DOCUMENT RESUME

ED 138 267 IR 004 683

TITLE Computer Networking in the University: Success and Potential. Proceedings of the EDUCOM Fall Conference (Boston, Massachusetts, November 10-12, 1976).
INSTITUTION Interuniversity Communications Council (EDUCOM), Princeton, N. J.
PUB DATE Nov 76
NOTE 327p.

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Chemistry; *Computer Oriented Programs; Computers; *Conference Reports; Health Education; Health Services; *Information Networks; Interinstitutional Cooperation; Library Administration; Library Networks; Management; Medical Research; Physical Sciences; Program Administration; *Universities
IDENTIFIERS EDUCOM; Interuniversity Communications Council

ABSTRACT

This book presents the proceedings of a 1976 Interuniversity Communications Council (EDUCOM) conference focusing on the successful application of computer networking to education and research in North American universities. Thirty-two papers share the common theme of the need for management and planning, user services and support, and facilitating organizations. Topics include: (1) issues in computer networking, with emphasis on political, legal, technological, and marketing problems; (2) applications of computer networking to library use and administration, as well as to education and research in the health, physical, and chemical sciences; and (3) practical aspects of administering computer networks. A directory of participants is included in the appendices. (Author/KP)

Computer Networking in the University: Success and Potential

Proceedings of the
EDUCOM Fall Conference

November 10 - 12, 1976
Boston, Massachusetts

EDUCOM, The Interuniversity Communications Council, Inc.

Library of Congress Catalog Card Number: 77-272

Further information and additional copies of this book may be obtained from EDUCOM, P. O. Box 364, Princeton, New Jersey 08540. The price per copy is $6.00 for members and $12.00 for non-members of EDUCOM.

Printed in the United States of America

CONTENTS

CONTENTS – Continued

PREFACE

Computer Networking in the University — Success and Potential

During the past five years, advances in computer and communications technologies have made possible the rapid growth of the use of computing services provided through networks. Recognizing this growth, the Fall 1976 EDUCOM Conference focused on achievements in networking in four areas of higher education: library services, health sciences, physical and chemical sciences, and the administration of computing services.

The networking successes described at the conference highlight opportunities for cost reduction and sharing of unique data and computing resources. Even more important, however, is the evidence of wholly new types of intellectual cooperation—ranging from collaboration on a journal manuscript in ophthalmology by five authors, to sequential processing of chemical data by researchers at five computing centers having different capabilities. The introduction to this volume gives a more complete description of papers included in the Proceedings.

Structuring the conference program, which included more than fifty speakers and panelists, was a complex and ~~difficult~~ undertaking. Conference Chairman Robert Scott and his committee devoted considerable time and effort to the task; we were rewarded by an outstanding program. Many thanks are due to: Robert Scott, Director, Office for Information Technology, Harvard University; John McGeachie, Director of Computing Services, Kiewit Computation Center, Dartmouth College; Weston Burner, Director, Information Processing

Services, MIT; Peter Meincke, Professor of Physics, University of Toronto; Warren Seider, Associate Professor of Chemical Engineering, University of Pennsylvania; and Eugene Young, Assistant Vice President, Rutgers, the State University of New Jersey.

James C. Emery
EDUCOM President

INTRODUCTION

For the last few years, it has been the practice for EDUCOM Spring and Fall Conferences to explore current potential and future opportunities for the application of technology to the improvement of education and of its management. During this same time period, two major EDUCOM activities have started to address in detail specific aspects of the application of computer networking to the university environment. These two activities are the NSF-sponsored Network Simulation and Gaming Project and the twenty-two member Planning Council on Computing in Education and Research.

The Fall, 1976 Conference sought to provide a forum for reporting on these and related activities and to indicate directions for future development. The theme selected for the Conference focused on success that has been achieved in the application of computer networking to education and research in North American universities. Development of this theme involved a survey of issues in computer networking, as provided by two major addresses, and two panel discussions addressing specific network activities, the degree of their success, the reasons for success or failure, and the problems that will be encountered in the political, legal, technological, and marketing areas in the future as inter-institutional computer networking develops further.

The majority of the Conference, however, was devoted to specific discussions of successes in the use of computer networking in education and research. Four concurrent sessions explored the application of computer networking to library administration and use, to the health and life sciences, to the physical and chemical sciences, and to university computing centers. The various papers that follow in this

Proceedings show that a wide spectrum of success has been achieved. Many common themes emerge including the need for management and planning, for user services and support, and for facilitating organizations.

During the Conference, 16 demonstrations of computer networking were conducted spanning a wide range of applications—library administration, value added service networks, commercial service bureaus, university computing centers, computer assisted instruction, and discipline specific facilities.

Robert H. Scott
November, 1976

Part I Survey of Issues in Computer Networking

Chapter 1

by Joe B. Wyatt

Computer Networking in the University — Success and Potential

The Airlie House conferences on Computer Networking, sponsored by the National Science Foundation and developed by EDUCOM, were held in late 1972 and early 1973 at a time when the ARPANET had already demonstrated the feasibility and some aspects of the usefulness of national computer networking for research. At those conferences, national computer networking for Colleges and Universities was extensively discussed as a potential contribution to higher education. The result of the Airlie House conferences were described quite well in the book *Networks for Research and Education*, by Greenberger, Aronofsky, Massy and McKinney. A number of goals and a plan were established for achieving the potential of networking for higher education and research. Some of those goals have been met, others still dangle like carrots on long sticks.

Another important outcome was the establishment of definitions and concepts which could be used to communicate about computer networking. Three types of networks were defined: *a user services network* which is people and task oriented and dedicated to the delivery of specific services to specific users, a *transmission network* which engages in the transportation of signals from point to point, and finally, there is the new concept of the *facilitating network*. The facilitating network conceptually includes the bridging functions necessary to join users to network resource types. Into this category of facilitating services falls standards of protocol, conventions for accounting and billing, documentation and user support, and marketing functions. Facilitating services also include the necessary interfacing hardware to accommodate the multitude of components interconnected in such a network. Two important concepts in the facilitating network definition are the wholesale—retail analogy uses to describe the relationship between major resources who supply "wholesale" services and specialists in user services which deliver "retail" services to individual users and groups of users. The other is the "distributed system" which focuses centralization and decentralization issues about hierarchical structures for computer resources.

One of the "largest" ideas which came out of the conference was that a new national Planning Council be established to provide a focal point for the organization, development, and funding of national networking concepts for higher education. In essence, the Planning Council was to concentrate on the development of two of the three forms of networks, namely, the facilitating network and the user services network with the hope that the user services network could begin to organize while the facilitating services are developed and refined. In addition, the Planning Council was to deal with the important questions of organization and structure for national networking in higher education and research.

Where are we now? In quantitative terms, there are several significant numbers by which we can measure our progress since the beginning of 1973, a period of almost four years. In order to establish a point of reference, we can now examine some statistical information which was not available then. In 1974, for example, the Information Industries Association already had 71 members and the estimated gross annual revenues of the members of this organization ranged from a low of $477,000,000 to $618,000,000 for the year. In short, there was already operating in 1974 a commercial version of the user services network with an operating volume which generated a cash flow of about one-half a billion dollars a year. Hardly a trivial activity! Other data for the period indicated that the total annual revenues for the telephone industry in 1973 was $25.5 billion dollars. The total annual revenues for the software and service industry was $3.7 billion dollars for the same year. Moreover, at this point in 1973 the annual growth rate of these two industries was compounding at a rate of slightly in excess of 15% per year. It is also important to note that in all of these cases, the costs for these services was being born and justified by individual users, most of whom presumably had made the decision that the value received from these services was worth the cost.

Transmission networks are now flourishing. In fact, the commercial packet-switched network has developed since 1973. It is interesting to note that Dick Bolt, who was a major participant in the Airlie House conferences, became convinced that the packet-switched network was a viable commercial venture. He recruited Larry Roberts from ARPA, put up substantial funds from his own corporation and managed to raise additional funds to establish and begin to operate the Telenet Communications Corporation which is now the major supplier of commercial packet-switched network services. In addition, other carriers including TYMESHARE, MCI, Southern Pacific, Western Union and a host of others have continued to flourish with transmission networks operated in special market sectors. Other than the much publicized demise of Datran, there seems to be no hesitation in the development of a viable industry around transmission networks for computer usage. In fact, with Western Union picking up much of Datran's resources, it is conceivable that the microwave network envisioned by Datran will become operable.

So transmission networks are alive and well—how about our facilitating network? From my point of view, the progress has been encouraging but not yet sufficient. The Planning Council for Computing in Education and Research has been established with twenty-two actively participating institutions. The participating institutions represent a widely diverse set of colleges and universities located throughout the United States. All of them already devote major attention to the use of computer resources in education and research. The staff of the Planning Council headed by Dr. James C. Emery is now well established in the supervision of the major tasks for the Planning Council. The governing organizations represented by the Policy Board and the Technical Committee have begun to operate as intended with participation and involvement from both the executive level and the senior technical level at each of the member institutions. It seems appropriate to examine some of the major projects of the Planning Council as illustrations of the facilitating network concepts described previously.

An idea which emerged from the Airlie House conference was the development of a "network simulation model" which would allow the systematic evaluation of the impact on a school of participation in a network. A small group of people, encouraged by the prospect and experienced in the development and use of simulation models met in Cambridge late in 1973 and concluded that it was feasible to develop a simulation model that would at least contribute to the knowledge about the financial impact of network usage, if not answer the other questions as well. Very fortunately, this group received enthusiastic encouragement and support from the National Science Foundation and the work was begun with a group of cooperating investigators to study the characteristics of such a model and determine its feasibility. This effort has resulted in a major research project for the development of a network simulation model which will allow individual institutions to simulate various conditions of network use to support decisions about computer resources. It is encouraging to report that the project has already achieved its initial goal for the development of a basic simulation model. This model now simulates the activities of a group of twenty-two institutions. Current efforts are underway to convert the simulation model to a basic decision support system for colleges and universities in evaluating computer networking trade offs. I should add that the current version of the simulation model is performing both more effectively and more efficiently than most of us expected and we are encouraged to hope that a significant decision support system might ultimately result from this research activity.

The Planning Council has initiated a project directed toward user-oriented network services and dealing with the problem of delivering computer resources to computer users at a substantial distance. One way to characterize the user network services envisioned by the Planning Council is to compare it to the ARPANET. The typical ARPANET user (if there is such a thing) was and continues to be a person skilled in the use of computer resources. The researchers who use the ARPANET and do not fit these characteristics usually require the services of an "interpreter" who is so skilled. As a result, the ARPANET users require little in the way of facilitating services and, in fact, seem quite pleased to be able to "do their own thing." The facilitating services which characterize the ARPANET are those which deal with hardware and system software interfaces, e.g., the ANTS system at the University of Illinois, the ELF system at the University of California and the UNIX system originated at the Bill Laboratories. By contrast, the user-oriented network services which the Planning Council is attempting to develop will facilitate the access to remote computer resources by users not skilled in computer science and technology.

The enhancement of discipline oriented Communities of scholars may be one of the most beneficial aspects of computer networking by enabling the concept of computer-based publishing of research and teaching materials. There are already substantial computer-based materials which are difficult to transfer from computer to computer. One example is large models which are used in chemical engineering, economics, computer science and a number of disciplines. Another example are large collections of data which are updated with sufficient frequency to make it inconvenient if not very difficult to transport from computer to computer. In fact, for much of this material, institutions can simply not justify the cost and the effort required to move the materials to the local computer so that effective access for any purpose is denied.

A particular illustration of this kind of material which EDUCOM continues to support is the distribution of materials in law teaching by means of a computer network. The materials are represented by computer-based exercises of two to four hours in duration which test students on their knowledge of various fields of law. The exercises basically simulates a one-on-one interaction between a student and a professor of law. A number of these exercises have already been developed and operate with the support of a large scale computer system at the University of Minnesota. Authors of the exercises are located at Minnesota as well as other institutions including Harvard. The cost of moving the whole system from computer to computer is significant. In addition, frequent modifications and improvements to the system at the University of Minnesota would have to be laboriously communicated and installed. The net result is that with lowering communications costs it is more cost effective to directly access the computer at the University of Minnesota than to deal with the transfer. These law exercises may be typical of the future in computer-based publishing:

they are an adjunct to course material as opposed to replacing specific techniques, or materials; they are produced in concert with printed cases which the law school students study and from which lectures are developed. As a result, they provide a new and valuable dimension to law teaching and have been received with enthusiastic support by a sizable number of law teachers. In fact, it is likely that these exercises will be used nationally by forty or fifty law schools within the next two years. The broad use of these materials is almost certain to encourage a larger group of authors to develop cases in the new computer-based medium. If this is a valid example of the concept of computer-based publishing, it appears to be self sustaining in the richest sense. I am personally hopeful that this model can be extended to a host of other disciplines for which remote access to computer-based materials will provide a new and valuable dimension to learning.

An Important Potential—The Library. Probably the most important and the most critical new function of a computer network for the future is as an alternative information resource. Although the use of networks as information resources has long been tested, it is now time for some tangible and useful manifestation of the promise.

In considering the problem and the potential, the relative economics of the status quo and technological innovation are revealing. This perspective of the library problem is illustrated in The Economics of Academic Libraries where the results of a study of 58 libraries over the years from 1964 to 1969 are portrayed showing that during this period the operating cost for the 58 libraries, calculated as a unit cost per student, approximately doubled. Over this same period the wholesale price index remained relatively flat. In fact, the library costs *increased at a rate seven times* that of the wholesale price index. Similar data for computer and communications costs expressed as a function of a unit of processing and computation showed that during the same period these costs were dropping more sharply than library costs were increasing. In fact, computer and communications units costs *decreased by ten times* relative to the wholesale price index over the period.

Some specific illustrations may help focus the point. In a recently published *Fact Book* on the Yale libraries, Dr. Rutherford Rogers reveals some striking quantitative data. In 1975 the Yale library held 6,518,848 volumes excluding serials. These holdings were 15% greater than the holdings of 1970, measured similarly, and the holdings in 1970 were 28% greater than the holdings in 1960. The growth rate in serials during the period is approximated to be in excess of two times that for books represented by the statistics. The cost for Yale library operations during the year 1974-75 was $10,733,000, the third largest on the continent, exceeded by the Harvard libraries at $11,496,000 and the Toronto University libraries at $11,997,000. One of the more distressing aspects of these cost data, in addition to the size of the numbers, is the rate of growth. For example, the Harvard library costs have compounded at a rate of approximately 9% for the past five years.

One reason that it isn't higher is that the funds devoted to acquisitions are now less than 30% of the library budget. If funds for acquisitions are squeezed to less and less as both the cost and quantity of printed materials increase, the result is both obvious and undesirable.

A costly feature of printed material is the physical space that it consumes. For example, at Yale there are 190 miles of linear shelf space in the University libraries housed in over 15 1/2 acres of floor space. In today's economic environment space costs are aggravated both by the increased costs of construction and the increased cost of operation and maintenance spurred by the energy crisis. This problem will likely continue for the coming decade as all these costs, particularly energy costs, move upward. The impact of space costs on libraries was the subject of a recent article in the *Chronical for Higher Education* (July 12, 1976) by Robert Wilson in which the author used space costs to argue that a large cost savings can result from an increased use of microfilming technology coupled with the sharing of library collections by regional warehousing of less frequently used materials. However, the concept of decreased self-sufficiency for university libraries, on which the conservation of space argument depends, represents a responsiveness problem to the librarian as well as a limitation on the user's ability to browse. Moreover, the storage of materials on the basis of frequency of use, a concept which has become very familiar to those involved in computer technology, is a relatively new concept to the librarians and one which plans stress on the book publishing market evident from the testimony before CONTU. This is but another illustration of the fact that technological solutions may create behavioral and policy problems.

The purpose of combining these data is, of course, to support the claim that the promise of computer and communications technology has some financial basis. It is virtually a certainty that unless that promise of computer and communications technology begins to provide alternatives in access to information, there is likely to be a substantial shortfall in the availability of information to scholars, a matter of some particular concern to those libraries which support research activity.

Scientists have for a number of years envisioned the day when computer and communications technology can be used directly as a substitute for visiting the library to plow through stacks of printed materials. The typical scenario envisioned by such library futurists includes a "conversation" with a computer terminal about the kind of information desired and the terminal responding with samples of information from the computer system which could then be used to refine the inquiry to the desired point. Moreover, it is envisioned that this conversation can occur in a fraction of the time required for a visit to the library of printed materials and would provide more "up to the minute" information as well. Such a system could not replace materials but that kind of information resource might reduce serials costs and reduce the labor intensive parts of library searches. It is obviously a scenario in which there is a good deal of promise when one considers the alternatives. Making a collection of information accessible through new technology, approximating completeness and making the service affordable in cost are equally as significant as making the human interface to such a service comfortable to individual users.

I submit that this concept is not far fetched. In fact, such a system already exists in microcosm. The legal data system offered nationally by Mead Data Central to practicing lawyers, called LEXIS, contains complete opinions for cases in particular areas of law from federal cases and several states including New York, Michigan, Ohio and Texas. The user interface to LEXIS has been carefully designed with the objective of allowing a senior partner in a law firm to use the system directly without an operator/interpreter. The theory is that the conservation of user time in the use of such a service for researching case law material will be cost effective in comparison to traditional use of law clerks and law libraries. It is also theorized that such a system will be equally or more effective in finding all pertinent cases for a particular proceeding given the time constraints that usually apply. This application has several characteristics which make it feasible for contemporary technology. First, the amount of data, although large, is sufficiently small to be compatible with contemporary memory technology for fast response devices. Second, the application is sufficiently specialized that a user language vocabulary can be targeted directly at a manageable set of terminology. Third, user time is valued very highly, often in excess of $50.00 per hour. Hence, the time saved can be used to justify a relatively expensive service. The combination of these factors makes the service successful and as technology becomes more powerful and less expensive, the methodology can expand to some other professions and disciplines, perhaps the sciences. The use of LEXIS and its competitive system offered by West Publishing Company is mushrooming in law firms in the states for which case data is maintained and for federal cases throughout the United States. These services also appear to be achieving to some degree the important objective of having practicing lawyers use the computer terminal directly (albeit computer terminals which has been carefully engineered to encourage such use). In short, these law systems may be early illustrations of the kind of technology-based information resource which can deal with a significant part of the library problem.

There are, of course, a number of immediate difficulties. The largest of these is the cost of providing such a service for disciplines not restricted to small volumes of information and specialized vocabularies. The challenge, then, is to concentrate on those areas of development which can improve the capability for specific disciplines at a reasonable cost. For the future, we should assure that the results of research in information structure and information technology are applied to areas of need where capability and financial factors are a match. Applied research, the marriage of theoretical concepts to very real information problems, must be performed in order for the computer and communications technologies to have a substantial impact on the library problem which has been described. Early indications are that the theoretical concepts are sound when applied to real problems. However, the development of computer based information systems as a meaningful alternative to the library in the 1970's has a large number of gaps, gaps which must be filled in order to avoid a real shortfall in the traditional role of the library for college and university scholarship.

Chapter 2

by Peter A. Alsberg

Computer Networking Experience at the Center for Advanced Computation

1.0 INTRODUCTION

The Center for Advanced Computation (CAC) is a research unit in the Graduate College of the University of Illinois. The CAC was originally formed in 1970 to support the Illiac IV computer. In 1972 we were forced to rely upon the ARPANET for our computer services and computer operations.

2.0 THE BEGINNING – 1972

CAC research activities. In 1972 the Center for Advanced Computation was an ARPA-dominated computer science research institution specializing in the development of operating systems, compilers, statistical systems, graphic systems, data management, and large application programs. The CAC had one research contract from ARPA at a level of 1.4 million dollars per year. Of that money, $480,000 per year was spent for our own Burroughs B6700 computer and support staff.

ARPANET status. In 1972 the ARPANET protocols were in their infancy. The few protocols that did exist were all first version. The ARPANET itself was a radical and very ambitious step into a new concept of computer communication and service. No one had any experience with multi-machine networking. There were very few server sites on the ARPANET. The only sites that could be depended on for a service orientation were the UCLA 360/91 and the Bolt Beranek and Newman (BBN) PDP-10. The Center for Advanced Computation was pioneering the mini-host concept for network access with its ANTS,† system. But at that time the mini-host concept was new and not well funded.

ARPANET transition. ARPA withdrew support from our computer and forced us to go onto the ARPANET for all of our computing needs.

The transition to the ARPA network provided for no overlap in local B6700 services. On June 30, 1972, the computer was removed. On July 1, 1972, Center research programs were transferred cold turkey to the ARPA network.

Arrangements were made with UCSD to provide B6700 service on the ARPANET. There were extreme problems with staff rejection of the ARPANET..

Immediate results. Within two months of the ARPANET transition, the staff attitude had completely reversed. There was virtually unanimous support for networking. The reliability, availability, and production orientation of the UCSD installation was superior to the facility we used to run ourselves. We were the largest single account at that installation. Our dollars were important to UCSD. In the fall of 1972 a B6700 was available on our own campus at the Department of Civil Engineering. UCSD knew this, and aggressively supported our requirements. We also saved money. During that first year, services at UCSD cost approximately $108,000. Support of our ANTS mini-host, local staff, and ARPANET IMP cost approximately $82,000. This annual cost of $190,000 was approximately 40% of our previous year costs.

3.0 THE MIDDLE – 1973 to 1975

Growth of heterogeneous computing. During the 1973 to 1975 period more service sites came onto the ARPANET.

There were many PDP-10's on the ARPA network, and ARPA had a well-funded graphics program for ARPANET PDP-10's. This provided the opportunity to exploit multiple machines on the ARPANET to solva a single problem. One of the first examples was in support of the Laboratory for Atmospheric Research. The PDP-10's on the network were superior interactive machines to the 360/91. Thus a PDP-10 was used in an interactive mode to prepare jobs which were subsequently sent to the 360/91. The results from the 360/91 were sent to another PDP-10 where graphic displays were prepared. Finally, the graphic results of the 360/91 run were sent to our ANTS mini-host using network graphics protocol. ANTS was then able to display the results on a variety of devices at the CAC and at the Laboratory for Atmospheric Research.

By 1974, it seemed clear that multi-machine operation offered not only performance but also cost advantages. To test this, we ran a series of benchmarks in late 1974 and early 1975. We found that machines varied by factors of 10 or more in the cost and the speed with which they could perform various tasks. Nevertheless, it was found that through networking we could exploit small differences in machine performance or pricing policy. Furthermore, network overhead could be absorbed even for very small tasks (tasks that cost only pennies to run or took only seconds to execute). The conclusion was obvious. Significant price and performance advantages were available if instead of attempting to solve an entire problem on a single machine, applications programs were split to take advantage of the best machine available for each component of a problem.

Cost growth. During this period, our cost for ARPANET services grew. We were doing much more computing than before. We were spending approximately $300,000 per year for service at remote sites plus $85,000 to access the network (ANTS, support staff, etc.). But, as competitive service sites became available, we were able to barter for preferred rates. We had increased our overall computing load several-fold and still were spending $100,000 less than when we tried to operate our own machine.

CAC transition. With the ARPANET, inter-institutional co-operation became common place. The researchers at the Center began to use the ARPANET mail facility to collaborate with other investigators. One result has been the frequent appearance of multi-institutional publications. Another has been that researchers who are dependent on the computer found that they could move from one institution to another and still have access to the same computing facilities. The CAC became a natural focal point for researchers who required sophisticated computer tools. By 1975, sixty percent of our contract support was now in disciplines that used computers only as research tools.

4.0 THE PRESENT – 1976

Current CAC research activities. The CAC currently has approximately one and one-half million dollars of contract research in energy, labor and economic impact studies, remote sensing, and land use planning. We still do computer science research for government and industry.

This research mix is likely to continue. Multi-machine expertise has become the rule for CAC staff. This expertise and the impressive resource of the ARPANET can place computing problems into a perspective more appropriate for these researchers.

Current cost. At present we are spending $300,000 per year for network computing services. Of the ARPANET services consumed by the Center contacts, approximately 30% are related to computer science and the hard sciences; approximately 40% are related to government policy support; and approximately 30% are related to energy and psychology research.

5.0 TECHNOLOGICAL CONCERNS

Resource access networks. At present, computer networks support resource access. A terminal attached to a computer in London or California can access a computer in Illinois as if it were local. A user can transfer files from one machine to another. All of the communications protocols are telephone surrogates. They provide point-to-point services in a manner that could be done with a telephone were a communications network not available. This mode of operation does not even begin to realize the potential of networking.

Resource sharing networks. Resource sharing networks are cost and performance superior to resource access networks. Resource sharing networks will require multi-point computer-to-computer protocols. So far, little has been done to develop those protocols. In addition there are major technological problems with resiliency, security, and accounting. Again, little has been done.

Resiliency. Today, most network systems assume a very benign environment. They do not account for protocol violations on the part of the computer. Protocols assume that no process on the network operates with malicious intent. Finally, current protocols ignore the nagging problem of lost messages. All of these assumptions are violated daily on the ARPANET.

Security. Access control and security on a computer network are extremely difficult. Current work in the Department of Defense indicates that the problem is exacerbated by multi-point considerations. Resource sharing networks will have to secure proprietary software and data bases and will have to support privacy for records about individuals and organizations. Once again, little has been done.

Accounting. For resource sharing to become a reality, we will need transitive billing procedures (and hence transitive security). For example, a user might interact directly with only one service site while interacting indirectly with many others. The first site should be able to bill the user for all services consumed, direct and indirect. The current requirement for separate accounts at each site limits resource sharing to the most innovative and technically agressive users.

6.0 LESSONS WE LEARNED

The hidden costs of our own computer. Prior to moving to the ARPANET, the "right" machine for any computer project at the Center for Advanced Computation had to be our own machine. Our machine was a fixed cost item. A significant amount of senior management time was drained by the care and feeding of the machine. Even our choice of research areas was subjected to constraints imposed by operating our own machine.

Networking advantages. In substituting the ARPANET for our own machine, we reduced our direct and indirect expenses. The ARPANET allowed us to spread our computing needs across the network, using an appropriate machine for each of our problems. It is common for Center staff to be fluent in three or more computer systems and to apply the different perspectives and philosophies of those systems in their day-to-day problem solving activities.

The drain on our senior management time is gone. We have replaced a major fixed cost item with a flexible cost item. We also have reserve computing capacity whenever needed.

Networking really isn't for everyone. Networking is not a free ride. The Center's experience has been that we cannot trust the common carriers to develop networking technology. Parochialism abounds. In particular the common carriers tend to design for terminal communications or for a specific operating system. Any organization seriously interested in networking must make the commitment to understanding and advancing network technology. That commitment is the hidden cost of networking.

Chapter 3

by Anne W. Branscomb

The Future of Computerized Information Systems: The Legal, Political, and Regulatory Environment

1.0 INTRODUCTION

What is to be the nature of a university in an information oriented society? Computerized information systems have the capacity to precipitate substantial changes not only in the manner in which information is disseminated but fundamental changes in the institutions within which such information is generated, collated, packaged and distributed to the consuming public.

The future evolution of networking of computerized information systems is more dependent upon issues involving information content and the roles of the institutions that develop them than it is on the cost and complexity of the communications and computing facilities themselves.

What then are the legal and regulatory problems you are likely to encounter, and what is the political environment in which they will have to be resolved? More precisely, what likely opportunities may arise which might facilitate and simplify the exchange of chit chat from one computer to another and who are the cast of characters you are likely to encounter as you enter the arena in which the legal parameters of networking are likely to be forged?[1]

2.0 THE PROTECTION OF PROPRIETARY RIGHTS IN COMPUTERIZED INFORMATION SYSTEMS

Computerized information systems transmit intellectual products that are produced by a wide variety of individuals, groups, and institutions both public and private whether operating in a university environment or in a business environment. One of the most striking characteristics of the new computerized information systems is the difficulty of separating public from private sources. The rules of behavior to which we have subjected the two different types of producers, and which differ substantially in their basic thrust, are now incompatibly mingled. Moreover, the new and long awaited Copyright Act recently passed by the Congress (1) does not really address itself to the questions of proprietary protection of intellectual property produced within the context of computer networking (2).

Consequently, a long and protracted development of the law in this area is a likely prospect for the foreseeable future. Although some of the most pressing problems offered by new technological developments such as cable television (3) and reprography (4) have been addressed, the Copyright Act has set aside the determination of the law concerning computerized information systems. Specifically, the Congress awaits the reports of two national commissions which are functioning independently and simultaneously: The Commission on New Technological Uses (5) which contains representatives of the publishers and the information industries; and the Commission on Libraries and Information Science which contains librarians and information professionals (6).

The results of these deliberations will have a profound effect upon efforts to interconnect data banks. The use of copyright protection of proprietary interests in computerized information systems raises far more questions and difficulties than it solves. The pouring of a new technology into legal concepts specifically designed for an old technology may alter the course of history in an undesirable manner than had the legal concepts been permitted to develop in an otherwise unstructured fashion.

However, the copyright act, together with trade secrets and unfair competition law are the only tools that the producers, packagers, and procurers of information systems have. Unless and until we can forge something more appropriate to the technology, we will have to live with them.

There are many ludicrous consequences. Copyright law was designed to protect authors whose only compensation came from the sale of the manuscripts to either the publishers, or, through them, from royalties on the sale of copies of the published manuscript. The U.S. Copyright Office requires two copies of the original publication to be filed with the office (this has already been abandoned when trying to deal with computer programs). But query: what is the original in a computerized information system such as the New York Times service which is constantly being updated every twenty-four hour period? What is the copy when only a small portion of the "original" is being used? Is this fair use which is uncompensable when only a small portion is being used or should one expect to pay according to use?

There are pluses and minuses when one starts to deal with access to a centralized data bank through a computer network. The ability to charge by use rather than by published copy is vastly improved, since the machine is capable of doing its own monitoring and billing according to actual number of times of use and even by categories of use. However, the problem is vastly more complicated when one uses local small processors to retrieve portions of information from a data bank, shares them with a number of users, and then returns them to the central facility in the same or a different format.

Certainly the concept of payment according to the number of copies made becomes even more complex in the case of a group product in which contributors on widely dispersed campuses participate in the design of an end product. To add another parameter, what happens when the intellectual product is the work of both university professors and industrial research professionals? To make the matter even more complicated, suppose the computerized information system is marketed by a non-profit entity that uses some material produced by a for-profit company? Does one apply a non-profit rate or a for-profit rate in compensating? Do the rules for transmission time on a public switched network apply at the for-profit or non-profit rate?

This is just a sample of the obstacles which may inhibit computerized networking from developing rapidly in the university world. Please bear with your university attorneys if they seem slow in resolving difficulties, and please cooperate with them to educate them to a full understanding of the technological facts of life in order that they may make your work easier and more productive in the long run.

My predicition is that compensation of proprietary interests in computerized information systems will develop along some concept of compensable use rather than the more established royalty per copy. Ultimately all professional journals, for example, may exist only as a single original data file from which all users will have direct access.

3.0 PROBLEMS OF PRIVACY

One of the major inhibitions to the development of computer networks has been the Big Brother image which was projected by Orwell's 1984. Since the year is almost upon us the fear seems to loom larger and larger. Although the legal framework in which the individual rights and stored in data banks can be as easily developed to work with computers as with other forms of technology, it is a pity that we are plagued with this continuing problem. It is a factor with which we have to live and which we will have to work together as lawyers and technicians to solve.

It is my projection that the public simply will not accept the proliferation of interconnected computerized information systems unless and until a body of law develops or a regulatory system evolves in which individuals and institutions may protect and preclude access to information obtained from and about them. Thus, we must: 1) recognize that information is a commodity* which has economic value and for which reasonable compensation must be paid; 2) Information of a statistical nature which must be collected for the efficient and necessary functions of government and business, i.e. census data, or social economic and financial statistics, must carry the protection from individual identifiability; and 3) Information about individuals must carry rights of release and powers of correction for inaccuracy. The context in which these problems will arise in the universities is in the collection and use of research data as well as storage and removal of student records or medical records in university hospitals. Networks which attempt to exchange information of this kind will run into statutory restrictions which will require revision to accommodate to the changed technological environment.

The responsibility for working out the details of legal access cannot be left to chance but must be developed along with the technological and economic solutions. It is an expensive waste of time and resources to develop facilities which the public will not use. The public will not use computer data banks unless they have confidence in the regulatory system, and they will not allocate funds unless they see some utility. The new commission set up under the Privacy Act of 1974 (7) is a good start in the right direction for federal systems. However, the protocols of use and protection of both privacy and proprietary interest in the product must proceed in a reasonably accelerated manner in order to keep pace with technological progress.

4.0 COMPETITION VERSUS MONOPOLY IN PRIVATE LINE & DATA COMMUNICATION

There is some good news and some bad news about the politics and economics of data and private line communications which will help or

* As compared with copyright protection which covers packaging and procuring of information.

hinder the development of interconnected computer networks. The Federal Communications Commission has recently come out with a decision in docket no. 20097 (8) which will permit the sharing of private line services, and Telpak, Digital Data Service (DDS). The decision is 82 pages, but, unlike previous decisions of this agency, it is articulate, well reasoned, and fascinating reading. This new development permits opportunities for sharing of resources which will not only cut costs of high speed better quality transmission facilities for interconnecting data banks, but will also permit a greater participation in the network design and control of the hardware which will route and process the messages. This decision is consistent with previous decisions of the FCC (9) over the last few years to open up competition in the longlines common carrier service (as well as the peripheral equipment and services), and should permit greater choice and freedom of action.

The decision encourages the use of an unregulated intermediary entity to manage the shared use. Not only will sharing of facilities by customers with an affinity of interests be encouraged as permitted by the Bell system for the airlines, electric utilities and stock exchanges, but all customers will be free to determine which of them wish to amalgamate their interests for the purpose of economy or quality of transmission. In the future both sharing, brokering, and resale of communications capability will be permitted without restrictions dictated by the Bell system and its affiliated carriers.

On the other hand the Bell system and its affiliated carriers have determined to put up a fight to extend the monopoly characteristics of the telephone system to all longlines carrier systems and to the peripheral equipment and services as well. AT&T and its associates have drafted the Consumer Communications Reform Act of 1976 (10) which Ralph Nader calls a violation of the truth in packaging act. There are more than 200 sponsors in the Congress which expires at the end of this year, and no further action can be taken before the new Congress takes office in January. However, the number of sponsors indicates both the strength of the telephone companies and the appeal of their argument that competition in data communications and attached terminal services severely handicaps their capability to provide low cost local loop service to residential customers.

Although the FCC has rejected the AT&T argument that the entry of competition on the longlines and data communications services will ultimately cause the deterioration of carrying capacity and the disintegration of the nationally integrated and remarkably efficient message toll service (11), there is substantial support for the Bell system's argument and general recognition that the Bell system has produced the best telephone system in the world. Economists are still at odds with each other concerning the facts,* and thus have not yet definitively determined whether there is sufficient evidence to prove AT&T right or wrong in its allegations. The FCC determination is an interim finding which cannot be appealed to the courts (12), and AT&T management has turned to the Congress for support.

In addition, the FCC has itself instituted a new computer inquiry (13) the outcome of which may also change the course of recent history. Under the consent decree of 1956 (14) the Bell system was prohibited from engaging in any activity which did not qualify as common carriage under Title II of the Communications Act of 1934 (15). Under the decision of an earlier computer inquiry (16), "data processing" and hybrid data processors were determined to be outside the reach of Title II which would be construed to cover only communications carriage and "hybrid communications systems" whose major activity was transmission rather than processing. However, the line between the two types of hybrids has become more and more diffuse with practice (17). Although many types of communications which border at the line of distinction are very similar in practice, some seek authorization under the FCC authority and some do not. Consequently, the pressure for clarification has come not solely from the Bell system which would like to be allowed to compete with the data processors but can do so only if such activity becomes subject to FCC regulation.

The lines are being drawn for a very substantial fight in two very distinct arenas: one in Congress and the other in the FCC rulemaking proceeding. Both are likely to have substantial effect upon EDUCOM and the efforts of universities to interconnect their computer facilities. Unless you wish to let others decide for you what is in your own self interest, you too must decide where your interest lies and get into the fight.

There is substantial evidence that a real reform of the Communications Act of 1934, which is long overdue, will in fact begin in earnest within the foreseeable future. The VanDeerlin Committee has made a commitment, the new president elect has made a public statement to the effect that he favors a new act, and the departure of Senators Pastore and Hartke opens wide the appearance of a new cast of characters in the Senate side as well. Therefore, you should do your homework and decide how you wish to see the Act fashioned to protect and promote your own interests in computer networking. Determination of where your interest lies is not a simple matter. Some

* Bell representatives argue that the longlines service subsidizes the local residential service and other economists argue that the local residential service has been subsidizing the longlines and business service.

nonprofit educational institutions receive discounts for private line service from carriers and some receive discounts for equipment from manufacturers. This discount policy may be a strong incentive to favor the natural monopoly approach rather than rely upon the natural forces of the competitive market place to drive the prices down especially if it can be ascertained with any definity that such downward direction in one service will drive the prices up in a related but socially productive service. What happens in the political arena in the next few years will have a profound effect upon the architectural design of our national communications system for at least the next quarter century.

5.0 SATELLITES AND SAUCERS

There are several developments in satellite interconnection with which educators ought to be familiar and which present novel and unusual opportunities. One is the filing of the CPB on behalf of the PBS for a satellite system which would interconnect public broadcasting stations through the country, both radio and television (19). This application is being filed with an expectation that there will be excess capacity available on the system as designed, and that such excess capacity will be leased back to a third party (20).

The public broadcasting system, as presently constituted, receives substantial government subsidies at the national, state, and local level. Indeed, many of the licensees are the same universities which are interested in networking computer facilities and belong to the EDUCOM Planning Council (21). It seems logical that some marriage of interests might be consummated. Joint efforts between the computer facilities and broadcasting entities might permit these non-profit users to combine their resources to facilitate the networking of computers and to obtain increased transmit capability for the public radio and tv stations (most of which will be left otherwise with receive only satellite saucers). Although the CPB has shown an extreme reluctance to clutter up its lead application with any specifically contemplated shared uses which might delay realization of their system; nonetheless, the public interest dictates that the ultimate costs to the consumer be considered. An opportunity which now presents itself should be seriously studied by universities who have both computer facilities and public radio or television stations.

Another contest in which the interconnection of computer facilities by satellite may be pursued is the Public Service Satellite Consortium (22), which operates under a mandate from its membership to catalog the potential public use of an interconnected satellite system nationwide for non-profit users and to pursue some jointly determined strategy for obtaining satellite interconnected capability. Such a national capability for public service users would permit costs insensitive to distance and permit small users the advantage of piggy-backing onto the use by larger institutions.

Another filing with the FCC for satellite interconnection was made by Satellite Business Systems (SBS), a partnership of IBM, Aetna, and Comsat (23). SBS intends to use a high power satellite system with the 12-14 gigahertz frequencies that can be received using smaller and less expensive saucers than can be used with the Domsat satellite carriers with low power like Westar. The SBS plan is especially designed to accommodate the needs of both data processors and private line voice systems. It can also carry video signals and therefore is ideally suited to the needs of university computing facilities as well as business systems.

All three of these efforts will be pursued in the arena of the Federal Communications Commission's licensing authority. They will pass or fall by the wayside depending upon the arguments pro and con which are put forward by the interested parties, including the existing common carriers which provide existing longlines service and potential users. Thus the entry of new competitors to provide alternative interconnection service will depend in part on who is interested in helping whom to do what.

6.0 CABLE TELEVISION AND BROADBAND LOCAL LOOPS

Many of the library associations have been especially interested in the development of cable television throughout the country. There is now a renewed interest in cable television regulation within the subcommittee on Communications of the House of Representatives. A recent report suggests that cable should be unleased to provide the new capabilities which have been promised (24). Here again is an opportunity to choose sides and enter the fray. However, the development of cable television as an alternative interconnection for computer networks is much more complicated than the mere unleashing of the regulatory control by the Federal Communications Commission.

In the past few years, the FCC has diligently tried to regulate into existence broadband public channels for use by educational institutions as well as municipal and public channels. This effort has been notably unsuccessful for a number of reasons which are not entirely the fault of the regulatory policies of the FCC. First, the general economic condition of the country has not been one in which a capital intensive industry like cable television can flourish. Second, the industry has shown little or no interest in the development of the broadband networks which would service institutional users such as universities and libraries, and it has looked upon the requirement of a public channel more as a burden than as an opportunity. Finally, the users for which the channel was set aside, the educational institutions, have been remarkably diffident about putting forward any organized effort to make use of such channels even where they were available.

The interconnection facilities for such data or video networks can be provided as easily; or perhaps more easily, by telephone companies. By regulation telephone companies are allowed to provide new business services althouth they are prohibited from providing cable television service in the local company's area of service. Thus telephone companies are not likely to look kindly upon the existence of computer networking services provided by competing cable television systems without offering private line service to such institutions as are willing to pay for them at a competitive or possibly predatory rate. Certainly the bankruptcy of DATRAN is not promising for the future viability of institutions which attempt to compete directly with the AT&T facilities; and there are substantial reasons for determining that the local loop is far more characteristic of a natural monopoly than the long lines.

Therefore, the so called freeing of cable television to develop undeterred by FCC regulation may become a hollow victory. There may be more mileage in pursuing alternative routes of public investment in broadband local loops for all kinds of service (25), or some form of leverage on the Bell system and its affiliates to install optical fibres or other forms of broadband capability in the local loops as this becomes economically feasible as Professor Vivian of the University of Michigan has suggested earlier this year (26). Finally, the judicial and regulatory restraints might be removed which inhibit Ma Bell, and the public might rely upon a truly competitive marketplace.

7.0 ELECTRONIC FUNDS TRANSFER AND ELECTRONIC MAIL

Another area of interest is electronic funds transfer, electronic mail, and business systems. As usual the real impetus for reform of the system is coming from those who have the greatest economic interests—the banks, the credit institutions, and the retail merchandisers. Consequently, it is my prediction that the pressure upon the legal system will continue to come largely from these institutions where the sale of information as a commodity has high economic value. A recent decision by the Supreme Court to the effect that terminals in retail establishments constituted branch banks (27) (thus restricting their development across state lines and within states which prohibit branch banking) will have a substantial inhibiting effect upon the development of EFTS networks interconnecting banks and retail merchandisers.

This may seem remote from the self interest of educational institutions and libraries, but it is not. More than half of first class mail is directly related to the exchange of information about financial transactions. If this bulk of mail, which pays its own costs, is siphoned off by EFTS networks, the loss will have a substantial second and third degree effect on libraries and educational institutions which use the mail for exchange of books and magazines at a highly subsidized rate (or with a substantial public investment to use a more positive phrase). Therefore, the capability of the U.S. mail system of exchange of continuing to carry this burden will quickly be questioned.

What is happening with legal definitions of branch banks may have some comparable legal characteristics when applied to institutions which have a more direct impact upon your higher education self interest. Query: is a library which transmits an instantaneous copy of a manuscript by facsimile a branch post office? If the postal service had as much foresight or political clout as the Bell system, the postal service's lawyer might easily argue that case and might draft a Consumer Electronic Mail Reform Act of 1977, 8, or 9. Indeed, in economic terms alone, the transmission of copies by high speed facsimile will certainly diminish the use of the mails for postal delivery of hard copies, and thus diminish the economic base upon which the postal service even now so precariously relies. Colleges and universities will want to take a careful look at what is happening in both electronic funds transfer and in the postal service to determine where their self interest, and that of the public, really lies.

The banks and the telephone companies have the resources to amplify their message and self interest, but the legal frontiers which are just now opening up will also make substantial systemic changes in the modes in which educational institutions and libraries have been operating.

8.0 THE ROLE OF THE UNIVERSITY IN AN INFORMATION ORIENTED SOCIETY

Many scholars have observed that the U.S. has nearly passed the 50% mark wherein the economy has become dependent upon information exchange as its major source of strength (28). Thus, in the future, the transfer of Information is likely to be treated as a commodity on the open market for which a direct value is assessed and collected rather than as a public good which is deposited in public institutions and made available freely without direct cost to the consumer. What is this likely to do to the universities and libraries which have traditionally relied upon grants, subsidies, and appropriations that were only remotely related to direct use? It is very likely to cause some real soul searching and dislocation of personnel and economic resources. Universities and their libraries will have to accommodate themselves to the new electronic and economic environment in which they will be operating.

To illustrate the universities' dilemma, consider the silk stocking industry which perceived its mission as the production of silk stockings rather than the covering of legs or, more broadly, as the manufacture of clothing. Universities face much the same opportunity and hazard today in the context of the opportunities presented by computerized information networks and the hazards of defining one's mission too narrowly. Already there are substantial pressures coming from other sources than computer networking which question the viability institutions of higher education as they currently exist.

The ivy covered walls of academe into which students disappear for a four year stint or more, to be tutored by scholars of national or international renown, is being challenged by a student body which does not choose to be isolated from the real world for such a long period and by an economic environment in which fewer and fewer students and their families can afford to pay for such a sheltered and personal educational experience.

Therefore both the preferences of the students for a different life style, as well as the drive of the economics of education delivery, dictate a more flexible environment in which the educational process can thrive. Education will be looked upon more and more as an ongoing experience which takes place throughout one's life at times and places of one's own choosing.

Computerized information networks are capable of providing educational information at the fingertips of the users wherever and whenever users can obtain terminals to access the system or systems. How such access is to be funded presents a dilemma which must be resolved. Who is to package the programs that make available the accumulated wisdom of mankind is the $64,000 question (more likely, with inflation, the 64 billion dollar question) which the current generation of computer managers, computer programmers, librarians, and educational administrators must resolve.

Traditionally we have paid professors modest salaries to produce the information upon which the world turns. They were supposed to receive their rewards in heaven or elsewhere depending upon their religious preferences. The amount of direct compensation for the intellectual product of universities has been woefully small, and textbook costs reflect primarily the price of publishing and marketing the manuscripts. With computerized information systems the cost of input and output must be recaptured by someone, but the monitoring capability of the computer system also presents an opportunity to recapture the cost of originally collecting the data. This opportunity was not available to the publishing industry, nor were the educational institutions under such economic stress to recapture this cost.

Libraries which have traditionally operated in a mode of depository and free use will find increasing opportunities for pricing their product by use as well as substantial economic and political pressures to do so. Already there is substantial drive in the direction of recapturing the costs of research and development, the bulk of which is funded directly or indirectly by the federal government, and there is substantial pressure from the public in the direction of lowering the tax burden upon citizens. Thus the taxpayers themselves, as they become more sophisticated and organized, are rebelling against a system which subsidizes too many costs indirectly with accounting systems which measure neither economic nor social value. The universities and their libraries are going to be faced with a mounting pressure to assess and recapture direct costs of computerized information systems.

Moreover, universities and their libraries are going to be faced with an agonizing reappraisal of their function in society. They will have to decide whether or not they wish to be the production centers for the collection and packaging and programming of instructional and informational systems. Also they must decide whether they wish to charge non-profit rates, a rate which reflects funds available for reinvestment in new systems, or turn to the public coffers to seek subsidies for the distribution of their intellectual product. How this appraisal comes out will, in large measure, dictate the course of the future development of the networks for computerized information exchange. The dividing line between what is for profit and what is not-for profit is becoming more and more difficult to draw.

The challenges are great, the opportunities are almost unlimited, and the hazards of failure are considerable in terms of social cost to society. University computer centers cannot afford to sit by and wait for others more powerful to fight for them. The legal and regulatory environment in which universities will find themselves will be fashioned to fit the needs of the more powerful combatants if the universities do not fashion the environment to meet their own vision of the needs of university networks in the discharge of their social responsibilities. During the next 12 to 48 months you will have an opportunity to participate in a great debate about the legal and statutory structure of a national communications system. Those of you who are at the leading edge of the technological frontier must not let the decisions be made by others by default. You have a responsibility to yourselves and to your fellow citizens to help fashion the future.

References

1. P.O. 94-553, October 19, 1976
2. Indeed, Section 117 specifically leaves the law as it is prior to passage of the new act.
3. Section 111
4. Section 108
5. P.L. 93-573, December 31, 1974, Title 17, U. S. C. Section 201
6. P.L. 93-345, July 20, 1969
7. P.L. 93-579, December 31, 1974
8. FCC-76-641, 41492, July 16, 1976
9. For example: Carterfone, 13 FCC 2d 420 (1968); MCI, 18 FCC 2d 953; Specialized Common Carriers, 24 FCC 2d 318; and Domestic Satellites, 35 FCC 2d 844 (1972)
10. H.R. 12323, 94th Cong. 2d Sess.
11. Docket no. 20003 FCC-76-879, 42463, September 23, 1976
12. This is an unfinished proceeding in which no final order has been entered.
13. Docket no. 20828, FCC 76-745-41799, August 9, 1976
14. Consent decree, 13 P&F Radio Reg. 2143 (1956)
15. 47 U. S. C. Section 201 f.
16. Computer Inquiry; 7 FCC 2d 11 (1966); 28 FCC 2d 291 (1970); 28 FCC 2d 267 (1971) rec. den. 34 FCC2d 557 (1972).
17. Berman, "Computer or Communications? 27 Fed. Comm. Bar J. 161 (1974).
18. Broadcasting magazine, October 11, 1976, p. 30.
19. FCC File No. 453-469, DSE-P-76, filed August 16, 1976
20. Agreement Between Corporation for Public Broadcasting and the Western Union Telegraph Company, p. 30-34.
21. Many of the major universities, both public and private, have joint radio and television facilities as well as computers. For example, EDUCOM member institutions in this category include: the University of Alaska; University of Arizona; University of California, San Diego; San Diego State University; University of Illinois; Indiana University; University of Mid-America; Ohio State University; and University of Wisconsin.
22. The PSSC is funded by contributions from potential users and a grant from HEW's Office of Telecommunications, and NASA.
23. FCC File No. 1, DSE-P-74, 166 and 167, DSE-76, filed December 22, 1975
24. Cable Television: Promise Versus Regulatory Performance; 94th Congress 2d Session (January 1976).
25. 25 Journal of Communication 55-56 (1975).
26. Wes Vivian, "Fibre Optics: Prospective Public Policy Issues" Fourth Annual Telecommunications Policy Research Conference, April 23, 1976
27. 45 U. S. L. W. 3238 (October 4, 1976).
28. The most definitive study shows 25% of the GNP accountable from the production, processing and distribution of information goods and services and another 21% accountable by information services performed by non-information institutions. Porat, *The Information Economy*, Center for Interdisciplinary Research, Stanford University, August 1976.

Chapter 4

by J. C. R. Licklider

Future Directions in Computer Networking Applications

1.0 INTRODUCTION

A note on the perspective of the author may be useful to the reader of this paper. Jim Miller interested me in EDUCOM at the time of its founding and gave me the opportunity of being an active participant in EDUCOM during its formative years. After about six years of active participation, I was away from EDUCOM about six years. Larry Roberts gave me the opportunity of succeeding him as Director of the Information Processing Techniques Office of the Advanced Research Projects Agency and thus of gaining, over a period of about two years, intimate and intensive experience with the ARPANET and other technological advances in computers and communications that Larry and his ARPA colleagues have inspired and supported. With the perspective that a year back at an EDUCOM member institution has given me, I want to incorporate some of the ideas derived from the ARPA experience into comments directed to EDUCOM.

2.0 TRANSITION

In her paper (Chapter 3), Anne Branscomb emphasized the criticality of the next few years. They are critical. Higher education is in a transition between two eras in the development and application of computing, and the shape of the second era will be determined in large part by decisions made during the transition.

During the first era, now ending, the crucial factor was the thrust of computer technology, and the main problems that had to be solved were essentially technological problems relating to the development of capability. During the second era, now beginning, the crucial factor will be the reaction of society. The main problems that will have to be solved will be essentially social problems relating to acceptance, assimilation, and impact of computing technology.

During the first era, the main impact of computers was to change the way industry and government carried out basic, established functions, such as payroll, inventory, and process control. Computers did not enter directly into the life-spaces of most people; they did not interface directly with many people who were not computer students or professionals. During the second era, computers will create new functions, such as: teleconferences, in which the participants do not have to meet in either space or time; computer-based markets for consumer goods that mediate between buyers and sellers in the manner of the stock market; and knowledge bases that go beyond storage and retrieval into the organization and application of the world's corpus of information. Computers will interface directly with as many people as can acquire the requisite informational skills and adapt to the technology whether it is cold and impersonal or warm, friendly, and ultrahumanly helpful.

During the first era, most computers were readily identifiable as computers. During the second era, most computers will either be built into systems (such as automobiles and heating-cooling systems) or devices (such as toasters and cameras) or integrated with communications to form networks.

The decisions made during the transition between the two eras will shape and constrain, to some extent irreversibly, the most important parts of the future of mankind: the parts concerned with communication, information, knowledge, and the intellect. There are two main roles for EDUCOM to play in the decision process. The first, urged by Anne Branscomb, is to represent the interests of the discoverers, organizers, preservers, and disseminators of knowledge to try to ensure that those interests are at least protected and, if possible, advanced. The second is to contribute to the structuring of the decision process, itself, by analyzing and clarifying possible courses of decision and action and predicting their impacts upon society and mankind.

In the following pages, some observations about the future of computer and communication technology are made with the hope of convincing the reader that the stakes are high and that the technology is willing and able if people who use it are willing and wise. Then three very significant issues are presented briefly. Finally, EDUCOM is urged to play as active a part in the national decision process that is shaping the future of networking as it is already playing in the fostering of networking to serve the needs and purposes of education and research.

3.0 THE FUTURE OF COMPUTER-COMMUNICATION TECHNOLOGY

Although an essential part of this message is that it is high time to turn most attention from the technology to its organizational, disciplinary, and social impact, a few respects should be paid to the goose that lays the golden eggs: the technology that provides the necessary basis for the informational revolution.

3.1 Cost-Effectiveness

The old rule of thumb about the cost-effectiveness of computer hardware still holds: overall computing capability per dollar continues to double approximately every two years. If that exponential course is not followed through the rest of the century, it will be because technologists did not try intelligently enough rather than because they encountered fundamental limits. Viewed microscopically, the advance of computer hardware technology is the progressive improvement of a specific sub-technology: increasing the reliability of vacuum tubes, decreasing the size and cost and increasing the speed of magnetic cores, increasing the density of the elements on LSI* chips, and so on. Viewed macroscopically, the advance is the replacement of one sub-technology by another that is inherently more cost-effective: vacuum tubes by individual transistors; individual transistors by integrated semi-conductor circuits; and cathode-ray memories (Williams tubes) by core memories by semiconductor memories, *et cetera*. The long-term future of computer hardware lies somewhere in a combination of new processing technology and new architecture for information processing systems. New processing technology includes: electron beams, x-rays, nonlinear optical phenomenon, three-dimensional structures, and storage in and processing by individual molecules. New architecture includes: parallelism, intermixed memory and processing, cellular automata, distributed systems. Policy toward basic research and exploratory development is therefore more crucial in determining the long-term future than is the more massive, but less far-sighted, drive of the marketplace.

3.2 Consoles

During the last few years, computer consoles (terminals) have decreased in price by about 25 per cent per year (less rapidly than processors and stores), and the console market measured in dollars grew

* Large-scale integration of digital processing and memory elements on chips of semiconducting material.

about 25 per cent per year (more rapidly than processors and stores). If that trend holds, and if there are now (as surely there must be) one hundred thousand consoles in the United States, there will be one console per person in the U.S.A. well before the end of the century. Lest that prospect seem both unlikely and bizarre, consider the trend in pocket calculators, of which there may be one per capita by 1984. Future consoles and pocket calculators (i.e., pocket computers) may be capable of very significant processing and storage, and many may be programmable. Whether or not those potentials are realized will depend on social and educational factors more than on technical factors. At present, the average person is not much interested in the things that can be done with either consoles or computers; the instruments are not now capable of fulfilling functions or delivering services in which the average person is much interested. But interests can change, or be changed, and functions and services can be developed. During the next few years people will decide whether the processing of information will progressively displace the possession of matter and the expenditure of energy (e.g., automobiles and driving) from the focus of our attention. If they decide to move toward information, there will be consoles and computers everywhere.*

3.3 Processors and Random-Access Memories

At present, there exists a gross disproportion in computer economics. It is as though, at a restaurant, the entreé were ten cents, the cover charge a dollar, each vegetable ten dollars, and a large meal ten thousand dollars. A microprocessor or a block of random-access memory is an LSI chip, and an LSI chip costs ten dollars. A minicomputer is a few chips in a box, and it costs a few hundred dollars. A minicomputer system is a minicomputer plus a few peripherals and costs a few thousand dollars. A number cruncher is the logical equivalent of a few hundred chips in a big box plus more, and more potent, peripherals; this costs a few million dollars. It is almost inconceivable that one can get so much function per unit cost on a single chip. Further, it is almost inconceivable that a real, working computer system can be as expensive as it is. The catches are: 1) that the function on a semiconductor chip is cheap on the chip but dear to get off the chip; and 2) that cabinets, disk drives, tape drives, paper tape readers and punches, and (despite the downward trend of cost) even the simplest consoles are much more expensive than are the basic semiconductor essentials.

---

* Most consoles will contain computers, and most computers will be designed to be connected into networks.

At first thought, it seems obvious that effort should be diverted from further development of LSI processors and memories to solve the problems of chip interconnection, subsystem packaging, and peripheral stores and consoles. Before taking that course, however, one should consider the possibility that some of those problems will vanish if chips can be made denser and denser, as they can be (down to some as-yet-not-understood limit) with electron-beam and x-ray fabrication techniques.

There appear to be two very important near-term opportunities: 1) to increase the density of LSI to the point at which a very capable minicomputer with many thousands of words of memory becomes a microcomputer, fitting onto a single chip and thus into a matchbox (which would eliminate or solve the chip-interconnection and subsystem-packaging problems); or 2) to learn to make many identical microcomputers work together as a single megacomputer without elaborate interconnections (perhaps by building a pocket-radio system into each microcomputer, putting many microcomputers into a shielded box, and achieving the required cooperative interaction among them through programming). In any event, small processors and memories are doing extremely well.

3.4 Less Than Fully Random Access Memories

It seems to be a principle of nature that large, inexpensive memories (or storage systems) must fall short of fully random access. Typically, such memories are organized hierarchically, providing slow random access to a block of storage cells and then fast sequential access to the successive cells of the block. Promising technologies for such memories include magnetic bubbles, electron-beam-accessed memories, and charge-coupled devices. One additional approach is also being explored by ARPA.

The plastic-disk TV-playback devices now being introduced into the home information-entertainment system market are truly remarkable from an information-storage point of view. One disk about the diameter of a 33-rpm phonograph record, and much thinner, holds 30 minutes of video-plus-audio program (analog information) which has a time-bandwidth product of about ten to the tenth power and corresponds (according to a conservative engineering rule-of-thumb) to about ten billion bits. Ten billion bits can be visualized as 300,000 typewritten pages of non-redundant text. The TV-playback devices, mass produced, provide a mechanical system capable of supporting 10-billion-bit selectivity for less than a thousand dollars. The audio plus video disks cost about one dollar apiece. The ARPA project hopes to achieve digital read-write capability using this technology. If successful technically and in the market, this new technology could cost as little as 10^{-5} cents bit for the device and 10^{-8} cents bit for the disks. Thus the entire device would be less expensive than the typing paper required to hold the alphanumeric contents of one disk.

3.5 Transmission

Of the three basic informational operations involved in networking—processing, memory, and transmission—transmission is progressing the least rapidly. Future technology of transmission will include satellites, cables, waveguides, fiber optics, electronic circuit switching, and other technological advances. In this paper, discussion is focused on the fact that the future developments in transmission capabilities of networks will be in the field of telecommunications, rather than the field of computing. Telecommunications and computing are very different fields. They are different with respect to age, style, rate of change, monopoly, and operation or regulation by government. The fact that networking necessarily involves these two fields accounts in considerable part for the ongoing transition from technological domination to sociopolitical domination.

However dissimilar, the two fields are strongly interdependent. The telecommunication switching function, whether circuit or message or packet switching, is essentially a computer function, and the only reasonable approaches to communications security involve digital processing and storage. Furthermore, if computers are to reach out to users, or if computers are to respond to sensors and control effectors, they must do so through communication channels. The two fields must interact. It will be interesting, to say the least, to watch the interaction.

3.6 Software

Networking places new demands upon the most complex and least well mastered of the computer arts. The demands are most stressful in nonhomogeneous networks in which the interconnected computers and their operating systems, data management systems, and programming languages are diverse and designed for independent operation. Processes have to cooperate with one another despite geographical separation, variable transmission delays, and occasional failures of computer and communication hardware. Users want to understand what is going on, especially what is going wrong, without learning all the technical details of several different computer systems. If no system consultant is available, the software may have to tutor the user in its use. It would be helpful if it could answer questions about itself—as a few research-level programs now can.

As networking advances, the amount of diversity within networks* will probably decrease. Operating systems and data management systems will be designed or adapted to support networking. The several computer systems in a network will tend to adopt one or a few common programming languages, which themselves will be augmented with networking primitives. In many cases, users will be screened from dealing with the network *qua* network: they will simply interact with their consoles, perhaps not even being aware that processors thousands of miles apart are cooperating in their behalf.

Underlying all the foregoing is the question of how homogeneity and coherence will be achieved. Will it be achieved through: a "bottom-up" evolutionary process** in which people are driven to standardization by the frustrations of the Tower of Babel; or through a planned, designed, optimized "top-down" process involving "leading standards" and requiring a combination of genius and luck to avoid global mistakes;*** or through some compromise approach?

The evolutionary approach has the advantage of not having to be planned. Already there are dozens of networks,**** highly heterogeneous and mainly independent of one another, serving banks, airlines, time-sharing companies, stock markets, communications, and research. Progress in standardization has been made mainly at the lowest levels where, without standardization, there can be no communication at all. If all the various companies, consortia, and countries that want or need networks build independent ones with nonstandard protocols, the "software problem"—which is even at present being viewed with much alarm—will rise in a great crescendo of complex confusion.

4.0 SECURITY, PRIVACY, AND CONTROLLED SHARING OF INFORMATION

Probably the least satisfactory area of networking technology is computer system and network security. Security is in a bad state largely because it is difficult for people to understand its crucial importance in the network picture. The topic does not have *a priori* appeal to administrators and managers; it sounds negative and secondary. It is easy to tell oneself that the first thing to do is to get the network to function; Later there will be time to worry about making it secure. But the fact is

* Within the lowest-level networks if there is a hierarchy of networks.
** Note that evolution has no way of planning ahead or working "top-down".
*** Because top-level decisions are made on a project-wide basis before anyone has gained experience by implementing and operating low-level subsystems and components.
**** But not yet dozens of packet-switching networks.

that lack of attention to security in the basic design of a network not only jeopardizes the informational privacy of individuals and the security of proprietary property but tends to preclude intentional sharing of information and may even kill the network.*

That computer system and network security is in a bad state is evidenced by the frequent reports in the computer and general press of invasion of privacy in data banks and of computer fraud. During the last several years the Information Processing Techniques Office of ARPA has supported a series of "tiger team" attacks on the security of selected computers, attacks designed to test the degree of security provided by operational systems. In no case has the team been unable to penetrate the security of the target system.

Some research and development work has been done and more is being done in the security area. IBM has for several years supported a program at a level of at least $5 million per year, but much of it is focused on improving the security of existing computer operating systems. Unfortunately, it is essentially impossible to render secure a complex system designed without much consideration for security. ARPA and the three armed services have been supporting smaller programs aimed, in part, at testing and improving existing systems but mainly at: 1) radically simplifying, clarifying, and then perfecting the mechanisms of access control in operating systems; 2) developing communications security techniques for packet-switching networks; and 3) creating formal models of access control within which it is possible to prove security in approximately the same way as, in a few simple instences, it has been possible to prove the correctness of a computer program. In 1975, the National Bureau of Standards proposed a standard encrypting algorithm, implementable on a single LSI chip, for use in commercial applications requiring digital communications or storage security. The federal intelligence agencies have also dealt with both computer and communications security and with the problems of ensuring it and compromising it.

Despite the efforts, network security is so poorly developed that any organization with valuable proprietary information to protect is strongly inclined to keep it insofar as possible under its own control. Banks, for example, are creating their own networks for electronic funds transfer (EFT) rather than trust a general or common network. However, the lines are of course leased from common carriers. IBM operates a network connecting its branch offices to its data processing headquarters, a network dealing with orders, system configurations, and

* Cf. Congressional action in the case of FEDNET as proposed in 1975 jointly by the Government Supply Agency and the Department of Agriculture.

progress through the sequence of sales, manufacturing, testing and delivery. Again the company's own network is constructed out of leased channels. Whether or not such networks are actually secure, they have a great psychological advantage over multi-organizational networks until the latter can demonstrate convincingly a very high degree of security. Thus the present inadequacies of computer system and network security are contributing to a proliferation of independent, diverse, and nonconnected networks. The proliferation of independent, diverse, and nonconnected networks will constrain the future of networking, precluding the economies of scale, the efficiencies of the Law of Large Numbers, and the advantages of intential sharing of information that could be enjoyed or provided by a coherent, common, general-purpose network.

5.0 ALTERNATIVE FUTURES IN NETWORKING

To describe the future of computer networking, it is useful to consider three alternative futures, or three very significant issues which cover separate effects of future networking. First, what kind of network will best serve the nation and society? Second, what new possibilities for communication among people are offered by future networks? Third, what impact will networking have on the organization and distribution of knowledge.

5.1 Future Kinds of Networks

What kind of network, or what kinds of networks, will best serve the nation, the society, people? What is the best approach to achieving it or them? This issue embraces: 1) the competition between circuit switching and packet switching; 2) the controversy over entrepreneurship, monopoly, and regulation; and 3) the question of over-all network configuration.

There appear to be three main kinds of network configuration:

- *A multiplicity of independent, heterogeneous, and largely non-interconnected networks.* Separate networks might exist for EFT, travel reservations, electronic messages, home entertainment-information systems, interlibrary loan, and so on.
- *A multiplicity of independently operated networks governed by uniform standards and protocols that facilitate interconnection through "gatekeeper" computers programmed to provide two-way access control.* For example, many business organizations might wish to tie their networks into the airlines' network for the reservation and ticketing function but to preclude other information flow between the networks.

- *A coherent "network of networks" featuring dynamic interconnection and controlled sharing* under a global security system and global privacy protocols.

Obviously, intensive study and wise decisions are required in the near future to determine which way to proceed. With respect to the first kind of configuration, it is necessary to weigh the difficulty of intentional sharing and the inefficiency due to lightness of load and redundancy of facilities that characterize independent networks against their security advantages, real or imagined. With respect to the second kind, it is necessary to explore intensively the problem of interconnecting variously dissimilar networks in order to find the necessary and sufficient conditions for internetting with specified limits on delay, inefficiency, and overhead. With respect to the third kind, it is necessary to determine whether or not a coherent, common, general-purpose network with extremely secure control of access and with thoroughly satisfactory guarantees of informational privacy is feasible. Overall, it is essential to evaluate the alternatives in the future of networking before the succession of events in the marketplace has narrowed the range too far. Possibly only poor alternatives could remain.

5.2 Future Communication Among People

Networking opens up exciting new possibilities for communication among people. The simplest of these, essentially a rechanneling of established forms and procedures, is electronic mail. Others include: teleconferencing (both concurrent and nonconcurrent); communication via executable programs or models (which ask questions and accept answers or display dynamic charts and figures); dialog with reference to shared information or knowledge bases (a two-person journey through a structured information space, with one person leading first, then the other, then perhaps the computer); and cooperative creative effort on the part of geographically separated team members (for example, distributed team programming).

There is at present considerable interest in electronic mail in the U.S. Post Office and in several of the computer companies. It is likely that one or more organizations will move aggressively to market a simple electronic message service in the near future. Would such action preclude, or seriously postpone, the development and use of the other modes of communication which would introduce new dimensions of human interactions instead of merely speeding up a mode that used to be reasonably fast and inexpensive but has fallen into a serious regression? Would a foray into networking bail the U.S. Post Office out or sink it? Is message service a proper natural monopoly, or is it a proper arena for pluralistic enterprise? Should message service stop with the delivery of the message, or should it extend into storage and retrieval and facilitation of reply? What about security and privacy? What about security classification (secrecy)? Finally who, if anyone, will speak for the private citizens who, and the very small businesses which, would have to underwrite the cost of paper mail if large organizations with computers that can afford satellite antennas were to desert paper mail for electronic mail?

5.3 The Future of the Body of Knowledge

The computer has introduced new ways of structuring and storing information that do not suffer the one-dimensional constraint of text or the two-dimensional constraint of diagrams and pictures. In the computer, a multiplicity of interrelations can be represented at the same time, and static forms can be translated into dynamic forms through the execution of simple programs. As yet, although computer-program modeling is on the rise, the new capabilities have had an amazingly limited effect on the way the bulk of human knowledge is structured, stored, and accessed. Evidently, it takes a long time for even the most potent of ideas to move from feasibility demonstration to widespread use.

During the remainder of this century, however, there is time and opportunity for a revolution in the organization as well as in the distribution of knowledge. Almost surely, dynamic models and knowledge bases will be prominent in the armamentarium of knowledge techniques, but the new knowledge technology is not yet clearly enough defined to try to sketch it here. It will grow, with some input from modeling and data management, mainly out of the field of artificial intelligence.

6.0 ROLES FOR EDUCOM

EDUCOM's most successful roles have been: 1) facilitation of the transfer of computer and communication technology into the service of education and research in its member institutions; and 2) representation of the interests of education and research (especially of universities and libraries) in the national decision process and in the marketplace. In connection with the first role, the network arrangements involving member institutions and TELENET, which could not have progressed as rapidly or as well without a consortium, are beginning to have significant practical and beneficial effects. In connection with the second role, recall that, a few years ago, EDUCOM almost singlehandedly convinced Congressional committees not to move forward with what would have been very unfortunate copyright legislation relating to computers and networks, and note with satisfaction that the recent legislation was rather noncommittal in the computer and network area, where it would have been much easier to do harm than to do good.

The role that I want to urge upon EDUCOM is related to the second of the two roles mentioned; it would be played in the theatre of national decisions, but it is different from representing the interests of constituents. The role is to facilitate and to improve the national decision process, especially in the areas of computers, communications and networking. Perhaps some of the new techniques of those fields can be applied to the task. At the same time, the role is to ensure that fair weight be given to intellectual values, to the quality of life, and to educating and informing (as distinguished from propagandizing and entertaining) the citizens and to giving them input as well as output access in the networks of the future.

