DOCUMENT RESUME

ED 073 696 EM 010 883

TITLE	Networks and Disciplines; Proceedings of the EDUCOM Fall Conference (Ann Arbor, Michigan, October 11-13, 1972).
INSTITUTION	Interuniversity Communications Council (EDUCOM), Princeton, N. J.
PUB DATE	73
NOTE	204p.
AVAILABLE FROM	EDUCOM, P. O. BOX 364, Princeton, New Jersey 08540 ($5.00 members, $6.00 nonmembers)
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	*Computers; Computer Science; *Conference Reports; Data Bases; Higher Education; *Information Networks; Interagency Cooperation; *Interinstitutional Cooperation; *Networks
IDENTIFIERS	EDUCOM

ABSTRACT

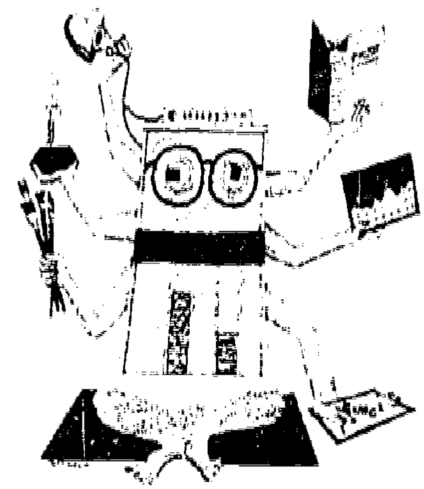
The Fall, 1972 conference of EDUCOM focused on problems, potentials, and current activities related to the development of interinstitutional cooperation through computer networking. In these proceedings, a number of representative networking attempts were presented, discipline-oriented workshops discussed possibilities in a number of different fields, more general findings about networking progress were reported, the impact of networking was discussed, and panel discussions attempted to provide an overview. (RH)

Networks and Disciplines

1973

EDUCOM
Interuniversity Communications Council, Inc.

Networks and Disciplines



PROCEEDINGS

of the

EDUCOM

FALL CONFERENCE

October 11, 12, 13, 1972
Ann Arbor, Michigan

EDUCOM
The Interuniversity Communications Council, Inc.

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 $5.00 for members and $6.00 to nonmembers of EDUCOM.

table of contents

III. NETWORK DEVELOPMENTS AND DATA BASES

IV. IMPACT OF NETWORKING: BANQUET ADDRESS AND WORKSHOPS

V. MAJOR ISSUES AND CONCLUSIONS: PANEL DISCUSSIONS

SUPPLEMENTS

foreword

The development of interinstitutional cooperation through computer networking has long been a goal of EDUCOM. With the technology available today, those purposes and goals for EDUCOM which were outlined in the book, *EDUNET*, in 1967 are capable of being achieved. Progress toward their achievement is being made and the fruits of real cooperation appear attainable once organizational and political barriers can be overcome.

The Fall 1972 Council Meeting and Conference, the fifth in a series of conferences focusing on computer networking for higher education, evidenced a striking consensus among computer users in higher education that the technology for computer networking is in hand. Representatives of various disciplines including chemistry, linguistics, economics and social sciences cited many specific examples of applications of computing in their disciplines which were available to be shared, expressed confidence that the technology which would permit sharing was also available and described a few cases in which sharing was already taking place. The obstacles which most speakers foresaw to widespread sharing of computer resources in all disciplines were primarily organizational, political and economic.

Over 250 persons attended the three-day conference which was held at the Chrysler Center for Continuing Engineering Education on the campus of the University of Michigan in Ann Arbor, Michigan. The Michigan Interuniversity Committee on Information Systems (MICIS), which has sponsored the development of the MERIT network, acted as host for the meeting. The conference followed the format developed in recent EDUCOM meetings of plenary sessions alternating with small group discussions.

The major speech at the banquet held in conjunction with the conference provided a change of pace for conferees. Dr. Edwin B. Parker, Professor of Communications at Stanford University, spoke on "Social Information: The Technology, The Need and the Challenge." He outlined his view of the possible and the probable futures for cable television, especially as related to higher education and challenged organizations like EDUCOM to help make possible the level of expenditure which will be necessary to move our society to the kind of education system which can be provided by means of cable television.

Panel presentations and speeches given during the conference have been edited by the speakers and are collected in the following pages. Summaries of the group discussions were drafted by a recorder in each group and have been edited for publication by the group chairman.

Further information concerning any of the systems or suggestions described by speakers can best be obtained by writing directly to the author of the presentation. Names and addresses of all conference participants are listed in the back of this volume.

Many thanks are due to the Program Chairman, Tom Kurtz, and the members of his program committee for the excellent program which they developed. A great deal of credit for the success of the conference is also due to each participant in the plenary sessions and the workshops.

Henry Chauncey

introduction

As the higher education community moves toward a posture of closer cooperation, many discipline groups — *logical networks* — have come into being. The EDUCOM 1972 Fall Council Meeting presented as examples several of these interinstitutional groups. It is asserted that such groups, acting in concert, can justify some sort of network utility connecting not only the institutions they represent but also most of the others as well.

The conference was organized about several panel presentations on current networking activities in six disciplinary areas and on the state-of-the-art in networking and data bases. These presentations are collected in the following pages. Peter Lykos outlines specific examples of the use of networks in chemistry: a large data bank with complex molecular structures, accessed by remote terminals, with a three-dimensional image capability; a collection of programs carefully constructed for general use permitting quantum/chemical calculations from remote terminals backed up by hot phone counseling; and a small college remote job entry station with graphics augmentation accessing a large-scale computer. Walter Sedelow, discussing networking activities related to languages and humanities, describes the Ce/NCoReL Study. He concentrates on some of the conclusions emerging from the exploration of a very wide range of types of language research which might be involved in networking. Some of the current independent networking efforts which are underway to serve the approximately 40,000 professional economists in the United States are outlined by Sanford Berg. Most of these activities are presently based in a variety of scientific societies, research institutes, government agencies and universities. Dr. Berg notes especially the lack of coordinating mechanisms for the independent efforts. Frederick Kilgour describes various online computer applications for libraries: MEDLINE, a remote subject retrieval system based at the National Library of Medicine; Basis-70 operated by the Battelle Memorial Institute; and internal library online applications for acquisitions, cataloging and circulation at the New England Library Information Network, Stanford University, The Ohio State University, and the Ohio College Library Center. The uses of computing in museums and related disciplines are summarized by Robert Chenhall and related to the existing museum data bank systems. He notes the necessity for current efforts in museum data bank systems to maintain a structural similarity in order to eventually develop regional and national networks of catalogs. James Davis, discussing the use of computer networking in the social sciences, outlines some problems faced in utilizing survey data which might be ameliorated through the use of networking.

Four presentations address from different points of view the state-of-the-art in network development and data bases. K. Leon Montgomery identifies and examines the current trends in machine-readable data bases. D. Don Aufenkamp describes National Science Foundation activities within the National Science Computer Network initiative as they relate to the development of the resouce sharing potential of such a network. Bertram Herzog gives an overview of the MERIT Computer Network project, explaining the original goals, technical design objectives and the organization of the project. Balancing individual institutional use over the network, as well as providing user support, are two of the primary concerns of the project at this time. Edward Weiss outlines the new aims and policies of the Office of Science Information Service of the National Science Foundation which relate to the interests and needs of the university community.

In the workshops' summaries and the following panel session, questions concerning the method of organizing interinstitutional computer communications are addressed. What kind of network organization makes sense for each group or discipline: national networks or regional networks? National centers or local centers? What kinds of computing technology are most important for each group or discipline using a network: raw computing power, large active files, large static files, programs and software, interactive time-sharing, graphic terminals, character terminals, communications? What kinds of support patterns are needed for each group or discipline in order to utilize computer communications networks: foundation grants, government subsidies, or institutional budgets? What are the advantages and disadvantages of each form of support for each group or discipline?

A demonstration of MERIT Network capabilities and discussions of related topics were conducted concurrently with the workshops focusing on disciplines. Summaries of these alternative workshops outline methods of support for faculty using computer applications in instruction and procedures and experiences in contracting with vendors of computer resources.

Members of the panel, "Networks and Computer Centers: Cooperation or Conflict?" describe their experiences in organizing and financing computer center operation in the context of a computer network. Problems faced by the individual computer center director in a network situation are the primary foci of these presentations.

In an address which was presented at the banquet during the conference, Edwin Parker challenges the conference participants to work for the utilization of television for education and outlines the kind of applications which might be possible if sufficient support were available.

The Fall 1972 Council Meeting and Conference provided an opportunity for faculty already organized in disciplinary groups and working through computer communications networks to bring representatives from other institutions into these working groups. For groups and disciplines which had not yet utilized networks, the conference provided an opportunity to learn of computer applications which might be available to them through a network and to develop working relationships in order to take advantage of these applications.

All of the speakers, workshop chairmen, and panel members at the conference contributed much time and effort to the program. Without their contributions, the level of discussion and wealth of information available at the conference and included in this volume, would not have been possible.

Thomas Kurtz
Program Chairman

CHAPTER 1

NETWORKING ACTIVITIES IN THE DISCIPLINES

Networks and Disciplines: A Point of View

Thomas Kurtz, Director Kiewit Computation Center
Dartmouth College

This organization, EDUCOM, had its origins in 1965, when a group of administrators in higher education had the vision to perceive that these institutions could better serve their purposes through cooperation and intercommunication than by going it entirely on their own. The major manifestation of their vision was the Boulder, Colorado study of 1966. This study, published in book form under the title *EDUNET*, explored in some detail the ways in which institutions could benefit from networking, and the ways in which networking could be technically achieved.

It was not in the cards for a physical network to come into being at that time. The actual use of communications to connect computers was rare and most of the now standard time-sharing systems were in their infancy. The vast majority of the higher education "public" had no more than a vague notion of the concepts of computing, communications, networking, and time-sharing. It was, therefore, not easy to convince this public that networking had any role at all to play in their lives.

Although a national network was not in the cards in 1966, EDUCOM pursued its fundamental goal to improve communication between universities. Its activities in this direction have been vigorous and varied: seminars and national meetings, studying the role of computing in libraries, especially medical libraries, exploring one form of program exchange (EIN), representing the computer point of view in the hearings for the copyright law and many others. All of these activities are legitimate and easily justifiable. But, the national network does not yet exist. The question is -- "Why not?" Although there are no headlines or fireworks to distinguish this particular Council Meeting, it is my belief that future historians of such matters will regard 1972 as the crucial turning-point year in the development of interinstitutional cooperation through networking. We should, therefore, examine some of the reasons why networking has proceeded so slowly in the past, and why I think that the pace will now accelerate noticeably.

We often hear that higher education is a "marketplace of ideas." Whether or not this is a good description, the marketplace model is useful for understanding our present concern. The two important ingredients in any marketplace situation are supply and demand. If there is some service that is generally available at some price, its success will depend on whether or not the community perceives using that service as being more useful than the alternative of not using the service. Conversely, if there is a perceivable demand for some form of service, then providers of such service will spring up. Many argue that the situation with networking is similar to the chicken-and-egg dilemma, that is, which comes first. As with the chicken-and-egg situation, I believe the only relevant answer that pertains to networking is: both. The fact is that there are both a demand and a supply for networking at the present time, albeit the demands and supplies are varied and diffuse, are sometimes difficult to identify, and are hard to bring together with the supply in the marketplace. While I believe the marketplace model is a useful aid to our understanding, it is unlikely that a conventional commercial marketplace can operate in this case. Large subsidies are needed, much experimentation needs to be done, and the capital costs are very high in terms of money, time, and institutional commitment.

The elusive national network that was conceived six years ago no longer seems as elusive. The ARPANET, as well as several commercial nets, clearly demonstrates the feasibility of large-scale national (and international) networks. Local and intrastate networks in many forms exist. The crucial matter that we face in 1972 is bringing together to produce a sale in the marketplace the many demands for networking that already exist, and the many sources of supply of networking that are coming into being.

The proceedings of the 1966 Boulder study are recorded in *EDUNET*. This book identifies three components (slightly paraphrased) to be considered:

- Needs and Applications,
- Network Facilities, and
- Organizational Structures.

The first obviously represents the *demand*, the second relates to the *supply*, and the third is an essential part of the marketplace mechanism.

Although the third component is the most difficult to solve, and should perhaps receive the greatest attention, we simply note that EDUCOM exists and that we are here at this conference. Progress is being made. Further progress should occur as a result of the General Working Seminars to be conducted by EDUCOM in the next three months.

Past and future EDUCOM conferences will be devoted to a discussion of the supply of network facilities. Indeed, on the program for this meeting, we will hear about one such facility in Michigan, the MERIT Network. But, I claim that the supply component is the easy one, and will say no more about that.

Finally, we come to the *demand* component in the marketplace of networking. If we were willing to be simple-minded, we might identify the major (not necessarily exclusive or exhaustive) types of application: ETV, CAI, and digital services and information. There has not yet been significant demand for ETV and related video activities which require physical networking, so we will not discuss that. CAI, especially massive CAI, is not operational. Even at its best, it is still experimental. We now consider the area of digital (or machine-readable) information and services, and the organizational structures that are already using such information –

the various disciplinary groups or *logical networks* that already exist. I believe that too little attention has been paid to this particular *demand* component in the marketplace of networking. This conference gives us an opportunity to look more carefully at the individual and independent demands for networking, and to explore whether or not, in aggregate, they achieve the critical mass needed to justify a network utility. The backbone of the program for these two days will be the examination of six example disciplines that exist in the higher education community, to determine or to estimate the extent to which they are already involved with networking, and to project into the future how they might benefit from an expanded network supply.

Networking and Chemistry

Peter Lykos, Office of Computing Activities
National Science Foundation

Computer networking[1] is a complex subject with technical, political, and cost-accounting dimensions. If it is going to have an effect on chemistry over the next five-to-ten years, leadership will have to come from the chemistry research community working closely with computer scientists and engineers. In the following pages three components are developed which bear on computer and computer network use in chemistry: some characteristics of chemistry as they bear on computer utilization by chemists; synopses of four complementary projects involving computers, research in chemistry and elements which are potentially network components; and a look into the crystal ball.

CHEMISTRY

Chemistry is an experimental science based on the laws of physics, using the language of mathematics, and addressing the properties of materials including those important to the life process. The information processing machine augments the doing of chemistry, in all of its dimensions, in a natural way.[2] Research in chemistry is based on: measurements of physical properties; data reduction, transformation, and representation (often graphical); mathematical and physical modeling and prediction; and on literature and data storage and retrieval. The computer plays a basic and comprehensive supporting role in chemistry through several types of applications:

- Real time data logging and experiment control
- Simulation and modeling
- Synthesis planning
- Analysis
- Information storage and retrieval
- An aid to pedagogy

Although chemistry is a large and widespread profession, it is served by a single professional society. Approximately 200,000 professionals are currently working at chemistry while another 100,000 have received basic training in chemistry at the undergraduate level as part of their preparation for engineering, medicine, and other fields. Approximately 100,000 chemists are members of the American Chemical Society. The ACS has a long history of effective service. In addition to journal publication and meetings which are local, regional, and national in scope, the society has a substantial and modern program in continuing education. The ACS employs many channels of communication, a wide variety of supporting services, and generally provides a *cohesiveness due in part to a general,* *common, and graphic language of molecular structure and properties.* The ACS Division of Chemical Education is concerned with the quality and content of chemical education.[3] There are 452 ACS accredited B.S. programs in chemistry and a similar number of non-accredited B.S. programs. These programs average about nine B.S. degrees awarded per school per year. The small number per class makes it possible to *routinely include research in chemistry as part of the undergraduate training* in chemistry. Recently that Division formed a new standing committee concerned with the computer and chemical education.

Chemical and allied products have a large impact on our economy. There is a large well-defined industry identified with chemistry. Sales of basic chemicals, drugs, refined petroleum, rubber and related plastics products, primary metals, stone, clay, glass, paper and allied products, amounted to $284 billion in 1971. Twenty percent of the national income is related to chemistry.

The chemistry profession can be described as a cottage industry. Except for chemical engineers, chemists typically work on a small scale. Individual researchers work with a few associated staff. In this regard, chemists are quite unlike their colleagues in Physics, who work cooperatively on large projects such as particle accelerators and astronomy observatories. Only in recent times, with time-sharing services and the massive wave of inexpensive, powerful mini-computers which the individual chemistry researcher can afford, have we seen widespread use and awareness of the computer in chemistry. There is presently no chemistry counterpart to the National Accelerator Laboratory at Batavia or the Very Large Array Radio Telescope under construction in New Mexico. However, a thrust toward a National Center for Computational Chemistry (NCCC) is gathering momentum. In early 1973 an NAS Report will be published which will outline the results of NCCC Feasibility Study. If the NCCC is established according to the recommendations of the feasibility study, its costs will be only 1/5 that of the National Accelerator Laboratory.

One of the oldest and most unique resources available to chemists could be even more fully utilized in the context of a computer network. Chemical Abstracts Service, a division of the ACS, operates the most comprehensive and largest computer-based abstracting and indexing system in science. CAS supports not only chemistry but, to a significant extent, the disciplines of Physics and Biology as well. The CAS operation is a model illustrating the smooth and orderly transition from a paper document based literature system to a computer based system with distribution taking place through regional information distribution systems. The widespread use of Chemical Abstracts Service is due in part to the common graphic language of molecular structure shared by all chemists.

FOUR COMPLEMENTARY PROJECTS INVOLVING CHEMISTRY RESEARCH AND COMPUTERS

As a demonstration of the "state-of-the-art" in computer resource sharing among research chemists, the following are presented.

A first example is the University of Kansas Chemical Physics Tri-Level Computer Network. At KU there are a number of physicists and chemists who share a building dedicated to chemical and physical research. Most of the research is experimental and heavily instrumented including dedicated mini-computers. Some of the research is theoretical making substantial use of large-scale computing resources. Professors Gilles and Culvahouse have designed and are implementing the Computer Network at KU as an alternative to the development by individual researchers of under-utilized computer centers built around mini-computers. Auxiliary equipment such as additional core memory or display devices broaden the capability of individual basic systems but frequently cost more than the original mini-computer and are used only a fraction of the time.

The hierarchical system at KU supports laboratory automation with computational services at three levels:

1. A dedicated mini-computer and appertances necessary are provided to particular experiments in each laboratory.
2. A nodal or hub computer physically located in the same building is interfaced to the dedicated mini-computers and to the central campus large computer. Attached to that hub computer is a comprehensive set of peripheral input/output devices, auxiliary storage equipment, and a graphics capability to support local massaging of results of large-scale remote computing.
3. The central campus computer provides large-scale remote computing services.

Many difficulties of computer networking are illustrated by the cooperation and coordination between different researchers sharing the same facility on a scale intermediate between a single laboratory and the campus-wide computer center.

All the problems of resource sharing are present and perforce need to be addressed including management and cost allocation. However, as the user community here is more homogeneous than that using the campus computer, cooperation and coordination should be less of a problem. The success of this hierarchical computing system for laboratory automation will provide a useful model for hierarchical computing in other areas such as the hierarchical system supportive to classroom computer use which was described in two previous EDUCOM publications.[4]

A second example of resource sharing through remote use is a project which is making quantum chemical programs available from the University of Utah. For many years, Indiana University has operated a Quantum Chemistry Program Exchange (QCPE). In addition to serving as a clearinghouse for applications software involving a minimum of testing, QCPE also publishes a newsletter and bulletin apprising its membership of the QCPE program content and activity, and of other relevant information such as the availability of remote access to the Lawrence Berkeley Laboratory CDC 7600, at internal rates, for federally sponsored research.

Professor Frank Harris of the Departments of Physics and Chemistry

NICKEL OCTAETHYLPORPHYRIN

Figure 1

of the University of Utah has undertaken the task of going beyond the services offered by QCPE by identifying those quantum chemistry computer-based methods which appear to be of greatest interest and utility to those who wish to apply quantum chemical methods. He is doing a careful job of designing and implementing corresponding algorithms such that the methods are accurate over the entire range of input parameters, contain a comprehensive set of alarms and diagnostics, and are transferrable over a reasonable variety of computer systems.

Since the University of Utah is a node on the ARPANET, the programs are tested on a variety of computer systems accessible via the ARPANET. In addition, a graphics terminal at the University of Utah is used to demonstrate the more powerful display option of the programs. Cooperating chemists desirous of using such programs may also access the Utah UNIVAC 1108 on a dial-up basis from TTY's. A telephone-accessible consultant at Utah is available to assist remote users at all times. Since Professor Harris has played a key role in the evolution of the National Center for Computational Chemistry (NCCC), the experience gained in his project will bear on the design and implementation of the NCCC.

A third example of resource sharing in chemistry research involves the Brookhaven National Laboratory, Texas A & M University, and the Institute for Cancer Research in Philadelphia. Dr. Walter Hamilton,[5] Chemistry Department, BNL, and Professor Edgar Meyer, Biochemistry Department, Texas A & M, have mounted a demonstration effort to discover the most cost-effective means for developing and maintaining a Protein Structure Data Base at BNL accessible over ordinary telephone lines on a dial-up basis from remote terminals with 3-D graphics display capability. The Data Base and supporting software are being developed at BNL. The terminal is being developed by Texas A & M University in cooperation with a local company.

When the prototype terminals are operational, one will be installed at the Institute for Cancer Research. Dr. Helen Berman at the Institute will help to hone and polish the overall system. Figure 1 shows a molecular structure which may be contained in the data base. When viewed through the stereoscopic viewers provided at the back of this volume, a 3-D image similar to that seen on the color-TV-based 3-D graphics terminal is visible.

An aspect of this project which should be highlighted is the fact that a National Science Laboratory is the focus. There are several National Science Laboratories around the country which could provide a focus for regional computer support systems for scientific research on a dial-up basis. AEC's Oak Ridge National Laboratory is well into a regional social data base system. Another example is the recently formed Institute for Computer Applications in Science and Engineering within the Universities Space Research Association of NASA's Langley Research Center, Hampton, Virginia.

The fourth example of remote computer use for chemistry research also involves a graphics capability. For his work with large molecules of biological interest, Professor David Beveridge of Hunter College CUNY uses a large-scale scientific computer to compute properties of large molecules and to generate pictures of those molecules which can be assembled into a dynamic moving picture display. Since Hunter College does not have sufficient computer-based research activity to warrant a large-scale scientific computer, Professor Beveridge has used a mini-computer as a remote job entry station to conveniently access the AEC-supported Courant Institute CDC 6600.

Recently Professor Beveridge started a project to modify that remote-job-entry station so that the heavy computing and graphic image generation may be done on the 6600 and the images transmitted in digital form to the local mini-computer for display and editing. In addition, the capability of the mini-computer is being expanded to enable Hunter College to access the incoming City University of New York IBM 370/168 which will be supporting RJE terminals at all 20 campuses of the 200,000 student CUNY system.

Two of the four examples have *international networking* implications. The Hamilton/Meyer/Berman effort involves cooperation with the Crystallographic Data Center, Cambridge, England. The Harris project could have impact on the Centre Européen de Calcul Atomique et Moleculaire, Orsay, France and on the large-scale Quantum Chemistry activity which links the University of Uppsala, Uppsala, Sweden, and the University of Florida, Gainesville, Florida.

A LOOK INTO THE CRYSTAL BALL

The chemist experimentalist will make increasing use of mini-computers and microcomputer components to further distribute "intelligence" in his experimental apparatus. The priority problem for real time computing will require that tightly-coupled computing, e.g., data collection and experimental control, be done locally while loosely-coupled computing, e.g., simulation and modeling, be done remotely. Local mini-computers may function as Remote Job Entry stations as well, probably channeled through the campus (company) computer center which will, increasingly, serve a broker function between the campus (company) user and the service and special purpose facilities available elsewhere.

The appearance of cost/effective graphics display devices provides an essential complement to the information processing machine, making it possible for the chemist to handle and to summarize a much larger collection of information than has hitherto been possible. *Dynamic graphical display raises to a distinctly new and higher level the chemist's ability to intuitize.* It is not mere coincidence that the four projects called out here all have a graphic component. Education in chemistry will be affected strongly as well by the increasing availability of graphics display. General purpose time-sharing computer service will be attained through the commercial sector. The Harvard Business School supports its graduate program from commercial time-sharing service. The IIT Research Institute released its in-house computer and leased its computer-conditioned space to a commercial vendor. IITRI's computer-using researchers, including chemists, operate on an open market basis with a variety of terminals accessing a variety of services seeking an optimal match between application need and service available.

Many literature and data bank services such as those being established by the U.S. Environmental Protection Agency[6] distribution centers handling CAS tapes, and others, will come to be used on a regional or national basis. Consolidation will be spurred by the increasing amount of data being generated, for example, by the NASA earth surveillance program.

Literature and data banks constitute an essential base for networking. Collections of basic scientific information constitute the strongest case for centralization of information bases and of information processing machines. The uniqueness of the common chemical language and the pre-eminent position of CAS suggest that chemistry will be an early participant in the movement toward national data centers accessible through computer networks.

A major trend in public higher education is another n... station of the Golden Rule.[7] States are beginning to self-organize regarding computer networking. Rich experiences already exist such as TUCC in North Carolina, MERIT in Michigan, and the three level system in California. A significant unknown factor is the extent to which the State Governments will be able to sense accurately the computer needs of higher education before proceeding with changes of a fundamental nature.[8] The State of Illinois Plan, discussed briefly at the April 13, 1972 EDUCOM Conference, "Networks for Higher Education," represents an extreme case in this regard.

Chemists in public institutions of higher learning are likely to be constrained to use statewide-accessible computer resources (where local mini-computers will not suffice) whose services will be organized around chemists' perceived needs and desires in direct proportion to the initiative they now exercise. Because many major private institutions have already fallen behind in their ability to sustain the largest and fastest computers available, they are more susceptible to participation in cooperative efforts. We can expect them to be among the first to hold at current on-site computer levels (or even reduce computer levels) and to increasingly use off-campus resources for researchers needing large-scale or specialized facilities.

The National Center for Computational Chemistry may play a role in fostering CNAR in chemistry. However, NCCC is at least 3-5 years in the future.

The evolving ARPANET (a whole new family of interface message processors under active development), with the promise and problems of ILLIAC IV and the trillion bit store, provides a fascinating and rich test bed for the technology of networking. As a functioning entity, it is at a stage where significant demonstrations in CNAR can be mounted. The ARPANET appears to have great potential for resource sharing of expensive and specialized facilities be they data banks, software, and/or hardware. For it to work as a research tool on a national scale, at least three requirements must be met:

1. The network must be operated by an organization responsive to the needs and desires of the scientific researchers with a high level of commitment by the major universities. The University Research Association, which has a 52 university membership and operates the National Accelerator Laboratory at Batavia, Illinois provides one model for such an organization. To date, the scientific community has not shown any initiative in this area.
2. A high level of federal commitment and coordination will be necessary which transcends the several agencies which support external (to the agency) basic research involving computer support. The principal agencies in this regard are ARPA, NIH, AEC, NSF, EPA, and NASA. Thus far no such program or plan has been announced although several of the individual agencies currently have studies underway.
3. An open market environment must exist within which the user can purchase the most cost-effective services available. Transition

to that state could be extremely difficult for many university computation centers.

Chemists could play a major role in mounting CNAR demonstrations on the ARPANET which would be scientifically sound, technically viable, politically comprehensive, and could provide a clientele to justify the scale of national experiments and resource centers.

REFERENCES

1 Computer Network Augmented Research = CNAR. A distinction is being drawn here between a *communication* network, which provides remote access to any *one* of a variety of computer services, and a *computer* network where two or more computers can work on portions of a given job in a manner reasonably transparent to the user. Remote access to a single major computer does not constitute a computer network in this sense. Currently at the National Science Foundation on leave from ITT. The opinions expressed here are those of the author and do not necessarily represent policy of the NSF.

2 There have been a number of activities focusing on chemistry and computers. The following list gives the flavor and constitutes a trace to earlier literature:

a. NAS Report on May 1970 Conference on Computational Support for Theoretical Chemistry. Available from NAS, Washington, D.C. includes a reprint of a good overview article on the ARPANET.

b. Proceedings of the NIU Conference on Computers in Chemical Education and Research, July 1971. The sixty papers span all the important areas of computers in chemistry. Available from Chemistry Department, NIU, DeKalb, Illinois, $10.00 postpaid.

c. ACM SIGCUE Bulletin, December 1971, pp. 321-6 lists 9 conferences of national scope relevant to chemistry and computers.

d. 1971-2 Annual Report, Division of Chemistry and Chemical Technology, NRC-NAS, summarizes three years of activity of the DCCT Committee on Computers in Chemistry.

e. Announcement of International Conference on Computers in Chemical Research and Education to be held in Yugoslavia, July 1973. Professor D. Hadzi, Ljubljana, Yugoslavia is the local organizer.

f. Two NAS reports will be published early in 1973. A report of the April 8, 1972, Conference on Computational Needs and Resources in Crystallography, and the report of the Feasibility Study of a National Center for Computational Chemistry.

3 "Special Report on Chemical Education Today," October 9, 1972, Chemical and Engineering News.

4 "Computing in Higher Education 1971: Successes & Prospects" EDUCOM, Princeton, New Jersey, pp. 11-14; EDUCOM Bulletin, Vol. 6, No. 3. FALL 1971, pp. 10-13.

5 On January 23, 1973 Dr. Walter Hamilton passed away. Questions regarding the Protein Structure Data Base project should be addressed to Dr. Edgar Meyer at Department of Chemistry, Texas A & M University.

6 National Environmental Information Symposium, September 24-27, 1972, EPA, Cincinnati, Ohio.

7 Golden Rule: He who has the gold . . . rules.

8 The NSF has published brochure NSF 72-16 entitled "Expanded Research Program Relative to a National Science Computer Network." Those concerned with local, statewide or regional computing should note the statement on page 5, "This announcement is not aimed at developing programs to enhance the computation or information capabilities on an individual campus or even in regional areas. Rather, it is directed to providing specific information as to the feasibility of a National Science Computer Network." Since this is an evolving program, which is being further defined through the EDUCOM series of General Working Seminars on the National Science Computer Network, further and current information may be obtained from the Office of Computing Activities or the Office of Science Information Service.

The Ce/NCoReL Study

Walter Sedelow, Professor of Sociology and Computer Science
University of Kansas

The goals of the Ce/NCoReL Study and some of the conclusions coming out of it are the basic topics in this presentation. My wife, Sally Sedelow and I have been working on the study for the past year and a half. Given the limitations of space here and the range of issues discussed in that study[1], this presentation can be no more than a kind of sampling or hors d'oeuvre session. The summary of the workshop session and the concluding panel discussion amplifies some details of the report and takes up additional matters.

The National Science Foundation provided funding for the Study of the concept of a possible Center or Network for Computational Research on Language (Ce/NCoReL). The National Science Foundation divisions that supported this research included not only the Office of Computing Activities but also, through that office, the social sciences division. We have met with scientists, administrators, and scholars in at least twenty states in the course of exploring this idea. An exceptionally broad definition of language was used for this concept study. However, there is no necessary implication that the definition of language which was used for these purposes of exploration will prove to be the definition of language that would be used if there comes to be a national center or network for computational research on language. At least in the early stages of such a center or network's functioning, a much narrower definition of language might be used. But, for the initial study at least, it was important to define the topic widely. In that connection one of the interesting and valuable aspects of this investigation was the emergence of common grounds for research which different types of language researchers in fact do share, sometimes without realizing it. Evidence of some significant progress toward research coherence is noted in this particular exploration of a possible center or network using the computer. Certainly, the opportunities made possible by technology and new policies are helping to bring some researchers together who hadn't considered interaction previously, at least not in the ways they are now thinking about them.

Languages and symbol system research included in the study have been drawn from many fields: anthropology, archeology, and art history; biology and biophysics; communications science, computer science, and engineering; a wide variety of humanities departments including Chinese, English, French, German, Greek, Latin, and Russian; information science, and library science; linguistics; mathematics; music; the neural sciences, psychology, psychiatry; social relations, and sociology.

The open notion of language which we employed for the purpose of exploring whether it would be attractive to have a national computer capability for research in this area was, essentially, "any kind of symbol system behavior." Language was defined as natural language (e.g., French or German) or manifestations of it in literature: language from the standpoint of the professional linguist sometimes concerned with rather abstract theories and models of the nature of language; language from the standpoint of the computer scientist where the emphasis might be on programming languages or on formal languages; or language from the standpoint of a person concerned with mathematical notation and the types of logical language. We were, in every case, interested in what kinds of useful interplay, possibly even synergism, there might be among different types of language researchers if there were a national research facility for language study using computer technology (network technology particularly). Such a center would involve any type of language researcher.

