It will probably be necessary for some time to come for difficult parts of the text requiring special characters, drawings, or layouts to be dealt with by facsimile methods: this could present difficult problems at the receiving station in the document delivery mode.

Storage, Transmission, and Reconstitution of Documents

The ARTEMIS study proposed that, once texts were digitized (either on an on-demand basis or to build up a stock of likely high-demand material in advance), supply centers, which could be publishers or full text host service organizations, would maintain full text databases which would then be addressed by online ordering systems so that relevant texts could be retrieved, electronically copied, and dispatched on the distribution network (EURONET in the first instance). The study examined the economics of the various possible storage mechanisms, including magnetic disk, optical disk, and more exotic systems such as laser memories. It concluded that with existing technology, storage costs would be acceptable (between approximately $.01 and $.06 per page). The main archival storage would not need to be instantly accessible to an order message in the general case, since orders would be accumulated during the day, batched in working storage, and transmitted over the network at night when interactive traffic will have practically ceased. It may be that a packet switched network is not ideal for electronic delivery of whole texts, since from an informatics point of view, this amounts to a file transfer, the data stream moving point-to-point at a steady rate. However, the main problem at the outset is more likely to be economic than technical; present network tariffs are geared to interactive traffic in which relatively few data bits flow between computer and user, and the main part of the charge is accounted for by the elapsed time the circuit is occupied. The ARTEMIS report suggests that, in the case of EURONET, existing tariffs would result in transmission charges of up to $.5 per page for facsimile, the most likely method at the start; transmission charges might be too high to attract document delivery traffic. (Charges for teletex transmission would be cheaper by a factor of about 5). A tariff more closely geared to the special features of electronic document delivery should therefore be negotiated.

There are two different philosophies in handling the reception of electronically-delivered documents. The first, stressed in the ARTEMIS report, suggests that local centers might be established at which high speed facsimile receivers would be installed, in association with high speed printers or intelligent copiers which could sort the output for dispatch by appropriate means to clients, for example by the use of messenger services. Major users would, of course, have their own document reception center, while small users would use services offered by third parties. A second system configuration involves user terminals capable of receiving document copies directly. Buffered teletex-type receivers could perhaps store documents received overnight, for display and printout as necessary the following day. Technically, the same could be done for facsimile systems, except that, currently, these would be costly if they present all the features necessary for high resolution, high quality unattended document receptions. The difficulty with the combined user terminal/receiving station is that documents will be transmitted through the network at high bit rates but small and medium-sized user organizations will normally be connected to the network by slow

speed devices. Ideally, the user should only need to purchase one terminal for both information retrieval and document delivery, and this may well be a slow speed device. Possibly an optimum solution might lie in intermediate reception centers for electronic storage of transmitted documents, to which users could dial to receive the text of their documents. A pilot experiment is needed to determine the optimum system configuration.

## Other Electronic Document Delivery Possibilities

Replacing the relatively slow speed terrestrial network by a satellite link would enable virtually instantaneous delivery of documents to large user centers possessing small ground stations. The American Institute of Physics/NASA experiment reported in Reference 10 employed an experimental communications satellite for delivery of full text and this turned out to be quite practicable. However, the experiment threw little light on the need for high speed, high volume full text transfer, since, for the most part, the documents available via the satellite were also available within the participating centers' own libraries. Even within the constraints of delivery by terrestrial network, real time document delivery could provide a browsing possibility which users might find of considerable value, although much would depend on the cost factors involved (time spent and volume transmitted). The possibility of retrieval using the full texts of documents, and the experimental system to be introduced by Bibliographic Retrieval Services in this respect has already been mentioned: it is clearly closer related to the browsing possibility offered by full texts online.