Anne Branscomb's paper presented a clear sketch of the decision process that, unless something is done to improve it, will shape the future of networking during the coming years. The FCC likes to see some competition. Ma Bell is moving the scene of battle to the Hill. IBM is trying to get into satellite communication without getting regulated. The publishers are concerned about what computers and networks will do to royalties on intellectual property, i.e., on works of authorship. The U.S. Post Office is worried about what electronic funds transfer will do to the volume of mail. Libraries ought to worry about that, too, for if the U.S. Post Office loses the check transportation business, the price of book transportation will go up. Every industry, every institution, and every organization has an interest to protect or to advance. Some are in there, vigorously protecting or advancing. Others, like universities, are not yet well represented in the fray perhaps because they have not yet realized that vital interests are at stake.

The suggestion that EDUCOM should formulate and uphold the interests of universities and libraries in the adversary decision process has much merit, and I concur in urging that EDUCOM redouble its efforts along that line. However, if the national decision process is to be successful in defining a national interest in networking, the process has to be more than a battle among special interests. It must be structured by modeling and analysis of alternative futures. It must give more weight to the public interest, if the public interest can be formulated in negotiable terms, than to any alliance of special interests. It must take into account the contribution of networking to national strength, on the one hand, and to quality of life, on the other. A great role for the universities, therefore, may be to advance the understanding of the social impact of networking and to introduce the advanced understanding as forcefully as possible into the national decision process. A great role for EDUCOM may be: to foster and to help organize a stronger constructive effort in the universities; and to facilitate, perhaps even to provide new channels for, the flow of the results of that effort into the national decision process.

Chapter 5

by Frederic G. Withington

Distributed Computer Networks: Prospects and Problems

1.0 CAPABILITIES OF DISTRIBUTED COMPUTER NETWORKS

There is broad appeal to all kinds of users in the concept of the distributed computer network, as contrasted with shared use of a monolithic central system. Users instinctively feel they can have more control over a local computing resource than over a central one serving multiple users. They can control its responsiveness to their needs, both for fast response operation and for the development of new applications. They can better control its integrity and security. They expect that they will be charged fairly for its use, and not be subject to unexpected charges resulting from errors in an obscure resource accounting and billing system. They will not have to pay for overhead resulting from the inefficiencies of systems programs designed to handle large numbers of concurrent processes in multiple modes of use. They will not have to pay for unnecessary communications. They expect to experience greater availability: while they rarely expect to have redundant equipment themselves, they hope that if their equipment fails, they will be able to connect their terminals to other similar equipment and share its use until their own is repaired. At the same time, they hope to avoid the excessive equipment investment needed for handling peak loads by borrowing capacity from others during peak periods.

Because of this broad appeal, there have been numerous experiments with distributed computer networks in the educational community, in large business organizations, and in engineering and scientific establishments. These experiments have produced both successes and failures, and it is possible to generalize about the relative capabilities of distributed computer networks as contrasted with those of general-purpose multi-mode systems, and about the problems and opportunities that lie ahead.

Distributed computer networks have proved capable of providing three classes of data processing environments:

- *For completely independent applications.* Each user implements his or her own application, and sharing is arranged informally. ARPANET is the prototype of this class of network.
- *For local portions of centralized applications.* There is already widespread use of interactive data entry systems, point of sale processors in retail stores, and similar local equipment clusters connected to central hosts. The local equipment does all it can, reducing the role of the central system and communications with it to a minimum. Usage is particularly widespread in Europe, where communications costs are high and the local manufacturers are best able to provide smaller systems. Such a network is hierarchical rather than fully distributed because it still depends on a monolithic general-purpose machine at the center, but it can provide a substantial part of the potential benefits of distributed processing.
- *For single applications in which the files can be segmented.* Generally of the resource assignment type, these include such applications as hotel reservations (where each hotel keeps its own reservations file in a minicomputer system) and banking (where each branch of a bank keeps its own depositor file). A transaction can be entered to the network at any point and can be routed to the proper machine for processing. However, so far such applications have been restricted to those in which the transaction itself identifies the location of its file (e.g., the customer's bank account number, or his Blue Cross-Blue Shield member number, identifies the location of the system in which his record is stored).

Distributed computer networks have proved unable to provide two other classes of environments:

- *The large-scale computational facility* needed for extensive mathematical work or for high-volume batch processing runs. A user can, of course, arrange for the connection of his or her local system to a very large remote computer for operation in the remote batch mode, but the resulting arrangement is hierarchical and subject to the congestion, overhead, uncertain service, etc. associated with use of a centralized general-purpose system.

- *For complex applications* whose files cannot be completely segmented. Airline reservation systems must display identical records of rapidly changing information to all terminals; it is almost impossible to conceive of a way to distribute such a file. Most manufacturing applications require simultaneous availability of multiple files (customer records, open order files, inventories, production plans, raw materials orders, accounting data). Such applications require the use of systems programs for data base management. A given process may require reference to any of the files, so it is not possible to distribute them without retaining (and maintaining) voluminous indices to the locations of all records, which is usually impractical. These applications exist in engineering (design automation) and computer science (shared libraries of models, programs, etc.) as well as in industry and government; they make up the largest part of the data processing market. Distributed computer networks will not constitute a general alternative to centralized general-purpose computers until they can handle the logical complexities of these applications at least as easily.

2.0 PROBLEMS OF DISTRIBUTED COMPUTER NETWORKS

Users of distributed computer networks have also found that the hoped-for economy and simplicity is not always achievable.

2.1 Costs

Costs are not always as low as hoped. If the local computer is to be versatile and support multiple modes of use (batch, time sharing and perhaps transaction processing), it cannot be below a certain size. It must have a large memory to contain the necessary systems programs and probably an extensive complement of peripherals. It typically turns out to be at least a Decsystem-20 or a Hewlett-Packard 3000 and usually costs over $250,000; it is much more than a "simple minicomputer." In fact, total equipment costs usually turn out to be about the same between a central general-purpose computer and distributed smaller systems: the small systems gain in computer power per dollar, but lose in total peripheral equipment costs where the larger devices still offer economy of scale.

2.2 Ease of Use

Smaller computers are often more difficult to use. The language processors, telecommunications systems programs, file processors and operating systems of the general-purpose systems have evolved to a high level of functionality over the years. Users often do not appreciate their merit until they try to use the more limited products offered with the smaller machines. This situation will improve with time, but the providers of systems programs for the smaller systems will obviously have a hard time catching up with the systems programs offered for large general-purpose systems, particularly since the large manufacturers' systems programs are continually evolving.

2.3 Freedom

Finally, users of distributed computer networks have rarely found the complete freedom they have hoped for. If they want to intercommunicate, they must abide by common communications standards. If they want to use one another's machines, they must abide by common programming language and configuration standards (at best) or use identical systems from a common manufacturer (at worst). If the applications they develop are to be understood by others, for their own use or for modification and maintenance, the applications must be prepared and documented in meticulous conformity with common standards. The proponents of a network which is to provide all these services—intercommunications, equipment sharing and application sharing—find themselves with a more difficult management problem than they had with centralized general-purpose systems, where users were automatically constrained by its standards.

Until distributed computer networks have the ability to handle complex applications conveniently and flexibly, the general-purpose system will remain the dominant data processing tool. However, its dominance will be steadily eroded as more and more users adopt distributed computer networks for the applications they are able to handle. Such adoptions are being fostered by several of the computer manufacturers, whose approaches are illustrative of the present state-of-the-art in computer networking.

3.0 STATE-OF-THE-ART IN COMPUTER NETWORKING

One approach is to provide the interfaces and systems programs for interconnecting small computers, and then invite users to develop their applications as best they can with the limited software tools available for the smaller machines. This approach is typified by Digital Equipment with its DECNET offering, and has been successful with users who are willing to be pioneers, such as EDUCOM, Citicorp and Bank of America (which employs General Automation minicomputers). Following this approach, a competitor hopes to get an early foothold in the market and later expand it as more versatile software tools become available.

A second approach is to concentrate only on applications which are within the state-of-the-art of distributed computer networking, providing complete software for them, and hoping to grow through early dominance of a market which (though specialized) is amply large. NCR is the main proponent of this approach, offering packaged networks for retailing and banking applications. Our studies indicate a great deal of growth in store for these application areas as electronic funds transfer networks evolve; if NCR could dominate the area, it should be able to grow substantially without having major positions elsewhere.

The third approach, represented by IBM's SNA, is to slowly move its systems programs for complex applications outward from the central general-purpose computer as the state-of-the-art permits, without ever offering systems programs of lesser functionality. This approach recognizes that most users are neither willing to be pioneers nor are in industries having only simple applications. Most users are believed to want the best software tools available for their applications, above all, and are expected to wait for the full benefits of distributed computer networks until the networks can offer the best software tools. A company following this approach is, of course, not precluded from following the other approaches in parallel as IBM may do.

To summarize, distributed computer networks have already made considerable inroads into the dominance of the centralized general-purpose computer, and substantial opportunities for further growth exist within the present state-of-the-art.

The general-purpose system will hold its own fairly well, however, until network systems programs are available for handling complex applications in a transparent, integrated manner. These will evolve only slowly and may not be perfected for five years or more. In the meantime, the users will be offered a confusing but rich diversity of alternatives by the manufacturers.

Part II Computer Networking in Library Use and Administration

by Peter P. M. Meincke

Overview: Cataloguing, Reader Services, Management, Future Developments

This set of papers covers current and future roles of computers and computer communication networks in libraries and the impact on resource sharing among libraries. Currently available services and activities, the impact on library operations including costs and benefits, and the increased opportunities for resource sharing among libraries are described in three broad areas: cataloguing; reader services; and management. The last set of papers is devoted to a discussion of possible future developments.

The greatest impact of computers on libraries has been in the area of cataloguing. The capability of sharing cataloguing information through computer network facilities has significantly reduced cataloguing costs and led to a much greater uniformity of bibliographic standards. The introduction of on-line terminals is also bringing about a major reorganization of cataloguing activities.

There have been significant effects on library services available to readers as a result of the introduction of commercial information retrieval services, on-line and microform catalogues, computerized circulation systems, and data banks. These effects have occurred in libraries of all sizes and a variety of stages of automation. Papers on this topic review the impact of these services on the library staff and users and the pros and cons of charging for various services.

Functioning library computer networks have run into many of the classic problems of resource sharing among the autonomous institutions; for example, management of the facilities, setting of priorities, charging for services, setting of standards, etc. The libraries also face new problems such as charging for catalogue data which, in many instances, has been regarded as a free good in the past. Loan information, holdings information, and other data captured during the cataloguing process are only a few of the many management information items that automation will provide, and which could have very significant impact on the administration of individual libraries. Papers on this topic discuss the problems of shared use of facilities and resources, and the impact of such management information on library administration.

As library operations are increasingly affected by computers and computer networks, alternative futures become clearer. Choices are being made now which may lock us into one or another of these futures. To explore possible futures, authors discuss: the potential impact of known technological advances (such as very cheap mass storage and communications); the potential impact on the ultimate users of the information, and the staff required to serve them; library organization and staff training; the cost of operations; and strategies for the introduction of automation.

Chapter 6

by Jack Cain

Standards and Identity in Cataloguing

1.0 THE PROBLEM OF STANDARDS

The chief difficulty with current standards which affect cataloguing is that they are not keeping pace with the developments and impact on the cataloguing process of automated techniques.

The process of arriving at standards is slow and arduous, especially since data exchange is becoming more and more international. Already there are five authors of the Anglo-American Cataloguing Rules (Library Association (Great Britain), British Library, Library of Congress, American Library Association, Canadian Committee on Cataloguing (representing three national bodies)), and more are knocking on the door for a voice in the revision, now underway, of this code.

Various standards, all of which affect cataloguing in some way, are being developed by separate bodies with often conflicting results: a series of standards for descriptive cataloguing for different categories of library material, called ISBDs (International Standard Bibliographic Description) is being developed by IFLA (International Federation of Library Associations); each country which has developed machine readable cataloguing has developed its own version of MARC—the format in which cataloguing data is expressed—an international format, UNI MARC, which hopes to reconcile all the national variations is slowly being developed by an IFLA working group; several filing codes exist and are in use for the printed displays of machine readable cataloguing data; several conflicting romanization schemes are in use internationally for the conversion of letters or characters into the Latin alphabet; two different standards exist for the creation of cataloguing data for serials or periodicals—those specified by the International Serial Data System, and those specified by the Anglo-American Cataloguing Rules.

At the same time as chaos reigns in the field of standards, large data bases of machine readable cataloguing records are being built with amazing rapidity. These bases are already counting their records in the millions, and more are springing up.

What are the implications for data exchange of the creation of these bases with their varying and variable standards? In fact, no one seems to know.

2.0 BIBLIOGRAPHIC IDENTITY CRISIS

At the core of the problem of standards is an issue which can be expressed by the philosophical question: "Is *this* identical to or the same as *that?*" Some large files are being created in which, through human intervention in the form of careful editing, duplicate records are not filed. Other large files either do not, or, because of their construction, cannot monitor the entry of duplicate records. Some do a partial job. As large files begin to communicate with each other for the purpose of exchange of cataloguing data or to create Union lists of holdings, the identity crisis will become more acute.

A unique identifier for each bibliographic item would, of course, solve this problem. To some extent ISBNs (International Standard Book Numbers) and ISSNs (International Standard Serial Numbers) as well as Library of Congress Card Numbers provide this function, but there are limitations. ISBNs, being publisher-assigned and not controlled by any central agency, are often in error; many monographs do not have ISBNs at all.

Once a unique identifier is present, the computer can decide that *this* is *that*, thus providing cataloguing data that is correct and holdings information that is accurate.

The identity crisis applies not only to records for bibliographic items, but also to names. Many large machine readable files are incapable of providing a list of works by a given author in such a way that the list is complete (for that base) and contains no extraneous information.

The identity crisis is not often noticed as a major problem . . . yet. However, the operative word is "yet". As files continue to grow and links begin to be established between them, this problem will, no doubt, become more and more apparent. Although few experiments have been done, it would appear that going back and trying to repair the situation can probably only be done with human intervention and is consequently very expensive.

Chapter 7

by A. H. Epstein

Shared Cataloguing: The Impact on the Library and the Patron

1.0 INTRODUCTION

There are several different levels or styles of shared cataloguing that are available today as on-line services. The first level allows a single version of a given title to be maintained in the on-line file. This is usually the Library of Congress MARC version of a record, or the first user's version of a record. Both OCLC and the Washington Library Network (WLN) use this form of service. The user is allowed to copy the record and make modifications. These modifications will appear on the user's catalogue cards and, in the case of OCLC, on a history tape. In the case of WLN, the user is permitted certain modifications and can receive catalogue cards.

The next level of shared cataloguing allows two versions of a title to be stored in an on-line file, the Library of Congress MARC version and the first user's version. This is the level that was implemented by the BALLOTS system until 1976. By separating the MARC file and the user's record, the MARC file could be updated with any changes and modifications made by the Library of Congress without affecting the user's record. This was done so the user's record would remain unchanged while other libraries would be free to use the latest version of the MARC record.

The third level of shared cataloguing service allows multiple versions of a title to be stored in the data base. This allows both the MARC version of a record and individual versions of records of all libraries participating in shared cataloguing to be maintained on-line. This third level of service was implemented in the BALLOTS system in 1976.

An extension of this level of service that improves computer and file storage efficiencies, would allow each of the variations of a title to be recorded within the same computer record. The University of Toronto has implemented the extended capability in their on-line system. Development plans at the University of Toronto, BALLOTS, OCLC, and WLN, include the provision of this extended capability in one form or another.

2.0 SHARED CATALOGUING—FIRST AND SECOND LEVEL

The library uses four basic sources of information in a shared cataloguing file. The first source is data catalogued by the Library of Congress and provided on Library of Congress MARC tapes. The second source is information catalogued by the Library of Congress, but keyed in by a library participating in shared cataloguing. The third source is data catalogued and keyed in by the participating library. The fourth source is data that is catalogued by another library and keyed in by a participating library generally from a source such as a book catalogue.

The obvious advantage of shared cataloguing from any of these sources is the elimination of duplication. Once LC supplies a MARC record or one of the participating libraries keys a record in for a given title, all other participating libraries may copy that record and make necessary modifications. Another advantage is the reduced cataloguing load. For every record that is found in the data base, the library need not perform the cataloguing activity for that record. Due to the shared use of a system, it becomes easier for all participating libraries to participate in the use of standards. Finally, since a library is taking advantage of the cataloguing efforts of other libraries, there is less work for a given library to perform. That is, there are far fewer titles that actually have to be catalogued by a given library. In return for finding so much work already performed by other libraries, it seems reasonable that a library that does catalogue a title which will be placed in a shared data base will do the best possible job in cataloguing to enable that record to be used by as many other libraries as possible.

3.0 ADVANCES IN SHARED CATALOGUING—LEVEL THREE

Implementation of the third level of cataloguing, in which each library maintains its own version of a title on-line, will have a profound effect on the library world. This level has been recently implemented at

BALLOTS, but it is too early to predict the full extent of the impact on libraries. One of the major advantages of the third level is the ability of each library to review any other library's version of a catalogue record. Also, the library that enters a particular version owns that record and has the ability to modify that record in the future. In late 1976, two large research libraries are participating in this level of shared cataloguing: the Stanford University Library and the University of California at Berkeley. Several large public libraries also participate including San Francisco Public Library.

As a result of this level of shared cataloguing, various libraries will most likely search different versions of the record until they find a pattern or discover one or two libraries with which they are quite satisfied. The searching library will then tend to search for the libraries of choice, and either accept the record as catalogued, or make very minor modifications. After such informal procedures are implemented, the work can be handled by a lower level classification of employee. For instance, at Stanford University and many other libraries, if a record appears on the MARC file, it is used with little or no modifications, since the cataloguing was done by the Library of Congress. Paraprofessionals are able to perform this level of work. Libraries will soon begin to treat specific other libraries with the same level of authority, and thereby obtain additional cataloguing and allow that cataloguing to be processed by the non-professional staff.

The University of Toronto has a similar capability implemented on-line. Any library in the system may review another library's cataloguing, and copy it or modify it to form a new custom version.

4.0 OTHER ADVANCES IN SHARED CATALOGUING

At Stanford University, there is an on-line acquisitions system that interfaces with the cataloguing system. The BALLOTS system provides an on-line in-process file of all books, from the time they are ordered until the time they are through cataloguing. Activities such as ordering, claiming, cancelling, keep track of the paperwork, and activities such as receipt of material, distribution of material, and non-purchase order receipt keep track of the item. As a result of this system, the bibliographic data is keyed in as soon as it is known, and updated as soon as new data becomes available. Thus the record attains different levels of maturity throughout its life cycle. One of the impacts of this on-line in-process file is that it has blurred the line between acquisition and cataloguing. It is not clear any longer where acquisition ends and cataloguing begins. Current plans at Stanford University include a provision to allow for a function to be called "Cataloguing in Receipt". This will allow a receiving clerk to not only acknowledge receipt of the library materials, but, if the material falls within certain categories, to catalogue the book at that point. The book would then be transported directly to the end processing area and by-pass the cataloguing area entirely. When this happens, librarians have to ask themselves a question. Was that actually cataloguing? There may need to be a need to redefine cataloguing.

Another impact on the catalogue department of the existence of an in-process file is the ability of one library to find out roughly what another library has on order. If two libraries have a good interlibrary loan arrangement, or a more sophisticated shared collection program, the second library need not order a particular item. If it is needed by a patron, it can be borrowed from the first library. Shared collection development becomes quite feasible when there is an on-line shared cataloguing file. Most large-research libraries have extensive collections in many different disciplines. However, all disciplines cannot be covered completely. Two or more libraries can get together and agree in certain areas where neither one of them have adequate coverage. They might agree to allow one of the libraries to maintain as complete a coverage as possible in a given discipline, or a given part of the collection, and the other will collect different parts of the collection. As long as the information is available to all of the libraries, and a good interlibrary loan operation can be implemented, each of the libraries has the advantage of a full collection in all of the disciplines under the agreement. The impact on the catalogue department in these libraries is quite beneficial. Rather than spread themselves among many different disciplines, cataloguing staff can concentrate and become more specialized.

4.1 Full Indexes

The BALLOTS System indexes each of the bibliographic records by all of the personal names in the record, each of the individual title words, the subject headings, the call number, the Library of Congress card number, and the corporate and conference authors. This allows the book selectors, reference staff, patrons, cataloguing and acquisitions staff of a library to determine what books are held by various institutions in a given subject area, call number range, or by certain corporate or conference authors. The catalogue department can search the file by call number or subject heading to determine what other titles in the file were catalogued under these terms. This is a valuable tool which should come into increasing use. The use of this type of searching by reference or public service staff members (or by library patrons) answers the question of what else the library has in this subject area, or what books have a certain word or words in the title. Furthermore, if this library does not have the title, what do others have in this same area? This is a powerful ability that is in use at Stanford University.

4.2 MARC Standing Search Requests

The BALLOTS system provides the ability for the catalogue department to institute an automatic repetitive search against new records coming in on the MARC file. If the library suspects that a record will appear shortly on the MARC file, the book is held aside waiting for the MARC record. Stanford University has approximately 3-4,000 titles in this category since it has a hit rate of somewhere between 10 and 20% each month. The impact upon the catalogue staff is that the cataloguers do not need to repetitively search on a terminal, for a particular title, since the search is kept in a separate file and processed against the MARC records on a monthly basis. Also, for every title that is found in the MARC file, the library need not catalogue that book.

4.3 Reference Files

One of the BALLOTS files available for searching is the Reference File. This file contains reference and cross-reference information for authors and subjects that is equivalent to the information found in a card catalogue on the "See" and "See Also" records. That is, if a new version of a corporate name appears in the file, there will be a cross-reference record, such as "EDUCOM" see "Inter-University Communications Council". If the searcher put a search in for "Find Corporate Name EDUCOM", they would see a note referring them to the Inter-University Communications Council. Since the files are automatically searched, the user need not specifically request a search of the cross-reference file.

4.4 Interlibrary Loan

As long as the item identification of a particular book at a particular library is maintained in the file (the local call number, the location of the book, and the copy number) it is possible to make a request for a specific book from a participating library. It is possible to copy the record by printing the record out on a typewriter terminal, or hard-copy terminal, and send that form in as a request for inter-library loan. Eventually, it will be possible to ask the system to produce an inter-library request form, since the system knows who the searching library is, which record is being reviewed, and what the local call number and copy number are. The system has enough information to print out a complete interlibrary request form. The advantage to the library receiving either a manually filled-out form which is done at the present time, or a computer-produced form, which will probably be produced within the next two years, is that the library need not search its own files to see whether it indeed has the book. The information is accurate enough for the library to determine whether or not the book is available for loan, and proceed directly to the proper shelving location and obtain the book. The searching step can be eliminated.

4.5 Magnetic Tapes

Both the BALLOTS and OCLC systems produce tapes for their users. Generally in the Library of Congress MARC format, these tapes are used for a variety of purposes. Some libraries process the tapes at their local data processing centers; other libraries run these tapes on to vendors or service bureaus who in turn produce book catalogues, microfilm and microfiche catalogues. These tapes will also be used in the near future to combine the holdings of different libraries and produce regional and statewide union catalogues.

5.0 ADVANCES IN SHARED CATALOGUING—FUTURE

5.1 A National Network

Each of several major on-line library automation networks throughout the U.S. and Canada has a separate data base. It is highly desirable that library users on one system have access to the database of the other systems. Yet such access would require an interface between the files and computers of the different systems. Although mutual access will take at least five to ten years to accomplish, it will have a major impact on all aspects of library work, including acquisitions, cataloguing, reference, patron searching, interlibrary loan, etc.

5.2 Authority Files

Authority files which define the approved version of personal names, corporate names and subject headings exist at the Library of Congress, the New York Public Library, and WLN. These authority files are currently implemented in those systems in the batch mode, and users submit records to the system. The records are checked overnight against the authority files and any variation of a name or a subject heading that does not appear in the authority file is rejected by the system. The ability to have access to an on-line authority file would have a great impact on the catalogue departments of all participating libraries. The combination of having, for instance, the Library of Congress authority files on-line, and have each of the networks inter-connected to the Library of Congress, means that the cataloguers can determine immediately if the Library of Congress has an authoritative version of a name or subject heading. If not, the cataloguer is required to submit a new name or subject heading. In most cases, however, the name is on the file and the cataloguer has a record that is validated. At a single terminal session, the cataloguer can provide an authenticated record. A record of this type is much more acceptable to most cataloguers, since each of the main entry points of a record will have been validated by the Library of Congress authority files.

Once authority files have been implemented, it will be possible to make bulk changes to the various files. For instance, consider the situation where a government agency changes its name. Rather than have a new cross-reference which goes from the old name to the new name, it may be possible to locate all records in the file that contain the old name, and update them to the new name automatically. This will have a great impact on cataloguers and patrons alike. In order to implement something like this, it may be a prerequisite to have abandoned the use of catalogue cards. It would be extremely uneconomical to reproduce all of the catalogue cards that had the name of a particular government agency, or major corporate entity, and refile in the catalogue or replace in the catalogue all of the new cards.

5.3 Patron Access—Other Catalogues

Sometime within the next five to ten years, the use of catalogue cards will be reduced if not eliminated entirely. There will be several forms of technology which will take the place of card catalogues in individual libraries. Book or microfilm catalogues are likely replacements that have been used already in many libraries throughout the United States. Some of these libraries do not have catalogue cards, but merely microfilm readers located throughout the different parts of the library and in different branches. To produce the microfilm records, one of the shared cataloguing users of the BALLOTS system obtains a magnetic tape rather than catalogue cards as a service. These tapes are sent to a vendor to produce a cumulative catalogue of all of the library's holdings on a bi-monthly basis, and the old microfiche file is discarded and replaced with the new one.

5.4 Patron Access—On-Line

Another form of catalogue card replacement is to maintain the files on-line as is done at the present time on the BALLOTS system. Patrons have the ability to search the file by all of the indexes described above. Within the next five to ten years, individual libraries will probably be obtaining their own minicomputers and maintaining a file of their collection on a local storage device. This will allow the use of patron terminals throughout the library, running off the minicomputer owned by the library, and makes the elimination of the card catalogue a practical idea.

At the beginning, patrons will be searching, using the existing indexes. Eventually, the library will probably begin to add more terms to be indexed in the bibliographic record. Since the amount of shared cataloguing has a tremendous impact on the elimination of duplicate effort throughout libraries, the library will have additional time and staff to turn to increased functions for the patron. It is likely then that the bibliographic record will be extended to include additional terms to be indexed like additional subject headings and alternate call numbers. Following that, the catalogue department will be responsible for generating a master thesaurus of terms which will allow patrons to go in under the vocabulary familiar to their discipline and have this translated automatically into terms that are indexed in the system. It should not be necessary for patrons to learn the library terminology. The computer should be able to provide the cross-referencing function automatically.

The next step could be the inclusion in the bibliographic record of the indexes at the back of the book. Many times the major terms indexed in the book are an indication of what the major topics are within the book. An index also provides a patron with a guide to which books contain descriptions of the terms the patron is looking for. In addition to that, it may be feasible for the bibliographic record to include the entire table of contents of the book. If the patron had the ability to browse the table of contents, it might be a great aid in determining the usefulness of a particular book.

Finally, one must consider electronic publishing. Already in 1976 several publishers, including West Publishing and Chemical Abstracts Service, have the entire contents of their publications produced through a computer and used for automatic typesetting. In the next ten to twenty years, the computer technology and storage technology will move in a direction where it may be economically feasible to store the entire contents of a book on-line. Perhaps the publishers will each own their own data base, and those users who wish to scan a particular book, or copy a particular chapter, would then be able to do so on their own terminal, and automatically pay the publisher and the author the appropriate royalties. Electronic publishing is far beyond the scope of this paper, but I believe it will have a major impact on libraries within the next twenty years.

Chapter 8

by William L. Newman

Characteristics of the DOBIS System

1.0 INTRODUCTION

The DOBIS System (Dortmunder Bibliothekssystem) is being developed by the University of Dortmund, West Germany. The system, currently being evaluated in Canada, employs a design philosophy usually not found in existing on-line cataloguing systems. Possible implications of these design features are discussed in this paper.

In 1971, the University of Dortmund contracted with IBM West Germany to jointly develop an integrated, on-line library management system. The resultant system, now called the German version of DOBIS, consists of cataloguing, catalogue search, and acquisitions subsystems in the testing stage, a circulation subsystem implemented at the University of Dortmund, and a serials control subsystem in a mid-development stage.

After examining a number of on-line systems, the National Library of Canada (NLC), the Canada Institute for Scientific and Technical Information (CISTI), and the Council of Ontario Universities (COU) decided to have a closer look at DOBIS. A DOBIS Working Group, consisting of library systems personnel from across Canada, drew up specifications for the introduction of Canadian requirements into DOBIS. A modification contract was negotiated between the University of Dortmund and NLC. During the period January to April 1976, DOBIS personnel, including the key IBM DOBIS developers, the McAllisters, modified the German versions of cataloguing and catalogue search to provide: 1) compatibility with what might be termed North American MARC; 2) a bilingual English/French system (rather than

German/English); and 3) a multiple library use capability. The resultant cataloguing modules have been called MARC DOBIS.

In June 1976, NLC, CISTI, and COU started a rigorous evaluation of various technical, functional, and performance aspects of MARC DOBIS. The evaluation should be finished in December 1976 at which time a submission will be made to Treasury Board with a go/no go recommendation. If the evaluation is positive, MARC DOBIS will probably be the base system for Canadian federal library housekeeping. If not, alternative systems will be evaluated because the reason for acquiring such a system—rising work loads, increasing costs, personnel shortages—certainly still exist. Note that although the MARC DOBIS evaluation has identified a number of problems ranging from trivial to significant, the results so far in November 1976 are positive.

In combination, references 1-4 provide a fairly complete description of DOBIS.

2.0 DOBIS CHARACTERISTICS

2.1 Machine Requirements and Program Characteristics

MARC DOBIS requires a machine capable of running the IBM VS operating system. It uses the CICS/VS communications monitor and approximately 300K real memory for MARC DOBIS and CICS/VS together with about 10K virtual for each terminal.

The basic mapping support feature of CICS/VS is employed in DOBIS which means that only IBM 3270 compatible terminals, with a screen display of twenty-four lines by eighty characters, can be used with the system.

As part of the DOBIS Project in Ottawa, a test data base is being created. Therefore, only crude measures of operational resources required are available at this time. These include about 100 disk I/O's per record and 4.4 IBM 370/135 CPU seconds per record to create a completely new record.

The MARC DOBIS system consists of about sixty-five programs. Eighty to ninety percent of the source statements are PL/I, the remainder IBM 370 assembler. The programs are modular, and largely self documenting.

2.2 Human Factors

Access points to DOBIS include LC card number, ISBN, ISSN, local call number, author names, subjects, titles, publisher, and miscellaneous numbers. The actual search term, or any truncation thereof causes fourteen retrieved access points to be displayed. The second line on the display corresponds most closely with the operator supplied search term. The operator can then see documents associated with any of these access points, or page forward or backward through the access point file.

Each display screen conforms to a standard three-part screen format:

- screen title (current function and subfunction) and error messages;
- communications area including retrieved information, and tables indicating possible replies to DOBIS, etc. and
- instructions to the operator and an area for operator response. Menus are employed extensively in the terminal dialogue to indicate possible operator responses to DOBIS. These features make the system easy to learn to operate. After an operator becomes experienced with the use of DOBIS, several commands can be combined into one. This last feature is called command chaining.

The DOBIS system has been designed to provide decentralized access to all library catalogues and files, and to permit one time entry of data. This means that the cataloguer can have on-line access to such vital cataloguing tools as the shelf list and all the authority files in one spot. The acquisition librarian, interlibrary loan location searcher, or reference librarian can perform one-stop searches of the in-process file and the catalogue. Also, if a long authority heading has been entered into the system once, it doesn't have to be rekeyed for each record that requires it.

2.3 Terminal Dialogue

In Canada, which is a bilingual country, it is important that information retrieval systems function in both French and English. Terminal operators can converse with MARC DOBIS in both languages.

2.4 Integrated System Design

From the beginning DOBIS was designed to be a fully integrated system incorporating the cataloguing, catalogue search, circulation, acquisitions, serials control, and information retrieval functions. The first screen with which a user is confronted after sign-on is a function selection. If the user has appropriate security clearance for the selected function, the required task can then be performed. All necessary information to perform the task should be available in one spot.

2.5 Data Base Organization

The DOBIS data base organization is illustrated in Figure 1.

Information stored in an access point file is not duplicated in the bibliographic files. For example, a name access point record contains the name, information about the name, and pointers to all bibliographic records that the name accesses. Bibliographic records related to the name contain a pointer to the name, and any information applicable to the name/bibliographic item relationship, for example, an indication of editor, illustrator, and so on.

The illustrated bibliographic file is actually two files. The primary bibliographic file contains sufficient information to identify the item. Since the secondary bibliographic file fills out the complement of bibliographic fields, almost any output, including full MARC records, could be produced from DOBIS.

The system holdings file outlines holdings by library.

Local bibliographic files and local access point files do not have to be established for all libraries using a DOBIS installation. If a library accepts the cataloguing established in the common files, then the local biblio and a.p.f. files are obviously not necessary. Even if a library does define local files, they only need be used for records and fields within records differing from the common records and fields.

The local holdings file is mainly a mechanism to get from the bibliographic file to records containing information applicable to each copy of the document. Each record in the local copies file contains information on status (acquisition, bindery, circulation, etc.) and location of the copy. Notes can be recorded in both the local holdings and local copies records.

Relationships between records in the bibliographic file, and between authority records in the name or subject files, are accomplished by control number linkages and type of linkage codes rather than by the repetition of information in both records. Examples of relationships between records in the bibliographic file include:

- preceding/succeeding,
- parent/analytic,
- parent/supplement, and,
- original language/translation.

Between authority records, relationships include:

- see/see from,
- see also/see also from,
- English form of name/French form of name, and,
- broader/narrower.

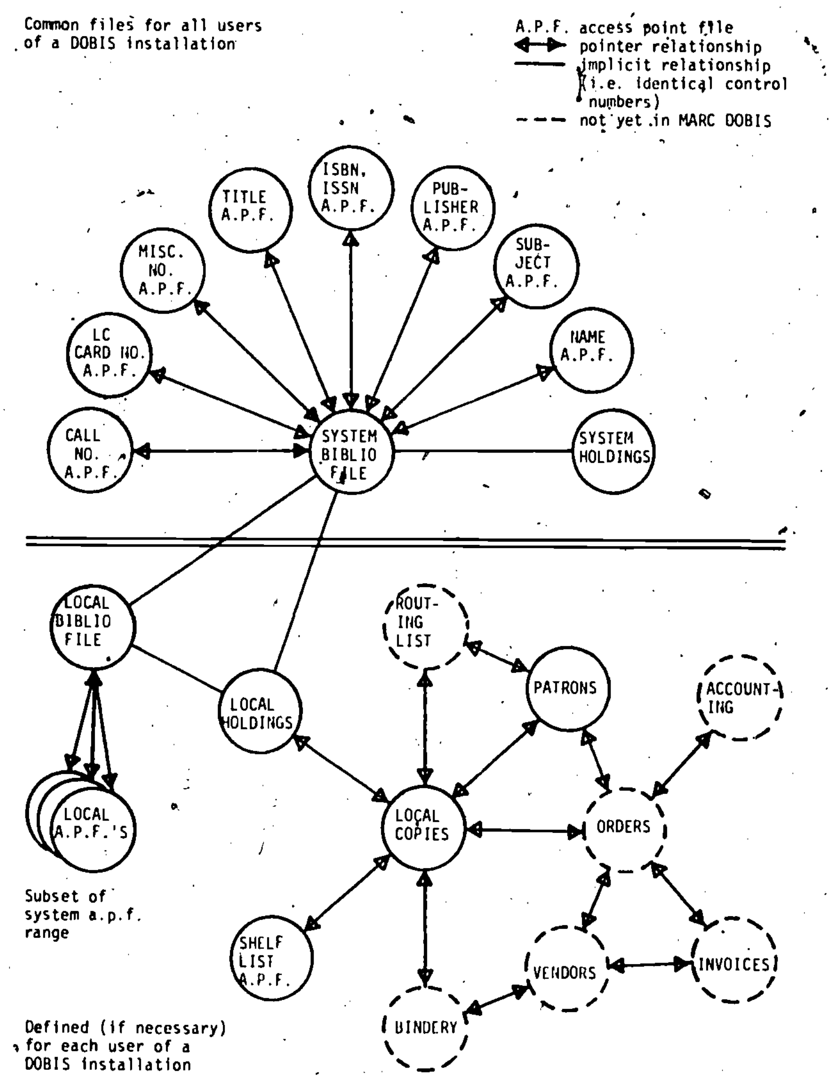


FIGURE 1. MARC DOBIS Data Base Overview

3.0 IMPLICATIONS

Through the use of authority files, only the authority record must be changed to introduce current usage, terminology and standards rather than each bibliographic description in which the authority is used. It's easy to see how authority information can be kept up-to-date and accurate with this type of file arrangement. Adjustments from ALA to AACR to AACR II can be made in practice instead of nominally. Malinconico and Rizzolo (6), and Buchinski et al (5) have discussed additional benefits of the use of authority files in cataloguing systems.

Provision of all cataloguer tools in one place should speed the cataloguing process significantly. Duplicate recording and storage of information for different functions or purposes is also virtually eliminated. This should enable all affected functions to be performed more efficiently.

The price paid for ease of file maintenance, elimination of duplicate recording and storing of information, and general system flexibility and expandability is mainly in terms of greatly increased numbers of input/output (I/O) operations to and from direct access storage. For almost any operation, information must be stored in, or collected from, several files rather than one. Can the increased system capabilities pay for the increased disk I/O's? That is one of the questions being addressed in NLC, CISTI and COU evaluation and impact studies.

4.0 CONCLUSION

In order to realize benefits discussed in this paper, card catalogues must be closed. As evidenced by the University of Toronto (4) and the Library of Congress (8) among others, the decision to close is becoming more commonplace.

In 1949, the ALA *Cataloguing Rules for Author and Title Entries* was published. *AACR* was published in 1967 followed by AACR revised 1970, AACR Chapter 6 (ISBD) 1974, AACR Chapter 12 (AV material) 1976, and AACR Chapter 14 (sound recordings) 1976. A major revision of AACR, abbreviated *AACR II*, is in the works and should be published in 1977. It is evident, even to the uninitiated, that there is accelerating change in the field of library cataloguing standards.

Laser and other technology holds the promise of a dramatic decrease in on-line storage costs.

With all of these trends, the closing of card catalogues, accelerating changes in library standards, and decreasing on-line storage costs, systems like DOBIS that employ data base management concepts, and an expandable, flexible design, if not feasible now, will be not only feasible, but necessary in the near future.

References

1. IBM. *System and Implementation Guide. Dortmund Library System–DOBIS.* in preparation. Order number GE15-6070. "Addressed to data processing management and systems analysts, this guide fully documents the operational stages of the system. . . ."
2. IBM. *Application Guide. Dortmund Library System–DOBIS.* Feltham, England, IBM, October 1975. Order number GE15-6069. Order through your IBM representative or branch office. ". . . addressed to the managers with responsibility for the major functions that are affected, and it explains the objective and benefits of the application in terms of those functions."
3. IBM. *Executive Guide. Dortmund Library System–DOBIS.* Feltham, England, IBM, October 1975. Order number GE15-6068. Order through your IBM representative or branch office. ". . . addressed to senior management of the potential user. It describes the areas of operation covered by the application. . . ."
4. Blackburn, Christopher, "New Ways to Store Catalogue Info" *Quill and Quire* 41:22-23 (September, 1975).
5. Buchinski, Edwin J., William L. Newman and Mary Joan Dunn, "The Automated Authority Subsystem at the National Library of Canada" *Journal of Library Automation* (to be published).
6. Malinconico, S. Michael and James A. Rizzolo, "The New York Public Library Automated Book Catalog Subsystem" *Journal of Library Automation* 6:3-36 (March 1973).
7. Sangster, Lesley. *DOBIS Notebook (Overview and Internals).* in preparation, sponsored by the DOBIS Project, National Library of Canada, Canada Institute for Scientific and Technical Information, and the Council of Ontario Universities.
8. Welsh, William J., *The Future of the Catalog,* paper presented at the International Federation of Library Associations General Council Meeting, Lausanne, Switzerland, August 1976.

Chapter 9

by Pauline Atherton

On-line Reader's Services

The Systems and Procedures Exchange Center (SPEC) of the Association of Research Libraries conducted a User Services Survey in 1975 to assess the directions ARL libraries were taking in the area of public service. More than 1/2 of the ARL libraries responded. The SPEC survey results, available in the SPEC Flyer No. 26 (dated April, 1976), indicate that more than 60% of the 69 respondents now conduct on-line searches to access such bibliographic data bases as ERIC, Psych Abstracts, COMPENDEX and MEDLINE via telecommunication networks to off campus computer systems such as those provided by the National Library of Medicine, SUNY-BCN (Biomedical Communications Network) and commercial vendors such as System Development Corporation, Lockheed Retrieval Service, and Bibliographic Retrieval Services.

The responding libraries are providing their users with literature citations at a cost. Few of the institutions provide searches as a part of the "free" or "hidden cost" services normally offered by the public service staff of the library.

This new development (less than five years old in all responding institutions) has had a tremendous impact on the library staff and the users of the library even though the service has been inaugurated at the same time that other user services have had to be reduced or eliminated because of staff and budget reductions. It is intriguing to try to understand why this is happening. The SPEC Flyer No. 26 contains details about other new services such as the loan of portable microform readers and book delivery systems, but no new development is as common as *computer-based reference services*. One reason for the prevalence of computer-based reference service may be because the service is relatively easy to inaugurate. Staff training and some capital outlay for a computer terminal and modem are the biggest initial expenses. Where the decision has been to pass charges on to the user, some accounting procedures have to be established, but relatively few costs are incurred. There is no computer time to budget, no software to write, no programmers or systems analysts to hire. Truly this use of computers is painless compared to other library automation projects. As the preliminary results of the SDC study, in 1974, indicated practically every academic library could, and probably would, begin such a service because it was an innovation so easy to adopt.

Thanks to a Council on Library Resources fellowship, I was able to visit more than 25 university and college libraries in the east, midwest, Canada and California to find out firsthand what impact this service was having on the more traditional reference service operation. The effect on the more traditional reference service was an important factor both to the successful adoption and continuance of the new service. I was curious to know how the staff viewed their professional role now that computer technology was a tool for giving user service and not just for performing technical operations better. I wanted to know how library administrators viewed the incorporation of the new user service, because academic library objectives have usually given priority to collection development. My interviews were not structured so my study should be considered impressionistic rather than scientific. Nevertheless, several impact statements were heard over and over again reinforcing my view about the direct impact of these services on the staff, the library itself, and the library's user.

Many staff members commented on the opportunity on-line searching services provided to be "more professional" for they must be an analyst before and after the search, getting a very clear understanding of the search request and the output results. Someone's comment summarizes this impact nicely: "I know and the user knows that all is in good hands."

The new service emphasizes the library's ability to give customized service, where the professional staff is available by appointment. This kind of routine is becoming the rule for consultative bibliographic service and library instruction. The staff's time, considered very valuable, is allocated and planned more closely now.

The computer search output is sometimes just the beginning of the service the library staff member can render. Other, older resources of the library are added to the list of current materials which the on-line bibliographic data bases cover. A tutorial on bibliographic search in the library may follow the search on-line. These are older services with a new twist.

Library administrators have noticed a different set of impacts on the library. Some of them expressed it in this way: "For the first time we are making a careful time check of our reference librarian's duties. We have reviewed their duties and their costs to the library and we have reassessed the priorities of these duties when our budgets have to be cut. When we can, we now pass on to the user a charge for staff time."

"Because the compilation of a bibliography is rather painless using the computer, we have found that our resources are used more and our ILL department is very busy. The weaknesses of our collection are immediately evident, causing us to make a greater effort to fill in the gaps. We now can use the computer searches and our analyses of the data bases on-line to check our holding in these areas. We also realized that we must streamline our document delivery systems and procedures to match the quick turnaround time of these on-line searching services. Both in the library's circulation procedure and our ILL were found wanting.

The cost/benefit rates for computer-based reference service is easily determined, and statistics are easy to keep. The administrators are trying to obtain comparable facts about other user services. Some said, "We never studied costs or benefits before. Current study should help us allocate resources better."

An unexpected impact on the library has been the need to realize that the library was now in the *marketing* business. As the service was available at a cost to the user, it needed to be promoted, demonstrated, and matched to the particular clientele best served by their use of it. This had a *direct effect* on the library's view of itself and a *cross impact* on the view of the library by others on campus. Almost overnight, large segments of the potential user population of the library saw the library as a modern service, the library staff as professionals interested in giving efficient service. Opinion leaders on campus began to speak of the library as being very sophisticated, more like the computer center than the archives department. In some instances this helped the library get a more favorable airing of their entire budget and also get some special funds.

Graduate students appeared to have more contact with the library staff now. A delayed result of this interaction may be their continued use of the library's custom services when they are teachers themselves.

The issue of elitism is a definite negative impact of these on-line services. Now these services in most libraries are available only to those who can pay and whose subject interests are represented by the on-line data bases. This could change as subsidies or charging patterns change, and as more data bases are added.

Overall, the introduction of computer-based on-line services has had a positive impact on library staff and library users. Beyond fixed costs, the costs involved are specific to the user and particular users. This has made it difficult to amortize all the costs over all users or to come up with a budgetary item absorbed by the institution rather than passed on to the user.

Eventually library managers will have to sort this out as have the computer center managers. I hope that a group like EDUCOM would bring the computer center managers, library managers and on-line bibliographic service vendors together for an exchange of procedures and charging policies. This may be a critical area affecting the future of on-line user services in academic libraries.

Chapter 10

by Judith Wanger
and Carlos A. Cuadra

On-line Impact Study

The 1970's have seen tremendous growth in the use of on-line technology for access to bibliographic-type information, i.e., citations, abstracts, and research summaries. As a long-time supplier of on-line retrieval services, System Development Corporation was aware that a wide variety of questions were being raised about these services that could not be answered with available data. To help provide such data, and address these questions, SDC developed ideas for an exploratory study of the users of the major on-line service suppliers and, in 1974, received a grant from the National Science Foundation to conduct the study.

Background on the study objectives and methodology and data from the study results are highlighted below, in the following nine sections:

Study Background and Methodology
Description of On-Line User Study Population
"Getting Started"
Selection and Training of Staff
How the On-Line Services Are Being Used
Costs of On-Line Searching
System Preferences
Problem Areas
Impact of On-Line Services

 This study was conducted with the support of the National Science Foundation (Grant Number SIS74-03465 A01).

1.0 STUDY BACKGROUND AND METHODOLOGY

The purpose of the study was to describe the community of users of on-line services, to learn about the kinds of preferences that users were developing and the problems that they were encountering, and to assess the kinds of impact that on-line literature searching was having on their operations.

SDC elicited the cooperation of 9 other major suppliers of bibliographic on-line services, to provide the broadest possible user and supplier base for exploring these areas. Each of the participating suppliers was asked to help in several areas: 1) to provide us with their user list, or to mail our questionnaires directly to their users; 2) to review our questionnaires and assist in revising items; 3) to review the preliminary data and the draft final report. The participating suppliers are:

- Energy Research and Development Administration
- Battelle Memorial Institute
- Canadian Institute for Scientific and Technical Information
- Defense Documentation Center
- European Space Agency
- Lockheed Missiles and Space Company
- National Aeronautics and Space Administration
- National Library of Medicine
- State University of New York
- System Development Corporation

Letters requesting participation in the study were mailed to using organizations identified by the suppliers, or directly from the suppliers to their users. Each organizational unit that indicated it was willing to participate received one Managers Questionnaire and up to 3 Searchers Questionnaires. Neither of the questionnaires posed any questions about specific systems or data bases, since our purpose was to understand the use and impact of on-line services, not to evaluate one system or technique against another. The total number of respondents was 1273: 472 Managers and 801 Searchers, representing 546 unique organizational units. We believe that these participants represent about two-thirds of the total on-line using community as of 1974 and early 1975. Of course, that community has continued to grow very rapidly.

The study data reported in the following sections are taken primarily from the two questionnaires; results are reported as Manager data and Searcher data.

2.0 DESCRIPTION OF ON-LINE USER STUDY POPULATION

Who are the Users of On-Line Services? The list below shows the general composition of the responding organizations using on-line services. Managers from 472 installations characterized their parent organizations to yield the following percentages:

- 32.2% Commercial or industrial organization
- 29.0% University or college
- 17.8% Federal government agency
- 8.9% Non-profit or not-for-profit organization
- 6.8% Other
- 2.1% State/provincial government agency
- 1.5% City or county government agency
- 1.3% School district (local, intermediate, coun[illegible])
- 0.4% Junior/community college
- 0% Public library

The final rounded percentages in these 4 major groups became:

- 32% Commercial
- 31% Educational
- 21% Government
- 16% Other

Two other on-line suppliers of bibliographic data bases also elected not to participate in the study. The authors believe that the remaining group represented at the time of the study—and still does represent—the world's major suppliers of on-line literature-searching services. These ten suppliers are listed in Figure 1, along with the names of their retrieval systems. As the reader can see from this list, the on-line supplier group, like its counterparts in other areas of information services, is very diverse and includes all sectors: private/not-for-profit, private/for-profit, federal, and intergovernmental. In addition, it is world-wide.

The differences among on-line suppliers in clientele and charging policies that stem from their placement in the economic sector probably account for many of the differences among the user groups discussed in the following chapters: commercial, government, and educational. For example, about 50% to 60% of the educational institution units in the study population are health-science-related libraries that are primarily users of the National Library of Medicine service, but may also be using one or two of the commercial services. About 20% are main reference libraries, which tend to use the commercially available services. The remainder are a combination of other specialized libraries and departments in higher education and elementary/secondary education units, which also use the commercially available services. Therefore, the data from educational institutions reflect two kinds of experience: the experience of MEDLINE users, who have had a history since 1970-71 of paying only nominal fees for these services, and the experience of other users, whose use of on-line services began only in 1973-74 and has most likely been with the commercial, non-subsidized services.

The commercial organizations, at the time the data were collected, were primarily using the commercially available on-line services; as we will see from the data, they are the least experienced of the three major organizational-type users. The government agencies represent a mix of users, some of whom have been using the federally subsidized services in their areas for up to six years, and others of whom are using only, or additionally, one or more of the commercially available services.

Each of the participating suppliers was asked to help in several areas of the study: 1) to provide their user list, or to mail our questionnaires directly to their users; 2) to review questionnaires and assist in revising items developed by SDC or adding items where coverage was needed; 3) to review the preliminary data and the draft final report. The suppliers were most cooperative in all areas and contributed many hours of their time in all of these areas. The authors believe that the first feedback conference with suppliers to review the preliminary study data marked the first time that representatives from all of the suppliers had sat around the same table to discuss the on-line services.

3.0 GETTING STARTED

Once the decision is made to institute an on-line literature searching service, staff members must be assigned and trained, equipment must be acquired, and the service must be announced. (Staff selection and training are described separately in Section D.)

3.1 Terminals

We asked the managers whether the purchase or rental of a terminal had been a major barrier. About 87% of the managers said that a terminal was already available to them or that it was not difficult to buy or rent one. The problems encountered by the other managers were mainly budgetary and included resistance from upper-level management because they did not understand the value of on-line services.

We asked managers where they decided to locate the terminals. The most popular location was in a separate "terminal" room within the library or information center. Presumably these rooms assure a quiet working area for the searchers and those end-users who work with the searcher at the terminal to provide guidance and feedback.

A large majority of the managers believe that finding a proper location for the terminal is important. For some, visibility to end-users was a prime consideration. The most frequently mentioned consideration was easy access by the staff and/or the end-users. This consideration does not necessarily result in placing the terminal in one fixed, central location. One manager told us that, in his organization, four or five different individuals use the terminal for different purposes every day, but none of them needs it for more than an hour a day. To facilitate sharing the terminal, this organization has installed phone jacks in all the appropriate offices and put the terminal on wheels.

About 6% of the managers keep the terminal under lock and key because of classified-information security, or security of the terminal itself from theft, or simply to keep staff members and end-users from "playing" with the terminal and perhaps damaging it.

3.2 Promotion of Services

We asked the managers several questions about the ways in which they informed users and potential users about the new on-line service. We learned that about 83% of the managers formally announced or promoted the new service in some way. The most popular method was the use of newsletters or flyers. Special presentations or demonstrations were also reported to be heavily used, either for groups of users or for important key individuals in the organization or company. Most of the other responses were a variation on the special-presentation theme. For example, in some organizations new faculty members are given one free search. Some organizations have developed audio-visual presentations. At least one library called a press conference, which resulted in their getting free advertising on television.

All of these methods appear to be successful, but managers report that the most successful ones are sample searches and special group meetings.

A few managers have reservations about fully promoting on-line searching. They are worried that the extra workload incurred by an increase in the number of users would be too much of a burden on their staff. Some other managers, particularly those representing organizations that do not try to recover the full cost of a search, are concerned that an increase in users could not be absorbed by their present budgets.

4.0 SELECTION AND TRAINING OF STAFF

4.1 Staff Selection

We asked the managers how many persons they had selected to be on-line searchers during the first 30 days of on-line service. The numbers ranged from one searcher to as many as 40, and the average was 2.8 persons. The larger numbers were primarily in government organizations. In slightly over half of the organizations, the manager was included as one of the initial searchers.

We also asked the managers why they selected particular individuals as searchers. We gave them 8 response choices. Results of their choice are shown below.

- 75.6% Trained librarians or information specialists
- 41.1% Familiarity with one or more of the data bases through searches of the printed products
- 29.0% Familiarity with computers or automation
- 25.0% "Right" personality
- 17.8% Only staff available at the time
- 7.8% Previous use of the system subscribed to
- 6.9% Other
- 1.0% I do not know

Across all organizations and units, over 75% of the managers chose individuals who were trained librarians or information specialists. Over 40% of the managers chose individuals who were familiar with one or more data bases, through coding for batch searches or manual searches. Having the "right" personality (25%) seems to include the searchers' having good communication skills and being able to conduct a good reference interview to find out exactly what the end-user wants.

About one-fourth of the managers met resistance or apprehension among the original staff members assigned to on-line searching. Some of these staff members seemed to have a pronounced lack of confidence. They were either fearful about computers or skeptical about the results of making use of computers. In some instances, they were clearly concerned about the possibility of being replaced by the computers. Some searchers also indicated that they had initial reservations.

However, most of them now believe that on-line systems are extremely valuable and say that they enjoy on-line searching.

We also asked the managers whether some individuals were better suited to be on-line searchers than others, and why. The most frequently mentioned qualification was having a reference-service background, a data base background, or subject background in the areas in which most of the on-line searches are performed.

The second important characteristic mentioned by the managers is having a logical, thorough, analytic mind. Other characteristics of a good searcher include curiosity and a desire to learn; adaptability; comfortableness with the machine interaction process; ability to cope with instances of system failure; and patience and persistence. Some people imagine that skill at spelling and typing is important. Very few managers report this to be the case.

4.2 Staff Training

Once the manager has selected one or more staff members to be on-line searchers, these searchers must be trained. We found that approximately 55% of the searchers received formal training for at least one system from the on-line system supplier. Searchers in commercial organizations are somewhat more likely to receive formal training than searchers in other kinds of organizations. Nearly 65% of the searchers in commercial organizations received such training, compared to 55% in government organizations and 50% of those in educational organizations.

In those instances where the searcher did not receive formal training, the two learning methods that predominated were 1) studying the users' manual, and 2) getting instruction from some staff member who had been trained by the on-line supplier.

One area in which there is still much controversy has to do with the number of data bases and systems that can be mastered by the novice searcher in the early months of his or her experience. Some persons believe that learning one data base and one system well enables the searcher to perform expert searches sooner and gives the searcher the confidence needed for learning additional data bases and systems at a later time. At the other end of the spectrum are those who advocate learning several data bases, on several systems, in the early stages of the searcher's training. The assumption behind this position is that, regardless of the data base or system being examined, the similarities of on-line searching far outweigh the differences, and being exposed to two more data bases or systems simultaneously helps in the general learning process.

About one-third of the searchers in our survey learned only one data base during the first three months of on-line experience. Another third learned 2 or 3 data bases and the remaining third of the searchers learned 4 or more. We asked the searchers who had learned only one data base whether they thought that they could have learned more at that time. We also asked the other searchers whether it would have been easier or better to have become thoroughly familiar with one data base before learning others. The answer of both groups supports the position that one can and should learn several data bases, not just one, in the early stages of training.

We also asked the searchers about learning more than one system in the early stages of training. The responses show that learning more than one system is more difficult than learning more than one data base. Almost 40% of the searchers who had early experience with several systems said that it would have been better to become more familiar with their first system before learning the second. However, the other 60% of these multiple-user searchers disagreed with this idea.

These responses show that there is not yet any one best way to learn to use several on-line systems. Some searchers will want to consolidate their knowledge and develop their confidence before beginning to learn a second or third system. Others will want to learn several systems within months of each other to help develop principles that apply to or differentiate the different systems.

Earlier we mentioned the advanced training or refresher sessions that are offered by some on-line suppliers. About 39% of the searchers had participated in such advanced training sessions and we asked them to evaluate their usefulness. Nearly 96% of the searchers who had had some advanced training believed that it had been useful. They felt that it enabled them to learn about new features that have been added to the system and that it provided a forum for exchanging ideas with other searchers.

5.0 HOW THE ON-LINE SERVICES ARE BEING USED

We asked managers how many on-line searches were conducted on the average each month in their organizations. The responses ranged from 1 to over 999, the upper limit for our coding system. The median number of on-line searches performed each month by the organizations we surveyed is 30. As shown in Figure 1, there are differences in the frequency of searching among various types of organizations.

Commercial organizations do the fewest searches. About 75% of these organizations do fewer than 50 searches a month. Educational institutions do the most searching: one-third of them perform over 100 searches a month.

Organization Type	1-49	50-99	100-999	No Resp.
Government users (N=101)	45.5%	15.8%	27.7%	10.9%
Commercial users (N=152)	75.0%	10.5%	5.3%	9.2%
Educational users (N=145)	44.8%	15.9%	33.8%	5.[illegible]
Other users (N=74)	60.8%	16.2%	13.5%	9.5%

FIGURE 1. Average Number of Searches per Month

The average amount of time per search at the terminal is reported to be about 18 minutes. Detailed analysis of the data shows that educational institution users spend the least amount of time per search. The longest search times are found among users in government agencies. We believe that these differences can be attributed to several factors, including the fact that more than any other group, educational institutions tend to use only one data base in carrying out a search. This in itself would help to minimize time on the terminal. However, educational institutions also show some other interesting differences. More than any other group, they pass charges on to users, and we know that where there is a high degree of cost consciousness, searchers often elect to get off the system quickly with unscreened search hits that will be reviewed for relevance later, off-line. In contrast, searchers who do not pay directly for on-line search time are more likely to evaluate and refine searches on-line and try alternative search strategies, and they are more likely to print long search results on line. We were told of one searcher in a federal agency who, after identifying relevant material and requesting an on-line printout, went home for dinner while the terminal printed out the search results.

We believe that the relationship between type of institution and speed of searching is not a simple one. It is important to remember that more searches are done in educational institutions than in the other types of organizations. Since frequent searching undoubtedly helps to sharpen search skills, it is possible that shorter search times reflect skill levels as much as they do organizational context differences.

There are several aspects of the search process for which data are now available. One of the most interesting has to do with the process by which data bases are selected for a given search. In some instances the end-user helps to select the data base(s): 13% of the managers report that *most* of their end-users request certain data bases; another 25% report that *some* of their end-users request particular data bases. In most cases, however, the searcher makes the basic decision, selecting in advance one or more data bases that can be expected to be relevant to the inquiry.

As mentioned earlier, some types of organizations tend very much toward one-data-base searches. This is shown clearly in Figure 2.

	Most	Some	Few	None	N.R.
Government users (N=151)	39.7%	25.8%	17.9%	7.3%	9.3%
Commercial users (N=215)	30.7%	32.1%	26.0%	6.5%	4.7%
Educational users (N=205)	69.3%	14.1%	3.9%	4.9%	7.8%
Other users (N=68)	57.4%	8.8%	10.3%	10.3%	13.2%

FIGURE 2. Frequency of Use of One Data Base for One Search

The group that most frequently relies on only one data base for a search is educational users. Even when there is no success on the first data base, they tend not to search other data bases, a practice that we believe is associated with the fact that about 50% of the educational users are using a system with one primary data base (MEDLINE).

	Most	Some	Few	None	N.R.
Government users (N=151)	17.2%	31.8%	23.2%	18.5%	9.3%
Commercial users (N=215)	15.8%	39.5%	26.0%	12.1%	6.5%
Educational users (N=205)	5.4%	32.2%	36.6%	16.6%	9.3%
Other users (N=68)	7.4%	29.4%	35.3%	14.7%	13.2%

FIGURE 3. Frequency of Use of Second Data Base When No Success on First

A second interesting aspect of on-line system use has to do with the way in which the end-user interacts with the retrieval system when he or she wants a search performed. In the traditional library or information center setting, the end-user discusses his or her search problem with the person who will conduct the search, or provides this person with a written statement of the problem, but is not further involved in the actual search process. Although this traditional interaction mode has been carried on by most organizations into their on-line service activities, other alternatives are also being tried.

The percentages below show how often various alternative methods are used.

	SEARCHES	
	Most or Many	Few or None
Preterminal Interaction (Verbal or Written)	79%	17%
User at Terminal with Searcher	45%	48%
User in Contact with Searcher, but not at Terminal	7%	80%
User Performs Search for Self	10%	78%

FIGURE 4. Alternative Searcher/User Interactions (Managers, N=472)

The top line represents the traditional interaction method, where the user provides a written statement and/or discusses the problem with the searcher. Managers report that *most* of their searches are done in this way. The bottom line represents the other end of the continuum. As can be seen, end-users are performing searches for themselves in only a very small percentage of the organizations.

These data are particularly interesting because they point to a second major technique that 18% of the organizations use for *most* of their searches and another 26% use for *many* of their searches. In this approach, the user works with the searcher at the terminal, while the search is in progress, taking full advantage of both the end-user's subject expertise and the searcher's system- and literature-searching skills. Many searchers feel that this approach is efficient and productive. As one searcher put it, "It lets the user narrow his search and doesn't require his being so specific and articulate in stating his problem initially." However, some searchers do not like to have the end-user present. It can be distracting and it makes them feel uncomfortable. One searcher provided this "quotable" objection: "The system behaves poorly in front of users but is always fine when I am alone."

There is one other aspect of system use that merits discussion; it has to do with the emotionally laden issue of whether end-users should be doing on-line searches. As mentioned earlier, fewer than 20% of the searchers in our study are end-users. We asked the librarians and information specialists whether end-users can do on-line searches in a cost-effective manner. Most of them said no. In contrast, almost all of the end-users believe that they can.

6.0 COSTS OF ON-LINE SEARCHING

We asked managers to estimate the average cost of on-line searches conducted across all systems and data bases. As the list below shows, the most frequently reported costs were in the $1-$10 range. The mean cost reported, across all systems and data bases, was $23.83.

26.5%	$ 1.00 - 9.00
20.9%	$10.00 - 19.00
15.0%	$20.00 - 29.00
8.1%	$30.00 - 39.00
5.9%	$40.00 - 49.00
12.8%	$50.00 - 99.00
10.6%	No response

There are several reasons why the reported average cannot be viewed as a completely accurate measure of costs.

- *Many organizations do not keep records of search times or costs.* Only two-thirds of the organizations involved in our survey provided us with data.
- *The average includes both federally subsidized on-line services and the computer-connect-time-based charges of the commercial services.* It is undoubtedly high for the subsidized services and low for the commercial services, but we do not know the relative proportions of searches from each kind of service involved.
- *The costs for on-line services have risen in the year* between the completion of the questionnaires and the writing of the final report. On the other hand, search efficiency has probably improved. The extent to which these have offset each other is not known.
- *Organizations do not generally include all costs in their estimates.* Most of the organizations we surveyed included the costs of off-line printing, computer-connect time, and communications, but only half of them included staff time or terminal costs, and only 15% included overhead.

- *Finally, there may be differences in the ways that the respondents defined a "search."* The project staff and most users think of an on-line search as encompassing all of the time and effort at the terminal necessary to fulfill an end-user's information request. After the questionnaire was developed, the authors discovered that some users tend to think of a search in more limited terms, such as only one terminal session or as a single formulation run against a single data base. Such an interpretation may be particularly prevalent in commercial information service centers, which charge users on the basis of per-data-base searches. It is not possible to determine from this survey data how frequently the more limited definition of search was used by respondents, but the reader should be aware of the possibility that the average per-search times and per-search costs may underestimate the true values.

The study also examined attitudes toward, and practices in cost-recovery. Managers were asked whether they charge end-users for searches and, if so, how. More than half of the organizations charge end-users and, for the majority of these, the charge is variable.

The organizations were asked whether on-line services should pay their own way in an organization. Opinion was about equally divided: 42% said yes, 37% said no, and 17% are still undecided. These results reflect the ambivalence in the library community on whether information services should be placed on a cost-recovery basis.

There are differences among the various types of organizations in their attitude about user charges and in their user-charging practices. More than any other group, educational users are against user charges. Yet, when we asked whether they charge their users, we learned that nearly 77% of the educational institutions have instituted user fees for on-line searches.

7.0 SYSTEM PREFERENCES

About three-fourths of the 339 multiple-system searchers reported that they do *not* use each system an equal amount of time. A few of their reasons for preferring one system over another are shown in Figure 5.

Additional analyses of these data suggest that the criteria for system preference vary somewhat by levels of user experience. For example, one of the more frequently given choices of new users (1 to 6 months experience) was "It uses an interactive language that is easier to understand and remember," while the more experienced (3 to 6 or more

	Percentage	Rank
Learned to use the system first; feel most comfortable with it.	45.1%	2
Offers access to more of the data bases needed.	78.0%	1
Greater range of capabilities.	30.1%	3
Uses an interactive language: easier to understand and remember.	29.3%	4

FIGURE 5. Percentages of Responses by Multiple-System Searchers: Top Reasons Why Systems Are Not Used an Equal Amount of Time

years of experience) users included "Greater range of capabilities" among their top-ranked choices.

Regardless of the number of systems that searchers use, they were asked to indicate the importance and usefulness of several system features, which were grouped into three categories: 1) search input features, 2) on-line aids, and 3) search output features. From the first group of features, the most important and most used were the Boolean operators (AND, OR, NOT), and the second-most important and used feature was the ability to combine terms with one or more operators in one instruction to the system. The least important features were word-proximity operators and relational operators. Although the combining capability was not necessarily a surprising choice, these responses run counter to arguments posed by some observers of the on-line scene who have said that on-line systems are generally being used in a very unsophisticated manner, and that simple author searches and single-term searches are more the rule than the exception.

For the second group of features—on-line aids—the feature that users considered most important was the capability to receive announcements on-line from the supplier about important system-related information. The feature that permits users to monitor elapsed time on-line was second in importance. Other important features included the ability to display on-line the alphabetical index or dictionary, or the list of related terms or hierarchical thesaurus. Among the least important and least used features were those that permit users to obtain on-line explanations of system features and to enter comments on-line to the suppliers.

For the third group—search output features—the most important and most used features were off-line printing capabilities and requesting standard or predefined print/display formats. The ability to receive citations in upper and lower case and to have a search strategy that was entered on-line run off-line later by the supplier were among the least important features.

8.0 PROBLEM AREAS FOR SEARCHERS

Several potential problem areas were explored in the two questionnaires, and the data for two of these areas are highlighted below: 1) multiple-system use; and 2) system reliability.

8.1 Multiple-System Use

One questionnaire item for searchers who use more than one system begins: "It is generally believed that using more than one system can be confusing. . . ." Their responses are shown below, first for the entire group of respondents and then, in the last three columns, by user group.

		Organizational Type		
Using More Than One System	All Multiple-System Users (N=339)	Gov't. Users (N=76)	Commer. Users (N=132)	Educ. Users (N=99)
Is a major problem and seriously affects my efficiency.	3.8%	5.3%	4.5%	2.0%
Is sometimes confusing but the problems are not insurmountable nor do they generally affect my efficiency.	47.7%	32.9%	50.0%	50.5%
Has not been a source of confusion for me.	31.3%	39.5%	30.3%	28.3%
No response	17.1%	22.4%	15.2%	19.2%

FIGURE 6. Degree of Confusion in Using Multiple Systems

Because this question is one on which many have already formed strong opinions, we call attention to the commercial user group data, in particular. These users are most likely to be using systems from different "families," i.e., the commercially available systems that have different interface languages and different repertoires of features and capabilities.

The 175 searchers who reported some or considerable confusion in using more than one system were asked to indicate the areas in which they were having the most difficulty. They reported moderate to great confusion in the following areas:

- 54.3%: relating capabilities or features with the correct systems;
- 46.9%: procedures for issuing printing instructions;
- 44.0%: procedures for entering searches.

Since this study was conducted, at least one user aid has been developed to show the repertoire of command names used on the various major systems. Additional aids are under preparation at this time, and the question will need to be pursued over the coming years as to how the remaining sources of confusion can best be minimized and eliminated without reducing systems to their lowest common denominator of features and performance. Although some confusion in multiple-system usage is evident, some users also report that learning one system actually helps in the learning of additional systems. As one searcher put it, much of the operation of the various systems is identical in theory, if not in form.

8.2 System Reliability

There are many links in the chain of access to a supplier's computer, including the people, the equipment, and the software. These links start with the user, his or her terminal, the characteristics of the room in which the terminal resides, the telephone, the coupler or modem, and in some cases, the switchboard. The chain moves through the telephone company and its equipment, and through the long distance lines, sometimes via a communications network with its many personnel, hardware and software, and telephone lines. Once the link is made to the supplier's computer, additional components are added to this system: the supplier's staff, teleprocessing hardware and software, and, finally, the many components of the supplier's computer and retrieval system. How reliable are some of these key system components? According to the 801 searchers in the study, most of these components are demonstrating high reliability. For example, only about 4% of the searchers reported that they have frequent problems with their terminals; better than 75% rarely or never encounter terminal difficulties. Some of the other data on problem areas are shown in Figure 7.