A further range of questions in the study deal with the relationship between the emerging studies of the functioning of the central nervous system vis-a-vis the studies of language. For example, would it be appropriate for the facilitation of research on language to also provide common resources for people who are interested in the study of the ways in which the human central nervous system acts as an information processor? Research-related agreements which President Nixon has signed with the USSR, call attention to various areas of prospective collaborative research and give special mention to the area of language study. Interestingly enough, there's a specific reference to the importance of the interaction between language research and brain research (or central nervous system research). In the USSR there is a particularly strong interest in statistical studies of language, many of which are relatively unknown in the United States. Scholars at the University of Michigan, notably Professor Richard Bailey, have just begun to make these studies known in the USA.

The results of this study are contained in the volume, *Language Research and the Computer*, which is being distributed to scholars who responded to questions or participated in four small conferences held on these topics. Additional copies of the report may be obtained by writing to us at The University of Kansas.

The report begins with a discussion of the Ce/NCoReL concept, i.e., the general ideas that guided our explorations. A second section includes an assessment of the research implications and emphases which, for different people within this broadly defined language research community, might be facilitated through such a center or network. A primary effort of the study was to elicit from researchers the specific kinds of things they would like to have done rather than to impose some notion of what should be done. We also sought to identify patterns of capabilities desired by particular people and types of researchers. The very broad and ample definition of language facilitated such open-mindedness, and prevented premature closure.

The report continues with a chapter specifically concerned with the software requirements necessary to meet the stated needs of these types of researchers. Another section is devoted to hardware: current demands for hardware capability, and particular properties in hardware which are not currently available but which would flow from a user orientation in providing computer facilities for researchers. The study first asks, "What do you want to do using a computer or kindred device? What properties should it have?", and only then asks whether currently available hardware meets those requirements, or whether there are certain sorts of hardware developments which it might be helpful to facilitate, or at least to explore in order to serve a certain type of user. Instead of having the hardware come first, the software second, and the user third, we started with the user.

Another chapter in the study is concerned with various ideas, advanced by letter, by conversation, and in conferences, for possible methods of organizing a national capability for computer based language research. There are several precedents for the organization of computer resources combined with network technology. The ARPANET is one model for the development of a distributed national laboratory or research facility. Project INTREX provides a model of a network designed to serve an online intellectual community. We are thinking more specifically about a dynamic capability for a particular type of researcher. A hypothetical model with special appeal to language researchers is the distributed laboratory or research facility in which people across the country could actively participate much as they might in a single building, by substituting a communications technology, particularly through and over computers and networks, for the physical movement of people.

A new type of national research laboratory, a multi-institutional, collaborative effort to articulate geographically separated research undertakings through a distributed center or network, is well suited to language research. In doing our exploring, we continuously heard about the importance of enhanced communication among the researchers. People didn't know of relevant work that was being done by others. In many instances, researchers wouldn't know of other's work even when it was published, because of the diffused character of language research and the great hererogeneity of types of researchers and journals. Even when work is widely diffused through publication, researchers often do share common problems and can make use of common research instrumentation, such as the computer.

Language Research and the Computer concludes with a discussion of recommendations for a possible center or network, and then amplifies some possibilities for research and applications in a series of appendices. One appendix, "Knowledge Systemics," is concerned with the implications for teaching at the university level of the utilization of computer technology in networks for research in language. Several provocative and interesting ideas are proposed concerning the possible relationship between a new approach to some aspects of education and the more effective response to certain major social problems, particularly problems which invite a systems approach, e.g., some of the problems of human beings in their environments. Additional appendices contain two detailed discussions of the current state of applications of computers in humanities research and in the fine arts.

As a result of our discovery of so much complementarity of activity among researchers who nonetheless do not share a common literature, extensive bibliography on the literature of networking as it bears on the use of computers for research has been included in the report. Another bibliography, which is concerned with attempts to measure language, follows. There is a great deal of research, although very badly scattered, on language as an object of study in a very narrow scientific sense.

Other bibliographies deal with literature on: information retrieval and computational linguistics, as distinct from the more specific verbal measures inventory; systems research, a topic of real importance not only to people participating in creating a system, but to an increasing number of people who will be oriented toward systems and cybernetics concepts in the substantive parts of the language research; and paralinguistic research including sociolinguistics, psycholinguistics, anthropological linguistics, etc. There exists a high social urgency for establishing a Ce/NCoReL because all areas of scientific and scholarly research involve the utilization of language to cope with research results.

The study of language could also contribute heavily to the understanding of the research process itself. To enhance research on a national level, effective communication must be achieved between researchers who start with different vocabularies and different models. Because of those different vocabularies and models, researchers may not see the common elements or transferrable theories in other's work. Language research also has a heavy bearing on information retrieval and information systems as they relate to science information. Some computer characteristics would be well adapted for studying language strongs.

First and most general among the research implications of a distributed national center for language research, is the importance of man's being able to model his environment more effectively to solve problems. Historically, an understanding of the components of complex social events has been especially hindered by the comparative inattention to lan age factors in social interaction. Since language looms so large in shaping behavior, it is curious that more attention has not been paid to language variables by sociologists, as well as by some other social scientists. We contend that the lack of attention is primarily due to the lack of the necessary instrumentation for effectively doing that kind of research in a rigorous way.

The true scope of the computer has probably not yet been located. It took quite a while before people locked onto ways to use the microscope. The occasional mismatch between the capability latent in apparatus and its usefulness is well illustrated by some pretty entertaining data about Leeuwenhoek and early microscopes — the sorts of things he used them for. *The computer will probably turn out to be the piece of instrumentation which*, like the telescope helping us understand the microscopic world and the microscope helping us examine the very small, *is meant for studying human symbolic processing.* We may experience a great sort of intellectual revolution as we shift more and more away from social scientists and humanists talking about 'ideas' to doing research on what's actually there language.

The computer is also critically important in looking at the way in which language is produced and processed. Coordinated central nervous system research at the University of California at San Diego Medical School is one example of this type of research. In trying to develop a coherent and comprehensive research approach to neural phenomena, researchers at UCSD are breaking out from the models that were produced by neurochemists, neuroanatomists, and neurophysiologists.

In the field of international relations, language research may aid in developing solutions to the tremendous difficulties resulting from the rising revolution of expectations around the world. While new high interdependency levels have been generated in part by technology, we still have the older illusion of social distance which is fostered by nothing so much as different 'natural' languages.

To address this difficulty we must move beyond the machine translation emphasis which has been based on the written word, to a machine-oriented (machine-aided) oral translation system. We might hope to have, with the aid of the computer, a capability for every man to pick up the telephone and talk with people whose language range is different from his. We need something on the order of a machine-assisted capability to provide what we now have in the 'manual' mode in the simultaneous translation schemes that are used at places like the United Nations.

If you have *m*achine-aided *o*ral *t*ranslation *s*ystems and they are very good, you not only have, acronymically, MOTS, but you have *BONS MOTS*. It seems to me that, for people who are interested in language, the use of a computer to produce *BONS MOTS* is a very, very good thing, indeed.

REFERENCE

1 *Language Research and the Computer*, Walter and [illegible] Sedelow, University of Kansas, 1972.

Networks in Economics

Sanford V. Berg, Assistant Professor of Economics
University of Florida

A number of essentially independent networking efforts are underway to serve the approximately 40,000 professional economists in the United States. Scientific societies, research institutes, government agencies, and universities are engaged in the computerization of bibliographical and biographical information, the development and dissemination of data banks, and the creation of computer systems for economists. It is essential that flexibility be built into these efforts to facilitate the evolution towards systems that will be compatible within the discipline and with systems developed in related social sciences. No coordinating mechanism is in sight.

Some standard results from international trade theory are applied to networking efforts in Economics. The presence of a low cost network could be viewed as a significant drop in "transportation costs"; alternatively, an easing of university regulations regarding outside computer usage could be viewed as a reduction in "tariffs" or an increase in "quotas" (permitted "imports"). In either case, dislocations may follow; but in general the resulting specialization and increased division of labor result in more efficient use of resources. Problems which will require coordinating mechanisms are examined in the context of what we have learned from past experience in international trade and finance.

As scientists who study resource allocation, economists have done remarkably little to improve the efficiency of their research efforts. Few have considered the potential impact of a relatively low cost computer network on resource use within the discipline. Certainly, other sectors of the economy have obtained productivity increases through the use of such capital intensive means of production. This paper examines how a computer network could pool users, systems management, hardware, software, and data to create a more efficient research environment in Economics. Concepts from international trade theory are used to show how institutions (nations) can benefit from trading the basic components of research through computer networks.[1]

USERS AND RESEARCH METHODOLOGY

The role of computers in Economics has been described elsewhere by Ruggles (1972), who identified three major factors which shape the methodology of economic research: the changing scope of economics, the industrial organization of the profession, and the evolution of information processing technology. With respect to the first factor, social and economic events such as the formation of giant trusts, the Great

MEMBERSHIP IN THE ALLIED SOCIAL SICENCE ASSOCIATIONS – 1969

American Economics Association	17,000
American Statistical Association	10,500
Biometric Society (ENAR and WNAR)	1,600
Institute of Mathematical Statistics	3,000
Econometric Society	3,500
American Agricultural Economics Association	4,100
Catholic Economic Association	400
Industrial Relations Research Association	3,200
Association for Education in International Business	250
American Marketing Association	18,000

Figure 1

RELATED SOCIETIES

International Association for Research in Income and Wealth
American Real Estate and Urban Economics Association
Association for Evolutionary Economics
Omicron Delta Epsilon
Association for Comparative Economics
American Risk and Insurance Association
Regional Science Association
Joint Council on Economic Education
Association for the Study of the Grants Economy
Public Choice Society
National Association of Business Economists
Economic History Association
Association for the Study of Soviet-type Economics
Operations Research Society of America
Institute of Management Sciences
Southern Economic Association
Western Economic Association
Midwest Economic Association
Union for Radical Political Economics

Figure 2

Depression, and the problems of developing countries and urban areas have stimulated new lines of thought within the discipline and in the last two cases, new fields of inquiry. The relatively new area of environmental economics illustrates how the pressure of the outside world changes analytical frameworks and data requirements. The second major factor, industrial organization, refers to employment patterns of the profession, with about three-fourths of those with Ph.D.'s in Economics finding jobs in colleges and universities. Academic economists have tended to work alone, often on long-range projects, while economists employed at various levels of government generally focus on the solution of problems: here the team approach is complemented by access to huge data resources.

The impact of the computer on economic research is a well documented example of how tools alter research patterns (as telescope and microscopes influenced astronomers and biologists). In the past, aggregation was facilitated by the introduction of punchcard technology. Now the analysis of disaggregated information has become possible with larger computer memories and improved input-output devices such as disks. Data processing, information retrieval, statistical calculations, and simulations characterize the activities of a growing number of researchers. Nevertheless, the diversity of needs, as reflected in degree of mathematical and statistical sophistication, type of modeling efforts, and data availability implies that a network effort will not be financially viable unless a broad range of capabilities is available.

SYSTEMS MANAGEMENT

Other disciplines have the resources and institutional backing which permit a wide range of communications activities: this may be contrasted with Economics, which is a fairly decentralized discipline. The major scientific society, the American Economic Association, has over 20,000 members and about 10,000 institutional subscribers to journals, but a number of specialized and regional societies also carry on important communications activities. (See Figure 1). If members of related societies are added, this number would expand. (See Figure 2). Government economists and business economists may not identify with these specialized groups, but they utilize the same analytic techniques and data bases. Thus, the relevant research population is probably on the order of 40,000 to 45,000 individuals in the United States.

The fragmentation illustrated in Figures 1 and 2 carries over into the other organizational affiliations of researchers. One implication of this degree of decentralization is that the management of communications systems is still in its infancy, although there are some centers which are likely candidates for key roles in any future network. The technical expertise at these centers would benefit a network, but perhaps their major contribution would be through organizational talent developed over the years. The management of risk, planning, coordination, and routine supervision are functions that characterize any large-scale activity, and a networking effort needs individuals who are involved in each of these phases. The specialized information centers that now serve different populations within economics will be described in connection with the software systems and data bases around which they formed. Suffice it to note that centers tend to arise because of the scale economies involved in some types of computer-related research. Systems need to be marketed and researchers taught how to use particular tools and techniques, even within the local research community. Other "overhead costs" stem from the need to prepare grant proposals, the cost of administrative record-keeping, the salaries of personnel involved in testing and evaluating the system, and "hot line" counseling regarding hardware, software, data, and error corrections. The scale economies in such organizations can be significant in terms of spreading overhead costs over a large number of activities, increased specialization, and a more efficient division of labor between men and machines.

HARDWARE

The physical linking up of computers is central to the network concept, although one could argue that it does not pose the greatest problems for Economics. Despite differences among computers in terms of manufacturers, models, and configurations, communications lines are now serving the research needs of some economists. Government economists have access to the federal telephone system, and the AT&T network permits access to a significant number of systems. The problems appear to involve limited financial resources and marketing efforts.

Of course, more sophisticated use of the networking concept would require a great deal of work on data documentation and interchange and computer interface standards. But I believe that such work would be stimulated more through the development of appropriate institutions than through the improvement of physical facilities. Physical networks involving terminals and communications lines exist. The costs are high, so a utility type network which could achieve scale economies and other advantages of large operations would certainly increase interaction among universities. However, the fundamental human networks are perhaps of greater importance. These invisible colleges would be improved through physical networks, but their character would depend very much on the institutional framework in which they evolve.

Some of the existing physical networks are essentially commercial in nature. And in the absence of a discipline-oriented network, these could bring many advantages to researchers engaged in applied areas, and who use standard data bases. Although many business firms have their own personnel who do economic forecasting, economics consulting firms operate physical networks and/or provide information for many of the largest firms. A recent *Wall Street Journal* article by Lindley Clark (September 20, 1972) stated that there were fewer than a dozen significant consulting firms, with the largest of these having only two to three million dollars in sales. Edie Economics, Townsend-Greenspan and

Co. Inc., Rinfret-Boston Associates, Data Resources Inc., and Chase Econometrics Associates are examples of such firms. The latter two not only provide estimates, but have models and data on-line and available for customers who wish to use their own assumptions in making forecasts. General Electric's Management Analysis Projection System (MAP) and Rapidata's Program Language for Economic Analysis (PLEA) are examples of systems which are used for time-sharing purposes. Customers of these last two systems may use the NBER Time Series Data Base which has over 1500 of the major time series on United States economic activity. (See Boschan, 1972).

When competitive networks appear to be financially viable, it would be difficult for an economist to argue in favor of a monopoly network, unless the "public utility" aspects of the network outweighed the resource misallocation that accompanies monopoly power. That is, if the scale economies are substantial, or if coordination yields system savings, a regulated monopoly could be efficient. The technical considerations are still not necessarily dominant, since the advancement of economic knowledge might be speeded through institutional changes accompanying the research utility. Academic researchers have a different set of needs than business economists, in terms of the degree of mathematical sophistication and emphasis on behavioral relationships. The budget constraint is also substantial, particularly for the use of the computer in teaching economics, since outside funding is drying up. Especially for small colleges, which may not have in-house interactive capabilities at present, networking which focuses on the teaching of economics may be very efficient. The same system could also be used by researchers at those institutions for simulation, estimation, and data analysis.

SOFTWARE

Software is not a central problem either, from the standpoint of the average researcher – although advances of the state-of-the-art which are embodied in readily usable programs will always be of benefit to the cutting edge of the research community. The basic statistical packages have provided batch capability for years, with Datatext, SPSS, and the BMD packages widely available. Interactive systems are now demonstrating the usefulness of exploratory work in data analysis. For example, TROLL is being further developed at the NBER Computer Research Center in Economics and Management Science so as to include cross section analysis as well as its present capabilities for time series analysis. Similarly, the Cambridge Project is developing a "Consistent System" of interactive computing tools. One methodological problem with interactive modeling is that data massage efforts too often involve not retaining a control population on which to test the resulting hypothesized behavioral relationships.

Another type of software package involves models which are already specified, and thus ready for re-estimation or simulation. The use of large-scale macroeconomic models (some with hundreds of equations) is widespread. The Brookings-SSRC, Wharton, and MIT-FRB models have contributed to our understanding of macroeconomic processes and sectoral interrelationships. And the availability of a network makes it possible to pick up where the initial model-builders left off. Equations could be re-specified, models subjected to stochastic shocks, and the results evaluated much more easily if the set-up costs of additional users could be spared. Large-scale microanalytic models, such as MASH (Microanalytic Simulation of Households), under development at the Urban Institute, also could be utilized from other locations (see Guthrie, et al, 1972).

Clearly, any networking effort would have to ensure the integrity of software (and data) file by limiting access. After all, an economist could work for months on a model, and he would wish to withhold it from general use until he could take advantage of the fruits of his labors. Patents provide a similar incentive for inventors. After a time, however, the benefits of wider availability outweigh the incremental addition to incentives for developmental work. It still will be up to a researcher to decide just when he will "release" a model, and proprietary researchers may never do this. However, for most economists working on basic research questions, release is in their own interests, and will occur in the process of communicating the substantive results of his efforts.[2]

It should be noted that special purpose programs have been developed at a number of major universities, and each is being generalized to do bigger and better things. Whether this activity is essentially duplication, or a response to diverse hardware requirements and data needs depends on one's perspective. Certainly, any networking effort cannot ignore either the drive for autonomy that such efforts represent or the desirability of multiple centers of initiative within the discipline. The economist would probably let the "research market" judge which system has the greatest potential (especially in terms of future funding). Thus, the existence of a low cost network would make it easier for a university administrator to allocate internal computer funds more efficiently, (that is, away from programming efforts that are truly duplicative), and observed demand would aid NSF and other agencies in making funding decisions.

DATA

The lack of satisfactory empirical research is often credited to inadequate data bases. After all, Economics is not typically an experimental science; rather, like meteorologists, we infer patterns from data which nature provides. The failure to achieve convincing tests of significant hypotheses about individual and social phenomena may be ". . . due to the inadequacy of the evidence brought to bear rather than to any great deficiency in our statistical, mathematical, and computer tools. In other words, data and data-related problems are at the heart of the matter" (Orcutt, 1970). Ed Kuh, in his comments on Orcutt's conclusions places the blame on the complexity of the behavior under examination, as well as the role of value judgments and lack of experimentation. Recently, a number of social experiments have begun, involving income maintenance, voucher systems for schools, and the location of medical services. As government policies begin to be geared to experimentation and evaluation, such data should become more important to the research community.[3]

The data generating sector has been well discussed by other economists who have noted our dependence upon government agencies: United States federal, state, and local governments, as well as international agencies, such as the United Nations, the Organization for Economic Cooperation and Development, and the International Monetary Fund. All the above publish volumes of statistics from machine-readable files, yet researchers are continually having this or that series keypunched in piecemeal fashion.

The recent availability of large-scale data bases, such as the public use sample of the 1960 and 1970 censuses, has stimulated some cooperative efforts, since acquisition and processing costs for such files can be enormous. Similarly, the individual observations from the Current Population Survey (CPS), the source of unemployment statistics for the United States, are now available in machine-readable form. In both these cases the need for data experts becomes clear. Without such an intermediary, the individual researcher finds himself faced with a raft of problems, from error correction and information on sample biases to interpretation of data. In the case of CPS, the Urban Institute has taken the initiative by obtaining tapes created from surveys taken in the 1960's. Because of the *ad hoc* ways in which Census created the tapes, a number of formatting and other problems must be solved before this potentially rich data base can be widely used.

Similarly, DUALabs, a private nonprofit institution, was established (with Ford funding) to assist a consortium of universities in the acquisition and use of Census material — particularly, in providing derivative tapes and packing them more efficiently. With NSF aid and together with the Center for Research Libraries, it has created a Clearinghouse and Laboratory for Census Data (CLCD). The CLCD performs many of the systems management functions described earlier, including consulting, group training, and publications (*Data Access News* and *Technical Bulletins*). In addition, it can serve in an advisory capacity to Census and as the key node of an information system on users and uses of census tapes. DUALabs plans to create indexes and catalogues which will aid researchers in the area. These will not only be for retrospective searching, but will announce research in progress. The sharing of resources through the consortium has made possible research which otherwise could never have occurred.

There are a number of examples of university-based research efforts which have focused on particular data bases. The Survey Research Center at the University of Michigan has a data bank from its Economic Behavior Program, with the content ranging from detailed financial information collected by the annual Survey of Consumer Finances to quarterly economic attitude surveys and special studies. The OSIRIS system is used for the analysis of these data. The University of Wisconsin has also been engaged in the development of software for social science applications. In particular, SEOSYS was developed for the retrieval of information from the Survey of Economic Opportunity, and a more general Social Science Information Management System (SIMS) is under development. Work at the Brookings Institution using Internal Revenue Service tax files is another example of a large-scale research effort at a nonprofit institution.

Finally, there are a number of proprietary data bases of interest to economists, including those collected by Dun and Bradstreet, the Industrial Conference Board, F.W. Dodge Corporation, and McGraw Hill. The diversity of interests among researchers makes it difficult to generalize regarding data needs and possible trends. However, some things can be said about institutions now operating in this area.

Data archives serve a variety of functions related to research, research training, and data dissemination. Examples of major data repositories include the International Data Library and Reference Service at Berkeley, which has focused on Latin American and Asian survey data, and the Inter-University Consortium for Political Research at Michigan, which has collected survey data on United States politics as well as Congressional roll calls in machine-readable form. The major archives in Economics are closely related to on-going research at various institutions, with the discipline currently lacking a generalized data consortium. The National Science Foundation has funded a quarterly newsletter which announces acquisitions of data archives which would be of interest to social science researchers. Since 1971, *s s data* has published brief abstracts of data bases, and the editor has expressed an interest in data storage and retrieval systems and in computer related research techniques. However, another newsletter, the *SIGSOC Bulletin* (published by the Special Interest Group on the Social and Behavioral Science Computing of the Association for Computing Machinery) also covers the latter issues.

A number of other groups are also working on data problems. For example, the Urban and Regional Information System Association (URISA) consists of government employees, consultants, and academicians interested in information systems and small area data.[4] The various urban institutes around the country have researchers with similar interests, and of course, some users of Census data would fall into this group. The American Statistical Association has several divisions which are interested in data problems, including the Social Statistics Section and the Business and Economics Section. The establishment of a Statistical Computing Section in 1971 reflects the increasingly interdisciplinary nature of computer problems. Also, the National Bureau of Economic Research has established an Interinstitutional Conference on the Computer in Economic and Social Research which brings researchers together for informal workshops and formal conferences. Whether this organization can serve as a fulcrum for change within Economics is as yet unclear.

AN APPLICATION OF TRADE THEORY

Without wishing to force the trade analogy, the current absence of computer networking in the economics discipline seems to be a perfect example of autarky: many nations (universities), trying to be completely self-sufficient in the production of goods (knowledge within particular disciplines),[5] as though transportation costs were so great that trade between nations is not feasible. Then introduce the analogue to speedy, low cost shipping, computer networking. Universities, like nations, could mistakenly react by setting up tariffs reducing the volume of trade, in order to avoid dislocations caused by instantaneous free trade.[6]

In addition, university administrators would resist what might be perceived as "colonial exploitation." A department may feel threatened at not having its own in-house computing facility. It may believe itself doomed to remain on the periphery of methodological developments in the field if another university in the region has all the supporting personnel and data on its premises. Designers of any computing network should anticipate such reservations within departments. The theory that "trade follows the flag" (and vice versa) conjures up visions of foreign domination resulting from trade. If departmental fiefdoms will not cooperate, the university itself will have a difficult time agreeing upon some resource sharing arrangements with other universities.

An alternative theory of trade is more widely accepted today: the theory of comparative advantage. Basically, it takes the resource base of nations as given, and explains what countries will export and import in terms of what products use most intensively the input which they have in relative abundance. Let us take two products from Economics to illustrate what the theory has to say. One field, consumer behavior, might require minimal computing power but a substantial amount of personal services (survey data collection, data entry, correction, counseling); (in point of fact, the storage requirements for consumer studies may be substantial). The other field, simulation models of the economy, draws from a data bank already in existence but requires substantial computing power to facilitate the analysis of alternative specifications, estimations of effects of various policies, and simulations with different types of stochastic shocks. Assume that there are two universities, each of which attempts to remain self-sufficient. If one university already has the large computer, and the other already has additional budget lines for personnel, there would be room for a deal which could make both departments better off. Institutionally speaking, this switch (sharing of resources) is easier than if one university agreed to handle "all" economics computing while the other handled all "all" biology computing.

There are a number of conditions which invalidate the basic conclusion that movement towards "free trade" (and greater specialization within countries) expand the potential output of all nations together. First, there may be distortions in the existing pattern of prices. Some universities may have had computer facilities subsidized in the past, and universities may continue this in the future. If this advantage is financial, rather than reflecting real resources, then adjustments might be inefficient. Similarly, there may be immobilities which prevent the movement of personnel and computers to where they can be used most efficiently. For example, if a univeristy has purchased a machine or has a long-term contract, it will not find purchasing computer time from another university an attractive idea.

The noneconomist will more easily recognize the "infant industry" argument found in connection with trade policies of developing countries. These nations do not want to take the initial resource endowments as given; rather they wish to restructure their economies through export promotion or import substitution policies. Thus barriers may be raised to permit an industry to gain a foothold, when exposure to external competition would doom the industry. Of course, some infant industries never grow up, so protection may generate no benefits – other than those which accrue to the protected industry. Resources are not released for the development of other sectors.

A key lesson to learn from trade theory is that the form which the protection takes affects efficiency. For example, a university which feared for the financial viability of its computer operations, might attempt to put quotas on the use of outside computing facilities. Such a policy has some very detrimental side effects. First, who is to be allocated the outside time? Whichever researchers are granted the quota have an advantage over other researchers from that department. If a first come-first served policy prevails, then a rush to use the super foreign facility will occur, with obvious inefficiencies accompanying the artificial speed-up in activity. The same limitation on foreign computing could be attained through a tariff (or tax) on outside usage. The users who place the highest value on the outside facilities will be willing to bear the higher real price. In addition, as demand for such computing grows, it will be supplied by the outside facility. If this threatens the financial viability of the "home" computer, the "tariff" could be revised upwards, but at least the administration would have a gauge of the perceived value of the outside facility.

The notion of financial viability brings up the problem of under-employment of the home facility. This situation is analogous to attempts of governments to cut back imports (through higher tariffs) to expand employment at home in a self-defeating "beggar my neighbor" policy. When other nations retaliate, employment in export industries falls – leaving everyone worse off. Nations pay higher prices for products, and overall employment is not improved. Clearly, a better policy would involve using internal policies to expand utilization of existing resources.

International Economics can also suggest techniques and institutions for facilitating the adjustment process. The formation of regional customs unions is not as threatening as the establishment of free trade with all nations, particularly if the production structures of the participants are complementary, rather than competitive. There may be trade diversion from countries outside the customs union, but there is also trade creation – and improved division of labor within the common market. There are also a number of dynamic advantages from such arrangements. Thus, regional consortia within Economics (and other disciplines) may prove to be one way of moving to more efficient production of scientific knowledge.

In addition, some national cooperative arrangements may prove to be useful for cases when all participants understand the advantages of specialization in production. GATT (General Agreements on Tariffs and Trade) set institutional rules limiting artificial barriers, and a similar institution would help delineate proper and improper (inefficient) ways of "protecting" departments. The basically liberalizing influence would be useful. Also, an institution corresponding to the International Monetary Fund would have a role to play. That is, some units of account need to be developed so that "trade" can grow efficiently. Since interinstitutional payments would put a strain on finances, some accounting procedure could be developed which facilitated multilateral, rather than bilateral, balancing of credits and debits. Perhaps a university undergoing a severe adjustment would be permitted to borrow for a period of time. Three problems (again analogous to those in balance of payments) arise: adequacy of existing interinstitutional arrangements, ways to facilitate the future growth of transactions, and ways to facilitate the adjustment process. As key personnel shift affiliations, as computer technology changes the economics of networking, as new fields grow and old ones decline in terms of grants or networking requirements, universities will need to be able to adjust and make the proper investments in technologies and personnel. One should not underestimate these problems, but neither should they prevent "institution-building" in this area.[7]

CONCLUSIONS

The limitations of present computer networks are beginning to be acknowledged by economists whose research needs are inadequately met. In response to the crisis in communications, the American Economic Association has established a committee to investigate information dissemination practices and to recommend improvements in those practices, with particular emphasis on the *speed* of availability of new information, the *accessibility* of that information to the research community, and *compatibility* with systems in related disciplines. It is clear that economists should begin to design and plan a more comprehensive and modern communications network for the discipline. At present, we have barely scratched the surface of determining communication requirements in Economics. Achieving consensus with respect to priorities will involve at least three types of activities (1) examinations of existing computer and communications networks in economics, (2) evaluations of innovations which are beginning to affect these networks, and (3) meetings with researchers who are working in the area.

Studies are needed to provide for Economics some basic technical and behavioral data before movement can be made towards a more integrated information system. Decision-makers in scientific societies, major universities, and research institutions need a clear picture of current information networks in Economics. Thus, studies of the functioning of information channels, including the causes and effects of inefficiencies in empirical research are needed to ascertain the potential market for a computer network. Pilot projects are also essential if a networking effort is to truly change research patterns through changes in the relative prices of research inputs.

We also need to survey and evaluate the major innovations affecting communication in Economics. A number of independent efforts are beginning to change information-gathering practices within sub-fields in Economics, but until recently, we have not tried to compare these fragmentary advancements nor formulate a picture of how economists might build a unified information system, using the strong points of current development efforts.[8] (See Appendix B for a brief survey of some activities).

Finally, although meetings are sometimes fruitless exercises, they facilitate the exchange of information described above. A series of workshops should be held to discuss existing channels, evaluate innovations, and establish a minimal set of communications requirements for the discipline. One such exploratory workshop was held by the NBER's new Conference on the Computer on June 4, 1971 to examine innovations in media and computer technology (see Beharic, 1972). This meeting served a catalytic role by focusing attention on fragmented developmental efforts. A follow-up working session on the *Journal of Economic Literature*, the production of handbooks, the dissemination of workshop papers, and other topics was held in September 1971. Some consensus with respect to communications priorities, was achieved, although networking, *per se*, was not emphasized.

In a recent survey, Sadowsky (1972) identified the major problems within social science computing as follows:

1. inadequate standards for data documentation and problems of data transfer;
2. the low level of computational knowledge among social scientists and the lack of adequate training available;
3. the slow rate of diffusion of computing innovations into social science computing, especially in government;
4. problems of program inaccuracy, documentation and transfer;
5. lack of adequate software tools for many types of complex processing operations; and
6. the low level of professionalism in social science computing activities generally.

He concluded that "Initially, the existence of network communication links will be far more important than how the network is implemented," (p. 882). The main theme of this paper is not exactly in line with his conclusion. If one emphasizes how the existence of any network will affect current research patterns, Sadowsky is no doubt correct. However, the institutional framework does matter in the long run – and the observations from trade theory were meant to illustrate how rules, organizations, and shared ideals can have an impact on efficiency in the use of computer resources and manpower. For example, in 1970 Holt recommended the creation of a National Series of Information Centers, to be coordinated (or held together) by a National Institute on Information

Services and Standards. Besides trying to avoid duplication of effort through communications activities, the Institute would engage in program validation, the establishment of computer networking standards, and the creation of research directories. If existing centers, such as DUALabs, the Urban Institute, the Brookings Institution, the National Bureau of Economic Research, and various university-based organizations, evolved into a comprehensive network, patterns of resource allocation would be very different than if the computer network left institutional arrangements unchanged. Furthermore, if the network does lower costs, and methods of financing its use affects the location and extent of computing, there will be severe dislocations unless advance planning eases the strains. Without adequate institutional cooperation, Economics could find itself with several duplicative networks – none of which adequately meets the needs of economists.

REFERENCES

[1] George Sadowsky helped clarify some of the ideas presented here, but he should not be held responsible for the final product.