## Electronic Document Delivery and Developing Countries

The possibilities discussed in the preceding paragraphs are of potential importance to rapid information transfer to and between developing countries; they would be able to rapidly accumulate electronically stored copies of wanted documents, to be reproduced in paper form as and when required. The present overlong delays in document delivery by mail would vanish, and providing that a solution to the copyright problem could be found, many of the current severe problems in document supply would be solved. Unfortunately, the same factors that inhibit the use of online database services located in industrialized countries by centers and individuals in developing countries apply here also: electronic document delivery to developing countries requires an extension of present computer networks to those areas, or the renting of expensive long distance circuits to connect them with existing network nodes. We have already seen that while there are almost certainly no technical problems in establishing ARTEMIS in Europe, there could be an economic problem in applying existing tariffs to document transmission by facsimile methods. The cost situation in full text transmission using intercontinental networks would presumably be much less favorable, unless and until the total volume of traffic was sufficiently high to produce overall economies of scale, and therefore cheap tariffs.

Woodward [1] points out that the new informatics technologies, in which we can presumably include those associated with electronic document delivery, although potentially able to provide a breakthrough in information transfer, may not turn out to be an unmixed blessing for developing countries. Each such advance in industrialized countries will absorb skilled manpower for local exploitation in developing countries, and may in any case be dependent on the existence of technical and information infrastructures, which are very weak. He concludes that further steps towards the paperless information society in science and technology might, in the worst case, have the effect of further isolating developing countries from information sources elsewhere.

# CHAPTER V

# STRATEGIES AND RECOMMENDATIONS FOR APPLICATION OF INFORMATICS TECHNOLOGIES

## INTRODUCTION

The previous three chapters have been concerned with current problems in information transfer, both general or universal in character, and those which are specific to developing countries and regions. In this analysis each STI area, information capture and ordering, information retrieval, and information distribution, has been dealt with in isolation, both as concerns the problems themselves and the informatics technology which might assist their solution. In the present chapter we are less concerned with the solution to technical problems than with suggesting practical strategies at the policy and administrative levels within which the relevant and appropriate technologies might be encouraged at the national, regional, and international levels.

Strategies will, for the most part, cut across the functional divisions of the STI process in which the material in the previous three chapters was established; nevertheless, it seems advisable at this point to briefly list the problems already discussed in each functional area as a type of checklist for the discussion of possible strategies.

### Summary of Problems by Functional Areas

Information capture and ordering - delay in information capture (publication): delay and difficulty of ordering the growing resource to permit effective retrieval; the where-to-look problem resulting from the recent rapid increase in the numbers of bibliographic databases and databanks of all kinds. In particular for developing countries, the problem of collection, recording, ordering, and dissemination of locally produced information not part of the formal STI stock.

Information retrieval - problems arising from the imperfections of the user/system interface: more and better intermediaries and intermediary training tools. Improved user friendliness: more multilingualism. In particular for developing countries, minicomputer database and retrieval systems, and technical, operational, and financial problems associated with provision of direct access to online information systems in industrialized countries.
Information distribution and document delivery - lack of timeliness in present systems: organization of document delivery, including request processing, location, and transmission of documents. The copyright problem, especially cross-border problems caused by differences in national legislation. Especially affecting developing countries, funding problems associated with building up adequate collections on a local basis, or purchase of copies from overseas: impact of possible copying royalties; technical, operational, financial, and legal problems associated with direct access to future electronic document delivery systems.

### Independent and Dependent Strategies

Solutions for problems in information transfer for developing countries, either already rising or likely to arise in the near future, tend to group themselves in two main directions. Taking the information retrieval area as an example, one strategy might take the form of concentrating on means to exploit the resource offered by the existence of online bibliographic database and databank services in industrialized countries. This is attractive, because it would enable a developing country having direct access to avoid the expense and difficulty of building up its own information retrieval service requiring database copies, the necessary computer facilities to operate them, and provision of the necessary skilled personnel to maintain a service. On the other hand, the means to obtain direct access (intercontinental leased circuits or extension of computer networks) are either very expensive or otherwise very difficult for most developing countries. However, the analysis in the preceding chapters tends to show that the dependent or direct-access strategy and the independent or local-facility strategy should not simply be regarded as alternatives from which a choice has to be made; instead, the situation seems to require the combination of elements from both, and such a combination will be discussed in the next section.