	Frequent Problems	Occasional Problems	Rarely/Never Problems	No Response
With supplier's computer system	13.1%	37.8%	36.9%	12.2%
Intermittent transmission of garbage characters	12.2%	21.1%	60.3%	6.4%
Disconnection from host computer	10.7%	36.5%	45.8%	7.0%

FIGURE 7. Reliability of System Components

In many instances, the symptom that a user experiences at a terminal may be traceable to any one of several of these components, which makes the diagnosis of problems by users fairly difficult. However, only 4% of the searchers reported having difficulties in diagnosing what appear to be system-related problems; over 50% rarely or never have problems.

All is not perfect, of course, and the searchers' responses in the open-ended questions suggest that system-response problems—where the cause of the delay is not clear—and getting and staying connected to the supplier's computer are the major areas of concern. A few searchers described the "finger-pointing dilemma" that they sometimes encounter in trying to diagnose problems. As one searcher put it: "Have checked the terminal, the coupler, the phone, phone lines. Tymshare and suppliers all say their segment is working fine." Although this searcher did not mention contacting the on-line supplier, we believe that the on-line suppliers can often be helpful in coordinating with other suppliers to diagnose problems.

9.0 IMPACT OF ON-LINE SERVICES

We first asked what benefits the organizations expected when they started using on-line services. The list below indicates the managers' responses to a number of choices provided in the questionnaire. Later we asked the managers whether all or most of these benefits had been realized. Nearly 92% of them said Yes. Responding to the question, What did your organization initially anticipate would be the *main* benefit of using on-line systems?", 472 managers responded as follows:

- 72.0% Faster turnaround time
- 67.5% Access to additional sources of information
- 56.9% Reduction of staff time
- 47.2% Greater precision in searching
- 44.4% Serve more users
- 43.0% Capability for literature-searching services
- 5.0% Other
- 2.9% Didn't know

In commenting on their success with on-line searching, managers most often noted a reduction in search time, the ability to serve more users, and access to information previously not available. They also mentioned increased user satisfaction. One manager reported that his organization, which had previously provided only minimal literature searching services, was now doing 200 to 300 searches per month. Another manager wrote that, in the days when all of their searching was done manually, literature-searching services were available only to a few faculty members and administrative personnel. With the use of the on-line system, this organization can now provide service to all students and staff, as well as to outside hospital personnel. Other managers reported similar changes.

With respect to increased access to information, our survey responses showed that more information is now available to users because the staff now knows what is available and where to look for it. In addition, on-line searching has permitted organizations to perform literature searches of a depth and quality that were not possible—or, at least, were not practical—with manual searching.

Both managers and searchers were asked about the contribution of on-line searching to staff productivity. We posed this statement: The productivity of staff is greatly increased by the use of on-line services. Over 75% of the managers agreed or strongly agreed with the statement, and only about 11% disagreed. The results were even more positive from the searchers. About 77% agreed with the statement and only 7% disagreed. Further analysis of the data shows that the more searches an organization does, the more likely it is to report that staff productivity has increased.

We asked the managers whether there had been any significant changes in staff attitudes or morale because of the introduction of on-line searching. Nearly half of the managers reported significant changes, and most of these changes were positive. The reason mentioned most often for such changes was what the managers perceived as an increased sense of professionalism among the staff. Some managers reported that the change in attitude stemmed from increased interest in and enthusiasm toward literature searching, because the on-line system had taken much of the drudgery and tedium out of searching.

A few managers—15 out of 200—had complaints about on-line services. The most often stated complaint—and one that, interestingly enough, actually attests to the success of on-line searching—was that the on-line system had proved so popular with the clientele that the staff had been deluged with requests for on-line searches. Unfortunately, the size of the staff had not grown to offset the increase in the use of the facility and the same staff was faced with more work.

We asked searchers to compare the quality of searches obtained from an on-line system with searches performed manually or by batch processing. The responses were predominantly in favor of on-line searches. We also asked about the cost-effectiveness of on-line searching, in comparison with both manual and batch searching. Both searchers and managers responded in favor of on-line searching.

Many of the searchers and managers did not respond to the questions asking for comparisons of on-line with other types of searching. A major reason for this is that at least 24% of the organizations in our survey had never done any kind of literature searching before. Thus they had no basis for comparison. It is obvious that, when we stop to consider what the impact of on-line searching has been, we must include the introduction of literature search services in organizations where they have never existed before.

What about user reactions to on-line searching? We asked the managers about end-user reactions, compared to those for the literature searching services previously available. Of those managers who were able to make a comparison, over 65% reported users' reactions to this new literature-searching tool to be favorable. Only 3% reported unfavorable reactions. The managers also reported some interesting changes on the part of their users: Users have come to expect faster service from the library or information center, and some users are irritated when manual searches cannot be provided as quickly as computer searches. As one harassed manager put it, "The users seem to think that we are automated in everything."

One interesting comment on user satisfaction came from a manager who said that engineers, scientists, and researchers more readily accept the results of on-line literature searching than they do the results of manual searching, even though the quality of the search is obviously still determined by the adequacy of the search prescription formulated by the librarian or information specialist.

The major signs of the acceptance of on-line services are not simply the expressions of satisfaction but the indirect measures of success. We asked managers this question: How has the on-line service affected the number of users of your information services? Over half of the organizations (52%) reported that the number of users has increased. In 17% of the cases, the on-line service allowed the organization to serve more users, even though it was not the only reason that more users were being served. In only 15% of the cases was there no effect on the number of users.

An increase in end users naturally means an increase in the number of search requests that a library must be prepared to process. Searchers in one organization saw a ten-fold increase in the number of search requests over a 12-month period after the on-line service was instituted. Although the experience of this organization may not be typical, many managers did claim impressive gains in the number of searches being performed for their users.

We asked the searchers to assess the impact of on-line searching on their workload and their productivity. For 49% of the searchers, their workload increased with this new service. Interestingly, the increase in workload does not necessarily mean that new staff must be added. Presumably some of the increase is offset by greater staff productivity. About 76% of the managers in the survey reported that the increased speed and efficiency of on-line searching over other modes of searching increased staff productivity. Only 22% of the managers believe that the introduction of on-line services requires adding new personnel to the staff. As might be expected, the managers who believe additional personnel are required tend to work in organizations that perform a large number of searches. In those organizations that perform fewer than 20 on-line searches a month, only 16% of the managers feel the need for additional personnel.

Many managers indicated that the on-line service caused an increase in the use of other resources. For example, in nearly 56% of the organizations there has been an increase in the number of requests for full-text copies of documents. This increase is, in turn, having an impact on acquisitions budgets and on interlibrary loan requests.

10. SUMMARY

This summary report has touched on some of the highlights of the study of the impact of on-line bibliographic retrieval systems. Such systems have been in existence for nearly 15 years, but only in the past

5 years has their full potential begun to be realized. What the SDC study has shown is that these systems are highly regarded by those that have used them and that, in spite of growing pains stemming from as-yet-imperfect technology, on-line searching is growing in importance and usefulness.

On-line services are having a profound impact on library/information reference service, not only in terms of the speed and comprehensiveness of service, and the improved self-image and morale of the information professionals involved, but also in terms of the philosophy of service. Many institutions that have begun using on-line services have had to modify or compromise a free-service philosophy to cope with an unexpectedly high demand for on-line service, a demand that has strained or overwhelmed existing budgets and required either the curtailment of service or the imposition of user charges.

Most of the current users of on-line bibliographic retrieval services are information intermediaries, but the number of scientists, technicians, businesspersons, and others who are doing their own searching is growing, and this growth presages a vastly larger community of users than we see today. The SDC study has provided a wealth of data to help information professionals and users take fuller advantage of this new technology, and to help data base developers and suppliers of the on-line services continue providing a responsive service. For information science researchers, the study provides insights into issues that warrant further examination. The importance of this technology to the library and information services community has been clearly established; responsibility for the further development of this technology and the maximizing of its potential rests with all of these audiences.

Acknowledgements

The SDC project staff would like to acknowledge the support and assistance of the many persons who helped us throughout this project. Some of them are identified in the Final Report. We would like particularly to use this opportunity to express our appreciation to the main contributors to the study: the many current users of the on-line systems who invested a great deal of their time to complete the lengthy questionnaires and were so willing to participate in our on-site interviews.

The on-line suppliers whose users we surveyed also gave much of their time and effort at several stages during the study. We would like to thank all of our key contacts in these organizations: Energy Research and Development Administration; Battelle Memorial Institute; Canadian Institute for Scientific and Technical Information; Defense Documentation Center; European Space Agency; Lockheed Missiles and Space Company; National Aeronautics and Space Administration; National Library of Medicine; and State University of New York.

The views expressed in this report are solely those of the authors. The conclusions, opinions, and recommendations do not necessarily reflect the views or position of the National Science Foundation, nor of the other participating on-line suppliers.

Chapter 11

by George Ember

Reader Services of CISTI

1.0 INTRODUCTION

One of the consequences of large scale library automation, computer networking and mechanized information retrieval is the sad fact that it becomes increasingly difficult to isolate and talk meaningfully about a *single* level or type of service. Very often this represents only a component of an interrelated system which supports and incorporates the particular service as a module, a sub-system or perhaps a purposeful assembly of these.

This paper will reduce the many facets and activities of a national system into the direct functions of reader or user services. Only brief references are made to the grand design, the setting and the organizational aspects of the Canada Institute for Scientific and Technical Information (CISTI) which offers these services to Canadians. CISTI is a division of the National Research Council of Canada that was created in October 1974 by the merger of two organizations: the National Science Library, which has existed since 1924, and the Technical Information Service, formed in 1945. These two unified organizations operate as the central node in the evolving Canadian scientific and technical information network. CISTI maintains an interface with various types of end-users and also reinforces the service capabilities of other Canadian libraries and information centres. Thus, CISTI is now retailer, wholesaler, importer, and, perhaps in the future, exporter of information as well.

2.0 CISTI SERVICES

Reader services, in the traditional sense, are insignificant in the CISTI organizational environment. There are few people in the reading areas or browsing in the open stacks where some one million volumes are held and 18,500 serial titles collected. Researchers of the National Research Council laboratories, of the two local universities, and of industrial firms borrow volumes and request information from the Reference Department, but the bulk of the services is provided to individuals and organizations which are located outside of Ottawa. For example the Interlibrary Loan Department receives a daily average of 400-600 requests for photocopies, monograph loans and microfiche blow-ups.

A large portion of the lending and duplicating transactions back up, with hard-copy literature, the retrospective or alerting information services which CISTI provides nationally. The collection building and acquisition policy is directed toward the subscription of journals and the purchase of monographs which are not held elsewhere in the country. The collection is checked quarterly for 1-week test periods to rank-order the usage of periodicals and to list the disciplinary subjects in order of usage frequency. Once each year periodicals usage is evaluated by plotting the titles per subject against the requests per subject to see whether supply and demand are harmonized.

The retrospective and alerting information services, which are backed up by the collection, can be divided into three categories. All of these are computer-based and operate in a network fashion. Two of them—an on-line enquiry service and a selective dissemination of information system—were created and are maintained by the National Research Council. The third functions by hooking up Canadian users, via a telecommunication network, with a system in the United States.

3.0 CANADIAN ON-LINE ENQUIRY SYSTEM

The first, CAN/OLE (Canadian On-Line Enquiry System) has been in operation since March 1974. It was designed to provide retrospective searching of large bibliographic reference files in all major fields of science and technology. In 1976 over four million references of six data bases can be accessed through approximately 240 terminals across the country: *Biological Abstracts; Chemical Abstracts; Engineering Index;* the British *INSPEC* service covering physics, electrical and electronic engineering, computer and control; the *NTIS* data base; and the *Union List of Scientific Serials in Canadian Libraries.* Is CAN/OLE a unique system? To some extent, it is unique. The user does not get only references to published papers when s/he is searching the various files. S/he can also switch immediately to the Union List, one of the available data bases, to find out which of the 250 science libraries holds the required serial title. S/he will know whether the journal is held in the library of his or her own organization and, if not, which is the closest holding library. All libraries participating in the Union List project of CISTI are equipped, and willing, to provide lending or reprography services. Another unique feature of CAN/OLE is its bilingual, English and French, interactive conversation capability. The third is a supporting communication system by which CAN/OLE can be reached from 11 large cities of Canada through a self controlled communication network. The cities are Vancouver, Calgary, Regina, Winnipeg, Toronto, Kitchener, Ottawa, Montreal, Quebec, Halifax and St. John's in Newfoundland. CISTI control of the communication network enables the Institute to offer CAN/OLE on an "equalized rate basis" so that access to the system costs $40 per connect hour across the country; in other words, the charge for OLE services is the same in Vancouver or Halifax as it is in Ottawa where the system is operated.

CAN/OLE is a growing system both in terms of users and of searchable files. The next data base scheduled to be added is a modest Canadian version of the Smithsonian's Science Information Exchange, SSIE. The file is produced by the Information Exchange Centre of CISTI which has a strong liaison with the 31 granting agencies of the federal government. According to a Cabinet mandate, these funding bodies supply CISTI with a fact-sheet of each research award they give to academic institutions, including fiscal data and a brief abstract of the project. Consequently, the IEC file contains roughly 10,000 entries annually plus approximately 30,000 project descriptions referring back to the last three years. CAN/OLE is not only a bibliographic search system, but also a location tool and a national directory of subject expertise.

As an unconventional reader service CAN/OLE is very significant. Training courses for users are held monthly, and the link between local terminal operators and CISTI is rather strong. Just recently the Institute conducted an extensive national user survey which supplied invaluable information on the overall reaction of subscribers; the findings of this study will assist us in further refinement and extension of CAN/OLE.

4.0 CANADIAN SELECTIVE DISSEMINATION OF INFORMATION SYSTEM

The second large system I wish to introduce is CAN/SDI, Canadian Selective Dissemination of Information, which, in batch format, serves some 2100 subscribers with personalized bibliographies. At the moment 15 data bases are utilized; BA Previews, CAIN, Chemical Titles, Chemical Abstract Condensates, COMPENDEX, ERIC, Geo. Ref, Government Reports Announcements, INSPEC, the MARC II of the Library of Congress, METADEX, Psychological Abstracts, and tapes produced by the Institute for Scientific Information in Philadelphia. Some time ago the Institute attempted to explore in a nationwide study how many people were benefiting from CAN/SDI subscriptions on a regular basis. Staff learned that, on the average, 2.8 people participated in the formulation of a search profile. This is the number that is directly served by the output. This result means that the present 2100 subscriptions are covering a user population of roughly 6,000.

Since the accuracy of the search profile (a list of keywords arranged in a Boolean formula) is the pre-requisite of relevant output, CISTI has made special efforts since April 1969, when CAN/SDI became a national service, to improve and maintain a high quality of the individual search strategies. The system-user interface rests on two levels. The Institute continuously trains users to produce their own profiles and also organizes monthly seminars in Ottawa and elsewhere for search editors who keep direct contact with the scientists and researchers whom they serve. Thus, over 90% of all search profiles are formulated and monitored either by the ~~end-users or by their~~ appointed co-workers or librarians. With scientific and technical subscribers CISTI has established a direct interface. However, some of the data bases are outside the natural sciences and engineering, and CISTI staff does not possess adequate subject expertise and terminological competence to support them intellectually. Therefore, CAN/SDI is offered also: through the National Library for users in the social sciences and humanities; through the Canada Centre for Mineral and Energy Technology for users in the mining technology and mineral sciences; through the Geological Survey of Canada for geologists and earth scientists; and through Agriculture Canada for the agricultural communities. This indirectly offered portion of the service is supported by CISTI only in a technical sense (the Institute runs and searches the data bases). All user contacts are maintained by the participating federal departments and agencies.

Beyond its national significance the CAN/SDI system has become an internationally utilized mode of service within the UNESCO/UNISIST programme. CISTI is assisting UNESCO ~~member~~ states to establish national SDI services modelled after CAN/SDI. Assistance consists primarily of providing, without cost, the complete CAN/SDI software package, staff expertise and on-site training of operating staff. The recipient country, with the help of UNESCO funds, is responsible for sending appropriate staff to CISTI for a minimum three-month training period and for ensuring that appropriate equipment and resources are available in their country. To date, the Institute has assisted in developing national SDI services in Australia, South Africa, Argentina, The Netherlands, India, Mexico and the United Kingdom.

5.0 AN ON-LINE MEDICAL INFORMATION PACKAGE

The third category of these national systems consists of an on-line package which Canada receives directly from the National Library of Medicine in Bethesda. The package includes MEDLINE, TOXLINE, CHEMLINE, CATLINE and the smaller standard files. The bilateral agreement between NLM and CISTI placed the whole Canadian service under the Institute's administrative and fiscal supervision. Thus CISTI provides the funds to cover the cost of indexed input into MEDLINE which is a Canadian responsibility in the *quid-pro-quo* arrangement. In addition, CISTI provides user and operator training and monitors the connect hour use of each MEDLINE centre. At the moment 37 Canadian MEDLINE centres have been established in 16 medical schools, in governmental and industrial research organizations, and in the CISTI Health Sciences Resource Centre. The CISTI centre offers not only on-line search services but also SDI services in a batch mode for 93 Canadian medical researchers. The charge for MEDLINE is $28 per connect hour which includes communication charges through the Telenet and TYMSHARE networks. Both of these networks have access nodes in most Canadian cities.

6.0 CISTI CHARGING POLICY

CISTI is maintained by public funds and a whole range of conventional library services are provided free to the public. For example, the Institute does not charge for lending monographs or providing reference services to the user. However, it does charge for photocopy processing and for related staff-cost in reprography. The charging rate for photo duplication is 22¢ per page with a minimum of $2.20. The standard rate for CAN/OLE is $40 per connect hour which includes communication cost. CISTI does not retrieve the cost of data base storage, file maintenance and staff salaries. CAN/OLE users are charged for the processing cost, stationery, postage and royalties. The same scheme, by which CISTI absorbs the fixed cost and retrieves the variable cost, applies to CAN/SDI as well. The Institute does not recover staff salaries and data base cost. However, SDI users are charged for processing their recurrent searches, stationery, postage and data base royalties.

Chapter 12

by Harriet Velazquez

Computer Networking and Reader Services

1.0 THE REFERENCE LIBRARIAN'S ROLE— LEVEL ONE

Reference librarians at the University of Toronto Library are working on two levels with computer based bibliographic files. Each level presents a wide range of problems and demands a variety of skills from staff members.

On the first level, librarians are users of both commercially available data bases, through such retrieval systems as DIALOG and ORBIT, and a very large machine-readable file of the University of Toronto's library's holdings. The Reference Department has phased the use of machine-readable files into the existing manual mode of operation. Reference librarians regard these data bases as additional, or alternative, reference sources for library researchers, and they advise users regarding the suitability of files as they would for manual sources, based on the library user's requirements, content of the file, and available access methods. Whether the researcher ultimately looks for materials in a published bibliography, printed periodical index, library catalogue, on-line data base, or a combination of these, s/he still requires that the reference librarians be able to advise him or her based on their knowledge of the available sources and, most importantly, their ability to help him or her define information requirements.

This level also requires that reference librarians do a considerable amount of what, in other kinds of institutions, might be called marketing. Besides explaining the use and potential of automated indices and catalogues at the subject oriented bibliographic seminars that they teach to students, they feel the need to actively publicize the new ways to find information. Interestingly, success in attracting people to the new technology has made it necessary for the reference librarians at U of T to do quite a bit more general library orientation.

2.0 THE REFERENCE LIBRARIAN'S ROLE— LEVEL TWO

The second level of reference staff involvement with computer based systems is participation in the design of products and services from the University of Toronto Library's computer based catalogue. The UTLAS Catalogue is a machine-readable file in MARC format of about 1.2 million records, representing most of the library's collection. Primary access to this data base is currently through computer output microfiche and microfilm. As of July 1976, the University of Toronto Library "closed" its card catalogue, and the microform alternate is now the only on-going record of holdings for library users. Reference staff participated in the design of this alternate which differs in form from the traditional card catalogue. It consists of short index entries (authors, titles, and subjects) pointing to a full bibliographic record sorted by call number. The micro-catalogue is now available in over fifty locations on campus, providing users with forty-nine extra places from which to access the total collection.

Reference librarians have also experimented with on-line access to a shortened form of each record in the file through a browseable call number index. In both the on-line and microform projects, reference librarians have worked closely with technical services and systems staff. They expect to continue this working arrangement as the librarians monitor user reaction to the micro-catalogues and to an expansion of on-line access in 1977.

3.0 CONCLUSION

Reference librarians at the University of Toronto Library have had, over the past two years, a rather unique opportunity to become familiar with automated bibliographic systems, not only as users, but also as innovators in systems design and application. This has already shown positive results for the library in the relative ease with which the COM catalogue has gained acceptance by library staff, and users, and in the on-going constructive dialogue which has been established between reference staff and the technical services section of the library. Most important for library users, reference librarians are building the expertise necessary to exploit computer based bibliographic files to their full potential.

Chapter 13

by Ralph E. Stierwalt

Ontario Universities' Library Cooperative Systems

1.0 THE ONTARIO UNIVERSITIES LIBRARY COOPERATIVE SYSTEMS ORGANIZATION

There were several principles established by the first members of the various projects in the Ontario Universities' Library Cooperative Systems (OULCS) organization. These were and in 1976 still are:

- Adherence to a set of standards for each project developed by the members;
- Requirement that all cooperative projects must have the potential for being cost beneficial;
- Desire to integrate any cooperative system with existing work procedures and routines or replace procedures and routines used in the earlier services and/or processing procedures;
- Requirement that each project belong to the members of the project;
- Equitable sharing of certain costs by the members in order that an institution should not be penalized because of geographic location or size. Some of these costs are for lines, terminals, membership fees, Office of Library Coordination support, etc.;
- Maintenance of open membership in all of the computer-based projects for all kinds of libraries in Quebec and Ontario. In fact, non-university libraries should be encouraged to join;
- Coordination through the Office of Library Coordination whose authorities and powers are granted to it by the members of the various projects;

- Reliance upon the goal of cooperative development only until it becomes ineffective. Cooperation should not spread poverty on a wider scale;
- Purchase, rental or adaptation of systems, documents and techniques required by the projects whenever possible.

The four major goals of OULCS are listed below in priority order:

1. To provide improved information services to users of libraries.
2. To provide the means for the greater and more effective sharing of research resources among libraries.
3. To develop the means for and promote collection development rationalization.
4. To design, develop, operate and utilize cooperative library support systems which will assist the members of the various projects to reach goals 1 through 3 above.

To a large extent, the goals and principles of OULCS have determined the group's approaches to managing the cooperative projects and evaluating the use of the various systems. The Office of Library Coordination is responsible for gathering and analyzing use and cost data and preparing reports for the members. The Office is also responsible for interpreting data for individual members and the membership at large.

2.0 MANAGING SYSTEMS

2.1 Computer Based Systems

Much of the information essential for helping one operate efficient and economic automated systems is of use to the individual user as well as the group of users. For example, useful information includes but is not limited to:

- connect time
- c.p.u. seconds
- page milliseconds (core utilization)
- disk accesses
- number of records displayed
- number of records derived
- number of titles added
- number of records processed

- number of terminals on-line and at what time during the day for the individual user and for a group of users
- growth of the union file
  how many derived records; source?
  how many original records; source?
  how many union file records used?
- how many record processing commands?
  hold
  file
  store
- how many kits (spine and book pocket labels and punched circulation cards) required per title? per record?
- the percentage of connect time used for searching
- the percentage of connect time used for cataloguing
- the length of time it took "x" institution's terminal operator "c" to input an original record; a derived record; to use a union record.

2.2 Application of System Use Data to In-House Management

Many of the members of the various OULCS projects were relatively unaccustomed to computer-based systems in the beginning. Most of the first members spent a great deal of time learning about how computers operate and how the planned systems were likely to perform. Since all staff likely to be affected by or have an effect on the computer-based system were involved in the learning process, this learning took place at every level of staff in the library. Observation and analysis make it clear that those people who admitted that they knew nothing learned the most and have made the most efficient and economic use of the computer-based systems. A careful study of the machine and its capabilities, and of the software systems and how they work, is essential to organizing a library staff to take advantage of the machine capabilities. When the computer does everything it can do, people are free to use the capabilities they have, like brains, that the machinery doesn't have. A sound knowledge of the hardware and software also enables the members to advise developers on how the software can be improved to perform better and to provide more services for them.

The members also require regular management reports in order to manage effectively. These reports are reviewed and used by staff at all levels including the terminal operator, the Chief Librarian, the cataloguer, the bibliographic checker, the division head, and so on. Although it sometimes indicates inefficiency, people are not threatened by the data. OULCS staff have tried to introduce a spirit of challenge (man over machine), individual management of one's work and participation in planning, budgeting and decision making.

2.3 Systems in General

It is important to point out that responsibility for efficient and economic use of the systems rests with the OULCS individual members. Neither the members nor the office staff want any higher central office or agency telling them the what, how, when and why. Adherence to this principle assures that each member will have quite a bit of flexibility in deciding how the systems are used and which aspects of the systems will be used. However, peer group pressure is brought to bear on those members who abuse facets of the system that affect the economy or efficiency of the group, such as non-adherence to cataloguing or format standards, or misuse of computer resources, which may drive unit costs up for the entire membership.

The Office of Library Coordination is authorized by the members to undertake in-house surveys of member libraries. These surveys record organization, workflow and management, techniques. In a wrap-up session, always held with the library's management team after the survey, Office personnel emphasize both the good points and the problems and offer advice on how the problems may be corrected. The member library is not obliged to accept or use the advice. After a report of the survey is written and flow charts are prepared, the report and charts are approved by the library and then distributed to all members of the project for their information. All of the members have learned some useful tips from each other which have been introduced into the individual's library organization, workflow plan, or management techniques. Many of the members also consult the Office of Library Coordination for advice on problem solving, system expansion, etc.

2.4 New Members

Prospective new members receive advance documentation, are encouraged to visit member libraries to observe the particular system in operation, and receive a visit from Consortium Office staff. Often a member or members of the project visit the potential member campus to give a full explanation of the particular system and a survey of the library system. Obvious potential problems are pointed out during the wrap-up session. A report of the survey and a cost picture are prepared for the prospective member and for the membership. The survey team then recommends to the membership whether or not the prospective member should be accepted into the project. If the membership approves acceptance, on-site training in the new member's library is arranged. Cataloguers or coders are trained in the coding techniques of the system, and terminal operators are taught to use the CRT and printer.

It took the first members in the UNICAT/TELECAT Project (*Union Catalogue System*) approximately eight to ten months to learn to use the system efficiently. In November 1976, it now takes new members about two months to reach an acceptable level of efficiency.

Management data are used extensively to explain the systems and teach people to use the systems to the best advantage.

2.5 The Information Industry

All of OULCS major projects are self-supporting. The consortium does not have grants from governments, foundations or agencies to subsidize its work or the use of any of the systems. Therefore members are encouraged to be efficient and to discover ways of doing library work and operating systems economically.

The members of the OULCS projects subscribe to the principle that information, and the many products necessary to gather it, record it, and preserve it, have commercial value. OULCS members sell everything they can in order to acquire funds for research and development, and to offset fees for the use and/or the operation of systems. Although acceptance of this fact works to the members' advantage in several ways, perhaps the most important influence is that people are encouraged to do better work because of the sales potential.

Items for sale include:

- union lists in print form and on microform;
- coding, cataloguing, and procedural manuals;
- cataloguing, editing, and bibliographic standards;
- bibliographic records; and
- consulting services.

Items available for lease include:

- software packages; and
- machine readable data bases for use in information systems.

On occasion, OULCS will trade products when members of the participant projects are satisfied that what the consortium receives in return will be of benefit to the members of the project. The OULCS consortia recognize that there is still much to be learned about cooperative development and reaching goals of efficiency and economy. One of the main concerns is that we, collectively, remain flexible, keep analyzing, and continue to find solutions to problems in order to reach the group's goals.

Chapter 14

by Gordon H. Wright

Management Once Enmeshed

1.0 INTRODUCTION

"In reviewing 1876, one senses an excited gathering of librarians' concerns and an excited move toward co-operation in dealing with them. The need for cooperation today is at once grimmer than in 1876 and easier because of new technological support. In 1976, one feels a shimmer of excitement on the edge of new areas of co-operation, and again, librarians approaching them willing to face the hazards to each library's autonomy which change will bring."

So says Helen Tuttle in her survey "From Cutter to Computer" . . . (Coll. & Res. Lib. Sept. '76) The shimmer of excitement may be a shimmer of horror as the ability to manage wanes and the forces for economy in scale and mass production technique jockey for position. Indeed today's tragedy is that the power politics of networking is obscuring the real problems that have to be solved before effective local library service can exist within a national network.

2.0 POWER POLITICS IN LIBRARY MANAGEMENT

It isn't, of course, all due to power politics. The library manager innovator rarely has a free choice of computer or computer management. Too often, once enmeshed in the computer, innovation is blunted, changed, and redirected either through computer personnel, network directors or other participants. In the past the library innovator could distinguish and set priorities as perceived for the library, apply and adapt new techniques in-house for the mutual benefit of patrons, administrators, and themselves. Today the library innovator feels at the mercy of the computer network innovator who may have similar goals—similar but not the same! A network innovator looks to the number of *users* he can service whilst the library innovator looks to the number of *systems* he can service. While these may not be mutually exclusive, they probably are exclusive in time scale.

Unfortunately, to meet the needs of the mass market, the computer innovator has looked to the degree of sameness that can be found in order to create, develop, and exploit the product. To establish the market this naturally leads to an extension of the services beyond the immediate confines of the Region served by the computer. It is a further short step to pander to the needs of the mass market by concentrating on a cheap production. Cost reduction becomes more important than cost efficiency. Innovative products such as PRECIS, which are developed by re-structuring library requirements around the methodology of the computer, are further at risk and take even longer to implement than the average change implementation process.

However, in the functional environment, the library manager and library innovator are controlled by the degree of differences that exist both in priority and application. To generalize,

Universities have tended to consider these goals as—

- collection building, bibliography, delivery

public libraries as—

- public service, freedom of choice, simplicity of access

school as—

- curriculum integration and stock relevance.

The differences that exist between these goals must be carefully considered in network planning.

3.0 PEOPLE WHO IMPLEMENT CHANGE

It is said that people create the greatest threat to the success of co-operative networking. E. M. Rogers and F. F. Shoemaker (3) have defined the characteristics of those involved in innovation adoption. Innovators are venturesome and willing to accept risks even if they have to absorb financial loss. Since early adopters are much more an integral part of the local social system and role models for other members of the community, opinion leaders are often drawn from their ranks. The trouble is, that they are not only of higher status but are on the move in the direction of a still higher level of social status. In fact, they are more likely to use innovation as a means of getting there!

There are three other categories of persons involved in adoption of innovation: the early majority who deliberate carefully before adoption and rarely ever lead; the late majority, the skeptics willing to proceed when persuaded and then only when the weight of evidence in favor is overwhelming; and, finally the laggards who are more likely to adopt the idea once it has been superseded by another. Laggards are suspicious of innovation, innovators and change agents.

This is the conglomerate that fashions the future of any profession. This is the mix one can expect once the local innovator endeavors to involve colleagues in a network. How does one make the mix palatable? Can a network develop which facilitates the dissemination of information and achieves creativity in the process? If so what are the planning requirements?

4.0 CHANGE IN EDUCATION AND LIBRARIES

It has been perceived that innovation has a slower adoption rate in education than any other area. Libraries often function as educational institutes. Indeed, within the community they serve, they provide the cheapest form of personalized education available. Mort (2) stated that "the average American school lags 25 years behind the best practice." There are obviously a number of reasons for this observation. Perhaps the most likely reason for this behavior pattern is that teacher and librarian are more concerned with the dysfunctional aspects of the innovation. However, from a variety of studies, one conclusion has emerged. The communication process involved in implementing a decision in education is either short circuited by an authoritative decision or confused by the complexity of collective decision-making!

Since most educational institutions are formal organizations, deliberately established to achieve predetermined goals, they are characterized by prescribed roles, authority structures and members governed by rules and regulations for behavior. In this type of organization, the decision unit has more authority, whether it is individual or collective, than the adoption unit. On the other hand, the successful adoption of the change will depend on the degree of communication and trust between units. As a generalization, Rogers and Shoemaker state that both the individuals' acceptance and satisfaction towards an authoritative innovation decision are positively related to their degree of participation in innovative decision-making.

In the normal university library, the innovator experiences constraints imposed by the institution as a whole. Thus, successful innovation will be dependent upon the nature and degree of participation that has been encouraged amongst library and university staff. As stepping into a network diffuses the communication process still further, the participating role is often ignored beyond the senior management level. Worse still, because the goals of the network may not be the same as those of the institution, the collective or authoritative decision which may stem from the network may not be in the best interest of the library either to adopt or to implement at that moment of time because of the personalities or economics involved. The task of the manager, to encourage adoption, is more onerous and resistance is likely to fester and grow.

No wonder, then, that network cooperatives have been ultracautious in approach. Resource sharing has been limited to areas where only small adoption units have been necessary such as units concerned with establishing union serial holdings, resource collection sharing, union list of non-print media or in a single system development such as cataloguing. Perhaps this is the tragedy. User co-operatives have been fearful of the management consequences of their entanglement, whilst computer directed networks have been fearful that local involvement may preclude their ability to capture the mass market.

Surely the innovative concept of the computer provides enormous flexibility to the library manager in library administration and service. In the past, the manager could only arrange the resources in a standard structured array and with limited access points through card catalogues. These limitations imposed codes of practice and regulations that bore heavily on staff and patron alike. Local co-operative endeavor, based on the need to avoid needless duplication of material, depended more on personal contact than objective study. Co-operative exchange of information between institutions with similar goals depended on professional collaboration.

The computer has made nonsense of much of this. Standards may be necessary, not for filing nor for shelf arrangement in individual libraries, but for:

- codes to facilitate retrieval especially to meet the needs for higher selectivity and relevance;
- carrier communication;
- cost recovery procedures;
- utilization and exchange of data; and
- interface between commercial and public sectors using the same network.

Certainly it is time that the disparate groups—universities, colleges, public libraries, schools, government agencies, and so forth examined:

- the needs of their patrons;

prescribed

- the information requirements to meet these needs;

specified

- the systems necessary to achieve these needs; and

established

- areas for co-operative and individual endeavor;

based

- on the computer and a computer communication network.

Failure to examine the problems at the grass root level, and to appreciate the advice that can be gleaned from those who have studied the change mechanism that innovation requires, leads to power politics, and to the exploitation of the worst rather than the best in any profession. Regionally, the formation of councils, that provide the opportunity for the disparate groups to examine these needs and to establish their communication mode with each other, would do much to facilitate effective planning.

Perhaps, if this were done, one might discover that the constant search for cost economy in shared cataloguing is leading librarians into a blind alley. Perhaps innovators have left out important contributors to the network like authors, publishers and patrons. Perhaps search codes are inadequate to select relevant information from the store, and speedier transmission techniques require even greater skill and user reliance on the purveyors and communicators. As libraries have grown bigger, librarians may have lost sight of their major role as purveyors and communicators of information to the society they serve.

5.0 LOOKING TO THE FUTURE

The future is indeed grimmer in 1976 than 1876 because the tools now in our hands are more powerful than ever before. We should be concerned to see that they are applied to the common good of humanity rather than manipulated for the few. Above all we shall have to consider carefully whether to agree that, in the future, the weight of gold held by an individual in society should be the basis upon which to establish access to knowledge.

References

1. Miles, M. B., ed., *Innovation in Education*, N.Y. Columbia Univ., 1964.
2. Mort, P. R., "Educational Adaptability," *The School Executive*, 1953, 7111-23.
3. Rogers, E. M. & Shoemaker, F. F., *Communication of innovations; a cross cultural approach*, 2nd ed. N.Y. Free Press, 1971.

Chapter 15

by J. C. R. Licklider

Library Networks: Should They Deal With Containers or Contents of Knowledge?

1.0 INTRODUCTION

It is gratifying that libraries are eagerly adopting and exploiting computer and communication technology, but it is distressing to observe the heavy concentration on—indeed, the almost total preoccupation with—not knowledge itself but the apparatus of bibliographic control.

Library people are often bored when intermediating between a library card and the shelves of numbered books. They are involved and eager—the more expressive way to put it is "turned on"—when interacting not only with a user but with the meaning and structure of his or her problem and when mediating between that problem and its solution through reference to the body of knowledge. It is therefore somewhat saddening to hear, in discussions of library networks, so much about computerization of what is on the covers of books and so little about the use of computers and telecommunications to organize, store, retrieve, process, and deliver what is on the inside.

In this brief paper, therefore, the author's purpose is to direct attention toward the greater and more profound opportunity for libraries that is being made ripe by continuing advances in computers and communications: the opportunity of organizing, interrelating, and disseminating not just the containers of knowledge (books, reports, journal articles, and so on) but the contents of knowledge. Even beyond the knowledge that is strings of alphanumeric characters, librarians can work with the ideas that are at least a little closer to the "deep structure" of the meaning of publications. This opportunity is not quite imminent as the one that is now so vigorously being pursued, but it is not so very far away, either. Two randomly selected items illustrate the imminence of full text manipulation by computer: 1) Digital storage technology makes it possible to record the alphanumeric contents of one thousand books on a sheet of plastic the diameter of a phonograph record; 2) the idea of a network of aerostationary regional microwave platforms at 70,000 feet is being explored by a group of advisors to NASA. These regional stations may outdo communications satellites by opening up plentiful transmission bandwidth to inexpensive ground stations with 2-foot-dish antennas. Library people must move rather soon into the research and development required to exploit such technologies in order to deal with contents as well as containers of knowledge. If librarians do not move quickly, in a few decades, the computer center and not the library will be the dwelling place of knowledge.

2.0 FULL TEXT, DATA, AND MODELS

The first steps in the direction of content-sensitive librarianship must necessarily be to get the full text of documents, the data of data bases, and the programs of computer-program models into computers and to get the computers, functionally (via networks), or physically, into libraries. As word processing and office automation proceed, more and more documents will be created in digital form. Even though the digital records may for many years be printed for distribution, libraries should acquire or locate the digital records, and learn to use them (with the aid of computer processing) to organize the body of knowledge. Computerized data bases and computer-program models already exist in computers, of course, but computer people have not been much interested in organizing data bases and models for widespread use. That is an essential library function which is both complex and very difficult to master. It is high time that librarians reach out into computer networks to create order and functionality out of what is now chaos. (The author would estimate that 90 per cent of the computerized information in EDUCOM universities "trickles down" to back-up or dead storage tapes within two years and that less than one per cent ever "perks up" again.)

2.1 Computer Understanding of Natural Language

Natural language viewed as alphanumeric strings is a very complex and not fully understood encoding of deep-structure meaning. Nevertheless, the work of Terry Winograd at M.I.T. and Stanford, Roger Schank at Stanford and Yale, Bill Martin at M.I.T., and others at other EDUCOM universities suggests that computers can decode and understand full text well enough to penetrate into the meaning, interrelate the parts of documents, and relate the contents of one document to the contents of another. The implications of that idea for the future of libraries are revolutionary. Inevitably, if there are libraries in the distant future, they will be transforming meaning from natural language form to other forms more amenable to efficient computer processing and back again. The nearer-term question is, can somewhat shallower computer processing of the text of documents be helpful in organizing and retrieving knowledge in the meantime? Advances in artificial intelligence and computational linguistics indicate that computer processing of text will be highly exploitable by libraries long before they are prepared to exploit it.

2.2 Knowledge Bases

A knowledge base is a data base or information base in which the contents are structured and organized to support intelligent processing efficiently. The semantically interrelated parts of a knowledge base are structurally and functionally interrelated. A store of meaningful text, whether in a computer or in a book, is not a knowledge base because the linear chaining of the characters does not do justice to the rich interrelation of the ideas represented by the text. In the field of artificial intelligence, complex storage structures have been developed that do more justice to the complexity of the intellectual content of a data base. Various forms have been called "pointer structures", "relational nets", "semantic nets", "procedural nets", and so on. Such structures, and knowledge bases built on such structures, will be indispensable to librarians who want to build the library networks of the future. In short, it is an essential responsibility of librarians to learn how to exploit computer and communication technology to remove the fundamental constraint imposed upon knowledge by the technology of print on paper: the constraint that knowledge cannot be processed in the form in which it is stored.

3.0 LIBRARIES AND PUBLISHING

There is, as many have no doubt noticed but perhaps have been too polite to mention, a fundamental structural defect in the relationship between libraries and publishing. The defect has always been there, but it did not seem irrational until the arrival of networking made it obviously so. Briefly, an author creates a document, and the document is then processed for some months or years. Then it is published (perhaps concatenated with other documents, replicated many times, and distributed to way stations such as libraries and bookstores). The many copies occupy shelf space for varying lengths of time. Some of them are, in due course, borrowed or bought, and finally some parts of them—rarely all, except in the case of fiction—are read. The early months or years of a document, when it would be most useful to readers, are thus spent in a state that is inaccessible except to editors and referees. When the document is replicated, it is put into a state in which it is intrinsically expensive to store and distribute. The cost of storing and distributing the document is far greater than its actual use warrants.

When documents are created in a computer network, they are, in principle, as available as their authors want them to be even while they are being created. For example, Marvin Minsky recently spent one year writing an epochal and influential paper on "frames" which are sophisticated abstract information structures related to semantic nets and knowledge bases. Word spread that the paper was in process, available, and significant. Through the ARPA Network, dozens of interested colleagues read and commented on one or more versions of the paper, and influenced subsequent versions, while the paper was in preparation. Comments were made while the paper was a dynamic, living, thing while it was most important for critics to read it. The paper has been referenced many times in the past year in papers that are not yet "out" in journal form. "Frames" will appear as a journal article any month now.

A contribution to knowledge should be created in a computer network and, as it is created, should come under the library's bibliographic control. The library should manage access control from the beginning, in agreement with the author(s). The library should monitor the pattern of use and, at an appropriate time and if indicated, advise the author with respect to publication-in-print. Meanwhile, the library would have the contribution to analyze with the aid of a computer and to put into relation with other parts of the body of knowledge. During preparation, the library could also disseminate the contribution subject to agreed access restrictions via an electronic network.

4.0 SELF-SUPPORT VERSUS SUBSIDY

One principle seems to have been adopted widely by libraries and even made a point of pride; that advances by libraries into the realm of computers and networking should be supported by income from charges to users on a pay-as-you-go basis. If acceptance of that principle were a last resort, a way of telling the foundations, the government, and society that libraries are determined to advance into the future even if the rest of the world wears blinders, the acceptance should be applauded. However, the rest of the world is not totally and permanently blinded to the prospects offered by the new technology, and acceptance of the principle is not truly a last resort. The basic problem is to raise capital to develop a far superior way of fulfilling the most essential one of the functions underlying civilization: upgrading and making effective the world's corpus of knowledge. Most business people would seek government aid in undertaking such a project. It may be harmful for libraries to shift, at this crucial time, from government funding for research and development to a policy of evolutionary change financed out of current income from the systems that are being developed.

Chapter 16

by Lawrence G. Livingston

The Near Future of Library Networks in the United States

1.0 INTRODUCTION

The year 1980 is useful as a starting point for a description of the library network that is emerging in this country, because by that time the Library of Congress will be putting essentially all of its current cataloguing into machine-readable form. In addition to making almost everything catalogued at the Library of Congress available in the MARC system, the Library of Congress will provide access on-line to its in-process file and its authority files, both name and subject, together with the necessary cross-references. The availability of this vast amount of useful material will have significant impact on the way libraries, especially network libraries, operate in the future. This paper will define the relationships between the nation's networks and systems and with the Library of Congress.

Before describing the network that will develop over the next decade, it is probably useful to state the assumptions on which the conclusions reached in this paper are based.

1. The principal objectives of the network will be:

a. To make the bibliographic resources of the United States widely known and accessible to all who require them.

b. To reduce the rate of increase in the per-unit costs of library operations.

2. For the period addressed in this paper, books and other materials printed on paper will constitute the overwhelming majority of the storage medium used by libraries. Microforms will play an increasing role, but, in the time frame under discussion, the computer storage of complete texts of monographs will not.

3. Network development in the next five to ten years will mostly be based on technology now available.

4. The National Library of Medicine and the National Agricultural Library will both be important on-line components of the national library network. The National Library of Canada can be expected to participate as an equal partner.

5. The national library network will require interface with international library networks and other U.S. national information networks as these develop, but primary emphasis will be on developing the national library network.

6. The library component of the national information network will be a very significant part but only a part of the overall network envisaged by the planning document of the National Commission for Libraries and Information Science (1).

7. The library component of this emerging national information network may be treated separately from the rest of that development because, among other reasons, the library component is a finite piece and is developing much more rapidly than the other parts of the network.

8. The bibliographic files and apparatus of the Library of Congress are absolutely essential to the bibliographic segment of the national network. These files and that expertise will not be duplicated anywhere else.

9. The national library network will be built upon the network systems components and libraries already in existence.

10. What will evolve is a confederation of the library networks and systems, a confederation of essentially equal components, each retaining a high degree of internal autonomy, and all looking toward the Library of Congress for a significant part of the total bibliographic information needed by the libraries affiliated with each component.

11. It is not yet possible to describe where and how or even by which organization the national library network will be managed. The roles of the Library of Congress and the other major components in the network do not change with the eventual locus of that management organization. Therefore, much can be done in the definition of the network before it is known which organization will manage it.

12. Since it is not possible in this confederation of more or less equal components to standardize internal formats, processing, query languages etc., it will be necessary to standardize on the queries and responses which flow along the communications lines which connect these diverse components.

13. Network components will use and adhere to bibliographic standards wherever these exist. Network components will cooperate

with each other, with Standards Committee Z-39 of the American National Standards Institute, the National Bureau of Standards, and others in the definition of requirements for and the promulgation of new standards.

2.0 DEFINITION OF A NATIONAL LIBRARY NETWORK

Based on these assumptions, the definition of the network and a description of its development may begin. It will be seen at once that this network will develop from very diverse organizations. Among those prominently to be considered in this development are: OCLC with more than 800 affiliate libraries and 2.5 million records on-line; the Washington Library Network with the libraries of a single state cooperating; the BALLOTS system of Stanford just beginning a network; the University of Chicago system, at this writing essentially designed for a single large research library but examining the possibilities of networking; the Research Libraries Group (just starting a network interaction with the Library of Congress); other kinds of state and regional consortia, such as SOLINET, NELINET, AMIGOS, and others; the Federal Library Committee through the Federal Library Network (FEDLINK); and, finally, systems and networks which are as yet undeveloped.

It is manifestly impossible to do all that needs to be done in network development between now and 1980 at the same time. This fact requires that some very difficult choices must be made, by funding agencies, for example, as to the priorities for action. As painful as this decision may be, it must be made. Having selected certain of the major network components for priority action, developers will proceed like this: the overall development of the national library network will be guided by a national network planning document now in preparation by a group of representatives of the major active network components in the United States. This group is sponsored by the Library of Congress with funding assistance from the Council on Library Resources. It has had two meetings, and its network design paper is fairly well along on its way to becoming a draft which can be circulated generally among libraries and librarians for comment. Under the design umbrella in this document, each of the major network components will proceed as follows: each will look at the products and services which will be available from the Library of Congress in the 1980-81 time frame. Each component will come to its own conclusions as to what this means for that component; what impact this widespread availability of machine-readable bibliographic information will have on the component; what changes in organization and operations the component will need to effect, and so on. Having decided in this fashion how each component should look in the 1980-85 time frame, responsible people in component management will enumerate the developmental steps necessary to get that component from its present state to the posture desirable in the fulfillment of its role in the national network beginning in 1980.

2.1 Financing Development of the Network

The funding for the accomplishment of these developmental steps will, as it has been in the past, be based on a combination of efforts. As it becomes available, money will be used from federal and state programs, service charges to users of the components, and from funding made available by private foundations. As was mentioned earlier, funding agencies will have to select among the elements of the national network those which are to receive priority consideration for funding. It is emphasized however, that any network component, any library, any system, may, with its own resources, accomplish the developmental steps, which will be further described below, for itself and on its own schedule. Ideally, the development to be undertaken by any network component will be decided upon after due study of the national network plan and consultation with personnel of the Network Planning Office of the Library of Congress.

2.2 Library of Congress

The Library of Congress has made a major statement of policy with direct pertinence to national library development (2). LC, in that statement has described the products and services to be available, on-line and otherwise by 1980. LC has also said that, as a general rule, on-line products and services will be furnished to libraries via regional and other networks. The Library of Congress has also said very clearly, however, that everyone realizes that some libraries will not have on-line access to a network for a long time, perhaps never. Therefore, the products and services of the Library of Congress (which satisfy requirements in these libraries) must continue. These products must be available in several forms: printed on paper, catalogue cards, microforms, and perhaps in other ways.

When each network component has defined its 1980 network role and enumerated the steps in development necessary to achieve that role, each will, to the extent resources are available, develop its own capability to contribute to the network and take full advantage of the contributions of others. As these developmental steps are taken, a system comprised of a confederation of autonomous and largely equal systems will take shape. This statement of the confederated nature of the library network is repeated because of its signal importance to the understanding of what will happen.

Each major network component will have on-line access to the bibliographic files of the Library of Congress. This will include the MARC file, the in-process file, and the authority files; perhaps others. Each major network system will have its affiliate libraries on-line to the regional and possibly other network data bases. In addition to being connected on-line to the Library of Congress, each major network component will be connected to other network components. It's easily seen how OCLC, BALLOTS, Washington Library Network, the University of Chicago, the Research Library Group, and so on, need to be connected on-line to the Library of Congress. One must not overlook, however, the fact that other types of network components will be involved; for example, the bibliographic centers such as the Bibliographic Center for Research and the Pacific Northwest Bibliographic Center. As a general rule, individual libraries will not be connected on-line to the Library of Congress. This is not, however, an absolute; there may be exceptions for cases of special expertise, unique collection capabilities, etc.

12.3 Required Institutional Self-Study

Let us turn now to the kind of study each network component will have to undergo in determining its 1980 role. It will begin, of course, by assessing the impact on the regional or other network of the availability of the products and services from the Library of Congress outlined above. For example, component management will need to compare its requirement for current cataloguing with the current cataloguing available from the Library of Congress to determine the effect that the availability of this material will have on the staff, resources and procedures of the library members of the component. This significant change in the way cataloguing is done and the places where it is done as the burden of placing new cataloguing into machine-readable form shifts to the Library of Congress, will obviously have a serious impact on the cataloguing staff in the network library. The network librarian will have to plan to provide for the absorption of that impact. The librarian will do this by allowing some positions to be vacated by attrition as people retire and leave and he or she will reallocate saved salaries to pay for network services. The responsible librarian will also need to consider the requirement to reassign staff: some to reference service, for example; some to helping the users to understand and use the system; and in other ways. The personnel turbulence which will result from more effective library networks and their impacts on the library will provide, perhaps, the greatest challenge to the network librarian that s/he has ever faced. Obviously, s/he must plan ahead and prepare to train staff and to reassure them that the network will not replace them entirely.

After having decided what the impact will be on the network and its affiliate libraries, each network component will need to look at hardware, software, and staff resources which will be required in order to take full advantage of the situation as it will exist beginning in 1980. Many difficult questions will need to be answered. For example, what is the best organization for the management of the network component? There are probably about four or five possible alternatives for component management: a not-for-profit corporation; a for-profit corporation; a governmental agency at the federal or state level; interstate compacts; and, finally, a quasi-governmental corporation.

Perhaps the most difficult series of questions that the network component management will need to face revolve around the data base of that network component. Let us begin with the national aspects of the data base. Obviously the Library of Congress will continue to produce a significant portion of that national data base. The Library of Congress and indeed all network components will each need to maintain a certain number of records to control the collections in their various areas of responsibility and to serve their users. If one begins with the assumption that each network component, to include the Library of Congress, will need to maintain a certain number of bibliographic records for its own purposes, it becomes apparent that the national bibliographic data base, by definition, will be a distributed data base. It is useful in this context to think of the national data base from both its logical and physical aspects. Logically, the national data base consists of a single bibliographic record for every item of recorded information in any medium held in any network component. Ideally, that logical bibliographic record was created only once and as near the source of the bibliographic item itself as is possible; converted to machine-readable form once, and made available as needed throughout the national network. Ideally, there would be only one copy of each bibliographic record in the national network. Physically, of course, the national data base is another matter. It is not possible, at least in the short run, to have only one copy of each bibliographic record. It will be necessary for each network component, and indeed each library, to maintain sufficient records to control its own collections and satisfy its users. The actual physical distribution of these records in the logical data base, how much redundancy, where and how back-up and security must be provided, and so forth, require serious and detailed study.

Some other important questions each network component management must ask about its own data base include the following: what proprietary rights exist in a machine-readable regional, state or local data base? Are any of these rights exclusive? May a data base be copyrighted? It is evident that, to be truly effective, the national distributed data base will require the cooperative creation and exchange of bibliographic records. This mutual exchange of records requires in turn adherence to standards of national acceptability if expensive reworking is to be precluded. Is network component management prepared to make the compromises necessary to bring its records up to national levels of acceptability for exchange? The managers of the various components of the national bibliographic data base will also be faced with the questions of duplicate records. It is not yet easily seen how unnecessary duplication in this national bibliographic data base will be eliminated. The question requires study. The final solution of this problem will probably not occur until there is a unique identifier for each bibliographic record. Neither the Library of Congress card number nor the International Standard Book Number, in its present form, is sufficient as this unique identifier. A new identifier, perhaps based on one of these two numbering schemes, will have to be devised.

The other major problem that data-base managers will have to confront has to do with the bibliographic quality and consistency of records in their part of the data base. It is short-sighted to say that records in a shared data base don't need to be consistent and of good quality; or that the speed and flexibility of the computer obviate the requirement for consistent bibliographic quality in records and permit individual users to upgrade them if they find them deficient. It is short-sighted because unnecessary and added cost accrues every time the staff of the second to the nth library work on a record that should have been brought up to nationally acceptable standards at the time of entry into the data base or immediately thereafter. A firm requirement for bibliographic consistency and quality will obtain in the data base under description. Since this is to be a shared effort with many people inputting records to the data base, it is beyond doubt that some sort of centralized bibliographic authority and control must be agreed upon and established.

The Library of Congress, in the bibliographic network development under discussion, will be principally responsible for bibliographic consistency and quality in the data base records. However, not all records in that data base will originate at the Library of Congress, so the responsibility for, and the work to be done in the preparation of, nationally acceptable bibliographic records will have to be shared. Certainly the input will be shared, and the center of responsibility concept, developed in the Conversion of Serials project, points the way. In the CONSER project, a dozen or so excellent libraries input serial records and have varying degrees of assigned authority to upgrade the records as they find them deficient, with the ultimate responsibility for the bibliographic quality of the data base residing in the national libraries, which have been designated as centers of responsibility in this context. Something analogous to this will undoubtedly need to be developed for the larger network.

2.4 Component Technical Considerations

Having decided its 1980 role in the national library network; having enumerated the developmental steps necessary to achieving a capability to fulfill that role; and having confronted and solved the very difficult problems with the bibliographic quality of the records in its data base and the size and distribution of that data base, network management will then be able to proceed to look at the systems, hardware and software, required for each component of the national network. Component management will quickly come to the conclusion that not every function requiring support in a large number of network affiliate libraries can be done on the central computer system. Circulation control is perhaps the most obvious example of a requirement for decentralized computer support. Central circulation control is probably out of the question in a large system with many affiliates because of the sheer volume of transactions.

If this concept is accepted, one may look for ways to decentralize the circulation control function. One approach certainly would be to install minicomputers in the network libraries for circulation control. A very efficient prototype of this minicomputer in circulation control has been running at Bucknell University for several years. In that system, major advantage of the capabilities of the central system for data base building and maintenance is taken in a system that has the mini doing the day-to-day circulation transactions and periodically updating the central files. This can, of course, have any number of permutations, and will undoubtedly be done in various ways in different places.

Some of the major questions to be asked are: what is the best relationship of the circulation minicomputer in the network library to the bibliographic data base held centrally in the system? Is the most efficient circulation control system one which keeps on-line only live circulation transactions; or is it one that keeps a bibliographic record sufficient for item identification on-line at all times? A corollary question is, how much bibliographic search capability must be supported by the circulation record in the network library? Does the circulation librarian or the user depend on the central network system for all bibliographic search capability; or is the circulation record sufficient to support bibliographic search for identification? The question of subject search in the network generally will be covered later.

If one can agree that the network library will probably require a minicomputer for the circulation function, one can then proceed to an examination of other functions that might be better handled in decentralized fashion like serials check-in, binding and routing control. It is obvious that the level of detail required for an individual library to check in its serial issues will be overwhelming when considered from the point of view of the network system handling many hundreds of libraries. Given the irregularity of serial issue arrivals and problems of title identification and different cataloguing procedures, maintenance of the serial check-in file would overwhelm the central system's capability for file maintenance in short order. So, serials check-in and internal control may form the second function to be handled in decentralized fashion in the network library. Next, fund accounting logically will remain in the network library; not only for technical reasons but also for reasons of a different kind. It seems evident that the individual librarian, responsible to his or her own bosses for the funds, will be entirely reluctant to release any control of these funds, and the accounting of them, to some distant network center. There are technical reasons for the decentralization of fund accounting as well. There are probably as many different fund accounting systems as there are different libraries. Some libraries maintain two sets of books, one internal to the library itself showing how the book funds are allotted and how they are spent, and another, often the aggregate of the first, maintained for the institution controller. To do fund accounting centrally would require one of two things, both of which are probably impossible. In the first instance, the network center would provide a different fund accounting software module for every difference in fund accounting in the network libraries. Conversely, the network center would build a single accounting module and expect every institution and every librarian in the network responsible for funds to change his own accounting procedures to fit the standard module. A little reflection will indicate that neither of these is likely to happen. Fund accounting will be run in decentralized fashion, probably on minicomputers acquired for the circulation function.

Minicomputers in the network library may be justified on the basis of the circulation function, and then may, using time otherwise available on the mini, support such things as fund accounting. Are there other functions that logically fit the availability of the mini in the library? Yes, there are. Given the computing power available in the mini and given the fact that circulation systems as a rule don't run at night, there will be excess capacity available on the circulation control minicomputer that can be very effectively used. For example, it will be possible for the local librarian to tailor bibliographic products and services very closely to individual, faculty, and departmental requirements. Such things as lists of new accessions built against individual and departmental profiles will be easy and comparatively inexpensive to produce. The local librarian will be able to tailor these services and products, undoubtedly to include some not possible at all in manual systems, be able to deliver them in very rapid order. The potential capabilities of the circulation mini to cooperate with the central system in the provision of management statistics to the local librarian must also be considered.

2.5 Search Strategies and Techniques Required for the Network

Some new search strategies and techniques will be required and available in the network under discussion. Bibliographic search keys derived from the data in a very limited number of fields, a concept so brilliantly pioneered in the OCLC system, will become progressively less efficient as file size and user library population grow along with the volume of transactions. When serial records are added and must be searched on corporate author with the search keys initially designed for searching in monograph files on personal author, the use of derived search keys is complicated further. The solution, of course, is to develop search keys which discriminate enough to produce a manageable number of responses in a reasonable time for viewing and selection by the user.

Further network development and systems growth will probably prove that a continuation of the concept of providing subject search along traditional lines, for example by the use of subject terms tacked onto the bibliographic record at creation time, will not be possible in a very large network with many millions of records on-line and many libraries as customers. Because of the immense size of the central network files, the number of times the disc memory must be accessed in subject search and the short response time that must be maintained to satisfy the user, subject searching along traditional lines will probably not be feasible in a very large on-line network. If this proves to be the case, what is to be done? New approaches will have to be taken. A combination of search capabilities must be devised that will permit searching on many more data fields than just title and combinations of author-title. In this way, using Boolean operators, it may be possible with very large files and large user populations to provide the discrimination that will be required. The problem of providing a subject search capability in a very large network furnishing comprehensive functional support to many libraries *and* updating files and indexes will not be solved simply by adding subject terms to the bibliographic records to be searched, expecting to produce subject specificity sufficient to be manageable in the system while maintaining acceptable response time and still satisfying the user. The probable solution to this dilemma is to provide a combination of search capabilities; for example, the capability of using Boolean operators to 'and' together key words from several places in the record. Some useful help may be found in language of publication, date of publication, perhaps place of publication, and the traditional author and title fields. In order to avoid spelling errors and get at key words wherever found in the record in a most efficient manner a word truncation capability will be required. It would also be highly desirable to have a capability to specify the spatial relationship in the record and indeed within the individual data field between two or more search terms. It should be possible to say whether these word stems or word fragments should be found contiguous to each other, separated by so many words, etc.

A national library network more distant than 1980 will turn for its content search capability to the unabridged dictionary approach. Just as one goes, with a subject discipline in mind, to Webster's unabridged dictionary, and comes out with the definition of a term in the context of that subject discipline, so will the user many years from now have what amounts to an unabridged dictionary of all the significant terms. Such a dictionary, edited by stop lists, in all the languages represented in the user's bibliographic file, will permit finding the subject term in the proper context, given the discipline in which the user is interested.

2.6 Library Network Technology 1980-1985

Turning now to the technology likely to be used in this network development from 1980-1985, one can expect that the network will be based on technology largely available today. Books and other materials printed on paper will be the principal medium used by libraries. The library without paper will not occur by 1985 or 1995.

Library network components will be dependent upon the kinds of main-frames (including minicomputers), line printers, discs, tape drives and so forth that are now available. Machine-sensible equipment, both optical scanners and magnetic stripe, will find increasing use, in circulation and interlibrary loan systems. In the time frame under discussion, advancement in size and speed, and a relative reduction of costs of mass memory devices can be expected. Librarians will see experimentation at least, in time-division multiplexing systems using video transmission of the kind very recently announced in England (3). There may be some message traffic over satellite links, but these, of course, will be transparent to the librarian and the user. For all practical purposes, communications technology will be that already available.

2.7 Required Standards

Much standards work by Standards Committee Z-39 of the American National Standards Institute and its peer agencies will be required in this developmental period. Several kinds of identification numbers will be required, for example. The requirement for a unique piece I.D. has already been mentioned. This will probably develop by building upon the International Standard Book Number and the International Standard Serial Number and perhaps international standard identifiers for other recording media, as these develop. An individual I.D. for library patrons will be required, as will an individual I.D. for every library component active in these networks.

Since the standardization in the network generally will focus on the message traffic between the components of the network as opposed to the internal functions of these components, it seems evident that the messages and their protocols will need to be standardized. This will result in an array of standard messages and queries of various kinds, responses containing bibliographic records or other information, switching instructions, 'mail to' types of instruction, etc.

It is evident that in a network of autonomous equals, it is not going to be possible to standardize internal procedures or query languages. Therefore, the only solution to meaningful communication among these diverse systems is to standardize all the messages that flow between them. Having done so, each network component will require only the capability to take such a message in, act upon it with an internal program and output a response in the standard form that will be equally acceptable to the software in any system which has agreed to provide the capability to handle the various standard messages.

A standard numerical code for name authorities will probably be required because national and international acceptance of a single form of each name to be used by all cataloguing agencies is nowhere in sight. The international code could be stored and transmitted as the single form of the name which could be recognized and accepted by all systems. The code for each name would have a one-for-one relationship with a single name entity, no matter in how many forms that name appeared in the various catalogues and authority lists. National bibliographic agencies would designate the nationally acceptable form of the name and all variant forms would be related to it and to the appropriate code by an accepted content designation scheme. If users of the system could find reference to the numerical code off-line, say in printed lists, on-line searching with the number would be more efficient than any key derived from one or another form of the name could possibly be.

3.0 SUMMARY

By reiterating the salient features of the network described in this paper, one can see whether it presents a picture of a network sufficiently clear so that individual components may begin forthwith to prepare for the definition of their own roles.

Based on the assumptions listed at the outset, the evolving national library network will be, in effect, a confederation of equal and autonomous components of many diverse kinds. Each component will retain a high degree of internal autonomy. All components will look to the Library of Congress for a significant amount of the bibliographic information they require. The national data base will be distributed in nature. There will be decentralized cataloguing to a degree, decentralized input, but centralized responsibility and authority for bibliographic consistency and integrity in the national data base. Some major network components will be selected for priority treatment by funding agencies and others in this network development; but any network component is free, as soon as the planning document is available, to assess its own relationship to this network, and, in coordination with the Network Planning Office of the Library of Congress, to plan the development necessary to permit it to fulfill that role.

The major library functions in the national network will be enhanced and supported by a series of diverse network component centers, including bibliographic service centers such as OCLC, brokering centers such as the bibliographic center at Denver, regional consortia of various kinds, perhaps interstate compacts, certainly state systems, and so on. There will be no single blueprint for network components. Funding will be a combination of federal, state, local, and private funds. The possibilities for network management at the regional level and at the national level include governmental agency, quasi-governmental corporation, for-profit corporation, and not-for-profit corporation. At the national level, an attractive alternative is quasi-governmental corporation analogous to the Tennessee Valley Authority. Participation by components in the national network will be voluntary. The trend in bibliographic practice over the long term will be toward closer adherence to Library of Congress practice and standardization.

References

1. *Toward a National Program for Library and Information Services; Goals for Action.* The National Commission for Libraries and Information Science. Washington, D.C. 1975.
2. *The Library of Congress and the National Bibliographic Center.* The Association of Research Libraries. Washington, D.C., February 1976.
3. Prinski, Robert, "Esmerelda Is New TV Show," Reprinted from the *Wall Street Journal* in *The Washington Post,* Sunday, November 7, 1976.

PART III Computer Networking in Education and Research in the Health Sciences

by Eugene Young

Overview: Networking in the Health Sciences

The papers that follow focus on the development of collaborative efforts between the computer sciences and the medical and life sciences for the improvement of medical education, research, and treatment. Currently, several national networks are users in these areas for collaborative efforts between computer science and clinical research groups. Authors discuss the ways in which the successful use of computer networks can be evaluated and improved in the future to increase the degree to which computing can be applied to research and clinical consultation and to the education of medical personnel.

The Health Education Network, Inc. has evolved from a federally-supported experiment to establish a user-supported and user-directed operating network. In 1972, the Lister Hill Center of the National Library of Medicine established the Experimental CAI network to: 1) assess the technical feasibility of inter-institutional sharing through a telecommunications network; and 2) determine whether materials developed at one institution would be accepted and used at others. In 1975, when federal support terminated, users of the Network and its hosts took actions to ensure continuation of nationwide access to the largest single data base of computer-based educational materials in health sciences. During the 4 years of its operation, the Network has been used by more than 150 different institutions in over 40 states. Materials are available for allied health sciences, nursing, basic medical sciences, clinical medical sciences, and continuing medical education.

Several projects that use artificial intelligence methods for consultation, and are accessed or developed via computer network are described next.

An evolving computer program is being developed at Stanford University to assist non-specialist physicians in the selection of therapy for patients with bacterial infections. MYCIN attempts to model the decision processes of medical experts and provides input, interrogation, modeling, and suggestive output. Computer networking allows wide access for experimentation with, and use of, this developing system.

Diagnosis in internal medicine is frequently complicated by the need to discern multiple diseases in a single patient. In order to enable computer-based diagnostic systems to deal with these complex clinical problems, it is necessary to devise problem-formation heuristics that can effectively partition the set of disease hypotheses evoked by a given set of clinical manifestations into coherent subsets, within which diagnostic problem solving methods may be applied. The problem-formation heuristics underpinning performance of the INTERNIST system are presented here along with examples of the system's performance.

A computer model of expert reasoning in digitalis administration has been developed at Tufts New England Medical Center. It assesses the patient's clinical response to the drug and uses these responses as feedback to guide subsequent therapy. A clinical trial has demonstrated the feasibility of using this program in a wider range of clinical settings than has been possible with earlier programs. Later trials and use of this program will utilize networking facilities to provide access from across the country.

The CASNET computer consultation system, incorporating an explicit model of disease processes and expert reasoning for the diagnosis and treatment of glaucoma, has been developed jointly at Rutgers University and the Mt. Sinai School of Medicine. Computer networking has played an important part in developing the program beyond the prototype stage into a proficient clinical tool. The Ophthalmological Network (ONET) consists of clinical investigators at five glaucoma research centers (Mt. Sinai School of Medicine, Washington University, Johns Hopkins University, University of Illinois at Chicago and the University of Miami) who collaborate in the testing and further development of the consultation programs. A data base of glaucoma cases, together with computer programs for retrieval and analysis, has also been established to facilitate joint clinical studies among the ONET members.

The ONET computing is carried out on large PDP-10 time shared computers at Rutgers University and Stanford University. Both universities are centers for research in Artificial Intelligence in Medicine (AIM). The Rutgers Research Resource for Computers in Biomedicine was established in 1971 as a focus for modeling and problem solving research applications in biomedicine, while SUMEX (Stanford University Medical Experimental Computer) was established in 1974 as a first national shared computing resource for medical research. The MYCIN, INTERNIST and CASNET programs are all accessible at SUMEX-AIM over computer networks. The ONET is the first example of a geographically dispersed group of collaborating clinical researchers that make use of these shared resources.

Chapter 17

by Ruann E. Pengov
Barbara B. Farquhar
and Robert G. Votaw*

The Health Education Network, Inc.

1.0. INTRODUCTION

The Health Education Network, Inc. has evolved from a federally-supported and federally-directed experimental network to a user-supported and user-directed operational network. In 1972, the Lister Hill Center for Biomedical Communications of the National Library of Medicine established an experimental network to 1) assess the technical feasibility of inter-institutional sharing of computer-based educational (CBE) materials** through a telecommunications network and 2) to determine whether CBE materials developed at one institution would be accepted and used at others. Both the technical feasibility of networking and user acceptance were established in 1974, after 18 months of network operation. Federal support continued for an additional 17 months, until May of 1975, during which time the users and hosts took action to ensure continuation of nationwide access to the largest single data base of computer-based education materials in the health sciences in the world.

Currently, the Health Education Network makes it possible for institutions to share computer-based education materials at a fraction of the cost of developing and maintaining them. The Network also

---

*Ruann E. Pengov is affiliated with the Ohio State University College of Medicine. Barbara B. Farquhar is affiliated with Massachusetts General Hospital. Robert G. Votaw is at the University of Connecticut Health Sciences Center.