[2] Since 1969 the National Science Foundation has funded a National Program Library and Central Program Inventory Service for the Social Sciences (NPL/CPIS). Abstracts of computer programs are available to permit potential users to make a preliminary judgment regarding the acquisition of a program source code. Information on the program's function, usage, source language, hardware configuration, and distribution source is provided, although the extent to which the service meets computing needs is not clear. For more information, contact: NPL/CPIS, Social Science Building, University of Wisconsin, Madison, Wisconsin 43706.

[3] See Morton (1972) for a survey of some data sources in Economics.

[4] For further information, contact Donald S. Luria, Treasurer URISA, 901 Elizabeth Avenue, Suite 100, Charlotte, North Carolina 28204.

[5] Of course, researchers do use the output of research conducted elsewhere in the production of their own research – science is a cumulative process. However, that production may not be as efficient as it could be if a further division of labor could be devised. Note that for a small minority of researchers, the benefits of working with scientists at other universities or using systems developed elsewhere are great enough to outweigh the costs of doing so.

[6] Fortunately, as well as realistically, universities do not have to fear having to cope with such a radical change overnight. Even if the network were to appear tomorrow, researchers would not be able to utilize it – perhaps it would take a decade before it became a standard research tool.

[7] Note that we have focused on using trade *theory* to determine commercial *policy*. Constraints and deviations from assumptions complicate the policy issues. Also, the entire "balance of payments" of a particular computer center could have a deficit on the trade of computer services. Some universities may find it cheaper to have no in-house computing capability since access to other facilities may be more economic given the structure of demand. Some National Science Foundation "Special Drawing Rights" would smooth the adjustment process and provide incentives for more efficient use of our computer resources.

If universities did want to balance "trade," a "floating exchange rate" would be the theoretical ideal – permitting all computers to be used to capacity. If one center became overloaded with demand, prices there should rise causing the amount of computer time demanded to fall. Users are responsive to price and would substitute more labor intensive processes for computer time. Eventually the higher prices would justify additional investment in facilities. Determination of the location of such facilities is no simple matter.

[8] Although the stress has been on data, the development of machine-readable information files has had a similar pattern. At present there are a number of efforts for indexing references to the economic literature, or abstracting articles. Such information bases will be just as important in the development of information systems within the disciplines as traditional data sources. In fact, the development of such files will be stimulated by networks, since the printed form in which such indexes now appear may be superceded or supplemented by the on-line information. Here, the cooperation with reference liberaties will be essential. And again, interdisciplinary cooperation is essential to ensure the compatibility of resulting systems.

Library Networks

Frederick G. Kilgour, Director
Ohio College Library Center

Library and information networks already exist. A library network may be a regional net of a group of participating libraries or a larger net consisting of a group of regions. Information networks are different from library networks in that an information network usually employs a data base that a single institution has constructed and to which queries are put.

Library networks have two major objectives. First, a library network should attempt to make resources throughout the network area available to users of each participating library. Second, the network should decelerate the rate of rise of per-unit costs – in the case of academic libraries, the rate of rise of per-student costs. In general, computerized library networks can reduce the rate of rise of per-unit library costs by continuously increasing the productivity of library staff. Libraries are labor-inflexible institutions wherein rising wages in the community force up library salaries even though there is not an increase in productivity.

In addition to these two principal objectives, both information and library networks should establish as a goal the furnishing of information to a user when and where he needs it. It may not be until the 1980's that networks will be furnishing users with textual information, but they are now furnishing bibliographic information to users when and where they need it. An outstanding example of such an achievement is the Ohio State University Libraries remote catalog access and circulation control system.

Three important information networks are the MEDLINE system of the National Library of Medicine[1], BASIS-70 of the Battelle Memorial Institute[2], and the ISIS system of the International Labor Office in Geneva[3]. Indeed, MEDLINE is the one nationwide information system in operation. The MEDLINE data base contains nearly a half-million journal article entries. Primary access to this data base is by subject. Physicians, health workers, and biomedical researchers can gain access to the system from typewriter terminals by dial-up to the nearest node of the national net.

BASIS-70 has about thirty data bases that various remote sponsors have put into the system at Columbus, Ohio. These data bases range in size from those that are relatively small to those that have tens of thousands of entries. Access to the system can be either via structured subject indexing or free-text searching.

The ISIS system of ILO is similar to BASIS-70 in that it also provides for subject searches via subject indexes and free-text searches. The ISIS data base is largely in the social sciences. The Swedish government has leased the system from ILO and has replicated ISIS in Stockholm where it is currently operational.

The State of New Jersey has developed the CAPTAIN system library network based on a computer at Rutgers University. The CAPTAIN system is a remote batch entry system and an on-line input system. At the present time, all products are batch products. The system is designed to computerize the acquisitions and cataloging activities of a group of New Jersey libraries. It recently began operation.

Stanford University is currently about to put BALLOTS[4] into operation, a system whose functions are similar to those of CAPTAIN. The Stanford system will also computerize acquisitions activities and catalog production. At first BALLOTS will operate for Stanford alone, but it is Stanford's intention to expand it to a group of academic libraries in the Bay region.

The New England Library Information Network (NELINET) originally developed a shared cataloging activity based on the MARC II data base that the Library of Congress makes available.[5] Remote entry employs teletypewriters. The system has been in operation for several years and continues to service NELINET members. However, NELINET is at the present time working in the direction of replicating in New England the Ohio College Library Center's on-line system. This past spring NELINET simulated the OCLC system and found that with additional equipment, the OCLC system would be able to handle some 249 New England libraries.

The first on-line cataloging system to become operational was that at the Shawnee Mission Public Schools, a school district in the suburbs of Kansas City.[6] The system began operation in the spring of 1968 with several operational terminals.

The Swedish government's LIBRIS system is designed to be a full library on-line system encompassing acquisitions, cataloging, and circulation.[7] It employs CRT terminals, and in August of 1972, LIBRIS was operational for two hours a day at one Swedish university. Managers of the system plan to implement it for several Stockholm libraries and the library at the University of Upsala in the autumn of 1972. The LIBRIS system is basically a shared cataloging system wherein libraries can remotely use cataloging data that already exists to perform their own cataloging.

The Ohio College Library Center at Columbus, Ohio is in overall design, a system[8] similar to LIBRIS. OCLC contemplates development and implementation of six subsystems: (1) an on-line union catalog and shared cataloging system; (2) an interlibrary loan communications system; (3) serials control; (4) technical processing with initial emphasis on acquisitions; (5) remote catalog access for users and circulation control; and (6) user access by subject and title. The on-line union catalog and shared cataloging system began to operate at the end of August 1971. Forty-eight libraries in Ohio academic institutions participate in the system, as well as a sprinkling of libraries in other regional centers. The data base consists of MARC II records from the Library of Congress and OCLC MARC records input by participating institutions. In early October 1972, there were over 400,000 catalog records in the on-line file, of which two-fifths were OCLC MARC and three-fifths Library of Congress MARC II. The system is operating at an annual rate of over a half-million titles being cataloged for which the Center produces over 3.4 million catalog cards yearly.

The OCLC *Annual Report* for 1969/70 contained an estimate of net savings that participating libraries would experience when the system will be in full operation. At the time of the estimate, the Center calculated that member libraries would be able to average six titles per hour cataloged on each terminal, and that an average of 1460 titles or more per day would be cataloged. From January through June 1972, OCLC member libraries used existing cataloging information at 82.6% of the rate in the original estimate; presumably after another year has passed, and the on-line catalog has grown, use of existing cataloging information will attain the originally estimated rate. Moreover, OCLC Members have reported rates of cataloging on terminals extending from 5.9 to 20 titles per hour per terminal, so that it is clear that the original estimate of 6 per hour was low. Hence, it appears that if the system is not now cost beneficial it soon will be.

The Ohio State University Libraries remote catalog access and circulation control system[9] is primarily a user service in contradistinction to CAPTAIN, Stanford, Shawnee Mission, NELINET, LIBRIS, and OCLC whose presently operating systems are primarily concerned with the computerization of internal library operations. It should be pointed out however, that several of these systems have plans to develop and implement user-oriented subsystems. OSU has the fourteenth largest library in the country with a stock of 2.6 million volumes, for which there are over 900,000 machine-readable records. Access to the data base is by author and title, and by title.

Users place calls to CRT terminal operators equipped with headphones. A user need know only a relatively small amount of author and title information about the book in which he is interested because truncated search keys are employed that consist of only the first four letters of the author's name and the first five of the first significant word of the title. When the operator has called up the record on the screen, he is able to inform the user whether or not the book is available, and if it is available, he can arrange to have it charged to the user in the OSU libraries possessing the book or to be sent by mail to the user if the user has a university address. The service is equally available to students, staff, and faculty.

The OSU remote catalog access and circulation control system is clearly the first major, computerized breakout from classical library practices that vastly improves availability of the library to the users.

From this brief review, it can be seen that information and library networks exist and operate, and that in the case of MEDLINE, there exists a national information network. However, to bring a national library network into being will require much more detailed planning than has been undertaken to date and leadership that no one has yet assumed. In particular, it will also be necessary for regional networks to conform with national and international standards. Unfortunately, some major library computerization projects do not conform to these standards and thereby cannot, at the present time at least, participate in networking at the national level.

Up to the present, no one appears to have done any serious work on the design of the organization of a national library network. Various options are open for such a design, and serious work on the problem should have begun by now.

Another obstacle in the road toward a national library network is the unavailability of personnel, particularly library personnel, to develop, implement, and operate such a net. It is clear that extensive reeducation and self-renewal must occur before a national library net can operate. Librarianship has not been a profession in which rapid development has occurred, but now that it has begun to occur, librarians will need to train themselves to be perpetual students.

Finally, library networks both regional and national must establish the goals of increasing productivity of library staff members in the future. For example, members of the OCLC staff hope to begin in the near future an experimental study of computerization of descriptive cataloging, for it now appears that such a technique can be invoked when library catalogs are entirely on-line. The work of Salton[10] and others is leading in the direction of computerization of subject indexing and subject classification. When the three processes that are now entirely manual in libraries are activated in the future, libraries will no longer be the labor intensive operations requiring increasingly larger expenditures without a corresponding increase in production.

REFERENCES

1 Davis B. McCarn, "MEDLINE"

2 David Penniman, "BASIS-70"

3 William D. Schieber, *Technical Manual on ISIS.* (Stockholm, Statskontoret, 1972)

4 A.H. Epstein and Allen B. Vearner, *A User's View of Ballots.* (Stanford, California, Stanford University, Ballots Project, June 1972)

5 Lawrence F. Buckland, et al., *NELINET – The New England Library Information Network.* (Maynard, Mass., Inforonics, Inc., December 31, 1972).

6 Ellen Wasby Miller and B.J. Hodges, "Shawnee Mission's On-line Cataloging System", *Journal of Library Automation*, 4 (March 1971), 13-26.

7 Sweden. Statskontoret. *LIBRIS Handbok, del 2.* (Stockholm, Statskontoret, 1972).

8 Frederick G. Kilgour, Philip L. Long, Alan L. Landgraf, and John W. Wyckoff, "The Shared Cataloging System of the Ohio College Library Center", *Journal of Library Automation*. In press.

9 Gerry D. Guthrie, "An On-line Remote Catalog Access and Circulation Control System", *Proceedings of the American Society for Information Science*, 8 (1971), 305-309.

10 Gerard Salton, ed. *The SMART Retrieval System; Experiments in Automatic Document Processing.* (Englewood Cliffs, N.J., Prentice-Hall, 1971).

Networks for Museums and Related Disciplines

Robert G. Chenhall, Executive Director
Museum Data Bank Coordinating Committee
University of Arkansas

A museum, by dictionary definition, is any place or building in which works of artistic, historic, or scientific value are cared for and exhibited. In many instances, the word "works" implies the output of an artist or artisan, as for example, in "the works of Verdi" or the ceramic vessel produced by an unknown potter 2,000 years ago. However, museums are also repositories for the works of nature – e.g., geological, botanical, paleontological specimens, etc. – which are considered to be of scientific value. In most instances the works of artistic, historic, and scientific value which are cared for in museums are the physical objects themselves which have been produced either by man or nature. There are, however, maritime museums, architectural museums, archeological museums, and perhaps others which are concerned with classes of physical objects that often cannot be physically transported to a central building for care and display. The very size of ships and historic buildings makes their movement both difficult and costly. In addition to the size factor, archeologists usually must destroy their sites in the process of excavation. In cases such as these, the "works" which are cared for and exhibited in museums consist not only of the objects themselves but of photographic representations, maps, drawings, blueprints, diagrams and written records.

When a museum is small, it is possible to maintain an adequate control over the collections by physical inspection and memory. For example, if one owns a collection of not over 30 or 40 dolls, or bells, or any other class of objects, and has these displayed on shelves in his living room he can readily remember where each object was acquired and probably a good deal of its history. He would be able to note immediately upon walking into the room if any of the collection was missing, and he could readily put together any of the objects which were similar either historically or descriptively. However, when a collection numbers in the thousands, perhaps with some objects on display and others in storage, control by visual inspection and memory becomes a problem. Some form of written record is necessary, even for purposes of controlling the inventory of objects in the collection, to say nothing of providing historic and descriptive information that may be useful for research purposes. It is this written record of a museum collection, however simple or detailed it may be and regardless of form, which constitutes a "museum catalog."

The advantages of using a computer for the cataloging of verbal data concerning physical objects are perhaps obvious. Once the data have been recorded in form for computer entry, it is possible to retrieve from the file all available data on selected classes of objects or selected classes of data on any or all of the objects, to sort the data in any sequence desired, and then to either print listings with a wide range of format variability or to count the numbers of items which fulfill certain descriptive parameters and perform a multiplicity of statistical manipulations.

In general, the purpose of linking together a number of computerized museum catalogs in a network form of organization is much the same as it is in other fields: namely, that it makes available to all of the participants in the network the data which have been stored in any of the individual catalogs. In the museum field, the advantages of such a network are many. To site just two examples, in the preparation of a comprehensive exhibit, perhaps on something such as the works of Verdi, one could readily obtain a listing which showed, among other things, the location of all such works which might be borrowed for the exhibit. Likewise, for any definitive research in the natural or historical sciences it is essential that a person be able to locate, as readily as possible, all of the objects that may be of interest to him. In most research of this kind an inordinate amount of time is usually required just to find the applicable corpus of data.

The first network of computerized museum catalogs was formed in 1967 when, following several months of discussion, 15 museums in New York (largely though not entirely in the art field) and the National Gallery in Washington formed the consortium known as the Museum Computer Network. This began as a pilot project and its first mission was to examine the feasibility of a computerized catalog of the combined holdings of many museums. Programming for the MCN was done by Dr. Jack Heller, a computer scientist now at the State University of New York at Stony Brook. This consists of a series of programs, written in PL/1, known as GRIPHOS (General Retrieval and Information Processing for Humanities Oriented Studies). Largely because of the financial difficulties facing all museums today, most of the original participants in the pilot project have not continued as active members of the Museum Computer Network. However, the Museum of Modern Art has cataloged its entire collection within the MCN system, the Metropolitan Museum of Art recorded a large number of objects before it was forced to curtail its cataloging program, the National Gallery is now ready to begin its computerized catalog, and, perhaps of even greater importance, the MCN programming package has now been installed as a nucleus for regional data banks in Florida, New Mexico, Arkansas, and Canada, and it has been utilized extensively in museums that are quite unrelated to the field of art.

In addition to the Museum Computer Network at least four other information systems are currently being used in museum related activities. The Information Systems Division of the Smithsonian Institution has developed a set of programs known as SELGEM (SELF-GEnerating Master). This system is written in COBOL and it has been used to record a wide variety of museum specimens, largely in the natural sciences (mammals, conodonts, foraminifera, nematodes, crustacea, etc.). Although the SELGEM system is centered at the Smithsonian Institution the program package has recently been installed as a basis for regional data banks in Florida and Kansas.

A third system known as TAXIR (TAXonomic Information Retrieval) was developed by Dr. David J. Rogers at the University of Colorado primarily for the storage and retrieval of biological specimen data. The facilities of this system have recently been expanded greatly and it has now been adopted as the basic network vehicle of the Gulf Universities Research Consortium.

Another system developed at the University of Oklahoma a number of years ago is known as GIPSY (General Information Processing SYstem). This program package has been installed at the University of Missouri as well as the University of Oklahoma and it has been used to record ethnographic museum specimens in Oklahoma, Missouri, and Arizona.

Finally, there is a program which is somewhat different from the other four in its objectives and organization. This is the Flora North America Program, a large-scale *centralized* data bank that is designed to collect, analyze, maintain, and disseminate diverse kinds of information about the plants of North America. FNA uses the General Information System (GIS) that was developed by the IBM Corporation and it is different from the others in that it is a centralized catalog designed primarily to record the authoritative botanical type specimens of North America, wherever they may be housed; rather than all of the objects contained in any particular museum or group of museums.

In order for a true interactive network of museum data banks to become a reality several problems will have to be resolved. As a matter of convenience these problems may be discussed under the two general headings of financial considerations and matters of compatibility of record structures.

It is perhaps common knowledge that museums today are caught in the same kind of financial bind that troubles universities and most other institutions in our society that are dedicated to long-range cultural, educational, and research purposes. Costs continue to rise and financial support, whethe. from public or private sources, is increasingly difficult to obtain. In the face of such financial difficulties it is extremely difficult for most museums, out of their own budgets, to finance the changeover from an antiquated card file type of cataloging system to a computerized cataloging system. No matter what the benefits to be derived ultimately from such a changeover, the money is just not there to accomplish the necessary keypunching or other form of data entry. In a number of instances both public and private foundation support has been made available for initial experimentation but very few museums so far (the Museum of Modern Art is a notable exception) have had the resources to catalog any substantial portion of their collections in a form that would permit them to eventually participate in a museum network.

The financial problems faced by museums, however, cannot be considered as separate and apart from the other problems I will discuss in a moment. In at least one instance a grant request which would have assisted a consortium of six museums to establish a regional cataloging network was turned down mainly because the funding agency was not convinced that the work these museums planned to undertake would eventually fit into and be a compatible part of a larger nationwide effort.

The second and perhaps even more important problem which museums face — i.e., the compatibility of record structures — is something which can be and is being actively worked on at the present time. Some of the major facets of this problem are:

1. Achieving an interactive capacity among the several extant computer systems being used for the cataloging of museum-type data.
2. Achieving an interactive capacity among the data that is considered to be important by the several disciplines represented.
3. Achieving an interactive capacity among catalogs that are maintained for different purposes e.g., catalogs maintained entirely for inventory control purposes as opposed to more detailed catalogs that are maintained for research purposes.

The first of these problems is not as difficult to resolve as it might appear. All of the five information systems which I have described were designed as general purpose storage and retrieval program packages which provide for the storage of data with either leading or trailing tags to indicate the content of each data string. Both in theory and in practice, the conversion of data from one of these systems to another can be done automatically by means of tag look-up tables. In one recent experiment, for example, a program necessary to convert data which had been recorded in the SELGEM framework into a framework of the GRIPHOS (MCN) system was written, debugged, and run in five hours. This, of course, did not include the building of a complete tag conversion table, but it does indicate that a compatibility between the different program packages can be achieved without too much work.

It is perhaps apparent that the major problem which museums must resolve in order to eventually develop an interactive network is the problem of data compatibility. I have expressed this elsewhere (Chenhall 1972) in the form of a simple question, addressed to the scientists and museum directors themselves: "What do you want to record?"

The use of computers multiplies the reliability of our memories and the effectiveness of our processing capabilities many fold. However, computers are very demanding in that they force us to work at a level of precision that for most museum people is not comfortable.

Let me give just one example. In the research design of a consortium known as the Southwestern Anthropological Research Group, there is a data category tag definition which is called "Site Type." On the surface, it would appear as though this would not present any problems, because in a given geographic region, there are only a finite number of types of archeological sites, even allowing for our lack of consistency in the use of the English language. However, in the SARG design, it is desirable to obtain a dichotomy between "habitation sites" and "other use areas." The question then arises as to whether one of the Southwestern Pueblos is always a habitation site. If so, one might ask the further question of whether a "Pueblo-type" structure which is next to a framing area, has no interior walls (i.e., a one-room structure) and contains no evidence of ever having been used as a dwelling (i.e., obviously a storage structure) should also be considered as a "habitation site."

There are answers to these questions, of course. For example, the system perhaps should be set up with two data categories, one for "Structure Type" and one for "Site Usage." The point I am making is that these are questions which can be answered only by archeologists or other museum related scientists. From our experiences to date with computerized museum catalogs, we have found that there are literally hundreds of questions similar to this, but that each time we encounter a new data bank or a new general class of data the problems are easier to resolve than they we· the last time. Apparently there are only a finite number of such questions to be answered even though at times the number seems to be infinite.

The kinds of questions discussed above must each be resolved in a framework which considers the data categories (i.e., the tag definitions) that are employed in the system where the information will be stored, the recording conventions that have been used to enter previous data (e.g., the use of common names *versus* Latin names for plants and animals, the use of abbreviations, etc.), and the terminology employed with previously recorded data. The number of data categories that are used is open-ended. When it is found that new categories are necessary, either because of inadequacies in the prior structure or because a new type of data is being recorded for the first time, new categories can be added to a master list. However, the master list of data categories must be controlled so that, for example, one person entering data in a catalog does not use a single category for site type and another person two categories. The control of data categories is a major problem, at least at the regional level, for data recorded in different catalogs cannot be brought together unless these data were recorded initially within a framework of compatible data categories.

The recording conventions and terminology that are used can be much more flexible than the data categories and still permit data recorded by different scientists, in one or more catalogs, to be synthesized in a meaningful manner. It is true that if one searches a composite computer file under a category of "Materials" and he asks only for the Latin names because that is what he used in recording his own data, he will not have a complete listing if someone else has used common names. However, manuals of recording conventions and thesauruses of terms used in particular disciplines are being prepared and constantly updated for at least some of the museum related disciplines. It is hoped that these will serve as standards for others to follow but there will always be room for additions.

The third area of data compatibility which museums have yet to resolve is related to, though not identical with, the second. It is the fact that different museums and often different parts of the same museum maintain catalogs for different purposes. A catalog maintained entirely or primarily for inventory control purposes, for example, can be built from records which contain no more than ten to fifteen categories of data in each record. That is, the entire system can contain no more than ten to fifteen tags (or "annotation classes" in the terminology of the Museum Computer Network system). The cataloging of a human or animal skeleton for inventory control purposes, however, is quite different from the cataloging of that same skeleton for purposes of comparative research by physical anthropologists or zoologists. In the latter case, several hundred annotation classes may be necessary in order to build a network of data that will be usable to the scientists involved.

In a recent critical survey of the uses of computers in archeology, Whallon (1972:36-37) contrasts what he calls "general-purpose" data files with smaller, more specific files of data which are constantly being maintained by archeologists for particular studies or analyses and which normally contain much more detailed information on a relatively limited set of variables. In the latter files the information is usually more rigorously defined and coded in special ways, dependent upon the particular exigencies of the intended analysis. Whallon concludes that what he calls the *ad hoc* solution of general purpose data banks probably will not prove suitable in the long run.

> If archeology continues to progress as a science, it will inevitably lead to the analysis and re-analysis of data in new ways, often requiring the measurement or observation of different variables or attributes of the same data. . . . It is easy to foresee that the neat and logical structure of catalogs now in use for data banking (e.g., Chenhall, 1971:11-16) will eventually break down. Catalogs and thesauri will become too large, unwieldy, complex, and internally inconsistent for effective and efficient use in analysis long before they will have reached their theoretical maximum [size] (Whallon 1972:37).

The problems discussed above are all very real deterrents to the development of networks for museums and related disciplines. However, those of us working closely with computerized museum catalogs do not believe that they are insurmountable. In March of this year a dozen of us closeted ourselves for two days in Hershey, Pennsylvania in an effort to develop a coordinated attack on these problems. As an outgrowth of the Hershey Study Group, a new organization, the Museum Data Bank Coordinating Committee, has been created, with funding for the first two years provided under the provisions of the National Museum Act. Some of the specific functions of this new organization are:

1. To develop comparative descriptions of the general information systems that are presently available so that a potential new user would have an objective basis for deciding which system was most appropriate to his needs.
2. To serve as a clearinghouse of data categories and minimal standard recording conventions for all museum data banks, so that data recorded in one of the systems will be compatible with that recorded in other systems.
3. To coordinate and disseminate information to all interested parties concerning new developments in the use of museum data banks.
4. To supply information and (for a fee) consultants to work with potential new users.

5. To serve as a central point for the communication of information to and from other data bank organizations around the world.
6. To coordinate the development of programs for the conversion of data from one system to another and the collection of data that have been recorded in more than one of the five systems.
7. To coordinate future system refinements, so that data recorded in any one of the systems can be processed substantially unchanged in any of the other systems.
8. At a later date, when sufficient information has been gathered in data banks across the country, to coordinate or contract for the synthesis of actual data for specific disciplines on a regional or national basis.

The Museum Data Bank Coordinating Committee began functioning September 15,1972, and we hope that, within the next two years, a sufficiently sound base will be provided so that true interactive networks of computerized museum catalog data may eventually become a reality. The problems today are not the limitations of available computer technology but rather the lack of adequate funding and, particularly, the lack of cooperative effort among the museums which are in a position to offer leadership in this area. Hopefully, the Museum Data Bank Coordinating Committee will be able to make substantial contributions toward the resolution of these problems.

REFERENCES

CHENHALL, ROBERT G.

1971 *Computers in Anthropology and Archaeology*. IBM Data Processing Application No. GE 20-0384-0.

1972 Reliable data control for multi-site research. Paper presented to the Society for American Archaeology meeting at Bar Harbour, Florida, May 4, 1972.

WHALLON, ROBERT, Jr.

1972 The computer in archaeology: a critical survey. *Computers and the Humanities* 7:29-45.

Microseconds and Multi-Months: Turnaround Time in Social Research

James Davis, Director
National Opinion Research Center

The National Opinion Research Center (NORC) is one of four agencies in the United States that is capable of gathering national sample data and not making any money off it; that is, NORC, the University of Michigan Survey Research Center and the Temple University Survey Research Center are academic, non-profit, national survey organizations. The fourth, the United States Government, you may have heard of.

There are three basic problems in our business, and three facts of life that seem to cause them. In the following pages some ideas are presented about how networking might help, if not to eliminate them, at least to mitigate them.

Let's begin with the three facts of life about surveys. The first is that they are extremely expensive at least by the standards of humanities and social sciences. There are hard scientists in the crowd who would consider our budgets petty cash, but, by the standards of the social sciences and humanities, the sort of data we gather are very expensive. When people meet me they ask two questions: "Is McGovern going to win?," (I don't know) and "How much does a survey cost?" There is no single answer. It's like saying, "What does a car cost?" I think it's fair to say that a rock-bottom cost for data for a national survey these days is about seventy-five thousand dollars. If you really want a foxtail and an FM radio and air conditioning and stuff like that, one million dollars is quite common just for collection of the data. It does not include analysis and planning.

Surveys are conducted almost entirely with hand labor. The interviewer is a human being who has to be sent to a place to talk to another human being. Field costs are the bulk of survey costs and although there have been a few Buck Rogers kind of experiments, it seems to be basically a hand-labor performance. Just like the economics of hair cuts, you can see a progressively deteriorating economic future for us unless, like the libraries, we can find ways to cut costs. First, these surveys are very expensive sets of data.

Second, surveys generate a high volume of information. There's no such thing as a typical survey. However, I will now tell you that the typical survey runs about one hour, collects about seven IBM cards worth of information per respondent, generates something like five hundred variables therein, and gives one the possibility of running something like a hundred and twenty-five thousand correlations.

There are two reasons for this high volume of information. First, social scientists do not have good theories. We don't know what to ask, so we ask everything. My favorite is a question in the NORC study about six years ago, a massive study about pressing urban problems. I was reading the questionnaire, which went on and on, and found this question: "Does anyone in this household take tapdance lessons?" I conducted research among the team of six or seven investigators and their assistants and *no one* could remember who had put that question in or why.

The second, somewhat less facetiously, is that the marginal costs of additional time are small. It would cost just as much to go out and ask people two questions as to ask them an hour's worth. This, indeed, has been studied by people trying to introduce cost efficiency. You might as well gather a lot of information since the marginal cost of additional questions is very small.

Our third fact of life about survey research is not economic but cultural. Data tends to be proprietary. Somebody owns the data. The data are either collected by an investigator, for an investigator who has a grant, or for a federal agency that has a particular mission. The United States decennial census is about the *only* non-proprietary large-scale data collection. The census collects its information for everybody.

These three facts of life — expense, high volume of information, and proprietary social structure — lead to three results: (1) a very low usage of the data; (2) delayed usage of the data; and (3) usage of bad data.

I don't have any hard evidence on the low usage rate, but I would guess the average survey generates either a book, a monograph, or two or three articles. Furthermore, I would guess that if you went through each report, it would be very rare to find that the analysis boiled down to more than twenty-five correlations out of a possible hundred and twenty-five thousand. Thus the data cost — I'm just making up numbers but I don't think they're far out of line — run about ten thousand dollars per correlation. The remaining correlations just sit there. Ours sit in IBM cards in something called Petersen's warehouse on the south side of Chicago. We have *very* low usage but Petersen is getting rich. You may laugh, but low usage is why we are headed for economic extinction. The second result is delay. It takes about two years to get a report. Last year I had a very crucial interview with a fairly prominent official in Washington. I was asking him for some money and he said, "Do you know why sociologists have a bad reputation in Washington?" I said, "Because we're stupid," – I wanted to beat him to it. He said, "No! Everybody in Washington is stupid. You're too slow."

It takes you about two years to get a report out and *then* most of the analysis is done by a trickle-down effect. The original investigator completes his study and writes his report. A couple years after that somebody hears about it. They then write us at NORC to ask if they can have copies of the cards. If we can find the cards and the code book, we send it to them. Again I don't have hard data, but I would guess the articles coming across my desk to review for journals are essentially on data collected between 1960 and 1965. The Interuniversity Consortium for Political Research (ICPR) at the University of Michigan is trying to solve this problem.

Because we have low usage of the expensive data, and great delay in getting reports out, most of my colleagues substitute by using bad data. You can do a survey of your freshman class quickly and cheaply. Given certain motivations to publish research in our business, people publish research on the data they can get. Frankly, most of the data published in the better sociological journals is pretty bad in terms of the samples and measurement. That is, in the ability to generalize to larger populations and concepts.

What is the solution to this sort of problem? For this audience in this setting it is obviously clear, and it boils down, I think, to two things. First, we need to collectivize data gathering. We need to have more groups of investigators or groups of institutions or whole professions getting together to subsidize the collection of data for everyone. This contradicts the "Leonardo theory" where the great genius gets a grant and everybody else hopes someday they can get their hands on the data. I don't think this theory is paying off and I don't think it's economically viable. We have to collectivize and socialize data so that the handful of organizations with the technical ability to collect data can be acting more in the general public interest. This of course is notoriously difficult given our traditions of granting and funding. Second, once the data are centralized, we need to get them out to the customers through something very much like computer networks. It seems to me an ideal operation would have two or three large centers gathering data for groups organized by content, social problem, discipline or whatever, and then have these data freely available immediately to everybody who is interested.

I should like to sketch for you two reasonably concrete examples of prototypical projects of this sort. The first is barely relevant to the question of computer networks but I think it illustrates the idea. NORC has received a grant from the National Science Foundation, Social Sciences Division, to conduct an annual survey on topics of general interest to sociologists. We went out in January 1972; we'll be going out again in 1973, 1974 and hopefully afterwards with a standardized schedule. The questions do not change; they are kept fixed every year so you can detect time trends. The questions are set by an advisory group of fat cat sociologists. We do not analyze the data at NORC. Rather, we deposit it with various depositories, in particular, the Roper Center at Williams, Massachusetts. Anybody who wants a copy of these data this year, next year, any year, can obtain them for twenty-five dollars a deck. The 1972 survey ran to two decks so it cost fifty dollars. That is strictly the cost of duplicating it. For fifty dollars anybody, student, faculty member, researcher, can obtain national survey data, five or six months old, on topics relevant to everybody. These materials are also being input into a Dartmouth government time-sharing system and thus will be available to people at Dartmouth and to all the customers of Dartmouth time-sharing.