## ESTABLISHMENT OF NATIONAL AND/OR REGIONAL FOCI FOR INFORMATION TRANSFER

Even if the problems of direct access could immediately be solved, it would be necessary to create a focus at the national level (or possibly in appropriate cases at the regional level). This focal point would constitute the interface between end users and the online system used; it would provide intermediary services including particularly education of users in what was available and how best to use it.

It would be natural that such focal points, once established, should also become involved in organizing information locally, i.e., its collection, recording, storage, and processing to provide (initially) local access, and possibly later, access on an exchange basis with other countries.

In some countries or regions it may be appropriate to set up more than one focal point, particularly if there is more than one main development activity requiring an information backup, with a corresponding division of responsibilities at the national level. Probably, it will be more usual to concentrate available resources in a single center at the outset, with the possibility of creating specialist subcenters as the need arises; with relatively small resources, better results will be achieved by concentrating rather than dispersing them. The best location for a focal point will naturally vary from case to case: a major library in a government agency or in a university having close connections with industrial or agricultural development would provide the kind of basic information infrastructure which could be built upon. Access to an existing computer center, or the possibility of establishing and staffing a small stand-alone computer operation would also be important.

Such focal points, having responsibilities in access to overseas online services, contacts with other information services having information relevant to local problems, and the organization and management of locally produced information, fit well into the concept of National Information Systems (NATIS), already formulated by UNESCO. Major international cooperative information programs managed by other UN agencies such FAO and IAEA have the effect of building up centers for scientific and technical information in developing countries, usually on a specialized basis; such centers might also form the nucleus of the kinds of foci suggested here, if the situation in a particular country calls for a measure of specialization in the organization of the services.

*Recommendation 1*

(a) On the national (or regional where appropriate) level, steps should be taken to establish, within the the NATIS framework, information services as foci for exploitation of overseas online services, organization and management of locally produced information, and liaison with other overseas sources of relevant information. Such centers should be integrated with the management of development programs, and be organized to provide the necessary interfacing functions between users, and information and informatics resources, including local computer services.

(b) On the international level, all practical steps should be taken to further the establishment of local information centers as summarized in (a) above. These steps could include the study of local situations, in

conjunction with personnel in the country concerned, to advise on the most appropriate organization and structure of the centers: help with training of potential information and informatics staff, and other pump-priming activities, including initial help with funding the establishment of such centers through technical assistance and aid programs. Coordination between international organizations having international cooperative information programs and national agencies in industrialized countries having bilateral technical assistance programs with developing countries should be furthered.

## SELECTION OF APPROPRIATE INFORMATICS TECHNOLOGIES

This report suggests that the pace of development of new informatics technologies is very rapid; new technologies are developed for a variety of applications, not solely or even mainly for STI transfer. Technology mechanisms, and STI systems and services are well enough developed in industrialized countries to ensure that technologies developed for a non-STI application (for example digital facsimile, Teletex, and the electronic office) are applied as necessary to STI transfer. Technology review and selection for this purpose may not, however, be quite such an automatic process in developing countries. Part of an overall strategy for applying informatics technology to information transfer in developing countries would therefore be the development of mechanisms to promote the intelligent matching of problems and the appropriate technologies to help solve them.

The strategies should embody several different approaches. The first involves straightforward reviews of informatics developments and STI problem areas (general and specific) as in this report. This approach can, however, only give rather general indications and needs to be extended in other directions. Second, the potential value of particular developments, for example the use of regional communications satellites, may be obvious, but the practicability of applying them may require in-depth studies of other factors, such as the possibility of other traffic, regulatory and infrastructure problems, and these may differ in different regions. Thirdly, possible application may be missed unless the particular situation and the problems arising in developing regions and countries are fully appreciated and understood by those with the task of reviewing new technology. In essence, the strategy proposed is directed towards a more systematic exchange of information about new technologies on the one hand, and about problems on the other.