**CBE (or computer assisted instruction [CAI]) refers to situations where students are engaged in interactive learning dialogues at a computer terminal. In these situations, rather than being the subject of instruction, the computer assists in providing instruction.

provides opportunities for medical educators and computer scientists to explore innovations in education. In these ways, the Health Education Network has achieved success; its potential use and future directions are limited only by the imaginations of those involved with it.

This paper is a case history of the development of the Health Education Network, Inc. which discusses the need for sharing, planning for networking, the experimental network, the user community, the Network hosts, the impact of the Network, utilization of the Network, the costs and benefits of networking and the key elements in network planning and operation. It is our hope that the information offered in the following pages will offer insights for the development of networks in other disciplines.

2.0 NEEDS FOR SHARING

The impetus for resource sharing arises from several points:

1. Sharing is a commonly accepted academic responsibility in the health sciences.
2. Sharing is implicit in the federal funding which supported many of the developmental efforts in computer-based education.
3. The front-end capital investment in the development of these computer-based education materials is high. Estimates of developmental time range from twenty to several hundred hours of effort on the part of faculty authors to produce one hour of interactive computer-based education. This depends on several variables including author experience and the complexity of the program being created. High development costs make sharing a necessity; no one institution can develop courseware to meet all its needs.
4. Courseware maintenance costs are high; these materials do not "live and breathe" on their own once developed—they require continuing "tender loving care". Faculty authors must insure the ongoing relevance and accuracy of lessons. Technical and educational strategy members of the course development team want materials to continue to be responsive to, and reflective of, the most dynamic educational strategies currently available. It is not cost effective or even possible to perform such maintenance functions for all lessons in each institution which uses them. Via networking, the most scarce resource, the author, is effectively shared.

These factors make resource sharing necessary if computer-based education is to be utilized in the health sciences. The Health Education Network was conceived, has evolved, and currently operates to help meet these needs.

The paragraphs which follow describe the key elements which were critical for the conception, survival, growth and development of the Health Education Network; many of these elements are still critical in current Network operation.

3.0 PLANNING FOR NETWORKING

There were several key elements in the early development of the Network. It was recognized generally within the health sciences that there was a need for sharing expensive and scarce resources. A national professional medical organization advocated support of networking. There was a federal sponsoring agency with a mandate to support sharing through networking and they had funds to implement a program. There was a fortuitous overlap of goals within the sponsoring agency because they were already networking in the area of medical libraries. There were some willing users of networks and some willing providers (contractors or hosts). The hosts had well-developed and recognized data bases. Further, they understood the importance of user services to successful network operation. Each element appears to have been essential to the birth of the Network.

In August of 1968, the sponsoring agency for the initial Network, the Lister Hill National Center for Biomedical Communications (LHNCBC), was established by Congress as the research and development arm of the National Library of Medicine. Their charge was to design, develop, implement and manage a biomedical communications network.

The Association of American Medical Colleges (AAMC), a national professional medical organization, worked with the Lister Hill Center to explore possibilities for networking (17). The Council of Academic Societies of the AAMC and the Lister Hill Center sponsored a conference entitled "Potential Education Services From a National Biomedical Communications Network." Conclusions reached at that conference were 1) that there was a need for a national network and 2) that there were many services that might be provided. This led to a contract with the AAMC to provide more specific recommendations for what might be shared in a network. Dr. Eugene A. Stead, Jr. chaired a task force for the Association of American Medical Colleges, which visited 10 medical schools and talked with over 100 individuals. In July of 1971, they published their report "Educational Technology for Medicine—Roles for the Lister Hill Center" in the *Journal of Medical Education* (16). There were fifty-three recommendations in that report; it was clear that there was some need to establish priorities.

The National Library of Medicine already had a national network for medical library searching—MEDLARS. LHNCBC staff recommended that the Library find other data bases that might share that network. They noted that there were several institutions that had data bases for undergraduate and continuing medical education. No attempt had been made to offer these jointly as a service to potential users. LHNCBC staff recommended that a small number of medical education programs be made available through the planned biomedical communications network.

In September of 1971, LHNCBC convened a meeting of potential hosts to discuss making programs available for networking. The LHNCBC thought it could obtain programs from the hosts, put them on a central computer, and distribute them to users. However, the programs are dynamic, "living" and do need regular maintenance by the author team. Hosts were willing to share the programs, but they insisted that the programs reside on their own computers. Another difference in perspective was that LHNCBC wanted to demonstrate programs to a wide group of users and use only a small number of programs, while the hosts, on the other hand, wanted to work in depth with a small number of users. Both perspectives were embraced by having two classifications of users, trial and operational. Hosts agreed to provide access to their programs through TYMSHARE, a commercial time-sharing and telecommunications company which was to support the medical library searches through its national network, TYMNET (1). The decision to share through distributed hosts postponed issues of standardization and transportability.

In March of 1973, LHNCBC sent out Requests for Proposal and held discussions with hosts regarding costs of services, the extent of services to be provided, and the basis on which costs would be determined. It is important to note at this time that LHNCBC estimated that a fourth of the cost would be for personnel; the hosts, on the other hand, estimated that two-thirds of the cost would be for personnel. In June of 1972, the National Library of Medicine allocated $145,000 for a ten-month period for distribution of materials from Massachusetts General Hospital, The Ohio State University College of Medicine and the University of Illinois Medical School. LHNCBC staff were told to make it explicit to both users and contractors that this was to be an experiment of limited duration.

In May of 1972, LHNCBC sent out 26 letters to potential Network users who had been selected from the membership of the Association for the Development of Computer-based Instructional Systems (ADCIS). It is not clear why LHNCBC did not return to the Association of American Medical Colleges (AAMC) or to the Council of Teaching

Hospitals to solicit users. This initial mailing was the only publicity sent to the medical community at large noting the availability of free computer-based education programs.

4.0 THE EXPERIMENTAL NETWORK

In July of 1972, Massachusetts General Hospital became the first hospital to offer materials on the Network. Two months later, The Ohio State University College of Medicine was officially connected to the Network, and in January of 1973, the University of Illinois Medical Center was joined to the Network. Connection involved installation of a TYMNET mini-computer and development of software to interface with the host computers. Massachusetts General was an original test site for TYMNET node connection to Digital Equipment Corporation hardware, and Ohio State University was an original test site for TYMNET node connection to IBM 360/370 series equipment. Given the "state-of-the-art", a period of some months was required to achieve smooth transmission of technical messages and to establish a stable and reliable network connection for users.

After almost a year of operation of the Network, the LHNCBC Center called a meeting for its operational users. The user meeting was simply a means for the LHNCBC network staff to receive feedback from the users. Until this time, LHNCBC's primary interface had been with the hosts in technical matters involving Network start-up. The May 1973 meeting was the first effort aimed at serious user involvement.

Initially the LHNCBC experimental network provided free service to institutions willing to support their own computer terminal(s) and connection to the nearest TYMNET network node. LHNCBC did not anticipate the high user acceptance which was, in fact, so great that LHNCBC was forced to call upon participating institutions for partial cost recovery. In February of 1974, user charges of $2.50 per connect hour were instituted. Still, user interest and Network success outreached budget allocations. In order to continue the Network, a plan evolved for gradual increase in user charges and for beginning the transition from a federally-supported network experiment to a user-supported network.

Due to local conditions, the University of Illinois Medical Center ceased to be a host on the Network in May of 1974. The Lister Hill Center made a major decision to transfer a portion of that data base, specifically the CASE simulation materials (6, 7) to The Ohio State University College of Medicine.

In July of 1974, less than a year after the first user fee announcement, the Lister Hill Center increased the usage fee to $5.00 per connect hour. Still, the fee covered less than one-third of the total cost for Network operation. The economic pressure brought to bear on the LHNCBC by the success of its experimental network was very real. At one point, LHNCBC estimated that almost one-half of their total expenditures in research were allocated for this Network. This circumstance is perceived as a major reason for the decision and subsequent announcement by the Lister Hill Center in January, 1975, that it would withdraw support from the Network in May, 1975. Another reason for termination of LHNCBC support was that the Network was no longer an experiment—it was a success.

In May of 1975, the National Library of Medicine officially terminated its support of the Network; this ended the experimental network. It is significant to note that the very next day the operational user-supported network began. The many events leading up to this smooth transition are discussed below.

5.0 THE USER COMMUNITY

Most Network users and their institutions were new to the concept of computer-based education. User institution administrative contacts, faculty and students, all were inexperienced. And so, each, according to his own time sequence, recapitulated a developmental process of initial awareness of, exposure to, experimentation with, and increasing insight into the potential of computer-based education. Each user had a different approach, but many were able to effect change and make significant instructional impact in their curriculums through the use of CBE. The time frame for this process was and is variable, and the rate is dependent upon many factors. The subject and process are in much need of further research.

At the August, 1973 meeting* of the Association for the Development of Computer-based Instructional Materials (ADCIS), a few Network users sought to begin planning for an orderly transition because it seemed improbable that Lister Hill participation would last indefinitely. These Network users, who were members of ADCIS, asked ADCIS to petition the Lister Hill Center for continuing support of the Network; ADCIS felt it was not able to do so because only a small part of its constituency was involved.

Six months later, at the January, 1974 ADCIS meeting in Washington, D.C., Dr. Harold Schoolman, Assistant Deputy Director of the National Library of Medicine, made the announcement that the

*Prior to LHNCBC announcement of the pending termination of its financial support for the Network.

Lister Hill Center had no plan to support the Network in its (then) current configuration beyond May of 1975.

Now, in reaction to this clear signal, a loose amalgam of Network users sought assistance from one of the interest groups in ADCIS, and a tentative plan was agreed upon for exploring continuation of the Network. Over the next four months, however, no viable plan was forthcoming from ADCIS. An eastern regional meeting of Network users was held at the University of Connecticut in May of 1974 to consider courses of action. The 14 institutions represented voted to seek means to continue the Network and to form a National Users' Group. Three task forces were established in the areas of networking, evaluation and faculty development. It is noteworthy that this user's group started its existence expressing concern about development of faculty literacy in the use of computer-based education strategies and concern about the evaluation of CBE programs through peer review and through research on educational effectiveness.

To meet the mandate set by the regional group in May, a national organizational meeting was held in August of 1974 in conjunction with the ADCIS meeting in Bellingham, Washington. The National Users' Group voted to reaffirm the conclusions of the eastern regional meeting. They also voted to organize as the Health Education Network Users' Group of ADCIS. This proved to be a crucial decision. It afforded an instantaneous frame of reference for governance and communication. Much time was saved. It also afforded a legitimate recurring focus for developmental activities and working sessions which were so necessary during the early phases of the Users' Group existence. The Users' Group elected officers and a Steering Committee* and modest funds were provided from ADCIS for the Steering Committee, which was spread across the United States, to meet for planning sessions. This action proved critical to the ultimate survival of the Network.

Many factors undoubtedly contributed to the apparent success of the Users' Group. Above all, the impending loss of LHNCBC support for the Network in May, 1975, only nine months hence, provided a problem orientation in which lesser differences were ignored. The full attention of the Users' Group was focused on finding means to continue the Network. Concern with evaluation and faculty development was forced into abeyance.

Throughout the next months, the Steering Committee attempted to develop, in close cooperation with the Lister Hill Center and the hosts, a plan whereby the Users' Group would assume direction and support

* Appendix B lists the Steering Committee members.

of the Network over a two to three year transition period. Quite apart from money matters, the Users' Group, as an interest group of a professional organization, was not capable of carrying out the management and contractual (i.e. billing) aspects previously borne by the Lister Hill Center. This plan proved unacceptable to LHNCBC, largely due to the User Group's lack of legal status as an organization. Planning then focused on an interim approach in which the hosts, Massachusetts General Hospital and The Ohio State University College of Medicine, would perform much of the management and contractual work.

In the spring of 1975, the hosts executed contracts with the TYMSHARE Corporation for communications and arranged to bill users directly. The hosts, in negotiating contracts directly with TYMSHARE, assumed the responsibility for any losses which might be incurred as a result of the contract. At OSU, this had perhaps less impact than it did at Massachusetts General Hospital since Massachusetts General Hospital was much more susceptible to fluctuations in the market than was OSU. Both institutions, however, assumed much financial risk due to lack of the capital which had previously been borne by the federal allocation to the experimental network.

The Steering Committee communicated with the Lister Hill Center concerning a need to continue to provide a central Network coordinator. A half-time coordinator was funded for ten months and was located at the University of Connecticut Health Sciences Center.

During this same burst of activity in the spring, the Steering Committee investigated affiliation and alliances with EDUCOM (Inter-university Communications Council, Inc.), NERCOMP (New England Regional Computing Program) and SECOS (Shared Educational Computer Systems). NERCOMP and SECOS were aware of the Network concerns and approached the Steering Committee with proposals. Considerable counsel and assistance was received from the AAMC (Association of American Medical Colleges) and, later in the spring, involvement was initiated with the American Association of Dental Schools (AADS).

The Steering Committee recognized that it would probably be unable to solve Network business and management problems by affiliation at that point in time. The rates of development in the other organizations were simply on a different time scale, and the Network had an urgent problem which could not wait. On this basis, in February, 1975, a task force was appointed to explore incorporation. A second task force was appointed to explore enlargement of the Network data base with particular reference to materials specifically relevant to dental education. For both educational economy of scale reasons, there was interest in enlarging the user base and, hence, a task force on marketing was appointed. To simplify the day-to-day running of the Network, an Operations Group was formally brought into being, recognizing, in part, what had already evolved—primary support of Network operations was being rendered by user services coordinators at the hosts who exerted TROJAN efforts in improving communications, administrative procedures and user services.

The LHNCBC experimental network terminated on May 31, 1975. Due to the efforts of the Steering Committee, of the Operations Group, and of many dedicated users, the Health Education Network began on June 1, 1975—the very next day. There was no apparent difference to the user; the technical configuration remained the same. A remarkable transition had taken place. The Network was transformed from a federally-supported and federally-directed experimental network to a user-supported and user-directed operational network.

The Incorporation Task Force was quite active, also. Proposals for by-laws and for articles of incorporation were developed and were reviewed by the Steering Committee. Incorporation papers for the Health Education Network, Inc. were filed in the District of Columbia in October, 1975. The Health Education Network, Inc. formally came into being at its organizational meeting in January, 1976. The Board of Directors and officers were elected,* and shortly thereafter, the Board of Directors established a subcommittee on financial planning. Certification as a tax-exempt, not-for-profit corporation was sought and granted by IRS in June, 1976.

The immediate problem facing the new Network was still development of a sound economic approach to network which would insure long-term Network survival and growth. The subcommittee on financial planning is at work. The Board is also actively exploring new approaches to enlarging its user base through innovative combination of certain marketing and faculty development ideas. When its business concerns have been addressed, it is likely that the users of the Network will turn their attention to the other two areas identified initially at the Eastern regional users meeting—faculty development and evaluation of CBE materials.

6.0 NETWORK HOSTS

6.1 Massachusetts General Hospital

The data base at Massachusetts General Hospital (MGH) consists of a library of about 30 simulations (3). Appendix D offers an overview of

* Appendix C offers a list of the names and institutional affiliation of the first Board of Directors of the Health Education Network, Inc.

the MGH simulations. There are multiple cases within each of these simulations, and the programs thamselves are designed to teach and test the process of clinical problem-solving. Programs do teach facts, but the emphasis is on the *process* of clinical problem-solving. Programs are used by medical students to practice decision-making and by practicing physicians for self-assessment, for continuing education and for certification of clinical competence. The Harvard Department of Continuing Education has approved many programs for Category I Continuing Education Credits toward the Physician's Recognition Award of the American Medical Association (AMA). It is most likely that the programs will be used in the future for the assessment of clinical competence, licensure, certification and recertification.

During the past 4½ years, MGH has served over 150 different institutions in more than 40 states with approximately 50,000 hours of computer time for medical education. Usage has varied. It was as low as 20 hours per month, in the very beginning, and has reached a high of 1750 hours per month. An average of approximately 50 institutions use Massachusetts General Hospital CBE materials via the Network at any one time, with an average of 700-800 usage hours per month. About 68% of the users are medical students in their clinical years. Another 18% of the users are practicing physicians, and this percentage is growing. Only 14% are nurses and allied health personnel. The institutions MGH serves are medical schools or university medical centers (68%), hospitals (16%), national medical organizations (11%), and physician groups (5%). Usage by physician groups will probably grow in the future.

6.2 The Ohio State University College of Medicine

The Ohio State University College of Medicine (OSU) data base is, in part, complementary to that at Massachusetts General Hospital. Over 350 hours of interactive CBE programs are available to a variety of audiences. Table 1 indexes the programs by audience and Appendix D provides a subject index of the materials. Fuller description of the OSU system and data base is available in other publications (2, 13, 14).

The primary emphasis in CBE usage at OSU is integration of the materials into on-going curricula in medicine, in nursing and in allied health. As shown in Figure 1, however, curricular usage within the College of Medicine constitutes only 50 to 60% of the total usage of OSU's program library. The statewide Computer Assisted Instruction Regional Education Network (CAIREN) (5) and the Health Education Network each account for significant percentages of total CBE usage at OSU.

TABLE 1
OSU Data Base Indexed By Audience

Audience	Number of Programs
Clinical Laboratory	27
Dietary	26
Dental	6
Emergency Medical	14
Environmental Services	8
Management	6
Medical Records	8
Medical (Physicians, Residents, Interns and Students)	113
Nursing	51
Occupational Therapy	18
Optometry	5
Patients and Families	10
Pharmacy	6
Physical Therapy	20
Respiratory	15
Radiology	17
Secretarial	11

Over 150 institutions in more than 40 states have logged approximately 36,000 hours of Network usage of the OSU CBE materials. On the average, 50 institutions access OSU via the Network at any one time. They log 500 to 600 usage hours per month. Users on the national network vary from small hospitals to major medical centers; the type and extent of usage varies as much as does the type and size of institution (15). By far, the largest number of users are presently reviewing available materials for potential curriculum incorporation. But several institutions have made significant commitments for utilization. The University of Washington in Seattle has accessed The Ohio State University Independent Study Program (ISP) materials (4, 12) and has altered them to meet local objectives and needs. As a result of initial access to CAI facilities at The Ohio State University, Washington

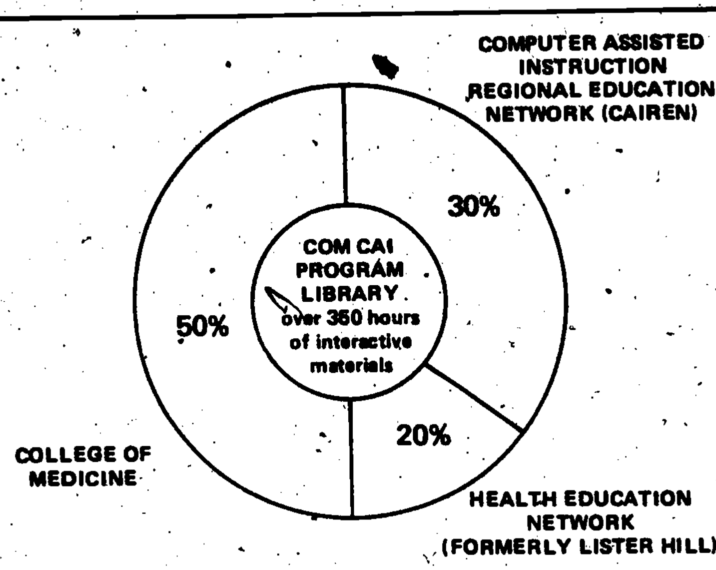


Figure 1. The Ohio State University College of Medicine CAI System Data Base & User Group Overview

is currently operating its own altered version of OSU's ISP. George Washington University was the next institution to author CAI materials remotely; its materials were the first user-developed materials to be released over the Network for general usage by other user institutions. The University of Pittsburgh, in association with Pittsburgh Eye and Ear Hospital, is using ophthalmology materials and has developed its own CAI materials with the assistance of CAI support staff at OSU. The University of Oregon logged significant usage in the nursing area and Ft. Worth Osteopathic Hospital was an early continuing education user. In a joint effort with the University of the Pacific and The Ohio State University College of Medicine, the American Board of Internal Medicine is refining, revising and generating new CASES, and exploring their use in the recertification process for internists. The project, which has assumed the acronym MERIT (A *M*odel for *E*valuation and *R*ecert-ification Through *I*ndividualized *T*esting), has successfully completed two regional field tests and is currently planning a national field test using the Network (7, 8, 9). In general, OSU sees a bright future for the use of its CBE materials in assessment of clinical competence, licensure, certification and recertification.

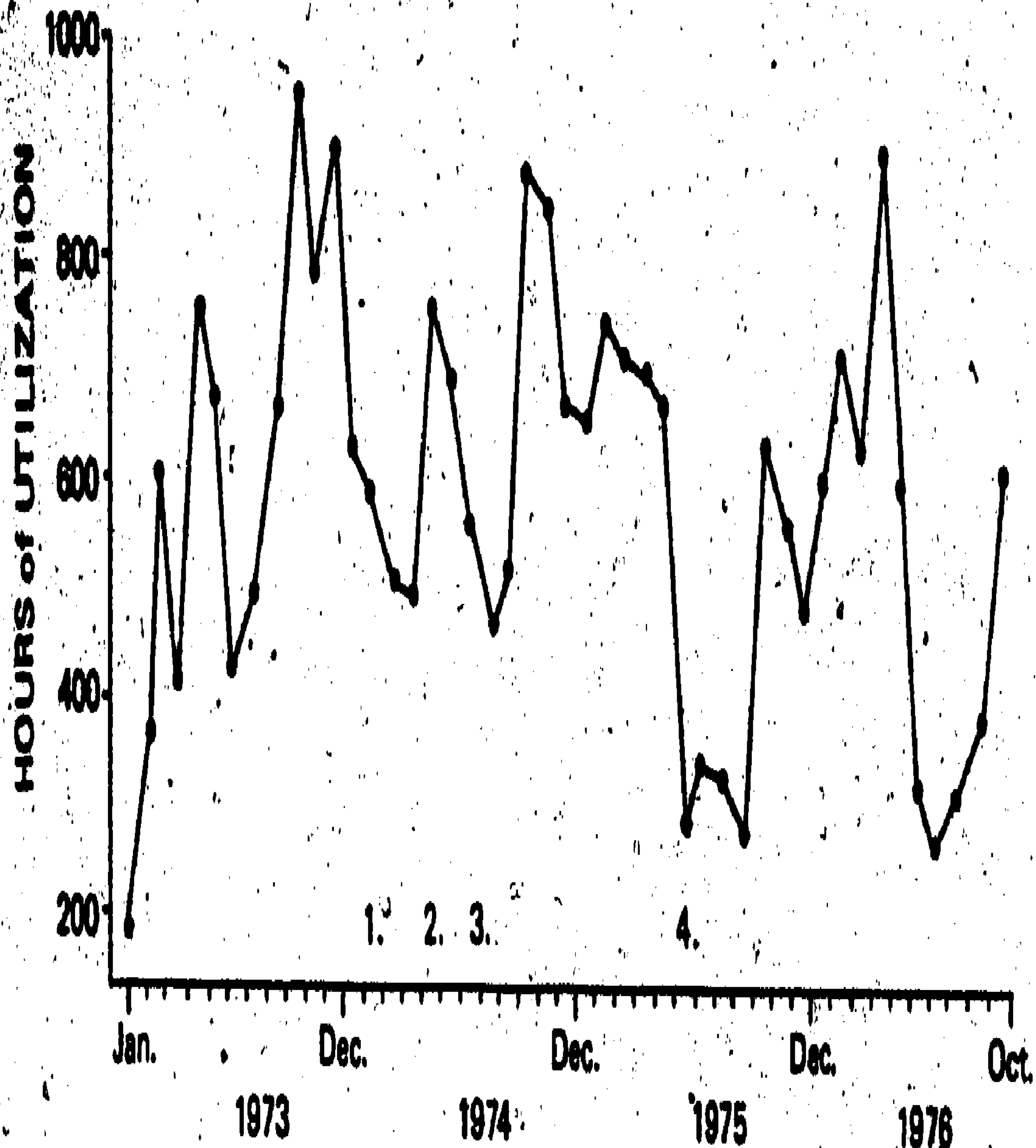
7.0 IMPACT OF THE HEALTH EDUCATION NETWORK: A SUMMATIVE VIEW

Combining the OSU and MGH data bases, the Health Education Network provides access to the largest library of health related computer assisted instruction in the world. According to two recent studies (10, 11) over one-third of existing health sciences computer-based education materials are accessible via this Network. To date, over 150 institutions have accessed either or both hosts. The largest group has been medical schools and teaching hospitals. The most common physical location for terminals (1/4 of the total) has been the biomedical library. Some 60 institutions are now members of the Network. Currently, 33 of the 54 United States Medical Schools using networking are members of the Health Education Network. Cumulatively, the Network has provided over 200,000 interactive user sessions.

8.0 THE UTILIZATION OF THE NETWORK

It is difficult to say with certainty, at any point in time, where Network utilization is on a hypothetical growth curve. If one knew, then other decisions would be easier. Figures 2 and 3 are graphs of Network usage of OSU and MGH materials. Viewing utilization, one can see four periods in the natural history of the Network when fluctuations in usage were apparently related to specific situations other than the cyclic, variable pattern of the academic year.

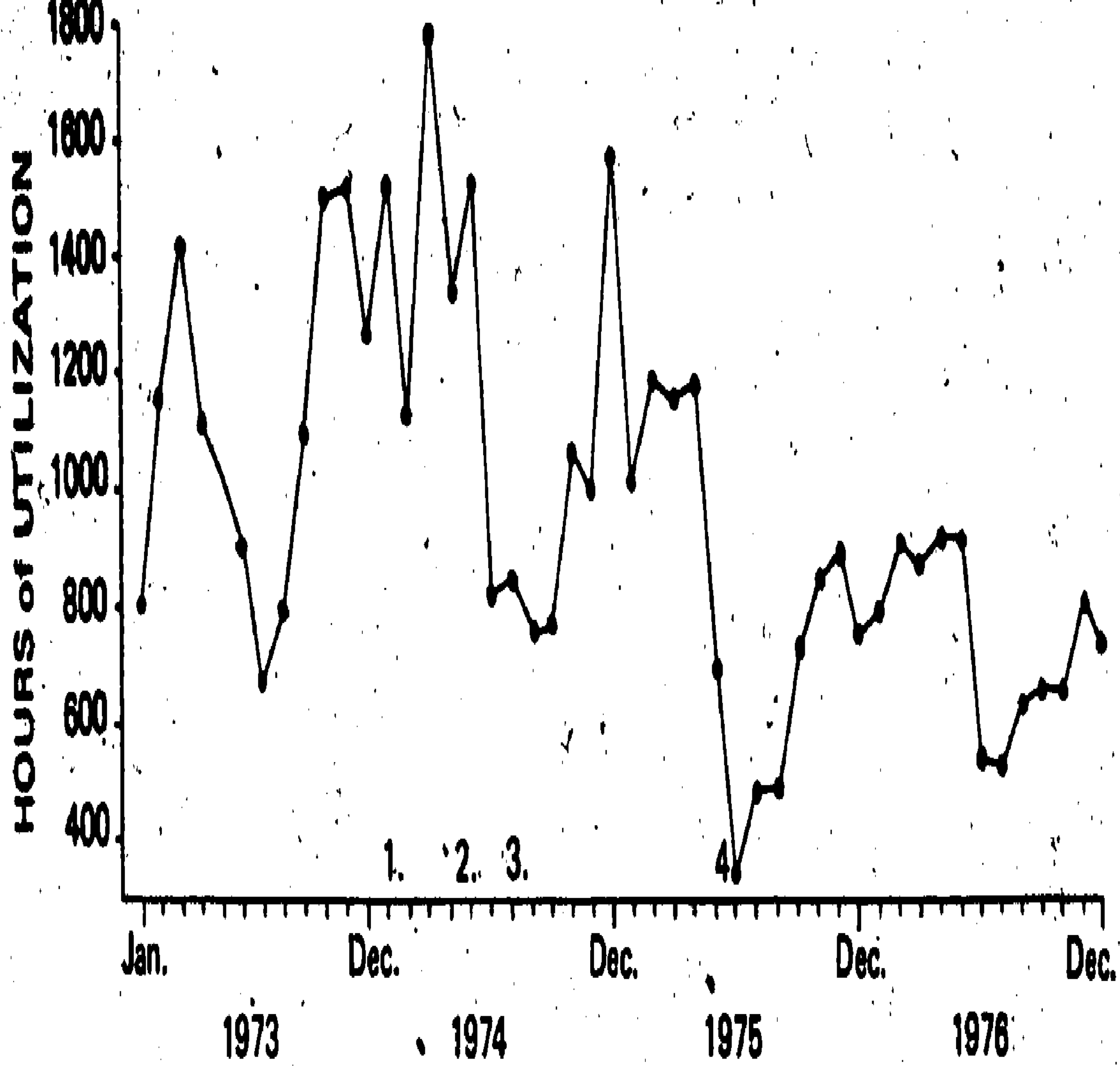
The first two periods relate to the inception of, or increase in, user fees that occurred in February and July, respectively, of 1974. In each instance, usage fluctuated somewhat following the rate change, and several institutions left the Network. The majority of institutions who left the Network when charges changed were academic institutions. The third period of fluctuation was when the CASE materials were transferred from OSU to MGH. The fourth period, involving a decline in usage, occurred immediately subsequent to the change over from Lister Hill to Health Education Network operation. Superimposed on the typical lower usage pattern of summer, was the major problem of delay involved in execution of new bilateral contracts between user institutions and hosts. Reestablishment of all new user codes was a lesser problem. Except for academic year fluctuations, usage growth seems to be upward or at least at a stable curve.



Key Events

1. $2.50/hr charge instituted
2. Case stimulations added to the O.S.U. data base
3. $5.00/hr charge instituted
4. Experimental network ends-health educ. network begins

FIGURE 2. Hourly Utilization of the Ohio State University CAI Programs by the National Network



Key Events

1. $2.50/hr charge instituted
2. Case stimulations added to the O.S.U. data base
3. $5.00/hr charge instituted
4. Experimental network ends-health educ. network begins

1976
The Ohio State University
College of Medicine

FIGURE 3. Hourly Utilization of the Massachusetts General Hospital CAI Programs by the National Network

Perhaps the major point in this discussion is that the Network has continued to grow in the face of major obstacles: imposed rate charges out of sync with budget cycles and major reorganization, in which both users and hosts had insufficient lead time for transition to a user-supported and user-directed network according to a carefully-phased and adequately-supported plan. One still cannot say with certainty where the Network is on the growth curve, but the authors feel that large untapped markets exist for health related computer-based education materials, which will continue to direct the growth curve upward.

9.0 COSTS OF NETWORKING

The current cost to the user of the Health Education Network ranges from $10 per connect hour to $4 per connect hour and depends upon the time of day during which usage is desired (prime time vs. non-prime time) and the level of commitment which the user is willing to make. The higher the commitment in usage hours the lower the unit cost to the user. The Network supports ASCII compatible terminals which can be bought for less than $2,000 and rent for as little as $50/month. The user provides his own computer terminal(s) and phone connections to the nearest TYMNET node. TYMNET, the communications network, has over 86 nodes throughout the United States in major cities. Locations with direct dial, foreign exchange or WATS line connection to node cities bear no incremental communications costs for Network access.

Hosts provide one set of documentation free with additional sets available at a nominal charge. Supplemental materials such as slides and tapes, used in conjunction with some of the OSU computer-based educational materials, are available at a nominal charge. Both hosts provide on-call assistance (technical and educational) as part of the hourly charge. If an institution requires special reports or extraordinary time from the hosts, there is a nominal charge.

Costs to the hosts include local computer costs, costs for personnel to support and maintain the hardware and software systems, and costs for personnel to provide support to users (i.e. troubleshooting, manuals). Costs are relatively low to simply maintain Network access. Costs for development and evaluation of CBE materials and for assisting users in curricular incorporation of the CBE materials are large. If hosts provided all services from which users could benefit, the user costs would have to be very high indeed. In the ideal, hosts would like to be able to provide more individual consultation on how programs could be used at each institution. Hosts would like to provide additional assistance in innovative uses of existing programs in development of new materials, in peer review and evaluation of CBE materials, in collaborating with specialty groups to devise teaching and testing materials, and in remote authoring. These activities require resources above and beyond those currently supported by Network income. Due to their congruence with the overall goals of the Network, however, they will receive continuing emphasis in Network planning and development.

10.0 BENEFITS OF NETWORKING

Dollar costs of networking have been delineated. Benefits are less easily discussed in dollars. The most important impacts of the Network have been those things which can never be quantified. In addition to providing educational benefit directly to the student, the Network has stimulated instructional computing development in the health professions and increased awareness of the potential of CBE. Faculty, students and staff users have had an opportunity to try out a new learning resource without a large front-end investment in hardware, software or personnel. They have had access to recognized materials and to the works of accomplished CBE authors. They have had opportunity to develop new lessons or alter lessons via remote authoring. And they have had access to expert technical, user services and education consultants at the host sites.

Authors have benefited via national exposure of their CBE materials which has resulted in accelerated refinement of programs. Additionally, peer awareness and approval are significant incentives to authors. Widespread peer review, although currently still an emergent phenomenon, has accelerated review and refinement of programs. Reliable, multi-institutional access to, and use of, lessons has fostered embryonic developments analogous to publication in the print domain.

Via the Network, the hosts have a relatively trouble free mechanism for sharing their CBE materials with others. This benefit of sharing forced other changes which are of long term benefit to hosts and to users. Uniform procedures for user access were implemented to make the Network as transparent as possible to the user. Standards for documentation, ancillary materials, and user services were refined to cope with the problems of interacting at a distance. Reporting systems were generalized. Methods of billing and contracts management were implemented. Program development procedures were modified in the direction of increasing generality vis-a-vis the national character of the Network leadership. Remote authoring caused additional refinement of the course development process and added new materials to the CBE library.

This particular Network has opened the door to new outreaches and projects which never before were even conceivable in the minds of medical professional groups, such as the Connecticut and Ohio Academies of Family Physicians, the National Board of Medical Examiners, the American Board of Internal Medicine and the American Academy of Orthopedic Surgeons. These groups can now consider the use of computers in continuing education, in recertification and in testing programs because the Network *is* in existence and *is* available for their use.

11.0 KEY ELEMENTS TO THE PLANNING AND OPERATION OF THE HEALTH EDUCATION NETWORK

Throughout this paper, an attempt has been made to identify variables which were of special importance in the development of the Health Education Network. In this section, an attempt is made to isolate principles for Network planning and development which appear to be generalizable.

11.1 General Principles

The LHNCBC experimental network came into being for the reasons outlined in the first section of this paper: there were hosts with programs which they were willing to share, users willing to try them, a federal agency charged with research and development in this area with dollars to support the initial experiment, and sanction by a national professional medical organization. It is safe to assume that the overlap in goals of these groups provided the environment necessary to start the Network at the time and place of its beginning.

Two other general principles are relevant. The first of these states the importance of beginning Network experiments with other than user access to resources. People value what they pay for. Free use, even for an initial start-up period, followed by a carefully phased transition to user support, although of some value, has the basic deficiencies that 1) cost ineffective uses will be tried, and 2) the upheaval brought about by cost recovery will outweigh the advantage of early free use. If costs reflect real costs accurately, it is more likely that cost effective uses will be generated from the outset. Curricular integration will be encouraged, and CBE will become a line item in the educational budget rather than a supplemental resource.

The second principle relates specifically to computer-based education. Early in the experiment, both hosts and LHNCBC management recognized the need to work with the decision-making sectors of the academic community. At the national level, this meant involvement of associations such as the Association of American Medical Colleges and the American Association of Dental Schools. At the institutional level this meant involvement of the faculty and administration who ultimately make decisions upon incorporation of this medium into the curriculum.

11.2 Philosophy of Operation

In providing a new medium of instruction, it is important to recognize the unique needs of each participating institution and tailor services to the needs of individual institutions. Institutions experimenting with CBE for the first time probably follow a common developmental sequence, but they do it in their unique time frame. It is important that institutions have access to proven courseware for demonstration purposes until such time when their faculty are ready to incorporate programs into the curriculum and such time when budgets can be rearranged to support large-scale student use.

11.3 Interaction Phenomena

In the establishment of a user-directed and user-supported Network, several important attitudes were created by people interacting over time. These included establishment of host credibility in the minds of users, development of a "trust" bond between users, development of a user community and development of a system for rapid formal and informal communications. A certain amount of group process appears essential to these attitudinal modifications, and planners should be sure that sufficient opportunity for face-to-face interaction is created in order to facilitate solution of the problems of working at a distance.

11.4 Host Dependent Variables

Critical to the continued viability and development of this Network is the presence of hosts committed ideologically to sharing scarce educational resources. Host commitment made the difference when Lister Hill Center support was withdrawn before the user group could organize an adequate independent business base to support continuation of the Network. In addition, an important element in permitting close cooperation of the hosts was the fact that the content of the data bases was essentially complementary. Hosts, therefore, did not compete for the same user resources. The third important variable was the high priority given by both hosts to the provision of user services. In the early phases, hosts proposed budgets and devoted the majority of resources to user services in contra-distinction to the perception of Lister Hill Center management whose budgets allocated the major share to hardware. The resounding success of the Network in the face of considerable adversity is in large part a function of well-developed and delivered user services.

11.5 User Variables

As an outgrowth of the interaction phenomenon, development of a user community was an obvious, critical factor in developing an ongoing operational network. Representatives of user institutions have been willing to participate and to serve on task forces, on the Steering Committee and on the Network Board of Directors. The evolvement of an active and capable cadre of leaders from the user community has permitted in large measure the progress made to date. This user community has the potential to help meet national needs for faculty development and program evaluation.

Inquiry into the factors affecting successful introduction of this new mode of instruction in any given institution revealed a single common variable: identification and involvement of a change agent. Individuals who acted as change agents were found in every area of the health community: in the biomedical library, in the computer or data processing unit, in departments or divisions of research in health education, in academic departments and divisions, in the academic administration, in community hospitals, in private practice and in professional groups. The common element was that these individuals could see the potential of the medium, had the respect of their colleagues, and could act as catalytic advocates for adopting a posture of active experimentation. Anyone seeking to innovate in the academic sector would do well to study closely the identification and nurturance of such individuals.

References

1. Combs, Bill, "TYMNET: A Distributed Network," Datamation, July, 1973.
2. Division of Computing Services, *The User's Guide to Computer Assisted Instruction at The Ohio State University College of Medicine*, 376 West Tenth Avenue, Columbus, Ohio 43210, 1973.
3. Farquhar, Barbara, et al, a chapter in Edward C. Deland's *Information Technology in Health Sciences Education*, (in press).
4. Folk, Robert L., et al, *Individualizing the Study of Medicine*, The Ohio State University College of Medicine Independent Study Program, 1976, Westinghouse Learning Corporation, New York, New York 10017.
5. Forman, Mark H.; Pengov, Ruann E.; and Burson, Jeanne L., "CAIREN—A Network for Sharing Health Care Learning Resources With Ohio Health Care

Facilities and Educational Institutions", in Edward C. Deland's *Information Technology in Health Sciences Education*, (in press).

6. Harless, William G.; Drennon, Gary G.; Marxer, John J.; Root, Judith A.; Wilson, Linda L.; and Miller, George E., "CASE—A Natural Language Computer Model", *Comput. Biol. Med.*, Pergamon Press, 1973a, Vol. 3, pp. 227-246 (Printed in Great Britain).
7. Harless, William G.; Drennon, Gary G.; Marxer, John J.; Root, Judith A.; Wilson, Linda L.; and Miller, George E., "GENESYS—A Generating System for the CASE Natural Language Model", *Comput. Biol. Med.*, Pergamon Press, 1973b, Vol. 3, pp. 247-268 (Printed in Great Britain).
8. Harless, W. G., "MERIT Progress Report—Ohio Field Study", MERIT Progress Staff, Pacific Medical Center, San Francisco, California 94120, October, 1976a.
9. Harless, William G., "A Proposal for a National Pilot Test of MERIT", prepared for the American Board of Internal Medicine, April, 1976b.
10. Kamp, Martin and Brigham, Christopher, "Current Status of Computer Assisted Instruction in the Health Sciences", *Journal of Medical Education*, Vol. 49, No. 3, pp. 279, 1974.
11. Lefever, Dale R., Ph.D. and Johnson, Jenny K., *AAMC Survey on the Use of Computers in Instruction*, 1976, Department of Academic Affairs, Association of American Medical Colleges, Washington, D.C.
12. Merola, A. John; Pengov, Ruann E.; and Stokes, Bradford T., "Computer Supported Independent Study in the Basic Medical Sciences", in Edward C. Deland's *Information Technology in Health Sciences Education*, (in press).
13. Pengov, Ruann E., "The Computer as an Aid in Medical Education", *Proceedings of the EDUCOM Fall Conference*, October 16-18, 1974, Toronto, Ontario.
14. Pengov, Ruann E., "The Evolvement and Use of Computer Assisted Instruction (CAI) in Health Sciences Education at The Ohio State University College of Medicine", in Edward C. Deland's *Information Technology in Health Sciences Education*, (in press).
15. Pengov, Ruann E., et al, "OSU Participation in the LHNCBC Experimental Network, June 1, 1972—November 30, 1976", *Final Report to the Lister Hill National Center for Biomedical Communications*, National Library of Medicine, National Institute of Health, Public Health Service, and Department of Health, Education and Welfare, 1977.
16. Stead, E. A., Jr.; Smythe, C. McC.; Gunn, C. G.; and Littlemeyer, M. H., "Educational Technology for Medicine: Roles for the Lister Hill Center", *Journal of Medical Education, 46*, July, 1971, Part 2.

Chapter 18

by Casimir A. Kulikowski
Sholom Weiss
and Aran Safir*

Computer Networking in the Development of a Glaucoma Consultation System

A computer-based consultation system for the diagnosis and treatment of glaucoma has been developed in collaboration with a network of investigators from five glaucoma research centers. Computer networking enables members of this Ophthalmological Network (ONET) to access the consultation program, which runs on two large time shared computers at Rutgers and Stanford Universities.

The Ophthalmological Network was established in 1974 to promote the development and testing of computer consultation and research support programs for clinical investigators in Ophthalmology. This organization evolved from a collaboration that began in 1971 between Dr. Casimir Kulikowski of Rutgers University and Dr. Aran Safir of the Mount Sinai School of Medicine within the framework of the Rutgers Research Resource on Computers in Biomedicine. The goal of the Resource is to support interdisciplinary research on computer-based modeling in biomedicine, particularly in the application of artificial intelligence techniques for problem solving.

The glaucoma consultation program is the first prototype of a medical decision-making system that uses a novel method for describing diseases in the form of a causal associational network (CASNET). Many of the subtleties and complexities of physicians' reasoning can be incorporated in a systematic yet flexible manner into such a descriptive structure. The clinical course, pathophysiological mechanisms, classifi-

* Casimir A. Kulikowski is affiliated with the Department of Computer Science, Rutgers the State University of New Jersey, New Brunswick, New Jersey. Sholom Weiss and Aran Safir are affiliated with the Institute of Computer Science, Mt. Sinai School of Medicine of CUNY, New York, New York.

cations and treatments of disease are all contained in the CASNET model. The findings of an individual patient can be interpreted through a CASNET computer model using various reasoning strategies that take into account the uncertainties of clinical information. By representing in the computer detailed patterns of disease evolving with the passage of time, these reasoning strategies are able to deal with the problems of patient management over multiple follow-up visits. Sequences of therapies for the various types and stages of progression of glaucoma have also been incorporated into the model. As a result, the program makes specific recommendations for further testing and for medical and surgical therapy.

Another innovative feature of CASNET is its ability, for a particular case, to present alternative opinions and reasoning derived from different expert consultants. The ONET members share in the development of the program by suggesting how such knowledge is to be incorporated into the glaucoma model.

After proving the feasibility of the initial CASNET model, the ONET was formed when Drs. Steven Podos and Bernard Becker of Washington University and Drs. Irvin Pollack and Lawrence Viernstein of Johns Hopkins University began collaborating with the Rutgers-Mount Sinai group to help develop the glaucoma consultation program into a proficient clinical tool.

In 1974 an added impetus was given to this work when the Mount Sinai-Rutgers Health Care Computer Laboratory was established by HEW to promote the further development of computer consultation systems for health care delivery. Shortly afterwards, ONET was expanded to include Dr. Jacob Wilensky of the University of Illinois at Chicago, and Dr. Michael Kass who now coordinates the ONET activities at Washington University. Most recently, the ONET has been enlarged to include Dr. Douglas Anderson of the University of Miami.

Since March 1975 the ONET members have been testing the consultation system and suggesting changes and improvements in the structure of its knowledge and its logic. Most recently, a concerted effort has been made to incorporate the results of various research studies and alternative opinions into the glaucoma mode: A data base of glaucoma cases, together with computer programs for retrieval and analysis, has also been established with the goal of facilitating joint clinical studies among the ONET members.

Chapter 19

by Barbara B. Farquhar

Sharing Computer-based Simulations for Clinical Problem-Solving: A Host's Perspective

1.0 INTRODUCTION

Development and maintenance of computer-based educational materials are expensive, and few individuals or institutions have extensive experience in production and distribution. When medical educators recognized the potential benefits in sharing scarce and valuable resources in 1971 (9), they recommended that the Lister Hill Center of the National Library of Medicine establish a biomedical communications network. Lister Hill's Experimental Computer-Aided Instruction Network (12) was a success, and it served as the precursor for the Health Education Network (10). The computer-based educational materials available on the Health Education Network vary in the audiences they are intended to serve, the instructional strategies used, the hardware and software which support them, and the methods by which they were developed and are maintained. Currently, there are two hosts on the Network: the Ohio State University College of Medicine and the Massachusetts General Hospital. A thorough description of Ohio State's philosophy and materials appeared in the proceedings of the 1974 EDUCOM Fall Conference (8). This paper is intended to complement that one and provide a view of networking from Massachusetts General Hospital's perspective. An expanded version of this paper will be available in 1977 in the book *Information Technology in Health Sciences Education* which is edited by Edward C. DeLand and published by Plenum Press, New York.

2.0 CLINICAL MEDICAL EDUCATION

The purpose of clinical medical education is to prepare students to practice independently. Traditionally, clinical education has taken the form of a clerkship in which students are expected to learn to imitate the behavior of their preceptors. There are dangers and limitations in teaching students solutions to problems and expecting them to devine the process by which solutions are reached. Usually, the focus is on the product rather than the process. Some students, who are able to discriminate critical elements of clinical problem-solving, develop clinical competence (4) readily. Other students, who may have superior knowledge and ability to recall facts, may need explicit, systematic instruction and supervised practice in problem-solving (11) in order to apply knowledge to solve patients' problems. Medical educators agree that supervised practice in clinical problem-solving is the best way to develop clinical judgment, but many factors work together to make systematic, supervised practice costly and difficult to achieve in a hospital setting (6).

3.0 COMPUTER SIMULATIONS IN CLINICAL EDUCATION

Computer-based simulations can supplement other formal instruction and do offer a number of advantages for students, clinical preceptors, and directors of clinical education. Simulations can be used at each student's pace and convenience since access to the simulations is available continuously and repeatedly. In a simulation, the computer can serve several functions: it provides information about a simulated patient; it permits the student to interact with the patient, to make decisions, and to take actions to which the patient responds; and it assumes the role of the supervising physician by providing consultation and guidance, and by relating one case to another. The computer can also provide recent, relevant references for a problem. The computer can score the student's performance objectively and compare the student's performance to that of others, provide the student and the instructor with a critique of the student's performance, and suggest remedial work if that is indicated. The student has responsibility for solving the patient's problem, but the learning process is without risk to real patients and is truly individualized. The student is supervised through the computer by the teaching physician who wrote the medical content and specified the instructional strategy. The teaching physician can select a simulated patient with the clinical problem(s) and constellation of signs and symptoms that the physician wishes the student to manage. The computer also gives feedback to the physician responsible for teaching the student. After reviewing the performance of their students, physicians may wish to alter either the content of their courses or the instructional techniques they are using.

In addition to providing individual records of performance, the computer can prepare a composite or can aggregate data for use by a curriculum committee or clinical education director. Objective records of student performance at the clinical level and patterns of behavior common to a class are difficult to obtain with traditional educational methods, but records of many students' actions in identical clinical problems can provide hard data to help evaluate both the students and the clinical education program in which they are participating. A thorough assessment can pinpoint strengths and weaknesses in a clinical training program and permit its director to modify the program until it meets its objectives as measured by student performance (7).

Costs to develop computer-based patient simulations are relatively higher than those to prepare other self-instructional materials, but these costs must be related to potential benefits. Simulations give students supervised practice in decision-making, and this is difficult for other media to achieve. Simulations can save time for practicing clinicians because simulations obviate the need for some repetition. Simulations can overcome some of the limitations of teaching in a hospital setting and ensure that all students have encountered "patients" with a sufficient variety of clinical problems. The computer's relatively high cost as an educational tool can probably best be justified when: 1) its unique capabilities are used as in the case of simulations; 2) it is used to support a kind of education that is expensive in its own right, such as clinical medical education (1); and 3) the programs are shared among many institutions.

3.1 Massachusetts General Hospital Computer-Based Simulations

Between 1966 and 1976, Massachusetts General Hospital has developed and tested a library of over 30 computer-based simulations to teach and test (3) clinical problem-solving. Each simulation of a clinical problem includes many cases, but all cases within a simulation share a common instructional strategy and focus on a general problem such as jaundice, abdominal pain, or coma. The simulations were developed in the Laboratory of Computer Science, a unit of the Department of Medicine, where professionals in medicine, computer science, and education work together on problems related to patient care and medical education. The cost to develop the library of computer simulations is difficult to estimate precisely, but it is well over $300,000.

Development was possible because the simulations were created in a facility where there were other on-going projects. The development costs cited are merely the incremental costs to support staff engaged in the production and developmental testing of the programs.

3.2 Development of a Simulation

Before an author develops a patient simulation, the author considers many factors and makes a number of determinations. The author defines the users, describes any prerequisites, describes terminal behavior, defines measures of performance, and establishes minimum levels of acceptable performance. Instructional objectives guide the author in developing the program and serve as criteria for its evaluation. Authors developing programs to be shared by users at many institutions avoid special procedure-oriented programs and controversial content because procedures and accepted practice vary from institution to institution.

After the author has determined whom he is going to teach with the patient simulation, what students should be able to do after they have used the simulation, and what kinds of patient simulations will help the users develop the desired competencies, the author must select an instructional strategy or program format for the simulation. In selecting an instructional strategy, the author considers the following:

- the purpose of the program (assessment, problem-solving practice, continuing education);
- the audience (medical students with limited vocabulary, advanced medical students or house officers with adequate vocabulary, or specialists with extensive vocabulary in limited areas);
- the user's task (diagnosis or management of a chronic or acute problem);
- structure of the task (unstructured, as in the diagnosis of a complex problem, or structured, as in the management of an acute, well-defined problem);
- sequence of actions (critical or not important);
- timeliness of actions (critical or not important);
- reaction of the simulated patient to actions taken by the user (patient static or dynamic);
- passage of time (one encounter or multiple encounters); and
- method of interaction (multiple-choice, vocabulary list, or free text).

If some cues are tolerable, the author may select a multiple-choice or prepared vocabulary list for the user to employ during interaction with the program. If cueing will weaken the program, the author may select either an extensive index or free text for user interaction.

The instructional strategy selected will determine, to a great extent, the amount of time required of the author in developing the patient simulation. A program author must specify the content and logic of the program, which means that he must describe how the program and patient should respond to any action that may be taken by the user at any and every point in the simulation. If the author selects a structured, multiple-choice instructional strategy, the number of actions available to the user is limited, and the author's job is minimized. If the author chooses an unstructured format in which the user may use free text and type any response to the simulated case, the author must anticipate the response and describe the actions to be taken by the program and patient as a result of the user's responses. In the latter case, the author's job is complex and time-consuming.

Many physicians do not want to learn about computers or computer programming but do want to create patient simulations. MUMPS (*M*assachusetts General Hospital *U*tility *M*ulti-*P*rogramming *S*ystem) "driver" programs permit authors to create and modify complex simulations without writing MUMPS program code by using a program format or instructional strategy that already exists in the form of a driver program. A driver is a set of special purpose utility programs designed to execute a repetitive task. Definition of the task is independent of the data, but the data "drive" or control the execution of the program. Drivers reduce the amount of program code that must be available on the system; one copy of a driver can be shared simultaneously by many users. A driver is slower on execution than a specially coded program, but it is more economical than a custom-made program in terms of reduced storage and a shorter time to create a new simulation. The driver includes definition of options, format, branching, and file structure of the data on which it operates. An author who is not a programmer can develop content and logic, specify branching, and enter his data base for a driver using an editor program. An editor is an on-line author aid that elicits the data base and logic from the author, checks syntax, and verifies that new data or modifications are consistent with the structure of the driver. The specifics of the computer and the language it uses are transparent to the author.

The author also prepares program documentation for the user's manual, describing the program and including illustrations of interactions and any supplementary visual materials that are needed to operate the program. After the author has a rough draft of his simulation and associated materials, and the program is technically correct and operates the way the author intended, developmental testing begins.

The cycle of program testing and revision is repeated until the patient simulation consistently produces the instructional objectives established by the author. For the first review, the author invites local experts and members of the target population to use the program individually. The author is present when these users try the patient simulation to observe their interaction with the program and to collect comments and suggestions. The author notes non-verbal responses and any apparent content or logic problems. For members of the target population, the author administers measures of performance to establish whether or not the target population meets prerequisites for program use and whether or not the program achieves the instructional objectives set for it. After collecting data and evaluating performance measures, the author modifies the patient simulation accordingly. S/he may then invite the same or a different group of experts and intended users to take the simulation, and s/he repeats the processes of data collection and revision.

When the author is satisfied that the program functions properly and meets the standards s/he has set for it, the program is released or made available within Massachusetts General Hospital for use by its medical students and staff. They comment freely on the content and structure of the program, and their suggestions form the basis for further revision and refinement. Next, the program is released to Countway Medical Library at Harvard Medical School, and additional suggestions are collected. The program is revised frequently at each step. Next, the author may arrange for field test by inviting target populations at remote institutions to use the program and the performance measures associated with it,

In many instances, a specialty board or nationally recognized panel of experts has been invited to review a new or revised patient simulation. Sometimes, the members of the specialty board or panel review the simulations together at Massachusetts General Hospital or at their headquarters. At other times, individuals review the programs at their home institutions using the Health Education Network. The programs are revised to conform to recommendations of specialty boards. When national standards are revised, simulations are modified to meet the new standards.

When the program performs adequately to meet objectives set for it, the program is released to the Health Education Network. Program documentation is sent to all users, and a notice is put on the Network that a new simulation is available. Users are asked to try the new simulation and to comment on it. Comments and suggestions are collected, and the author makes revisions to the simulation. A national network of users helps to keep the author informed of recent research and publications related to his simulation, and users' constructive criticisms are of inestimable worth in refining and updating programs.

As long as the simulation is available on the Massachusetts General Hospital computer systems, the author (or a designee) is responsible for following and maintaining it. The author answers users' comments, keeps program content and logic current with generally accepted practice and revised national standards, and keeps references in the program up-to-date. The programs undergo continuous revision as long as they are in use.

4.0 MUMPS

Development and maintenance of the computer-based patient simulations are facilitated by a powerful yet simple programming language, MUMPS (*M*assachusetts General Hospital *U*tility *M*ulti-*P*rogramming *S*ystem). Program authors do not need special training in mathematics, electrical engineering, computer science, or computer programming to create patient simulations in MUMPS (5). The language is easy to learn and use; people with no previous training are able to write and debug a basic MUMPS program in less than a week with no other aid than a MUMPS programming manual. Learning to use the full intricate capabilities of MUMPS, however, and to use its associated global file structure optimally takes time and practice.

MUMPS is well-suited to creating medical education programs because it is an interpretive system which provides program flexibility. All programming is done on-line, and it is easy to detect and correct errors. The corrected version can be run immediately and tested. Medical education programs undergo changes continuously in the process of their development and refinement, and MUMPS allows these changes and additions to be made easily in operating programs without hampering their use.

Currently, all of the Massachusetts General Hospital simulations are written in the hospital's dialect of the MUMPS language (2). A standard MUMPS language has been specified (13), and the hospital plans to convert its simulations to standard MUMPS. Versions of MUMPS also run on Burroughs, IBM, Artronix, and Data General equipment as well as on several models of Digital Equipment Corporation computers. A MUMPS Users' Group and a MUMPS Development Committee exist to facilitate communications among users and to promote language commonality for program interchange.

The text processing capabilities and file structure of MUMPS make it especially suitable for patient simulations. The language provides elaborate string or text processing features in addition to standard arithmetic and Boolean capabilities. MUMPS handles string data in symbolic form and performs syntax checking by pattern-matching to determine whether a string contains a predetermined sequence of numerics, alphabetics, punctuation, or combination of these. Strings can be assembled, disassembled, modified, or searched easily. MUMPS uses a common, random-access file structure. Data files are stored in dynamic tree-structured form. Nodes of the tree are referenced symbolically using N-dimensional subscripts. Files are created and manipulated dynamically, and space is allocated as it is needed. A datum in the file can be changed without disturbing other data in the array. Together, these features provide capabilities which make it possible to write sophisticated simulations that are easy to modify on a relatively inexpensive computer system.

An experienced computer programmer who is creating a patient simulation may elect to use a driver but insert MUMPS code or call other programs to perform certain special functions. Most computer languages distinguish between programs (executable code) and data. They permit programs to create, reference, change, and exchange data but do not allow programs to modify their own or other programs' code. In MUMPS, a program may create, modify, or expunge part or all of its own executable code. A driver usually interprets data and executes the option indicated. When a driver finds MUMPS statements, it creates a program and executes commands.

MUMPS simulations run at Massachusetts General Hospital on any of six PDP-9 and PDP-15 computers. Each computer has a minimum of 32K of core, a 3-million character fixed-head disk, and two 30-million character movable-head disks. The interpreter uses 12K of core, and the remainder is divided into 20, 1K user partitions. Programs are core-resident, and they are interpreted rather than compiled. Extended core supports more than 20 simultaneous users on a single computer. The system supports ASCII terminals at the rate of up to 2400 baud in-house. Systems are half-duplex and asynchronous. Average response time is 0.5 seconds, but response time varies as a function of the task to be performed, particularly as a function of disk activity and of the number of other users of the system. Response time may be as long as 2-5 seconds when the system is loaded heavily.

5.0 SYSTEM FEATURES

There are several features of the patient simulation system that permit users to ask for and receive assistance. As the system has developed and become more sophisticated, features which have been found useful have been incorporated into new programs. Not all programs have all features, but some desirable system capabilities are described below.

The patient simulations undergo constant review by their authors and users. Users communicated with authors through a system COMMENT feature; at any point in a program, the user may type the word COMMENT and follow this with a question or suggestion. After he has made the comment, the user returns to his place in the program sequence. Each day, both the comments and the names of individuals making them are printed for program authors. Authors review and respond to comments individually by addressing their answers to the individual users who made them. The next time the user logs on the system, he receives a response to his comment. The COMMENT feature provides authors with continuous feedback from users. The information makes it easy for authors to detect and correct problems with program logic or content, so it accelerates program development. Users find that answers to comments make the simulations truly individualized, personal experiences. Faculty at remote institutions appreciate having the programs maintained and updated by responsible faculty authors at Massachusetts General Hospital.

During the course of a patient simulation, the user may need assistance; in this case, the user may type the word HELP. The program will ask whether he needs assistance with the format of his answer or if he requires the kind of review and discussion of the patient's problem that a teaching physician or consultant might provide. After the program provides the help requested, it returns the user to his place in the simulation. The user may call the HELP function repeatedly within a single patient simulation.

Each call for assistance and the type of help requested are noted on the daily output for program authors. When authors find many calls for procedural assistance, they know that directions are not clear. If there are numerous calls for medical guidance from the target population or individuals with more advanced training than the target population, the author may reconsider and revise the characteristics or behavior of the simulated patient. Alternatively, the author may enrich the kind of medical guidance the program provides to users when they request advice or consultation.

Frequently, a user finds it necessary to interrupt a patient simulation. Many users of clinical simulations have responsibility for patients and must terminate interaction with a simulated patient to take care of a real patient. The user may type the word STOP at any point in the simulation, and the program will end and print THANK YOU.

At times, there are interruptions in the computer or the communications network, and the user may lose connection with the host computer system. If there is an interruption in the patient simulation either because the user typed STOP or there has been a technical problem, the user may wish to RESTART the simulation at the point at which he left it. A few programs give the user the option to RESTART in the middle of the simulation, to go on to another case, or to go back to the beginning of the interrupted case.

When the patient simulation is long and complex, or in the case of an interruption and RESTART, the user may wish to have a summary of the findings to a given point in the simulation. To get this summary, a few programs allow the user to type the word REVIEW, and the findings are printed so that he can review the data he has collected before he continues the simulation.

6.0 SIMULATIONS AND THEIR USES

The simulations use combinations of static patient models for collecting information leading to diagnosis and management as well as dynamic models which represent a physiologic process or disease which charges over time. In the dynamic models, the simulated patient responds to actions taken by the user. Much of the program content is oriented to emergency medicine. The simulations do provide guidance and facts, but they emphasize problem-solving skills.

Programs are used in a variety of ways. Some programs are required; others are recommended, and still others are available for *ad lib* use. They may be used in conjunction with independent study. During the past 4 years, more than 150 institutions in more than 40 states have used approximately 50,000 hours of time on the Massachusetts General Hospital system for medical education by accessing programs through the Network. Major users are medical students (68%) and practicing physicians (18%); the remaining usage (14%) is by nurses, dentists, and allied health personnel. Institutions who have access to the programs consist of medical schools (68%), hospitals (16%), national medical organizations (11%), and physician groups (5%). Medical students in their clinical training use the programs to practice clinical decision-making. Practicing physicians use the programs for self-assessment and for continuing education. Many of the programs have been reviewed and approved by the Harvard Department of Continuing Education for Category I credits toward the American Medical Association's Physician's Recognition Award. Potential users of computer-based materials for licensure, certification, and recertification are great.

7.0 NETWORKING

Massachusetts General Hospital has received many requests from other institutions to share the patient simulations. Sharing educational resources is a commonly accepted academic responsibility, and sharing the end products of federally-sponsored research and development is implicit in grants and contracts. Sharing can be accomplished in a variety of ways. To date, Massachusetts General Hospital has found that sharing through networking permits a small staff at the host institution to provide its best, most up-to-date programs to the most recipients at the lowest cost.

In networking, the program resides on the computer at the originating institution, and recipients are provided access to the programs through a telecommunications network. The originators maintain and update one copy of the program, and they supply recipients with extensive documentation on the *use* of each program. User documentation, as opposed to program documentation, changes less frequently and does not need to be modified extensively for each recipient. Recipients share the cost of program distribution. A small staff at an originating or host institution can maintain and update an entire library of programs and make these available to a wide audience. Networking appears preferable to program transfer when programs are changed frequently and the resources at either the originating or receiving institutions are very limited. Costs to both the host and recipient are frequently less for networking than for program transfer for an equivalent end product.

7.1 Host Advantages and Costs

There are several advantages to the host in networking. By providing access to programs through a network, the host fulfills his obligation for sharing the end products of research and development responsibly and economically. Networking accelerates the development and refinement of programs because many more people use and comment on programs than they would if programs were restricted to use in-house. Authors have extra incentive to keep programs up-to-date and non-parochial when they know their materials are used nationwide by an audience of students and peers. Documentation required for in-house maintenance of programs is relatively simple. A small staff at a host institution can maintain a library of programs and provide access to that library at a reasonable cost.

Network host costs are for computer resources, communications services, and personnel. Costs for computer resources may be fixed if a portion of a computer is dedicated to networking, or costs may vary with usage if the computer is capable of serving others when not being accessed by a user of the network. Similarly, communications services may be at a fixed rate with unlimited usage of a fixed number of lines, or they may vary with the locations of users, number of characters transmitted, or total amount of usage (measured in time). Personnel are required at the host institution to maintain computer programs and systems, develop and update documentation, respond to users' questions, and handle billing. Personnel costs to perform maintenance functions are relatively low. Personnel costs to provide network users with the technical, educational, and faculty development support could be very high if the host provided users with all the services from which users could benefit. There are economies of scale; in general, the higher the level of usage, the lower the unit cost to the host.

7.2 User Advantages and Costs

Network users enjoy many benefits. Accessing programs through a network permits people who have an interest in trying computer-based educational materials to do so without a large capital investment. The programs they receive are the most current available, and the recipients need not be concerned about transforming them for local computer requirements. Network users do not require a local computer installation, technical staff, operation staff, or faculty commitment to update programs. Users share the resources of a complex facility such as Massachusetts General Hospital, and they have access to both the programs and the people responsible for them. Networking also allows the potential for fostering remote authoring of simulations.

Two disadvantages some users attribute to networking are cost and control. Renting programs is often perceived as more costly than running programs on a local computer system. However, cost considerations tend to be deceptive because users usually consider and compare costs assuming that program transfer has taken place. Costs of transferring and updating may well be greater than those for networking. Then, too, users do not always compare costs of equivalent end products. Massachusetts General Hospital has redundant computer systems, can guarantee access through a reservation system, has a full-time operations staff, has personnel on-call for educational and technical assistance, and has faculty who respond daily to users' comments and revise programs. Providing equal services at a remote site can be extremely costly, especially if the sole use of the computer were for computer-based educational materials.

The other disadvantage to networking that some users perceive is lack of absolute control over program content. Control is exerted on program content by external sources such as nationally accepted standards, generally accepted medical practice, and comments from users. Complete control over program content could be exercised by each using institution if MGH permitted remote authoring and allowed each institution to tailor and store a copy of each program. However, currently the MGH avoids writing procedurally-oriented programs and controversial material which tend to have little general applicability over a range of institutions.

Initial and ongoing costs to network users include charges for computer terminals, local and long distance communications, computer usage at the host institution, documentation, other supplementary materials, and assistance from host personnel. A computer terminal must be compatible with the network to which it connects; not all computer terminals will work on all networks. Some sophisticated terminals with capabilities to handle graphics, color, microfiche, film, slides, and sound must be purchased and cost over $10,000. Basic alphanumeric terminals can be purchased for under $2,000 and rent for as little as $50/month. Communications charges may be as high as a long-distance call to the host on a specially conditioned telephone line or as low as a local phone call to a network node on a regular phone line. Charges for use of the host computer may be based on the time that the user is connected to the host's system, or they may be figured on the amount of central processing unit used, characters transmitted, or storage required. Documentation and supplementary materials, such as slides or tapes to be used in conjunction with programs, may be provided to users or available at a fee. The host usually makes staff available for on-call assistance to answer users' questions. If the host provides special services for individual users, there may be additional charges for these services.

The costs to recipients for networking with Massachusetts General Hospital are limited and under the user's control. Prices for accessing the programs range from $10/connect hour to $4/connect hour; rates depend on the user's minimum monthly commitment for usage and the time of the day or night the person uses the system. Users provide their own terminals, which rent for as little as $50/month, and pay local telephone charges between their sites and the nearest TYMSHARE nodes. Users have limited liability and may change their levels of service or terminate service on 60 day's notice.

8.0 SUMMARY

One of the most important goals of clinical medical education is the teaching of the problem-solving process. The process is best taught in clinical settings, yet the constraints imposed by patient care and the unpredictability of the clinical problems that will be available for study may limit the effectiveness of such educational experiences. Computer-based patient simulations can serve as a valuable adjunct to education in clinical settings. Simulated clinical problems allow medical students to receive uniform instruction by computer and to learn from their mistakes in clinical judgment in a variety of cases without risk of harm or harassment to real patients, without undue repetition of materials by their clinical preceptors, and without regard to the vagaries in clinical problems presented by a particular patient population. Patient simulations can also be used by graduate physicians for continuing education, self-assessment, and for certification and recertification.

The Massachusetts General Hospital has developed a series of computer-based simulations that stress the process of decision-making and that supplement practice in clinical problem-solving. Programs are written in MUMPS, a language well-suited to clinical simulations. The programs, which are carefully tested and continually revised, are available locally and throughout the United States via the Health Education Network.

References

1. Association of American Medical Colleges, 1973, Undergraduate medical education: elements, objectives, costs, Report of the Committee on Financing of Medical Education, Association of American Medical Colleges, Washington, DC.
2. Barnett, G.O. and Greenes, R.A., 1970, High level programming languages, *Computers and Biomedical Research* 3:488-494.
3. Farquhar, B.B., Barnett, G.O., Goldfinger, S.E., and Dineen, J.J., 1973, Computer-based examination: a technique to evaluate clinical competence, *in* "Instructional Technology in Medical Education" (J.P. Lysaught, ed.), pp. 91-99, Rochester Clearinghouse, Rochester, NY.
4. Feinstein, A.R., 1967, "Clinical Judgment," Williams and Wilkins, Baltimore, ML.
5. Greenes, R.A., Pappalardo, A.N., Marble, C.W., and Barnett, G.O., 1969, Design and implementation of a clinical data management system, *Computers in Biomedical Research*, 2:469-485.
6. Hoffer, E.P., Barnett, G.O., Farquhar, B.B., and Prather, P.A., 1975, Computer-aided instruction in medicine, *in* "Annual Review in Biophysics and Bioengineering" (L.J. Mullins, ed.), vol. 4, pp. 103-118, Annual Reviews, Inc., Palo Alto, CA.
7. Mankin, H.J., 1975, Development of educational characteristics for orthopaedics, unpublished manuscript, Massachusetts General Hospital, Boston, MA.
8. Pengov, R.E., 1974, The computer as an aid in medical education, *in* "North American Perspective" Proceedings of the EDUCOM Fall Conference pp. 141-155, EDUCOM, The Interuniversity Communications Council, Inc., Princeton, NJ.