Next I'd like to describe a project we are beginning for the RANN division of the National Science Foundation. The project is a literal response to the statement that social sciences are too slow. Beginning in 1973, NORC will draw what would be a national cross section sample of the adult population, size nine thousand. This is fairly big as surveys go. We are then going to divide the sample into fifty equal "mini-samples" and interview one batch a week, starting from January 1973 and continuing —

we hope for a long time — but surely for one calendar year. This means that we will be in the field continuously. This is a distinct advantage in terms of getting fast output. When a man calls up with a problem instead of saying, "Well, let's get a grant, design a study, and so forth," we can say "All right, we'll stick your question in next week." We will be interviewing continuously. Furthermore, this research will be for a consortium of federal agencies. (Again, this notion of collectivizing the whole thing.) No particular agency needs this much information so fast. However, we have a collection of six agencies: the Department of Agriculture; HUD; the Department of Transportation; OEO; NSF; and the Office of Management and Budget. We hope that we can proceed interactively with these people, giving them information, and, if they don't like what they are getting, changing the question the next week. We're not only trying to get data out fast, we're trying to get a feedback loop built back in so we can improve our data collection. Finally, as the true computer network aspect of it, we have arrangements with Dartmouth to put these materials into Dartmouth time-sharing, in the IMPRESS program. IMPRESS is the only shared- time social science program that could handle this much material. IMPRESS is being changed so it can be up-dated continuously, which means that our federal users in Washington, or anybody, can call up Dartmouth time-sharing and run his own data. We are hoping to train our federal customers to run their own data because once we've got this thing running so fast, it is a waste of time for them to write us and ask us to make tables and mail the tables back. This is our network. How many points do you have to have to have a network? I guess two will do. We'll have three. The data will be entered in Chicago, maintained in Hanover, and accessed by our various users in Washington, along with users on our staff. The data have a pure science sort of mission too, and in tape form, card form, maybe through access to terminals, are available to any investigator anywhere. This, of course, is a very small network. Today it consists of telephone lines from Chicago to Hanover. But, I think, it does provide a prototype for a way of reorganizing social science data collection in such a way that two things will happen, one noble and one less noble. The noble result is that the nation will be getting more return for its investment in social science. The less noble result is that survey houses will be able to stay in business.

CHAPTER 2

DISCIPLINE ORIENTED WORKSHOPS

Museum and Related Disciplines

Chairman: Robert G. Chenhall, Executive Director
Museum Data Bank Coordinating Committee

John H. Beaman, Curator
Beal-Darlington Herbarium

David Vance, President
Museum Computer Network Inc.

Forrest McGill
University of Michigan

Recorder: Helen Ianni
Wayne State University

Dr. Chenhall opened the workshop with a summary of computer based activity at the University of Arkansas Museum. He began building a computerized data bank at the University of Arkansas for the Arkansas Archeological Survey in 1969. The Museum Computer Network software package was used to record both archeological sites and artifacts. The same package is now being planned as a basis for a catalog of the entire collection of University of Arkansas Museum.

Mr. Vance, Registrar of the Museum of Modern Art in New York which holds a permanent collection of approximately 25,000 objects, described the computerized file format which the Museum uses. Using the computer based system, one file rather than four are maintained at the same price and with the same amount of work. This file can be accessed according to any of 40 data categories, or any combination or permutation thereof. The Museum is looking forward to exchanging information through the Museum Computer Network with other art museums, and museums representing other disciplines. To do this, they must work closely with other museums, establishing a system whereby information is organized in the same way. The network performed a crucial experiment in 1968-69, testing to see if data selected at random from files which were in poor condition and not intended to be used together could form a coherent data base which could be used effectively. The experiment was successful. The few museums who can afford it are using programs of the Museum Computer Network. The programs are continually being developed and expanded, although they are useful in their present state.

The Museum Computer Network has worked with other museums which have developed their own programs. Some science museums were unwilling to come under the wing of the Museum Computer Network, since the Network has become associated with one particular program package which is tied, at least in the foreseeable future, to IBM (due to the limited development of PL/1 compilers). To bring together these museums and others, a super-organization, the Museum Data Bank Coordinating Committee, has been formed which represents museum groups in the United States and in Europe. Each museum has developed almost identical systems, using completely different programs. The committee plans to sponsor some logically trivial programming which will enable these museums to exchange information.

Mr. McGill outlined a project which he directs that has been designed to work out a system of coding the variables for a very large group of images of the Buddha from Southeast Asia, particularly Thailand. The information will be retrievable in several different configurations.

Dr. Beaman, speaking as a member of the Editorial Committee in the Program Council of Flora North America, described FNA plans to develop a new computerized flora of all the plants in North America. This program was to have been implemented October 1, 1972 through the Smithsonian Institution, involving Michigan State University, the New York Botanical Gardens, the Missouri Botanical Gardens, and the University of North Carolina. Flora North America (FNA) has developed a set of files: lists of species, morphology files, etc., and has developed batch processing with the IBM General Information System. At this time, FNA has no real time-sharing capabilities.

Another project at Michigan State University which Dr. Beaman directs is developing data bases in the form of matrices with taxons on one axis and characteristics on the other. Through different program packages, a user can: (1) make a direct identification by inputting characteristics; (2) have the computer construct keys to a particular group or write descriptions based on the information in the matrix; or (3) have the computer quiz a student on his knowledge of characteristics of a particular group of organisms.

In the open discussion which followed, several points and observations were made. One advantage of using the computer is the ability to eliminate the occurrence of human error in misclassification of plants, art objects, etc.

The problem of classifying archeological artifacts is a more complex one, due to the vastly different levels of description which are not organized into a widely accepted method of classification. There are three ways of weighting characteristics of plants and art objects in a museum classification scheme. The simplest system is not to weight them at all. A second is to subjectively weight them, inserting in an extra row the respective weights of the characteristics. The third is to calculate the probability for frequency of occurrence, as of a plant species in a given locale. However, the mathematics for this have not yet been worked out. The third weighting scheme has several disadvantages when it is applied to classification of art objects and archeological artifacts. Many artifacts are one of a kind, but the scheme definitely would be useful in the case of the 10,000 Buddhas, many of which were probably made by different workshops, whose work is distinguishable. A classification scheme for art objects like the Buddhas must include indicators of the relative importance of style, and iconographical details. It is virtually impossible to create a hierarchy beforehand. For archeological artifacts, for example, there exists a space/time problem: groups of morphological characteristics may define certain geographical areas, while a group of design-style changes through time may cross several of these areas. A solution, greatly facilitated by use of a computer, is to run sorts according to each characteristic before classifying, compare the sorts, and see which sorts tend to bring the same groups together. Through the computer, one can also discover new relationships, e.g., between plant species and the selective nature of the environment.

The economics of using computerized files is an issue for museums today individually and as members of networks. Many have relied on computers only for better performance of tasks which have been traditionally done manually. Mr. Vance estimated a cost of $1 billion to catalogue each of the roughly half-billion museum objects in the United States for inclusion in computerized files. Drs. Beaman and Chenhall felt that the expense of such cataloging would be prohibitive whether one considered plants, archeological sites, or artifacts. They advocated a plan to computerize new classifications and new data, which would yield excellent records, perhaps, 10 years from now. It was noted that the Mexicans have been using computers for many of the above purposes since 1964.

The next step is to implement networks, realizing the cost of implementation, and what kind of economic support can be expected.

Computers and Networks in Chemistry

Chairman: Harrison Shull, Vice Chancellor for Research and Development
Indiana University

Recorder: S. P. Singh, Department of Chemical Engineering
University of Michigan

The discussion section began with the presentation by Harrison Shull of different ways in which computers can be used in the discipline of chemistry. The following uses of computers in chemistry are examples.

1. Computers can function as a computation tool for theoretical chemists to solve complex problems.
2. Computers can control laboratory equipment enabling experiments to be modified as they proceed.
3. Computers can be used in on-line data acquisition.
4. Computers can be used for information retrieval and storage. Through Chemical Abstracts Service, in these systems there can be access to the structure and substructure of chemical compounds. Graphical systems can also be used for information retrieval.
5. Computers can be used in the area of inventory control problems. In the discipline of chemistry, computers are used for inventory control of chemicals and equipment of daily use.
6. Computers are being used in many areas for the purpose of computer augmented instruction (CAI). In the area of chemistry, computer aided education can be especially useful.
7. Chemistry departments are complex. Computers can be used for management and resource allocation problems.
8. Computers can also play an important role in message switching services. Communication between different departments of different institutions can be made much more efficient.

Dr. Shull pointed out that if we can solve networking problems in the area of chemistry, we can solve problems for other disciplines. Chemistry is a well-defined discipline and provides an ideal example of the kinds of problems which are faced in building networks.

Open discussion covered many of these applications of computing in chemistry. It is very important to decide in what manner one is going to interact with the computer. To make computers more useful, it is necessary to develop a natural language, the language which is used by chemists on a daily basis. It is important to have the computer adapt to the vocabulary of chemists.

Graph theory can be used in chemistry for representing the connectivity of atoms in molecules as well as for indexing Chemical Abstracts.

Graphical terminals can be used to display and store complex biochemical structures. New structures can be edited, stored, examined and manipulated.

There was some discussion on the use of networking in the discipline of chemistry. All of the uses which Dr. Shull outlined do not require a national network. Dr. Kriloff remarked that, when the University of Illinois chemists get programs from some other place, a great deal of time is spent in changing the programs to function in the new computing environment. There are also problems caused by poor documentation. If a computer network is available to chemists, these problems will not arise. Dr. Wyatt said that two types of costs can be avoided by using national networks: (1) conversion costs, and (2) storage costs.

National networks will make clear and detailed documentation a necessary part of the everyday life of a chemist.

Networks for Languages and the Humanities

Co-Chairman: Sally Yeates Sedelow
Professor of Computer Science and Linquistics
University of Kansas

Co-Chairman: Walter A. Sedelow, Jr.
Professor of Computer Science and Sociology
University of Kansas

Joyce Friedman
University of Michigan

Hans E. Lee
Michigan State University

Victor L. Wallace
University of North Carolina-Chapel Hill

Recorder: Kurt F. Lauckner
Eastern Michigan University

The workshop on networks for languages and humanities began with a series of presentations, led by Dr. Lee. The full text of Dr. Lee's paper appears as Appendix C in this volume.

NETWORKS FOR WHOM?

Of particular significance to the humanities is the fact that many forms of research in the humanities have been funded by the National Endowment for Humanities while the development of computer networks has been supported primarily by the Office of Computing Activities within the National Science Foundation and similar agencies. Scholars in the humanities who use computer services may find it very difficult to influence decision making about the allocations of computing services in a particular campus.

Most computer services so far have been developed by people who either have been trained as systems programmers or are numerically oriented. However, the user, for whom the system should have been designed, is the person who wants to do some research in the humanities or in the social sciences but who is not a computernik. Hospitals provide a nice analogy. Sociologists for a number of years have been pointing out that patients in hospitals are treated as objects. In effect the whole hospital structure and its processing operations are designed for the convenience of the nursing staff, the physicians, and the administration. The same thing carries over into the provision of computer services for humanists.

COMPUTER PROGRAMS AS NETWORK RESOURCES

Dr. Friedman described two primary goals of researchers in the humanities: accessing large programs available in other places; and making local programs available to others. A number of large linguistics programs which have been written in other places have been brought to the University of Michigan with some effort. They probably aren't going to be used very much. It would have been easier to use the programs in the original installation.

Groups of programs that people want to look at could be distributed over a network. Often one would like to try a program but doesn't because the time and effort necessary to obtain a tape and run it are too great.

Dr. Friedman has large linguistics programs at the University of Michigan which are designed to be used by linguists who are not necessarily computer people. One program takes a particular form of grammar and grinds out sentences in order to help people to write grammars. Users get to see what the output is going to be. There are a number of users, but it's a large program and it's fairly hard to send out. Dr. Friedman would like to be in a situation where people could try it out and decide whether they really wanted a program that big. She would like to be able to work back and forth with her users, to demonstrate her programs, and to have people demonstrate their programs.

The Association of Computational Linguistics will meet in Ann Arbor in the summer of 1973. Individual programs will not be brought to the University of Michigan for demonstration even though the machine facilities are available, because the time and effort necessary for the transfer cannot be justified for a short demonstration. Networking would make demonstrations like this possible.

REASONS FOR NETWORKING

Dr. Wallace noted that creation of the appropriate kind of network is very much dependent upon the reasons why people use networks. One of the reasons people use networks is for economy of scale. A large machine with twice the capacity will not cost twice the amount of money. Through economy of scale the user gets more computer power and more attention from a well-trained staff while the supplier gets smoother demand. Capability of staff is the second reason for using networks. A person centered in a small university may have a small computer but also needs to have access to a larger machine and the associated staff.

A third reason why people want to have networks is to share resources, especially nonportable programs. At the present time the nonportability of programs unfortunately is the rule rather than the exception. Researchers also want to share hardware resources. There is only one ILLIAC. If a user needs to get at it on a frequent basis, as part of a research project but is located in the other end of the country, he wants to be able to communicate better. Data sharing is another form of resource sharing facilitated by networks. Some computer applications require the use of more than one processor simultaneously. In computer graphics one processor is needed for the graphics terminal and another for related servicing. The two processors hooked together form a basic network. One last, entirely different, reason for networking is to facilitate getting messages back and forth from one place to another.

The network that is appropriate to each reason is a different network. To achieve economies of scale you want heavy bandwidths over short distances. To utilize a more capable staff one needs, essentially, a terminal processor kind of network. Again it is going to be fairly local because computer power is not terribly unique. For resource sharing one may be forced to use a national network. Rarely, however, does it need to be a wideband network.

DISCUSSION

The open discussion which followed covered several topics. One additional reason for networking is increased availability. Very often with a small machine it is not worthwhile to run it more than one shift, whereas the big remote machine can be accessed from home, very often six or seven days a week.

Expertise sharing is very important. As the demand level for access to expertise picks up from all over the country one cannot expect the expert to operate on a shoe ing and to fulfill functions for which one actually needs a technical staff. Outstanding researchers may be unable to continue to grow as researchers and scholars if they must become programmers and technical advisors to others who wish to use their programs.

Many people who are doing language and humanities-related research feel that they are not in settings where there is a sympathetic computing environment. In some places it may not be feasible to try to achieve the level of understanding, interest, and sophistication of expertise that make the environment sympathetic. Through a network a user can sometimes get an approximation of a sympathetic computing environment. It is very easy if one is in a computer-rich environment to feel that little needs to be done to change the current situation. People who are well-situated raise all sorts of objections about what would happen if computing capability became available to more researchers. Yet, the failure to open computing resources through networks denies many people who might do creative research the opportunity to do so.

A network does not necessarily mean one center. A network could mean something more ARPA-like, or have regional nets within a larger net. In the physical sciences it may be most economical to establish a single national computational research center. At the National Center for Atmospheric Research (NCAR) it has been cheaper to bring people in by plane, in fact, than to bring them in by telephone lines to use a computer. Because a system has been so finely tuned, it may be better to have researchers work on site. In linguistics and humanities the situation is very different. It might be a great mistake to settle upon a given computer or a given set of programming languages. There are so many different kinds of capabilities, to which different sets of language research users should have access, it would be a great mistake to try to consolidate them all.

Any plan for networks should fit into the overall system, the dynamic system, of our society. Nonportability of programs is going to continue to exist if for no other reason than the United States mail is getting worse all the time. By the time the network exists the United States mail will not be a competent mechanism for distributing software. Transportation may improve in the next 100 years so that people are transportable themselves. We should be addressing the question of networks in these terms.

A key problem with networking is communicating the protocols to potential users of the network, either regional or national.

User support must be available at each network node for that system and for other national facilities. One possibility is to build the documentation into the interactive programs. A user wouldn't have to see it each time, but if he makes an error, documentation could explain the legal options at that point. Another possibility is to have para-professional or a para-pedagogical staff to fill these very important informational roles, especially to decipher and reinterpret documentation that's received from other centers. The protocol explicator has to be a very intelligent, available, and sympathetic person. Otherwise computing will really be discouraged, especially in the humanities and social sciences.

One can distinguish between technology development, resource network development, and human network development. A network is more than a bunch of equipment tied together by wires or by radio. The real contribution of the whole network program will be an increased ability to look at various national problems: the urban problem; the poverty problem; the pollution problem; and the transportation problem. Networks will help various people in various places to work together and will encourage an inter-disciplinary approach.

In summary, resources of interest to computer network users in languages and the humanities are: computer programs; data banks; expertise; computer power and, in some cases, computer availability might be a part of a computer power. User support services should include two kinds of documentation (reference materials of a very first rate quality and tutorial materials) and, possibly, audio-visual materials as well. It may be useful to have full-time librarians at each node.

One mechanism to encourage sharing would be to give credit to humanities researchers who prepare material for a central data bank or software bank. It is also important to establish a central clearinghouse for documentation of resources available on the network, and to guarantee fast access to these resources. It might be necessary to have at least one node on a general purpose network dedicated to humanities use in order to provide fast access to computing power.

In an area like the humanities, which is so broadly diffused, one center cannot effectively take care of all functions: calling conferences, providing research fellowships, having seminars, serving as a clearinghouse, providing the staff, providing program maintenance, and providing area specialists. If there is one central organization for the network, it should function as a communications utility; providing a service for whatever use people want to make of it.

There are some natural functional groupings among humanities researchers, already developed today, which evidence a synergistic or cross disciplinary trend in computer based research; e.g., the use of some research on natural language for theoretical development of formal languages; the comparison of programming languages with natural languages in the study of semantics; and, the study of semantics as it relates to cognition. Other kinds of groupings have formed, ironically, by dislocation. Occasionally a research project which once consisted of graduate students and a number of faculty members in an institution has broken up because those graduate students took their degrees and went off somewhere else. Now they're completely cut off. The researchers aren't able to put the system up at their institutions and don't have access to the data. Research careers, which have been focused upon a particular type of research, are lying fallow. A network to link these researchers to previously used systems and data would be most welcome.

A network governing body for a center or centers serving humanities or language research areas, should have a board of directors perhaps similar to that of NCAR which represents many institutions in a consortium. There ought also to be public members, people who don't have vested interests, to avoid the kind of inbreeding we all fear.

Computers and Networks in Economics

Chairman: George Sadowsky, Senior Research Staff
The Urban Institute

Sanford Berg
Assistant Professor of Economics
University of Florida

Mary W. Hook
Senior Staff Economist
Council of Economic Advisors

Mark Eisner
Technical Director
National Bureau of Economic Research,
Computer Research Center

Recorder: Gene Raymer
Wayne State University

The workshop was opened by Chairman George Sadowsky leading the panel with a statement of opinion, involvement, and ar of concern. Given the fact that the networks exist and that the technology is available to create them, what kind of a network environment or a computing environment is going to be most helpful in the future to practitioners in economics? In other words, if we h. l a free hand in creating an ideal computing environment for economis , what would it look like?

It would have to include service for a variety of types of actors, people who are economists or who do things that are related to economics. They're widely scattered in business, in non-profit research organizations, in colleges and universities, and in all levels of government. In terms of location, these actors are national in scope and move in national patterns. At present, although economists are highly mobile, their programs and data are often not. There are many cases of people who have moved geographically or institutionally two or three years previously but who are either still computing on their previous computer or have spent several years getting started again in their new environment.

Networks appear to offer some kinds of solutions for some of those situations. If the actors are national and move nationally, then the network facilities offered to them probably ought to be national in scope also.

There are two basic concepts of network. The first concept of network is a general switchable utility concept much like the telephone network. If you want to talk with anyone in the United States, you pick up the phone and call them. The line connecting you carries whatever you decide to put on it, i.e., whatever you decide to say. If you're dealing with a remote computer, the line carries whatever kinds of characters or messages you transmit between you and your terminal and the computer.

The other kind of network is considerably more specialized and is more intimately connected with a specific service or set of services. Library networks and airline reservation networks are examples of this concept. Messages that are transmitted on these networks are often highly encoded for a specific purpose. It's important to distinguish between the two concepts of network because they result in two very different scenarios.

The concept of a general utility network results in an important reorganization of computing use patterns. The network may now be used to make one or more product markets, where the products are semi-independent of network characteristics. The network allows this by providing very cheap transportation for data, programs and other information between buyers and sellers who have heretofore been isolated. This expansion of markets results -- at least in economic theory -- in increased specialization and incentive for increased investment. In concrete terms, there is much less likelihood of 100 more people writing 100 new regression programs if a network existed that allowed easy access to existing regression programs. So one of the primary benefits of networks expanding markets for particular products is eliminating duplication and increasing competition. The network becomes a vehicle through which products compete in larger markets.

Mark Eisner spoke as a frustrated applications systems developer. Many people who deal with funding agencies like NSF find themselves in the situation where there are millions of dollars available for developing very good and sophisticated tools, but no money available for supporting the dissemination of those tools.

In quantitative macro economics, one is not dealing with a great deal of data. Almost all the macro economic data that people would want to use could fit on eight 2314 packs and be on-line all the time.

There is a strong need for good and sophisticated tools to deal with these data. Current computer technology can serve a large group of users. The technology for networking also exists. Certainly in macro economics and econometrics, the appropriate professional community could get together and use the tools that are already provided. We already have the basis for a network using existing telephone switching systems and circuitry, but its costs often make it prohibitive to use. If telephone charges were made in accordance with their real costs -- not just a mileage charge -- and if the difference between voice and data communication was recognized operationally, we might have an adequate network today. The primary barrier that prohibits economists from using a remote computer facility is the communication costs. There are many other solutions to the network problem of providing cheap shipping -- some of which already exist; for example, the GSA network. The distance makes no difference and the charge is computed on the basis of the amount of information that is sent.

A logical community of users in macro economics exists today. The big drawbacks are political and institutional problems. Who pays for the computer time is a basic concern. A funding agency gives a university money, but can the researcher in that university spend money elsewhere? Systems that can support an entire large section of a discipline are large and complicated. However, if they have been developed in the public domain, no one owns them and their maintenance cannot be supported. There's a myth that if something is developed in the public domain then anyone can take the software and make it operational on their computer. This is not true for systems containing 100,000 to 500,000 lines of code.

Network concepts have developed from broad, very large dreams: yet people discuss immediately how one computer can talk with another computer. We ought to concentrate on a much simpler goal, and move very slowly in an evolutionary way. It is important right now and for the next five years to have users "talk" to one computer and get information back from that computer. Let's not worry now about computers talking to one another.

With a simple network that gives users the ability to run a 600 line per minute printer, then remote batch entry and production of bulk output is feasible for a reasonably small amount of money. That network should be produced completely independent of any discipline. It should be a service that's available and that's supported so that all disciplines can get on and, in a free and competitive way, people can use the resources made available by the network. Through competition, better products, better organization, and better ways of dealing with institutional problems will emerge.

Mary Hook described her major interest as using economic data and information and gaining access to information quickly. The volume of macroeconomic statistics is, on an absolute scale, relatively small. It's not small at all, however, when one speaks in terms of collecting it, transforming it into machine-readable form, and using it on short notice. The Council of Economic Advisors is basically interested only in the major macro series, but there are problems if we want to get information to the council members quickly, i.e., within one to three hours after the data have been produced by the originating agency. Most working economists and econometricians get their data from the standard publication sources. We just can't wait that long.

The Council, in addition to doing short-term reporting and short-term forecasting, is constantly initiating studies of special problems. Last year the Council promised the Joint Economic Committee that it would, for the 1972 annual report, produce another study of unemployment. Work has begun at the Council now.

In terms of computing facilities, we have available to us a 360/75. We haven't used it because it has no data bank. Any data that we want to operate on must be keypunched and entered into the machine. The state of the programs available to manipulate the data is less than advanced. Only a primitive version of TSP and little communication on how to use the control language of the computer are available. While the computer is a free good, it's not an attractive free good.

On a commercial time-sharing network there are available a very sophisticated language for manipulating data and a quite large data bank, containing close to two thousand quarterly time series. We are most interested in those series. The service gives us the ability to transform variables and has all of the standard statistical regression capabilities. We've also recently begun to use the Brookings Institution computing center, which also has a small data bank.

A recent special study, required a large subset of BLS employment data, which is already in machine-readable form. BLS does sell copies of its files. Unfortunately the data format is not directly machine-readable by any of the computers to which we have access, since they are written by old 7070 programs. The gold mine of data exists in relatively unusable form. We have the pity of knowing that in all too many places that same data are being laboriously keypunched, entered into a computer, and used. It's a very sad situation.

Sanford Berg noted that some institutions are beginning, because of involvement with new computer research and large data bases, to try to cope with the information and cooperation problem. With the 1970 Census tapes at DUALabs, literally just one group must perform the function of disseminating these data. It's almost imperative in that setting that cooperation occurred. This has also happened with several other large-scale microdata bases. The Survey of Economic Opportunity, for example, is now housed at the University of Wisconsin. A machine-readable bibliographical system describes everyone who is doing research in the poverty area and how to get access to those people.

An additional inefficiency in the utilization of computer services for economics is due to the use of block grants rather than project grants. Universities would like to charge high prices for those local demands that are quite inelastic, e.g., those supported by government grants, but if they obtain sufficient revenue from project grants they can't then give extra time away free for their internal uses. Instead, the product has to be differentiated to justify differential prices, such as charging less for overnight or third shift use of the facility. Often, this results in computers not being used to capacity. Block grants awarded to universities, would probably be a better method for the Federal government to maximize research output at universities.

One project that this group and other groups in economics can take responsibility for is to study the economics of the current and future computing situation and point out what can be done to increase the efficiency of use of these resources.

George Sadowsky summarized the consensus that what's really needed are ways of accessing individuals, computers, and other people's data and programs. The domain is national with special focus on special data and program repositories such as Washington. Elaborate user oriented service centers are not as useful as access to what already exists. Once we learn how to use existing resources, we can proceed, but the immediate task is to break down the barriers between local computing enclaves.

The second question posed by the conference organizers seems easy to answer: what kinds of technology make sense for the group or discipline:

raw computing: large files, active or static: programs and software: interactive terminals: graphics? It seems that the answer is "some of each" and it depends upon the application being executed and the purpose of the research.

What we seem to need now is to make full use of what we have. We are overdeveloped and there's not enough user testing of products that exist.

Regarding the kind of support patterns that make sense for economics we can only list issues. What are the pros and cons of institutional budgets and grants from foundations, government subsidies and is there a case for government subsidy of a network? Is there a case for specific conditions to be tied to grants so that computing money is freely floating money? Is that the way to restore consumer choice once a network exists? How do we finance computing if the scenario is going to shift to increasing use of networks and how do we change the financial structure of computing to the extent that public policy at some level provides a more efficient allocation of resources, and better services to researchers. There are many choices and a great deal for us to learn.

Library Networks

Chairman: Fred Kilgour, Director
Ohio College Library Center

Davis McCarn
Acting Associate Director
for Science Communications
and Computer Engineering Service
National Library of Medicine

W. David Penniman
Associate Chief, BASIS-70
Battelle Memorial Institute

Ronald Miller
Director, NELINET
New England Library Network

Recorder: Carolyn Landis
EDUCOM

MEDLINE

Davis B. McCarn described MEDLINE, the on-line, nationally available literature search and retrieval system of the National Library of Medicine. Through this system, a computer at the National Library of Medicine (NLM) can be accessed from forty cities. The service reached its present state of operation in October of 1971. Information available for each title in the data base is standard bibliographic information: author, title, language, publication date, etc. Access made by subject yields a response within four seconds, listing all titles available in the system pertinent to that subject. The user can request that all titles be printed if he wishes. Average search time for most users is approximately ten minutes.

Four modes of access are available: Western Union, Datacom, TWX, and Tymshare telephone networks. The data base includes over 1200 journals indexed with over 400,000 citations with 10-12,000 new citations being added to the system each month. The data base covers publications cited in *Index Medicus* since January 1969.

MEDLINE service runs 43 hours per week using primarily the Tymshare telephone network. Charges for service are made on the basis of resources used. Each ten minute search costs about one dollar. Medical schools and research libraries are the major users. Beginning with approximately 106 searches per month, the MEDLINE service has grown substantially. In each of the most recent months, approximately 3000 off-line printouts have been provided which resulted from approximately 10,000 searches. Approximately twenty users, of 120 potential users, access the system simultaneously. Average response time with twenty simultaneous users is three or four seconds.

A parallel service run by NLM on another computer at the Systems Development Corporation (SDC) is the AIM-TWX service which consists of a reduced rate base of approximately 200,000 citations in addition to the resource base from the Educational Research Information Center. On a trial basis, this service has been available free to MEDLINE users during September 1972. The NLM date base TOXICON and Chemical Abstracts Service CBAC tapes are available through the same network on a commercial basis. In December 1972, Chemical Abstracts Condensates tapes will be available to MEDLINE users for $45.00 per hour. With each of these parallel services, whether available without charge or at commercial rates, the onus is on the user to learn the differences in methods of access and of use for each system.

Several constraints have prevented extremely wide use of the MEDLINE network. Some commercial network time-sharing services are available for $10-$15 per hour while retrieval services may cost $45-$50 per hour through the NLM network. Tapes used in MEDLINE were designed for old-fashioned batch processing systems and are not really tailored for the current information retrieval system.

Problems encountered in library use of the data bases are primarily attitudinal. Many libraries measure the performance of information retrieval services with precision and recall ratios which do not accurately measure additional advantages available through the on-line systems as opposed to manually operated systems. A second problem is the lack of readiness in libraries to make use of an on-line system. The tradition of the free library makes acceptance of fees for on-line services difficult for most libraries.

Dr. McCarn predicted a vast proliferation of on-line information retrieval systems utilizing data bases on individual computers which will become available in many subject areas and many geographic areas. The NLM and other networks have already laid the groundwork for such expansion.

In answer to questions from workshop participants Dr. McCarn explained that MEDLINE uses the Tymshare network for communications only. A data deconcentrator, which is attached to the NLM computer, funnels into that computer communications from outlying terminals which arrive via the Tymshare network. MEDLINE does not utilize any of Tymshare's computer resources. The cost of a single search on MEDLINE ($5) is considerably less than the cost of a similar search on MEDLARS ($85) which is the batch operated processed search service also available through the NLM. The user pays for: long distance charges to the nearest Tymshare node with a data concentrator; the terminal located at his site: staff time required to become familiar with the NLM data base; and increased service commitment to the local user community. No payments are made directly to the NLM. Some user institutions, e.g., the College of Physicians in Philadelphia, must charge users additional fees to recover costs of providing access to the network.

New data bases which will be incorporated into the system will include serials catalogs, and additional entries of journal citations.

BASIS-70

David Penniman described the BASIS-70 information retrieval system which is operated by the Battelle Memorial Institute in Columbus Ohio. BASIS-70 is an on-line, interactive system tailored for the user who is not a computer expert. Over thirty files are available through BASIS-70 including literature references, management information systems, social systems information and files on materials selection. Clients who have contracted for use of the system include: the United States Department of Transportation; the United States Department of Environmental Protection; the United States Department of Defense; the Price Commission; and the Copper Development Association. If a particular sponsor wishes to make his data available to others, Battelle Memorial Institute will cooperate and assist outsider users to access the data base. Commercially available data bases like Chemical Abstracts Condensates will be available on the system in the near future.

The BASIS-70 system operates on a CDC 6400 using the intercom/scope operating system with programs written in Fortran and Compass. Terminals operating at 10, 30, and 240 characters per second can be supported. BASIS-70 will begin using Tymnet in November 1972 to make the system available to a widely dispersed set of users on a cost effective basis. The service is now available from 9 a.m. until midnight each weekday, from 8 a.m. until 5 p.m. on Saturdays and for four hours each Sunday. Most terminals accessing the system are CRT terminals with capability for producing hard copy printout.

BASIS-70, although primarily an information retrieval system, has been used as a general computer resource by some sponsors.

NELINET

Ronald Miller outlined the problems faced by the New England Library Network (NELINET) and the solutions derived by the network during the past few years. NELINET originated in 1967 with the support of the Council on Library Resources, the advent of the MARC data files, and the technical assistance of Inforonics, a computer service bureau for libraries operating in Maynard, Massachusetts. NELINET functions through an executive committee which represents twenty-seven participating members. The New England Board of Higher Education, which is an instrument of six area states, has veto power over some activities of NELINET especially those related to the use of facilities of the Board. NELINET has grown from six founding state university libraries to include twenty-seven potentially on-line libraries in universities and state agencies.