*Recommendation 2*

To ensure that new technologies match to actual problems in developing countries, studies and reviews should be carried out at both national and international levels. Such studies should be integrated; that is, the technology

reviews and the problem-oriented studies should be undertaken in concert, with full interchange between the teams or individuals on both sides.

## DIRECT ACCESS TO INDUSTRIALIZED COUNTRY INFORMATION RESOURCES

Whether the information resource to which access is required by developing countries is the present day online systems available over computer networks in Europe and North America, or whether it is tomorrow's online document delivery system, strategies must be devised to enable users in developing countries to access such systems at prices they can afford. Access possibilities range from telex and the public-switched telephone system to leased circuits and extension of commercial and other computer networks to bring them within reach of users outside the present areas of coverage. The study has shown that regional communications satellite systems working with small earth stations offer an attractive alternative, perhaps not requiring highly developed communications infrastructures in developing countries to support them. However, it is doubtful whether either the extension of terrestrial networks or the establishment of special satellite links could be economic in present circumstances: the volume of traffic generated will be insufficient to justify the costs involved. While every effort should be made to improve the economics, for example by promoting regional rather than national arrangements and by planning diverse use of the proposed new communications facilities, it seems clear that market forces alone will not bring about the desired result, at least in the near future. International telecommunications carriers, whether they are corporations or government bodies (PTT's) must operate as quasi-commercial organizations.

The problem of providing direct access for developing countries to facilitate information transfer is one facit of a much larger problem, that of providing improved telecommunications facilities for a variety of purposes on a global scale. The relation between better telecommunications and the pace of development in developing countries is far from clear in general, but in the specific case of scientific and technical information transfer, there appears to be a close connection; the STI community would, therefore, have a contribution to make to any major study directed towards a global policy for improved telecommunications as an aid to development.

In the more limited perspective of possible actions in the near and intermediate term to improve direct access to information resources in industrialized countries, a practical strategy might be to provide direct financial and technical aid to countries and regions able to take an immediate advantage of this type of assistance. This aid could take many forms depending on circumstances. One possibility would be to provide financial help to rent circuits so that simple multiplexer connections could be made between a number of user locations in a developing region and an online service or a network node. For reasonable economy, multiplex network extensions should be organized on a regional basis where this is practicable and the local telecommunications

administrations invited to collaborate by providing national circuits free of charge. Multiplex extension of this type could only be provided where the right combination of potential user concentrations, adequate telecommunications facilities, and reasonable proximity to an existing network node existed. Other possibilities would be the direct subsidy of users by assisting them to meet bills for direct access to online systems where this this was practicable by international dialed telephone or telex circuits but where the combined charges for telecommunications and database access were prohibitively high.

*Recommendation 3*

(a) On an international level, means should be explored to provide centers in developing countries with cheap direct access to online services. The means could include direct subsidy of the use of online systems by international dialed telephone and telex circuits, and the leasing of circuits to provide multiplex network extensions in particular regions. The collaboration of national telecommunications administrations should be sought in this respect.

(b) On the international level, the STI community should offer its collaboration in telecommunications policy studies delineating the connection between improved telecommunications, including regional satellite systems, and development, and the means by which such facilities could be financed.

(c) At the regional level, collaboration between telecommunications administrations should be furthered, with a view to optimizing the use of multiplexer links to online services overseas at minimum costs to the user.

## THE DEVELOPMENT OF STAND-ALONE SYSTEMS

The short term possibilities for providing direct access just discussed are not applicable across the board: moreover, it could be argued that they should be seen as ameliorations, rather than cures for the situation in which many developing countries find themselves in their inability to take advantage of the online information resources. It therefore seems necessary, as a complementary strategy, to further the development of stand-alone systems, especially in view of the priority which must be accorded to developing and exploiting local information resources which could be of direct and immediate importance in development programs.