9. Stead, E.A., Smythe, C., Gunn, C.G., and Littlemeyer, M.H., 1971, Educational technology in medicine: roles for the Lister Hill Center, *Journal of Medical Education*, 46:Part II.
10. Tidball, C.S., 1977, The Health Education Network, *in* "Information Technology in Health Sciences Education" (E.C. DeLand, ed.), in press, Plenum Press, NY.
11. Travers, J.F., 1972, Types of learning III - problem-solving, *in* "Learning; Analysis and Application", (2nd ed.), pp. 128-151, David McKay Company, Inc., NY.
12. Wooster, H., 1977, The LHNCBC Experimental CAI Network, 1971-1975: an administrative history, *in* "Information Technology in Health Sciences Education" (E.C. DeLand, ed.), in press, Plenum Press, NY.
13. . . . . "MUMPS Language Standard" (ed. J.T. O'Neill), 1976, Systems and Software Division, Institute for Computer Sciences and Technology, *National Bureau of Standards*, Washington, DC.

Acknowledgements

Development of the MUMPS language and the simulations was supported in part by Grant No. 5 R01 HS 00240 from the National Center for Health Services Research and Development. Development of the simulations was supported by the National Fund for Medical Education. Distribution of programs was supported initially by Contract No. N01-LM-2-4716 from the National Library of Medicine.

Chapter 20

by Charles Oleson

INTERNIST: A Computer-based Consultation

1.0 INTRODUCTION

INTERNIST is a project being developed at the University of Pittsburgh. The principal investigators on the project are Harry Pople, a computer scientist/business professor in the Graduate School of Business Administration, and Jack D. Myers, a University Professor in the School of Medicine.

The primary objective of the INTERNIST project is to provide an expert computer-based consultation capability for diagnosis of the diseases of internal medicine. In order to attain this objective the INTERNIST program would have to be able to:

- work effectively on cases where more than one disease is present since this is true of many clinical cases;
- take as input history, signs, symptoms, and lab data in any order;
- partition the manifestations into coherent subsets in order to focus effectively on the actual individual problems in the case—in some cases this also requires that misleading input be dealt with as well;
- pursue through interrogation further information about the case in order to substantiate or disconfirm the impressions created by the initial manifestations.

In order to explain how project staff attempt to satisfy these requirements, this paper describes the structure of the medical data base and INTERNIST programs and discusses areas slated for further development.

2.0 DATA BASE

The INTERNIST data base requires approximately 600,000 bytes of disk storage and is estimated to contain approximately 75% of the diseases of internal medicine. These diseases are incorporated in a tree structure organized, typically, according to organ system. For instance, LIVER DISEASE is a top level node in the disease tree under which come HEPATOCELLULAR DISEASE, NEOPLASMS OF THE LIVER, CHOLESTATIC DISEASE etc. Similarly HEPATOCELLULAR DISEASE is broken down into HEPATOCELLULAR REACTION, TOXIC HEPATOCELLULAR INJURY, and ABNORMAL IMMUNITY HEPATOCELLULAR INFECTION. HEPATOCELLULAR INFECTION is then reduced to HEPATITIS A and B, INFECTIOUS MONONUCLEOSIS and LEPTOSPIROSIS WITH HEPATIC INVOLVEMENT. These last are the terminal or diagnosis level nodes. The hierarchical structure of INTERNIST files is illustrated in Figure 1.

Associated with each of these diseases is a list of manifestations that describe the disease. For instance, some of the manifestations associated with HEPATITIS A are JAUNDICE, FEVER and VOMITING.

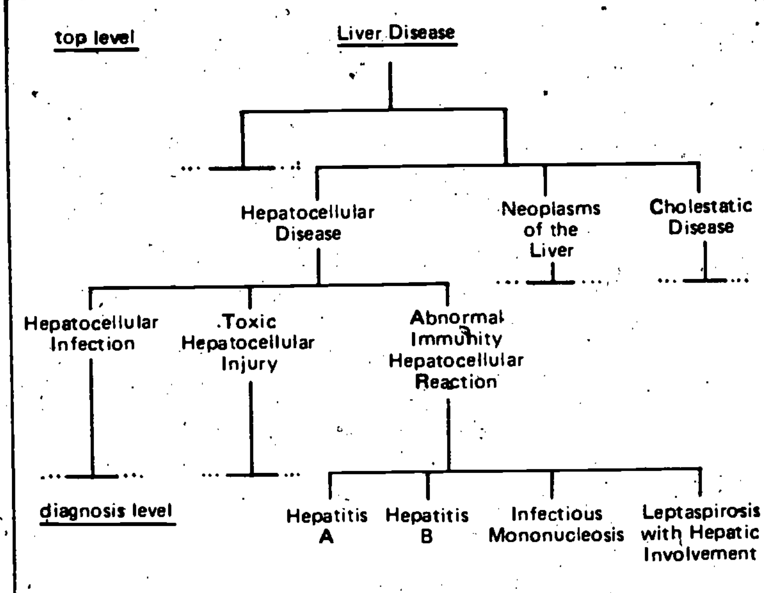


Figure 1. Hierarchical Tree Structure of INTERNIST Diseases

Two values are also associated with each manifestation under a disease. The first number is called the evoking strength. It is a value from 0 to 5 which indicates the impact the manifestation has in causing clinicians to think of the disease in question. A value of 0 means the manifestation is very non-specific and could be caused by any number of diseases (e.g. FEVER) or indeed may require no explanation (e.g. SHELLFISH INGESTION RECENT HX). A value of 5 means the manifestation is pathognominic with the disease. In other words, if the manifestation is present the disease is present.

The second number is called the frequency. It has a value from 1 to 5 that indicates the frequency of occurrence of the manifestation in the disease. A value of 1 here indicates that the manifestation occurs but infrequently in the disease. A value of 5 means that the manifestation is always present in the disease.

With each of the manifestations a number of properties are associated:

- diseases in which the manifestation occurs
- import of the manifestation
- implications of the manifestation
- type of manifestation, including source and costliness

The first and most important of these is the list of diseases in which the manifestation occurs. A second property, the import of the manifestation is intended to be a measure of the degree to which the manifestation need be explained. So for instance, FEVER or JAUNDICE have imports of 5 whereas SHELLFISH INGESTION HX has an import of 0. Another property which may be associated with a manifestation is a list of implications. For instance AGE GTR THAN 55 implies AGE 26 TO 55 and AGE LESS THAN 26 are negative. The "type" refers to the source and costliness of the manifestations. The possible types in order of costliness are history, sign, symptom, Lab, Lab1 and Lab2. The "prerequisite" property indicates manifestations which must be known to be positive (or in some cases negative) before the manifestation may be requested. For example, before doing a liver biopsy one should know that the patient has sufficient platelets to stop any bleeding induced by the biopsy.

3.0 THE PROCESSOR

The INTERNIST program takes the initial input manifestations and from these evokes the diseases which potentially explain them. It then ranks the diseases according to the degree to which they explain the manifestations and chooses the most highly ranked disease and its competitors as the subject for investigation. A disease is considered to be a competitor if, to a significant degree, it and the most highly ranked disease explain the same subset of the set of known manifestations.

The mode of investigation varies according to the number of competitor diseases and the level of interrogation. Four alternative modes are available:

- RULEOUT
- NARROW
- DISCRIMINATE
- PURSUING

If there are more than five diseases and the level is history, sign, or symptom then the mode is called RULEOUT. If there are more than five and the level is Lab, Lab1 or Lab2 the NARROW mode is used. If there are between two and five diseases then the mode is DISCRIMINATE and if there is only one disease, but it does not satisfy threshold criteria, the mode is PURSUING. If it does satisfy the criteria then the disease is concluded.

The type of questions asked depend upon the mode of investigation. RULEOUT mode looks for high frequency questions in order to eliminate diseases by getting negative responses. NARROW and DISCRIMINATE look for manifestations which differentiate between the diseases. And PURSUING asks about manifestations which strongly evoke the disease being pursued.

Consequently in the second phase the model proceeds by ranking the diseases, choosing a group of competitor diseases, selecting the mode of investigation, asking questions, and then recycling until criteria are satisfied. When a conclusion is reached, linked diseases are given bonuses; explained manifestations are marked "accounted for"; and the processor recycles.

4.0 DIRECTIONS FOR DEVELOPMENT

There are a number of directions in which the INTERNIST project might develop. Implicit in project leaders' goal for broad-based use of the system for medical consultation is access to it via computer networks. By November 1977 such access might be available. Inherent in this service will be evaluation of the system as it is being used by people outside the framework in which it was developed.

Large scale use of the programs through networks, and the resulting evaluations, will contribute to work on the systems interface with the user. Specifically project staff expect to make it easier for the user to input manifestation names. Another potentially significant tangential project will be to use the medical data base in conjunction with the INTERNIST programs to provide educational and testing programs for internal medicine. Of course, efforts in data base development and program refinement will be continued.

In conclusion, it is interesting to note that all of the research and development work on INTERNIST over the past two years (1974-1976) has been done, via computer networks, at the SUMEX facility in Stanford. From the point of view of project staff this ability to make viable the highly specific collaborative use of large machines is one of the valuable capabilities of computer networking.

Chapter 21

by Edward H. Shortliffe

The MYCIN Project: A Computer-based Consultation System in Clinical Theraputics

1.0 COMPUTER NETWORKING AND MYCIN

Computer networking has been an important component of the ongoing research involving the MYCIN Program, a computer-based project developed primarily at Stanford University, but also involving researchers in Arizona, Massachusetts, and at other sites across the country. MYCIN is an interactive consultation system designed to help physicians select antimicrobial therapy for patients with infections.

1.1 Characteristics of MYCIN

The program's novel approach to medical decision making models clinical reasoning by using empirical judgmental knowledge, acquired from experts, rather than formal statistical analyses of diagnostic or therapeutic probabilities. The system's design also includes features intended to enhance its usefulness and acceptability to physicians such as an ability to explain the reasoning behind its advice and automatically to correct spelling or typographical errors.

Although the program focused initially on a discrete area of infectious disease therapy, the treatment of patients with bacteremia or meningitis, its rule-based mechanisms should be generally applicable as the knowledge base expands to include other problem areas of infectious disease therapy. The validity of MYCIN's bacteremia advice is being formally evaluated in late 1976. However, the program's implementation in the clinical setting will occur only after the knowledge base has been expanded to provide useful advice for a broad range of infectious disease problems.

1.2 Development of MYCIN

As system development progressed, the number of individuals interested in the project and involved in its growth increased substantially. What began as a small project involving 4 or 5 physicians and computer scientists has become an extensive collaboration of physicians, pharmacists, biostatisticians, and computer experts. Furthermore, development of a system such as MYCIN requires constant testing and feedback from interested parties that make up a potential user community. The program's existence on the SUMEX computer, accessible via both the ARPANET and TYMNET computer networks, has enabled a large number of individuals to test the program and give developers their immediate comments regarding its strengths and limitations. Not only has this interchange via networks generated a widespread interest in the MYCIN program, but it has forced the developers to provide program features that keep it from becoming too regional or provincial in character.

The networks have also allowed true collaboration at a distance. For example, the current author, formerly actively involved in the research at Stanford, has managed to remain involved on a day-to-day basis despite his move across the country. Simple bookkeeping tasks, such as message sending and receiving, greatly facilitate this kind of collaboration. Another senior investigator on the project has moved from Stanford to Arizona but is also able to remain involved in system development by using computer networking technology.

2.0 COMPUTER NETWORKING FOR SYSTEM USE

A serious concern in the development of powerful computing systems for hospital use is the difficulty in packaging the programs on machines that are small enough to be purchased by hospitals. Computer networks may therefore provide important alternatives whereby complex computer programs such as MYCIN (which currently requires the computing power of a DEC PDP-10 when all its features are enabled) may be mounted on a large central computer and shared by several hospitals via modern communications equipment. Another advantage of such a scheme is the ability to update and maintain the program's data in the central resource on a daily basis rather than to send occasional "updates" to a number of peripheral small computers.

References

1. Davis, R., Buchanan, B.G., and Shortliffe, E.H. Production rules as a representation for a knowledge-based consultation system. To appear in *Artificial Intelligence* (1977).

2. Shortliffe, E.H. *Computer-Based Medical Consultations: MYCIN*, Elsevier/North-Holland, New York, 1976.
3. Shortliffe, E.H. and Buchanan, B.G. A model of inexact reasoning in medicine. *Math. Biosci.* 23, 351-379 (1975).
4. Shortliffe, E.H., Davis, R., Axline, S.G., Buchanan, B.G., Green, C.C., and Cohen, S.N. Computer-based consultations in clinical therapeutics: explanation and rule-acquisition capabilities of the MYCIN system. *Comput. Biomed. Res.* 8, 303-320 (1975).
5. Shortliffe, E.H., Axline, S.G., Buchanan, B.G., and Cohen, S.N. Design considerations for a program to provide consultations in clinical therapeutics. Proceedings of the 13th San Diego Biomedical Symposium, pp. 311-319; San Diego, CA; 4-6 Feb. 1974.

Chapter 22

by Stephen G. Pauker, Peter Szolovits,
Howard Silverman, William Swartout,
and G. Anthony Gorry*

A Computer Program which Captures Clinical Expertise About Digitalis Therapy and Provides Explanations of its Recommendations

1.0 NEED FOR THE DIGITALIS PROGRAM

The administration of digitalis in the treatment of congestive heart failure and cardiac arrhythmias is attended by substantial risk. Studies have shown that 20% of patients receiving the drug may develop digitalis toxicity; among these toxic patients the mortality approaches 30%. Recently, a computer program to advise doctors concerning digitalis therapy has been developed which differs from earlier ones in two important respects. First, it constructs a patient-specific model, reflecting the program's knowledge of pharmacokinetics and special features of the patient's condition which may alter his or her response to therapy. This model is used to construct the initial dosage recommendations. Second, the program makes assessments of the toxic and therapeutic effects which actually occur in the particular patient to formulate subsequent dosage recommendations. A clinical trial was performed in which the program "followed" a series of patients managed by clinicians on a cardiology service. That trial demonstrated the program's ability to recommend appropriate therapy in acutely ill patients. Each patient who developed toxicity had received more

* Stephen G. Pauker, M.D., F.A.C.P., F.A.C.C., is Assistant Professor of Medicine at Tufts-New England Medical Center, Boston, Mass.; Peter Szolovits, Ph.D., is Assistant Professor of Electrical Engineering, Laboratory for Computer Science, Massachusetts Institute of Technology; Howard Silverman, M.S. and William Swartout, B.S., are affiliated with the Laboratory for Computer Science, Massachusetts Institute of Technology; and G. Anthony Gorry, Ph.D., is affiliated with The Program for Health Management, Baylor College of Medicine, Houston, Texas.

digitalis than would have been recommended by the program. The program anticipated each episode of toxicity before it was recognized clinically.

Although the program outlined above seems to be successful in formulating dosage recommendations, to be widely accepted, any program which attempts to make clinical decisions must be able to explain its reasoning processes to the clinician. An explanation capability provides several important benefits. First, physician acceptance will be more easily obtained if the clinician can assure him/herself that the program makes reasonable deductions and recommendations. Second, an explanation feature may serve a valuable pedagogical function. Students or practitioners may improve their own abilities by learning expert techniques from the program. Finally, the program's ability to provide explanations can serve as a valuable debugging tool. Since explanations are produced in English, the clinician can locate errors of medical knowledge, although a good programmer will still be required to find other errors.

2.0 EXPLANATION FACILITY OF THE PROGRAM

To provide an explanation facility, a new version of the program cited above has been written in OWL, an English-based computer language being developed at MIT. The use of OWL allows the program to provide English explanations of its reasoning steps. The program can explain both the methods it uses and how those methods were applied during a particular session. In addition, the program can explain how it acquires information and tell the user how it deals with that information either in general or during a particular session.

Most explanations are produced directly from the code used in prescribing digitalis, and from information which is generated by the OWL interpreter as it runs. The ability of the program to translate its internal structure to an English explanation is provided by structuring the program in semantically meaningful procedures (SMP). Each SMP attempts to represent a single concept or idea that should be meaningful to the physician using the system. Just as a general idea may involve smaller, more specific ideas, so too a semantically meaningful procedure may be composed of more specific SMPs. As an example, the procedure that checks for factors which may make the patient excessively sensitive to the toxic effects of digitalis includes a call to a procedure which checks for sensitivity due to hypoxemia. Because the program is organized in this way, the explanations produced by the system tend to relate well to ideas with which the physician is already acquainted.

In many current systems that ask the user a series of questions, a problem occurs if the user wishes to change his or her answer to a previous question. The system accepts the change, but must recompute all the results computed subsequent to that question to insure that none of them are affected. Clearly, this may involve a considerable amount of unnecessary recomputation. By using OWL, one can obtain the data structures necessary to avoid this problem. The same mechanisms also allow answering hypothetical queries and investigating what the effect of changing some answer would be on the therapy recommendation.

3.0 AVAILABILITY OF THE PROGRAM

A computer program that implements the capabilities outlined above currently exists at MIT and can be accessed via national network facilities of ARPA.

Part IV **Computer Networking Physical and Chemical**

by Warren D. Seider

Overview: Networking in the Physical and Chemical Sciences

This set of papers explores the role of computer networks in the physical and chemical sciences. Authors were asked to consider materials to be distributed, standards, user services, technical staff, and financial considerations with particular concern about the modular structures of programs and data being distributed. Papers included here describe important issues relating to the use of computer networks in the chemical and physical sciences and suggest how the sharing of software and data bases might be accelerated. The potential role for EDUCOM in this process is also explored.

The following recommendations of the Physical and Chemical Sciences group at the Fall Conference are intended to help the EDUCOM Planning Council accelerate the use of computer networks in higher education. By accepting these and related recommendations the Planning Council will well serve the community of physical and chemical scientists, as well as benefit by participation of people in these disciplines.

1. Processor to Processor Protocols. There is concern that, in the next few years, when ARPANET is phased-out, that other communication networks develop protocols for exchange of files, including data and programs. Hence, we recommend that the Planning Council work with the communication networks to define protocols. In addition, the Planning Council should consider funding a project to identify the advantages of exchanging files among processors.

2. Cooperative Construction of Data Bases. There are several examples where computer networks have enabled researchers at different sites to cooperatively construct, access, and maintain data bases;

for example, in the areas of earthquake, oil production, and economic research. We recommend that the Planning Council support one or more projects, where researchers at different universities participate in building a data base using computer networks. Two areas where independent data collection activities exist today are thermodynamic properties and geotechnical data.

3. *Friendly Interfaces.* We recognize that throughout the physical and chemical sciences many programs exist that require minimal work to create "friendly interfaces" for usage on a computer network. The Civil Engineers, in particular, have been concerned with the problem of testing software and upgrading programs for distribution, and would welcome an opportunity to use computer networks as a means for more active communication among program contributors and users in the field. One possibility would be for the Planning Council to fund a small project in which the Civil Engineers and EDUCOM prepare a proposal for a joint project to be submitted to NSF. The Planning Council is advised to work with other disciplines in a similar manner.

4. *Bibliographic Information Services.* Information services such as *Chemical Abstracts* have been more widely available through magnetic tapes and communication networks. Yet, we are told that the level of utilization continues to be relatively low. We recommend that the Planning Council consider projects to catalyze usage of these services in research and teaching. A second recommendation is that EDUCOM work with distributors of these services to arrange for reduced rates for EDUCOM member institutions.

Prof. Richard R. Hughes, University of Wisconsin, was particularly helpful in developing these recommendations by sharing his insights concerning the work and goals of the EDUCOM Planning Council. His contributions deserve special recognition.

Chapter 23

by Donald M. Austin

Computer Networks in the Physical Sciences

1.0 INTRODUCTION

The ERDA (Energy Research and Development Administration) research community consists of eight multipurpose laboratories, more than a dozen single purpose laboratories and energy research centers, and over 100 universities distributed throughout the United States (see Fig. 1). In addition, a large number of private contractors undertake development and demonstration projects for ERDA. The nature of ERDA's mission requires the development of large scale modeling systems, analysis programs, data bases, and hardware facilities. These supply powerful computational techniques to a wide spectrum of research activities—from the investigation of the fundamental laws of nature to energy policy analysis. Computer science research in ERDA has created a variety of unique facilities, ranging from innovative computer hardware configurations, through sophisticated algorithm developments, to high level human-machine interface techniques. This research has brought computational and data management techniques to the ERDA community which form the basis for most of the theoretical and experimental analysis necessary for scientific research.

Since the ERDA community is widely distributed geographically, it has been difficult to provide the complete spectrum of computational and data management facilities equally to all researchers. The practical consequence has been, in some cases, duplication of effort; or, in other cases, lack of computational facilities. The expense involved in transporting specialized computer software among diverse operating systems is appreciable. It is widely acknowledged in the computer science community that the ratio of software development costs to

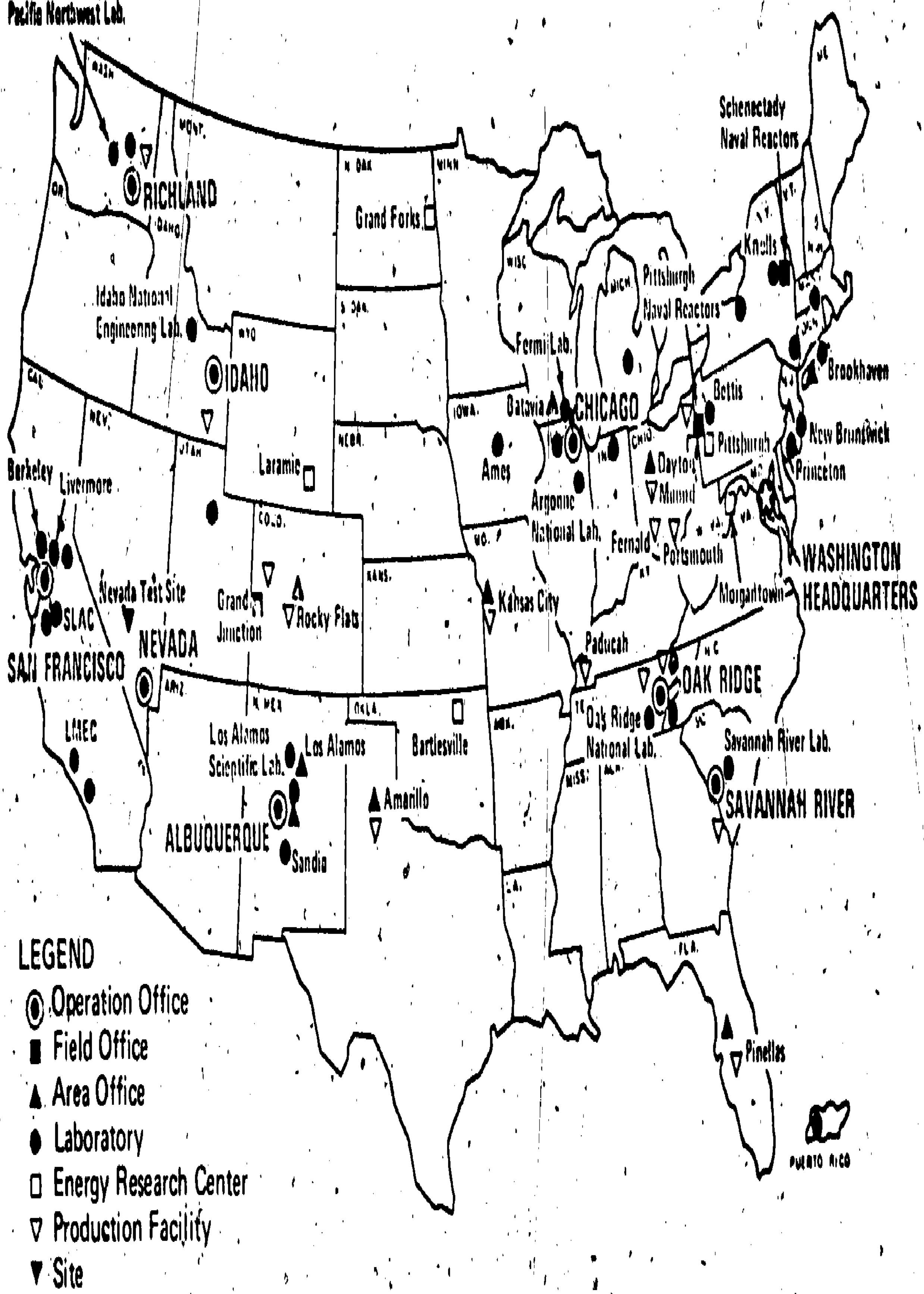


FIGURE 1. ERDA Field Organization

February 1976

hardware costs is approaching four to one. Yet even the smaller portion of computing costs, specialized hardware, is much too expensive to distribute on an equal basis to all ERDA sites. To achieve a balanced distribution of computer resources throughout the ERDA community, one solution is a distributed computer network.

There are several kinds of networks being developed today. Among these are the star network, where one host computer controls all network activity in a centralized configuration with all other nodes connected directly to that host. The most basic kind of star networks consists of a large computer center supporting remote job entry (RJE) stations and interactive terminal ports. Most modern computer centers provide this basic level of network service. The RJE's typically have card readers, line printers, and perhaps a tape drive. Interactive terminal ports can be hardwired directly to the front end of the host computer, or provide dial-up modems for remote access over telephone lines.

Another interesting configuration is the homogeneous network, consisting of two or more similar mainframes hooked together running a single operating system. This arrangement provides multiples of the usual facilities found in a single system, usually in a manner transparent to the user. This configuration can also be the center of a star network as in most commercial time sharing services.

The most interesting and versatile configuration is the heterogeneous distributed network, where different computing services are made available to a distributed community of users as in the ARPA network. This configuration is somewhat more difficult to use, since the nodes may have different operating systems, accounting algorithms, and so forth, but it does provide access to unique resources, including hardware, software, data bases, and personnel.

The ERDA Network Experimentation Project (1) was established to explore the use of distributed computer networks as a potential long term solution to the problems of reliable access to unique computational facilities throughout ERDA. The primary test vehicle for this research is the ARPA network (see Fig. 2). ERDA sites currently connected to the ARPA network are depicted in Figure 3. Investigators from each site use the ARPA network, and sometimes TELENET (a commercial "value added" packet network), to access remote time sharing systems and to transfer data files between sites. This activity has been increasing during 1976, and has produced many examples of uses of the network for conducting scientific research (2).

2.0 CAPABILITIES OF A DISTRIBUTED NETWORK

The important capabilities inherent in a distributed computer network are the sharing of unique hardware, software, data bases, and

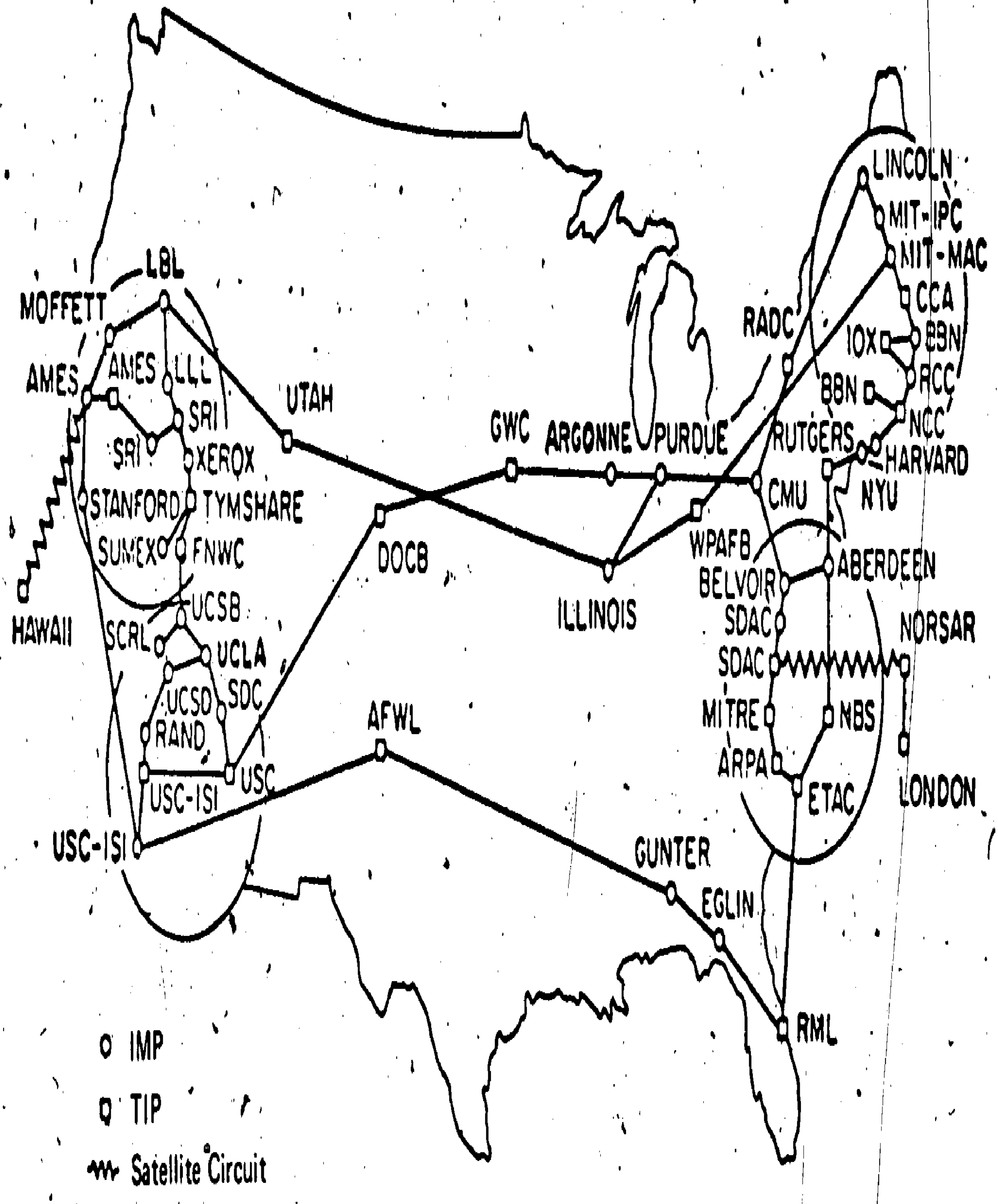


FIGURE 2. ARPA Network, Geographic Map
June 1975

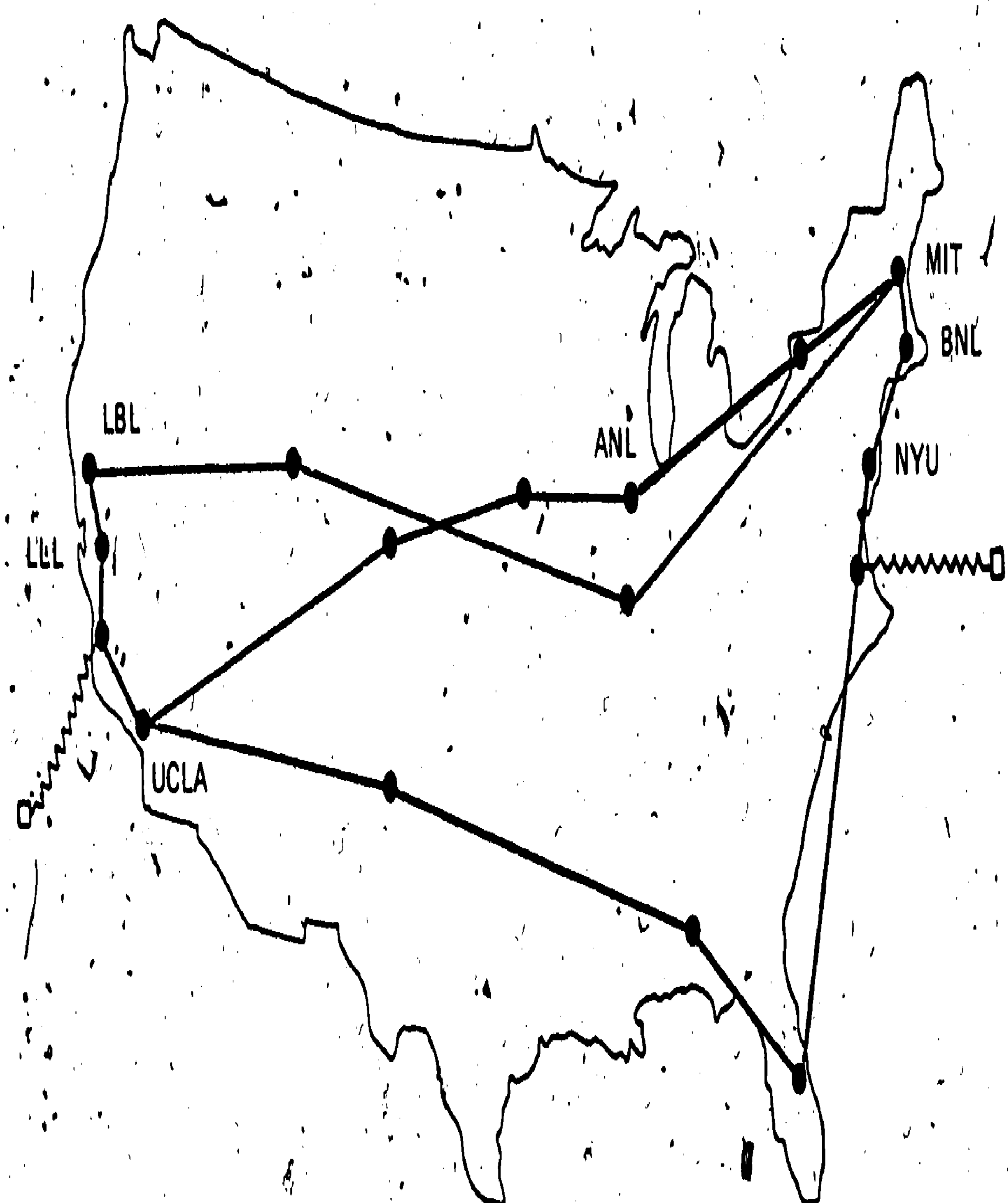


FIGURE 3: ERDA Sites on the ARPA Network

personnel. The possibilities of remote scientific collaboration using networks are increasingly exciting as the development of device independent graphics systems, teleconferencing systems, and network mail facilities continues. The richer scientific environment offered by networks should provide a new style of scientific research in the near future. Some examples of these resources are given below.

2.1 Hardware Resources

The 60 or so nodes on the ARPA network offer a wide variety of hardware capabilities which have never been collected in a single site accessible to the general scientific community (Fig. 4). These include, for example, the ILLIAC IV parallel processor machine at NASA-Ames, the CDC 7600's at LBL and BNL, IBM 370/195's at UCLA, ANL, and RHEL for large scale computation. There are terabit mass storage devices available at CCA (the DataComputer), at LBL (the IBM Photo-digital Chipstore), and at NASA-Ames (the Unicon device). Computer output to microfilm (COM) devices can be accessed at several sites (with appropriate arrangements for mailing the output back to the user), providing high quality graphics output for network users.

2.2 Software Resources

The cost of software development is often four times the cost of hardware. One might expect that the standards for FORTRAN, PL/I, ALGOL, COBOL, etc. would obviate the claim for unique software, just as Esperanto was supposed to make Mr. Berlitz obsolete. For some time to come, I foresee that network access of special software has as good a future as Berlitz. After all, the international standard ASCII character codes have been adopted by every major computer manufacturer—except, of course, IBM, CDC, UNIVAC, and so forth. It is a fact of life that really large scale software efforts tend to take advantage of machine architectures and will thus be relatively difficult to transport. If they can be used over networks, there is no need for transporting. Figure 5 represents the concept of network access of special software.

2.3 Data Bases

The strongest case for distributed processing can be made for access to unique data bases. Consider, for example, the very large data bases containing the decennial census (on the order of 20 billion characters). This represents a vast amount of data of interest to a broad spectrum of users. To transport such a large data base to thousands of computer systems is a horrible and wasteful task. Few people ever require the entire data base. Instead they need to sample freely from the data as

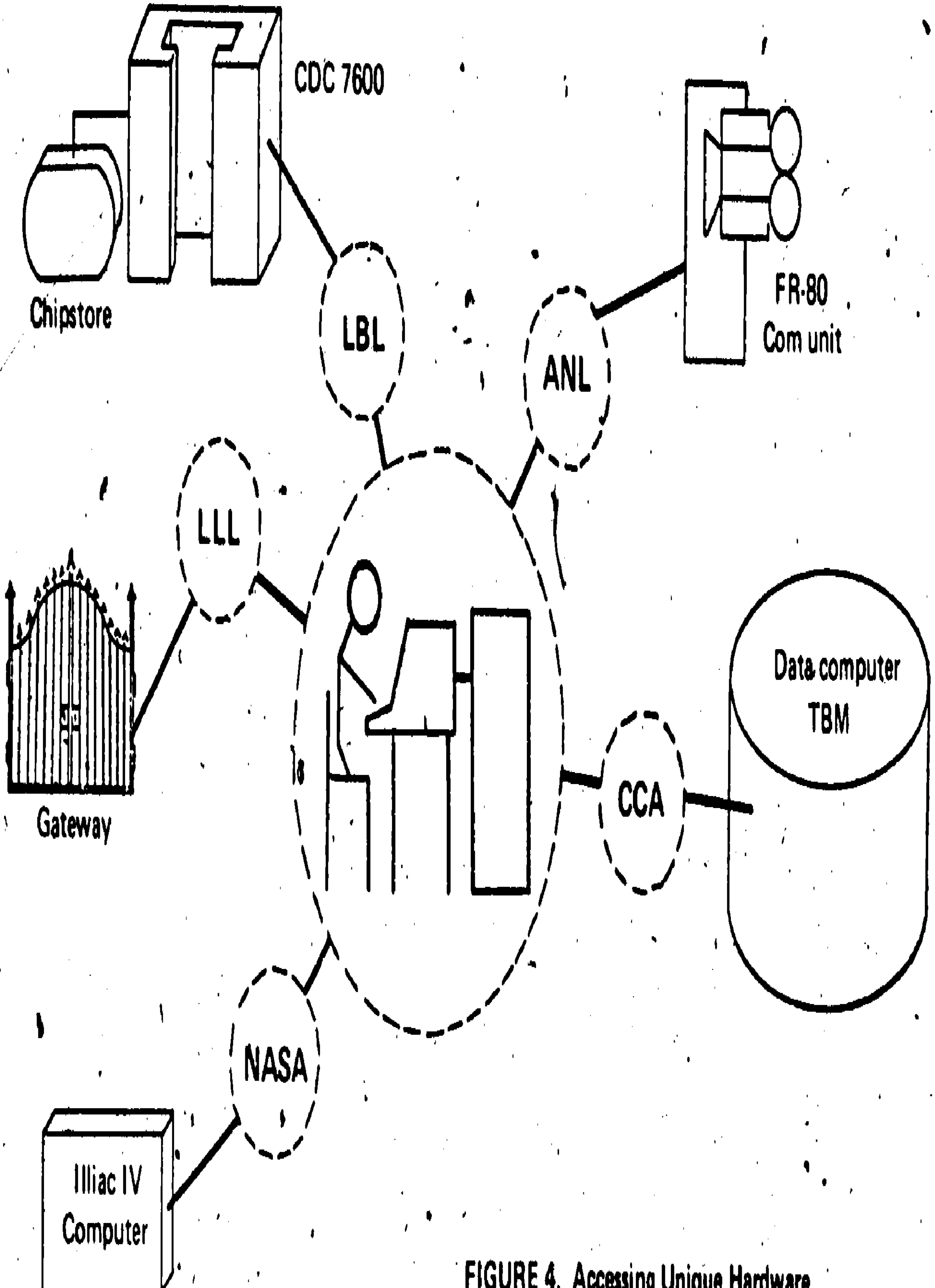


FIGURE 4. Accessing Unique Hardware

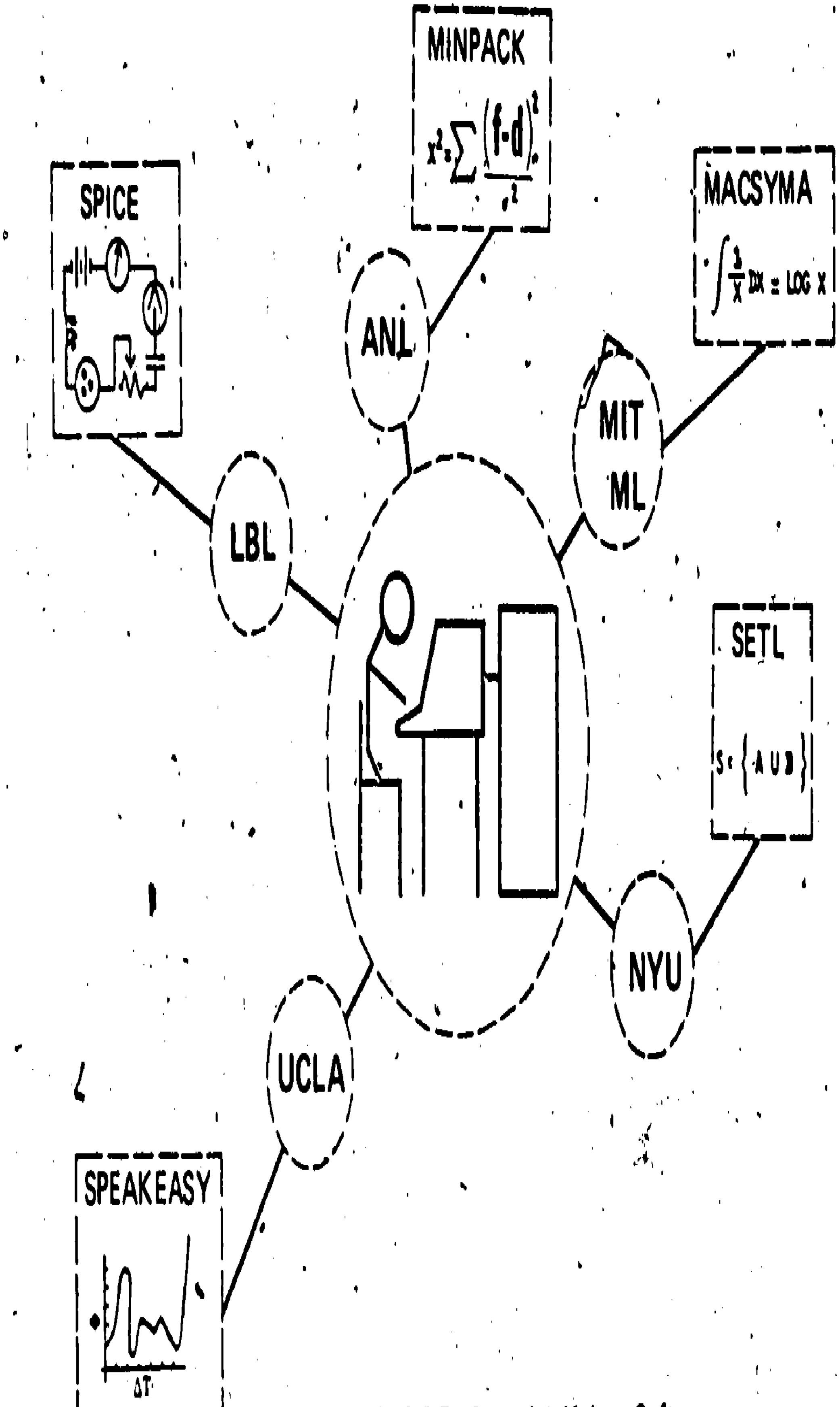


FIGURE 5. Accessing Unique Software

the need arises. Clearly, if one could store the data efficiently on one system, and provide easy and efficient access to all users, a real savings in time and effort would accrue—perhaps even enough to finance the development of human oriented query languages, high level network protocols, archival mass storage devices, and other such developments required to make distributed networks function smoothly. Although a somewhat smaller set of users exists for scientific data bases, such as the elementary particle data base, the table of isotopes, neutron cross sections, and the physical properties data base, the principle is the same.

2.4 Research Collaboration

Perhaps the most exciting consequence of computer networks is the possibility of enhancing research collaborations by scientists from several institutions. The field of high energy physics is particularly affected by this requirement, because there are only a few large particle accelerators available (such as Fermi Lab, SLAC/LBL, CERN, DESY, Rutherford, etc.). These machines must be used by all experimental physicists to do their work, and the expense and lengthy times involved in most experiments demand that the work be collaborative. Similar statements obtain for magnetic fusion energy devices, such as TOKAMAKs, TORMACs, Baseballs, and mirror machines. In order to do physics experiments, groups from several institutions join their efforts in setting up the experiment, taking data, and analyzing the data. The advent of computer networks linking the host computer systems of the home sites of these researchers' has already had an impact on the quality of research collaboration. When several sites are involved in collaborative data analysis from a single experiment, it is mandatory that good communication be established between the groups, and that the analysis techniques are compatible and accurate. Heretofore, this function has been satisfied only through travel, telephone, and the U.S. mail. By sharing data, software, and comments over networks, this process has increased the quality and timeliness of scientific collaborative research.

In other fields, the comparison of models will become practical over networks, allowing consistency checks and model expansion to become feasible. Regional energy models can be hooked together to form a national model; local models can be "aggregated" to provide input to regional models; and such matters as the appropriate units of quantities can be decided upon collaboratively. These aspects, including the sharing of model data bases, make computer networks an exciting medium for doing scientific research in the very near future.

3.0 EXAMPLES OF APPLICATIONS

The following examples were taken from experiments already conducted or in the planning stage under the ERDA Network Experimentation Project. The difficulties involved have been deliberately underrepresented for the purpose of this illustration. These difficulties fall mainly in the areas of low effective transmission rates and the lack of adequate system documentation available. Nevertheless, the examples serve to illustrate the potential of computer networks for scientific research.

3.1 Partial Wave Analysis

Collaborators at Lawrence Berkeley Laboratory (LBL) and Rutherford High Energy Laboratory (RHEL) are just beginning to analyze data on p̄p - 2 pi experiments at 16 different energies. The analysis proceeds in a series of six more or less distinct steps involving two network nodes (LBL and RHEL) and another two potential nodes (CERN and BNL) (fig. 6). The experiment proceeds as follows:

- *i.* Data is sent to RHEL from LBL over the network, CERN over an RJE link and from BNL currently, by air mail.
- *ii.* The first phase of the analysis is carried out on RHEL's IBM 360/195, producing a large set of coefficients for the partial wave amplitudes.
- *iii.* The file of coefficients is transferred to LBL over the ARPA network. At LBL, a cluster analysis program divides the solution set into several groups. This is a lengthy process requiring human judgement, so during this phase the collaborators use network mail facilities to communicate on a daily basis.
- *iv.* The results of the cluster analysis are subjected to a continuity fit program on LBL's CDC 7600 for further selection of solutions.
- *v.* These results are then transferred back to RHEL, where they are used as weak constraints for a pole extraction fitting program (to select resonance states).
- *vi.* The final solutions are input to a graphics display program at RHEL. The displays are sent to LBL over the telnet connection to a storage scope display terminal.

Clearly, step *i* is a potential network process, since BNL is almost ready to connect the central computer facility to the network, and the amount of data contributed by each site is small enough to transmit over the network.

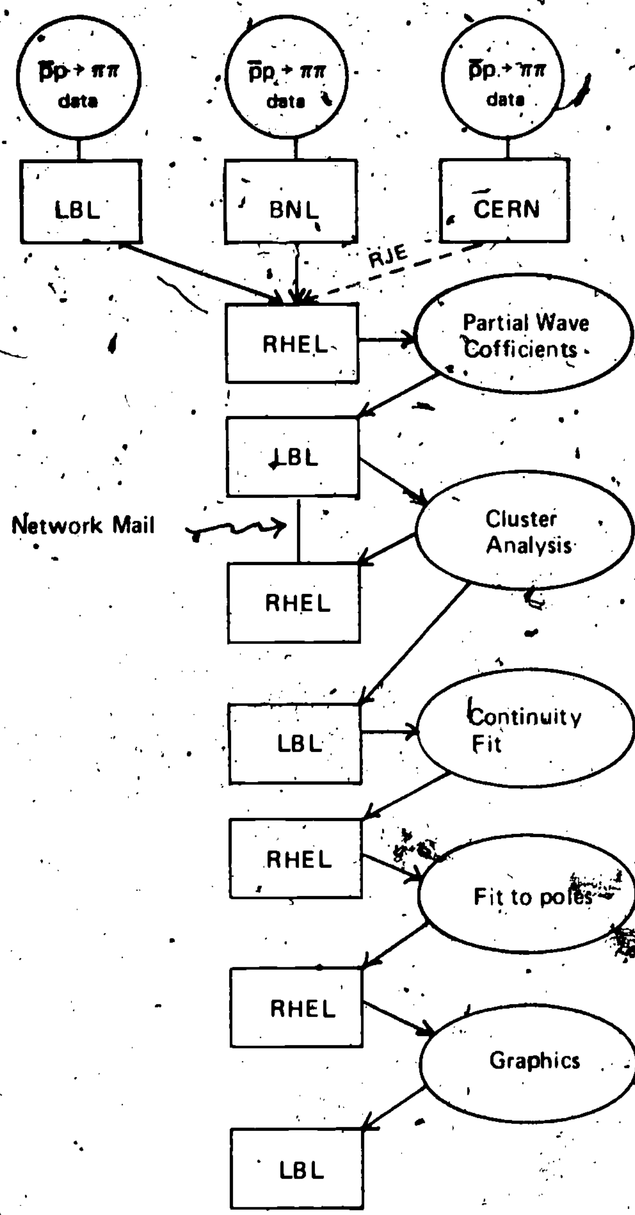


FIGURE 6. Collaborative Analysis via Networks

In several of the steps, the potential exists for using available analysis software wherever it exists on the network. For example, a cluster analysis program developed by a chemist at LLL could be used for that portion of the analysis.

3.2 Track Sensitive Target Experiment

The Track Sensitive Target (TST) is (3) a perspex box of hydrogen immersed in a neon-hydrogen bubble chamber at RHEL. The TST provides isolated protons for a pion beam to interact with, and the outgoing gamma rays are detected by the neon-hydrogen chamber. This experiment was a collaboration between LBL, RHEL and Turino.

The film from this experiment was divided three ways, with LBL, RHEL, and Turino each analyzing a third of the events (Fig. 7). It was necessary to determine any biases produced by the three separate analysis systems, so close collaboration of results was required throughout the analysis process. The long delays inherent in the postal system (a minimum of ten days) tend to produce undesirable short cuts and oversights. Using the network mail facility and the capability of transmitting graphics displays over the network, LBL and RHEL were able to maintain daily communication. In contrast, Turino had to send two physicists to RHEL for several months to complete their work. This exercise, which was one of the first research experiments carried out over the LBL link to the network, demonstrates one of the most powerful applications of the network. It does not speak, of course, to the loss to the researchers of two months in the lovely Oxford country-side.

3.3 Mathematical Software Portability

The Numerical Software Group at Argonne has been developing reliable, transportable mathematical software for several years (Fig. 8) (4). The EISPACK package for solution of eigenvalue problems is perhaps the best known of their efforts. During the development of mathematical software packages, considerable effort goes into making the software portable and accurate to a certain precision on several machines (IBM, CDC, UNIVAC). By gaining access to other computers via the network, this task can be greatly facilitated. The first effort to test this concept was a particular routine in the minimization package, MINPACK. The routine, Davidon's Optimally Conditioned Optimization Routine, was compiled and run on LBL's CDC 6600, and the output transferred back to Argonne for appraisal. It has been estimated by the director of that project that the network would have saved up to eight months in the development of EISPACK, and an even greater potential savings existed for the development of MINPACK.

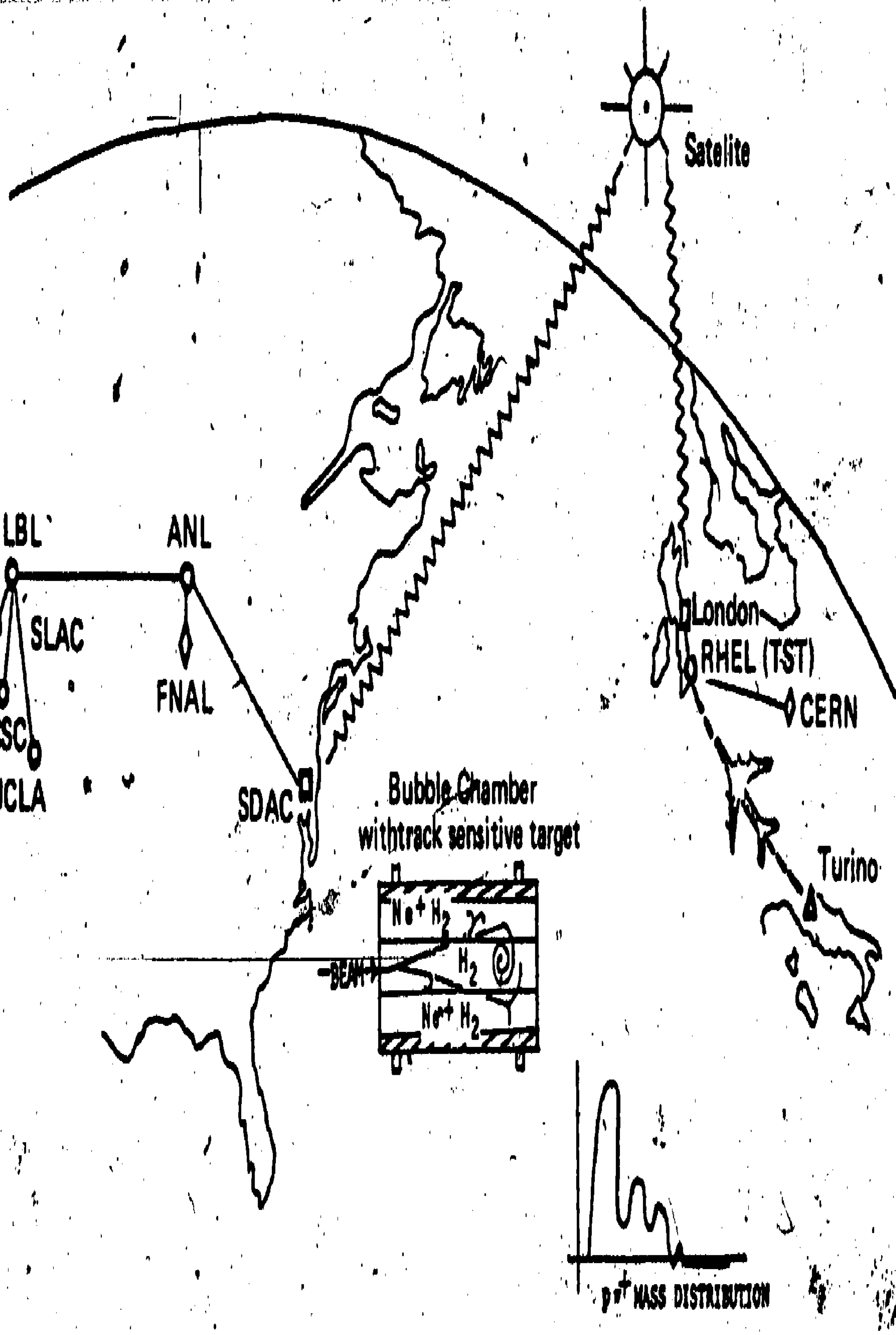


FIGURE 7. Scientific Collaboration

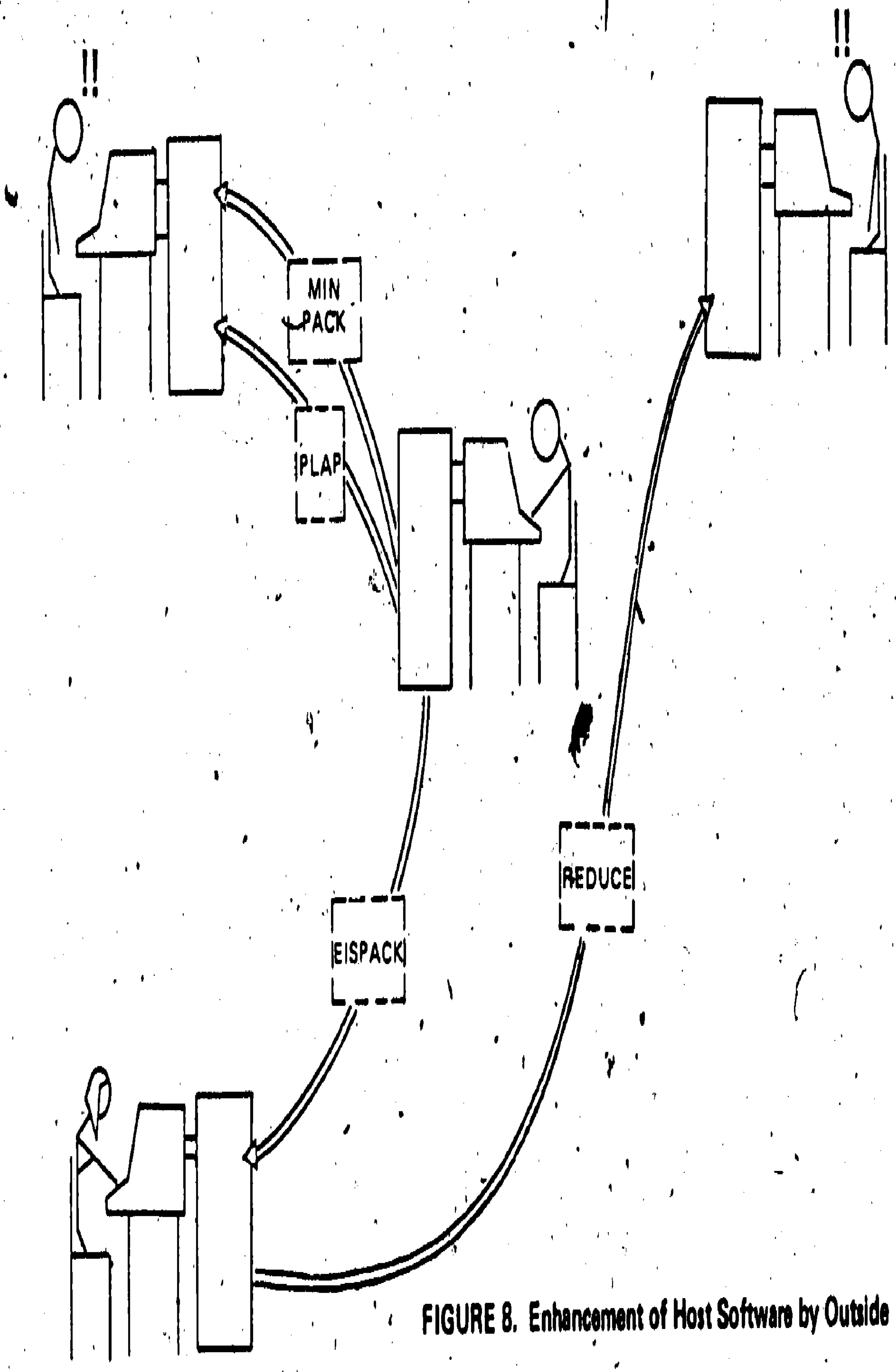


FIGURE 8. Enhancement of Host Software by Outside Users

3.4 Gateways Between Networks

The Magnetic Fusion Energy network (5) (MFEnet, formerly called the CTRNet, Fig. 9) has been established by ERDA to provide a central facility for the support of research in this area. This star network provides access to major participating laboratories and universities throughout the country. However, there remain many sites with somewhat smaller MFE programs, particularly universities, which do not have reliable access to the MFEnet.

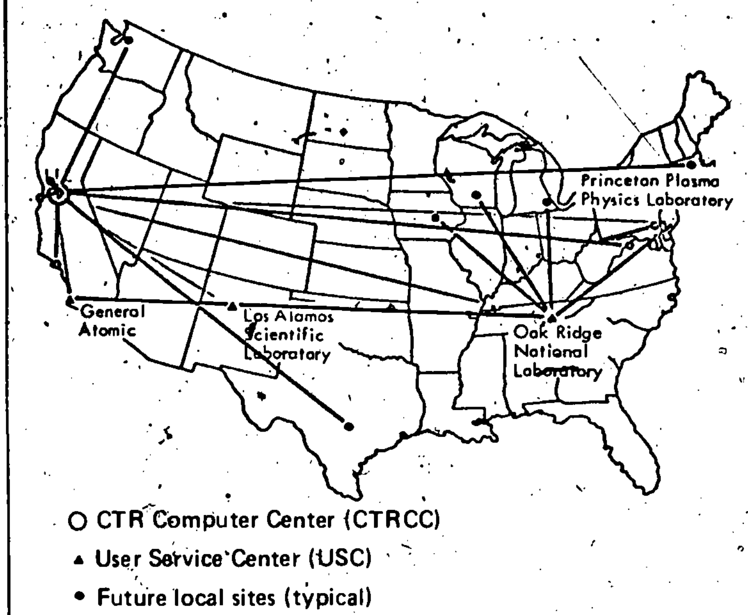


FIGURE 9. The Magnetic Fusion Energy Network.

A network gateway is a host having access to two or more networks, and which supplies the necessary protocol translation software to allow messages to flow from one network into another (Fig. 10). An example of such a gateway is the RATS (6) system developed by LLL. RATS is an ARPA network host, and through its dial-out facilities, has access to the MFEnet centered at LLL (as well as to commercial networks, such as TELENET and TYMNET).

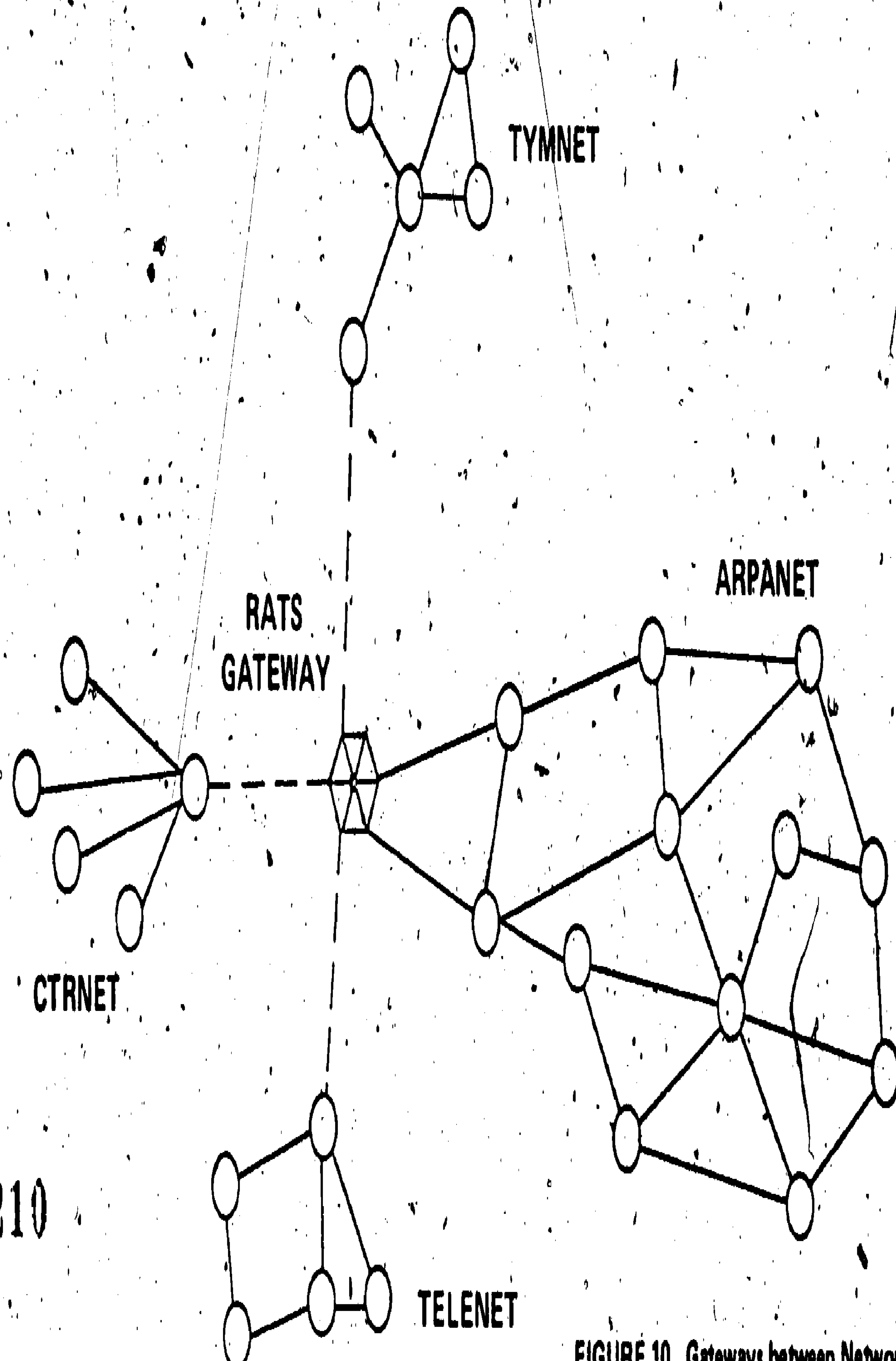


FIGURE 10. Gateways between Networks

Network investigators at NYU and ANL (4) were able to gain access to the main MFE CDC 7600 through the ARPA network by using the RATS gateway. Although this was done on a test basis only, it proves the feasibility of internetwork communication as a mechanism for resource sharing on an even broader scale. Clearly, any system which supports more than one network has the potential of linking these networks together for an even broader set of resources.

3.5 Teleconferencing

The PLANET-II teleconferencing system (7) developed by the institute of the Future is being used for daily communication by nearly a dozen collaborative groups within the ERDA community (Fig. 11). PLANET provides a bookkeeping service for messages, a survey or vote-taking service, a review service, and several other handy features which make communication among large groups easy. Since a printed transcript is always available, it far outweighs the telephone conference for many purposes. A sample of the groups using the teleconference facilities includes:

- *i.* The ERDA Interlaboratory Working Group for Data Exchange, involving researchers from eight laboratories and ERDA Headquarters;
- *ii.* Three panels on the ERDA Network Experiment Project, including people from six laboratories and universities, discussing the implementation, application, and objectives of network use;
- *iii.* An international group developing a transportable data management system (the Berkeley Data Management System);
- *iv.* An international group developing robust mathematical software.

It is significant that these groups, which formerly met at most on a quarterly basis, are now keeping in daily contact over the network. The facilities available to conferees include the ability to enter print files or data files into the conference, collaborate on group reports, broadcast results to all participants immediately, communicate asynchronously (at any time, independent of time zone differences), and to save transcripts of the conference proceedings for further distribution.

4.0 THE FUTURE

There are several serious problems to be faced in existing networks. Primarily, the transmission rates are too slow for many applications, the heterogeneous operating systems are unfathomable to the casual user,

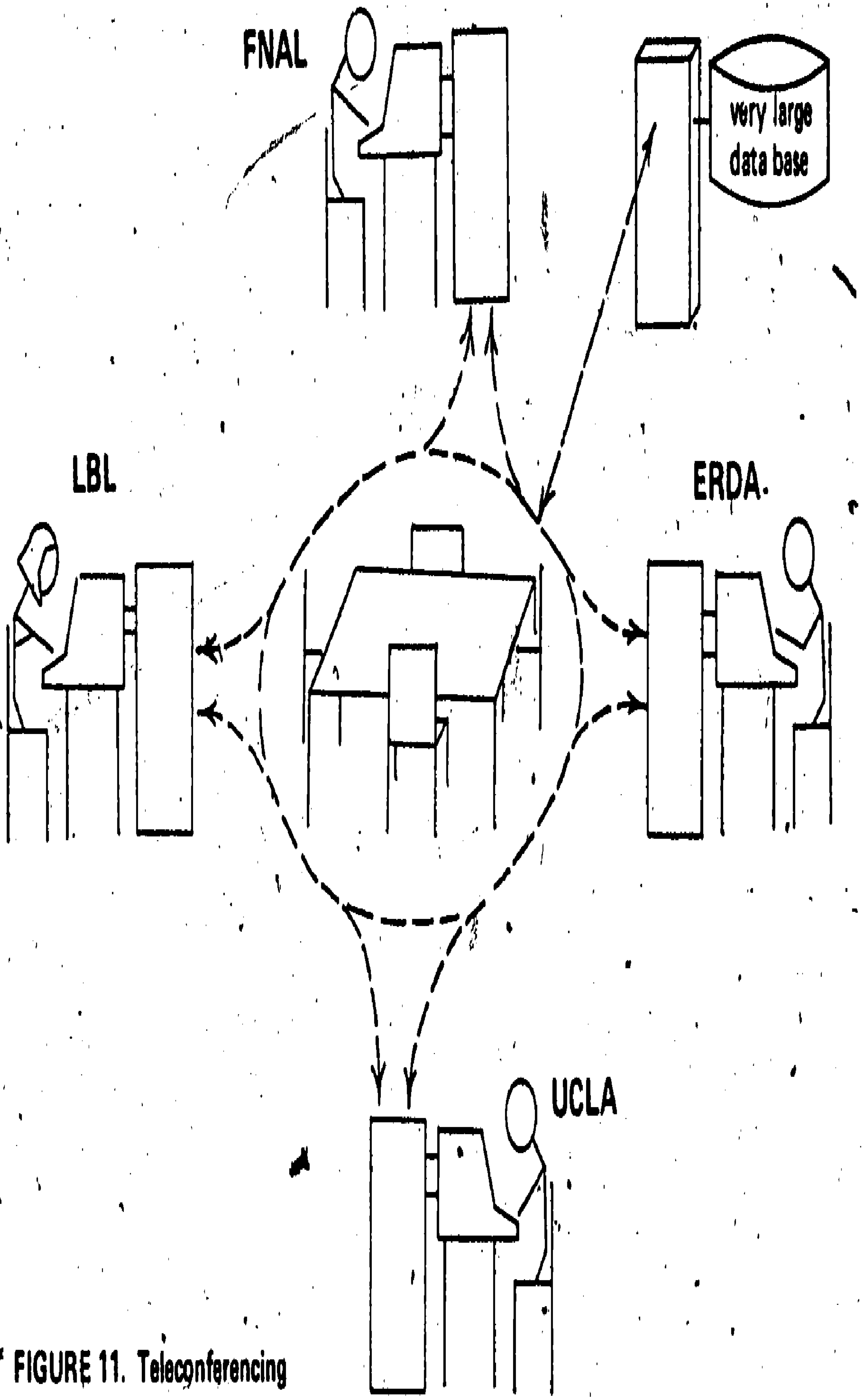


FIGURE 11. Teleconferencing

and the availability of resources is not well documented. The course of computer science research in networking must take all these impediments into account if networking is to ever reach its full potential to the scientific community.