Cataloging, one of the most labor intensive activities in libraries, was chosen in 1967 as the first area for design of a computer based system. A technical processing system which produced labels and catalog cards utilizing paper tapes and batch processing was initiated with the assistance of Inforonics. After experimenting with and abandoning a teletype to teletype connection, users began sending paper tapes to a central location for processing. This method continues to be used today. Other library activities which were designated in the original plan as future computer based activities are still to be incorporated into any NELINET operating system: circulation control; reference services; management information services; and acquisitions systems.

In 1970 NELINET began to investigate whether the Ohio College Library Center (OCLC) system could serve the 250 libraries in the NELINET area. The decision to begin the investigation rather than to develop a separate nearly identical system was made after a preliminary evaluation indicated that the goals of the NELINET and the OCLC systems were almost the same. At this time, NELINET and OCLC have agreed to begin the process of replicating the OCLC operation in the Northeast. Inforonics will continue to provide off-line cataloging services to NELINET members until all member institutions have been incorporated into the OCLC system. A second step in replication will consist of transferring OCLC services for New England institutions to a NELINET computer.

Thirty-three terminals located at twenty-seven member institutions will be connected to OCLC during the first quarter of 1973. Other area libraries plan to use the off-line services of the OCLC and still others have contributed to NELINET but have not yet made a commitment to participate in either on-line or off-line activities.

NELINET's focus on higher education may entitle it to preferred rates on telephone communications. Negotiations are now underway regarding the availability of preferred rates. Support for administrative expenses are provided to NELINET by member institutions. Support for specific activities is sought from several sources: for transition to networking from the federal government; for training new personnel from the federal government and from foundations; for communications costs from member institutions, from the federal government and from state governments; and for research and development from the federal government or from foundations. In general NELINET has approached the federal government for support of functions which it considers potentially national in scope and has asked member institutions to support functions which provide direct services to those institutions.

Some problems faced by NELINET have resulted from a dependence upon service bureaus (Inforonics and others with which Inforonics has subcontracted) for computing resources. The use of pricing algorithms based on experience with commercial users leads to constantly increasing costs with constantly growing data bases. In replicating the OCLC system, NELINET will seek to avoid this pitfall. Another major goal of NELINET is replicating the OCLC system will be to build into the system a mutual dependence of member institutions. In the past, the ease of withdrawal of member institutions has caused instability in financial support and in provision of services within NELINET.

The existing agreement between NELINET and OCLC is a three year contract which includes a planned transition to independent operation by NELINET. Mr. Miller forecast a requirement of a minimum user community of seventy member libraries to make an independent operation by NELINET cost effective. At this time costs, implementation schedules, and even the technology of the node are all unstable. However, there is consensus between NELINET and OCLC regarding priorities for implementation and no serious problems are foreseen in implementing the OCLC system in New England and making the transition to independent operation within the three year time limit.

Networks for the Social Sciences

Chairman: Richard Hofferbert, Executive Director
Interuniversity Consortium for Political Research
University of Michigan

James Davis, Director
National Opinion Research Center
University of Chicago

Ronald Anderson
Department of Sociology
University of Minnesota

Recorder: Edward J. Schneider
University of Michigan

This session opened with the description of two organizations, first SIGSOC, the Special Interest Group for Social and Behavioral Science Computing, and then ICPR, the Interuniversity Consortium for Political Research.

Ronald Anderson explained the nature of SIGSOC and informed the workshop of SIGSOC activities. A symposium will be held in June, 1973 at Michigan State University; it will focus on a critical evaluation of the state-of-the-art of social and behavioral science computing. SIGSOC produces a newsletter to disseminate information among its 500 members. Anderson expressed his hope that SIGSOC could provide the vehicle which allows social science computing types to get together and discuss common problems such as the problems inherent in networks.

Richard Hofferbert explained the organization of ICPR. It represents a partnership between the University of Michigan and approximately 170 colleges and universities around the world. The object of the partnership is to pool resources to produce services which could not otherwise be afforded by most members. Membership in ICPR is by institutions, not individuals. Services which the ICPR provides include the collection, documentation and dissemination of data, a software package (OSIRIS), and a training program for social scientists in advanced quantitative analysis techniques. Three archives exist: the Survey Research Archive, the Historical Archive, and the International Relations Archive. Hofferbert explained that merely the existence of these widely used archives have raised the sense within the Political Science and History communities that data collected with public money should be readily available to the scholarly community. Use of archived datasets has risen exponentially over the past five years.

Hofferbert posed the question "We have an organization which uses people to distribute data by mail, do we really need a computer network?" Davis and others argued that the thirty day delay which is common in a system where a user must request data and then wait for a tape in the mail is an intolerable delay for many users of social science data. Davis and others argued that a network would be instrumental in changing styles of research. For research to become truly interactive the data archives must be on-line and available nationally without a substantial time lag. Some argued that a computer network was a necessity to avoid duplication of effort. Only, they argued, with a network, is one freed from having to make copies of data to bring to your local center and freed from the necessity of rewriting programs everytime one moves from location to location.

A computer network was seen as facilitating the sharing of resources. Some noted that before sharing is possible, users must know what exists elsewhere. John Kolp described SSDATA, a newsletter published at the University of Iowa which disseminates announcements of recent acquisitions by thirty social science data archives to eight or nine hundred users. Ed Glaaser thought that users would also like to know what programs are available elsewhere as well as who else has requested the programs.

Networks were also seen as a method for overcoming a growing resistence on the part of social scientists who are unhappy because they must re-tool whenever they move from installation to installation. A network might overcome the need to re-learn the interface.

On the second day the workshop was joined by the economists. George Sadowsky summarized what had gone on in their group the day before. Hofferbert noted that the economists had reached more of a consensus than the rest of the social scientists.

Although no consensus about a computerized network had been reached, there was a common feeling that a range of useful services could be provided by such a network. These services include data storage and transmission, analytic capacity, and information retrieval. Wrigley suggested that only once we have a network will we be able to create uses for it. Hofferbert suggested that a network might help solve three problems. First, it might ease the difficulties involved in getting the data/programs; and it might increase interactive analysis of the data.

Several incremental steps toward implementing networks were suggested. Sadowsky suggested an increase in the number of terminal nodes on the ARPA net. If NSF or some other institution was the sponsor of the ARPA net it might facilitate cooperation with the network. Development of an organizational structure such as the ICPR might facilitate the development of methods of monetary exchange for support of the network.

A consensus was reached on one item — there is confusion among social scientists as to what is meant by a network.

CHAPTER 3

NETWORK DEVELOPMENTS AND DATA BASES

Current Trends in Machine-Readable Data Bases

K. Leon Montgomery, Assistant Professor
Interdisciplinary Doctoral Program in Information Science
University of Pittsburgh

In the past decade the use of the phrase "machine-readable data base" has often referred to particular files such as the Chemical Abstracts Services Condensates file. A somewhat broader use of the phrase seems to be emerging in this decade. This article briefly reviews the trends that seem to have developed and attempts to summarize data bases in relation to today's critical issues in information networks.

NATIONALLY AVAILABLE DATA BASES

Perhaps most attention has been given to nationally distributed data bases. Available from governmental, commercial, and professional society sources, these machine-readable data bases have usually been distributed on magnetic tapes. An important survey of scientific-technical tape services was published in 1970 by the American Institute of Physics (AIP).[1] This survey reported a directory of the ten current information sources utilizing magnetic tapes. A one-page summary including the data base characteristics, frequency of tape issue, average number of source items cited per tape, subscription cost or leasing details, software availability, type of in-house service offered, and publications produced from this data base by the originator was prepared for each tape service. The fifty tape services included in this survey were primarily commercial and professional society sources.

In addition to the cost of the original service, three issues are considered in relating data bases to information networks. These are data base growth trends, geographical distribution of data base producers, and data base overlap. From a geographical distribution point of view, the tape services from the AIP survey are concentrated in New York (15), Ohio (13), Pennsylvania (4), and Maryland (4). Again, it should be noted that a number of governmental information sources were not included in this survey. From a subject matter point of view, chemistry was the predominant subject in twelve tape services, and engineering in seven.

A broader ongoing survey is reported in the *Directory of Computerized Information in Science and Technology*[2] which began in 1968. This directory attempts to create an instrument for the announcement, description, indexing, and dissemination of computerized information collections. It attempts to include all known collections which might be of interest and value to the science and engineering community and which exist either solely in machine-readable computer form or simultaneously in computer and traditional bibliographic forms. Approximately three hundred such collections are described in the current directory. Figure 1 illustrates the geographical distribution of these information sources.

Another important directory is being prepared for data bases in the social and behavioral sciences.[3] This directory is a compilation of information about the existence of, and user access to data bases in social and behavioral sciences throughout the world. Data base as used in this directory is not limited to machine-readable data bases. It is scheduled for publication by Science Associates/International.

GROWTH TRENDS IN MACHINE-READABLE DATA BASES

It is important to comprehend the magnitude of periodical, document, and book production by the world's publishers. *Ulrich's International Periodical Directory* lists over 12,000 titles and estimates world-wide magazine publication at 60,000 to 100,000 titles. Approximately 16,000 of these are published in the United States. It has been estimated that there are approximately 2,000,000 articles published each year relevant to science and technology.

To provide some insight on the control of the scientific and technical periodical literature I have chosen to discuss specifically two major information sources utilizing machine-readable data bases. These are Chemical Abstracts Services Condensates tape files and the Engineering Index-Compendex tape files.

In 1970 the Condensates tape file contained 276,000 articles. In 1971 it contained 308,000. It is estimated that it will contain 340,000 in 1972 and 360,000 in 1973. This source alone then is providing machine-readable access to approximately 18% of the periodical literature of interest to the scientific and technical community. The percentage of English language periodical literature is, of course, much higher.

The Engineering Index-Compendex file has produced access to 44,000 articles in 1969, 58,000 in 1970, and an estimated 85,000 in 1972. Although this amounts to less than 5% of the world's production of articles, it is again much higher in terms of the English language articles.

Biological Abstracts provides machine-readable access to approximately 250,000 articles per year. Summing the coverage provided by these three services, it can be seen that a large portion of the scientific and technical articles can be accessed through these machine-readable tape services. Thus, a continuation of these and similar efforts can provide in time machine-readable access to most of the world's scientific and

COHAN'S DIRECTORY

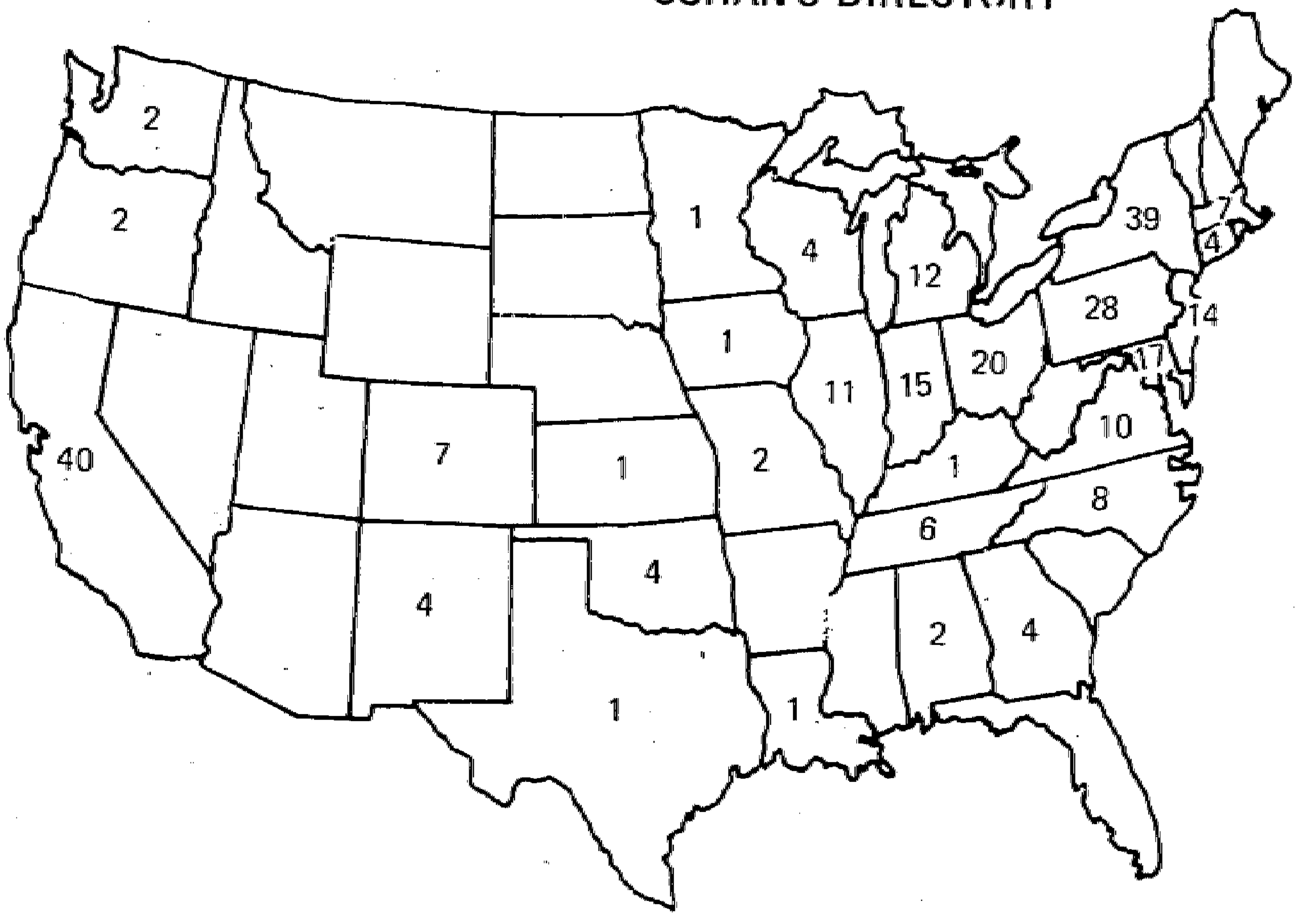


Figure 1

technical periodical literature. Other subject areas are being stimulated to provide similar access, but less optimistic predictions can be made at this time.

UNESCO has compiled the following estimate, which is reported in *The Bowker Annual*,[4] of world book production over the past few years -- 285,000 in 1955, 364,000 in 1960, 450,000 in 1965, and 475,000 in 1970. The Library of Congress receives approximately 400,000 books and pamphlets per year.[5] Of these, approximately 70,000 per year can be accessed through the Machine-Readable Catalog (MARC) tape. Coverage has been limited to English language materials in the past. However, work now in progress at the Library of Congress promises to remove this limitation. Other nations are cooperating in this effort as well. Thus, the percentage of books and pamphlets being made accessible via machine-readable tapes within the next decade should become a significant portion of the world production.

PROBLEM AREAS IN MACHINE-READABLE DATA BASES

There are serious problems confronting these efforts however. A number of these were enumerated in the 1969 SATCOM report.[6] The problems of overlap among major abstracting and indexing services and cooperative efforts among these services are focused on here. Figure 2 illustrates the overlap among Chemical Abstracts Services and seven other major services. These data are quoted from the SATCOM report.[7]

Duplicate coverage among these services appears to be wasteful effort inhibiting yet more comprehensive coverage or improved operational characteristics. Some of the factors militating against extensive cooperation and coordination have been listed in the SATCOM report as:

1. Basic conflicts in the goals, incentives, and constraints which influence the various producing organizations.
2. Fair trade, antitrust, and other legislative acts which can inhibit cooperative efforts to apportion coverage or to reach agreement on pricing policies.
3. Lack of the incentives or resources necessary to effect cooperative arrangements.
4. Inertia or pride in service traditions, which inhibit the discontinuation or merging of services.

It seems probable that some degree of overlap will continue into the future but that the scope of coverage will also continue to widen. Perhaps it will be left to the user community to cope with these and other technical problems such as differing tape formats. Martha Williams, for example, discusses the approach taken by the Illinois Institute of Technology in operating a computerized retrieval system for searching a variety of data bases.[8] This institution is but one of several concerned with this type of problem. The University of Georgia, for example, has been involved in similar activities for a number of years.

% OVERLAP BETWEEN CHEMICAL ABSTRACTS AND SEVEN OTHER SOURCES

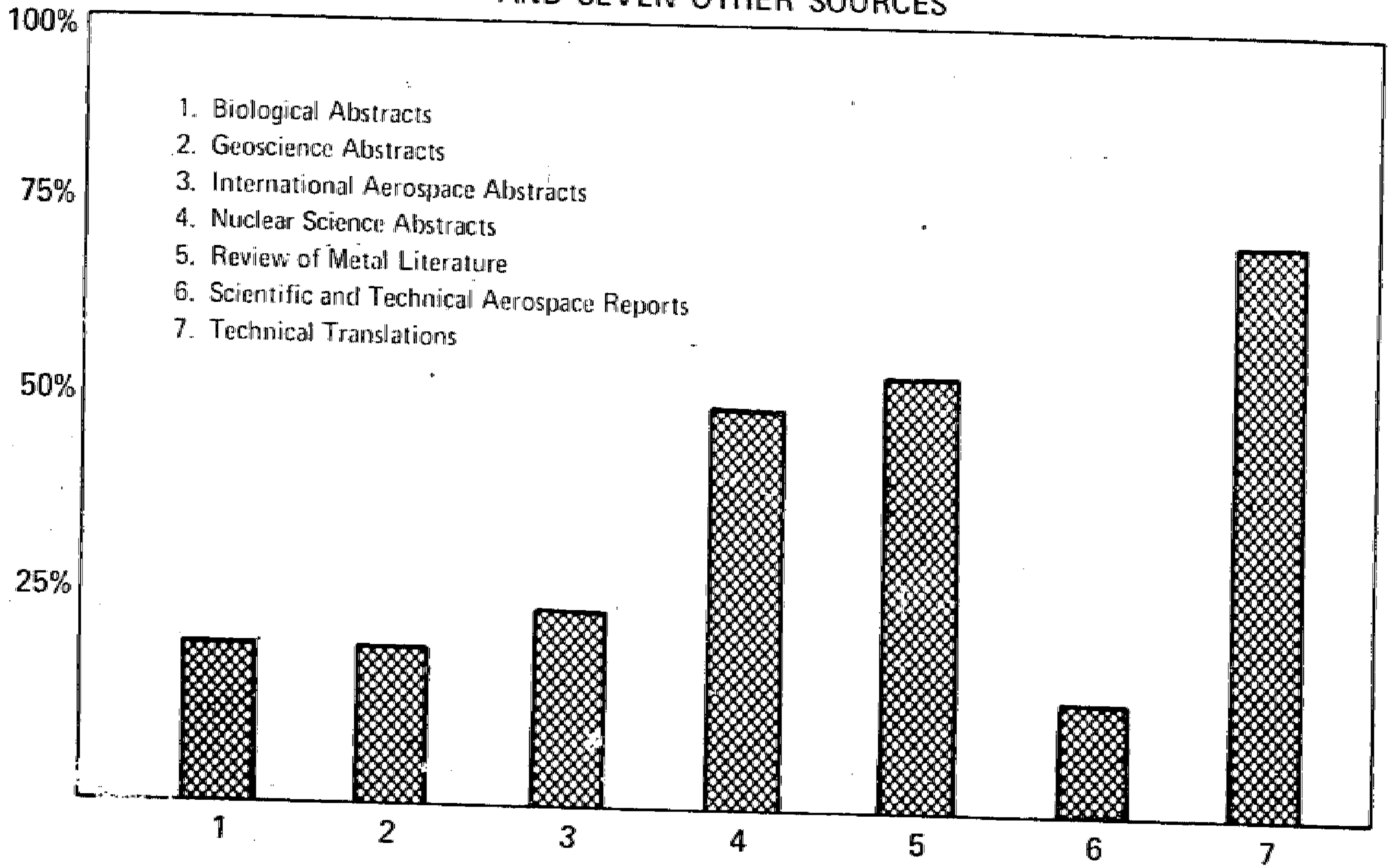


Figure 2

DATA BASE MANAGEMENT

Flexibility of access is perhaps the most pressing problem in dealing with data bases. They are costly to produce and hence are a valuable resource. Modern computer networks will provide a major mechanism for accessing these data bases by many users.

In his recent article Shubert[9] enumerates the three elements of a data base as seen from a data base management point of view. These are: physical storage structure, data and control information contained within a data base, and logical relationships among data stored within the data base. He provides in this article a brief overview of some of the detailed terminology and definitions established by the Committee on Data Systems Languages (CODASYL) Data Base Task Group.[10] This group has attempted to specify an approach for future data base management systems. The present issue is whether to accept the recommendations provided by CODASYL.

A detailed review of the data base management issue is provided in a recent issue of the EDP Analyzer.[11] In this article, data base management is viewed as the following group of activities:

- Storing and retrieving records in a data base.
- Inserting, deleting, and modifying records in a data base.
- Protecting data by denying access to unauthorized persons.
- Guarding the data against errors and system malfunctions.
- Providing for error recovery.
- "Cleaning up" the data base after numerous insertions and deletions have occurred.

Most of these activites can be software oriented.

In their April 1971 report, the CODASYL Data Base Task Group proposed the following main features for a data base management system.

- A data descriptive language for defining a data base. This language is dependent of any programming language.
- A COBOL data descriptive language for interfacing data bases with COBOL programs.
- A COBOL data manipulative language for storing, retrieving, inserting, deleting, etc. of records in a data base.

The details of the proposed languages are beyond the scope of this article, but the debate over data base management has been superbly summarized in the EDP Analyzer article cited.

The importance of the data base management problem is perhaps best highlighted in the recent National Academy of Sciences' report on *Libraries and Information Technology – A National System Challenge.*[12] One of the findings on which attention is focused is quoted as follows:

> The primary bar to development of national level computer based library and information system is no longer basically a technology feasibility problem. Rather it is the combination of complex institutional and organizational human-related problems and the inadequate economic/value system associated with these activities.

Data bases and data base management problems seem to be at the heart of the complex institutional and organizational human-related problems. EDUCOM certainly provides a suitable forum for the consideration of such complex issues.

REFERENCES

[1] Kenneth D. Carroll, *Survey of Scientific Technical Tape Services* (New York: American Institute of Physics, 1970), Report No. AIP ID 70-3.

[2] Leonard Cohen (ed.), *Directory of Computerized Information in Science and Technology* (New York: Science Associates/International, Inc., 1968).

[3] Vivian S. Sessions, *Director of Data Bases in the Social and Behavioral Sciences* (New York: The City University of New York, University Graduate Division).

[4] *The Bowker Annual of Library and Book Trade Information 1972* (New York: R.R. Bowker, 1972.

[5] *Annual Report of the Librarian of Congress: For the Fiscal Year Ending June 30, 1971* (Washington, D.C.: Library of Congress, 1972).

[6] *Scientific and Technical Communication: A Pressing National Problem and Recommendations for Its Solution* (Washington, D. C.: National Academy of Sciences, 1969).

[7] *Scientific and Technical Communication: A Pressing National Problem and Recommendations for Its Solution* (Washington, D. C.: National Academy of Sciences, 1969), p. 146.

[8] Martha E. Williams, "Experiences of IIT Research Institute in Operating a Computerized Retrieval System for Searching a Variety of Data Bases," *Information Storage and Retrieval*, 1972, p. 57-75.

[9] Richard F. Schubert, "Basic Concepts in Data Base Management," *Datamation*, July 1972, p. 42-47.

[10] *Report of the CODASYL Data Base Task Group, April 1971* (New York: Association for Computing Machinery, 1971).

[11] "The Debate on Data Base Management," *EDP Analyzer*, Vol. 10, No. 3, March 1972, p. 1-16.

[12] *Libraries and Information Technology: A National System Challenge* (Washington, D.C.: National Academy of Sciences, 1972).

NSF Network Initiative

D. D. Aufenkamp, Office of Computing Activities
National Science Foundatio

The concept of a National Science Computer Network was advanced at the EDUCOM Spring Conference.[1] Such a network, it was argued, could provide its users with access to computing facilities, information systems, data banks and other computer-based resources without regard to geographical considerations. The implications for resource sharing in research and education, would, indeed, be far-reaching. From a single point of access to the Network, a user would have available -- literally at his fingertips -- a vast array of resources and services. The possibility of a much closer coupling of computing and information systems would open new worlds for both the conduct of research and the timely dissemination of its findings.

In July, a Foundation brochure[2] set forth a program designed to undertake exploration and evaluation of the many dimensions and facets of the resource-sharing potential of a nationwide computer network. This program is being mounted jointly by the Office of Science Information Service and the Office of Computing Activities and has as its objective the providing of specific information as to concept and feasibility of a National Science Computer Network.

Such a network in terms of support for research and education is much more a question of people and resources than it is one of computer communications. Thus, it must be stressed at the outset that this program is, for the most part, complementary to the technology. The Foundation is not proposing to implement and operate a major computer communications network. The technology has been advanced to the point where feasibility of a national network of computing resources, transparent of the user, is no longer an issue. Today it might be implemented with land lines, tomorrow with satellite-based communications. But, a great many problems remain to be resolved of how institutions as well an individual researchers and instructors could avail themselves of network resources in the current complex computing and science information milieu. These problems, as it is well-known, are not necessarily technical in origin although clearly the technology plays a key role; they are organizational, political and economic. This Foundation initiative is addressing these and related problem areas.

To illustrate by example, it might be interesting to speculate on the extent to which anticipated savings in institutional computing expenditures would prompt an institution to reduce its campus facilities in favor of network services? Would 10 per cent be sufficient? Perhaps not. Would 20 per cent suffice? Possibly. But how would a decision be reached? Most of the discussion one hears seems to be concerned with how to get on a network. What happens if an institution chooses to withdraw from a network or can no longer make its resources available, particularly if users elsewhere have become dependent on these facilities for special systems or data? What impact would there be if an institution that has been providing outside services to produce needed income to maintain facilities locally, loses this work to a "competitor" in the marketplace of a public network?

There is no need to belabor these questions. The point is that while the technology affords us with a new dimension to improving the sharing of resources, this improvement is not accrued automatically with the advent of the advanced technology. We hope to address the many organizational, political and economic issues in a way that the technological advances in computer networking can be exploited to the fullest extent possible.

The Series of General Working Seminars on a National Science Computer Network[3] to be conducted by EDUCOM is indicative of the thrust of the Network initiative. It is vital to gain the active participation of all who might have individual or institutional interests in exploring the resource-sharing potential of a national network. This series of seminars is but one step in developing a structure for a continuing dialog to assist in gauging the community's sense of direction, interests and requirements. Clearly, there will have to be many more in the months and years to come to address specific issues raised.

The Foundation brochure delineates the scope and objectives of the Network thrust. The approach is one of undertaking a comprehensive and interrelated set of research project activities to address the relevant problem area. How one addresses the issues related to utilizing a national network is crucial. We have already argued that one should not assume a "solution" in terms of a specific network technology before the user characteristics and needs are known. We would also argue that it would seem unlikely that questions of network management or resource sharing could be addressed on the basis of examining two or three potential network nodes in a way that would permit meaningful extrapolation to the national scene. It would also seem especially important, in order to make any project experiences as useful as possible, to avoid creating artificial environments which would have a marked impact on the validity of any observations made.

The network brochure suggests that a trial National Network would be integral to these activities in the nature of a test vehicle. To the participants in the trial Network the experience should be "real" to permit meaningful observations to be made. To the rest of the world, however, it should be in the nature of a well conceived and functioning demonstration. The immediate challenge can be viewed as that of formulating this comprehensive and interrelated set of project activities to explore fully the resource-sharing potential of a National Science Computer Network. These project activities must incorporate a rich and whole hearted collaboration of institutions and individuals and, in effect, they constitute the proposed trial Network. They must be structured to permit a thorough and convincing evaluation of the experiences from a multifaceted perspective to establish the boundaries of this approach to resource sharing. Clearly it is an exceedingly complex undertaking. It will not come to pass merely because of a Foundation announcement. If the undertaking is successful it will be so because of a combination of efforts on the part of many. This, then is the challenge that confronts us here today.

REFERENCES

[1] Proceedings of the EDUCOM Conference: Networks for Higher Education, EDUCOM, April 1972.

[2] NSF 72-16: Expanded Research Program Relative to a National Science Computer Network.

[3] See announcement of A Series of General Working Seminars on a National Science Computer Network issued by EDUCOM, September 1972.

An Overview of The Merit Computer Network

Bertram Herzog, Director
MERIT Computer Network Project

Let me make only a few remarks to give you some background about the MERIT Computer Network. Compared to others I have the advantage that you came to the network. We have arranged several demonstrations to let the Network tell its own story.

The Michigan venture into networking began in 1966 when representatives from the three large institutions met to discuss the prospects for forming a network. The State Legislature supported these actions by appropriating funds. These funds were provided subject to the condition that matching funds be obtained. The National Science Foundation provided the matching funds and a project was launched, July 1969.

Today, just over three years later, the network is operating. To reach this point, we spent the first year getting to know each other better. This phase was aided considerably by the interest and influence of the principal administrative representatives of the universities, Drs. Hubbard, Muelder and Smith. I can only reiterate Vice President Smith's comments noting the favorable evolution of our cooperation from the early days to the present days of very frank dialogues.

During the first year the specifications for equipment and software were refined. Negotiations to purchase or build equipment followed. Later in that year these matters were resolved, software developments began, followed by frustrating delays until the equipment was installed permitting the first network testing. Thus by December 1971 we had what engineers, system programmers and project directors call a network. By April 1972 we had what users might call a working network of just two nodes. Happily, just in time for this meeting we have the complete network in operation.

How did we organize ourselves to get to this point? The central network staff location of Ann Arbor was selected largely due to my personal affiliation as professor at the University of Michigan. An integral part of the project is the set of groups of system programmers loaned by the participating computing centers.

In addition three associate directors are appointed by the universities, one at each university. Their primary involvement is not in building the network. They look ahead to address questions relating to how the network will be used and who will disseminate information. They have worked hard to find answers to these important issues.

Further details about the network will be offered in the workshops. Eric Aupperle will present a description of the network and show how it can be used. Another workshop will discuss the mechanisms for exchanging money between universities to pay for computer resources exchanged. Karl Zinn, one of the associate directors will lead a discussion on user problems.

May I take a couple of moments to leave a few thoughts about networks with you. What a network is becomes even clearer as one builds one. Still we have difficulty describing the network, and more important its behavior, to visitors. Frequently we must start with a demonstration of one of the time-sharing systems to serve as a basis for the network demonstrations. Please do take the opportunity to see these demonstrations; I think you will find them illuminating. They show how to access the network for general computing, the use of data base, and process to process demonstration.

This last is a difficult, interesting but most challenging use of networking. Process to process communication occurs when a program resident in one computer talks to a program in another computer. The authorship of these programs is not important. The quality of authorship is a factor. We will demonstrate how a process in computer A can talk to a process in computer B. One may extrapolate from this that the richness of the process in computer A and the richness of the process in computer B depends upon the available resources and how well a user knows to use them. In other words: as one wishes to use the resources of computers in a network today, so you must know the detailed mechanisms to access those resources. Simply, if you want to use a CDC machine, you had better know its operating system protocols or how to get at those resources. Were I to state anything else, I would be deceiving you. Obviously one would like it otherwise.

Time is running short. Permit us to tell you more about the MERIT Computer Network at the workshops, demonstrations, and at coffee breaks. Thank you.

Science Information in a Changing World

Edward C. Weiss, Office of Science Information Service
National Science Foundation

This presentation will cover several of the major thrusts of the total OSIS program with emphasis on support for networking and resource sharing in the university community. Both our research and university systems programs are highly dependent upon networking as a means for the accomplishment of their various goals and objectives. In fact, networking will provide an important interface between these two programs by way of reducing the lag-time in the introduction of new techniques and methodologies to existing and future systems.

In the 1960's OSIS supported both the research and development which led to the existing first generation information systems. Networking is the culmination of these efforts. At the same time it opens the door to the technology of the 1970's by providing a real world environment for the identification of new problems and the validation of their solutions.

As many of you are aware, OSIS has been supporting system development efforts which are primarily based in the scientific professional societies and the university sector. The society-based systems produce the information bank (i.e., abstracts and indexes of the world's scientific literature) and supply tapes recording its content. The university-centered systems provide the distribution outlets to the user community.