It has already been noted that practical database and information retrieval systems for use on minicomputers are already almost state of the art. It would seem quite practical to carry out the necessary further development so that total packages could be offered, including hardware, compressed versions of the database, and the appropriate retrieval system tailored to the hardware

and database, together with training aids. Further, a database management system could be incorporated for the processing of local information. International organizations operating major cooperative information programs would be particularly well placed to provide such facilities for the databases they themselves create, while full advantage should be taken of commercial initiatives in developing minicomputer-based database and retrieval systems. In the longer term, use of regional communications satellite systems for operation with small ground stations would provide another dimension to the use of minicomputers for information retrieval by the dynamic distribution of databases.

*Recommendation 4*

(a) On the national level, existing computer centers in developing countries should be encouraged to experiment with database and databank systems to provide a local information retrieval facility using externally produced databases of importance in development projects, to the extent that the necessary trained manpower can be be made available. [See also Recommendation 1(a).]

(b) On the international level, the circumstances under which minicomputer database and retrieval systems could provide a useful stand-alone facility in developing countries should be investigated. Consideration should then be given to designing complete packages comprising hardware, software, databases, and training facilities where these are appropriate. [See also Recommendation 1(a).]

## EASING DEVELOPING COUNTRY PROBLEMS IN INFORMATION MANAGEMENT

The report has identified a number of problem areas in the capture, ordering, and supply of information sectors, which, while general in character, are likely to particularly affect the development of information systems in developing countries. In so far as these problems can be resolved by the application of informatics techniques, they will probably be resolved in industrialized countries first: electronic delivery of primary document copies to overcome the delays in present delivery systems is an example. In such areas therefore, the most appropriate strategy for UNESCO might well be to closely monitor what is being done in industrialized countries and indicate how developing countries might take advantage of the work being undertaken when this seems appropriate to their needs and state of development.

Aside from such areas as electronic publishing, the use of broadband satellite systems for document transfer in real time, and electronic document delivery, there are a number of less revolutionary developments which can more directly help developing countries in information management, either of their own local information resource or of the external resource typified

by online systems. Several of these (an obvious example is multilingual tools) are of great importance also to industrialized regions, and in these cases also, the appropriate strategy would appear to be to monitor developments and assist their application to developing country problems. In other cases, the problem, while universal, is less of a barrier to information transfer in industrialized countries because existing infrastructure and the skills available can do much to overcome its effects. An example is the where-to-look problem: intermediaries in information centers in the U.S.A. and Europe may find the absence of a sophisticated, continually updated online referral system (database on databases) a hindrance in their work, but the problem can be overcome in other ways. The same is not necessarily true for their colleague in a developing country, who may have little previous experience and no help conveniently at hand. In such cases therefore, the appropriate strategy must be to try to inject greater priority in the development of solutions to the problems concerned, and to ensure that they are directly applicable in developing countries.

It should perhaps be stressed in this connection that any action to ease the problems of information management can only be effective to the extent that an infrastructure and the requisite skills exist: implicit in the recommendations which follow is the necessity to provide a minimum level of information and informatics infrastructure at the national level. This is probably the most important single problem in information transfer to developing countries; informatics technology can only make an indirect contribution to its solution.

*Recommendation 5*

(a) At the national level, high priority should be given to the selection of qualified personnel as intermediaries. Similarly, at the international level the organization of training schemes for such personnel should be given high priority.

(b) At the national and regional levels, priority should be accorded to participation in international or regional collaborative programs for the collection and ordering of locally produced information.

(c) At the international level, developments in informatics technology of direct value to the build-up of information services in general, and intermediary services in particular in developing countries, should be actively furthered. These include:

- -- simplification of retrieval languages: extended application of a common command language,
- -- online, in-depth guide to the contents of databases and databanks,
- -- research-in-progress systems: multilingual tools.