The future of communications technology seems particularly bright today. The use of satellites, helical wave guides, and the like make megabit/second transmission a five year goal. Other possibilities include the piggybacking of data, voice, video, and facsimile signals on high bandwidth media. It seems certain that the large scale use of high bandwidth communication media will bring the price to within reasonable limits very soon.

The human-machine interface is perhaps the most troublesome problem faced by network researchers. After many years, the computer is still a variety of dumb beast which simply refuses to bow to human languages. The availability of intelligent terminals, with local storage and powerful microprocessors, can alleviate this problem somewhat, but what is still needed is the universal job control language (another Esperanto).

Machine-machine interfaces show some signs of being solved. The international CCITT X25 (8) protocol is an encouraging sign of cooperation on this problem. Probably the next step is to interface machine directly with the communication medium (when that medium becomes capable of handling speeds comparable to other peripherals).

There is much work to be done to make distributed computing transparent to the user, but I am confident that the next few years will provide great strides in this direction.

References

1. Austin, D.M. and Hall, D.E., "Computer Science Research in Resource Sharing over Networks," in Selected Examples of ERDA Research in the Mathematical and Computer Sciences, Ed. M.E. Rose, ERDA-76-118, 1976.
2. Hall, Dennis E. and Austin, Donald M., A Summary of Findings of the ERDA Network Investigators Panel, LBL-5322, June, 1976.
3. Austin, Donald M. and Hall, Dennis E., Status Report on ERDA Network Experiments, UCID 3817, February, 1976.
4. Amiot, L.W. and Bare, R.A., Status of Argonne ARPANET Implementation and Experiments, ANL Internal Report, September, 1976.
5. Killeen, John, "The CTR Computer Network," in Proceedings of the Berkeley Workshop on Distributed Data Management and Computer Networks, LBL-5315, May, 1976.
6. Donnelley, James E., "Controlling Transactions Between Distributed Computer Resources," in Proceedings of the 5th Texas Conference on Computing Systems, University of Texas, Austin, October, 1976.

7. Vallee, J., Johansen, R., Lipinski, H. and Wilson, T., "Pragmatics and Dynamics of Computer Conferencing: A Summary of Findings from the FORUM Project," in Proceedings of the Third International Conference on Computer Communication, Ed. P.K. Verma, Toronto, August, 1976.
8. CCITT-COM CII Proposal for Revised Draft Recommendation X.25, France and UK Post Office, November, 1975.

Chapter 24

by Peter Lykos

Computer Networking in Chemistry

1.0 INTRODUCTION

This EDUCOM session is a sequel to the Fall '72 EDUCOM Conference "Networks and Disciplines." The proceedings (1) of that conference contain the paper "Networking and Chemistry" replete with a stereographic picture of a molecule and a viewing device! Accordingly this paper constitutes in part an update spanning the elapsed four years. Also during this period the American Chemical Society (ACS) formed the Division of Computers in Chemistry (COMP). The 18 symposia COMP has since organized include one on Computer Networking and Chemistry. Published proceedings constitute a recent primary reference for all aspects of that subject except for commercial information services.

In order to give some perspective to those interested in computer networking in chemistry, it is important to present an overview. The point has been made that the importance of a discipline-oriented focus in computer networking has not been adequately recognized (1) and indeed *should* be emphasized as the third dimension of what has been viewed as a problem only along two dimensions, namely, those of computer network technical support and administrative structure (3).

Chemistry is an experimental science, based on the laws of physics, described using the language of mathematics augmented by the graphics of molecular structure, and dealing with bulk matter from a molecular perspective. Hence the computer in all its manifestations, plays an important role in all of chemistry. Indeed a commemorative 13¢ U.S. Postage Stamp issued in April 1976 as part of the ACS centennial celebration depicts a collection of chemical flasks against a background of punched paper tape.

2.0 ACTIVITIES IN COMPUTERS IN CHEMISTRY

Six selected computers in chemistry activities highlight important issues related to the use of networks and suggest how the sharing of software and data bases might be accelerated.

2.1 American Chemical Society Computer-Based Service and Activities

First, one must recognize the American Chemical Society (ACS). The ACS is 100 years old, has 110,000 members, and has a long history of effective and modern service to its members and to society as a whole. Only three relevant ACS activities are described here.

A. ACS Chemical Abstract Service (CAS), an operating subsidiary of the ACS, is a $25,000,000 per year operation. It examines 14,000 different primary journals published in 125 nations, patent reports from 26 countries, monographs, books, conference proceedings, government research reports and university theses, all of which generate more than 1,500,000 items per year in 50 languages, and of which about 25% contain information of interest to chemists. CAS produces about 1,500 abstracts per day in machinable form including appropriate code for retrieval. The process includes use of an algorithm for taking a computer graphics terminal image of a molecular structure and incorporating it, with redundancy checking, into the basic CAS File. A steady stream of corresponding computer readable material is distributed via magnetic tape to CAS Information Centers all over the world as well as to end users. Computer readable services delivered via magnetic tape that are listed in the 1977 brochure "Information Tools" published by Chemical Abstract Service include:

- C.A. Condensates (weekly)
- C.A. Subject Index Alert (biweekly)
- C.A. Service Source Index (Master file update quarterly)
- Chemical Biological Activities (biweekly)
- Chemical Industry Notes (weekly)
- Chemical Titles (biweekly)
- Ecology and Environment (biweekly)
- Energy (biweekly)
- Food and Agricultural Chemistry (biweekly)
- Materials (biweekly)
- Patent Concordance (two volumes per year)
- Polymer Science and Technology (biweekly)

In addition CAS has designed and is using the Chemical Registry System whereby a unique identification number is assigned to every known compound indexed to a supporting file. Over 3,000,000 substances have been registered and the number increases at the rate of 300,000 per year.

There exists a tremendous gap between this huge and important resource and the potential user community. Education of potential users as well as the cost-accounting and technical problems of access are all significant barriers (4) to the realization of that potential.

At ACS headquarters in Washington, D.C., a major demonstration project, CEDS (Continuing Education Delivery Systems), is intended to go beyond the hitherto very successful ACS efforts using audio-tapes and traveling short courses. CEDS is an attempt to use computer augmented instruction via a national commercial computer network. A comprehensive survey of the CEDS intended audience, the industrial bench chemist, revealed seven areas of interest. It is anticipated that establishment of a computer network delivery system for continuing education will lead to expansion of that delivery system to include remote data access as well (4).

The ACS Division of Computers in Chemistry (COMP) came into being in April, 1974, with several objectives including (1) to serve as a forum for computational chemists and (2) to serve as an interface between Chemistry and Computer Science and Engineering. Relative to the thrust of this EDUCOM session, COMP's national symposia have included "Program Certification and Transportability", "Computer Networking and Chemistry", and "Algorithms for Chemical Computation". In order to improve the chemistry/computer technology interface COMP is in the process of affiliating with the American Federation of Information Processing Societies (AFIPS).

2.2 The NIH-EPA Chemical Information System (CIS)

The National Institutes of Health and the Environmental Protection Agency have undertaken a cooperative project, the creation of the Mass Spectral Search System data base (MSSS). As the Federal Government cannot be in competition with the private sector and as products produced using public funds are in the public domain, NIH and EPA negotiated with commercial time-sharing companies to make MSSS available to geographically distributed scientists. The MSSS effort now serves 125 separate organizations in North America and Western Europe daily. In 1976 the NIH-EPA provides a variety of chemical information services in addition to MSSS. Fourteen data bases and application programs are now part of CIS, and seventeen different organizations in six countries have committed manpower and/or monies toward the expansion and maintenance of CIS (5).

2.3 The National Resource for Computation in Chemistry (NRCC)

Ever since September 1951, when a group of quantum chemists met at Shelter Island in Long Island, there has been a growing organized interest in and concern about the role of computers in chemistry. In May 1970, the then recently formed National Academy of Sciences-National Research Council Committee on Computers in Chemistry organized a Conference on Computational Support of Theoretical Chemistry (6). The proceedings of that conference, distributed nationwide as a National Academy of Sciences report, carried a detailed description and recommendation for creation of a National Resource for Computation in Chemistry.*

Both the Energy Research and Development Administration and the National Science Foundation became interested in the concept of an NRCC. In 1976 four ERDA laboratories are now generating proposals to become the NRCC site. Those proposals are due 3 December 1976 and it is expected that within one year Phase 1 of the NRCC will be started (7). One of the aspirants is Argonne National Laboratory which contains the Argonne Code Center, an ERDA-wide computer program exchange and information center for computer programs and systems developed in ERDA program areas. As that Code Center is likely to play an important role in one of the NRCC's principal functions, program certification and transport, a brief summary of its functions, organization and activity is given as a footnote (8).

The importance of access to NRCC by geographically distributed remote users has been stressed repeatedly at each step in the evolution of the NRCC concept. An important consideration has been the fact that no university has on its campus the largest and fastest scientific computer extant. Only the federally-sponsored laboratories enjoy that privilege.

Indeed, as was revealed at a recent National Oceanic and Atmospheric Administration (NOAA)-sponsored briefing of cognizant federal agencies by computer vendors, even the scattered federal laboratories may not be able to tap the potential of computer technology because the vendors are no longer willing to invest large sums in research and

* The proceedings of that conference also included a comprehensive article on the ARPANet emphasizing both the feasibility and relevance of computer networking, a description of a corresponding European effort already in operation (CECAM) and a list of concerns about the impact of such discipline-oriented centers on university computer centers.

development for a high-risk market. Scientific computers 100 times more powerful than the CDC 7600 may be technologically feasible by 1980, but a more coherent and organized analysis and demonstration of need will be required by OMB and the U.S. Congress before funding to realize that potential will be made available. The NRCC, and other similar discipline-oriented grassroots demonstrations of relevancy and need, can play an important role in bridging the credibility gap that now exists between Congress and super computers.

2.4 The Quantum Chemistry Program Exchange (QCPE)

QCPE was created many years ago at Indiana University and was initially operated with some federal support as a minimum-level program exchange. QCPE became self-supporting via an NSF grant in the early seventies. The testing done to programs contributed to QCPE is minimal. Under standard procedures the program is run with author provided sample input at a presumably compatible computer center, and QCPE staff compare the actual output with the corresponding author supplied sample output. QCPE has hundreds of members from all over the world. In addition to exchanging programs, QCPE serves as a communication network via its newsletter and other publications. As an indication of the trend toward complexity of programs used by chemists, a recent analysis of the 16 programs requested most often during the first 10 months of 1976 (from the QCPE program library of about 300 programs) showed that they were complex major systems. The most popular of them is Gaussian 70 (QCPE #236) developed by John Pople and collaborators, a program that consists of over 13,000 symbolic cards.

2.5 Journals That Support Computational Chemistry

Coincident with the creation of ACS's COMP, Pergamon Press announced "Computers and Chemistry, an International Journal". Its guide to contributing authors includes the statement:

> "In general, complete programs will not be published, but FORTRAN or ALGOL sub-routine listings will be welcome to afford a clear illustration of the algorithm."

In addition, the ACS publication, The Journal of Chemical Documentation, formerly the principal vehicle for publication of articles of primary concern to the ACS Division of Chemical Literature, changed its name to the Journal of Chemical Information and Computer Sciences and broadened its Editorial Board to span the full range of computational chemistry. Simultaneously the Division of Chemical Literature changed *its* name to the Division of Chemical Information in an attempt to emphasize more the computer-related aspects of chemical information handling. The ACS Journal of Chemical Education has published over 200 articles or notices dealing with specific algorithms supportive to chemistry including a reference to availability of the supporting computer program.

2.6 The National Science Foundation (NSF)

The National Science Foundation is an agent of change for science in the USA. Within the NSF, however, the transition from channeling federal support for all computer-based discipline-oriented enhancements through a single NSF office to working directly through the appropriate discipline-oriented NSF programs has moved slowly. Computer-based activities in the disciplines directed to problem-solving and decision-making has moved much more rapidly. Consequently insufficient funding for computer-based activities in chemistry is available. An NSF anachronism (9) which continues to inhibit advances in Chemistry, is the continued low-financial-level activity of NSF's Information Science Division without a concomitant thrust evolving in the NSF discipline-oriented programs.

3.0 CONCLUSIONS AND RECOMMENDATIONS

In the process of accelerating the sharing of software and data bases, EDUCOM might assume several possible roles.

- *Study the CIS network* as a model of inter-institutional cooperation involving government laboratories, industrial laboratories and higher education vis á vis the most obvious and least sensitive computer-network-sharable resource, namely, published data.
- *Expand the EDUCOM facilitating network* in order to incorporate more of the on-going computer-network related chemistry activity, particularly access to CAS material.
- *Promote a better coupling between professional societies and higher education* in the use of computer augmented continuing education delivery systems.
- *Encourage discipline-oriented curricular revision* to incorporate design and use of data bases, access to them, their linkages to and modularization of computer programming; particularly given the growing movement to chemistry department owned and operated complete general purpose computer systems (10).
- *Explore the credibility gap* regarding the need for larger and faster computer hardware and the role of the U.S. Congress, OMB and the computer vendors.

- *Examine the relevant activities of the NSF and NIE* and make corresponding recommendations.
- *Cooperative efforts with higher education elsewhere*, particularly given that international networking is already here, would be in order (10) as the USA presumably leads the world in the design, manufacture and application of computer technology.

References

1. "Networks and Disciplines", EDUCOM Fall '72 Proceedings, 1973.
2. P. Lykos, Editor, "Computer Networking in Chemistry," ACS Symposium Publication # 19, 1975, Pub. Am. Chem. Soc., Washington, D.C.
3. P. Lykos, unpublished Letter to the Editor of Science in reaction to "Computer Networks: Making the Decision to Join One", SCIENCE *186*, 414-420 (1974) by Wm. Massey. Copies available on request.
4. R. Cairns, "The second century-new directions for ACS", C & EN, p. 46, 18 Oct. '76.
5. R.S. Heller, G.W.A. Milne, R.J. Feldman and S.R. Heller, "An International Mass Spectral Search System (MSSS). V.A. Status Report", J. Chem. Info. and Comp. Sci. *16*, 176 (1976).
6. "Computational Support for Theoretical Chemistry", National Academy of Sciences, Washington, D.C., 1971.
7. W. Spindel and M.A. Paul, "Planning the Proposed National Resource for Computation in Chemistry", J. Chem. Info. and Comp. Sci. *15*, 137 (1975).
8. The functions of the Argonne Code Center include:

   (1) Collection, review, testing, packaging, maintenance, and dissemination of computer programs and program documentation in USERDA program areas.
   (2) Preparation and publication of program abstracts and bibliographies of computer programs.
   (3) Compilation and preparation of the USERDA Computer Program Summaries report.
   (4) Answering of inquiries concerning computer programs and their application; training and assisting in the implementation and effective use of library programs.
   (5) Maintaining communications and exchange agreements with other U.S. and foreign computer program information centers.
   (6) Designing and implementing storage, transmission, and retrieval systems for the Center's information resources.
   (7) Initiating, encouraging, and assisting in the development of standards and practices to facilitate technology transfer.
   (8) Coordinating proposed acquisition of non-government software by USERDA offices and contractors.
   (9) Maintaining the American Nuclear Society Mathematics and Computation Division's benchmark problem files and compiling reports describing the collection.

   During the Center's existence, over one thousand program packages have been processed for inclusion in the Center library. These packages, while primarily written in FORTRAN, included FLOCO, MAD, RPG, PL/1, and COBOL programs as well as a variety of assembly language routines. Programs prepared for

more than twenty different computer systems have at some time been reviewed for inclusion in the Center's library. The collection currently numbers over 650 packages.

Over the past 13 years methods end means for promoting effective program interchange have been investigated and developed including program documentation standards, packaging guidelines, and recommended programming practices. Much of this effort has been in cooperation with the European counterpart, the OECD Nuclear Energy Agency Library, representatives of associated "registered installations", and the American Nuclear Society.

In a single year, fiscal 1973, the Center staff responded to over a thousand requests for program package material or authorizations for use of library programs. About 300 installations including USERDA offices, ERDA and other government contractors, universities, and industrial establishments have filed Computing Facilities reports and appointed representatives to serve as their liaison in the Center's registered installation program, a program designed to foster technology transfer of ERDA-funded software development.

9. Public Law 864-85th Congress, Title IX, National Defense Education Act of 1958, establishing within the NSF a science information service. This very small part of NSF activities was created by a separate act of Congress and has survived intact as an administratively distinct unit even though the remaining structure of the NSF is almost continually being revised.
10. P. Lykos, "Computer's Role in Undergraduate Chemistry Curricula", 1976, 155 pp. Report, UNESCO, Paris.

Chapter 25

by Robert L. Schiffman
and David M. Jubenville*

Engineering Software Coordination

1.0 INTRODUCTION

In recent years, the engineering profession has successfully integrated a new tool to facilitate analysis and design. The computer is a sophisticated tool with significant potential to aid the engineer in problem solving. It has been successfully employed on many projects. Currently, however, under-utilization of computer techniques is more the rule than the exception (4, 8, 20, 21). In great part, this lack of use is due to the engineer's slow development of computer applications.

An engineering software community can be defined as an aggregation of individuals, institutions, and/or industrial establishments which have a common interest in engineering software. Several interest areas within the community are well defined. Other areas of interest are underdefined, overlapping, or newly emerging. No attempt is made here to rigorously define the various entities in the software community and their complex interrelationships. Rather, a fundamental model of the community has been constructed to demonstrate the role of the organization within this community.

A model of an engineering software community is illustrated in Figure 1. The entities defined in this model are users and potential users, program originators, engineering professional committees, professional computer standards committees, the computer and information

* Robert L. Schiffman is Professor of Civil Engineering and Associate Director for Research, Computing Center, University of Colorado, Boulder, Colorado, 80309, U.S.A. David M. Jubenville is Research Associate, Computing Center, University of Colorado, Boulder, Colorado, 80309, U.S.A.

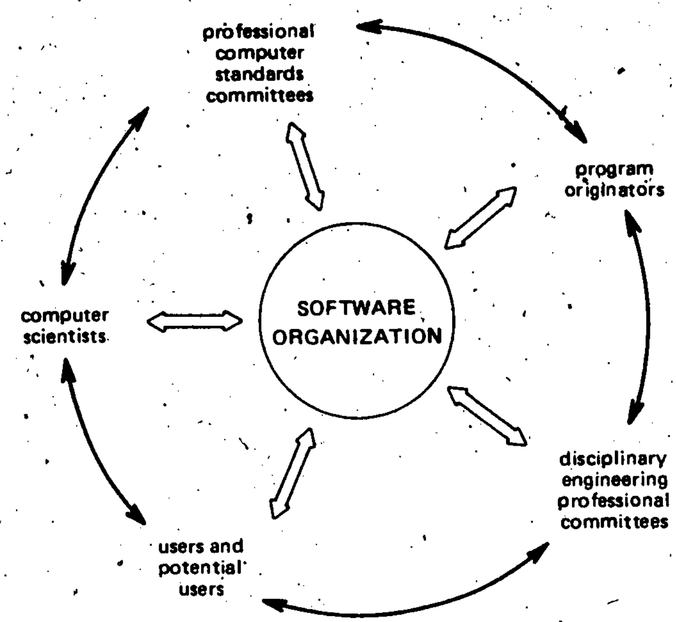


Figure 1. An engineering software community

scientists (software engineers), and a software organization. Each of these groups has different needs.

Software users and potential users are concerned with broadening the use of computer-based techniques. They are well-equipped to define potential areas of application, demonstrate the need for modifications to current computer practice, and provide valuable information derived from first hand experience.

Publication and distribution of engineering software is the prime concern of many program originators. They must be informed of software needs. They also must be kept abreast of the latest developments in the computer and information sciences as these developments affect the production of useful programs. In addition, they must be aware of the activities of the professional computer standards committees with respect to the establishment of hardware and software standards.

Engineering professional committees serve the engineering software community by formally defining needed engineering software, and by determining the elements of software which have professional utility. These committees form a bridge between the computer professional and the engineering professional. For example, the Computer Applications Committee of the Geotechnical Engineering Division of the American Society of Civil Engineers has published documentation and distribution standards for geotechnical engineering computer programs (7, 25).

Professional computer standards committees establish the minimum current level of computer programming standards for the software community. For example, the American National Standards Institute (ANSI) prepares and publishes a variety of hardware and software standards, such as a standard for the FORTRAN language (1, 2). The counterpart international standards group is the International Organization for Standardization (ISO). This group has adopted FORTRAN standards which are similar to the ANSI standards (16). The standards set by these groups are widely accepted and can be viewed as a minimum base standard for computer language.

The computer and information scientists (software engineers) contribute expertise in new computer-based techniques to the engineering software community. These techniques cover a variety of areas including numerical analysis, programming techniques, and algorithm verification. In addition, software engineering concepts and methodologies are being used to assist the process of software coordination (11, 12) by providing software tools and procedures that analyze programs for software coordination attributes.

The functional areas discussed above currently exist in the engineering software community. That is, engineers use and write computer programs, and professional engineering and computer standards committees are functioning. However, minimal communication among the various community interests has inhibited the increase of computer utilization.

There have been a series of studies in the United States to investigate the desirability and feasibility of a national effort in software coordination (3, 8, 11, 12, 13, 21, 22, 27, 29). These studies have recommended the establishment of such a national effort, and have indicated that it would be technically feasible.

It is postulated that an organized engineering software coordination effort will benefit the engineering profession by stimulating the use of computer-based techniques and facilitating the exchange and use of engineering software. It is further postulated that the coordination effort should be on a subdisciplinary basis, e.g. geotechnical engineering. An organized disciplinary coordination effort will productively channel efforts into areas of common interest and create a mechanism for the formal dissemination of software and related information.

2.0 SOFTWARE COORDINATION FUNCTIONS

An engineering software coordination effort could include a variety of software coordination functions related to specific user requirements and/or software attributes. The following functions are considered essential:

- Information exchange
- Review
- Verification
- Distribution
- Enhancement
- Maintenance
- Education
- Development

A software organization would probably undertake one or two of these functions during its early development. The mature software organization will actively pursue each of the functions and extend the scope of its investigations in each area.

2.1 Information Exchange

The most basic function of a software organization is information exchange. The software organization coordinates information exchange between software originators and users. First, an information abstract for each software package (program and documentation) is prepared by the author and maintained by the software organization. Then, the abstract serves several purposes for the potential user:

- *Awareness*. The existence of a program and its documentation is made known.
- *Applicability*. The applicability, including limitations, of the software in solving engineering problems is clearly defined.
- *Availability*. The conditions of availability and administrative restrictions on the software's use are stated.

2.2 Review

The review function requires that the engineering software organization and/or the appropriate professional committees review software and associated documentation. This review process, which fosters interaction within the software community, engineering software organizations and professional engineering committees must consider the following attributes:

- *Capability*. The algorithms utilized by the program must be appropriate for the problem which the program solves.
- *Completeness*. The program must fully reflect the problem area. Documentation must completely describe the algorithms used. Documentation must be available for each stage of program development.
- *Credibility*. The ability of a program to perform as advertised must be established. Establishing credibility includes verifying that the code correctly translates the algorithms.
- *Proof*. A mathematical argument sometimes can be developed to demonstrate that a program will perform a defined function. Within the criteria of proof, a program must execute correctly (6, 10). Program proof can be either formal or informal. Engineering software is not generally amenable to formal proofs. It is necessary, however, that software be quality assured, by empirical testing, to provide a high level of confidence in the integrity of the software.
- *Reliability*. The degree of stability demonstrated by a program over a long period of time. A reliable program will function properly for a long time (9, 31).
- *Usability*. Software must be easy to use. This implies conditions for both the program and its documentation. Input/output must be simple and straight forward. The program must be fully documented both internally and externally.
- *Intended Use*. The extent and content of software documentation is dependent on the intended use of the software. Documentation requirements may be minimal for a program written by the person who is its exclusive user. At the other extreme, documentation must be extensive for software developed for a general user community with access to a variety of computers.
- *Established Need*. Software must fill a well-defined and stated engineer's need.

Initially, the review process would be performed exclusively by engineers within the subdiscipline of the software. However, the mature engineering software organization would add software engineering dimensions to the review process. In an advanced stage of review, the software organization or its consultants would apply a set of software tools (24) to each program to analyze each of the attributes mentioned above.

2.3 Verification

Program verification (quality assurance) is the most important function of a software organization. Through this function, the user of a software package can be assured of the integrity of the program and documentation, and of its applicability to a given engineering problem.

The verification function is an advanced review procedure. Verification of a software package by an engineering software organization would indicate that the program and documentation have fully attained all the attributes defined under "Review."

The verification process must be complex and lengthy. An engineering software organization must establish stringent rules for the verification function. Certification would have to have legal standing and absolute professional credibility.

2.4 Distribution

The distribution function, unlike the previously described functions, requires that the engineering software organization maintain a program library and have access to a hardware resource. In addition, the software organization must have software and programming experts who will maintain the library programs in a form readily implementable on a variety of computer systems. There are two attributes to the distribution function:

- *Dissemination.* A mechanism must exist for the software organization to accept the program and related publications, store the software package and distribute it to users. This mechanism should be clearly defined and maintained (7, 17, 29).
- *Implementation.* A distributed program is most useful when it is readily implementable on a specific target machine and operating system. Since the software organization cannot be expected to have direct access to all hardware types and operating systems, achieving this capability implies that a means exists or can be found to make programs portable.

2.5 Enhancement

The engineering software organization ought to maintain an active computer science staff to improve library programs. Program enhancement would include modification of an incoming program to render it adaptable, portable, and consistent in style. Continuing modification in response to user needs would also be accomplished by the program enhancement staff.

- *Adaptability.* Programs and documentation should be constructed so that adaptation to suit a user's needs can be accomplished with a minimum of programming effort (11, 23).
- *Portability.* The program made available to a user must be constructed so that the code can be moved from machine to machine with a minimum of programming effort. Program portability should be thoroughly documented (5, 23), and should address five distinct aspects: the portability of the source language syntax and semantics (i.e., FORTRAN, COBOL, etc.); the portability of arithmetic operations (i.e., precision); the portability of the operating system commands necessary to invoke information flow within the system (i.e., the job control language); the portability of the run-time operations used in an operation (i.e., the system routines used by an application program); the portability of the data files which the application program reads, exercises, and updates.
- *Style.* Program and documentation style is continually subject to review and modification (18, 19, 32).
- *Modification.* Any program of significance and the associated documentation will require modification during its life. Modification generally involves improvement of program capabilities, completeness of documentation, algorithmic changes, or increasing code efficiency.

2.6 Maintenance

To fulfill the maintenance function the software organization must notify software users of changes in a program or its documentation. Initially, the organization will simply act as an intermediary between authors who maintain their programs and the program's users. At a mature stage of development, many of the software changes will occur within the software organization. Two types of maintenance can be defined.

- *External.* External maintenance consists of checking errors and/or enhancements in the capabilities, algorithms, and code which originate within the user community. Valid errors are fixed and the program changes are reported to the user community.
- *Internal.* Internal maintenance consists of implementing and reporting program and/or documentation changes which result from the enhancement process. The maintenance originates

within the software organization and is reported to the user community.

2.7 Education

There are two attributes to the education function:

- *Consultation.* A mechanism should exist which permits a user to receive adequate, decisive advice on the use and operation of a program. Ideally, the user should be able to communicate with the software organization by telephone using his own language.
- *Instruction.* The software organization should provide materials and an educational opportunity for all interested parties from novice user to expert. Special texts, printed materials, and short courses, should be utilized in this effort.

2.8 Development

A natural process in the evolution of a software coordination effort will be the development of new software. There are two attributes to this function:

- *Engineering software.* A coordinating organization will be in an excellent position to develop needed programs, or to define areas of needed development.
- *Software aids.* The feasibility of performance of the coordination functions described above will depend upon the availability and usability of a variety of software aids (14, 24, 30). The tasks of coordination, if performed by human labor alone, would be inordinately time-consuming, costly, and subject to substantial potential error. Thus an automated procedure, as free of human labor as possible, is required. Software which analyzes software can fulfill this objective. The recursive nature of software permits a software aid to automatically test the permutations and combinations inherent in tested programs with speeds and accuracy which are beyond human capabilities.

Thus a vital function of software coordination is the development of new, and the adaptation of existing, software tools which have utility in the coordination process.

It is essential that software development be controlled by individuals with disciplinary engineering expertise. Furthermore, the process of software development must be performed in the context of the coordination functions described above.

3.0 STRUCTURE OF AN ENGINEERING SOFTWARE ORGANIZATION

An engineering software coordination effort requires an extraordinary amount of interaction between the engineering subdiscipline involved and the software interests of the community. The complexities involved dictate that the scope of software being coordinated should be limited to specific engineering subdisciplines (8, 29). In addition, the software organization should be nationally or regionally based. This structure will permit easy communication and allow the software organization to operate efficiently with regard for the differences in software and engineering practice.

However, the facilities which comprise a software coordination effort could be organized in several ways. If organized on a disciplinary basis, each facility would represent a particular, professionally identifiable aspect of engineering. For example, the software organization could be specific to structural engineering, geotechnical engineering, earthquake engineering, etc., as relatively unique branches of civil engineering. If organized on the basis of hardware, each facility would maintain software implementable on the hardware of a particular manufacturer. On a geographical basis each facility would be located in a particular part of the nation or region served by the overall organization. Finally, if organized on a functional basis, each facility would perform one of the functional tasks. For example, one facility might distribute information while another might perform software maintenance.

To be successful, a software coordination effort must interact closely with the engineering profession it serves. A passive software library effort, even though it may use the most sophisticated tools of the computer sciences, will have a minor influence on the engineering professions if it lacks professional engineering identification (3, 28). On the other hand an exclusively engineering organization will not be able to provide its community of users with the software engineering expertise required to properly and efficiently coordinate engineering software.

Interdisciplinary activities within the engineering professions, and between computer scientists and engineers are desirable. They are only feasible, however, when such interactions contribute to their respective professions, and when these contributions are mutually desired and accepted. Developments in network technology could be helpful in promoting interdisciplinary aspects of engineering software coordination (15, 26). Networks could provide message switching capabilities which would enhance communication. The use of distributed data bases, and on-line access to a variety of hardware configurations, and program libraries, could relieve many of the problems associated with portability and adaptability. In addition the linking of disciplinary centers, via a network, could enhance interdisciplinary cooperation. Networking would also provide for efficiencies by removing duplications of effort in those areas which are common to two or more engineering subdisciplines.

4.0 FINANCIAL ASPECTS OF SOFTWARE COORDINATION

Engineering software coordination involves two fundamental financial questions. How much will it cost to operate a coordination organization? How will the costs be paid?

Operational costs of an engineering software coordination effort are dependent on several factors. First, the costs will be directly proportional to the number of coordination functions undertaken and to the extent to which the organization engages in each function. Other factors related to the cost of operation are the size of the user base, the number of software items being coordinated, and the technology available for software coordination. An exclusively labor intensive effort will be significantly more expensive than an effort which relies heavily on software aids.

There are two widely divergent views concerning the financing of an engineering software effort. The first view holds that since the engineering profession benefits from software coordination, the profession should bear its cost (27). The other view holds that software coordination enhances national productivity, and therefore it is a valid activity for government sponsorship. Proponents of this view argue that since most engineering software is developed in the course of government sponsored research, the government has a continuing responsibility to promote its use by supporting coordination (3).

Traditionally, software coordination has been an add-on cost to development. Thus, at this time, costs associated with software coordination are a significant percentage of the expenditure for development. In addition, coordination costs are often more visible than the cost of development, which is often buried as a small percentage of a larger project. This real, or apparent, imbalance is a transiency. As software is developed with a view to coordination, the activities, and thus the costs, required to bring the software to a state amenable to coordination will be significantly reduced. In addition, costs will be reduced as the coordination process evolves from reliance on labor to reliance on machine procedures.

Despite individual views on the ideal method of financing, evidence to date indicates that engineering software coordination cannot be totally supported by engineering users. This situation will exist until the volume of computer use increases to a point that the costs of coordination can be recovered.

5.0 THE GEOTECHNICAL ENGINEERING SOFTWARE ACTIVITY

The Geotechnical Engineering Software Activity (GESA) was established at the University of Colorado Computing Center on January 1, 1975. GESA is a voluntary, nonprofit, pilot program to test the operational feasibility of software coordination within the geotechnical engineering community. GESA was established to test the hypothesis, stated previously, that professionally based software could be successfully coordinated by a joint effort of engineering and computing professionals operating within a computing environment.

GESA develops, collects, enhances, distributes, and maintains computer software items which have professional utility. Internally developed software is maintained in a machine independent and adaptable form and distributed at the cost of reproduction. Software contributed to GESA is reviewed and enhanced. The enhanced software may be returned to the contributor or maintained and distributed by GESA.

GESA has undertaken most of the coordination functions of an engineering software organization which were discussed above. Briefly, these functions are: information provision, review, enhancement, distribution, maintenance and development. Information is provided to potential users in the form of detailed abstracts of all the software and documentation in the GESA library. Review is performed by a combination of examinations by experienced geotechnical engineers and the application of software tools developed by computer scientists. For example, software is used to analyze candidate programs for the adequacy of the accompanying test data (14). If the analysis indicates that the test problems incompletely exercise the code, the GESA geotechnical engineering staff and the author collaborate to develop complete test data. Distribution is accomplished by supplying card decks, magnetic tapes, or perforated paper tape according to the standards established by the Committee on Computer Applications, Geotechnical Engineering Division, ASCE (7). Enhancement is accomplished by creating an adaptable, portable, and stylized program. Adaptability is provided by manually separating the input, output, algorithms, and data structure of the program. In addition, a main program is used as an executive to call subroutines within the structure program. This modularization provides facile adaptability to a user's needs. All programs are checked for machine dependencies. Software tools are used to identify and convert non-ANSI standard FORTRAN constructs (24, 30). Non-portable constructs outside the ANSI subset are determined by careful hand examination. Machine dependencies are eliminated or replaced with parallel machine dependent code. Machine dependent code is clearly commented for machine or operating systems applicability. Program and documentation style are checked and modified when necessary.

Maintenance is performed by GESA in concert with the author for programs distributed by GESA. Programs distributed by their authors after GESA enhancement are maintained by the authors.

GESA program development has been primarily in the areas of consolidation and settlement. Some analytical software tools have been adapted and modified for GESA use (14, 24, 30).

An analysis of the costs associated with eighteen months of GESA operation indicates that the coordination functions can be divided into fixed and variable cost areas.

Approximately 40 percent of the total operational cost is overhead that is evenly divided between fixed and flexible costs. Fixed overhead consists of mechanisms to undertake the coordination effort like clerical and administrative procedures, that are relatively inflexible. Flexible overhead costs vary slightly with the size of the organization. Information exchange, distribution, maintenance, and education will increase with the size of the coordination effort. However, the economy of scale is small.

The remaining costs associated with the GESA coordination effort are variable and can be divided into two categories. Software review, verification, enhancement, and development comprise approximately 30 percent of the budget. Preparation of documentation comprises the final 30 percent of the budget. Software and documentation costs are almost directly related to the number of programs which are involved in the coordination effort.

Careful examination of the information above indicates that further increases in the GESA budget will significantly increase the productivity of the coordination effort. Assuming overhead costs remain relatively static, it is estimated that a 50% budget increase will approximately double productivity.

Development of more sophisticated software tools would further increase productivity. For example, the single most time consuming element in the GESA processing is the preparation of documentation.

Appropriate self documenting software for engineering software would significantly increase productivity.

The use of software tools is critical to the process of coordination. In addition to increasing productivity by reducing the time of processing, there are functions such as verification which cannot be accomplished without software aids. The use of these aids provides a substantial increase in the overall quality of geotechnical engineering software.

6.0 CONCLUSION

An objective of engineering software coordination is to raise the level of software preparation and testing in the engineering professions. The accomplishment of this objective will permit users to economically secure and implement quality programs. It will also provide authors with the tools necessary to produce quality programs. The production and use of quality software, in turn, will facilitate the solution to the myriad problems associated with verification and certification of engineering software. This will increase the use of computer-based techniques, and thus, the quality of engineering.

References

1. American National Standards Institute, (1966a), "USA Standard FORTRAN," United States of America Standards Institute, X3.9-1966.
2. American National Standards Institute, (1966b), "American National Standard Basic FORTRAN," United States of America Standards Institute, X3.10-1966.
3. Baron, M.L., Schiffman, R.L., and Fenves, S.J., (1974), "A National Software Center for Engineering," *Structural Mechanics Computer Programs* (Edited by W. Pilkey, K. Saczalski, and H. Schaeffer), University Press of Virginia, Charlottesville, Virginia, pp. 615-631.
4. Beckreck, L.N., (1973), Private communication.
5. Brown, P.J., (1974), *Macro Processors*, John Wiley & Sons, London.
6. Caplain, M., (1975), "Finding Invariant Assertions for Proving Programs," Proceedings, International Conference on Reliable Software, IEEE Cat. No. 75CH0940-7CSR, pp. 165-177.
7. Christian, J.T., Chmn, (1976), "A Standard for Computer Program Distribution," Journal of the Geotechnical Engineering Division, ASCE, *102*, No. GT10, Proceedings Paper 12472, pp. 1049-1057.
8. Civil Engineering Program Applications, (1975), "A Proposal for a National Institute for Computers in Engineering," Civil Engineering Program Applications, Inc., Report.
9. Culpepper, L.M., (1975), "A System for Reliable Engineering Software," Proceedings, International Conference on Reliable Software, IEEE Cat. No. 75CH0940-7CSR, pp. 186-192.
10. Elspas, B., Levitt, K.N., Waldiner, R.J., and Waksman, A., (1972), "An Assessment of Techniques for Proving Program Correctness," Computing Surveys, *4*, pp. 97-147.

11. Ewald, R.H., Goodspeed, C.H., Jones, R.A., Smith, R.R., Rehak, D.R., Voreades, J., and Wong, Y-C., (1975), "Computer-Based Technology Transfer," Engineering Issues - Journal of Professional Activities, ASCE, *101*, No. E12, Proceedings Paper 11228, pp. 175-187.
12. Fenves, S.J., (1973), "Methods Appropriate to an Engineering Software Center," Paper presented at ASCE Annual and National Environmental Meeting, October 29-November 1, 1973. Also APEC Journal, *7*, No. 5, pp. 1099-1107.
13. Fenves, S.J., Schiffman, R.L., and Baron, M.L., (1972), "A Position Paper on Computer Based Technology Transfer in Civil Engineering," Colloque International Sur Les Systèmes Intégrés en Génie Civil, CEPOC, Université de Liège, Belgique, 3.
14. Fosdick, L.D., (1974), "BRNANL, A FORTRAN Program to Identify Basic Blocks in FORTRAN Programs," Department of Computer Science, University of Colorado, Report No. CU-CS-040-74.
15. Herzog, B., (1973), "Organizational Issues and the Computer Network Market," MERIT Computer Network, University of Michigan, Report No. MCN-0273-TP-12.
16. International Organization for Standardization (1972), "ISO Recommendations R1539, Programming Language FORTRAN," ISO/R 1539-1972 (E).
17. Jubenville, D.M., and Schiffman, R.L., (1975), "The Role of the Geotechnical Engineering Software Activity in the Software Community," Geotechnical Engineering Software Activity, Computing Center, University of Colorado, GESA Report No. S-75-2.
18. Kernighan, B.S., and Plauger, P.J., (1974a), *The Elements of Programming Style*, McGraw-Hill Book Co., New York, New York.
19. Kernighan, B.S., and Plauger, P.J., (1974b), "Programming Style, Examples and Counterexamples," Computing Surveys, *6*, pp. 303-319.
20. Kraus, H., (1973), "Attitudes Toward Computer Software and Its Exchange in the Pressure Vessel Industry," Hartford Graduate Center, Rensselaer Polytechnic Institute, Report.
21. Medearis, K., (1973), "Report on an Investigation of the Feasibility of Establishing a National Civil Engineering Software Center, American Society of Civil Engineers, Report.
22. National Academy of Sciences, (1975), "The Proposed National Resource for Computation in Chemistry: A User-Oriented Facility," National Research Council, Assembly of Mathematical and Physical Sciences, Office of Chemistry and Chemical Technology.
23. Poole, P.C., and Waite, W.M., (1973), "Portability and Adaptability," *Advanced Course on Software Engineering* (in Bechmann, M., Goos, G., and Kunzi, H.P., *Lecture Notes in Economics and Mathematical Systems, 81)*, Springer-Verlag, Berlin, pp. 183-277.
24. Ramamoorthy, C.V., and Ho, S.F., (1975), "Testing Large Software with Automated Software Evaluation Systems," Proceedings, International Conference on Reliable Software, IEEE Cat. No. 75CH0940-7CSR, pp. 382-394.
25. Reti, G.A., Chmn, (1973), "Engineering Computer Program Documentation Standards, by the Subcommittee on Program Documentation of the Committee on Computer Applications of the Soil Mechanics and Foundations Division," Journal of the Soil Mechanics and Foundations Division, ASCE, *99*, No. SM3, Proceedings Paper 9614, pp. 249-266.

26. Roberts, L.G., and Wessler, B.D., (1970), "Computer Network Development to Achieve Resource Sharing," American Federation of Information Processing Societies Conference Proceedings, Spring Joint Computer Conference, *36*, pp. 543-549.
27. Schiffman, R.L., (1972), "Report on the Special Workshop on Engineering Software Coordination," Computing Center, University of Colorado, Report No. 72-2.
28. Schiffman, R.L., (1974), "Current Software Dissemination Practices and Organizations," *Structural Mechanics Computer Programs* (Edited by W. Pilkey, K. Saczalski, and H. Schaeffer), University Press of Virginia, Charlottesville, Virginia, pp. 591-614.
29. Schiffman, R.L., and Jubenville, D.M., (1976), "Software Coordination in Geotechnical Engineering," *Numerical Methods in Geomechanics* (Edited by C.S. Desai), American Society of Civil Engineers, *2*, pp. 1228-1240.
30. Trowbridge, S.J., (1975), "The ANSI-izer," Computing Center, University of Colorado, Report No. 75-7.
31. Tsichritzis, D., (1973), "Reliability," *Advanced Course on Software Engineering*, (in Bechmann, M., Goos, G., and Kunzi, H.P., *Lecture Notes in Economics and Mathematical Systems, 81)*, Springer-Verlag, Berlin, pp. 319-373.
32. Van Tassel, D., (1974), *Program Style, Design, Efficiency, Debugging, and Testing*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Acknowledgments

The work for this paper was supported in part by a research grant entitled "Collaborative Research—Computer Based Technology Transfer in Civil Engineering," funded by the National Science Foundation. The collaborative institutions are the Department of Civil Engineering of Carnegie-Mellon University, the University of Colorado Computing Center, and Weidlinger Associates.

Chapter 26

by Warren D. Seider

Aspects of Software Dissemination in Chemical Engineering

1.0 INTRODUCTION

This paper describes experiences in using computer networks for coursework in chemical engineering and in developing new systems for usage by the chemical engineering profession. It is intended to provide background information for a discussion to consider methods for accelerating the sharing of software and data bases.

2.0 COMPUTER NETWORKS FOR COURSEWORK

To the author's knowledge, the use of computer networks in chemical engineering courses has been coordinated solely by the CACHE Corporation.

2.1 CACHE Structure

CACHE, Computer Aids for Chemical Engineering Education, is a non-profit corporation whose Board of Trustees is comprised of twenty professors. CACHE has representatives at approximately 150 universities in the United States and Canada. Its goal is to accelerate and coordinate the introduction of digital computation into chemical engineering education.

CACHE was originally organized in 1969 as a committee within the National Academy of Engineering. In 1971, the commuter received Funding from the National Science Foundation (10). By 1975, the last year of NSF support, CACHE had reached a self-supporting basis, and the non-profit CACHE Corporation was established.

2.2 Large-Scale Systems Task Force

The Large-Scale Systems Task Force, one of several CACHE task forces, began its work in 1971. Its objective was to find a mechanism for chemical engineering departments to share large-scale programs and data bases for simulation and design of chemical processes. During 1971 the groups developed plans for on-going activities.

2.3 CACHE Guidelines for Large-Scale Computer Programs

In 1972, the task force prepared "CACHE Guidelines for Large-Scale Computer Programs" (1). These guidelines were intended to help in the evaluation and selection of potential programs. The first section of the report considers desirable features of large-scale programs and offers guidelines concerning execution cost, data input (free format), program output (annotated) and error messages. A second section considers program standards which are to follow "CACHE Standards for FORTRAN Computer Programs (2)" (prepared in 1971) with exceptions such as: 1) Programs may exceed 32K words; 2) Mixed-mode expressions are permissible; 3) Magnitudes of constants are not restricted, etc. The earlier standards were prepared for distribution of small FORTRAN programs (approximately 1-500 statements) to be run on many computers. The 1972 guidelines were to apply to large-scale programs (e.g., 30,000 statements) to be run on a single network computer.

Additional sections provided for evaluation of documentation, for program installation and maintenance, for providing consulting services, and for review of programs when determining their applicability for coursework.

As the "Guidelines for Large-Scale Computer Programs" were being completed and evaluation of potential programs began in early 1973, the task force turned its attention to a search for communication networks.

2.4 Search for Communication Networks

Four networks were considered: ARPANET, INFONET, CYBERNET and United Computing Systems (UCS). CACHE representatives were repeatedly discouraged in attempts to arrange for usage of ARPANET by students from 30-40 schools. In 1973 ARPANET was still in its early stages of development, and network administrators were more concerned with the problems of interconnecting computers for research. After carefully examining INFONET, CACHE rejected it when the task force learned that FLOWTRAN (the most promising program) could not easily be installed on UNIVAC computers. Both CYBERNET and UCS were evaluated in detail.

To aid in the selection of computer networks, the task force prepared "CACHE Guidelines for Computer Networks (3)" in 1973.

2.4 CACHE Guidelines for Computer Networks

The CACHE guidelines concentrate on the host computers and communication networks. In a first section, authors considered desirable features for host computers: 1) simple sign-on procedure; 2) reliable accounting system; 3) remote job entry for communication with typewriter terminals; 4) printing and mailing service; 5) availability of systems programmers for system maintenance; 6) well-documented literature; 7) costs comparable to university computers; and 8) a history of reliable performance.

The second section of the guidelines considers desirable features for communication networks, such as: 1) 300 bits per second transmission speed for typewriter terminals; 2) 2000 bps transmission speed for card reader-line printer terminals; 3) cost less than standard telephone rates; 4) access without typing-in a computer; and 5) a local phone call for at least 30 chemical engineering departments who expressed interest in using a program.

It should be mentioned that these guidelines were prepared to help the task force locate a computer network suitable to house a large-scale, batch program for computer-aided design, such as FLOWTRAN.

2.5 Monsanto Grant

In the Fall of 1973, Monsanto Company arranged to make FLOWTRAN available for chemical engineering courses. FLOWTRAN (FLOWsheet TRANslator) is a computer system for simulation of chemical processes in the early stages of process design. The Monsanto grant included: maintenance of the program on a single computer by Dr. Allen C. Pauls, and 7,000 dollars to aid in providing user services covering secretarial assistance, telephone costs, computing fees, preparation of text materials, and organization of a user group.

The task force welcomed this grant, as FLOWTRAN was judged to be the most satisfactory program in meeting the "CACHE Guidelines for Large-Scale Computer Programs." It had been used previously by engineers at approximately 70 chemical companies.

2.51 FLOWTRAN Simulation – An Introduction

In the Spring of 1974, Monsanto conducted a three-day workshop for the CACHE Task Force to teach the basics of FLOWTRAN. Task force members concluded that the Monsanto User Manual was poorly organized for student use and launched a project to prepare text material to introduce computer-aided design concepts in the chemical engineering curriculum. Professor J. D. Seader and W. D. Seider teamed with Dr. Pauls to write *FLOWTRAN Simulation – An Introduction* (9), which was completed in August, 1974, printed in Ann Arbor, Michigan and distributed by Ulrich's Book Store. Over 1,300 copies have been distributed in two years.

2.52 Selection of United Computing Systems (UCS)

In mid-1974, the CACHE task force decided to install FLOWTRAN on UCS. Because UCS was available to more chemical engineering departments for the cost of a local phone call, it was the logical choice. UCS offers a CDC 6600 system with back-up computers. It also provides 300 bps communications in over 100 cities for the cost of a local phone call, and 2000 baud communications across the United States using a toll-free number (800-821-7660). Charges for use are $15.00 and $10.00 per connect hour, respectively, computation costs that are roughly comparable to university computer centers. However, they are not directly comparable due to a complex charging algorithm involving CPU time, memory used and channel time. UCS agreed to waive the charges for storing FLOWTRAN ($660 per month), if CACHE spent $15,000 per year, and to charge only $3,000 for storage in the event of zero utilization. In addition, UCS has a mail service that permits users to submit card decks and receive printed output by mail.

2.53 CACHE Use of FLOWTRAN on UCS

In the Fall of 1974, to assist new FLOWTRAN users, Prof. R. R. Hughes prepared *CACHE Use of FLOWTRAN on UCS* (7). This document shows UCS procedures for executing the programs in *FLOWTRAN Simulation – An Introduction* (9) and describes procedures for entering FLOWTRAN programs using typewriter terminals and scanning files to retrieve selected portions of lengthy output files. It is an important manual that concentrates most of the information concerning UCS, required by FLOWTRAN users, in a single document. The CACHE task force on network use concluded that such a document is a requirement for any computer network to be well used by professors and students.

2.54 Workshops and User Groups

The task force decided to require professors to attend a four day workshop to learn approaches to introducing computer-aided design methods in courses, and keeping costs reasonable. By late 1976, four workshops have been held, attended by 79 professors.

A User Group (5), co-chaired by Profs. J. Peter Clark and Jude T. Sommerfeld, meets biannually. Activities of the User Group include publication of "FLOWTRAN News" three times yearly, preparation of a book of problems and solutions, an Instructors's Manual, and organization of technical sessions at AIChE (American Institute of Chemical Engineers) meetings.

2.55 Consulting Networks

The User Group has organized a consulting network to, in as much as possible, insulate Dr. Pauls from simple questions concerning FLOWTRAN. Six professors who cover six different sections of the country serve as first contacts for other professors who have questions. The organization is:

Section	Consultant	School
Northeast	W.D. Seider	Pennsylvania
Mid-Atlantic	J.P. Clark	Virginia Poly.
Southeast	J.T. Sommerfeld	Georgia Tech.
Great Lakes	R.S.H. Mah	Northwestern
Midwest	R.R. Hughes	Wisconsin
West	J.D. Seader	Utah

When these consultants cannot answer a question, they contact Dr. Pauls. Fortunately, FLOWTRAN runs reliably, UCS is extremely reliable, and the documentation prepared by CACHE answers many questions. As a result, the consultants receive very few calls.

2.56 Use of FLOWTRAN through CACHE

Usage figures are given in Figures 1-3. Figures 1 and 2 show detailed school-by-school utilization in 1974-75 and 1975-76. This information has been summarized in Figure 3. It is interesting to note that average expenditure per school ($780) and average cost per student ($39) held constant over the first two years. In 1976-77, CACHE Trustees project that 35 schools (out of a possible 125) and 700 students (out of a possible 3,000) will use FLOWTRAN. The rate of growth of use should increase as the cost of communications decreases and universities become accustomed to utilizing programs on external computers.

2.57 Experience with UCS

UCS service has been outstanding. It is available around the clock. In addition, sales personnel in many cities have been helpful in providing information for new users.

School	1974	1975 January	February	March	April	May	June	Total Dollars
A	$112.66	129.04	5.00	5.00	279.62	94.62	3.50	$ 629.72
B	64.31	664.05	222.25	166.51	55.78	0.50	0.50	1173.90
C	803.36	157.13	35.30	125.41	208.28	34.02	29.60	1393.10
D	2.07	63.83	14.00	14.00	217.91	127.50	84.97	524.28
E		11.08	31.15	5.00	5.00	5.00	5.00	62.23
F		1370.26	46.08	48.80	223.57	0.75		1689.46
G		0.39		110.37	18.31	144.09		273.16
H		0.40	10.85	15.57	45.79	533.61		606.22
I		0.40			29.15	480.78	99.29	609.62
J		0.74	113.43	506.95	155.07			776.19
K		168.83	421.70	358.53	738.51	8.64		1696.21
L		1.81						1.81
M			75.55	120.48	477.28	187.52	11.00	871.83
N	0.42			7.14	117.64	-8.00	8.00	141.20
O				39.97	30.96	59.81		130.74
P				10.62	301.08	71.90		363.60
Q				0.44		57.66	723.03	781.13
R				0.44				0.44
S				23.74	430.40	928.76	1507.21	2890.11
T				0.50	31.39	673.03	(8.56)	696.36
U					72.71	281.61	6.70	361.02
V						3.20	15.08	18.28
W						0.74		0.74
X							3.81	3.81
Y							0.84	0.84
	$982.82	2567.96	975.31	1559.47	3438.45	3701.74	2489.97	$15715.72

Summary of Expenditures 1974-1975

Total User Schools	$15715.72
Maintenance	616.58
Accounting	340.24
Workshop	2895.95
Systems Manual	1982.60
TOTAL	$21551.09

FIGURE 1. FLOWTRAN Usage July 1, 1974 – June 30, 1975

Disadvantages are that connect-charges are very high and the accounting system is not sufficiently automated. Currently, user schools specify the maximum dollars to be spent and UCS is responsible for invalidating an account when funds have been expended. However, this procedure is not automated and it is possible for a school to exceed its allocation before UCS discovers the situation. Although UCS does not hold the schools liable for overexpenditures, and the schools agree to monitor usage carefully, embarassing situations have arisen.

2.6 New Programs

In 1975, the CACHE task force turned its attention to installation of new programs. One program of particular interest that models for

School	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	Total
A	$ 3.50	3.50	3.50	332.59	1517.26	13.39	.32			14.71	252.70	79.72	$ 2221.19
B	211.93	123.22	55.67	113.49	196.00	133.68	25.10	.50	.50	.50	.50	.50	861.58
C	.83			30.33	88.03	237.47	4.10	203.45	283.06	221.25	212.80	43.30	1324.62
D		.10											.10
E	5.00	33.50	13.81	5.00	5.00	5.00	25.37	52.24	74.17	62.04	163.69	154.91	599.73
F				.26									.26
G													
H		.81							17.16	58.87	103.04		179.88
I	5.50	5.50	5.50	5.50	5.50	5.50	5.50	52.05	5.60	5.60	5.50	5.50	112.75
J		22.28	2.60	2.00	488.50	84.01	.72	.50	.86	.50	.50	.50	602.95
K				283.91	456.81	329.68	26.00	26.00	300.69	491.40	730.83	88.45	2733.77
L													
M	6.30	100.72	1.03										108.05
N	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00	15.74	9.00	9.00	9.00	106.74
O						16.02							16.02
P									32.69	275.32	154.60	16.21	478.82
Q	8.47	3.48		60.58	368.33	50.16	60.23	91.73	4.17	151.20	247.53		1045.88
R	40.48		2.12	.26									42.86
S	201.60	24.67	.50							581.16	3126.23	5.77	3939.93
T											440.34		440.34
U				62.92	124.75	195.28	53.50	6.15	3.75	17.41	515.08		978.84
V	40.87	19.51	90.96	26.17	31.52	14.50	14.50	14.50	4.69			.21	267.43
W			1.69	55.84	235.28	149.33	17.00	17.00	17.00	10.82			503.96
X	58.27	1.50	.50	.50	.50	.50	.50	.50	.50	.50		.50	64.27
Y	150.28	31.08	8.00	7.01									196.37
Z			145.06	60.03	18.56	109.67	43.34	11.55	2.41		141.67		532.29
AA				.16		781.82	56.89	336.96	291.21	578.64	164.94	6.42	2217.04
BB				.16	10.45	1.00	1.00	21.15	95.91	25.50	24.16		179.33
CC					.26	26.34	10.05	4.50	4.73	4.50	8.05		58.43
DD						19.08	48.39	139.44	31.16	71.63		.82	310.52
EE									.44	375.87	971.84	.33	1348.48
FF											226.05	15.62	241.67
	$741.03	377.85	338.94	105.71	3554.75	2180.43	400.51	986.22	1186.44	2956.42	7499.05	427.76	$21704.11

Summary of Expenditures 1975-1976

Total User Schools	21704.11
Maintenance	281.79
Accounting	544.93
Workshop	319.63
Systems Manual	426.84
TOTAL	$23277.30

FIGURE 2. FLOWTRAN Usage July 1, 1975 – June 30, 1976

	1974-75	1975-76
No. of schools	20	28
Monies spent	$15,700	$21,700
Avg. per school	$785	$775
Approx. no. of students	400	560
Avg. cost per student	$39	$39
Book sales	~900	~400

FIGURE 3. Summary of FLOWTRAN Usage 1974–76

automatic synthesis of separation flowsheets, was developed by Roger Thompson, a graduate student at the University of California, Berkeley.

In November 1976, the task force is negotiating with companies to develop a university version that will accompany a proprietary industrial version. The CACHE Trustees believe that it is advisable to work with programs that are well-maintained for usage in industry and appropriate for instruction in chemical engineering. They conclude that only well-maintained programs can be used successfully in course work and that usage by industry assures good maintenance.

3.0 NEW COMPUTING SYSTEMS

Professor Warren D. Seider and Lawrence B. Evans are coordinating a new Energy Research and Development Agency research project entitled "Computer Aided Industrial Process Design." The project is centered at the MIT Energy Laboratory and Chemical Engineering Department under the direction of Prof. Evans, with a subcontract at the University of Pennsylvania under the direction of Prof. Seider. It began in June 1976 and is scheduled for three years with funding projected at 3.3 million dollars.

The goal of the project is to develop an advanced computing system for process engineering. The system will be used for analysis of fossil energy conversion processes and energy conservation in industrial processes. The tools developed by the project will be used by process engineers in the chemical industry, at universities and working for the government.

3.1 Problems with Existing Systems

There are several problems with existing systems, three of which are considered here:

- *They are not publicly available.* Even in those cases where systems have been marketed successfully using service bureaus and computer networks, their utility is limited because portions of the source code are not provided for examination and alteration by the engineer or installation on an in-house computer.
- *The systems are large, self-contained, and somewhat cumbersome.* It is difficult to extract portions of the systems, modify them, and tailor them to specific applications. Large systems for computer-aided design are often too extensive to be executed on small computers and not sufficiently modular to permit decomposition into subsystems that can be executed independently.
- *Lack of integration.* Most of today's systems were developed independently. The net result is that programs for process simulation, equipment design, and economic evaluation often do not share common data. Often they do not use the same program modules for estimation of physical properties. This is usually the by-product of insufficient modularity and does not promote program-sharing and integration.

3.2 Challenge

The challenge in developing a next generation system is to create a framework that permits: easy insertion of program modules into large systems, and data records into data bases, on network computers; and utilization of program modules and data records on mini-computers.

Three recent papers discuss system structures with emphasis on data structure and programming language (6, 8, 10). Because space does not permit a technical discussion of these aspects of system design, interested readers should consult these papers.

The results of experiments thus far are encouraging, and indicate that it will be possible to develop a framework that meets the challenge above. With this in mind, consider some of the practical aspects of disseminating program modules (and systems) and data records (and the bases).

3.3 Center for Process Industry Computing

Probably the best mechanism for accomplishing a dissemination mission is to concentrate a staff in a "Center for Process Industry Computing." The existence of such a center is assumed in the next sections.

3.31 Standards

To facilitate easy dissemination, standards for program modules and data records will be developed utilizing existing standards as guidelines. For program modules, standards will include, for each programming language, restrictions to permit easy transfer of programs among computers. Documentation requirements will also be specified and will include descriptions of: 1) data to be supplied; 2) algorithms implemented; 3) assumptions; 4) linkages to subordinate modules; 5) results to be printed; 6) execution times; 7) memory size; and 8) interpretation of error messages.

Standards for data records will specify data formats to permit easy transfer of data among computers. Documentation will include descriptions of: data source, range of applicability (e.g., temperature and pressure ranges), and error estimates.

3.32 Dissemination Vehicles

Two dissemination vehicles are envisioned: the mails for small collections of data records and program modules, as well as data bases and problem-oriented languages; and communication networks for access to these resources on distant computers.

Because problem-oriented languages and information retrieval systems contain extensive word-processing programs for translation, they are not easily transported. These systems are often accessed using communication networks. However, component parts, such as program modules and data records should be more easily transported.

3.33 User Services

The provision of user services is especially important due to widespread geographical distribution of users. The mix of services should include person to person services, on-line automated services, and printed materials (4). Workshops and individual consulting will be particularly important to assist persons in industries where modeling technologies are not commonly used. The center's staff should work through vendors, industrial associations, consultants, and soon, as well as directly with users. Technical persons should service users and suppliers of data and programs, match system resources to user needs, provide system maintenance, and develop new systems.

3.34 Financing the Dissemination Operation

Founders need to develop an ongoing mechanism for financing the dissemination operation such as a compensation-for-services structure and guidelines for setting usage fees and providing royalties to suppliers of data and programs.

The federal government will provide funds to create a framework that permits companies to insert and acquire data and programs, and to develop standards, user services, etc., that will facilitate dissemination of the data and programs.

4.0 SUMMARY

The experiences of CACHE working with Monsanto Company to provide FLOWTRAN on the UCS network for usage by chemical engineering professors and students may serve as a model for other disciplines sharing of an instructional tool. Plans to develop a next generation system for process design, with emphasis on the problem of disseminating program modules and data bases for use at remote locations, may provide a similar framework for sharing computer resources for research and development in process engineering.

References

1. "CACHE Guidelines for Large-Scale Computer Programs," CACHE, 77 Mass. Ave., Cambridge, Mass. 02139.
2. "CACHE Standards for FORTRAN Computer Programs," CACHE, 77 Mass. Ave., Cambridge, Mass. 02139.
3. "CACHE Guidelines for Computer Networks," CACHE, 77 Mass. Ave., Cambridge, Mass. 02139.
4. "Support Services for the Remote User in a National Computing Network," EDUCOM, Princeton, New Jersey, 1976.
5. Clark, J.P. and J.T. Sommerfeld, "Use of FLOWTRAN Simulation in Education," *Chem. Eng. Educ.*, Spring, 1976.
6. Evans, L.B. and W.D. Seider, "An Advanced Computing System for Process Engineering: I. System Structures," submitted to *AIChe Journal* for publication, available from authors.
7. Hughes, R.R., *CACHE Use of FLOWTRAN on UCS*, CACHE, 77 Mass. Ave., Cambridge, Mass. 02139.
8. Joseph, B., L.B. Evans, and W.D. Seider, "An Advanced Computing System for Process Engineering: II. The Use of a Plex Data Structure," submitted to *AIChE Journal* for publication, available from authors.
9. Seader, J.D., W.D. Seider, and A.C. Pauls, *FLOWTRAN Simulation – An Introduction*, Ulrich's Bookstore, Ann Arbor, Michigan, 1974.
10. Seider, W.D., L.B. Evans, B. Joseph, E. Wong, "Routing of Calculational Modules in Process Simulation," available from authors. 1976.
11. Seider, W.D., L.B. Evans, and A.W. Westerberg, "Computers in Education: How Chemical Engineers Organized the CACHE Committee," *EDUCOM Bulletin*, 8, 2, Summer, 1973.

Acknowledgements

The contribution to CACHE of persons who have been active in the Large-Scale Systems Task Force and the Program Distribution Task Force is gratefully acknowledged.

Large-Scale System and Program Distribution Task Forces

Co-Chairmen

J.D. Seader	University of Utah
W.D. Seider	University of Pennsylvania

Present Members 1976

J.P. Clark	Virginia Polytechnic Institute
R.R. Hughes	University of Wisconsin
R.S.H. Mah	Northwestern University
J.T. Sommerfeld	Georgia Institute of Technology
A.W. Westerberg	Carnegie-Mellon University

Past Members

L.B. Evans	Massachusetts Institute of Technology
E. Phillips	Tufts University
P.T. Shannon	Dartmouth College

Part V Practical Aspects of Administering the Use of Computer Networks

by Weston Burner and John S. McGeachie

Overview: Administering Computer Networks

In Part V, papers explore the ways in which computer networks have been used to advantage by university computing centers to increase the diversity, richness, and quantity of services that are made available to the user community. Problems addressed include organization, management, control, user services, economics, privacy and security, and the buyer-seller interface. These problems must be thought through and resolved before network can be an effective means for the exchange of computer resources.

Papers have been contributed by representatives of four different types of successful networking activities that have achieved substantial success in resource sharing. To represent single server cooperation, the Harvard-MIT Experience is described; for multi-server cooperation, the ARPANET Experience; for private regional cooperation, the Dartmouth NERComP Experience; and for public regional cooperation, the state network experience of the Illinois Educational Consortium. These papers describe the ways in which various problems of cooperation were addressed and resolved and the success and potential for each partic..'ar type of network.

251a

Chapter 27

by Thomas E. Byrne

Private Regional Cooperation: The Dartmouth-NERComP Experience

1.0 DARTMOUTH'S COMMITMENT TO REGIONAL COMPUTING

For nearly twenty years Darthmouth College has maintained a high degree of commitment to the concept of regional cooperation in computing. First, as a member of the steering committee of the New England Regional Computing Center (prodecessor of NERComP) and as a user of M.I.T. computers from 1957 through the early 60's; then, as a provider of computing services to regional colleges and secondary schools under a succession of NSF grants (1967-1972); and independently, from 1971 to the present, through the Darthmouth Educational Time-Sharing Network and the NSF supported NERComP network. Several factors motivated Darthmouth to participate in these regional ventures.

In the early days, as one of the majority of "have not" institutions, our needs for computing resources were met by NERCC's offering at M.I.T.

In 1964, with the support and encouragement of the Federal Government through the National Science Foundation, easy-to-use, general-purpose time sharing, with the simplified BASIC language, was developed at Dartmouth. It is no secret that computer scientists at Dartmouth believed, *with missionary zeal,* that this approach offered the best means of integrating computing into the college curriculum! Since public funds were instrumental in the development of BASIC and time sharing, college representatives felt that they should make these services widely available. Dartmouth computer center staff were also

251 b

motivated to share resources with north country neighbors, as had others at the college throughout the 200-year history of Dartmouth College.

Additional funding by the federal government guaranteed the early success of the regional consortium. Dartmouth continued to support this cooperative effort, which had been initiated and sustained with public funds, long after those funds were exhausted because the college felt a moral obligation to do so.

As the center gained experience with regional sharing, another motivating factor emerged—the enrichment of the computation center and its products as a result of relationships with regional colleges. There were two kinds of enrichment. First, there was *product* enrichment. Programs developed by regional users were added to the Kiewit library, and documentation was developed and improved to meet the needs of remote users. Staff also improved skills in planning, managing, and maintaining an economical communications network, which led to increased accessibility and reliability of the systems.

There was also *financial* enrichment. Outside revenues enabled Kiewit Computation Center to add staff to support the remote user and to invest in software development and documentation which further improved the product. This revenue became an essential part of the funding of computing at Dartmouth and benefited all of the members of Dartmouth's community of users.

2.0 THE DARTMOUTH-NERComp RELATIONSHIP

Dartmouth's institutional commitment to NERComP has bee strong in recent years. When Dartmouth speaks of NERComP, it with the interest and concern of a "major stockholder." At present, f example, nearly 85 percent of all services distributed by NERComP a supplied by Dartmouth, and a large portion of the corporatio accumulated indebtedness is to Dartmouth. A chronological review the parallel development of the Dartmouth Regional Network and t NERComP network points out the mutual successes both have realiz and indicates why problems have arisen.

In 1966 Dartmouth was one of eight institutions in the Unit States which operated a general-purpose time-sharing system. At tl time Mount Holyoke College and several secondary schools w already using time-shared computing services supplied by Dartmo via private telephone lines from Hanover. By 1966, Dartmouth I ceased to be a user of NERComP's M.I.T. services, although m other colleges in the Boston area continued to use those services

1967, the federal government supported the establishment of Dartmouth's regional consortium.

In 1968-69, Tom Kurtz, then Director of the Kiewit Center at Dartmouth, was a member of the NERComP steering committee, which at that time was "considering incorporation." Meanwhile, Dartmouth, through the National Science Foundation financed regional network, was already providing time-sharing services to 2,000 students at 11 colleges throughout New England.

By 1970-71, NERComP served nine universities in the Boston area over three dedicated Dartmouth "ports." These NERComP members used 800 terminal hours during that fiscal year. That same year, Dartmouth's consortium served over 3,000 students whose terminals were connected to the Dartmouth Time-Sharing System (DTSS) for a total of nearly 28,500 hours.

Figure 1 shows the parallel growth of Dartmouth regional activities and NERComP network activities between 1970 and 1976. Despite the exhaustion of federal support, use of DTSS, as measured in "terminal connect hours," continued to grow. NERComP's efforts, primarily in the greater Boston area, brought new users on to the DTSS system.

As Figure 2 shows, about 22% of all terminal hours on the DTSS are used by off-campus educational users. Today about 25% of this off-campus use is by NERComP members using 14 dedicated DTSS ports. Dartmouth personnel are active on the Board of Directors and on each of the advisory committees of NERComP, and Dartmouth's technical staff has worked closely with NERComP staff in planning, designing and maintaining the data communication facilities of the network. Over the years, many institutions, and tens of thousands of college students, have received their introduction to computing through the Dartmouth and NERComP networks. This initial experience has also spawned many local time-sharing systems and several mini-consortiums in New England. Although the seeds, which were sown as terminals across the Northeast, often grew to maturity in the form of cost-effective mini-computers instead of the terminal clusters which had been envisioned, this end result achieved the goal of integrating computing into college curricula in New England on a large scale.

The Dartmouth-NERComP relationship came full-circle in 1975, when a Dartmouth chemistry professor made use of M.I.T.'s IBM 370-168 computer for research which could not be done economically on the Dartmouth system.

While Dartmouth and NERComP are pleased with these successes, mutual experiences have not been without problems. Dartmouth, as the major "stockholder" and supplier of the NERComP network, has a unique knowledge of these problems.