The university-centered information systems have had substantial computer facilities, monetary resources, and interested faculty and staff available to permit the development of such projects. They have also made significant dollar investments in the projects during the course of their development. The university administration, in every case, is committed to providing the necessary funds for the long-term operation of the system past the period of NSF support for the development effort. But not all universities and colleges have comparable resources, nor is it reasonable to expect the NSF to fund the development of such systems at every institution of higher learning in the country. Other means must, therefore, be found to provide these resources to the faculties and staff of institutions which presently lack adequate access to modern scientific literature sources.

Our experience with the support of university-centered science information systems over the past several years has indicated that they rapidly expand to service regional areas. In this way, the faculty, research staff, and graduate students of smaller institutions are provided access to the same scientific and technical literature as their colleagues at the larger universities. But more importantly, the future will see a strong emphasis on networking which will incorporate such regional centers as major nodes. Moreover, the costs associated with system development and operation are rapidly making it clear that such activities will become an economic necessity in the future. Much work remains to be done to foster these types of cooperative arrangements, but much has already been accomplished. The opportunity to interface OSIS-supported information systems with OCA-supported regional computing networks has been a major factor. The necessity for library/information system interfacing in order to solve the problem of document delivery has also accelerated the process.

The problems which will pose the major obstacles to successful efforts in networking of information resources are not judged to be primarily technical in nature. The areas which require immediate attention and where it is intended that initial support be focused are concerned with administration, user requirements, economics, and legal matters. For example, the optimum mechanism for the administration of a network which may consist of a loose consortium of universities cannot at this time be specified with any degree of precision. When this problem is compounded by the factors of geographic dispersion and regional interests on a nationwide basis, the ramifications of this concern only begin to emerge.

Thus, the communication system of science and technology is a melange of procedures, media, production techniques, services, centers, and people. The system has evolved by way of necessity and expediency, and in its current stage of development exhibits many different types and levels of complex interactions. Many services and institutions, aside from the university-centered information systems, are operational at the present time, and these will also function as nodes in whatever network will evolve. Such a network undoubtedly will be evolutionary, in the sense that it will develop as circumstances permit, rather than being systems engineered *a priori*. The present resources in operational components of the network represent an investment too large to scrap in favor of an optimal design even if it were possible to design the optimal network. Because of this, it becomes necessary to plan a course of action flexible enough to accommodate different strengths and relationships among the existing nodes which include the tape suppliers and publishers, as well as the distribution centers and libraries.

UNIVERSITY-CENTERED INFORMATION SYSTEMS

We have been supporting the development of multi-disciplinary, literature-search systems at six institutions — Lehigh, Georgia, Pittsburgh, Ohio State, Stanford, and UCLA. Lehigh, Georgia, and Stanford have now received final funding.

We have also supported the development of two specialized information systems: the Treaty Information Center at the University of Washington, and the Arid Lands Information Center at the University of Arizona. Although these are also located at universities, their purpose is entirely different from the six just discussed. Their purpose is to accumulate all information in a specific field, regardless of its type, discipline origin, or form and to provide a comprehensive computerized reference service to research workers in that field wherever they may be located -- on campus, throughout the nation, or even in another country.

Objectives

The objectives of this program, which were formulated as early as 1968, derive from the fact that the main locus of the overall research support functions of the Foundation is located in the university environment. Thus, our *immediate* objectives are threefold.

First, it is necessary to enhance the dissemination of information among academic research scientists and the graduate students they are training to become the scientists of the future.

Second, it is necessary to establish campus-based terminals to accept the end products of the discipline-oriented systems based in the professional societies, as well as those from federally-based systems and commercial suppliers.

Third, it is necessary to support the development of major nodes for an emerging national science information system.

The *ultimate* objective is to support the development of information systems which will be much more highly interactive with the research process than the first generation systems whose development we are currently supporting. Such systems will eventually expand the function of information retrieval to form the basis for true science communication systems of the future. As such, they will combine the elements of information retrieval, data reduction and manipulation, and communication.

However, the role of the Foundation has been greatly expanded since this program was initiated in 1968. It is no longer solely concerned with the support of basic research in the academic environment, although this indeed remains the primary emphasis. But obviously such programs as RANN and the newly announced Research Incentives and Technology Assessment programs portend, perhaps, an even more expanded scope for the future.

Such changes are in recognition of the fact that there has been a shift in national priorities which calls upon science and technology to aid in the solution of major problems which confront our society. Thus, we must now begin to include among the objectives of the university-centered information systems program, as well as our efforts in networking, and the Research Program, of which I will speak shortly, a focus on the information transfer linkage between science and society, utilizing to the maximum the potential provided by the rapidly emerging communications technology.

Future Trends

We are making some substantial changes in regard to the future support of multi-disciplinary, literature-search systems. We, of course, intend to meet our commitments to the three – Pittsburgh, UCLA, and Ohio State – which still require further funding. But in the future, we will explore some new directions. The existing systems all used different approaches, custom-fitted to the specific requirements of their own academic communities. Each has been successful in its own way, but none was purposefully designed for widespread transferability. Henceforth, working with other universities, we plan to emphasize the prototype intent – that is, the development of systems or components as basic models or patterns that can be adopted or adapted by any university wishing to do so, without Federal subvention being an essential.

We also intend to broaden the base of such activities to include all aspects of the information-transfer process on the university campus. The computerized science information systems enhance the alerting and retrieval functions, but there remains the need to improve the document delivery system in order to complete the service. University libraries are today facing critical functional and economic problems in meeting the demands for traditional reference materials. To help relieve these problems, we plan to extend our support to such efforts as improved document delivery methods and other library operating and management requirements. With such improvements, libraries will be able to integrate the computerized search systems and function as "knowledge service centers" within the constraints imposed by existing first generation technology.

Further, it should be noted that the techniques developed by university libraries are largely transferable to public, private, and industry libraries. The university library community expects to extend its resource-sharing capability with the non-academic library community. Thus, our support promises substantial benefits to the total United States library community.

In regard to the future aspects of our support for specialized information systems such as the Treaty and Arid Lands Centers, we intend to limit our support, by concentrating on a few areas that evidence high priority in the view of the research community and the Foundation, and which are outside the scope of other Government Agencies and beyond the interest of commercial organizations.

Recent Developments

The Environmental Protection Agency (EPA) has recently announced a program of cooperation with science information centers developed under OSIS grants. This cooperative agreement was undertaken with the firm conviction that existing systems and services should be used whenever possible, rather than generating new ones.

Under this program, the information centers of three universities will be coordinated for EPA use. Lehigh University has been selected to provide on-line services, the University of Georgia will be responsible for batch and retrospective services, and Ohio State University will produce selected SDI searches on a trial basis. This program will permit EPA users to have access to over 25 data bases, while determining usage requirements of the various files and services. These data bases include those in the traditional scientific disciplines and engineering, as well as environmental tapes supplied by EPA which cover such areas as water resources, air pollution, and solid waste disposal.

This arrangement will serve a network of some 40 EPA libraries and information centers, with the Library at EPA's National Environmental Research Center in Cincinnati, Ohio, designated as access point to the information centers. Eventually this network will be expanded and EPA has already held discussions with Stanford University in regard to providing services in the San Francisco region.

As a result of this EPA action, the National Oceanic Atmospheric Administration (NOAA) has approached Lehigh University toward establishing a method of indexing or bibliographic control for the mass of raw data now being collected via remote sensors in the atmospheric and ocean sciences.

This is an important first step, but the full impact cannot be realized until one recognizes the type of snowballing effect which must inevitably follow. These data are not being collected as an end but rather as a means for new research. They will be manipulated to produce new understandings which will eventuate in new data and reports. We are now confronted by a situation which, by virtue of its sheer magnitude is qualitatively different from the information explosion as we have come to know it. It literally poses the question of what to do when even the capacity of a trillion bit storage device is exceeded. We hope to act before our vision once again becomes hindsight.

THE RESEARCH PROGRAM

In the face of the proliferation of scientific information combined with the cessation of almost all other Government-supported information research and urged on by the rapid advances in computer technology, the Research Program is confronted by a two-fold challenge: (1) to develop the basic technical knowledge necessary to build more effective information and data retrieval systems than those currently envisioned; (2) to extend our fundamental understanding of the information-transfer process.

Objectives

In order to meet this challenge, the objectives of the Research Program are summarized under two major categories. The first objective is aimed at supplying the essential knowledge, methodologies, and techniques to facilitate on-going system development efforts. The second objective is to provide the theoretical framework on which future developments will depend.

Program Implementation

In order to implement these objectives, we are proposing a program of research which will cover the broad spectrum of activities ranging across basic research through applied studies and pilot demonstration models. These categories are delineated as follows:

Under *Fundamental Research* we are considering such things as:

- Theoretical Studies
- Communications Patterns
- Information Flow
- Policy Studies
- Behavioral Studies

Under *Applied Research* we have listed such areas as:

- Quality Control
- Requirements Analyses
- User Studies
- Large-File Management
- Retrieval Strategies
- Library Functions
- Evaluation Measures
- Systems Modeling
- Resource Allocations
- Management Studies
- Data Manipulation

Under *Technology Applications* we foresee support for projects dealing with:

- Micrographics
- Pilot Demonstrations
- Systems Configurations
- Terminal and Display Devices
- Mass Storage Devices

Program Definition

We have recently begun to define a thrust and direction aimed at the integration of these various categories toward a specific goal. There is an increasing recognition and awareness that the research and development efforts of the 1960's have yielded a first generation technology which is primarily focused on the application of computer technology to existing practices and procedures. Such a statement is not intended in the pejorative sense. These efforts which will culminate in networking were an essential first step.

But now that we have achieved our current vantage point, a critical examination reveals that we are still manipulating a document bound system which has persisted since the invention of the written symbol. We must now begin to examine the possibilities of an information or knowledge transfer system as opposed to a document transfer system.

In our initial thinking, this seems to imply the need for a reexamination of the way in which scientific and technical information is packaged, stored, and accessed with a special emphasis on the real desires and requirements of the user. Further, in line with the expanded scope of the Foundation, we must begin to focus on the needs of various classes of users. First, of course, is the established researcher who has been the primary recipient of our past concerns. Second, are the students who will serve the nation in various capacities in the future. Third, is the class of practitioners who range from engineers through city planners and social workers. The fourth class consists of decision makers of various persuasions such as legislators at the Federal, state, and local levels, as well as managers in industry and those responsible for the administration of other types of institutions. Finally, the time is fast approaching when the average citizen will also need access to such information in the course of his daily pursuits as consumer, voter, etc.

The implications of the above in terms of the future role of information specialists at both the scientist and technician levels are enormous. Whole new curricula will have to be developed in order to develop the skills and insights necessary to perform the subtle correlations and transductions on the input material to produce the kinds of outputs which will be required.

Each of these points merits a more elaborate and detailed discussion which time does not permit. My opening remarks dealt with the networking efforts which we are beginning to explore jointly with OCA. The goals which I have outlined for the Research Program will be dependent upon breakthroughs in such areas as hierarchical programming, artificial intelligence, hardware technology, etc. Thus, there is the requirement for another important interface with OCA and we are beginning to examine the possibilities.

THE DATA SYSTEMS PROGRAM

I will conclude my presentation with a brief discussion of our newly organized Data Systems Program. Present literature-based abstracting and indexing systems identify the many documents containing the "how" and "why" of research and development, but provide no access to the "what" – the factual and quantitative data which the documents may contain.

This access is a prime requirement for scientists and engineers and is becoming increasingly essential as data accumulate, both in traditional fields and in broad new areas such as social indicators and environmental quality. The data are now widely scattered in different forms throughout millions of articles and reports, making the retrieval from literature systems extremely difficult, sometimes impossible.

With the publication crisis now confronting many professional societies, this problem will be intensified. Publishers are reducing the size of their journals to keep costs within feasible limits; with smaller publications and shorter papers, they are omitting graphs, tables, and charts whose data content may well be lost if no systems exist for their storage and retrieval.

While the development of systems to meet this need is of essential importance to scientists and engineers, it is equally critical from an economic standpoint, since the presence of effective data-handling capabilities will relieve literature and libraries of a vast amount of processing and reference work now required to satisfy, however ineffectively, the demands for specific data.

It is in recognition of these pressing needs that OSIS has created a program dedicated to data systems, rather than continuing to treat them as a relatively undefined subset of information systems generally.

Establishing OSIS in a position of leadership in the sense of its being the dominant force in this broad realm is clearly unrealistic -- fiscally, functionally, and politically. But establishing it as a central catalyzing agent -- a focal point for the identification and recognition of widely varying interests and programs, and for supporting efforts to fill gaps in a total framework -- seems not only realistic, but essential, since no other organization is now attempting this vital function in any comprehensive manner.

While the program is still so new that many of its dimensions have not yet been precisely spelled out, we intend to try to do for data systems almost exactly the same thing we are doing for information systems. Our efforts must encompass both disciplines and subdisciplines, as well as problem-oriented endeavors in many areas of technological concern -- in short, any field of science and technology where discrete accumulations of significant data exist and are not adequately available without the development of computerized systems for their organization, manipulation, and retrieval.

In approaching this broad requirement, OSIS intends to concentrate its limited resources upon the development of methods and techniques, drawing wherever possible upon existing technology, including that already established for documentary systems. As with information systems, our emphasis will be on the creation of improved software or processing techniques that can be transferred to related systems or applied against requirements in other fields.

CHAPTER 4

IMPACT OF NETWORKING: BANQUET ADDRESS AND WORKSHOPS

Social Information: The Technology, The Need and The Challenge

Banquet Address
October 12, 1972

Edwin B. Parker, Institute for Communications Research
Stanford University

I'd like to talk about cable television and education, with the emphasis on education rather than on cable television. I found as I listened to the proceedings today that some of what I'm going to say tonight has been anticipated. Vice President Allan Smith, in his remarks, talked about the glittering futures of communication technology and the possibility of an outward look for universities, to "Open Universities" or "Universities Without Walls". It's that possible glittering future that I want to talk a little bit about tonight. The other way I was anticipated somewhat was in Ed Weiss's remarks this afternoon. He was talking about changing national priorities and how those interested in science information should also look at how science can be related to society and how our science information systems could provide science information for the whole society, not just scientists.

Some of what I'm trying to say was adapted from an article in the June 30th issue of *Science* in which, along with my co-author, Don Dunn, we had an article entitled "Information Technology: Its Social Potential".

Before I discuss the glittering future, I'd like to carefully distinguish between possible futures and probable futures – because my message on cable television tonight is that I have some good news and some bad news. The good news is that the glittering future of communication technology for expanded educational service at reduced unit cost is really possible. The bad news is that it isn't likely to happen.

I think a lot of us who have written on the subject of cable television and its great possibilities are being used. We've been creating a myth about cable television, a myth that the National Cable Television Association and the people who stand to gain profits from that industry are finding to their advantage. They like the myth that all of this glittering future is going to happen because there's profit in it for them. But when you look underneath the stone, it's unlikely that the kinds of optimistic things we're talking about, the educational potential, will really happen. So when we talk about the glittering future of cable, we ought to distinguish carefully and remember that that's a possible future that is very unlikely to happen unless we do something to bring it about.

There are three themes to this talk: The Technology, The Need, and The Challenge. And since I'm not a technologist with a solution in search of a problem, and am a social scientist with a vision of a different kind of society, I'll start with the need first and then go on to talk a little about the technology and conclude with the challenge. The challenge is to EDUCOM because I think this is the kind of an organization that can make possible some of the things that some of us are dreaming about.

I think we have a tendency to supress needs that we don't know how to meet, or we call them something else other than needs until we perceive some way of meeting them; and I think in our society now we are beginning to articulate a vast, unmet educational need because now we can begin to see a glimmer of how to meet that need. So let me start with my three goals for what I would like to see the educational system of our society doing in, say, the year 1985. I have three goals for the education system of our society. One is equal opportunity of access – and I mean equal opportunity for rich and poor, brown and black and white, high IQ and low IQ, men and women; I mean equal opportunity for everyone in the society.

My second goal is in a sense a subset of that first one and that's for lifelong learning. Too often these days education is the province of the young, and yet the society is changing at such a rate that we all need continuous retraining. We need to go on learning through our entire lifetimes, and our educational systems really haven't been designed with lifelong learning in mind.

My third goal for education is sufficiently wide diversity of content such that anyone who wants to learn can learn whatever he wants to learn. The world is complex and growing more so. There are many subjects and many topics, not all of which are accredited kinds of topics in our educational institutions but are things that people want to learn. What I'm saying is that I'd like to deprive no one in our society of educational opportunity. I have no patience with those who say, "Don't waste your time with education for the masses – they only want to drink beer and watch football games on television". I'm saying we ought not to deprive them of that choice or assume that most of them will make a different one. If ninety per cent of them choose to watch entertainment television, that still doesn't entitle us to deprive the other ten percent or whatever fraction would be interested in something different. In any case, I suspect we would have a different kind of society in another generation after we've brought up a generation of children who have grown up in an environment where they could learn whatever they wanted to learn from earliest childhood without having a kind of system that unfortunately in all too many cases tends to punish people for curiosity rather than rewarding them.

When I think of that kind of goal, that kind of educational system, and look at the present, of necessity labor-intensive educational activities, and try and project present services and present institutions onto that need, the projections don't make sense. You can project on to the year two thousand when you might wake up one morning and find that half of us are teachers and half of us students and that we'd have to reverse roles in the afternoon. But even that understates the problem because given the variety of topics that we would like to be able to learn, we would like to have a different private tutor for different subjects; and so the number of teachers necessary is n times the size of the population, and the projections quickly become really absurb. And here's where I think our technology — our *potential* technology, the electronic tutor that doesn't yet exist in economical operational form – really has a promise.

In talking about educational technology I'm not talking about replacing teachers in present institutions. I think it's going to be very difficult for institutions of higher learning to fire enough of us tenured professors to really reduce costs and replace us by machines. But I think the new technology will permit us to serve the unmet needs for education in the society.

What I'm talking about is not the present generation of cable television, not one-way television service distributed by cable into homes rather than by over the air. The present generation of cable television is like any new medium of mass communication. The first content that gets put into the new medium is the content of the old. It takes some time for the content to evolve to fit the form of the new medium. And in any case, the cable medium as we now know it is in the next ten years going to undergo a rather dramatic conversion, a real metamorphosis, I predict. What I'm talking about is two-way cable television with digital response capability from each television set such that the television set is a terminal for computer aided instruction and information retrieval, and a terminal for student-response television. In student response television, the television lecturer, although presenting the same motion video message to a large class, is able to have each student in the class with his response keyboard actively practising responses during the course of the presentation rather than passively viewing. The answers can be aggregated simultaneously so that the instructor can see how badly his message is getting across instead of just inferring from the fact that he's focusing on the one bright face in a classroom that maybe it's getting through somewhere. Those same responses that can be available on an aggregate basis in real time can be the basis for a computer managed instruction program where supplementary materials, either remedial or enrichment, can be selected to be distributed to individuals on the basis of their unique pattern of responses to the instructional material.

When we add video cassettes with the potential of bringing motion, color, stop-action and instant replay under the student's control, and add communication satellites for interconnection such that small audiences, classes that are too small to be economically viable in one community can be aggregated across the nation in order to create a reasonable market size, then I think we're talking about a quite different kind of technology for education to individuals than we have now. And I'm inclined to suspect that that kind of educational technology will create as much a difference in our society as did the introduction of printing into the oral and scribal culture that preceded it.

Obviously, in presenting this kind of vision of the kind of educational service that our society could be providing by 1985, I don't think I'm talking about our present institutions and how we could save a dollar here or fifty cents there or how we could offer a slightly improved service to our present clientele without raising the cost too much. I'm talking about a society in which people can learn what they want to learn, without hassle, without necessarily meeting formal prerequisites or selection processes. If they think they can hack it, let them try; let them find out. Let them demonstrate by their performance or nonperformance. When they're not preempting a scarce teacher resource, we don't need to be so selective about who can try. Let's have a kind of system whereby people can get credit for what they demonstrate they know, with failures unrecorded so there's no stigma attached to trying, and a system in which people have a chance privately to learn from an electronic medium to which it doesn't matter that they expose their ignorance. Often we're afraid to ask questions because we don't want to show our ignorance to other people. That can be changed. So people can learn what they want to learn in their own homes or in neighborhood store fronts, in a wide variety of different places.

Now I'm not predicting that this will in any way reduce the total cost of education. But we can dramatically reduce the *unit* costs of education such that it's really economically feasible to provide education for the whole society. I think we'll be able to afford a level of education that's unthinkable now. One analogy is printing. The printing technology certainly did not reduce costs for the educational enterprise. The costs of education have been dramatically affected by that technology. They've gone up astronomically. I think this new technology will also increase the total cost of education, but by reducing the unit costs dramatically can make it more widely available.

Another much closer to my home analogy comes from the development of a computer information system at Stanford called SPIRES that I had something to do with. Five years ago when I was getting this project started, I went to Dick Lyman who is now our President (at that time he was Provost) and explained to him that this was a research grant that the university ought not to accept unless they realized the implications of it. This was a project that, if we succeeded, had severe budget implications for the institution. He needed to know that before the institution accepted the grant in the first place. And sure enough, about five years later, now that that project has been successful, he's in the position where Stanford's budget is larger by a small number of hundreds of thousands of dollars because of the demand that was generated within the institution from the kind of information retrieval service that we built. The benefits are not cost saving benefits; the benefits are those of being able to do things that you couldn't do before. These increases came at a time when declining income forced major cuts in expenditures in other areas. And this leads us to all kinds of economic problems. It's true that there are some examples where new computer technology in fact will save costs and I think Fred Kilgour's example at the Ohio College Library

Center is one of them, where it permits things to be done cooperatively at cost savings. But I think those examples tend to be the exceptions rather than the rule, and I think by and large we're talking about increases in budgets rather than cost savings. So how are we going to pay for them? How is our society going to pay for this kind of educational future that I've projected? Where is the money going to come from?

It's not going to come from the private sector of the economy. Education is already firmly entrenched in the public sector. It's not likely to come from existing educational institutions because their budgets are already pinched tight. In any case, the scale of present institutions is probably wrong for the kind of enterprise we're talking about. It may be that state funding will be available on a sufficiently large scale. States in lieu of expanding community college systems may find that rather than building more and more community colleges, it is more economical to build electronic colleges. But that's only if the technology is there and available to be operated as an operational service. But who's going to pay for the research and development and demonstration funds to get us to the point where it's possible for state legislators to even seriously consider that? I think that leads us to the necessity for extensive if not massive federal funding for research and development and demonstration to get us to the point where we have feasibility demonstrations for these kinds of education systems.

Let me switch from that to the technology of cable, and where it looks like cable is going – which is at this point a slightly different direction than education. As a result of the FCC rules on cable that went into effect on March 31st this year, that finally after many years unfroze or at least partially thawed the freeze on cable television, we're now embarked on a decade of growth in cable television that will be analogous to the growth of broadcast television in the decade of the 1950's. The new FCC rules have shifted the climate in the financial community such that the large financial institutions, the insurance companies and so on, are now for the first time, willing to loan the capital in order to get the nation wired. We're going through a period of consolidation of small cable companies into large conglomerates which is very reminiscent of what happened in the early days of the telephone, when we had many independent companies that were ultimately merged and gobbled up into one giant.

The FCC rules require two-way communication capability in all new construction in the top hundred markets (where 85% of the people are) which today have been for the most part not wired. That by itself would not ensure that cable had the two-way capability that some of us would like to see, because if there were not economic incentives as well as FCC rules, the industry would find a way to stall or get waivers. But the economic incentive is there, also. That economic incentive is pay TV. Cable operators want that digital response capability from each home so that we can push button A to say, "Yes, I'm willing to pay for that movie," or for the teleshopping functions where you can punch a button B that says, "Yes, I'm interested in that product; send me more information." And the major manufacturers of cable television equipment are now competing with two-way technology. Some of them have twelve button pushbutton pads; others have sixteen button pads; some of them are showing how you can interface alphanumeric keyboards with their polled digital response systems on two-way cable.

The Hughes aircraft company has got a satellite application in for a domestic communication satellite which they'd like to put up as soon as possible in order to interconnect cable television systems because they see a profitable commercial market in satellite interconnection of cable systems. We've already seen technically demonstrated the so-called frame grabber technology where at an individual television set a still picture, uniquely addressed to an individual set, is sent down a cable television system. This is one of the examples that was seen in the Reston, Virginia cable system, with the Mitre Corporation conducting the demonstration in which the standard home television set becomes a computer terminal. Television is presented with thirty frames (thirty still pictures) per second giving us the appearance of motion. In the Mitre system, they were uniquely addressing half-frames of television to individual homes, and so you can have individually addressed still pictures – the jargon is time division multiplexing – on this party line cable system in order to do computer aided instruction and information retrieval in homes as an add-on to a television set. I'm not predicting that that kind of technology is going to be widely distributed in the next two or three years. I think it's probably going to take until about 1980 for the widespread distribution of the frame grabber technology, but it's on its way.

So I think the technical potential is likely to be there – or likely to be *almost* there as a result of normal actions in the marketplace that we can predict already. The hooker is that those hardware systems and all the software to go on them will be designed for the first profitable commercial application, and the kind of polling response time that's necessary for participatory entertainment, for pay TV functions and some of the teleshopping functions. It is probably not going to support computer aided instruction. A slightly more expensive system, with small incremental costs, could permit a more general purpose system to be ·alled throughout the nation, but if there's no commercial incentive to put additional capabilities on for education, they're not likely to do it. And retrofitting is going to be expensive – to try and change it over to a different kind of system will not be cheap. So the technology will be almost there for the kind of educational vision I've presented, but it's more likely to be available for X-rated movies than education.

That brings me to the challenge. I've sketched a vision of an information utility for 1985. I've picked 1985 because it's the year after 1984 and I'm afraid we'll get there first. Let's contrast this educational vision of the kind of society that we can create by 1985 with what I think was the business we've been talking about throughout the day today, namely, computer networks linking our own institutions. I think those computer networks linking our own institutions constitute a highly significant part of that larger vision. Aiming at the larger vision is more likely to bring about the computer interconnection that we're talking about. If we limit our goals to that narrower vision of connecting our institutions with computer links, I think we'll probably fail. And I think we'll fail for the political fact that large amounts of federal money are going to be needed to bring about even the narrower vision. And I think those amounts of federal money are unlikely to be forthcoming in the scale required for what many sectors of the society will perceive as an essentially elitist enterprise, that will connect our elitist institutions (I can say that, I'm from Stanford) in a way that will widen the information gap between the information rich and the information poor. I think the changing social goals are going to make it very difficult for us to tap the federal coffers for self-serving purposes or what will be perceived as institutional self-serving motives. So I think we in our own self-interest need that larger vision if we want to accomplish the computer network. I'm not suggesting we do that cynically; I'm suggesting that we really do go after the larger vision, because the resulting society will be a better place.

What I'm talking about is the possibility of a social goal with the glamour of a moonship. However, instead of building a NASA to put man on the moon and trying to justify it in terms of national prestige and the spinoff and the secondary benefits, I think that we're almost ready for a kind of social goal where we use the advantages of our technology to bring real services, needed services, to every individual in our society in a way that everyone can perceive and can understand. The organizational means of implementing such a national goal will have some similarities to a NASA mission: clear-cut or reasonably clear-cut goals; a large organization to help us get there; and a fair amount of organizational innovation. NASA involved the universities and industry and the whole society in this grand venture by controlling one hundred per cent of the funds to buy it. In this instance, we're talking about a leverage operation where perhaps five per cent and possibly as much as ten per cent of the total are federal funds. The private sector of the economy will build ninety to ninety-five per cent of the kind of system we're talking about, anyway. I think that leverage operation is economically and socially feasible. Some estimates say that the cable television physical plant is likely to approach fifty billion dollars a decade from now, and that's on the basis of projecting what private enterprise might invest. Relative to that, a public sector investment aimed at the incremental cost, the leverage costs to make that system also serve education for the society as a whole, is a potentially feasible goal. Over the next two to four years we ought to engage in planning studies, development projects and pilot demonstrations needed to refine and test the kind of vision that's possible.

By the time of our two hundredth anniversary in 1976, it should be possible for the President to announce as a two hundredth birthday present, that within the first decade of our third century, the benefits of electronic technology will bring education to every home in the nation or at least, to every urban home and rural community. And I think that's a vision that it's possible to implement. I'm not making a prediction that it will happen that way because if we just go home and go about our usual business, it won't happen. But if enough of us work on the organizing, the lobbying, the working through, the refining, and the correcting, it's possible that over the next decade an organization like EDUCOM can be a potent force in making possible that level of expenditure to move our society to a new kind of educational system. Without that, the cable story is the bad news side that I mentioned earlier, and I'm sorry about that.

Paying for Exchanged Resources

Chairman: Bertram Herzog, Director
MERIT Computer Network Project

Einar Stefferud
Consultant
Einar Stefferud and Associates

Franklin H. Westervelt
Director, Computer Center
Wayne State University

A presentation of the model and policy for payment of resources exchanged via the MERIT Computer Network opened this workshop. After Dr. Herzog responded to questions, a lively discussion between the audience and panel members followed. The policy statement and supporting materials are reproduced here.

NECESSARY STIMULI TO ENCOURAGE NETWORK SERVICE

A policy of service exchange and service charge settlement has been developed in conjunction with the directors of the participating computing centers. This policy, documented under the title, *Network Resource Exchange Mechanism*, was approved by the Computing Center Directors on September 15, 1972.

To promote actual network flow requires two stimuli. First, it is necessary to educate users who could make proper and beneficial use of the network's resources. Second, there should be no impedance to such justified use due to the inability to provide the resources or pay for those used.

The issues are aroused by the existence of the network but are essentially *internal* to any university.

THE SITUATION BEFORE THE EXISTENCE OF THE NETWORK

The typical budgeting and operating considerations for a university computing center can be simply stated in terms of three variables: the computing center's expenses, A, the expected income to the computing center derived from general funds, B_1, and the income derived from other sources, B_2. For the purpose of this discussion no delineation is made between the complexities of priority rate structures or different rates for different classes of users. We believe the essence of the problem can be caught by the simple cases discussed here.

For the present purposes a computing center's expenses are considered to be fixed once equipment, personnel, and basic material and supply commitments are made. Any services delivered in excess of the annual estimate can be delivered at only nominal incremental variable cost, principly for paper, cards, supplies, etc.

At the beginning of any fiscal period, typically one year, the computing center director makes his plans based upon his expenses A, see Figure 1. The amount of this budget item A is related to and based upon the reliability of the anticipated income.

$$B=B_1+B_2$$

A successful year is one where the actual income at least equals the actual expenses.

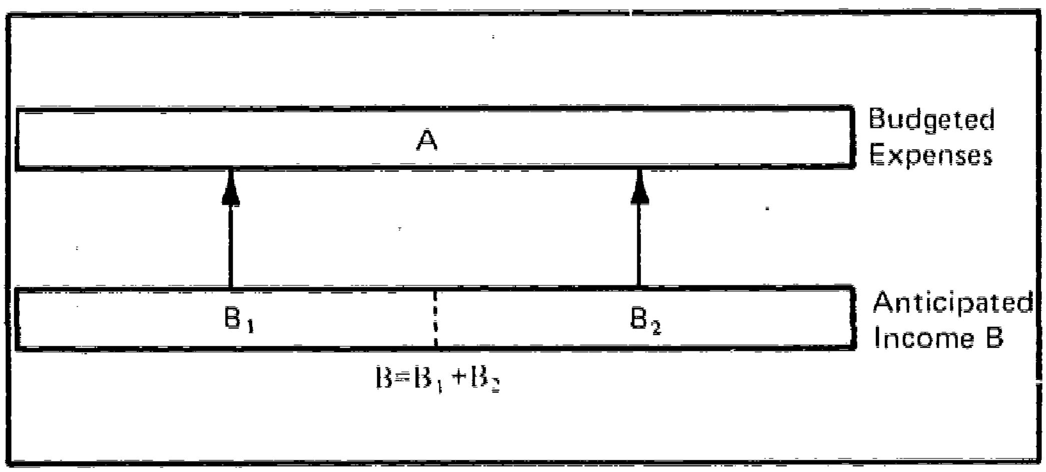


Figure 1

THE SITUATION IN THE PRESENCE OF THE NETWORK

Suppose a computer network, with its ability to exchange computing resources, is superimposed upon the situation described above. Suppose further, that the university buys computing services from other members of the network in the amount C*. Simultaneously the computing center

---

*Note that C may be composed of portions of B_1 and B_2; the users determine their source of funding.

delivers services to the network's other centers in the amount D, see Figure 2. Obviously if C and D are the same, the network does not disturb the balance between the expenses A and the anticipated income. Aside from this ideal situation what special cases arise?