(d) At the international level, monitoring of informatics developments in industrialized countries with a view to facilitating the transfer of appropriate informatics technologies to developing countries.

## LEGAL, REGULATORY, AND ADMINISTRATIVE-FINANCIAL ASPECTS

There are a number of this class of problems which could inhibit or restrict the process of information transfer: most have no connection with the application of informatics technology, but as actual or potential braking forces they cannot be ignored in any overall strategy for improving information transfer. On the legal side, the main difficulty is in the widely different practices in handling the copyright of scientific and technical publications; a user in Country X, where the "single copy for private study and research" principle is regarded as an infringement of copyright may order from Country Y, where the principle is accepted. Similarly, users will try to avoid sources where photocopying royalties are charged. A harmonization of copyright practice in the direction of royalty payments would add still further to the hard currency problems of users in many developing countries. We have also noted that difficulties in obtaining foreign exchange and the increasing costs of journals and books, are factors hindering the flow of information to many developing countries. Exploiting the online information resource will also involve new demands on foreign currency available for information and this could result in expenditure cuts in other areas, for example journal subscriptions.

It is difficult to suggest any general strategy to help solve these problems, all of which have the effect of increasing demands on the hard currency resources of developing countries. Whether the amounts in question are significant or otherwise is perhaps beside the point: exchange control procedures will apply to small as well as large invoices. From the point of view of the management of an information center in a developing country, block grants of foreign exchange for information materials, rather than having to clear each purchase through the central bank, simplifies the problem. Industrialized countries might consider the flow of certain types of information products on a no-billing basis in the bilateral aid programs.

Practical proposals and recommendations for action to solve these problems are clearly important, but are somewhat outside the scope of this report: a different type of analysis is required. Neverthless, a conclusion may be drawn from the material presented: it is that the application of informatics technologies to information transfer is likely to increase the administrative/financial difficulties faced by information centers in some developing countries; legal and regulatory factors may indirectly act as braking forces in information transfer. The situation requires further review and study.

REFERENCES

[1] Woodward, A.M. 1980. Future Information Requirements of the Third World. Journal of Information Science 1: 259-265.

[2] Newsfront. In Data Communications. 1978. December:15-18.

[3] Hanell, S. 1980. Telematics and Satellites. Part 1: Information Systems, edited by W.R. Burke. Paris: ESA Scientific and Technical Publications.

[4] EUSIDIC Database Guide. 1979. Oxford: Learned Information Ltd.

[5] Tell, B.V. 1980. The Awakening Information Needs of the Developing Countries. Journal of Information Science, Principles and Practice 1(5):285-289.

[6] Cuadra, C. 1980. Surviving the Eighties: New Roles for Publishers, Information Service Organisations and Users. NFAIS Newsletter:22(2).

[7] Dubois, C.P.R. 1979. Multilingual Information Systems: Some Criteria for the Choice of Specific Techniques. Journal of Information Science 1:5-12.

[8] Griffiths, J.M. 1980. Application of Minicomputers and Microcomputers to Information Handling (Draft). PGI-80/WS/24. Paris: UNESCO.

[9] Gillespie, P.D., P. Katzenberger, and J. Page. 1979. Problems of Document Delivery for the EURONET User. Report for the Commission of the EEC prepared by the Franklin Institute GmbH, Munich. Munich: K.G. Sauer.

[10] Lerner, R.G., and C.K. Mick. 1980. Database Searching and Document Delivery Via Communications Satellite. Final Report to the National Science Foundation, Division of Information Science and Technology. Grant No. DSI 77-18052 to the American Institute of Physics. AIP 80-1. New York: American Institute of Physics.

[11] Little, A.D. 1980. ARTEMIS: A System for Document Digitalisation and Teletransmission. Report to Directorate General XIII of the Commission of the European Communities. Luxembourg: EEC.

[12] Walshe, W.A. 1979. Electronic Mail Diversifies with Technological Innovations. Word Processing World, April: 15-22.