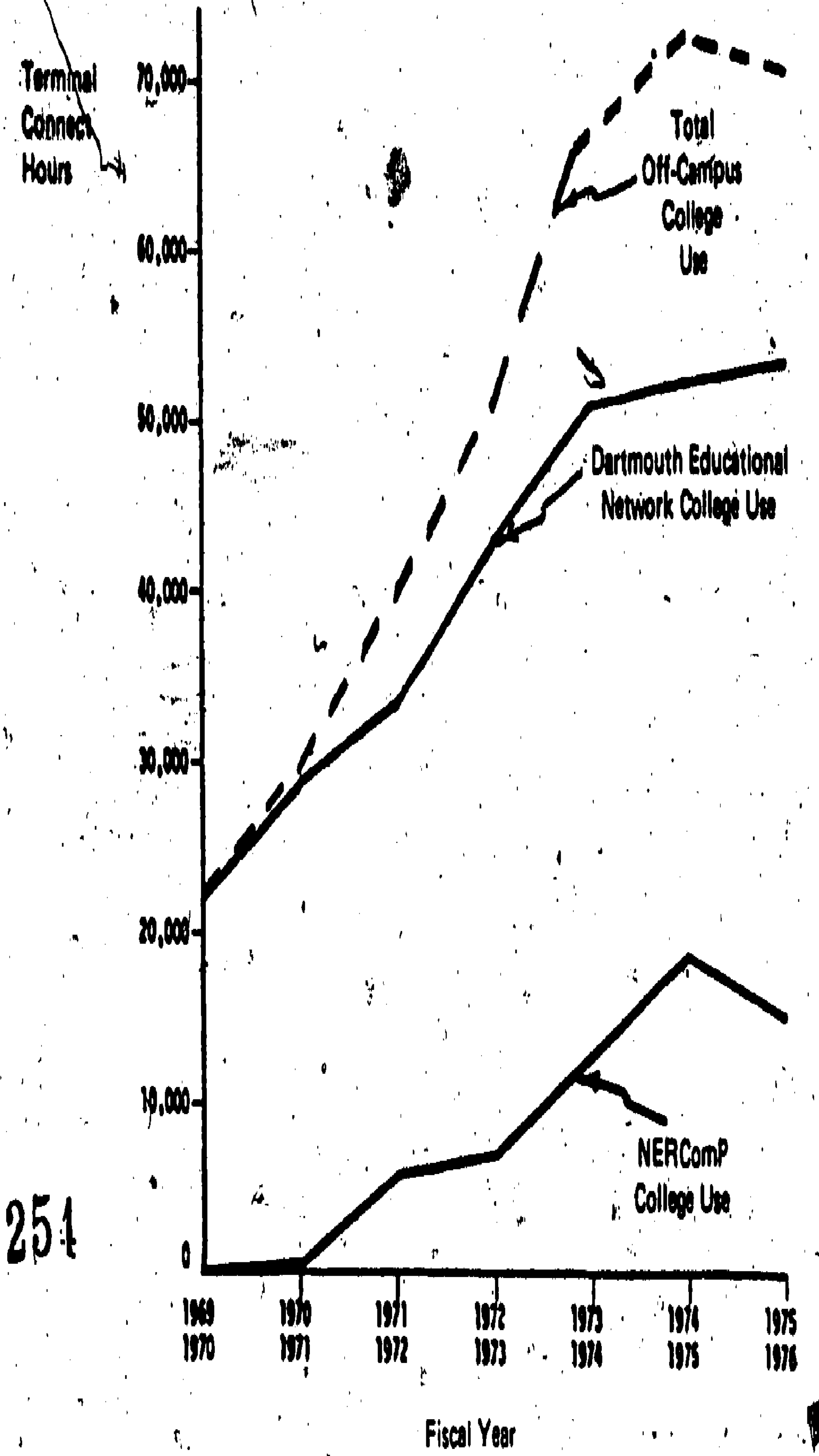


FIGURE 1. Use of DTSS Services by Off-Campus College Users — NERComP and DTSS Networks Compared 1969-1976

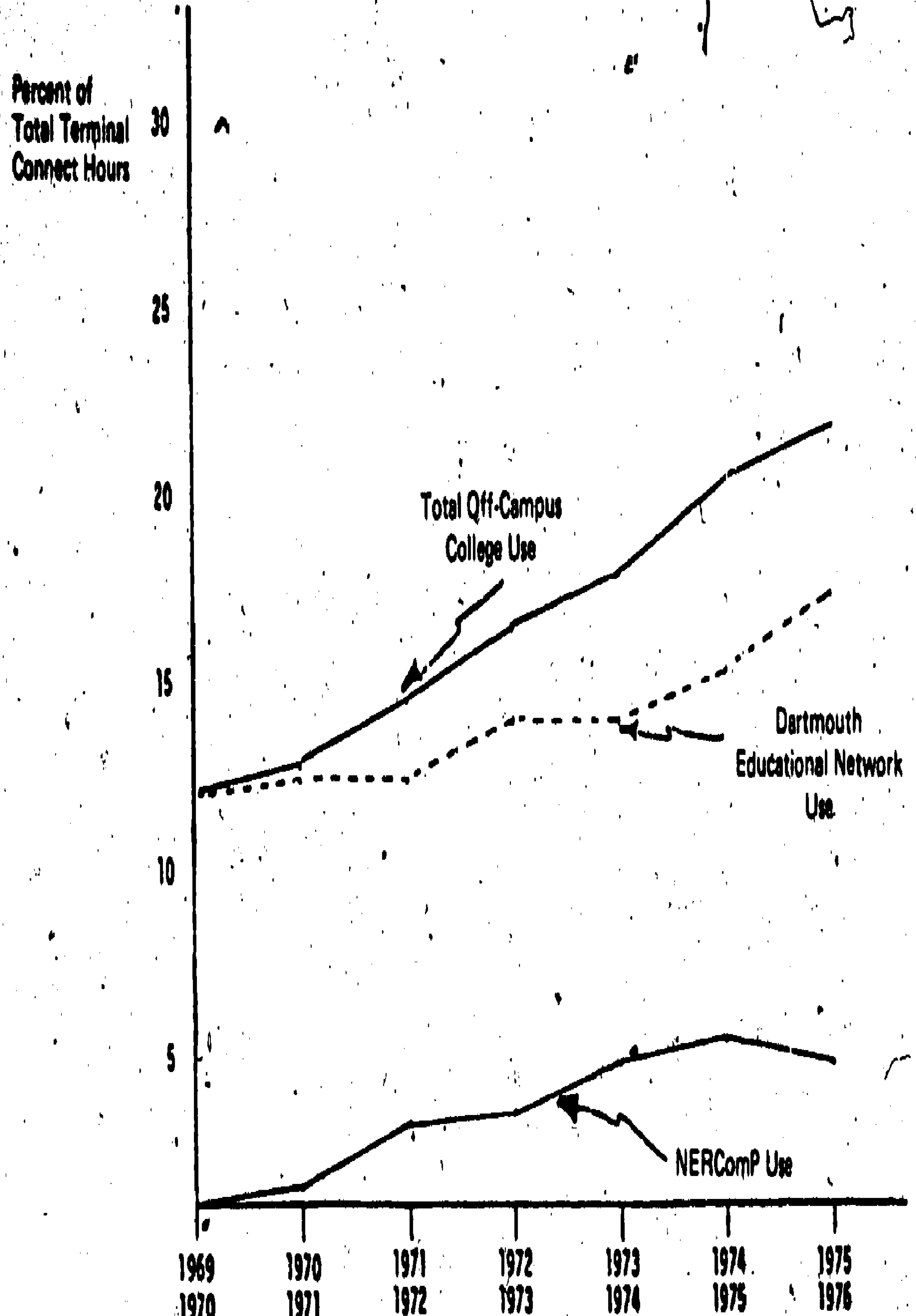


FIGURE 2. Off-Campus College Use as Percent of Total DTSS Use: NERComP and Dartmouth Use Compared 1969-1976

3.0 NERComP'S OBJECTIVES

Robert Gibbs, who has been NERComP's Executive Director since late spring 1976, succinctly identified NERComP's primary missions:

- To *market* computer services to NERComP members.
- To *connect* members to appropriate member-suppliers.
- To *support* member-users and member-suppliers, once these connections have been made.

He also mentioned educating members, consulting with members, and conducting research on their behalf.

Dartmouth staff are confident that Bob knows where NERComP should be heading, and what needs to be done to move in that direction. However, there are problems which impede NERComP's progress. Consider them within the framework of the missions of NERComP.

3.1 Financial Problems

The first obstacle to the accomplishment of these objectives is insufficient funding to staff NERComP adequately. Moreover, there is no solid financial investment on the part of all NERComP members; no equity in the organization; no sharing of risks; and thus, no common financial stake in the successes or failures.

Dartmouth and other major New England institutions have made significant investments in NERComP's operations, and are committed to the concepts of networking and resource sharing. However, this commitment and these risks are not shared equitably by all members.

To illustrate this lack of firm commitment one need only examine NERComP's membership roll. Out of nearly 50 members listed in the 1975 NERComP proposal, only 34 are included as members in November 1976, and, of these, only 23 are "active." Of these "active" members, only 13 participate as users, suppliers or distributors of the NERComP network. Among these 13, however, are some of the most prestigious institutions in New England. However, even among these major institutions the level of commitment varies considerably.

Dues paid to NERComP by the membership amounted to less than $20,000 last year. In comparison, many institutions pay as much as $10,000 per year in dues to the EDUCOM Planning Council on Computing in Education and Research, with no greater guarantee of a tangible return. Like the investment in the EDUCOM Planning Council, the investment in NERComP is an investment in *potential*. As Bob Gibbs points out, "NERComP was organized 'by and for' the membership," and the membership should fairly share the expenses and the risks.

3.2 Marketing

There are several serious problems to be overcome before the marketing objective can be accomplished. First, members must define specific responsibilities. Much thought and effort were devoted to the definition of the NERComP marketing structure. NERComP Central was to assume responsibility for identifying marketable products and for developing promotional materials and documentation. The Distributors, or Supplier-Distributors, were to market these unique products in their territories. But in fact in 1976, suppliers and distributors are not selling! There has been no acceptance of the marketing responsibility by the distributors. The lack of a formal commitment, compounded by the absence of a financial investment, fosters little concern among distributors for the success of the marketing venture.

As an example, Dartmouth, a major supplier of DTSS services, and a probable distributor, is not formally committed to marketing the services of other NERComP suppliers either on the Dartmouth campus, or in its self-defined sales "territory." Nor is Dartmouth equipped to do so! Kiewit staff do not possess detailed information about others' products, nor have they established policies and procedures pertaining to the purchase of off-campus computer power by Dartmouth's on-campus users.

Unless a formal proposal is submitted to Kiewit, staff are unlikely to address these questions. Furthermore, unless Dartmouth knows that all other member-suppliers are willing to accept similar proposals, the college is unlikely to accept such a proposal.

This lack of defined responsibility, along with the lack of acceptance of responsibility, places an overwhelming burden on NERComP Central personnel who must sell door-to-door simply to survive, and thus cannot attend to their primary tasks of product identification and promotion.

A second problem has been competition between NERComP and its members. NERComP suppliers have not entered into a formal agreement to grant NERComP the exclusive right to market computing services in New England. As a result there has been overlapping effort, and sometimes direct competition, between NERComP Central and its suppliers. With total marketing resources as scarce as they are, this situation cannot be allowed to continue.

Finally, NERComP services an unstable clientele. The principal product being marketed by NERComP is DTSS general-purpose time-sharing services. Experience shows that the mini-computer is a deadly competitor to this kind of service. It is unusual for DTSS Inc. to keep a customer for more than two or three years before the customer is convinced that there are economical advantages to acquiring an in-house time-sharing system.

Once the financial ills of NERComP are cured, the next highest priority item should be an all-out effort on the part of the organization to identify the network's unique products and promote them widely. NERComP must develop a stable clientele for unique applications to insure predictable and significant income to offset the high overhead associated with the fickle, experimental, short-term user.

3.3 The NERComP Connection

Despite problems that beset the technical staff of NERComP, the present network is more cost-effective than Ma Bell's offerings, and Telenet cannot compete with it in the New England area. NERComP is studying its emerging role in data communications where it acts as a coordinator, facilitator, or connector as often as it does as an operator. A reasonable expectation is that at some time in the near future, the data communications expertise and services required by New England educational institutions can all be provided by NERComP.

3.4 Support

The problems of remote user services have received much attention, perhaps too much, at the current stage of NERComP's development. The NERComP model suggests that user services will be distributed, that is the distributor will provide the interface to the user. Dartmouth's approach to remote user services is to provide an easy-to-use product which is virtually self-supporting and is backed up by complete documentation. Such a product can be easily supported by distributors.

NERComP's primary efforts should be directed toward establishing standards for product quality and evaluating each product offered by suppliers to insure that they will be easy to use and to support. A substantial investment of resources and effort at the beginning of preparing these products for the marketplace will significantly reduce the effort and expense required, on the part of the distributor, to support them. Since the network is communications-oriented, and, since the technology is there, there should be no reason why existing facilities cannot be utilized to provide user services directly from the supplier. This capability would also reduce the burden on distributors and would increase their willingness to market the product.

NERComP's efforts should also be directed toward coordinating a thorough review of the major suppliers' plans for applications and software development. The potential rewards for eliminating duplication of effort are great, and significant economies could be realized.

4.0 NERComP'S VALUE ADDED

In its new role as connector, facilitator, or even matchmaker, NERComP must give more thought to compensation for value added to network products. Originally several schemes were considered, all of which were some variation of slicing the (revenue) pie. None of the models considered produced a slice large enough to satisfy all recipients. This is still an area of discussion and disagreement.

NERComP will perform many important functions for which there is no immediate return for effort expended: for example, searching for new applications to meet known needs; developing promotional materials; coordinating seminars; and recruiting new members. The cost of these activities should be financed using income from membership dues, and such financing should be viewed as straightforward and necessary subsidization.

When NERComP provides communications services to members, the full cost of such services should be recovered from the users of those services. When NERComP provides computer services, compensation should be received at rates equal to those published by the suppliers of the services. Bulk discounts on the sale of these services to NERComP should leave a margin to cover the costs of providing the services, and NERComP should receive some percentage of the revenue generated to cover its marketing costs. When NERComP provides consulting services, consulting fees should be paid. When NERComP installs communications equipment, it should be paid an installation fee. All network participants should receive fair compensation for all services rendered.

However, network bureaucracy should not restrict members from entering into other mutually satisfactory agreements to accomplish resource sharing. NERComP should provide an atmosphere conductive to many kinds of agreements, all of which fall under the NERComP "umbrella."

To protect NERComP's interests, suppliers should agree not to enter into any agreement with another college or university in New England, unless that institution is already (or is willing to become) a member of NERComP. Furthermore, perhaps NERComP should be entitled to the same percentage of revenue generated from such agreements as from agreements where it is the principal marketing agent. The point is that, when NERComP Central's resources or facilitaties are not used, NERComP's compensation should be minimal.

5.0 SUMMARY

In this paper, the author suggests that the business plan which guides NERComP's future operations should include:

- A substantial financial investment and equitable risk-sharing by all NERComP members.
- The formal acceptance of marketing responsibilities by suppliers and distributors.
- The adoption of an agreement by all suppliers which gives NERComP the exclusive right to distribute computing services throughout New England under the NERComP umbrella.
- A formal redefinition of NERComP's role in the network, and a re-ordering of priorities. Emphasis should be placed on product development and refinement, promotion, communications and support in that order.

As a major supplier of services through NERComP, Dartmouth holds these expectations for the future:

- That NERComP will continue to locate new users for DTSS services augmenting Dartmouth's missionary effort.
- That NERComP will continue to offer the opportunity for small institutions without on-campus computing systems to meet their diverse computing needs.
- That NERComP will emphasize the identification and marketing of unique, quality computing services.
- That NERComP will assume an active role in the coordination of systems and applications development in New England to avoid duplication of effort.
- That NERComP will gradually assume responsibility for planning, providing, managing and maintaining data communications facilities for New England's educational institutions, and for interconnection to national networks.
- That NERComP will continue to plan and coordinate discipline-oriented seminars and conferences.

Dartmouth is encouraged by the initiatives taken by NERComP's new management in the second half of 1976. Kiewit is no less committed today than in the past to the concepts embodied in the NERComP organization. Both feel confident that, with the assistance of a financially sound NERComP organization, they can continue to have a substantial impact on the integration of computing into college curricula, and, at the same time, economically meet all of the computing needs which arise on the Dartmouth campus.

Chapter 28

by Robert Gibbs

Networking from a Regional Perspective

1.0 INTRODUCTION

NERComP is an action-oriented agency for promoting the sharing of academic resources, especially computer resources for educational purposes in Higher Education. The agency serves higher education in New England and operates for the benefit of its membership. NERComP has approximately 40 members, the majority of which are active in some capacity as suppliers, users, or advisors. Seven colleges are suppliers and two act as distributors.

2.0 NETWORK GOVERNACE

The direction of NERComP is by and for the membership. Member schools appoint Institutional Representatives who, in turn, elect a Board of Trustees which determines the policies and objectives of NERComP. Elections are held annually. From the Board of Trustees an Executive Board, a Finance Committee, and a Planning Committee are drawn that meet monthly. As Executive Director, the author implements and executes the policies of the Trustees, manages the daily affairs of the organization, and makes recommendations to the Board for improvements.

NERComP staff presently consists of four people besides the Executive Director. Staff conduct the principal activities of marketing, connecting and supporting member institution educational computing activities. One staff member is concerned with network operations and maintenance, one with user services, one with marketing, and one with administration.

3.0 NERComP FINANCES

The network is wholly supported by dues from members and by the margin between agency costs, for the agency computer and data link, and revenue from charges to users. Some revenue is also derived from grants and grant activities. Total revenue is approximately $300,000 per annum. As a not-for-profit organization, NERComP attempts to recover only costs from revenues. Since its founding in 1967, the agency has progressed toward a self-sustaining level while concurrently expanding the list of suppliers and conducting research and development in communications techniques and equipment.

In 1976, NERComP is phasing-down from an NSF supported User Services and Research grant and attempting, to become wholly self-supporting. This grant permitted the network to pioneer and implement the resource chaining concept whereby an interested and capable member takes on the responsibility of servicing local users and acts as a distributor for the network. Essentially they extend the agency's capacity as a distributor.

4.0 THE IMPACT OF NERComP

Staff believe that the impact of their activities is to minimize the costs to members of modernizing curricula and attaining computer literacy.

To a major supplier school such as Dartmouth, NERComP brings efficient distribution of computing power and effective use of staffing. The NERComP switching center at Wellesley, Massachusetts concentrates 11 - 22 lines onto one line to Dartmouth at Hanover, New Hampshire, using tailored hardware and software that imitates in part the operation of a Telenet TCO. By using NERComP services, the Dartmouth system can communicate with NERComP users in the same manner as with other Telenet users. The switching center provides error control, auto baud, and graceful degradation whenever the user loads exceed line capacity. The agency also provides users services support either directly or via the previously mentioned distributors.

Because NERComP has automated concentration to DTSS and manual switching to many other systems, it offers the service of connection upon demand to the system of the users choice. Port sharing and port contention are features of this switching equipment. Thus, the combination of features and some line timers permit the network to offer a unique services, overflow hourly rates in addition to part and micro rates. NERComP has built timers for specific users which permit them to have the advantage of having dedicated ports and occasionally overflowing to shared ports on an hourly basis. Terminal speeds and types range from 110 to 600 baud.

5.0 NERComP ACTIVITIES

The principal activities of NERComP are:

- Marketing
- Connecting
- Supporting
- Educating
- Consulting
- Research

The network actively solicits potential users for its many suppliers. In addition to connecting users and suppliers, either directly or through the NERComP switch or through lines shared with others, the agency supports user services by providing the hand-holding assistance required not only for terminal and line operation, but also for software use. Over 1,000 individual users use the network every month and receive Hot Line Users Services on demand.

The agency also conducts seminars and workshops especially on unique services, such as IMPRESS and PLATO. Staff continue to research alternate methods of communications, such as STDM and TELENET, and negotiate compensation with suppliers and establish rates for users. At member institution's request, staff will advise and consult on communications matters and do installation work. NERComP is also a clearing house for vendors wishing to sell communications equipment at a group discount.

Most of the agency's actions in arranging user supplier hookups follow guidelines established by the advisory groups which were established with NSF funding of research into user services. These groups were: A Governace Committee which identified relationships and responsibilities for suppliers, distributors and users; and a User Services Committee which defined the requirements for effective delivery of user services. These committees acted to create policies required for the resource chaining concept which we have now implemented.

In 1976, NERComP has two distributors both of whom support DTSS services. One distributor is Yale University. The other is Bates College. The communications link to Yale is via Telenet, to Bates via the NERComP switching center. These distributors provide support to local schools.

To guide the staff in multi-faceted dealings with schools which, on one hand may be a supplier, and on the other a user of another supplier, the results of the work of the Governace, User Services, and Technical Advisory Committees are available. The organizational problems NERComP management faces range from matching suppliers and users in a complementary yet not competitive fashion, being properly compensated for the services NERComP supplies both the suppliers and the users, and for implementing cost effective communications links.

NERComP management has attempted to reduce the problem of competing with NERComP suppliers by holding periodic suppliers meetings to outline central office business plans, to catalog unique software and to identify agency prospects, and by altering the method of charging for services to preclude any favoratism to a particular supplier. In the case of Dartmouth, NERComP has a de facto established territory in Higher Education for colleges not yet connected to Dartmouth. In the case of other suppliers and distributors the territory is not as well defined, and the few instances where central office activities overlap with the supplier are handled by negotiation.

The economics of the network, as presently configured and at the present revenue volume, are growing stronger. Approximately 40% of the suppliers standard charges are required to cover networking expenses. This would usually be paid by the supplier as a fee for services. The remaining expenses are borne by revenue from dues, grants and additional fees paid by the user.

6.0 LOOKING TO THE FUTURE

The basic challenge ahead of NERComP is twofold: *the simpler*, to deliver academic computing to the remaining have-not schools; *the other*, to uncover, catalog, promote and deliver unique services to any and all users. The agency must generate enough revenue growth to keep ahead of the rate of transfer of simpler services from time-sharing to in-house computing, or, alternatively, to find cheaper costs of communicating and interconnecting at a rate faster than the decline in costs of small systems.

A paradox of NERComP's success in sharing is that as user experience grows so does the demand for in-house computing. The cycle is completed when the minicomputer user (nee time-sharing user) returns to NERComP to acquire some of the sophisticated unique applications and to sell available bread and butter BASIC to the uninitiated. As the need for a sharing agency such as NERComP continues, the agency functions in a manner similiar to a library providing access and reference to methods and solutions. At NERComP, however, the text is never missing from the shelf but is always there waiting for the next seeker.

The search for unique products leads NERComP staff not only to deal with specific disciplines but also to consider sources of supply far outside of New England. The agency is challenged to act as a catalyst for discipline-oriented groups by bringing together foremost partic- tioners and their findings with newcomers. In academic computing, NERComP also sees the challenge of acquiring superior software resident outside of New England, either by transfer or through communications links. Through constant interaction with institutions of Higher Eudcation in New England, NERComP is able to determine not only the demand for certain software but also the capacity of local suppliers to deliver. Small colleges with under-graduate science courses now have an opportunity to gain access not only to the best in New England but also to software such as the PHYSICS DIALOGS developed at the University of California, Irvine, CA. This software, which operates on a Sigma 7, could conceivably be brought to New England via NERComP. Also excellent, but unpublicized, capabilities like MACSYMA on the MULTICS system at MIT should be brought to the attention of all network members. To determine the value of such specialties network staff make use of those members who act as advisors to NERComP who have special talents or interests. For instance, what better way to determine the desirability of MACSYMA for the have-not schools than by having a mathematician at a small rural New England College test it and give his opinion. The test by the remote probably ultimate user is an acid test. This cooperation by the members of NERComP contributes to the strength and vitality of the organization. It is in the best spirit of resource sharing that our members contribute time and advice.

In conclusion, NERComP is an action-oriented agency dedicated to academic resource chaining of computer resources; is organized to identify and measure the demand and supply for resources; is equipped to interconnect supplier with user; and is qualified to provide user services support.

Chapter 29

by Guy Ciannavei

Single-Server Cooperation in Computer Networking: The Harvard Experience

1.0 INTRODUCTION

The Harvard-MIT arrangement under which Harvard purchases IBM 370/168 computing services from MIT, is now starting its sixth year of operation. Within the first quarter of 1977, I expect that Harvard will formally announce its intention to continue the arrangement for year number seven. In the computing business, seven years is a fairly long time and this factor of longevity alone leads one to the conclusion that Harvard and MIT must be collectively doing something right.

In order to help gauge the true dimensions of that success, however, one must point out that not all of the computing needs of Harvard University are satisfied by the arrangement with MIT. Moreover, a very significant portion of Harvard's computing needs are satisfied completely independently of the Harvard Computing Center. To better understand why this is so, one needs to know how Harvard is organized and where the Harvard Computing Center fits within that organization.

2.0 HARVARD UNIVERSITY ORGANIZATION FOR INFORMATION TECHNOLOGY

The primary product of Harvard University is education. That education is produced by ten separate graduate schools and one undergraduate college. The deans of each of these schools operate with complete autonomy in academic affairs and with no less autonomy in fiscal affairs (which comes as a surprise to many people). The dean of each school is responsible for the financial stability of his organization. He discharges this responsibility first by generating revenue via tuition and other fund raising devices and then by exerting strong controls over the manner in which those funds are spent by his faculty.

The university does provide many centrally-managed services to the various schools. However, none of these services are provided free of charge. Some of these services exist as natural monopolies like personnel administration, student billing and the university police. There are other central services, such as the printing office, which do not enjoy monopoly status and must compete with external sources. Computing is a centrally provided service that must compete both with external sources and with internal do-it-yourself efforts.

The Harvard Computing Center is the duly constituted central source of supply of computing services. The center is operated as a branch of the office for information technology (OIT). The Office for Information Technology is part of the financial systems organization which is headed by the vice-president for financial affairs.

The primary mission of OIT, is to provide advice and guidance to all areas of the university in technological methods for processing and conveying information. In those areas where OIT perceives an unfilled need for services, it may provide for the production of those services by the establishment and maintenance of operational facilities. OIT currently is the organizational home for three such operational facilities whose services are available to all areas of the university on a fee-for-services basis. One is the video services center which supplies and maintains a large inventory of video programs and devices. The second operational facility is the applications development group, which consists of a staff of approximately 35 systems analysts and computer programmers. The third service organization in OIT is the Harvard Computing Center. Harvard University's highest governing body has assigned OIT the responsibility for the coordination of all computer activities at Harvard. The mission of the computing center is to provide those computing services for which OIT perceives an otherwise unfilled need.

3.0 THE HARVARD COMPUTER CENTER

The internal organization of the Harvard center is not significantly different from any other college computer center. The staff of 80 persons collectively provides the full range of classic computing center services. The staff is compartmentalized into six functional groupings: systems programming; machine operations; data preparation; data control; user services; and an accounting and billing office. The systems programmers maintain the IBM 370/145 operating system, and the Harvard end of the MIT Interface Link. They participate in analytic studies for future planning, and they provide a full range of user assistance services including development of user documentation and the conduct of user training courses in programming languages and available packages. Their user assistance functions require extremely close communication and liaison with their counterparts at MIT, since the majority of Harvard users who need assistance are using the MIT CPU.

The Harvard Computer Center organization is somewhat unusual since almost half of the staff is involved in functions subservient to administrative applications. The data preparation and control groups are almost wholly dedicated to the business end of Harvard Administration. The computing center, then, is both research oriented and administratively dedicated.

The center runs an IBM 370/145 with a VS1 operating system. This computer is used to process most of Harvard's administrative production jobs while simultaneously serving as a remote job entry device to MIT's IBM 370/168. The 145 has three card-readers, 348K of core, 4 tape drives, 4 mountable disks and four 1100 line-per-minute printers. The remote job entry function to MIT's 168 is performed by an IBM-produced HASP workstation package. Two other remote job entry stations at Harvard tie in via leased lines directly to MIT's 168. A REMCOM 2780 is located at the center for behavioral sciences at William James Hall. The other, a Data 100, is on the Boston side of the river in the Harvard Medical area at the Harvard School of Public Health. These stations are operated with financial support from the computing center but without the benefit of the computing center's managerial control. All revenue collected from users of these remote stations for 168 services accrues to the computing center. These stations do not access the 145.

The services of the center are available to all organizational components of the university and affiliated bodies such as the teaching hospitals and other non-profit organizations. The center does almost all of the central administrations' computing as well as most of the administrative applications required by the various faculties. There is also a large mass of research-based computing and this work represents the bulk of Harvard traffic with MIT. The most commonly used language over the MIT link is FORTRAN, and the most popular packages are DATATEXT and SPSS.

4.0 THE LINK TO MIT

Users of the center represent each of the schools, all areas of the central administration, and the affiliated teaching hospitals. In dollar terms, central administration spends the most providing 42% of total revenues on the center, with the Faculty of Arts and Sciences running second at 20%. On the MIT 168 Link alone, however, the places are reversed with the faculty in 1st place at 24% and Central Administration running second at 20%. It is interesting to note that, on the Boston side of the river, the School of Public Health, the Medical School, and the Teaching Hospitals provide collectively 26% of the 168 dollar volume while providing only 16% of the center's total revenue. School faculty representation is at the trivial level only at the schools of dentistry and divinity. Of the total revenues earned by the center, only 17% derive from federal contracts and grants, while 25% of the 168 computing is federally based.

The decision to assign work to either Harvard's 145 or MIT's 168 is made *in every case* by the user. The computing center provides advice and counsel only. Of course, the decision is seldom a difficult one, since the service characteristics of the two machines are so markedly different. The 168 will turn the average job around in twenty minutes; on the 145, twenty minutes of negotiations may be required with operations personnel to establish whether the job will be run before, or during, the third shift.

Since the preponderance of jobs processed by the 145 are administrative production jobs, the regularly recurring production cycles lend themselves to a high degree of advanced scheduling. Most of that work is shepherded through by the center's staff of data control technicians. The occasional 145 individual user, therefore, most often finds a long line of big jobs waiting ahead of his/hers.

It should be noted that on new administrative applications and projects still in the development or pre-production stages, programmers normally will submit all their debugging runs to MIT's 168, capitalizing on the vastly improved turnaround and justifying the possibly costlier service with offsetting savings in programmer productivity. After an application has been proven ready for production, it will be transferred to the 145 if running costs are lower than on the 168. Some debugging is not done on the 168. The vast majority of Harvard's general accounting programs were originally developed under the DOS Operating System on an IBM 360/30 that was formerly operated in the comptroller's office. Since the 168 does not have a DOS Emulator, none of those jobs can qualify for 168 processing.

In addition, there are considerations of security and privacy which tend to prevent certain sensitive applications from being run on the MIT machine. Such sensitive applications include the hourly payroll, the weekly payroll, the semi-monthly payroll and, of course, the monthly payroll.

The 145 is also designated for use as backup to the 168. With the exception of DOS jobs, any job that will run on one machine will also run on the other. The Harvard Computing Center has developed a program called SCAN that translates 168 ASP job control language to OS/VS JCL. Fortunately the 145 has had to perform a backup role only rarely, but occasionally during thirdshift hours, when MIT service is not available, the 145 will process urgently needed device-independent jobs that otherwise would have gone to the 168. This is never done, however, without the specific consent of the user.

5.0 THE COMPETITIVE ENVIRONMENT

Perhaps the most significant single unifying characteristic of the Harvard user community is that it is free to acquire its services anywhere it chooses. Indeed there are groups who purchase services from external sources completely independently of the center. For example, the college's chemistry department buys much of their computing from Columbia University, and the astronomy department uses heavily a CDC machine at the Smithsonian Astrophysical Observatory. The undergraduate college and the business school operate their own independent medium size computing facilities which satisfy most of their educational computing requirements. The computing center, therefore, operates the competition with the whole world, except MIT's 168.

As mentioned previously, the Harvard Computing Center provides its services on a strict fee-for-service basis. It is constrained to recover all of its operating costs from its users. There is no subsidy or other financial assistance granted to the center by the university. The center pays its own rent, heat, electricity, telephone and janitorial bills. The only services that it receives free from the university are personnel administration, accounting services and, of course, general managerial guidance. The center operates on a genuine full cost recovery basis. The schedule of prices charged for its services is intricately detailed so as to attempt to spread the burden as equitably as possible among users, consonant with the services they require. This results in a multi-page catalogue of services and prices which usually overwhelms the first-time user.

Working under these rules with no subsidy, complete cost recovery and competition with the world, financial considerations tend to play an overwhelming role in the process of deciding in what manner the computing center will produce computing services. The build-or-buy decision leans strongly on the fiscal facts of life. This is where MIT comes into the picture. The Harvard Computing Center has been able to buy wholesale services from MIT at a cost lower than the cost of building comparable services in-house. Although center management continually window-shop for substitutes, they have found none who appear able to compete with MIT.

One of those window-shopping excursions resulted in an arrangement with Princeton. Center management had determined that, for some CPU-intensive jobs, the Princeton retail prices were significantly less expensive than MIT rates. The center did establish the link with Princeton, announced the service to Harvard users, and prepared for the rush. The rush never came. Not one user has ever requested that his/her job be run at Princeton. That experience has taught center management that lower prices alone are not sufficient to cause user migration. Obviously the combination of providing only over-night service and requiring users to convert job decks to the Princeton conventions were strong inhibiting factors.

The MIT arrangement on the other hand, in addition to its generally favorable structure, provides an attractive combination of speedy service and stability. Any job that worked on MIT's 370/155 in 1971 will work today on the 168, with faster turnaround and, for the Harvard user, a lower unit cost.

The Harvard-MIT arrangement operates under a written agreement which is re-negotiated annually. For this year Harvard has agreed to buy monthly minimums of 168 services approximately as follows:

- 30 hours of central processor time
- 25 thousand kilo byte hours of memory
- 10 million tape & disk I/O's
- 75 hours of disk use
- 90 hours of tape use, and
- 8,000 device units

All of this will require, at Harvard, an input of 4 million cards resulting in 20 million lines of print-output, distributed across 12,500 batch jobs, and 600 hours of TSO terminal connect-time. These parameters constitute the minimum monthly profile. Actual usage will exceed this minimum guarantee, and there is a schedule of unit price decrements for corresponding increments in usage.

The amount that Harvard pays to MIT represents approximately one third of the total cost that the center expends to distribute these services to Harvard users. The end cost of the Harvard user is in no case any higher than that which would be paid by an MIT user for the same job. Also, the Harvard user is paying less for the average job now than he or she would have paid 5 years ago.

Although financial considerations have been or great importance in the decision making process for continued renewal of the MIT arrangement, there are also other considerations. For example, Harvard could quit the MIT service and upgrade the 145 to a larger machine. This would give the Harvard center a small reduction in overall operating costs at the expense of drastically increasing turnaround time for former 168 users. It is the considered opinion of center management, however, that most users would not appreciate a 5 or 10% cost savings if the waiting time were to grow two or three times longer than that to which they have become accustomed.

An additional consideration of no less importance is stability for Harvard users. The program conversion effort required to move to the 370's was highly traumatic for the Harvard community. In the present competitive environment in which the Harvard Computing Center operates any similar mandatory conversion would result in large scale migration of users to other sources of supply.

6.0 SUMMARY

Harvard prefers to use MIT-provided computing service for three major reasons; low cost, speed and stability.

MIT has proven capable of providing eminently satisfactory turnaround time, and MIT's recent long-term commitment to the 168 augers well for future maintainance of performance at no increase in hardware costs. The long term stability of MIT's services has been excellent. There may be some clouds on the distant horizon, however. MIT does plan to move to a virtual operating system, MVS, and they plan to implement the JES2 control language. Center management at Harvard perceive a resulting adverse impact on stability. Hopefully, any irritating effect from that change will function like the grain of sand around which the diligent oyster builds a shiny pearl.

Chapter 30

by J. R. Steinberg

Single Server Cooperation: The Supplier's Viewpoint

1.0 THE MIT IPS-ACADEMIC AND RESEARCH COMPUTING CENTER

The MIT Academic and Research Computing Center serves two roles; it is both a supplier of computing services to Harvard and NERComP and a distributor of computing services, to the MIT community. The fact that the center is both a supplier and a distributor is a very important consideration. A brief description of MIT facilities and services, that product that the center distributes and supplies may be useful to put the supplier/distributor role in context.

The Center operates two large-scale computer systems. The Honeywell 6180 provides the Multics time sharing service, and an IBM System 370 Model 168 provides a multi-processing batch service and a time-sharing service. MIT users are fairly evenly divided between the two systems. There is some preponderance in numbers on the System 370 side, but very few users are ambidextrous, that is, able to use either system with equal proficiency and ease. Although MIT Information Processing Services staff distribute services on both systems, as a supplier IPS provides System 370 services almost exclusively. Although IPS has been encouraging both Harvard and NERComP to broker Multics services to their communities, this is a fairly recent development and there is not a great deal of activity yet.

MIT-IPS is tied into three networks. The System 370 is tied into both the NERComP network and the Telenet network, the second as part of EDUNET. The Multics system is tied into Arpanet and Telenet. Either system is accessible through standard dial-up telephone lines from anywhere in the world.

The System 370 also has several direct lines connecting it to remote printer/reader stations. One of these lines runs through campus to the East Campus computing facility which is a Data 100 remote terminal. (There is also a PRIME computer on this site and plans exist for tying it into the S/370.) Through a coaxial cable connection, and IPS runs, the terminal at 9600 baud using homegrown modems which do not actually modulate or demodulate.

The System 370 is also tied in directly through two mechanisms to the Harvard University Computing Center which is approximately five miles away. One is the leased 40.8 kilobit telephone line using the Bell 303 package on each end. The other is a 40.8 microwave connection between the William James Hall at Harvard, Harvard's skyscraper, and Building 39 at MIT, the building that houses the computers. The microwave link is a line-of-sight, dish-to-dish connection. Since one of these high speed kilobit connections is actually in use at any time, the other functions as a spare or backup. In addition to the data transmission band, the microwave facility has capacity for several video channels, and these are used fairly frequently for telecasts between the two universities.

Finally, leased lines between the MIT System 370 and the other remote stations at Harvard and the System 370 also provides a facility for dial-up service to remote entry stations. Dial-up service is used by stations that have too little usage to justify the cost of the leased line, i.e., Tufts and Brandeis Universities.

2.0 DIFFERENCES BETWEEN MIT AND HARVARD USER INFERFACES

The System 370 operating system is OS/370 MVT ASP with the time sharing option (TSO). Although both the MIT communities and Harvard communities are running using the same operating system, in many ways the user interfaces to this system are different for the two communities. For example, the separator pages that surround the user's printed output have different formats for the two schools to conform to the different filing procedures used by the separate dispatching areas. At MIT the output is filed by jobname, actually a sequential number produced by the use of prepunched jobcards. At the Harvard University Computing Center the output is filed by programmer name. Therefore, on Harvard jobs, the programmer name is printed in large stencil like letters along the edge of the separator sheet. For the output of TSO submitted jobs, MIT has adopted this Harvard convention even though staff still file output by jobname.

There are a number of such differences in the user interfaces; the Harvard user runs at different prices than the MIT user; the Harvard user uses a slightly different account identification card than the MIT user; the Harvard user's account number is of a different format than the MIT user's; the OK2RUN checking process, though similar, is slightly different for each user and the checking is done against two different lists; the Harvard user and the MIT user have different sets of cataloged procedures; the Harvard user can specify a number of service requirement indicators which set prices as well as job priority which the MIT user is not allowed to use; and so on and so on. It is important to remember that both communities are using the same operating system, albeit an operating system that has been heavily modified to preserve both the MIT and the Harvard identities. It is almost safe to say that there is no reason for the Harvard users to suspect that they are using someone elses computer.

In some places, the users' interface is the same for both communities where IPS would rather it were not. However, it was too difficult to change. For example, at the end of each job's printed output there is normally an Announcements Page, a page which contains important notices such as changes in operating schedule, short course announcements, changes in software packages, and so on. These notices are usually prepared, using a TSO program, by the person most generally responsible for the area concerned. If that person does not recognize that both communities will be reading the same notices considerable confusion can result. This problem is amplified somewhat because most IPS staff are involved only, or primarily, in the distributing function.

3.0 MIT-IPS AS SUPPLIER

MIT is supplying computing services to a distributor, the Harvard University Computing Center, not to the Harvard users directly. Consequently MIT does not supply user consulting, user documentation, or even printed output. The MIT-Harvard arrangement is actually a very simple model of the chained resources network that NERComP has been developing. Because of the geographical consultants, user services are best distributed to the Harvard community by the Harvard University Computing Center, just as IPS distributes those services to the MIT community.

So what are the things that IPS supplies to Harvard University Computing Center? IPS does maintain a number of application programs on the system that are used by the Harvard community, just as Harvard maintains a number of application programs which are used by the MIT community. Although MIT maintains most of the compilers, loaders, system utilities, etc., the Harvard Computing Center does maintain COBOL and the associated COBOL–TSO tools. MIT-IPS maintains the operating system, and is responsible for the equipment associated with the central processor, selection, operating, maintenance and the physical environment in which it runs.

3.1 Contractual Arrangements

In these areas, where MIT-IPS provides a direct service, that most of the dialogue and problems occur. However, it is interesting to note that these services are not included or even mentioned in the final agreement which Harvard for the provision of computer services. That agreement specifies guaranteed minimum annual expenditures and estimated monthly usage where usage is measured by dollars expended at MIT rates.

The agreement between Harvard and MIT is quite interesting for a number of reasons. Remember that Harvard and MIT, at first, were partners in this venture and shared equally the cost of the central facility in which the central processor was an IBM System/370 Model 155. Note that sharing in the cost of the central facility, was for both MIT and Harvard expense in addition to the cost for all of the ancillary facilities including the online printers and card readers, the dispatching area, the user consulting and documentation functions, etc. As a result of the different cost bases and the different usage bases, Harvard and MIT arrived at different charging rates for their users. Since users found it difficult to understand why the same service on the same machine had different prices at the different centers, both MIT and Harvard staff spent a lot of the first year explaining.

After about one year of a partnership arrangement, MIT realized that a larger machine was necessary, particularly if TSO services were to be offered. At that time, however, Harvard felt no need for expansion and was not willing to share the cost of a larger central facility. Accordingly, the economic arrangement was changed although much of the spirit of the partnership still remains. Many of interface changes reflect those early partnership days and, of course, set precedents for the future.

Having decided that they did not want to share the cost of the larger central facility, but were still interested in retaining the service, Harvard agreed to set its fair share of the cost for the coming year, directly on the original partnership cost. Both institutions assumed that there would be no large increase in Harvard's usage, and that the standard measure of their usage would be MIT dollars. For this amount of usage Harvard pays a guaranteed minimum with additional surcharges if the monthly, or annual, usage exceeds certain leeways built into the profile.

Over the years Harvard usage, as well as MIT's usage, has increased slightly, and the Harvard profile includes TSO usage as well as batch usage now. However, it it still reasonably safe to assume that Harvard's usage is equivalent to approximately one half the capacity of a 155.

3.2 Mutual Trust and the Spirit of Partnership

There are several services that MIT provides Harvard like systems maintenance and development, operations, and some applications maintenance and development. What guarantee is there to Harvard that these services are up to scratch? What guarantee is there to MIT that Harvard won't demand more or better of these services than IPS is willing to provide? The answer is, and some people would laugh at this, mutual trust and the spirit of partnership. Examples of this trust are console privileges MIT allows the Harvard Center, and the use of the advanced scheduling service indicator that provides them an extremely high job priority. Abuse of either of these privileges could seriously impact MIT's access to the system. Conversely, they trust MIT-IPS with their tapes, their disks, and of course, their jobs. Harvard assumes, and reasonably so, that MIT will maintain these services at a reasonably high level. Of course, MIT does have an additional responsibility to its own community in these areas. MIT assumes that Harvard will not make demands on these services that are inconsistent with their financial contribution. It is very much an article of faith and pretty much of a gentleman's agreement. Does it work?

In each of the service areas a certain amount of dialogue must be maintained between people providing the function at MIT and people providing the similar function at Harvard. Perhaps the most continual dialogue is at the operations level. First of all, there is a hotline connecting machine room to machine room. Whenever something unusual is going to happen at the central facility Operations makes it a point to use the hotline to notify the Harvard computing room operators. Conversely, whenever Harvard is going to make an unusual demand on the system, the Harvard operators notify the machine room operators at MIT. Whenever there is trouble at either end, the hotline is busy, of course. In addition to the hotline there are other day-to-day operational lines of communication. Harvard has a messenger service which delivers and picks up in the MIT machine room about four times a day. Part of this delivery service usually includes signing in and out tapes and disks. Consequently, another daily contact occurs between the tape librarians in both installations.

There is also daily operational level interface between the user accounts offices of the two centers. Although Harvard can, and does, register batch users into the system because they have a separate

OK2RUN list, TSO maintains only one registration list. The MIT user accounts office must be informed whenever Harvard wishes to register a new TSO user. There is also communication if anything happens to go wrong with the accounting metering files. These files are a common set of records used by both centers to keep track of usage and to bill their users.

Harvard's systems programming staff's main function is the support of the VS1 system for the 145. However, as systems programmers they do have some ideas of how the operating system at the central facility could be modified to improve service and, they are more than willing to share those ideas. They also maintain some parts of the operating system, particularly those parts associated with the COBOL compiler and the Harvard catalogued procedures, and are quite involved in the review of specifications for the MVS system which MIT is now implementing. The interface between MIT system programmers and Harvard system programmers is busy.

At the applications programmer level and at the user services level, there is also some communication. Staff share information about problems and changes and occasionally collaborate on assisting a user.

3.21 Formal Management Consultation

Perhaps, the most important interface is at the management level. There is a weekly meeting of the management of the two centers scheduled, though lately this meeting has actually been held only about monthly. At this meeting Harvard usually will bring up complaints about system reliability, operating procedures, equipment failures, and all of those things about MIT services that they feel have not been up to snuff. MIT representatives will complain about the weather, the economy, the government and generally try to explain why problems have occurred. These meetings also provide an opportunity to share problems, and sometimes solutions, and occasionally to plan. Often these meetings are used as a vehicle for arguing out differences in philosophy and priorities. Although, without an arbitrator, these debates are often not revolved or even scored.

4.0 PROBLEMS

What are the problems in an arrangement such as this one? Although MIT is both a supplier and a distributor, most of the MIT center's staff are primarily concerned with the distributing role. They tend to think of the Harvard Computing Center as a strange and demanding user. There is occasionally some resentment when things must be done differently then they would like because of Harvard's involvement.

This shows up, for example, in our review of the MVS specifications. Harvard and MIT have different priorities and quite different views of their users. Because of this they have quite different views of what the system should look like and what changes are tolerable and which are intolerable.

Originally there was also considerable user resentment. Somehow or other if the system was down or turnaround was bad or anything at all wasn't just right, the MIT users were convinced it was the fault of the Harvard usage. I expect that the Harvard users had some choice words for MIT as well. After a year or so, however, the arrangement came to be accepted and IPS staff now hear few complaints from users about Harvard's use of the system.

Another problem, of course, is the additional work required to maintain the separate interfaces. The brunt of this work is carried by the systems people, though it occasionally reaches down into all levels. Clearly the perceived need to preserve the identities of both centers does create problems.

Other problems occur in communications. Occasionally operators forget to tell Harvard that MIT is planning on scratching a library until after they've done it. Sometimes Harvard neglects to tell MIT about a large number of jobs coming through that must be finished that evening. Perhaps it would be more truthful to classify this problem as an awareness problem rather than a communications problem. Certainly the communications lines have been set up, but staff on both sides of the river sometimes find it difficult to remain aware of each other's needs. This relates again to the fact that most of both staff are in the distributing function not the supplying function.

Suprisingly there is some competition between the two centers for a common customer base. This is hardly a day to day problem, but there has been at least one occasion in the past, with a third party involved, when this became a very serious problem. Fortunately, it was also a problem which could be straightened out and was.

In the area of standards for performance, there are occasions when MIT is willing to settle for a lower level of performance than Harvard. A recent case in point has been the reliability of the tape drives. Harvard felt quite strongly that too many tape jobs were having difficulty while MIT felt that performance in this area was more than satisfactory. The upshot of this, and hopefully the upshot of most such disputes, was a raising of the level of performance and reliability in this area. (The customer is usually right, anyway.)

5.0 CONCLUSION

It's an interesting arrangement, sometimes a trying one, but certainly economically advantageous for both institutions. If Harvard and MIT had it to do over again they'd probably *want* to do some things differently, but probably wouldn't.

Chapter 31

by Michael S. Sher, Roland D. Spaniol

Public Regional Computer Networking

1.0 THE ILLINOIS EDUCATIONAL CONSORTIUM

The Illinois Educational Consortium is a not for profit corporation whose members are the four Illinois public systems of higher education. Each member annually selects two representatives, one of whom is the member's executive officer, to sit on the Consortium's Board of Directors. Three non-voting Directors represent the Illinois Board of Higher Education, the Illinois private universities and the Illinois Community College Board. In November 1976, the Consortium has thirteen employees. The state universities provide additional staff support on a protracted or as-needed basis for specific studies or programs.

2.0 FORMATION OF IECCS

In September, 1972, the thirteen Illinois state universities, through their four governing Boards, established the Illinois Educational Consortium for Computer Services (IECCS). The creation of IECCS was in response to the recommendations made by the Illinois Board of Higher Eudcation and the Illinois Department of Finance to the Joint Council on Higher Education, which includes the four governing Boards' executive officers and the presidents and chancellors of all public universities. A need was identified for creating an organization to increase cooperation, coordination and sharing of computer equipment, systems, facilities and personnel among the institutions of higher education in the state of Illinois. The Consortium was incorporated as an administrative entity for cooperation between higher educational units.

The express purpose of IECCS was to advance the development and use of computing technology and information systems in institutions of higher education in order to achieve improved management and more effective education, to achieve improved cost efficiency, and to strengthen and enrich the public universities' computational capabilities in support of their educational missions and operations.

In its first two years of operation, IECCS explored the viability of cooperative programs to deliver computing services. Studies performed under the auspices of IECCS provided a clearer perspective of the desirability of resource sharing and cooperative programs. During this period, two networking service were created by the state universities. The Cooperative Computer Center was established by one of the four governing Boards' (the Board of Governors of the state colleges and universities) to consolidate administrative and instructional computing services at three of its universities. Also, the University of Illinois consolidated administrative computing at its three campuses. In addition to performing several studies, IECCS:

- *Developed pilot programs* to give several of the state universities remote access to the computer center at the University of Illinois at Urbana-Champaign.
- *Facilitated the transfer of computer equipment* becoming marginally useful at one university to other educational institutions where the benefit was greater and the cost relatively less. Proceeds from the transfer were credited to the original owner to reduce lease costs of new equipment.
- *Acting as purchasing agent* for several state universities in acquiring computer software, training aids, and support equipment where common needs could be identified and bulk discounts could be negotiated.
- *Explored financing of computer equipment lease/purchase* for universities as an alternative to renting or leasing the equipment from commercial vendors.
- *Fostered planning* among several state universities for establishment of a shared facility for instruction and research computing.

3.0 TRANSFORMATION FROM IECCS TO IEC

A broadened scope of Consortium activities was reflected in amendments to the corporation's Articles of Incorporation which becam effective April, 1975. The corporation's name was changed to th Illinois Educational Consortium (IEC). The expressed purpose of IE is to support and advance the collective activities of institutions higher education in the state of Illinois as a means of improved management.

IEC has encouraged, and assisted in establishing, interuniversity cooperative programs leading to more cost effective delivery of services through resource sharing, collective purchasing and economical financing arrangements. IEC is supported by membership dues, service fees and grants. Fees are charged by IEC for established services such as leasing. Members' dues support program development and IEC's role as facilitator and coordinator. Facilities and personnel are provided by members as required for specific activities.

In July, 1976 IEC was recognized by the Internal Revenue Service as exempt from Federal income tax as an organization described in Section 501(c)(3) of the Internal Revenue Code of 1954. This is a reversal of the IRS's adverse ruling of May 1974 and the result of the Consortium's court actions which began in 1975.

IEC expects to effect evolutionary, though not dramatic, changes in the management of several specialized state university service functions supporting their educational missions. During the fiscal year 1975-1976, IEC concentrated in six areas.

3.1 Consultant Services

In conjunction with the professional staff of the state universities, IEC has provided consulting services to member schools in the operation and management of academic computing facilities. IEC has also provided assistance with acquisition of used computer equipment and has profitably disposed of members' used equipment traded-in for newer equipment. Through its dealings with the State universities, IEC staff has become knowledgeable of members' experiences with vendors and products as well as members' plans to release or acquire equipment. This information has been used to the benefit of several member universities.

3.2 Distributor Services

IEC has purchased over 600 computer terminals for sale or lease to member colleges and universities at discounts of 20-40%. The staff expect to acquire an additional 300-400 terminals during calendar year 1977. Where the arrangement has been useful, IEC has also been used as a contractual vehicle to obtain significant discounts in the acquisition of mini-computers, computer peripherals, computer software and audio-visual training aids for its members.

3.3 Computer Equipment Leasing

IEC expanded its leasing operations by purchasing several computer systems which had been leased by manufacturers to Illinois state universities. These include an IBM 370/158 leased to Southern Illinois University-Carbondale, an IBM 370-145 leased to the Cooperative Computer Center, an IBM 370/145 leased to Illinois State University, a Hewlett Packard 3000 leased to Sangamon State University, a Data General Eclipse S/230 leased to Eastern Illinois University, and a Digital Equipment Corporation PDP-11 leased to Southern Illinois University-Carbondale School of Medicine. Over a five-year period these participating universities have estimated five-year savings of over a million dollars by leasing this equipment from IEC rather than leasing the equipment directly from manufacturers.

3.4 Cooperative Purchasing

During FY76, IEC collective purchasing activities were focused on initial organizational, planning and operation phases. The IEC Purchasing Advisory Council, composed of the purchasing directors of the thirteen state university campuses, joined together to judiciously expand (from 7 to 24) the list of generic commodities for which there was an indication of service and monetary advantage through collective procurement. In FY76, the dollar expenditures for commodities collectively bid increased from the FY75 level of $1.1M to $2.9M. The timing of the procurement schedules was also greatly improved. In FY75, the four governing boards of higher education adopted common purchasing procedures developed under the auspices of IEC. Amended resolutions providing for a further commitment to collective purchasing were adopted by the four governing boards during July, 1976.

3.5 Computer Services Advisory Council

The directors of computer centers that provide services sponsored primarily from state appropriations to the thirteen Illinois state university campuses have voluntarily formed an advisory council to the Consortium's president. Among the current activities of the advisory council is the development of a Directory of Computer Services. Members of the advisory council agree that some identifiable body of standards and practice, designed to help an institution determine its actual computing needs and measure the capability of its present resources, would be of significant value to both university officers and computing center managers. It is also recongized that adequate knowledge of existing facilities and conditions of operation is an essential prerequisite for any effort at coordination or planning. By taking the initiative to develop an appropriate reporting system, the Council hopes to prevent the application of arbitrary ad hoc standards by persons unfamiliar with the requirements of computing in an academic environment. To support these objectives, the advisory council will develop an appropriate set of questions facing academic computing administrators, conduct studies at the state supported universities, and compile the results into a Directory of Computer Services. This standard reference document which will cover a broad range of subjects, from institutional mission to equipment configuration, will be updated periodically. The current configuration of state funded computer systems at the state universities that will be included in the Directory is shown in Figure 1.

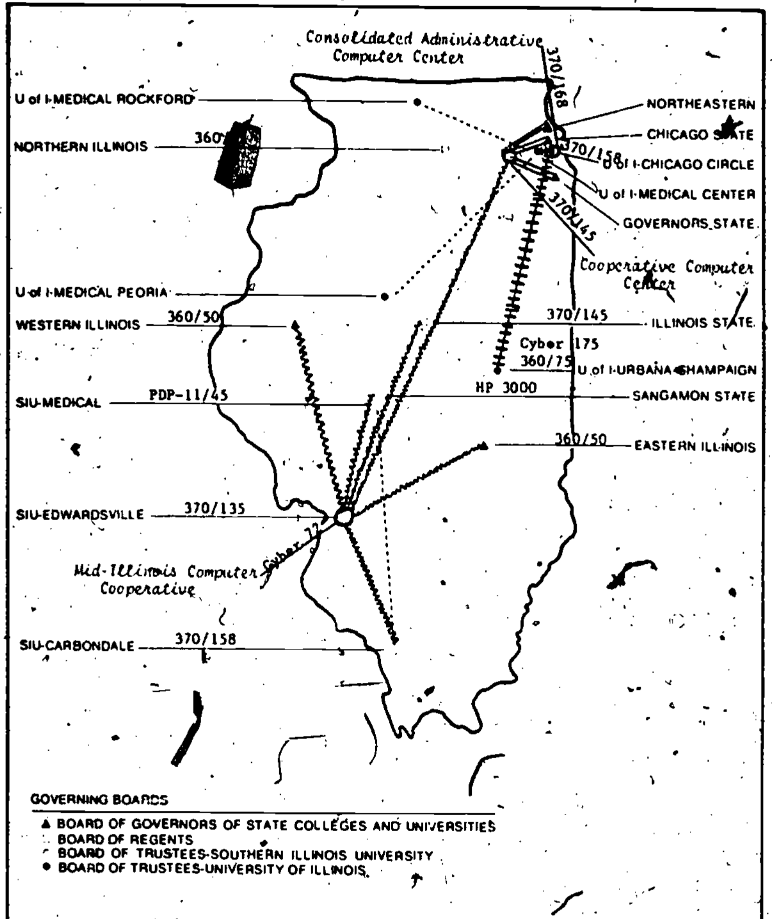


FIGURE 1. IEC Member Universities Computer Configuration

3.6 Mid-Illinois Computer Cooperative (MICC)

MICC (Mid-Illinois Computer Cooperative) is a cooperative activity operating under the auspices of IEC sponsored through a Higher Education. Seven of IEC's thirteen employees are MICC staff. MICC initially provided remote time sharing computer services for instruction and research programs at five state universities (See Figure 2). In February, 1975, MICC began providing services from a Control Data Cyber 72 computer system located at Southern Illinois University-Edwardsville. FY76 was the first full academic year in which the system was operated. Monthly system use was doubled compared to the first semester of operation in FY75. Southern Illinois University-Carbondale became the sixth member of MICC in September, 1976. Governors State University, Chicago State University, and Northeastern Illinois University will join MICC early in 1977 (the administrative computing requirements of these three universities are supported by the Cooperative Computer Center). Continued growth in service offerings to current and new members is expected (See Figure 2).

MICC was established on the premise that certain types of computer services, such as interactive timesharing and fast turnaround batch processing could be provided most effectively and economically to several universities on a cooperative basis. Although the technical feasibility of a computer network of this type had been previously established, there were organizational and administrative problems to be overcome. It was decided that these problems could be best solved if direct, centralized funding of the network was provided during the first crucial years of operation.

The significance of "Cooperative" in the title of MICC signals the the significance of the organizational and administrative problems to be overcome. First, the MICC is a non-legal entity and all actions such as hiring personnel and contracting for hardware are handled by IEC. Second, since the member institutions furnish the campus user coordinators, the MICC Network Manager must work with a major personnel resource over which he has little direct control. Thirdly, none of the member institutions were forced to join. If they can satisfy their users locally, they may withdraw from the cooperative. The members generally agree to the concept of economy of-scale and therefore have accepted the responsibility of retailing computing resources to educational institutions in their marketing area; there are over 100 higher education institutions in Illinois. Lastly, the members generally agree that the success of the cooperative will be based on the richness of software/courseware available and they have initiated cooperative software sharing efforts. In fact, the first MICC User Conference was held three months prior to delivery of the hardware.

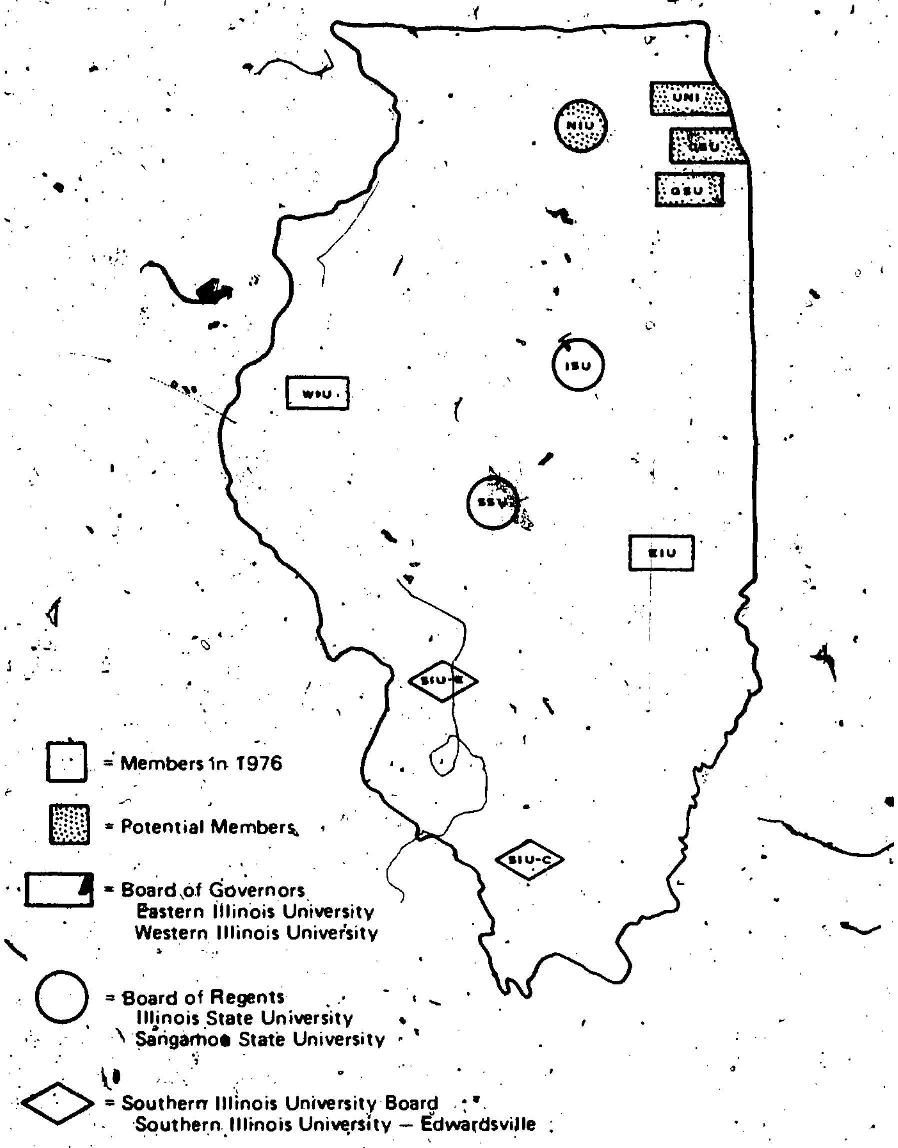


FIGURE 2. MICC Member and Potential Members 1976

Each state university that elects to join MICC makes substantial contributions to this program. One or two staff members from each university serve as representatives on the MICC Operating Board, the governing and policy making body for the Cooperative (Figure 3). In addition, each school provides released time for a professional staff member to act as interface between MICC and computer users on the local campus. Physical facilities for the MICC operation are provided by SIU-Edwardsville. Several of the member institutions have also elected to purchase computer services from MICC in addition to the basic level of support* which is currently provided without charge.

The program is supported over a 3 1/2 year period of direct funding to IEC (January, 1975 through June, 1978). Supplemental allocations will be provided to member universities for FY 1978-79 to allow them to continue receiving the basic level* of computer services and purchase additional services through a contractual arrangement with IEC. The 3 1/2 year period of centralized funding is intended to allow MICC, and its member universities, to reach a level of stability sufficient to continue to serve its members. To reach such a level of stability, MICC must pursue three major areas of development:

- *Operational maturity of the Cooperative* in mechanical, technical and procedural activities.
- *Maturity of the member institutions,* i.e., 1) informing users about the resources available and incorporating the use of these resources into user academic programs, and 2) developing a management philosophy regarding the use of locally controlled and remote shared resources.
- Maturity of management and administrative policies and procedures necessary to allow for smooth operation of the Cooperative after the termination of centralized funding.

Prerequisite to the success of MICC is the establishment of an organizational structure that deals with a dynamic evolution of technology to support cost-effective computer services. The optimum method of providing a full range of computational support for several institutions is a mixture of distributed and centralized facilities (See Figure 4). Operations best handled locally will be provided by individual institutions. Operations in which significant economies of scale can be realized will be centralized with costs shared by several schools. When institutions provide specialized services not available at other member institutions

* The "basic level" of support is defined as the access and use of system resources through seven asynchronous and one synchronous ports and the central staff support for the maintenance of the network and the implementation and coordination of software.

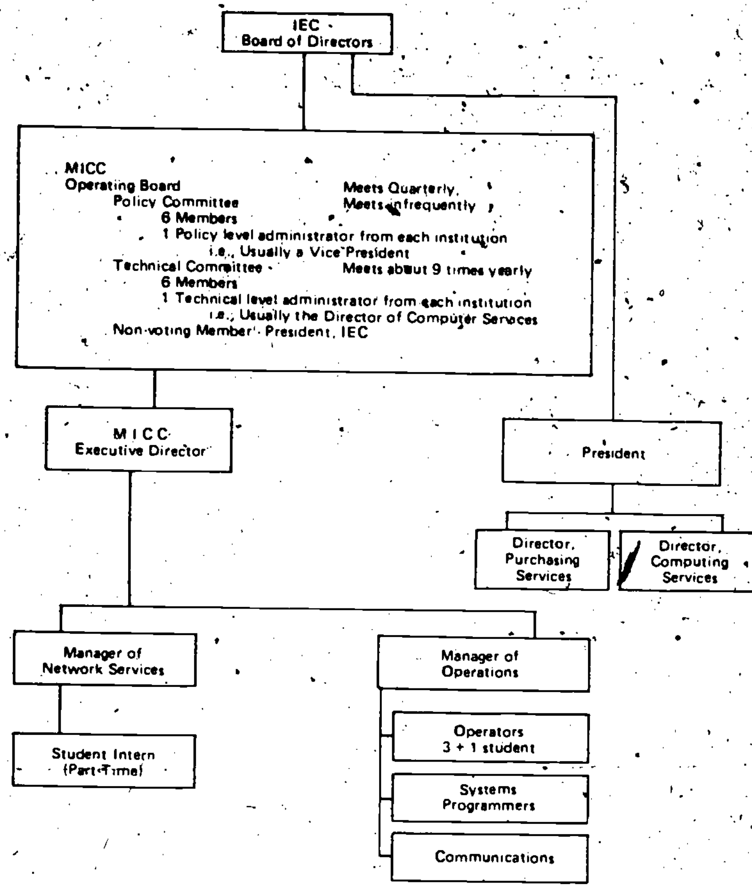


FIGURE 3. Structure of IEC/MICC

and of a nature not to be efficiently centralized, those services will be made available to the other members through the MICC network.

FTE On Campus Students			On Campus Computer
BOG.			
EIU	9,112		IBM 360/50
WIU	13,815		IBM 360/50
CSU	4,149		IBM 370/145
GSU	2,271		IBM 360/50
UNI	6,319		
		35,666	
BOR			
ISU	17,777		IBM 370/145,
NIU	18,463		IBM 360/67
SSU	2,073		HP 3000
		38,313	
SIU			
C	19,987		IBM 370/158
E	9,529		IBM 370/135
		29,516	
		103,495	

FIGURE 4. Profile of MICC Member Universities

The primary objective of MICC is to provide cost-effective time-sharing computer services to educational institutions. MICC provides reliable time sharing computer hardware and system resources. MICC coordinates the development and implementation of application software packages by individual institutions to be shared by all user institutions. Detailed cost accounting and resource utilization systems are being refined to aid the determination of cost effectiveness and system efficiency for the delivery of services. MICC's success will be measured by the extent to which costs of operations are recovered by charges to users and by the level of incorporation of MICC services into user academic programs. Figure 5 and 6 illustrate user participation in MICC in 1976.