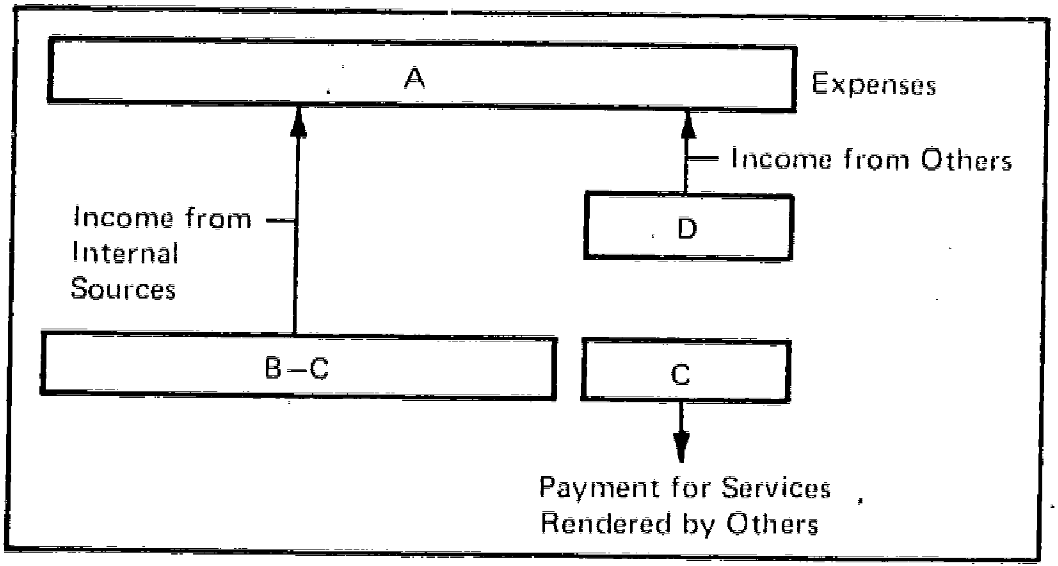


Figure 2

The Capacity Problem

Consider the issue from the computing center director's point of view. Assume that his center has reserved capacity in excess of the amount he expects to deliver to derive his income B. Assume then that he delivers service, via the network, to others equivalent to income amounting to D. Further suppose that of his originally expected income B, a portion C must be paid to others for service rendered via the network. Let D exceed C such that he receives a cash amount E in excess of his anticipated income, i.e.,

$$E = D - C$$

If, as a result, computer response, just one measure of instantaneous or long-term capacity, deteriorates then the center director is faced with a decision. One possible outcome is to limit D so that it exceeds C only by the available capacity. This could make the outside users feel like "second class citizens" who will question the quality and reliability of the offered service.

The Unbalanced Budget Problem

If the user community requires access to other centers in the network in the amount C, and this exceeds the amount D, the center delivers to

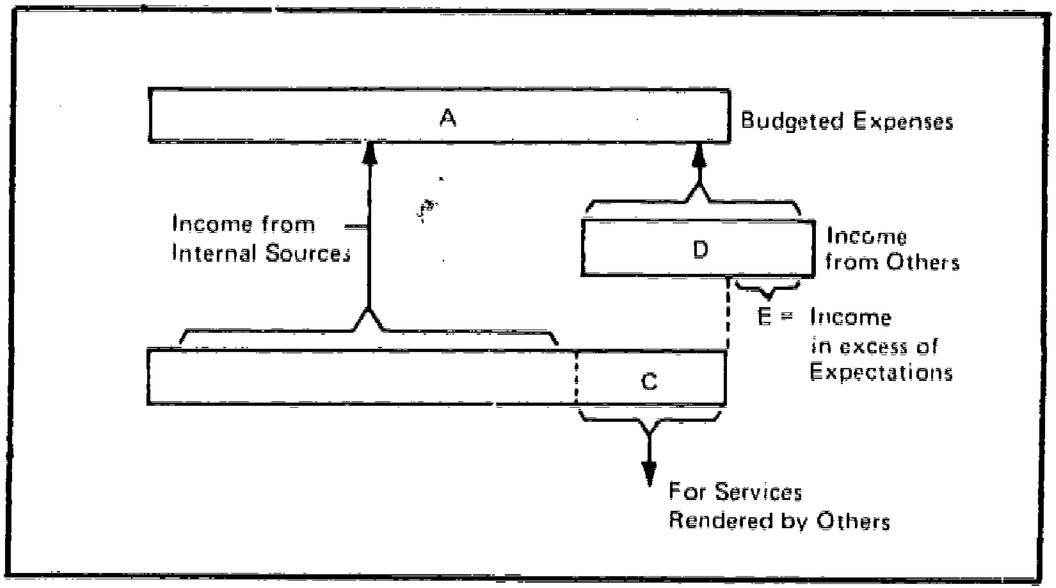


Figure 3

others, see Figure 4, then there will be a net outflow of cash in the amount,

$$F=C-D$$

Such a condition, in the absence of other stimuli, leads the computing center director to arrive at the only possible solution: the usage C must be restricted. Now the computing center director is unpopular with his university's users.

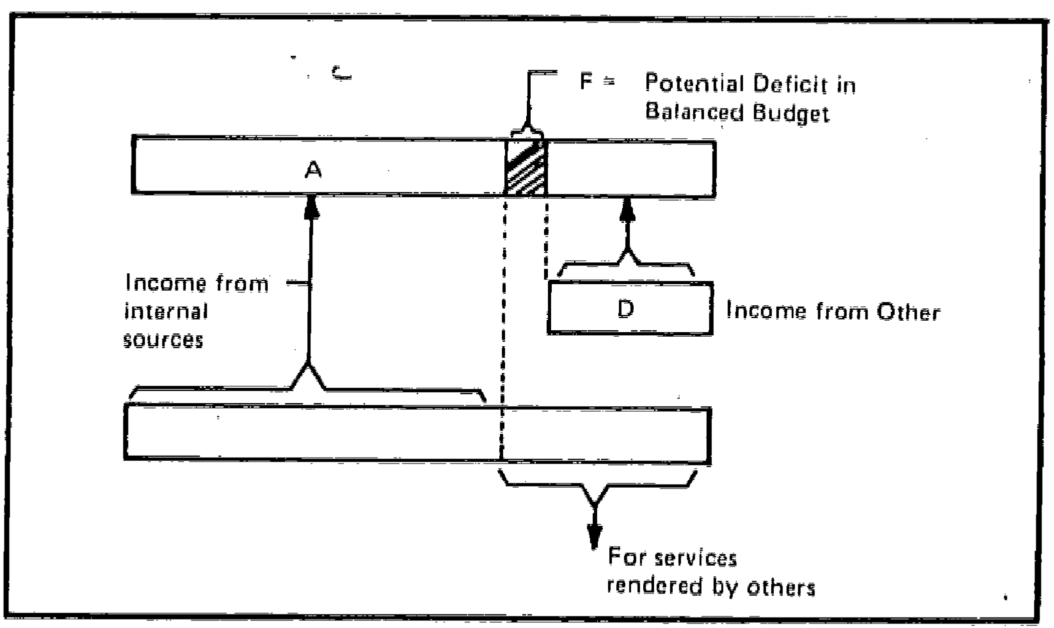


Figure 4

SOLUTIONS

For the *capacity problem* only honest marketing positions must be taken. The computing center must not be put in the position of having to promise delivery of service beyond its effective capacity. In this case, the mechanism of assigning the available resources, discussed elsewhere, permits the supplying computing center director to exercise appropriate control. From a long-term point of view he must be apprised of external demand estimates to be included in his future machine acquisition posture, etc.

For the *unbalanced budget problem* the computing center director must be given an explicit and separate guarantee that defines the amount of his *potential permissible* budgetary deficit, F.

It is important to note that the success of the network is directly proportional to the volume of exchanged resources. C and D each contribute to the volume of exchanged resources. Hence, it is important to provide conditions that permit C and D to be as large as possible while minimizing F.

Here again there is a long-range solution. Suppose that in its combined wisdom a university community decides that a special service offered by another center provides the best and cheapest way to obtain that service. Then C could always be expected to exceed D. That being the case, an appropriate adjustment must be made in the planning and budgeting process.

The program discussed herein is thus most aptly applied to unscheduled imbalances occurring after the budgeting process is completed.

RECOMMENDATION

Each university shall specify the amount F it is prepared to pay at the end of the fiscal year, if so requested. *This money should* **not** *be part of the current computing budget, namely A or B.*

The Merit Computer Network from the User's Point of View

Chairman: Eric Aupperle, Associate Director
MERIT Computer Network Project

The MERIT Computer Network currently connects three large general purpose computers (two IBM S/360, duplex Model 67's, and a CDC 6500), respectively located at the University of Michigan in Ann Arbor, Wayne State University in Detroit, and Michigan State University in East Lansing, Michigan.

These large machines are connected to one another over telephone lines and by means of three mini-computers, called Communications Computers. The Communications Computers (CC's), handle the interfacing problems, the data verification, and transmission tasks for the network. These smaller machines accommodate the system differences and therefore remove this load from the main processor.

Each of the three computer centers has an independent staff functioning autonomously from the others. The MERIT Network staff is composed of a different group of individuals drawn primarily from each of the schools. Although it is an independent unit, the MERIT staff does interact with the staffs of the three computer centers.

Referring to Figure 1 for a general depiction of the network, the circles and squares represent terminals and I/O devices at the individual centers, and the blocks marked "CC" connected by double lines represent the network hardware.

Diagram of the MERIT Computer Network —
A Symmetric Switching Computer Network

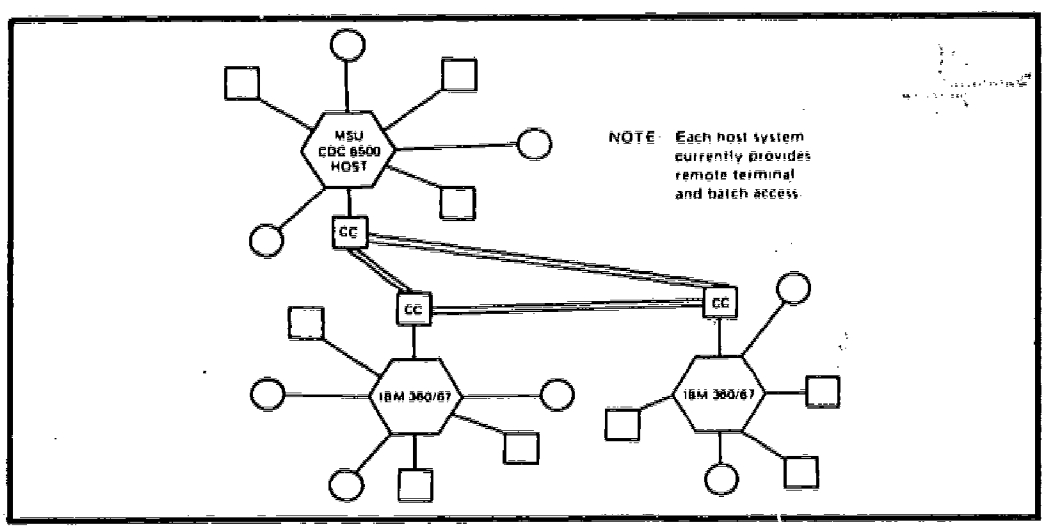


Figure 1.

In order to use the network a user first must be signed on to his local computer. Thus the MERIT network has no users per se, rather the potential users of the network are the total valid users of the computers of the three schools.

To obtain access to the network and use one or both of the other centers, the user requests an account number from his local computer center. The user's bill from his local center then includes services at all three nodes of the network. Insofar as possible, access to the network is kept administratively as simple as possible for the user.

The network was originally designed to provide three classes of service: (1) interactive; (2) batch; and (3) file access. At the time of the conference, interactive service was available and was demonstrated on each of the three systems over the network. Interactive service has been available between Wayne State University and Michigan since last April and between Michigan State University and the others just recently.

In the interactive mode the user signs on to his local computer and requests connection to the computer at another university. He may then utilize the resources at the remote computer as if he were an interactive user at the remote school. Ordinarily users initiate connections from a local terminal, but interactive network connections can also be initiated by a local batch job.

In batch mode, the user will be able to submit a job for batch processing at his local site specifying that the job is to be executed at a different network node. The output can be sent to still a third node. The steps now required to establish an interactive connection will be handled more or less automatically in batch. Work is underway on batch service, and it should be available by February 1973.

File access mode will allow a program executing on one of the computers in the network to specify that input is to come from a file located at another node of the network, or that output from the program is to be stored on a file at another node. In this mode, it will be necessary for the user to go through the several steps now required to open a connection.

Referring again to the interactive service that is currently available via the network, to access the network the user must first sign on or log on to his local computer as he normally would by dialing in via a Teletype or other terminal over local telephone lines. To access the network, the user can then attach it in much the same manner as he would attach another external input/output device, e.g., a tape drive and audio response unit.

Figure 2 is a copy of a listing of a terminal session in which the user initially signs on to the University of Michigan system. (On the terminal used, input typed by the user appears in lowercase, output from the computer system is in upper case.) In lines one through six of the listing, the user has dialed into his local host computer and signed on, giving his account number and password. At line seven, the user has asked that a program be run which mounts the network. (This is the same program the user would run to mount a tape, for example.) This program establishes a logical connection between the local host and the remote host; it then

```
1        MTS : ANN ARBOR (NC00-0422)
2   #signon kiys
3   #ENTER USER PASSWORD.
4   ?[redacted]
5   #**LAST SIGNON WAS: 14:01.51 10-09-72
6   #    USER "KIYS" SIGNED ON AT 14:02.50 ON 10-09-72
7   #run *mount par=mnet *net* dest=ws
8   #EXECUTION BEGINS
9   MNET *NET* DEST= WS
10       *NET*: MOUNTED ON MN00
11  #EXECUTION TERMINATED
12  #run mnet:newc par=*net*
13  #EXECUTION BEGINS
14       MTS : WAYNE STATE (MN10-0087)
15  #signon sbcm pw=
16  **LAST SIGNON WAS: 16:35:46 10-06-72
17       USER "SBCM" SIGNED ON AT 14:09:56 ON 10-09-72
18  #run * users
19  EXECUTION BEGINS
20  THERE ARE 27 TERMINAL USERS, 1 BATCH TASKS, 4
21  AVAILABLE LINES, AND 14 NON-MTS JOBS USING 710
22  VIRTUAL PAGES AND 171 REAL PAGES. HARDWARE IS CPU'S
    P2, STORAGE C D F G H, CCU'S 1
23  EXECUTION TERMINATED
24  #run * status
25  EXECUTION BEGINS

26  STATUS OF SBCM AT LAST SIGNOFF
                                      USED    MAXIMUM  REMAIN.
27  CUMULATIVE CHARGE      ($)        0.19    200.00   199.81
28  CURRENT DISK SPACE     (PAGES)    0         10       10
29  CUMULATIVE TERMINAL               0.04
         TIME              (HR)       0.04

30  EXECUTION TERMINATED
31  #signoff
32  OFF AT 14:11:41
33  ELAPSED TIME           104.916  SEC.
34  CPU TIME USED             .943  SEC.
35  STORAGE USED             9.886  PAGE-SEC.
36  DRUM READS               5
37  APPROX. COST OF THIS RUN   $.18
38  FILE STORAGE    0 PG-HR.   .00
39  *CONNECTION CLOSED
40  *NET*: DISMOUNTED
41  #EXECUTION TERMINATED
    #
```

Figure 2

returns control to the user for further input. To the remote host the process looks virtually identical to that of a user dialing directly in over a phone line.

To begin interaction with the remote computer, the user at line 12, requests that a program be run that reads output sent by the remote computer and displays it on the user's terminal. Further, this program will read input from the user's terminal and send it over the network connection to the remote host. While the user could utilize the network connection as if it were an ordinary I/O device (this is done for certain applications), the network copy program used here conveniently allows the user many of the same control functions, e.g., the setting of several options regarding record lengths and mode of data transmission.

At line 14 the sign-on request sent by the remote host when the connection was established is printed at the user's terminal. (It was stored in a buffer when initially received by the local communications computer waiting to be read from the network connection.) The user at this point may interact with the remote host as if he had dialed into it directly.

First the user signs on (15-17) to the remote host; he can then run programs, create files, destroy files, in short he can do anything he could do had he signed on directly. After signing on, the user runs a prog-am at Wayne State which indicates the activity and configuration at the remote host at that time (18); output from the program is printed at the user's terminal. A second program is run indicating the status of the user's account at the remote location (24-30). Finally, the user signs off the remote system. After the accounting information sent by the remote host is printed, the network connection is automatically closed and dismounted.

While using the network, the user is charged for the resources he uses on all systems to which he is connected, at specified rates for such services. There is currently no charge for the network itself; methods for this charge will be worked out later as the network progresses from its experimental status to an operational one.

Several other examples of use of the network follow:

Figure 3 illustrates the process of copying a file residing at Wayne State University to a file at the University of Michigan. (The network connection is established and the program to transmit records to and from the network has started.) Note the lines coming from and going to the remote host are indented one column from the margin relative to those being processed locally. A line beginning with the character period, ".", are commands to the network copy program.

The command ".mts" (5) returns the user to a local mode of operation, allowing him to interact with his local host. The command "restart" (10) returns control to the network copy program for further interaction with the remote host. Subsequent commands effect and verify the file transfer to the University of Michigan (11-33). In gereral an effort has been made to keep interaction over the network and commands to the network logically, and even formally, consistent with the user's local command language and conventions.

```
     MTS : WAYNE STATE (MN10-0087)
#sig sbcm pw=
**LAST SIGNON WAS: 14:19:54   10-09-72
  USER "SBCM" SIGNED ON AT 14:30:24 ON 10-09-72
#.mts
#get -temp
#READY.
#list -temp
#END OF FILE
#restart
#.sink -temp
#copy rw.s
#.sink *sink*
#.mts
#list -temp
>   1         DIMENSION ALPHA(100)
>   2         INTEGER*2 LEN
>   3    10   CALL READ(ALPHA,LEN,0,LNUM,5,&20)
>   4         CALL   WRITE(ALPHA,LEN,0,LNUM,6,&20)
>   5         GO TO 10
>   6    20   STOP
>   7         END
#END OF FILE
#restart
#list rw.s
    1         DIMENSION ALPHA(100)
    2         INTEGER*2 LEN
    3    10   CALL READ(ALPHA,LEN,0,LNUM,5,&20)
    4         CALL   WRITE(ALPHA,LEN,0,LNUM,6,&20)
    5         GO TO 10
    6    20   STOP
    7         END
  END OF FILE
#sig s
OFF AT 14:35:02
E   278.07
C       .594
S     4.583
D    16
    $.29
*CONNECTION CLOSED
***: DISMOUNTED
#EXECUTION TERMINATED
#
```

Figure 3

Figure 4 illustrates establishment of a connection to the CDC 6500 at Michigan State University from the University of Michigan.

Figure 5 depicts the process of transferring a file from the IBM S/360 at the University of Michigan to the CDC 6500 at Michigan State University. (The log-on and log-off procedures are omitted.)

Finally Figure 6 illustrates the compilation of a FORTRAN source program stored in a file at the University of Michigan by the FORTRAN compiler at Wayne State University.

A question and answer period followed the presentation. Several persons were concerned about the rate at which data could be transmitted via the network since they seemed to feel the data rate would greatly affect the manner in which the network could be used. Currently the network transmits data at the rate of 2000 baud with a design capability of 50,000 baud. The current rate was determined solely on economic grounds. For interactive use not involving the transfer of large data files this rate seems to be quite satisfactory.

The advantage to the user of going through the network, aside from savings on telephone line charges, is that users can jointly utilize resources at more than one site simultaneously. This is not possible otherwise.

With the data rate of the network at 2000 baud, doesn't the charge for core residency, while waiting for data to be transferred over the network, become prohibitive? As with most time-sharing systems, the user is not charged except for connect time while his programs are actually executing instructions. In fact, while the data transfer is taking place the user's program would be paged out onto secondary storage and would not be using real core storage. The only additional cost to a user running a program utilizing data at another node would be the connect time required for the transmission, plus a relatively small amount for the housekeeping required by the network interface routines.

After the question and answer period, a film of the network in operation was shown. It demonstrated a program at the University of Michigan which provided data to a program running at Wayne State. At Wayne State the data was processed and the results returned to the first program, which used it to construct a graphical display on a device connected to the UM machine. Professor Bertram Herzog, the narrator in the film, was able to interact dynamically with the two programs through the graphics terminal.

During the entire conference the MERIT network was available for use by conference participants who wished to get hands-on experience.

```
#run *mount par=mnet *net* dest=ms
#EXECUTION BEGINS
 MNET *NET* DEST=MS

    *NET*: MOUNTED ON MN00
#EXECUTION TERMINATED
#run mnet:newc par=*net*
#EXECUTION BEGINS


 0/09/72 MSU HUSTLER 2 L239 LSD 29.14 10/08/72
 TYPE PASSWORD, PN, AND USER ID.
 XXXXXXXXXX
@MMMMMMMMMM
@SSSSSSSSSS
@TTTTTTTTTT,
@
 [redacted],616467,aupperle
 SS83023, LINE 30
 LAST ACCESS: S 10/09/72 13:20
 RUNS: 13 BALANCE: $4989.99
 PRODUCTION INTERRUPTED AT 9 P.M. FOR SYSTEM
 DEVELOPMENT UNTIL 11 P.M.
 SYSTEM: FORTRAN LINE RANGE 0 - 0
 LENGTH = 72 MARGIN= LINES=
 READY 13.25.45


   files.

 — 1 SYSTEM FILES--
 —PRIVATE FILES--
 C*ZZZOUT EWFILE
 READY 13.27.04

 logout,t
 JOB COST: $ .23

 *CONNECTION CLOSED
 *NET*: DISMOUNTED
#EXECUTION TERMINATED
#
```

Figure 4

```
 ok.
OK·
 connect,input.
OK·
 copy,input,temp.
 .source demo(1,9)
 *eof
 *eof
OK·
 .cci=on
 connect,output.
 OK·
 rewind,temp. copy,output.
 *eof
 *eof
 OK·
 copy,temp,output.
*                  A-B=C EXAMPLES
EXAMPLE            CSECT X'20'
START              MOV A,C
                   SUB B,C,
                   HALT
A                  DC F"10"
B                  DC F"20"
C                  DC F"0"
                   END

 OK·
 .sink -t
 rewind,temp. copy,temp,output.
 .mts
#list        -t
>             1      *             A-B=C EXAMPLES
>             2      EXAMPLE       CSECT X'200'
>             3      START         MOV    A,C
>             4                    SUB    B,C
>             5                    HALT
>             6      A             DC     F"10"
>             7      B             DC     F"20"
>             8      C             DC     F"0"
>             9                    END
>            10
>            11      OK·
 END OF FILE
```

Figure 5

```
≠run mnet:newc par=*net*
≠EXECUTION BEGINS
   MTS : WAYNE STATE (MN10-0077)
  MTS WILL BE GOING DOWN AT 12 MIDNIGHT FOR
  RECONFIGURATION
≠sig sbcm pw=
**LAST SIGNON WAS: 17:18:33 10-10-72
  USER "SBCM" SIGNED ON AT 17:41:46 ON 10-10-72
≠run *ftn par=source-*source*, print=*sink*
EXECUTION BEGINS
 .source rw.s


 MICHIGAN TERMINAL SYSTEM FORTRAN IV G COM-
 PILER MAIN
    0001              DIMENSION ALPHA(100)
    0002              INTEGER*2 LEN
    0003    10        CALL READ(ALPHA,LEN,0,LNUM,5,&20)
    0004              CALL WRITE(ALPHA,LEN,0,LNUM,6,&20)
    0005              GO TO 10
    0006    20        STOP
    0007              END
   TOTAL MEMORY REQUIREMENTS 000344 BYTES
 NO ERRORS IN MAIN

 NO STATEMENTS FLAGGED IN THE ABOVE COM-
 PILATIONS.
≠sig s
OFF AT 17:44:51
E     185.676
C       1.854
S      56.053
D      38
     $.36
*CONNECTION CLOSED
*NET*: DISMOUNTED
```

Figure 6

Assistance for Faculty Users: Development of Computer-Related Instructional Materials

Chairman: Karl Zinn
Center for Research on Learning and Teaching,
and MERIT Network
University of Michigan

Ronald Code
Director
Northern California Regional
Computing Network

Ralph Deal
Department of Chemistry
Kalamazoo College

Joseph Denk
Curriculum Coordinator
North Carolina Educational
Computing Service

E. C. Hertzler
Department of Biology
University of Michigan

Dana Main
Department of Psychology
University of Michigan

Recorder: Mary Jill Ault
Center for Research on
Learning and Teaching
University of Michigan

PROJECT EXTEND

Karl Zinn noted that Project EXTEND ·s established within the MERIT Computer Network environment to bring demonstrations and trial experience with computing resources for instruction to small college faculty. It draws heavily on the computing resources, software, and documentation provided by the computing centers of the participating MERIT universities – the University of Michigan, Wayne State University, and Michigan State University.

The most important contributors to the success of instructional computing are the instructors in the various disciplines. Their ideas and judgments are crucial to the acceptance of new technology for learning. Not only is their judgment on the value of various applications crucial, but in addition their advice on documentation, user guides, and other support materials is quite important. The contribution of the disciplines should be handled through departments, professional associations, and authoring teams. One cannot depend entirely on computing centers, administrators, regional consortia, or publications on instructional computing, but must include the leaders in the teaching of each discipline apart from computers. The decisions will be made and the incentives will be provided by those people who are leaders and set standards of quality for what is important in the discipline and its teaching.

The Project makes a major effort to facilitate the transfer of ideas and programs between institutions and considers small colleges to be among the most important sources. It attempts to translate the services of a large university to small institutions in the area. Small college computing centers often must look for cooperative arrangements to assemble resources and expertise.

Some of the services and support functions which Project EXTEND provides are listed below.

1. *Information.* The files and technical memos of the Project provide information and advice about the capabilities of computers for use in instruction.

2. *Demonstrations.* A number of rather general demonstrations have been prepared, each one indicating a type of contribution to learning through the use of computing.

3. *Consultation and training.* Staff advise individual faculty and provide training through workshops and written materials, emphasizing means for carrying on effective instruction. The development of computer-related instructional materials is a primary subject of workshops and consultation; advice is offered on development procedures, personnel requirements, appropriate equipment, evaluation of outcomes, funding of further development activities, etc.

4. *Development and modification.* The Project offers programming assistance to adapt the demonstration programs to the specific and individual needs of participating faculty. The faculty member is encouraged to make suggestions, observe their implementation on the computer, and test them out with students. In addition the participating faculty are encouraged to implement their own ideas. Credit is given to program authors through the library of programs used by the Project and in the ON-LINE Newsletter, distributed to all colleges and universities in the state.

5. *Reproduction and distribution.* The staff assists with the editing and production of manuals. Credit is given to program authors through the publication of documentation packages which facilitate classroom use.

6. *Evaluation and reporting.* Careful attention to the evaluation of the activities is encouraged, including objective measures of performance or reports of student attitude wherever possible. Perhaps more important is the professional review by peers in the same discipline. The Project also collects data from the use of programs at other colleges toward the purpose of modifying these programs to make them more serviceable. A model for the documentation of programs and for the dissemination of those programs is being constructed.

Three of the panelists have been involved to varying degrees with the Project and represent different points of view about the use of computing in education. Dana Main has developed a flexible program for instruction in psychology for which documentation was published by the Project. Ralph Deal is with a college which is currently working to develop capabilities in instructional computing. E.C. Hertzler represents an organization in existence long before the Project which can be said to be a model of some relationships between a university and the surrounding community colleges. The two remaining panelists, Joseph Denk and Ronald Code, are involved with instructional computing through networks in other states and will comment on the problems of dissemination.

EXPER SIM: AN INSTRUCTIONAL COMPUTING PROGRAM

Dana Main described the development of EXPER SIM. This week the one-thousandth student used the EXPER SIM program to place an experimental design onto the computing system and to obtain simulated data by means of data-generating models. He was a college sophomore, thinking he might major in psychology, and had no knowledge of statistics, programming, or for that matter computers. He was able not only to design an experiment, as he was within the structure of the old elementary psychology laboratory course (traditionally a rat lab), but also to design a research program on subject matter as diverse as the etiology of schizophrenia, imprinting, and motivation in routine tasks in an industrial setting. He coped with cost-benefit programs, budget writing, as well as the design of experiments, hypothesis testing, and theory testing.

The student is able to do this because we have developed what we call the experimental simulation supervisor, written in FORTRAN for use on the Michigan Terminal System by Robert Stout, a graduate student in mathematical psychology. At this point the supervisor can generate data from one of three different files = one for each model. We are in the process of putting three more files onto the supervisor which are more complex in nature and involve different subject matter.

We do not have a laboratory computer or even a terminal. We use the public terminals and keypunches on the Michigan campus. A student is taught how to use the system in either batch or interactive mode, using a keypunch or a typewriter terminal as he wishes. The classroom itself is the simulation of a scientific community where each student plays the role of a social scientist and goes through all of the various roles which the time constraints of data collection eliminated in the traditional lab.

The models used in the system reflect the pedagogical as well as the research interests of our graduate students. They are based on a body of literature which in the behavioral sciences is often very contradictory. The model builder, the computer programmer, and the instructor are distinctly different roles, although occasionally one person will play two or three of these roles. We wanted to keep the roles distinct so that the model builder was involved purely with subject matter and did not have to deal with programming constraints. Our goal is to develop a library of models such that a potpourri of subject matter is available for investigation, and to develop with appropriate funding a language which would allow model builders to place their models onto the system without extensive programming ability.

We have operated without formal funding in the sense that we have not had a research grant. My salary was paid by the department as coordinator of this course, and graduate student teaching fellows have contributed tremendously to our resources. Their only reward is that credit as authors is given to them on every printout. We have also gotten some help in the development of these materials from Project EXTEND, from the Office of Research Administration, and from the Center for Research on Learning and Teaching, which supported some graduate students in the summer who developed the programming.

As funds for maintaining a rat lab were cut back, I found money for educational computer use in a different budget. I was able to persuade my chairman of the value of this kind of instructional method and he has been very generous in our portion of the computing funds allotted to the psychology department.

My plan of course is to extend the materials developed here to other courses within the department. To some extent this has been done, for example, in the statistics courses. Other instructors only very recently are seeing the potential of developing certain models, concentrated for more advanced courses.

The entire project has been developed at a cost of about $5000-6000. Our budget for student computer use is around $5000 a year, or $10 a student for approximately 12 experiments each.

ONE COLLEGE'S PROSPECTS FOR INSTRUCTIONAL COMPUTING

Ralph Deal spoke from experience at Kalamazoo College, a small private liberal arts college. In the attempt to make effective use of computers we have appointed Donald Stanat to a new faculty position Associate Professor of Computer Science. He initiates a new legitimate discipline in our liberal arts curriculum. This position is not the administrative role of computer center director, but a genuine faculty role. Don will teach a series of three courses during the year to provide background for students working on senior independent projects with a strong computer emphasis, and to provide enough background for them to consider graduate work in computer science. His remaining responsibilities lie in faculty development – talking with faculty members who think the computer might be useful to them and helping them to explore these possibilities. This is an area in which Project EXTEND will be very useful to us.

Kalamazoo College is really just coming into Project EXTEND. We hope to see several things happen at the college through our participation. A terminal will be connected directly through a leased telephone line to the Michigan State CDC 6500 and through MERIT to the Michigan Terminal System at the University of Michigan and Wayne State University. Thus we will be able to use at Kalamazoo College programs (such as EXPER SIM) which have been tested and well documented by Project EXTEND in a variety of disciplines. Because some CPU and connect-time funds will be provided, within a limited range we will be able to explore those programs.

We see our new program in computer science as providing a significant impact on education at Kalamazoo College and perhaps serve as a model for the study and use of computers in a small liberal arts college. Through Project EXTEND we expect to gain assistance in constructing strategies for faculty development and for the education of the administration regarding the possibilities for effective use of computing facilities at Kalamazoo. For when Project EXTEND terminates, we are committed to provide meaningful computing facilities for our students and faculty on a continuing basis.