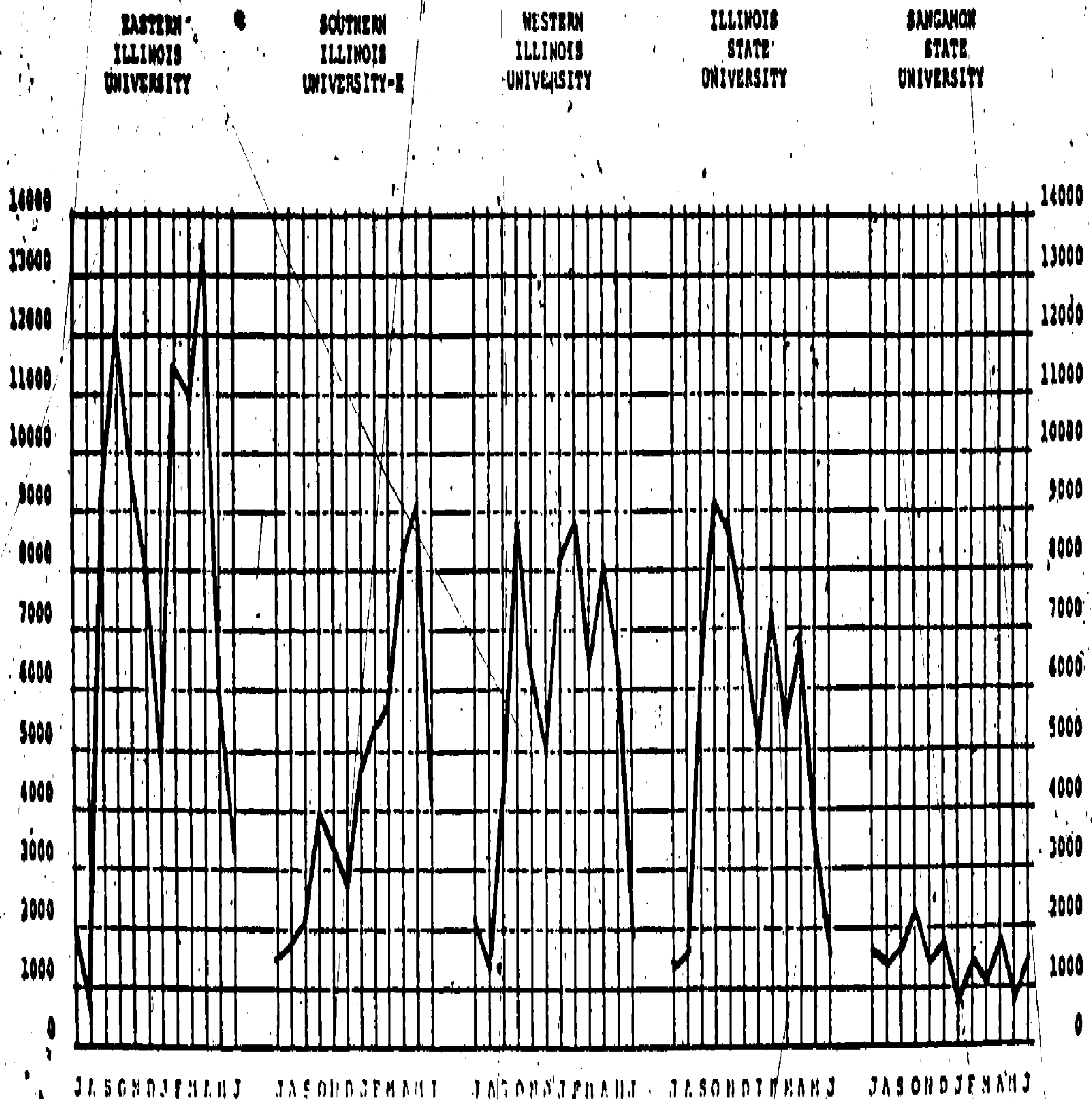


FIGURE 5. Number of Jobs Per Month, MICC 1975-1976

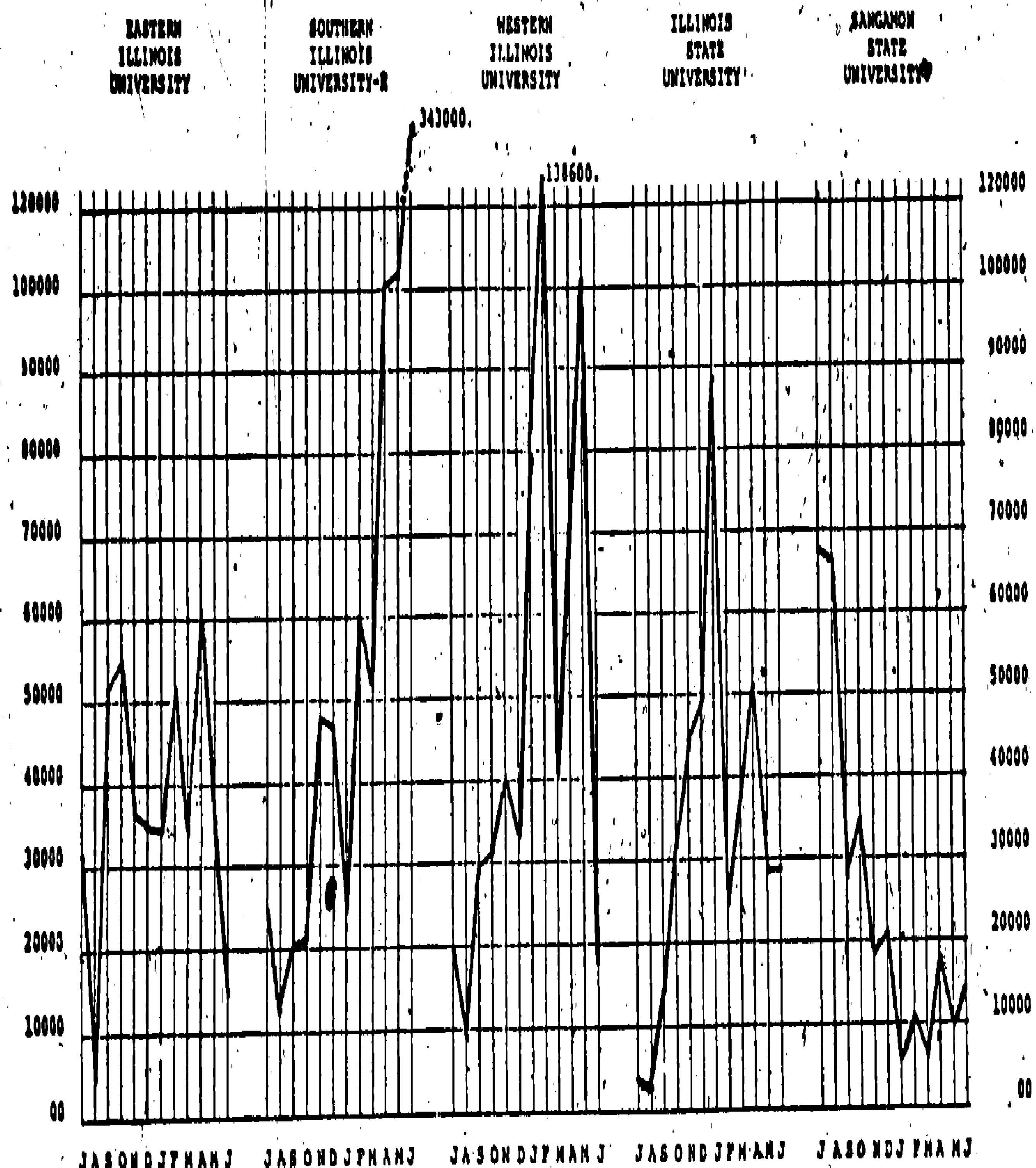


FIGURE 6. CPU Utilization in Seconds, MICC 1975-1976

Chapter 32

by George T. Williams

Multi-server Cooperation: The ARPANET Experience

1.0 THE ARPANET AT MIT

The ARPANET (Advanced Research Projects Agency NETwork) is, by any measure, a success of the first order at MIT. This paper deals with the ARPANET at MIT as it pertains to the Information Processing Services (IPS) user community at MIT; rather than, the technical aspects of how the ARPANET works. Whereas the ARPANET has been in operation at MIT for more than five years, the author has only been involved with it for the past two years. He must admit to an adolescent infatuation with the ARPANET: he is as impressed today as he was two years ago each time he sits down at a terminal in Cambridge, Massachusetts and logs in to a computer in Cambridge, England or when he consults with a user from Lawrence Livermore Laboratory who is using MIT Multics over the ARPANET.

Until recently the Laboratory for Computer Science (LCS) at MIT, formerly Project MAC, acted as the primary interface for all users of the ARPANET at MIT. That responsibility now belongs to the User Services staff at the IPS. Responsibility was transferred to IPS for two reasons: first to allow the personnel at LCS to concentrate on research projects rather than the everyday problems faced by most users; and second, to offer more immediate service to ARPANET users who need help or who have questions about how to use the ARPANET. After all, the primary function of a user services staff is to know what is available where it is, and how it should be used.

The areas to be discussed in this paper are:

- *User Services*. what, if any, unique problems do ARPANET users raise;
- *Privacy of Information*. how private is the information users entrust to the ARPANET;
- *Financial Arrangements*. there is no such thing as a free lunch.

2.0 USER SERVICES

User Support is a top priority item with the ARPANET. There is a wealth of information available about the ARPANET, i.e., what software is available at the various sites, what the various protocols look like, who to talk with at any particular site when a problem arises, etc. However, there is a lack of specific documentation available about the use of particular sites on the ARPANET.

At the top of the user support structure is the Network Information Center (NIC), housed at Stanford University, which serves as coordinator for the collection and dissemination of information about the ARPANET. In support of the ARPANET user community, the NIC publishes, among other documents of a more technical nature,

- *THE ARPANET DIRECTORY*. a directory of users and sites on the ARPANET
- *THE ARPANET RESOURCE HANDBOOK*. a very complete compendium of the resources available to the ARPANET community.
- *NICNOTES*. informal newsnotes which are distributed on-line by the NIC to the Liaison and other interested personnel.

In addition to the publications supplied by the NIC, there are also online reference files maintained by the NIC which are accessible to all ARPANET users. The most useful of these online files is a version of the Resource Handbook available through the NIC/QUERY program. Complete instructions on how to reference this information is contained in the ARPANET Directory as well as in the Resource Handbook.

In addition, members of the Network Liaison Group (NLG) act as advisors to ARPANET users, and to the NIC, on technical matters. Every ARPANET site specifies one individual to act as that site's Liaison Officer and, as such, to accept responsibility to keep the NIC updated on the hardware and software available at the site as well as personnel changes that affect the ARPANET. The Liaison is also expected to keep local users informed of network-related matters. Since Liaisons are the chief source of information about the ARPANET, they represent a very valuable network resource.

Obviously, the NIC has done its best to see to it that the users of the ARPANET are provided with the best support possible. The quality of support offered at any particular site depends upon the Liaison and management at that site.

MIT offers exceptional user support to all of its users, and ARPANET users are no exception. The Multics system has a variety of online user aids: a complete online facility for the retrieval of all source code for any non-proprietary program installed on the system; an online HELP facility, which describes the over 400 commands and subroutines available on Multics; and an online consultant. The online consultant is a member of the IPS technical staff who is available to answer questions for at least six (6) hours a day. Usually, someone is signed on as consultant 12 to 15 hours per day.

The online consultant is prepared to answer questions from novice users as well as the most sophisticated user of the system. In addition to the online consulting facility, there is an online trouble reporting system for reporting bugs found in programs installed on the system. System bugs are usually fixed or new code supplied for compiler problems within 24 to 48 hours. There is also an online graffitti file available for the airing of gripes, general inquiries about software that might be available in a private library, etc.

Network users do not pose any unique problems to User Services at MIT. The network users' question are generally no different from those posed by any other user. However, there are some areas that can cause problems. Althougt user services personnel can tell network users where to find something on the ARPANET, they cannot usually give instructions on how to use the particular package or how to use the host operating system. With so many operating systems available on the ARPANET, it is about impossible for one user consultant to be familiar with all of them; never mind the specific intricacies at any particular site. Figures 1, lists the sites on the ARPANET as of July 1976. Information on each site computer system is often available at the NIC or through the site Liaison.

A user problem particular to MIT is location of the TIP (Terminal Interface Processor) quite a distance from the computer center. Because of the distance, Liaison access to the TIP log of users who have been denied access is rather inconvenient.

MANUFACTURER COMPUTER	OP-SYST	HOSTNAME	HOSTADDR (Dec)	STATUS
BOLT BERANEK AND NEWMAN, INC				
PLURIBUS•	PLI	MOFFETT-PLURIBUS	109	Data encrypter
PLURIBUS•	PLI	NUC-SECURE	67	Data encrypter
PLURIBUS•	–	SDAC-CCP	39	Seismic data processor
BURROUGHS CORP				
B-4700	MCPV	GUNTER-ELF	13	Principal, USER
B-6700	–	I4-TENEX	15	Peripheral processor
ILLIAC-IV	–	I4-TENEX	15	Principal, SERVER
		I4B-ELF	79	
		I4B-TENEX	143	
CONTROL DATA CORP				
CDC-3200•	–	FNWC	33	Communications controller
CDC-6400•	SESAME	LBL	34	Communications controller
CDC-6500	SCOPE	FNWC	33	Principal, USER
CDC-6600	–	BNL	122	Principal, USER, Up ?
CDC-6600•	SESAME	LBL	34	Principal, SERVER
CDC-6600	SCOPE	EGLIN	53	Principal, USER, Up Winter 76
CDC-7600	–	BNL	122	Principal, USER, Up ?
CDC-7600	BKY	LBL	34	Principal, SERVER
CULLER/HARRISON, INC				
MP32•	–	CHII	182	Communications controller
MP32A	–	CHII	182	Peripheral controller
AP90	–	CHII	182	Speech processor
DATA GENERAL CORP				
DCU-50•	–	ROCHESTER	82	Communication controller
ECLIPSE	–	ROCHESTER	82	Principal, USER
NOVA 800•	–	PARC-MAXC	32	Communication controller
NOVA 800•	–	PARC-GATEWAY	96	internet protocol research
NOVA 1200	–	SRI-IA11	115	Principal, USER
DIGITAL EQUIPMENT CORP				
PDP-1•	–	BBN-1D	232	NCC, testing and diagnostics
PDP-6	–	MIT-AI	134	Peripheral processor
PDP-6	–	SU-A	11	Peripheral processor
PDP-9•	–	LONDON	42	Communications controller
PDP-10	KA DEC 10/50	CMU-10A	78	Principal, Limited SERVER
PDP-10	KA DEC 10/50	CMU-10B	14	Principal, Limited SERVER
PDP-10•	KA DEC 10/50	HARV-10	9	Principal, Limited SERVER
H-716	–	NUC-SECURE	67	–
INTERNATIONAL BUSINESS MACHINES				
360/40•	DOS/360	NORSAR-40A	41	Principal, USER,
360/40•	DOS/360	NORSAR-40B	233	Principal, USER, Up Winter 76
360/40•	DOS/360	SDAC-DP	103	Principal, USER
360/40•	DOS/360	SDAC-NEP	167	Principal, USER
360/44•	DOS/360 USC-PS	SDAC-44	230	Principal, Limited SERVER

• = Host Computer (i.e., a computer attached to an IMP)

FIGURE 1. Computer Systems on ARPANET, July 1976

360/44•	USC-PS	USC-44	23	Principal, Limited SERVER
360/67•	TSS/360	AMES-67	16	Principal, SERVER
360/91•	OS-MVT	UCLA-CCN	65	Principal, SERVER
360/195	OS-MVT	LONDON	42	Principal, SERVER
370/158•	OS-MVT	RAND-RCC	7	Principal, SERVER
370/168•	VM-370	LL	10	Principal, SERVER
370/195•	OS-MVT	ANL	55	Principal, SERVER
KONGSBERG VAAPERFABRIKK				
SM-3	–	NORE	105	Principal, USER
SIGNAL PROCESSING SYSTEMS				
SPS-41	DOS	BBN-SPEECH-11	50	Speech Processor
SPS-41•	DOS	ISI-SPEECH11	22	Speech Processor
SPS-41	DOS	LL-11	202	Speech Processor
SPS-41	DOS	SRI-NSC11	51	Speech Processor
SPS-41	–	SU-AI	11	Speech Processor
UNIVAC				
1108	–	NUC-SECURE	67	Principal, USER Up Winter 76
VARIAN ASSOCIATES				
73•	–	ANL	55	Communications controller
XEROX CORP.				
MAXC	TENEX	PARC-MAXC	32	Principal, Limited SERVER
PDP-10• KI	DEC 10/50	NBS-10	19	Principal, SERVER
PDP-10• KI	DEC 10/50	RUTGERS-10	46	Principal, SERVER
PDP-10• KA	DEC 10/50	SU-AI	11	Principal, Limited SERVER
PDP-10• KL	DEC 10/80	SU-AI	11	Principal, Limited SERVER
PDP-10 KL	DEC 10/80	DEC	137	Principal, USER Up Fall 76
PDP-10 KL	DEC 20/40	DEC	137	Principal, USER Up Fall 76
PDP-10• KA	ITS	MIT-AI	134	Principal, Limited SERVER
PDP-10• KA	ITS	MIT-DMS	70	Principal, SERVER
PDP-10• KL	ITS	MIT-MC	236	Principal, Limited SERVER
PDP-10• KA	ITS	MIT-ML	198	Principal, Limited SERVER
PDP-10• KA	TENEX	BBN-TENEX	241	Principal, SERVER
PDP-10• KA	TENEX	BBN-TENEXA	197	Principal, Limited SERVER
PDP-10• KA	TENEX	BBN-TENEXB	49	Principal, SERVER
PDP-10• KA	TENEX	BBN-TENEXD	113	Principal, Limited SERVER
PDP-10• KA	TENEX	BBN-TENEXE	5	Principal, Limited SERVER
PDP-10• KA	TENEX	CCA-TENEX	31	Principal, SERVER
PDP-10• KI	TENEX	I4-TENEX	15	Principal, SERVER
PDP-10 KI	TENEX	I4B-TENEX	143	Principal, Limited SERVER
PDP-10• KA	TENEX	OFFICE-1	43	Principal, SERVER
PDP-10 KA	TENEX	MOFFETT-ARC	45	Principal, Limited SERVER
PDP-10 KA	TENEX	SRI-AI	66	Principal, Limited SERVER
PDP-10• KI	TENEX	SUMEX-AIM	56	Principal, SERVER Dual processors
PDP-10• KI	TENEX	USC-ECL	215	Principal, SERVER
PDP-10• KA	TENEX	USC-ISI	86	Principal, SERVER
PDP-10• KI	TENEX	USC-ISIB	244	Principal, Limited SERVER
PDP-10• KA	TENEX	USC-ISIC	150	Principal, SERVER
PDP-10• KA	TENEX	USC-ISID	214	Principal, SERVER
PDP-11•		HAWAII-ALOHA	100	Communications controller Up Winter 76
PDP-11•	–	I4B-TENEX	143	Communications controller
PDP-11•	ELF	NCSL	117	Terminal concentrator
PDP-11•	ELF	NSA	57	Terminal concentrator

FIGURE 1. (contd.)

PDP-11•	ELF	SCRL-ELFDEVEL	54	ELF R & D
PDP-11s•	C.mmp Hydra	CMU-11	206	Parallel processing research, Up Winter 76
PDP-11•	ELF	J4-ELF	79	Experimental machine
PDP-11•	UNIX	ARPA	92	Principal, User Up Winter 76
PDP-11/10•	ELF	AMES-11	208	I/O controller
PDP-11/10•	ELF	ASL	112	Magtape mass data dump
PDP-11/10•	ELF	SRI-CBC11	130	—
PDP-11/20•	ANTS	ILL-CAC	12	Terminal concentrator
PDP-11/20•	ELF	SU-DSL	120	Internet protocol research
PDP-11/35•	ELF	GUNTER-ELF	13	Communications controller
PDP-11/40•		ARPA-XGP11	220	Peripheral processor
PDP-11/40•	ANTS, ELF	BELVOIR	27	Terminal concentrator
PDP-11/40•	ANTS ELF	BRL	29	Terminal concentrator
PDP-11/40•	ELF	ARC-RD	2	Front-end research for NSW
PDP-11/40•	ELF	NELC-ELF	131	I/O controller
PDP-11/40•	ELF	SRI-PKT11	243	Packet radio research
PDP-11/40•	ELF	BBN-SPEECH-11	50	Speech compression
PDP-11/40•	SP	CCA-SIP	223	Seismic input processor
PDP-11/40•	ELF	ISI-XGP11	52	Peripheral processor
PDP-11/40•	ELF	NYU	58	Communications controller
PDP-11/40•	ELF UCLA-VMM,UNIX	OT-ITS	89	Terminal concentrator. Simulation and standards
PDP-11/40•	ELF	SRI-ARC	179	Terminal concentrator
PDP-11/40•	ELF	SRI-NSC11	51	Speech compression
PDP-11/40•	RSX-11M RSX11-M,UNIX	SRI-IA11	115	Speech and graphics protocol research
PDP-11/40•	ELF,DOS, RSX11-M,UNIX	UCB	98	Speech and graphics protocol research
PDP-11/40•	TOPS 20AN	DEC	137	Communication controller
PDP-11/45•		SU-AI	11	Terminal concentrator
PDP-11/45•	ANTS.ELF UCLA-VMM,UNIX	UCLA-ATS	1	Network analysis and modeling
PDP-11/45•	ELF	AMES-11	208	Communication controller
PDP-11/45•	ELF	BNL	122	Provides access to graphics devices.
PDP-11/45•	ELF	ETAC	59	Terminal concentrator
PDP-11/45•	EPOS	ISI-SPEECH11	22	Speech compression
PDP-11/45•	ELF or RSX-11M	LL-11	202	Speech compression
PDP-11/45•	ELF	PURDUE	38	Terminal concentrator
PDP-11/45•	ELF	SCRL-ELF	118	Speech understanding and frontend research
PDP-11/45•	RSX-11D	HASKINS	201	Speech understanding
PDP-11/45•	RSX-11M	EGLIN	53	Communication controller
PDP-11/45•	UNIX	UCLA-SECURITY	129	Principal, User
PDP-11/50•	ELF	LL-ASG	108	Terminal concentrator
PDP-11/50•	RSX-11M	MCA	110	Communication controller, Front end research
PDP-11/50•	UNIX	ILL-NTS	76	Terminal concentrator, Frontend research
PDP-11/70•	UNIX	RAND-UNIX	71	Principal, SERVER
PDP-15•	DMS	ARPA-DMS	28	Principal, Limited SERVER
PDP-15	—	SRI-AI	66	Peripheral controller
HONEYWELL INC				
H-316•		BBN-NCC	40	NCC prototype
H-6180•	MULTICS	MIT-MULTICS	6	Principal, SERVER
H-6180•	MULTICS	RADC-MULTICS	18	Principal, SERVER
H-68/80•	MULTICS	MIT-DEVMULTICS	44	Principal, Limited SERVER

FIGURE 1. (contd.)

IDENT.	GROUP NAME	COORDINATOR
ASG	ANTS Support Group	Kelley, K C (KELLEY@ILL-NTS)
CBIG	Computer Based Instruction Group	O'Sullivan, T (O SULLIVAN@SI)
CAHCOM	Computer Assisted Human Communication	Farber, D J (FARBER@SI)
ERDA	Energy Research and Development Agency	Estrin, G (Estrin@MIT-Multics)
EWG	Elf Followers	Retz, D L (RETZ@SRI-AI)
IIG	Imlac Interest Group	Belleville, R L (BELLEVILLE@SIC)
INWG	International Network Working Group	Barber, D (KIRSTEIN@SI)
INWG1	International Packet Network Working Group	Barber, D (KIRSTEIN@SI)
INWG2	International Packet Network Working Group	Barber, D (KIRSTEIN@SI)
KWAC	Knowledge Workshop Architect Community	Norton, J C (NORTON@OFFICE-1)
NALCON	Navy Laboratory Computer Network	Brignol, F (BRIGNOL@BBNB)
NET-SPONSORS	Network Sponsors Group	Czahor, R (CCACODE535@SI)
NGG	Network Graphics Group	Michener, J C (JCM@MIT-DMS)
NLG	Network Liaison Group	Feinler, E J (FEINLER@OFFICE-1)
NMG	Network Measurement Group	Naylor, W E (NAYLOR@SIC)
NSC	Network Speech Compression Group	Kahn, R E (KAHN@SIC)
NSW	National Software Works Group	Postel, J B (POSTEL@SIC)
NWG	Network Working Group	Postel, J B (POSTEL@SIC)
PRC	Packet Radio Group	Frank, H (NAC@SID)
P11	P11 TENEX Group	Boynton, [illegible] (BOYNTON@SR)
SURG	Speech Understanding Research Group	Wolf, J J (WOLF@BBND)
TIPUG	TIP Users Group	Walden, D C (WALDEN@BBNE)
TRG	TENEX Research Group	Burchfiel, J D (BURCHFIEL@BBN)
UNIX	UNIX Interest Group	Holmgren, S F (STEVE@ILL-NTS)
WDBWG	Weather Data Base Working Group	Underwood, R (UNDERWOOD@4-TENEX)
XIC3	XIC3	Kahn, R E (KAHN@SIC)

NOTE The names given above are not considered to be a complete listing of all special interest groups on the Arpanet

FIGURE 1. (contd.)

Support of the network software is excellent at MIT. The biggest problem facing User Services at MIT, and probably everywhere else on the net, is the lack of clear documentation and general support from the various network sites. User support is not as complete at all ARPANET sites as it is at MIT. The solution to this problem lies with the local Liaisons and local users. Since a local organization is apt to be more responsive to the needs of local users, there should be considerably more interaction between Liaisons than there is at present. The ARPANET user community should put more pressure on their local Liaisons to ensure that their support needs are being met.

3.0 SECURITY OF INFORMATION ON THE ARPANET

How secure is the information that a user entrusts to the ARPANET? Security often depends upon the host. The communications subnet will ensure that messages are delivered to the proper host, and the non-deterministic routing of the individual packets provides some degree of security to the subnet. At that point it is the responsibility of the host to insure delivery to the proper user. File security is the exclusive responsibility of the individual hosts.

MULTICS is probably the most secure computer system on the network. In MULTICS all information is owned exclusively by its author, and the author must determine whether anyone else can have access to that information. Access to information can be gained only when the author explicitly allows it. Besides providing a security system that limits access to information stored on it, MULTICS offers a complete backup and retrieval system to its users. Data on MULTICS is automatically backed up on a regular hourly, daily and weekly basis. Any information that has resided on the system for three consecutive hours will probably make it to one of the backup tapes and be available for retrieval for at least three (3) months and for as long as (12) months.

4.0 FINANCIAL ARRANGEMENTS REQUIRED FOR ARPANET PARTICIPATION

Membership in the ARPANET requires an initial investment of $40,000 to $100,000 with annual charges in the area of $20,000 to $35,000 for maintenance and additional ARPANET privledges. However, that rather high fee gives an institution and its user community access to a wide variety of unique resources: the ILLIAC IV for large amounts of matrix arithmetic and array processing; the interactive graphics at the University of Utah; the MULTICS system; the MACSYMA system at the Articifical Intelligence Laboratory at MIT, etc. A need any of these unique resources can provide as significant amount of justification for ARPA membership.

The ARPANET has been a success at MIT; a number of users come in to use MIT facilities, and a number of users go out to use other facilities on the net. As users come and go from around the world on a daily basis, use is steadily growing.

APPENDIX A
Conference Participants

Dharam Ahuja
Marketing Supervisor
A T & T
1776 On-the-Green
Morristown, NJ

Robert M. Aiken
Computer Science Dept.
The University of Tennessee
Knoxville, TN 37916

Paul Aizley
Assistant to the President
University of Nevada, Las Vegas
4505 Maryland Parkway
Las Vegas, NV 89154

Peter A. Alsberg
Associate Director
Center for Advanced Computation
University of Illinois
Urbana, IL 61801

Neils H. Anderson
Director
University of Nevada
System Computing Center
P. O. Box 9068
Reno, NV 89507

Pauline Atherton
Professor
Syracuse University
School of Information Studies
113 Euclid Avenue
Syracuse, NY 13210

Donald M. Austin
Computer Science & Applied Mathematics Dept.
Lawrence-Berkeley Laboratory
Berkeley, CA 94720

Jack Austin
Consulting Group
EDUCOM

George Beatty
Associate Director Budgeting & Trust Studies
University of Massachusetts
Amherst, MA 01002

Warren Belding
Program Librarian
Dartmouth College
Hanover, NH 03755

Peter M. Bellinger
Research Foundation—SUNY
P. O. Box 9
Albany, NY 12201

Dan Bernard
Senior Consultant
EDUCOM

Howard L. Bleich
Associate Professor of Medicine
Beth Israel Hospital
330 Brookline Ave.
Boston, MA 02215

Ronald W. Brady
Vice President for Administration
University of Illinois
349 Administration Bldg.
Urbana, IL 61801

Robin Braithwaite
Assistant Director for Network Services
UTLAS
130 St. George Street
Robarts Library
University of Toronto
Toronto, Ontario, Canada M5S 2T4

Lt. Comdr. John Brandon
U.S. Merchant Marine Academy
Kings Point, NY 11024

Anne W. Branscomb
Vice President
Kalba Bowen Associates, Inc.
Hidden Oak Lane
Armonk, NY 10504

David C. Brown
Assistant to the Director
Madison Academic Computing Center
University of Wisconsin at Madison
1210 West Dayton Street
Madison, WI 53706

Weston J. Burner
Director, Information Program Service
M.I.T. Rm 39-565
77 Massachusetts Ave.
Cambridge, MA 02139

Tom Byrne
Kiewit Computation Center
Dartmouth College
Hanover, NH 03755

Jack Cain
Robarts Library
University of Toronto
Toronto, Ontario, Canada M5S 2T4

Elwin F. Cammack
Associate Vice President
The University of Wisconsin System
1530 Van Hise Hall
Madison, WI 53706

Frank Campanella
Executive Vice President
Boston College
Chestnut Hill, MA 02167

Elizabeth J. Campbell
Manager
Lab. on Nuclear Science Computing Facility
M.I.T. Rm 24-017
Cambridge, MA 02146

George A. Carroll
Director of Management Affairs/University Library
Alexander Library
Rutgers University
New Brunswick, NJ 08903

Kelvin C. Carstairs
Hematologist
Toronto General Hospital
Toronto, Ontario, Canada

Richard Carter
Director Computation Center
Northeastern University
Boston, MA 02115

Robert R. Caster
Assistant Vice President for Business & Administration Service
Director of SWORCC
3333 Vine Street, Rm 507
University of Cincinnati
Cincinnati, OH 45220

Gilles Chaput
Coordonnateur des services informatises-Service des Bibliotheques
University de Montreal
C.P. 6128 Succursale "A"
Montreal, Quebec H3C 3J7

Henry Chauncey
Trustee-at-Large
EDUCOM
R.D. 1
Ringoes, NJ 08551

Guy Ciannavei
Manager
Harvard Computing Center
1730 Cambridge Street
Cambridge, MA 02138

Hugh F. Cline
Senior Research Sociologist
E.T.S.
Rosedale Road
Princeton, NJ 08540

James B. Conklin, Jr.
Director
CIRCA, Rm 411 Weil
University of Florida
Gainesville, FL 32611

Ruth Corrigan
Director of University Libraries
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

George Cowan
Data Base Manager
University of Quebec
Postal Code 8888
Montreal, Canada H3C 1T8

Frances Bardello Craig
Network Services Consultant
Carnegie-Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Carlos A. Cuadra
General Manager
System Development Corporation
Search Service
2500 Colorado Avenue
Santa Monica, CA 90406
Mail Stop 5142

Effa Dalton
EDUCOM

Pierre Deslauriers
Project Officer for University Libraries
Conf. of Rectors & Principal of
Quebec University
2 Complexe Desjardins, Suite 181F
Montreal, Quebec, Canada H5B 1B3

Len D'Innocenzo
Data Dimensions, Inc.
51 Weaver Street
Greenwich, CT 06830

Howard L. Dockery
Director of Computer Center
University of Iowa
University Computer Center
Iowa City, IA 52242

James T. Dodson
Director of Libraries
University of Texas at Dallas
P. O. Box 688
Richardson, TX 75080

Ely E. Duncan
Coordinator CBIS
Office of Communications Program
University of Pittsburgh
Pittsburgh, PA 15260

Anne Ekstrom
Director Library Sup. Division
The Ohio College Library
1125 Kinnear Road
Columbus, OH 43212

George Ember
Chief Librarian Service
Canada Institute for Scientific
and Technical Information
National Research Council of Canada
Ottawa, Canada K1A 0S2

James C. Emery
President
EDUCOM

A. H. Epstein
Director
BALLOTS Center
SCIP-Willow
Stanford University
Stanford, CA 94305

David J. Farrell
Manager of Applications Development
Harvard Office for Information Technology
1730 Cambridge Street
Cambridge, MA 02138

Barbara B. Farquhar
Director, Comp-based Medical Education
Laboratory of Computer Science
Massachusetts General Hospital &
Harvard Medical School
Boston, MA 02114

Joseph Fennel
Sr. Staff Scientist
Office of the Provost
George Washington University
2121 Eye Street, Rice Hall
Washington, DC 20052

Russell W. Fenske
Assistant Vice Chancellor
University of Wisconsin-Milwaukee
Chapman Hall No. 310
Milwaukee, WI 53201

Brenda Ferriero
Computer Services Coordinator
M.I.T. Rm 39-575
Cambridge, MA 02139

Minerva S. Figgs
AIDP Coordinator
Kentucky State University
Frankfort, KY 40601

Mike Ford
Senior System Engineer
Telenet
10 Dale Street
Waltham, MA 02154

Raymond Frankle
Director of Library Services
Stockton State College
Pomona, NJ 08239

Greydon C. Freeman
Director
Yale University Computer Center
175 Whitney Avenue
New Haven, CT 06520

Harland Garvin
University of Iowa
128D Lindquist Center
Iowa City, IA 52242

Robert Gibbs
Executive Director
NERComP, Inc.
40 Grove Street
Wellesley, MA 02181

Bruce Gilchrist
Director of Computing Activities
Columbia University
612 W. 115th Street
New York, NY 10025

Don Haback
Data Dimensions, Inc.
51 Weaver Street
Greenwich, CT 06830

Dale J. Hall
Director of Computer Systems
Development
Indiana University Computer Network
Information & Computer Service
Bryan 07
Bloomington, IN 47401

Lowell H. Hattery
Information Retrieval of Library Automation
P. O. Box 56
Mt. Airy, MD 21771

John E. Haugo
Executive Director
Minnesota Educational Computing
Consortium
2520 Broadway Drive
Lauderdale, MN 55113

Peter J. Heffernan
Manager of Systems and Users Services
Harvard Office for Information
Technology
1730 Cambridge Street
Cambridge, MA 02138

Paul S. Heller
Deputy Executive Director
Planning Council
EDUCOM

Robert Hernandez
Harvard Law School
Cambridge, MA 02138

Richard R. Hughes
Coordinator of Computing Activities
University of Wisconsin-Madison
1500 Johnson Drive
Madison, WI 53706

Gerald Hutchison
Associate Dean
School of Natural Sciences &
Mathematics
University of Alabama in Birmingham
University Station
Birmingham, AL 35294

Jeremy Johnson
Director
Computing & Data Processing Service
University of Maine
Computing Center
Orono, Maine 04473

Ronald Jonas
Director of Planning
Youngstown State University
410 Wick Ave.
Youngstown, OH 44555

Walter J. Kanavy
Director of Data Processing
University of Scranton
Scranton, PA 18510

H. Eugene Kessler
Executive Director
EDUCOM

Michael P. Koda
Data Dimensions, Inc.
51 Weaver Street
Greenwich, CT 06830

Casimir Kulikowski
Associate Professor of Computer Science
Rutgers, Hill Hall-Busch Campus
Department of Computer Science
New Brunswick, NJ 08903

Tom Kurtz
Director
Office of Academic Computing
Dartmouth College
Hanover, NH 03755

Carolyn P. Landis
Corporation Secretary
EDUCOM

Arlene Larsen
Assistant Director
Dartmouth College
Kiewit Computation Center
Hanover, NH 03755

William H. Latimer
Director
Wharton Computer Center
University of Pennsylvania
210 Vance Hall/CS
Philadelphia, PA 19174

J.C.R. Licklider
Professor of Computer Science
Laboratory for Computer Science
M.I.T. Rm NE 43-218A
Cambridge, MA 02139

Arno Liivak
Law Librarian & Professor
Rutgers University
Camden, NJ 08102

Lawrence G. Livingston
Program Officer
1 Dupont Circle, Suite 620
Council on Library Resources
Washington, DC 20036

Peter Lykos
Department of Chemistry
Illinois Institute of Technology
Chicago, IL 60616

Michele Lynch
The Ohio College Library Center
1129 Kinnear Road
Columbus, OH 43212

Joseph J. McConkey
Director of Management Systems
State University of New York
Stony Brook, NY 11794

John W. McCredie, Jr.
Vice-Provost for Information Services
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Henry S. McDonald
Assistant Director
Electronics & Computer Systems
Research Laboratory
Bell Telephone Laboratory
Murray Hill, NJ 07974

John S. McGeachie
Director of Computer Service
Kiewit Computation Center
Dartmouth College
Hanover, NH 03755

Fran McKee
Staff Analyst
Harvard University
Cambridge, MA 02138

Colin McKirdy
Harvard University Library
Harvard University
Cambridge, MA 02138

Michael Maione
NERComP, Inc.
40 Grove Street
Wellesley, MA 02181

Clair G. Maple
Director, Computation Center
Iowa State University
Ames, IA 50011

Dillon E. Mapother
Associate Vice Chancellor for Research
University of Illinois at Urbana-Champaign
101 Engineering Hall
Urbana, IL 61801

George J. Maslach
Provost
Professional Schools & Colleges
University of California
Chancellor's Office
200 California Hall
Berkeley, CA 94720

Walter Matherly
Systems Design Consultant
Office of Vice President for Administration
Harvard University
361 Holyoke Center
1350 Massachusetts Ave.
Cambridge, MA 02138

Stuart Mathison
Vice President
Telenet
1050 17th Street
Washington, DC 20036

Peter P. M. Meincke
Physics Department
University of Toronto
Toronto, Ontario, Canada M5S 1A7

John P. Menard
Director
Computer Science Center
University of Maryland
College Park, MD 20742

Rudolph Menna
Vice President
Data Dimensions, Inc.
51 Weaver Street
Greenwich, CT 06830

George Meyfarth
Director Computer Services
Tufts University
Medford, MA 02155

Arthur Miller
Faculty office Building
Harvard Law School
Cambridge, MA 02138

James G. Miller
President
University of Louisville
Louisville, KY 40208

Robert C. Miller
Director of Libraries
University of Missouri—St. Louis
8001 Natural Bridge Road
St. Louis, MO 63121

Ronald F. Miller
Director
NELINET
New England Board of Higher Education
40 Grove Street
Wellesley, MA 02181

Leslie Mills
Yale University
New Haven, CT 06520

Ronald Mills
Yale University
New Haven, CT 06520

David Mintzer
Vice President for Research & Dean of Science
Northwestern University
633 Clark Street
Evanston, IL 60201

Hannah Mitchell
Programmer/Systems Analyst
Council of Ontario Universities
130 St. George Street, Suite 8039
Toronto, Ontario, Canada M5S 2T4

Kathie L. Morris
Assistant to the Director
Wharton Computer Center
University of Pennsylvania
210 Vance Hall/CS
Philadelphia, PA 19174

J. Charles Morrow
Provost
University of North Carolina
at Chapel Hill
Chapel Hill, NC 27514

John M. Nevison
Consultant
3 Spruce Street
Boston, MA 02108

Jerry Newhall
NERComP, Inc.
40 Grove Street
Wellesley, MA 02181

William L. Newman
Head, Data Processing Research & Planning Br.
National Library of Canada
395 Wellington Street
Ottawa, Ontario, Canada K1A 0N4

Gordon L. Nordby
Associate Professor of Biochemistry
University of Michigan
M5434 Medical Science I
Ann Arbor, MI 48109

Comdr. D. J. O'Connell
Systems Management Officer
Deputy for Management
U. S. Naval Academy
Annapolis, MD 21402

Dennis R. Ojakangas
Director of Computing
University of California—Davis Campus
Davis, CA 95616

Charles Oleson
Research Associate
University of Pittsburgh
1360 Scaife Hall
Pittsburgh, PA 15261

Beverly O'Neal
System Analyst
EDUCOM

Ed O'Neill
Yale University
New Haven, CT 06520

Steven Pauker
Asst. Prof. of Medicine (Cardiology)
Tufts—New England Medical Center
171 Harrison Street
Boston, MA 02111

Ruann Pengov
Director
Division of Computer Service for Medical Education & Research
Ohio State University College of Medicine
076 Health Science Library
376 West 10th Ave.
Columbus, OH 43210

William J. Pervin
Director
The University of Texas Regional Computing Center
5601 Medical Center Drive
Dallas, TX 75235

James F. Poage
Director of University Computing
Princeton University
87 Prospect Street
Princeton, NJ 08540

Edward L. Pocino
MIS Project Manager
Rutgers University, CCMS
Administration Services Building
New Brunswick, NJ 08903

Terry Prickett
General Terminal Systems
13777 North Central Expressway
Suite 416
Dallas, TX 75234

Charles B. Ray
Director of Booth Computing Center
California Institute of Technology
1201 E. California Blvd. 158-79
Pasadena, CA 91125

Pamela Reekes
Systems Librarian
Harvard University Library
Cambridge, MA 02138

Roy F. Reeves
Director, I & R Computer Center
Ohio State University
1971 Neil Ave.
Columbus, OH 43210

Martin D. Robbins
Vice President
EDUCOM

Lawrence G. Roberts
President
Telenet Communications Corporation
1050 17th Street, NW.
Washington, DC 20036

George B. Rockwell
Manager, Financial Industries
Arthur D. Little, Inc.
35 Acorn Park
Cambridge, MA 02140

Peter G. Roll
Special Assistant to the President
for Academic Affairs
University of Minnesota
217 Morrill Hall
Minneapolis, MN 55455

Walter A. Rosenblith
Provost
M.I.T. Rm 3-240
Cambridge, MA 02139

Joseph E. Rowe
Provost
Case Inst. of Technology
Case Western Reserve University
Cleveland, OH 44106

Lillian Rubinlicht
Senior Cataloguing Librarian
Concordia University Libraries
1455 de Maisonneuve Blvd. West
Montreal, Quebec, Canada H3G 1M8

John W. Rudan
Director, Office of Computer Services
Cornell University
209 Langmuir Lab, Cornell Research Pk.
Ithaca, NY 14850

Robb N. Russell
Director of Computing Services
Bryn Mawr College
Computing Center
Bryn Mawr, PA 19010

Jean Scarborough
EDUCOM

Roger R. Schell
Hq. ESD (MCIT)
U.S. Air Force Data System Center
Pentagon
Washington, DC 20330

Robert L. Schiffman
Professor of Civil Engineering
Associate Director for Research
Computer Center Rm 111
University of Colorado
Boulder, CO 80309

Robert H. Scott
Director
Office for Information Technology
Harvard University
Cambridge, MA 02138

Ronald Segal
Manager, Network Simulation Project
EDUCOM

Warren D. Seider
Associate Professor of Chemical
Engineering
Dept. of Chem. & Bio. Eng.
Rm 311-A
University of Pennsylvania
Philadelphia, PA 19174

Allen Seltzer
Senior System Engineer
Telenet
10 Dale Street
Waltham, MA 02154

Edward Sharp
Director Computer Center
University of Utah
3110 Merrill Engineering Bldg.
Salt Lake City, UT 84112

Patricia M. Sheehan
System Designer
M.I.T. Rm 14S-320
77 Massachusetts Ave.
Cambridge, MA 02139

Michael Sher
President
Illinois Educational Consortium
1 North Old State Capital Plaza
Suite 520
Springfield, IL 62701

John Shinners
Research Foundation–SUNY
P. O. Box 9
Albany, NY 12301

Chuck Shomper
Director of Regional Computer Center
University of Iowa
128A Lindquist Center
Iowa City, IA 52242

Edward Shortliffe
Intern in Medicine
Massachusetts General Hospital
Fruit Street
Boston, MA 20114

Donald R. Shurtleff
Director, Office of Computing Activities
830 Clark Hall
University of Missouri
Columbia, MO 65201

Patricia C. Skarulis
Director
Administrative Systems &
Data Processing
Princeton University
P. O. Box 352
Princeton, NJ 08540

Warner V. Slack
Associate Professor of Medicine
Beth Israel Hospital
330 Brookline Ave.
Boston, MA 02215

Robert L. Smith, Jr.
Assistant Professor of Computer Science
Rutgers University
Busch Campus
New Brunswick, NJ 08930

Steven R. Smith
Assistant Dean & Professor of Law
University of Louisville School of Law
Belknap Campus
Louisville, KY 40208

Martin B. Solomon
Director
University of Kentucky
Computing Center
Lexington, KY 40506

Roland D. Spaniol
Director
Computer Services
Eastern Illinois University
Charleston, IL 61920

Kevin Sperl
Analyst-Programmer
University of New Hampshire
Department of Computer Services
Kingsbury Hall
Durham, NH 03824

Patricia Stadel
Director of Computer Services
San Jose State University
San Jose, CA 95192

Richard J. Stanton
Manager of Administration
Harvard Office for Information Technology
1730 Cambridge Street
Cambridge, MA 02138

Joseph R. Steinberg
Assistant Director of Information
Program Service
M.I.T.
Cambridge, MA 02139

Charles W. Stevenson
Acting Manager of Computer
University of California
Systemwide Administration
Berkeley, CA 94720

Ralph Stierwalt
Director
Office of Library Coordination
Council of Ontario Universities
130 St. George Street Suite B039
Toronto, Ontario, Canada M5S 2T4

Richard Storer
Director—CCFS
Rutgers University
New Brunswick, NJ 08903

John C. Storlie
Director, Computing Center
University of Wisconsin—La Crosse
1725 State Street
La Crosse, WI 54601

George Sullivan
Manager
Library Computer Center
Rutgers University
Alexander Library
New Brunswick, NJ 07730

Sharon A. Sullivan
Systems Librarian
Harvard University Library
Cambridge, MA 02138

Peter Szolovits
Assistant Professor of Electrical
Engineering & Computer Science
M.I.T.
Cambridge, MA 02139

Joe K. Taylor
Assistant Planning & Research
Librarian
Carol M. Newman Library
Virginia Polytechnic Institute &
State University
Blacksburg, VA 24061

David A. Thomas
Professor of Law
Brigham Young University
393A JRCB
Provo, Ut 84602

R. M. Thompson
Director
Library Services (ACL)
Department of External Affairs
125 Sussex Drive
Ottawa, Canada K1A OG2

Sharon Tostevin
Analyst-Programmer
University of New Hampshire
Department of Computer Services
Richards House
Durham, NH 03824

Larry E. Travis
Director
Academic Computing Center (MACC)
University of Wisconsin-Madison
1210 West Dayton Street
Madison, WI 53706

Martha Traylor
Associate Professor
Seton Hall Law School
1111 Raymond Blvd.
Newark, NJ 07102

Mark Van Baalen
Staff Analyst
Harvard University
Cambridge, MA 02138

Mary Vassiliou
EDUCOM

Harriet Velazquez
Robarts Library
University of Toronto
130 St. George Street
Toronto, Ontario, Canada M5S 2T4

Frank Verbrugge
Director, Computer Services
University of Minnesota
143 Space Science Center
Minneapolis, MN 55455

Robert Votaw
Associate Dean
for Pre-clinical Medical Education
School of Medicine
University of Connecticut
Health Center
Farmington, CT 06032

Umesh I. Vyas
Reference Librarian
Faculty of Law
University of Calgary
Calgary, Alberta, Canada T2N 1N4

Harold B. Wakefield
Assistant Vice Chancellor
State University of New York—
Central Administration
99 Washington Ave.
Albany, NY 12246

William E. Walden
Director, Systems and Computing
Washington Higher Education
Computing Consortium
Washington State University
Pullman, WA 99163

John E. Walker
Director, Computer Center
Lehigh University
Computing Center Rm. 154
Packard Laboratory Bldg. No. 19
Bethlehem, PA 18015

Frederick Way
Associate Director Department of
Computing & Information Sciences
Case Western Reserve University
Cleveland, OH 44106

Ben T. Williams
University of Illinois
1400 W. Park Ave.
Urbana, IL 61801

George T. Williams
Programs Analyst
Information Processing Services
M.I.T.
Cambridge, MA 02139

Frederic G. Withington
Senior Staff Member
Arthur D. Little, Inc.
35 Acorn Park
Cambridge, MA 02140

Thomas E. Wolf
Chairman
Math, Computer Science & Engineering
University of Pittsburgh at Greensburg
1150 Mt. Pleasant Road
Greensburg, PA 15601

Bardie Wolfe
Law Library
Cleveland State University
East 24th Street at Euclid
Cleveland, OH 44115

Kenneth E. Wortz
General Sales Manager
Data Dimension, Inc.
51 Weaver St.
Greenwich, CT 06830

Gordon Wright
Director, Planning & Budgeting
University of Toronto Library
University of Toronto
Toronto, Canada M5S 1A7

Joe B. Wyatt
Vice President for Administration
Harvard University
Massachusetts Hall
Cambridge, MA 02138

Eugene P. Young
Assistant Vice President for
Academic Affairs
Rutgers University
New Brunswick, NJ 08903

APPENDIX B
Steering Committee 1975

Karen Duncan, Ph.D.
College of Dental Medicine
Medical University of South Carolina
80 Barre Street
Charleston, South Carolina 29401
803-792-3211

Barbara B. Farquhar
Lab Computer Science
Massachusetts General Hospital and
Harvard School of Medicine
Boston, Massachusetts 02114
617-726-3924

Gerald Fleischli, M.D.
Health Information Technology Services, Inc.
University of Nebraska
8073 Cedar Street
Omaha, Nebraska 68124
402-397-6122

Thomas H. Meld, Ed.S.
Director of Media
University of Maryland
Office of Medical Education
660 W. Redwood St., Rm. 257
Howard Hall
Baltimore, Maryland 21201
301-528-6613

Paul Jolly, Ph.D.
Association of American Medical Colleges
One Dupont Circle
Washington, D.C. 20036
202-466-5148

Ruann E. Pengov, M.S.
Director, Division of Computer Services
Health Sciences Library
Ohio State University
Columbus, Ohio 42310
614-422-6192

Paul G. Rehkopf
Design Engineer
Eye and Ear Hospital
230 Lothrop Street
Pittsburgh, Pennsylvania 15213
412-683-3500 X280

Charles S. Tidball, Ph.D., M.D.
George Washington University Medical Center
2300 Eye Street, N.W.
Washington, D.C. 20037
202-331-6547

Robert G. Votaw, Ph.D.
Assoc. Dean for Medical Ed./Preclinical
Dept. of Medical Education - AG069
University of Connecticut Health Center
Farmington, Connecticut 06032
203-674-2403

Bernice M. Wenzel, Ph.D.
Assistant Dean for Educational Research
UCLA School of Medicine, Dean's Office
Center for the Health Sciences
Los Angeles, California 90024
213-825-2716

APPENDIX C
Health Education Network, Inc.
Board of Directors and Officers, 1976-1977

Robert G. Votaw, Ph.D.
Associate Dean for Medical Education
Office of Medical Education AG069
U. Conn. Health Center
Farmington, CT 06032
(203) 674-2403
Term: 1976-1978

Charles S. Tidball, M.D., Ph.D.
Director of Computer Assisted Education
George Washington University Medical School
Ross Hall 257
2300 Eye Street, NW
Washington, DC 20037
(202) 676-2869
Term: 1976-1978

Charles M. Goldstein
Chief, Computing Technology
Lister Hill National Center for Biomedical Communications
National Library of Medicine
8600 Rockville Pike
Bethesda, MD 20014
(301) 496-4441
Term: 1976-1978

Martin Kamp, MD
Chief of Academic Computing Services
Information Systems (76U)
Room 76 University Hospital
University of California
San Francisco, CA 94122
(415) 666-4546
Term: 1976-1977

Robert C. Hickey
Associate Vice Chancellor of Health Professions
University of Pittsburgh School of Medicine
Alan Magee Scaife Hall
Pittsburgh, PA 15261
(412) 624-2845
Term: 1976-1977

Weston D. Gardner, M.D.
Professor of Anatomy
Medical College of Wisconsin
561 N. 15th Street
Milwaukee, Wisconsin 53233
(414) 272-5450
Term: 1976

Harvey S. Long
Instructional Systems Consultant
IBM Corporation
10401 Fernwood Road
Bethesda, MD 20034
(301) 897-3261
Term: 1976

William C. Brown
Head, Information Science Section
National Research Council of Canada
M-50 Montreal Road
Ottawa, Ontario
CANADA KIA OR8
(613) 993-2484
TELEX 053-4134
Term: 1976

Ruann E. Pengov
The Ohio State University
College of Medicine
3190 Medical Science Building
333 West 10th Street
Columbus, Ohio 43210
(614) 422-9063
Term: ex-officio

Barbara B. Farquhar
Laboratory of Computer Science
Gray 1323
Massachusetts General Hospital
Fruit Street
Boston, MA 02114
(617) 726-3924
Term: ex-officio

Karen A. Duncan, Ph.D.
Office of Computer Resources
College of Dental Medicine
Medical University of South Carolina
80 Barre Street
Charleston, SC 29401
(803) 792-3211
Term: 1976-1977

OFFICERS OF THE HEALTH, EDUCATION NETWORK, INC. FOR 1976:

President
Robert G. Votaw

Secretary
Barbara B. Farquhar

Treasurer
Charles S. Tidball

APPENDIX D
Available Programs at Massachusetts General Hospital

Programs that use a MULTIPLE CHOICE format:

Arrhythmias code: ARR	1. Simulates patients with variety of common rhythm disturbances; task is to diagnose and manage arrhythmia. 2. Target pop: med students in clinical yrs; physician continuing education; nurses in ICU's. 3. Avg time to complete 1 sequence is 20 min. 4. Four "patients" in case file.
Cardiopulmonary Resuscitation code: CPR	1. Simulates patient who may have suffered cardiac arrest; task is to resuscitate patient. 2. Target pop: med students in clinical years; physician continuing education; nurses in ICU. 3. Avg. time to complete 1 sequence is 15 min. 4. Five "patients" in case file; subsequent cases vary randomly
Hypertensive Emergencies code: HYP	1. Simulates patient with sever hypertension; task is to evaluate and treat patient. 2. Target pop: med students in clinical years; physician continuing education. 3. Avg time to complete 1 sequence is 15 min. 4. Three "patients" in case file; subsequent cases vary randomly
Hypertension management code: HTM	1. Simulates patient with hypertension; task is to manage patient. 2. Target pop: med students in clinical physician continuing education 3. Avg time to complete 1 sequence is 15 min. 4. Four "patients" in case file (several new cases being prepared)
Idiopathic Respiratory Distress in the Newborn	1. Presents cases and information on general neonatology and RDS and on pulmonary physiology

code: RDS

2. Target pop: all students of the neonate
3. Avg time to complete 1 sequence is 15 min.
4. 18 "patients" in case files

Programs which use a VOCABULARY:

Abdominal pain

code: ABD

1. Simulates patient with chief complaint of abdominal pain; task is to diagnose complaint
2. Target pop: med students in clinical years and beyond
3. Avg time to complete 1 sequence is 20 min.
4. Eight "patients" in case files

Coma

code: COM

1. Simulates comatose patient; task is to diagnose and treat cause of coma
2. Target pop: med students in clinical years; physician continuing education
3. Avg time to complete 1 sequence is 25 min.
4. Thirty-three "patients" in case files

GI Bleed

code: GIB

1. Simulates patient with GI bleeding; task is to diagnose and manage cause of bleeding
2. Target pop: med students in clinical years and beyond
3. Avg time to complete 1 sequence is 20 min.
4. Fourteen "patients" in case files

Pediatric Cough and Fever

code: PCF

1. Simulates child with cough and fever; task is to diagnose cause of complaint
2. Target pop: med students in clinical years and beyond
3. Avg time to complete 1 sequence is 20 min.
4. Forty-five "patients" in case files

Trauma

code: TRA

1. Simulates patient who has sustained multi-system trauma; task is to identify and treat immediate critical problems
2. Target pop: med students in clinical years

3. Avg time to complete 1 sequence is 25 min.
4. Seven "patients" in case files

Jaundice

code: JAU

1. Presents case history of patient with jaundice; task is to determine cause of jaundice
2. Target pop: med students, prior to clinical years
3. Avg time to complete 1 sequence is 25 min.
4. Six "patients" in case file

Joint Pain

code: JOI

1. Simulates patient with acute joint pain; task is to diagnose cause of the joint pain and begin treatment
2. Target pop: med students in clinical yrs, physician continuing education
3. Avg time to complete one case is 25 min.
4. Twenty "patients" in case files

Hypertension Diagnosis

code: HBP

1. Simulates patient with newly diagnosed hypertention.
2. Target pop: med students in clinical years, physician continuing education
3. Avg time to complete one case is 25 min.
4. Ten "patients" in case files

Programs which accept NATURAL LANGUAGE input:

Diabetic Ketoacidosis

code: DKA

1. Simulates patient in diabetic coma; task is to manage "patient's" condition until stabilized
2. Target pop: med students in clinical years; physician continuing education
3. Avg time to complete 1 sequence is 25 min.
4. Initial pt conditions set; course depends on user's management

Experimental Trauma Case

code: TRE

1. Simulates patient with multi-system trauma; task is to diagnose and manage immediate problems
2. Target pop: med students in clinical years; physician continuing education

3. Avg time to complete case is 30 min.
4. One "patient" in case files

Other programs:

Cardiac Simulation

code: CSM

1. Simulates cardiovascular system.
2. Target pop: students of cardiovascular physiology.
3. Avg. time to complete 1 sequence is 25 minutes.
4. User changes cardiac variables by entering new valves.

Digoxin Dosage Advisor

code: DIG

1. Computer suggests a reasonable dose of digoxin to administer based on patient data.
2. Target pop: pharmacists; med students in clinical years; physician continuing education.
3. Avg. time to complete 1 sequence is 10 min.
4. User enters values for patient variables.

Drugs in Renal Failure

code: DRF

1. Computer advises user on administration of drugs to patient with impaired renal or hepatic function.
2. Target pop: med students in clinical years; pharmacist, physician continuing education.
3. Avg. time to complete 1 sequence is 10 minutes.
4. User enters values for patient variables.

Anticoagulant Simulator

code: ACS

1. Simulates patient taking anticoagulant; task is to keep prothrombin time in therapeutic range.
2. Target pop: med students (2nd through 4th years).
3. Avg. time to complete 1 sequence is 10 minutes.
4. User enters dosage of anticoagulant to be given for each day.

NEW PROGRAMS
(1/75)

Digitalis Teaching Program
Code: DTP

1. Presents information on digitalis, and questions users' retention and understanding of this information.
2. Target pop: Med students in clinical years; physician continuing education
3. Average time to complete entire program is 1-2 hours, but it is designed to be taken in sections which take 15-20 minutes to complete.

Lab Test Simulator Program
Code: LAB

1. Give patients history and physical exam information; user orders lab tests toward diagnosing the cause/causes of the patient's symptoms.
2. Target pop: Medical students in clinical years; physician continuing education.
3. Average time to complete 1 sequence is 30 minutes.
4. Two "patients" in case files.

APPENDIX E
CAI Courses — Subject Listing

THIS INDEX LISTS ALL OHIO STATE UNIVERSITY COLLEGE OF MEDICINE CAI COURSES (RELEASABLE AND RESTRICTED) WHICH ARE CURRENTLY OPERATIONAL. (REPRODUCTION OF PRINT OUT)

TOPIC	CAI COURSE
ABBREVIATIONS	ABBREV
ABDOMEN, ANATOMY OF	GANATI (SEC 2)
ACETONE, TESTING FOR	PAMADI (SEC D)
ACID-BASE BALANCE	ABEL
	ACIBA
	PCHEMI
	SERAB
ANATOMY	
ABDOMEN	ABANG
	GANATI (SEC 2)
ANKLE	LOWEX (3)
ANTERIOR ABDOMINAL WALL	GANATI (I)
BACK	GANATI (B)
BONES	DIETAN (C)
BRAIN	VESALI (C)
	SITIN
CARDIOVASCULAR SYSTEM	DIETAN (H)
	CCNUR (I)
DIGESTIVE TRACT	DIETAN (C & D)
ELBOW	ELBOW
ANTERRACHIUM	ELBOW (C & D)
BONE LANDMARKS	ELBOW (A)
BRACHIUM	ELBOW (B)
VASCULAR SUPPLY	ELBOW (E)
VEINS	VEINS (A)
ENDOCRINE SYSTEM	DIETAN (I)
	ENDO
ENDOTRACHEAL INTUBATION	INTUBE (B)
EXTREMITIES	
LOWER	GANATI (6)
	VESALI (B)
	LOWEX
	DIETAN (B)
UPPER	GANATI (7)
	VESALI (A)
	ELBOW
	HANAT
	DIETAN (B)
EYE	GANATI (10)
	OPHTHA
GENITAL	SEXED
GLUTEAL REGION	GANATI (4)
HAND	HANAT
HEAD	GANATI (9)
HEART	CCNUR (1)
HIP	LOWEX (1)
INTESTINE	DIETAN (F)

MICROBIAL	PILOT (9)
MUSCLES	DIETAN (8)
OF ELBOW JOINT	ELBOW
OF SHOULDER JOINT	SHOJT
NECK	GANAT1 (9)
NERVOUS SYSTEM	DIETAN (1)
NEURO	NEURO
PELVIC, CLINICAL EXAM	GANAT1 (5)
PELVIS	GANAT1 (4)
SKULL	SKULL
SPINAL CORD	VESAL1 (C)
TERMINOLOGY	GANAT2 (1)
THORAX	GANAT1 (3)
VEINS	VEINS (A)
ANESTHETICS	AGENT
ANGIOGRAPHY	ABANG
ANKLE, ANATOMY OF	LOWEX
APPENDIX, DISEASES OF	PHASE4
ARRHYTHMIAS	
ATRIAL	STRIPS (B)
HEART BLOCK	STRIPS (C)
VENTRICULAR	VENTAR, STRIPS (A)
ARTERIAL SYSTEM	ABANG
ARTERIOGRAPHY	ABANG
BASIC PATIENT CARE	BUGOUT
	SIGNS
	REVIEW
	VEINS
	UREME
BLOOD PRESSURE	SIGNS (D)
CAI, GENERAL INTRODUCTION	CAI
CARDIOVASCULAR	CASE 1
	CASE 7
	CASE 9
	RESCUE
	RESTUP
	PACARE
	STRIP1
	STRIPS
	STROKA
	SHOCK
	VEINS
	VENTAR
	PILOT (8)
	CCNUR
	CVTERM
	DIETAN (H)
CLINICAL NUCLEAR MEDICINE	RADIO
	RADTEC
CLOSED DRAINAGE SYSTEMS	BOTTLE
COLON	PILOT (4 (D))
CRANIAL NERVES	SKULL
DIABETES	JUDT
	JUDTAR
	PAMADI
DIAGNOSTIC TESTS	ENDO
	ENZICS
	RADIO
	RADTEC
	SERAB
	UREME
	URAL
	VEINS
	THERM
DIET PLANNING	FOODS
DIGESTIVE SYSTEM	AUTRAC
	DIETAN
	FOODS
	NUTRO

	PHASE4
	QUAILS
	RADIO
	RECIPE
	REVIEW
	PEDS
DISEASES OF INFANTS AND CHILDREN	PILOT (3)
	A-THE FETUS AND NEWBORN
	B-CONGENITAL DEFECTS
	C-DEVELOPMENTAL DISORDERS
	D-DISORDERS OF NUTRITION AND HYDRATION
	E-ACCIDENTS (TRAUMA AND POISONS)
	F-ILLNESS AND THE CHILD
DRUGS	PILOT (S)
	AGENT
	PSYMED
DENTISTRY	ORALCA
	ORPAIN
	PHARMA
	TOOTHA
	TRETPA
ELECTROCARDIOGRAPHY	STRIPL
ENZYMES	ENZICS
ENDOCRINOLOGY	GANAT2 (6)
	CASE 5
	CASE 8
	PILOT (J)
	PILOT (5)
	A-ENDOCRINOLOGY
	B-ACID-BASE MECHANISMS
	DIETAN (1)
	ENDO
ENDOTRACHEAL INTUBATION	INTUBE
ESOPHAGUS	PILOT (N (A))
FIRST AID	ONEAID
GASTROINTESTINAL SYSTEM	PILOT (E)
	PILOT (N)
	A-ESOPHAGUS
	B-THE STOMACH AND DUODENUM
	C-THE SMALL BOWEL, MALDIGESTION AND MALABSORPTION
	D-THE COLON
	E-THE LIVER
	F-THE PANCREAS AND BILIARY TREE
	GANAT2 (4)
GENETICS	PILOT (K)
HEMATOLOGY	PILOT (R (H))
	PILOT (H)
	PILOT (U)
	A-HEMOGLOBIN AND THE RED CELL
	B-COAGULATION AND BLEEDING DISORDERS
	C-MYELOID ELEMENTS (GRANULOCYTES)
	D-LYMPHORETICULAR DISEASES (NON-FUNCTIONAL)
	E-FUNCTIONAL DISEASES OF IMMUNE ORGANS
	F-IMMUNOHEMATOLOGY
HISTOLOGY	HISTI
HISTOLOGICAL TECHNIQUES	TOADS
INFECTIOUS DISEASES	CASE 4
	PILOT (V)
	A-BASIC CONSIDERATIONS 10
	B-CHEMOTHERAPY - IMMUNIZATION 2
	C-MAJOR ANATOMICAL GROUPINGS 15
	C-QUESTIONS FOR REVIEW AND CASE STUDIES 41
INFECTION CONTROL	BUGOUT

IMMUNOLOGY	PILOT (P)
INTESTINE	DIETAN (F)
KIDNEY	DIETAN (G)
LOGARITHMS	REVIEW (B)
LIVER	DIETAN (F)
MATH	MATH
MEDICAL RECORDS	MEDREC
MEDICINE, NUCLEAR	NUCTEC
MEDICINE, SOCIAL & PREVENTIVE	PILOT (T)
MICROBIOLOGIC MECHANISMS	PILOT (R)
	A-MICROBIAL ANATOMY
	B-MICROBIAL GENETICS
	C-MICROBIAL AGENTS OF DISEASE
	D-DISINFECTION AND STERILIZATION
MUSCULOSKELETAL SYSTEM	PILOT (1,2)
	GANAT2 (3)
	MORPHO
MUSCLES	DIETAN (B)
NERVOUS SYSTEM	DIETAN (J)
CENTRAL	PILOT (D)
MUSCULAR AND PERIPHERAL	PILOT (H)
NUCLEAR MEDICINE	NUCTEC
NUTRITION	PILOT (E)
OB/GYN	OHTERM
	ORTOX
	ALTRAC
	BREECH
	CASE 2
	PILOT (L)
	PILOT (A)
	A- I. ANATOMIC CONSIDERATIONS
	1. DEVELOPMENTAL ANOMALIES
	B- IV. BENIGN TUMORS
	7. MYOMAS
	C- V. MALIGNANT NEOPLASMS
	11. CARCINOMA OF THE ENDOMETRIUM
	C-VII. GENERAL PRINCIPLES
	18. LOWER ABDOMINAL PAIN
	GANAT1
OPHTHALMOLOGY	CASE 6
	FUNDUS
	OPHTHA
	REDEYE
	REVIEW (B)
OPTOMETRY	OPTOM
	OPCLIN
ORTHOPAEDICS	CASE 3
	LUMBUM
	ORTRAC
	PROSHK
PATHOLOGIC MECHANISMS	PILOT (Q)
	A-INFLAMMATION AND REPAIR
	B-CONGENITAL DISEASE
	C-DEGENERATIVE PROCESSES
	D-NEOPLASIA
PEDIATRICS	PEDS
	TOTS
PH	REVIEW (C)
PHARMACOLOGY	AGENT
	TRUAMI
	VEINS
	PSYMED
	MATH
PHYSICS, NUCLEAR	NUCTEC (1)
PHYSIOLOGY	PHYSIO
	ENDO
	NUTRO

	PCHEM1
	REVIEW (E)
	SEXED (C,D)
	SHOCK
	STRIP1
	THERM
PHYSIOLOGICAL CHEMISTRY	NUTRO
	PCHEM1
PRIMARY CARE	ONEAID
PSYCHIATRY	CASE-5
	GROUP
PSYCHOLOGY	PILOT (N)
	A-INDIVIDUAL AND GROUP HUMAN BIOLOGICAL BEHAVIOR 19
	B-PSYCHOLOGICAL DEVELOPMENT 36
	C-LEARNING AND MOTIVATION 19
	D-MAN'S SENSORY WORLD 8
	E-THORACIC NEOPLASMS
	F-DIAGNOSTIC METHODS
	GANAT2 (7)
	ACIBA
	AHEL
	BOTTLE
	INTUBE
	PACARE (STOP SMOKING SECTION)
	REVIEW
RESPIRATIONS	SIGNS (C)
SEXUAL DYSFUNCTIONS	SEXED (G)
SEX EDUCATION	SEXED
SEXUAL RESPONSE	SEXED
SHOCK	SHOCK
SKIN	PILOT (G)
	A-SKIN AND CONNECTIVE TISSUES
	B-SKELETAL SYSTEMS 13
	GANAT2 (2)
STAINING	TOADS
STATISTICS	QUCOST
SURGERY	AGENT
	ASSESS
	BOTTLE
	QUAILS
	INTUBE
	ORTRAC
	PROSHK
TEMPERATURE	VEINS (A,B,C)
	REVIEW (SERVICE MODE - E)
TERMINOLOGY	
MEDICAL ABBREVIATIONS	ABBREV
BASIC PERCEPTUAL CONCEPTS	PERCEP
CARDIOVASCULAR	CVTERM
OBSTETRIC	OBTERM
MEDICAL	TERMS
URINALYSIS	URAL
VEINS	VEINS
VITAL SIGNS	SIGNS
UROLOGY	CASE 6
	E-THEORIES OF DEVELOPMENT 14
	F-CAPSULE CASES 16
	PILOT (4)
	A-THE SCOPE OF ABNORMAL PSYCHOLOGY
	B-CLASSIFICATION OF PSYCHIATRIC SYNDROME
	C-THE NATURE OF CAUSATION
	D-PSYCHIATRIC SYNDROMES
	E-DISORDERS OF CHILDHOOD
	F-EVALUATION AND TREATMENT
	G-PHARMOCOLOGIST MECHANISMS
	H-PSYCHIATRIC CASE CAPSULES

	1-PRACTICE INTERVIEWING
PULSE	SIGNS (R)
RADIOLOGY AND NUCLEAR MEDICINE	NUCTEC RADTEC RADID
RADIATION THERAPY	IONIZE
RENAL EXCRETORY SYSTEM	PILOT (D) PILOT (Z) A-CLINICAL ASPECTS B-INTRINSIC RENAL DISEASE C-SYSTEMIC DISEASES AFFECTING THE KIDNEY D-RENAL VASCULAR DISEASE AND HYPERTENSION E-ANATOMIC AND SURGICAL DISORDERS F-DISEASES OF THE MALE GENITAL TRACT
REPRODUCTIVE SYSTEM	PILOT (6) A- I. ANATOMIC CONSIDERATIONS 1. DEVELOPMENTAL ANOMALIES B- IV. BENIGN TUMORS 7. MYOMAS C- V. MALIGNANT NEOPLASMS 11. CARCINOMA OF THE ENDOMETRIUM D-VII. GENERAL PRINCIPLES 18. LOWER ABDOMINAL PAIN PILOT (L) GANAT2 (9)
RESPIRATORY SYSTEM	ASSESS PILOT (C) PILOT (Y) A-VENTILARY DEFECTS DUE TO FAULTY PULMONARY AND THORACIC MECHANISMS B-MISCELLANEOUS DISEASES OF THE LUNG C-RESPIRATORY EFFECTS OF INFECTIONS D-RESPIRATORY ASPECTS OF SYSTEMIC DISEASE OR DISEASE OF OTHER SYSTEMS
VITAL SIGNS	**SIGNS**
UROLOGY	**CASE 6**