Another advantage in our participation is that we will gain experience with an ongoing academic network. It is very unlikely that Kalamazoo College will receive federal funding for the purchase of computing hardware. If federal funding is forthcoming to support computing facilities, it will probably be through network activities. We would like to see a computing network between colleges and universities designed to enhance undergraduate learning. The development of new modes of learning at small colleges frequently suffers from faculty isolation; sharing program development over such a network should remove this hindrance. Experience from Project EXTEND will be invaluable in the development of an educational network.

A MODEL OF SOME RELATIONSHIPS BETWEEN A UNIVERSITY AND THE SURROUNDING COMMUNITY COLLEGES

At least in the state of Michigan, universities are surrounded by community colleges. In any particular area the number of students in the community colleges exceeds the number in the university. This difference is likely to increase because the community colleges are growing much faster than the universities. If computing facilities are to be networked in a given area, certainly the community colleges are going to become involved at some point.

The Association of Community College Biologists (ACCB) is one model of some relationships between a university and the surrounding community colleges. E.C. Hertzler described the model. The ACCB has grown up around the Dearborn campus of the University of Michigan and involves ten nearby community colleges with a total of about 100,000 students. The group's success is based on the fact that it was organized by users, not by administrators. I suspect that if networks develop successfully in metropolitan areas this will have to be the mode of organization. The usual university attitude toward community colleges tends to kill all kinds of cooperation. We began quite open-mindedly and were rather surprised to find the high quality of instruction and facilities in the community college.

The instructors at community colleges have often come from high school teaching experiences and feel that they have graduated from the professional organizations for high school science instructors. Yet they cannot join the professional university organizations that have a heavy commitment to research. Hence they are cut off from one another. The ACCB was organized to meet the real needs of the biology instructors.

We have found that community college instructors recognize the need for feedback from specialists who are frequently not available at their schools. In the university, specialists in many areas of a discipline are available as consultants for elementary courses. ACCB brings together the community college instructors and specialists from nearby universities.

Community college faculty want special graduate courses even though their institutions do not reward them for earning doctorates or for completing more than 30 hours of post-master's work. ACCB surveys individual needs, designs a special course, and finds a professor in a neighboring university who is willing to lead a course covering specific topics and taught in a specific way. For the instructor this is very different from choosing a course by title from a catalogue in the hope that it will meet his needs.

If computer use is extended to community colleges for the university, it will probably be modelled on this kind of tailor-made arrangement.

PROMOTING FACULTY USE WITHIN A NETWORK

Ronald Code noted that in many schools one department member seemingly has been assigned the role of computer user, and therefore he always offers the computer course. In some places a point of stabilization is reached where the computer is used by one physics teacher, one social scientist, one business teacher, and perhaps three instructors in data processing and programming. For several years it is difficult to progress beyond that point. At the Northern California Regional Computing Network (previously the Bay Area Network) we have tried to overcome this by funding a group of people for a project and "bugging the daylights out of them to make them produce." In other words, it is not voluntary; once someone is involved we won't let him go.

A more general problem is the resistance to the adoption of new methods. I normally find that the authors of most packages are about 1000 percent more enthusiastic about it than anyone else, at least at first. Gradually people react to these systems if they think that something is going to catch on. However if they feel that this is a specialized area, they can ignore it safely and perhaps can be better off by doing so. The whole department will not automatically adopt a new method of teaching once it is introduced. It is often a long tedious process.

The user at a remote school in the network sometimes feels that he will have difficulty in having the program adapted to his needs. However if the remote user has this problem, the user at the computer site has the same problem. Almost all of the widely-used programs have some data area that can be changed. If this were not so the program would not be flexible, and all users would suffer with it.

In our experience users very seldom seek to adapt the teaching strategy of a program; they either accept or reject it as a whole. With a question and answer program or drill the author finds the program of great interest and very relevant and no one else likes it at all. In the programs on our network the learning strategy has not been well developed and is quite simple.

Although our program is located at the Stanford Computation Center computer, we have no teachers from Stanford participating. The nine participating community colleges are not interested in supporting large research projects because they don't do any research. They are most interested in the number of students that can use the computer, given that each student does virtually nothing on the computer. Some of the colleges are finding that they are paying for things that they don't need when they buy into networks hosted by a very large university. Some of the most successful community college systems utilize small computers with time-shared terminals and 10-20 line BASIC programs. Their computing needs are in a different category than are those of a research center.

TRANSPORTING TEACHING PACKAGES

Joseph Denk outlined some problems in transporting computer-based teaching programs. As with texts where perhaps six books dominate 95 percent of the market in a particular discipline, of the many available computer programs for teaching only a few are in wide use. These few allow many pedagogical approaches. Most programs are bound by their pedagogy as described in the documentation. Frequently an instructor will be able to find only two or three programs out of a library that he can fit into his undergraduate course. There are as many approaches to education as there are faculty members. A few approaches are outstanding. When an outstanding program is documented and used by a large group of people within a user-to-user network, it is improved in the process. A highly structured CAI package does not often move away from its place of origin. Other more skeletal programs like the EXPER SIM package have been very viable, because they are tailored not to fit or replace a curricular package but to supplement it as homework activity. We at North Carolina have found that the flexible open-ended homework-oriented structure is easily fit into a particular instructor's pedagogy and research interests. As it is transportable between teachers in one discipline, it can also move between disciplines.

The technical transport of a computer program is relatively simple. The big problem is the pedagogy. A necessity is the presence of dedicated faculty members who have enough time to move the pedagogy. At North Carolina we are highly dependent on the total dedication of the staff who take care of user's technical problems, helping the faculty members implement programming support for his pedagogical purpose. The faculty user must be involved with the design of the program in order that its educational goals remain distinct. Yet there are too many technical barriers toward mounting adequate data bases and clean, reliable, and maintainable statistical packages for the user to assume this responsibility himself. Among the most valuable employees to a computer center is the one who knows the system, the user needs, and the disciplines. Although such a person is very rare, in some way all these functions must be implemented in the computing center if support is to be guaranteed to the user.

The user-to-user network is the only functional way that computing can be introduced into the curriculum and be disseminated properly. When computer networks, such as ARPA and NSF were set up, this aspect was overlooked in favor of hardware interfaces. It is hard enough to interest other instructors within the same department, let alone those in 40 widely spread schools in a network.

FURTHER DISCUSSION

Several problems arise in the dissemination of a computer program. EXPER SIM has operated on a first-come, first-serve basis when responding to requests for information. Because EXPER SIM was written in Fortran it is relatively simple to transport it to other Fortran systems. However at the University of Louisville two of the three programs have been rewritten in BASIC, and at another school they have been run successfully on a PDP-8 through the use of overlays to accommodate the smaller computer memory.

There is occasionally author resistance to giving away programs without a charge, after all the time spent on their development. However in California one author decided, after two years of trying to expand the use of his program, that it was hard enough to arouse interest even when the program was available without cost. Most programs for instructional use are now available at reproduction cost. For example when EXPER SIM documentation was reproduced by Project EXTEND, it was distributed free of charge with credit given to the author (who also holds the copyright).

But it is not enough to distribute the program. The faculty needs to be trained to use it (in Northern California by two-month-long summer workshops). Also after training the faculty needs support -- technical support, guidance on where to consider the use of the computer in their disciplines and released time from their institution to make the program operational.

To differentiate those schools using the computer successfully, the deciding factor does not seem to be the size of the school's computer. Instead there seem to be other factors. First if the people involved are influential, instructional computing is likely to catch on in the school. Secondly if the school feels that innovative and/or potentially cost-reducing methods of instruction are very important, then there is likely to be much activity in instructional computing.

The most transportable programs are those that are not tied to a particular philosophy of education. The teaching fellows using EXPER SIM run very different classes, some highly structured, some quite loose. But all instructors use hypothetical costs for the experiments each student runs, primarily because awareness of costs contributes to good experimental designs. Since only a skeletal program is provided, the course can be structured in many different ways.

There was considerable controversy about the responsibility of the instructor to give the undergraduate student experience with tools for the collection of real data. Some felt that with simulated data the student was missing a very important part of scientific technique.

Dana Main thought that this was a sequencing issue. EXPER SIM is used in the first laboratory course a student in psychology has. In a traditional lab the student must learn techniques before he is allowed to think. With EXPER SIM the sequence is reversed. He begins to formulate hypotheses about the subject area in his first course. Laboratory resources are used for the advanced student who has already been immersed in the subject matter. She did not at all advocate the elimination of the data collecting experience from the student's undergraduate career, but felt that the limited resources were better placed with the advanced student who has done some thinking about the field.

Assistance for Faculty Users: Exchange of Instruction-Related Computer Programs

Chairman: Karl Zinn
Center for Research on Learning and Teaching,
and MERIT Network
University of Michigan

Joseph Denk
Curriculum Coordinator
North Carolina Educational
Computing Service

Ray Geitka
Academic Computing Services
Oakland University

Edmund Goings
Instructional Computing Center
Eastern Michigan University

Jack Meagher
Computer Center
Western Michigan University

Recorder: Mary Jill Ault
Center for Research on Learning and Teaching
University of Michigan

NCECS

Joseph Denk opened his presentation by remarking that the NCECS network is known primarily through a "fishwrapper" called PALS, which is a current awareness bulletin in North Carolina, and which outside the state has been misread as a catalogue. PALS is the implementation of a curriculum model which will be described in some detail.

The regional network was born out of private funds and was a far-sighted project by the state of North Carolina. A totally unique computing situation was developed by three closely situated universities – the University of North Carolina at Chapel Hill, North Carolina State University, and Duke University. They formed a non-profit organization, the Triangle Universities Computation Center (TUCC), to purchase a large IBM computer. The Center is run by a board of directors composed of representatives from the three universities. In 1966 private money was provided to support computing in North Carolina's smaller colleges remote to TUCC. A terminal and a year's computing time was given to each college who wished to participate. Currently forty remote colleges with the three universities in TUCC form a regional network.

The Computation Center is run by an excellent staff including six systems programmers whose only responsibility is to keep the machine running. The primary function of the twelve employees of the North Carolina Educational Computing Service (NCECS) is to meet the instructional computing needs of the remote users and in particular to provide applications programming. This group has attempted to enlarge the scope of educational computing at the remote campuses. The educational service staff at NCECS supplies user service via WATS lines across the state, runs user workshops in job control languages and technical services, and keeps the network viable and functioning.

In 1969 I was teaching chemistry at one of the remote colleges and was the coordinator of the chemistry group that was trying to share programs and ideas among its members. In addition I was the manager of the college's terminal to the computing complex. Instructors in business, sociology, political science, and other disciplines came to me for useful instructional packages. I couldn't get my hands on enough materials to meet their needs. After much difficulty I discovered that the Computation Center was supporting only a couple of big statistical packages. I could not identify a clearinghouse for computing packages operating in any one discipline, with the exception of the quantum chemistry program exchange in Indiana which was geared toward the researcher, not the educator.

So I collected everything that I could find and set up my own informal clearinghouse. In the first two years I had collected 3,000 packages from conferences, journals, and any other place I heard about them. I organized the programs into three categories. Category C was for programs that I had heard about but hadn't been able to locate, either because I hadn't found their authors or because they didn't exist. Category B was for programs with low support. Category A was for those programs that were fully operational on the computing system and maintained by the staff.

When I had collected enough programs, I began to run workshops throughout the state. We (meaning myself and some 150 faculty members from North Carolina and other states) have run 41 two-day workshops in the last three years and have trained 1400 faculty members. During these workshops we expose materials in a semi-operable or limping state (Category B) to the faculty. If the faculty indicates that a program seems to be viable, we put it on the system with full support. Thus in this model for obtaining curriculum materials the faculty sifts out programs of interest in workshops, indicating which have an initial appearance of validity.

Several packages, which I thought were very good (for example IMPRESS from Dartmouth for social science data processing and an information retrieval system for infra-red spectra), were slipped right on through the workshop stage and put on the system directly. However the majority of the programs on the system have been chosen by the faculty. As of the first of January, of the 400 programs on the computer (with 340 in Category A) 312 are in classroom use. PALS lists 1000 and another 2000 have been collected.

Some disciplines are more reliant than others on canned or "black box" programs. For example, mathematics, biology, chemistry and physics have used very few canned programs. We have used the model to sift out those programs which are important to save for education and to support fully on the computer. As can be seen in PALS the largest number of programs are in chemistry, but the greatest acceleration in accumulation has occurred in the social sciences.

Arising from our public advocacy of the necessity of the Computation Center's support of program exchange, another organization, CONDUIT, was born. CONDUIT is composed of five regional centers who have come together to test the transportability of curriculum materials. The five centers are Dartmouth College, the University of Iowa, Oregon State University, the University of Texas (Austin), and NCECS in North Carolina. Computing center personnel have been involved from the beginning in order to achieve the most effective mix of programmers, computer scientists, and most importantly students and faculty. On January 1 we began to test the mobility of materials in six disciplines. A faculty committee in each discipline selected the materials to be moved.

CONDUIT has moved materials in all six disciplines among the five schools. In chemistry 40 educational packages were moved. In the financial sciences (really two disciplines) we moved 20 packages in operations management for business and seven simulations in micro and macro economics. In mathematics we moved the Iowa materials in linear algebra. In physics we moved the CO-EXIST materials for introductory physics from Dartmouth. (However in this package the student does all the programming, so I don't know if we have tested anything here.) In the social sciences we accomplished the most extensive work. Nine data bases were mobilized to fit the survey analysis systems of all five regional networks. We built translating systems to allow the future transport of data bases between the four basic statistical systems in operation – IMPRESS at Dartmouth; SPSS at Iowa, Texas and North Carolina; SIPS at Oregon State; and POISSON at North Carolina (a version of IMPRESS). This major piece of work was accomplished by disciplined, hard-working social scientists and could not have been as functional if it had been done by a group of computer scientists.

This summer CONDUIT conducted a series of workshops in four of the six disciplines to train the faculty in the new programs. The workshop on the use of the social science programs will be held in late October, and the workshop for biologists in December.

Because of the success of CONDUIT, NCECS has been able to add over 100 packages to the next issue of PALS.

In the beginning there was very little faculty interest in educational computing, but now we have 100 responses within ten days of the announcement of a workshop. There are 1000 active users in the state. (An active user is defined as someone who has used a program in his class during at least one semester.)

Supported by the National Science Foundation, we have spent about $110,000 in the past two years – half on workshops educating the faculty about possible uses of computing in education and half putting the selected programs on the system. This second body of work was done by fifty faculty members and the NCECS staff.

The NCECS model for assembling a large library of curriculum materials rests on the concept of transportability, the feasibility of moving programs onto the TUCC system in North Carolina. Because of financial limitations at many computing centers this model cannot be widely adopted. However its major value is that hopefully you will see the sifting process as a method realistically able to bring forth quality curriculum materials that we will be able to move as packages to your schools.

DISCUSSION

DENK: The critical point is that there must be someone who is an intermediary between the disciplines and the other computing center staff, someone who works on application programming and forms a bridge between the disciplines and the computing center. Any one computer center employee can support only a few packages. What is needed is a user-to-user linkage, making the discipline consulting group the actual users themselves.

MEAGHER: At Western Michigan University we have three application staff members working with particular departments, through the device of joint appointments. For example one man is appointed jointly at the Computing Center and the sociology department. A large percentage of the working educational packages on our system are statistical, because most research work on campus is being done by the sociologists and because the most vocal demand for computing services has come from them. However the sociology department actually wants a programmer to write packages for them and they will probably hire one soon. Another joint appointment is with the anthropology department, but this department does not seem to be very interested in instructional computing. The third joint appointment is with the mathematics department. This one has been our biggest success. The man filling this position is not interested in further research, but in consulting on applications. He has been an extremely useful statistician to the faculty on campus.

I agree very strongly with you on the approach of delaying implementation of a package until someone wants it. Otherwise you can spend weeks on a program and then find that it is not being used.

GOINGS: What criteria do you use to move a program up into the implementation range, or to Category A?

DENK: In preparing for a workshop we will get a number of programs in that subject area working, at least temporarily, on the system. The authors of important programs or their disciples are invited to the workshop to discuss with participating faculty the pedagogy involved and the operation of the program. After the meeting we conduct two informal surveys (one immediately following the workshop and the second about two weeks after) to ask participants if they want to have this package available on-line. If the program is called for by one or two people, we move it up onto the system.

We monitor user requests for usage of all programs on-line. Our on-line library is close to 500 programs dedicated to education, in addition to the statistical packages library which is formally supported by the computing center.

GOINGS: Our computing center, a small shop, sees its functions as keeping the machine operating and providing languages and statistical packages for the user. Although we do try to encourage the use of the computer by the faculty, really this is the task and responsibility of the departments. Two of them, sociology and education, have been working with Project EXTEND.

DENK: Ultimately you may be right, yet the computing center must provide the enabling means to allow the use of computer aids in instruction. The computing background of most faculty members is extremely limited. The computing staff has to provide the technical knowledge. When a good program is functioning, many more people may become interested.

GEITKA: At Oakland University work for the administration consumes most of our time. Until recently little attention has been given to faculty and students. However we hope that with our new terminal and assistance from Project EXTEND which brings us on to the MERIT Network, the faculty will gain experience with large-scale computing. Also we have recently acquired the SPSS package which we believe will increase the computing center involvement of the faculty user.

David McBLAIN (Austin University): Do you think that students can be successful in pressuring faculty to use the computer?

Ronald CODE (Northern California Computing Network): Although the faculty can be very good at self-defense, this can work to some extent. However some computing centers seem to work very hard at convincing the faculty and students to stay away from the computers.

Contracting for Computer Resources

Chairman: Harry B. Rowell, Director Operations
Carnegie Mellon University

Recorder: Timothy Zorka
Wayne State University

To open the meeting, Mr. Rowell distributed copies of a tentative checklist of terms and conditions for use in contracting for purchase of computers from vendors. He explained the objectives of the panel: (1) to exchange terms or conditions successfully used by universities, and (2) to establish a checklist of issues, terms, and conditions to deal with as an institution entering into contractional agreements.

Discussion focused on interinstitutional contracting for computer services as one variation of contracting with vendors. Especially with interinstitutional contracts, intentions and capabilities must be made. Verbal agreements and statements such as, "Well, we'll do the best we can" are often accepted in place of contracts.

There has to be "good faith" between parties in areas such as turnaround time and quantities of disk storage. However, standards of performance must be defined to mutual satisfaction and remedies provided in case a particular service isn't provided by the supplying computing center.

For example, in a contract negotiated between the National Bureau of Standards and Yale University specific statements are included on availability of service for a stated number of hours per day and days per week. If the computer is down 48 hours in succession, NBS can break the contract. If the down time is 49 or 50 hours, the probability of NBS leaving is very slim, but the option is there.

Institutions are inexperienced in interuniversity negotiations. These negotiations are causing many computer centers to examine the manner in which business is conducted internally. Computing center directors are not accustomed to dealing with users internal to the university on the terms of guarantee, regardless of bounds that are set.

To cover costs and needs of both negotiating parties, some guarantee of minimum availability of equipment to the user is necessary, as well as a guarantee of minimum computing by the user. Perhaps a daily "pot" should be created against which the user may charge computing time if used, or forfeit funds to the supplier if not used.

The University of Toledo is trying to work out agreements with Bowling Green University for remote access to the Bowling Green University Computer. Remote connection is scheduled for January;

eventually a regional center is planned which will be headed by a director independent of all institutions. Agreement is necessary between the presidents of the institutions negotiating interinstitutional contracts.

One major problem with convincing the presidents of the universities of the merits of such an agreement is that the people presenting the arguments on the merits aren't convinced themselves or don't understand them fully.

Center directors can be confident with network arrangements only as long as there is an external source of funds such as a National Science Foundation grant. If the center is wholly dependent on the institution's resources, balancing the center's budget must take precedence over objective evaluation of resources available remotely. Interinstitutional contracts must reflect this necessity. Objective consideration of the options for remote use of computing resources can be made only if the administration of an institution is willing to protect the computer center financially. It took two years to work out an interinstitutional agreement for computer use through the MERIT Network.

Each computer center director should be encouraging network use and providing network service. At the same time he should be held responsible for operating on a budget and meeting it. He needs to expect a certain income and anticipate a balanced budget at the end of the fiscal period. The administration at each participating institution must be prepared to cover a potential deficit.

If a long-term deficit occurs on the MERIT Network, perhaps due to a University of Michigan physicist using the CDC machine at Michigan State University, this expense should be budgeted separately.

A long run deficit will force the deficit university (the University of Michigan in this case) to reevaluate the facilities available locally to determine whether the resources should be enriched to fill the needs of the physicist, or whether it is more cost effective to let the physicist continue to use the remote Michigan State University facilities.

Regarding Future Task Force Activity Mr. Rowell offered to add examples of clauses dealing with inter- and intra-institutional negotiating problems.

CHAPTER 5

MAJOR ISSUES AND CONCLUSIONS: PANEL DISCUSSIONS

Common Themes and Consensus: Report and Discussion of Workshops

Moderator: Martin Greenberger, Professor of Computer Science
The Johns Hopkins University

Robert Chenhall
Executive Director
Museum Data Bank Coordinating
Committee

Richard I. Hofferbert
Executive Director
Interuniversity Consortium for
Political Research
University of Michigan

Frederick Kilgour
Director
Ohio College Library Center

George Sadowsky
Senior Research Staff
Urban Institute

Sally Yeates Sedelow
Professor of Computer Science
and Linguistics
University of Kansas

Harrison Shull
Vice Chancellor for Research
and Development
Indiana University

MODERATOR:

In this session, we shall hear a little about what happened at each of the workshops and see if we can begin to find some common themes across the various fields and disciplines covered in these workshops.

On the panel are representatives of the museum and library fields: Robert Chenhall of the Museum Data Bank Coordinating Committee and David Penniman of Battelle Memorial Institute; also representatives of the physical sciences, the social sciences, and the humanities: Harrison Shull of Indiana University (Chemistry), Richard Hofferbert of the Interuniversity Consortium for Political Research (Political Science), George Sadowsky of the Urban Institute (Economics), and Sally Sedelow of the University of Kansas (Linguistics).

To provide a framework for the workshop reports, the panel proposes to consider networking, not in the restricted physical sense of a national telecommunications network interconnecting computers and terminals all over the country, but rather in the more general and familiar sense of people working together in an organized fashion to share resources and meet common ends. In the more general sense, of course, we have had networks on the national and regional levels all along. What makes the subject of moment today are the technological advances that facilitate new ways of networking.

In the more general context, we can talk of what the fields and disciplines need in the way of resources and services, where they now stand, and where they would like to go if the necessary means were available. We can talk about opportunities for sharing resources and about incentives for finding these opportunities and taking advantage of them. Most of us will agree that higher education has fallen short in availing itself of such opportunities. We can talk about the kinds of organizations of users that would make most sense for a given field or discipline, and then, coming to the technological question, we can ask in which respects a computer-communications network is required or desirable, and what new possibilities it offers. Once having addressed the technological question in this broad context, we might then reexamine the separate issues of user needs, sharing, and organization in the light of the new conditions and possibilities that computer-communication networks create.

Our approach to the subject can be put into the schema shown in Figure 1.

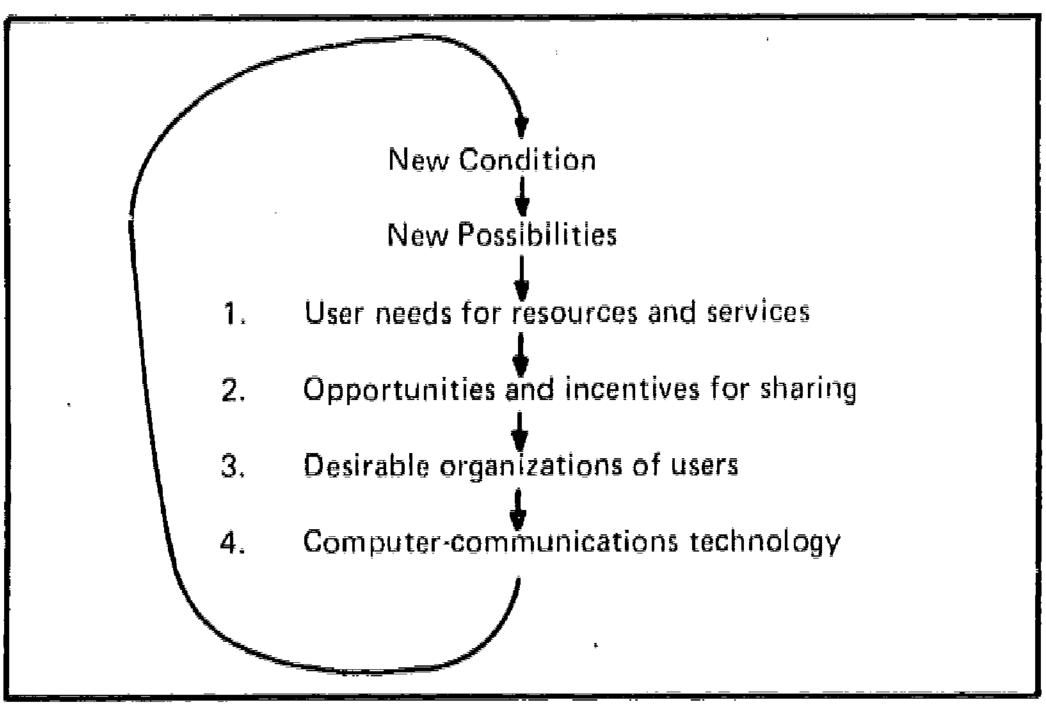


Figure 1

It is more than we can hope to accomplish in one session, but perhaps we can make a start.

Beginning with the first cateogry, the panelists will now draw on the workshop discussions to report on the kinds of resources and services that their field or discipline would like to see developed or improved.

DR. CHENHALL:

In the broadest sense, the resources and services that are presently available to museum related disciplines are virtually nil. There is no interconnected communications system in existence and we are many years behind the field of chemistry, for example, in this whole area. However, we have taken a major step, and have now reached the point where we do have a vehicle for communicating the needs and resources which are available to the widely diverse groups of people who come under the whole category of "museums." This new organization, The Museum Data Bank Coordinating Committee is, one of the greatest resources we have. We also have available several different systems for the computerized cataloging of museum specimens. However, an interconnected network of these systems is a long way off. At this stage, our network objectives would have to be considered far more in the nature of logical networks rather than physical networks. Whether the need for physical networks is there or will ever be there remains a moot point at this time. However, we have achieved this first step of an organization network which has to occur before there can ever be any further development.

There is a very real incentive for the people in the disciplines that I represent to share the information which each of them has locked up in the vaults of their various museums. The primary incentive is not as much an immediate *sharing* of the information as the cataloging of that information so that it is accessible to museum directors and to researchers -- on an intra-museum basis rather than on an inter-museum basis. Ultimately the opportunities will be there and the needs will be there for the true network kind of sharing of data. However, at this stage, the primary incentives are for developing intra-museum information systems. Opportunities for sharing common data structures are far more available than opportunities for sharing specific data, as in the social sciences, or than using physical facilities to interconnect the data.

At this stage in the museum related disciplines we are engaged primarily in development of techniques by which we can describe physical objects in machine-readable form. In essence this means being able to precisely describe physical objects in terms that are meaningful to another person. The opportunity that we have through the Museum Data Bank Coordinating Committee is simply to coordinate the various efforts to develop techniques of description so that eventually a true interchange of the information base itself will be possible. That interchange of information is a long way down the road.

I agree completely with Dr. Hofferbert's comment that the opportunities for sharing are in truth determined by the organization. In the particular field that I represent, we have to consider organization along several different lines. One of the reasons for the formation of the Museum Data Bank Coordinating Committee is the fact that organization along disciplinary lines simply did not seem to be the best way of approaching the whole problem of sharing museum data or resources through networks. An organization can be and sometimes is very effective along disciplinary lines, but networks must be organized also along the lines of common data needs. All museums are concerned with the inventorying or indexing of primarily verbal data concerning physical objects. This I consider as a common data need. Organization can also be along the lines of common needs for analytical capacities, but in the disciplines that I represent, the primary organizational need is for common data.

There presently exist two different levels of organization, both of which are important and both of which have specific functions to play in the overall problem of networks for museums and related disciplines. The first is organization along the lines of individual systems. We are presently using some four or five, perhaps more in individual instances, different general information system packages. There is communication among users of each of these packaged systems to improve the systems and to provide the documentation necessary to make them work. At a second level, is the felt need for an organization along the lines of what we are trying to accomplish with the Data Bank Coordinating Committee: organization of common data structures, common conventions of recording data and, to some extent, common terminology. Organization at both of these levels is being accomplished. Most of us who are deeply involved in this feel that the organizations that are presently available will provide opportunities for the sharing of resources.

DR. SHULL:

I represent the field of chemistry in this panel. You heard a good deal about the organization of chemistry from Dr. Lykos yesterday, and I think it is clear that chemists have an extended series of resources already available within the discipline that make sharing a natural phenomenon for us. The Chemical Abstracts Service is over 70 years old. It has become the largest abstracting service in the world. It is totally computerized in its journal production as well as in its abstract production and is now producing tapes for bibliographic searches. Chemical Abstracts provides a nucleus of resource sharing in chemistry that probably has few parallels in other disciplines.

In addition, chemistry is fortunate in being a fairly mathematically-oriented discipline. It has the resource of many individuals who have some computer expertise, not only in the rapidly developing fields of on-line data acquisition, interfacing to computers for laboratory experiment control, and in the use of computers as a tool in computational work and theoretical work, but also in the use of information retrieval and information processing.

I wouldn't want, however, to hide a phenomenon we recognized in our workshop as applying even to chemistry. People outside our discipline seem to think that all of us are computer buffs and that we know all about the computer. The fact of the matter is that we are very, very far from being computer saturated. Even in a relatively well-known and advanced department like my own, probably not more than ten percent of the faculty are real computer buffs in the genuine sense of the word. Those of you who are in disciplines other than the physical sciences may be a little bit surprised to realize how far behind some of the faculty are as compared to the rest of them. However, computers are a phenomenon of the last decade, and our faculty span a period of fifty years. Many of our older faculty have had no computer experience, and it is going to take some time before 100 percent of them are comfortable with computers. Computer expertise is not always inversely proportional to age, but there is a remarkable correlation.

Chemistry has many resources. Our resources can be developed and already are being developed through various types of sharing mechanisms. The real problem as expressed by the common chemist, is the same one the rest of you have -- money, and not just money, but the distribution of money. It is very, very difficult to redistribute money from old-fashioned technologies into new ones. Without question, the most general resource we all need is the ability to proceed with the ideas that we already have.

Sharing is not specific to chemistry, nor to any point of view represented on this panel. One of the prime reasons for sharing is to use our resources more efficiently. The overall reason for networking, if there is a single one, will probably be that it is more economical to share our computer power that way than it is in any other. Within the field of chemistry, we need and can make use of the most economical computing power available. However, economy is not, by any means, the only reason chemists need to share. I began a sharing program a decade ago, because of the realization that, despite very rapid technological advance, my discipline of theoretical chemistry was standing still in time. Each individual's achievements in every two-year period was being repeated in the succeeding two years, with a succession of rewrites of computer programs. Each individual program was rewritten with increased sophistication, of course. Science has a long, fruitful history of people standing on the shoulders of people who have gone before them, but theoretical chemistry arrived at a point in time where no one was standing on the shoulders of anyone before them.

We become involved in the sharing of computer programs in the hope that at least some of these programs would be useful to others. My idea was that, as machines were moved from one institution to other lesser ones, at least the programs that were good on the machine would be useful when the machine was moved to the next place. It subsequently did not occur in just that way, but the sharing has enabled people in a limited way to stand on each other's shoulders. The real need as I see it is a new organization designed to be a national center for sharing. We need a center for the sharing of expertise that leads to genuine, continued development of software that far transcends what an individual person, or any isolated portion of a discipline, can do alone. That is a sharing in progress. We do share data and computer programs, and we have always shared educational techniques in a limited way through the distribution of books, textbooks, reviews, and summaries. However, the ability to share educational programs is only just beginning. Herein lies a real opportunity for the future. I think we know very little, in fact, about *how* to educate. Two few of us are working on the very basis of what most of us are doing.

There's a well-developed way of sharing very large expensive resources. It extends not only to things like computers, but also to large machines like accelerators in physics. But I don't think one should forget about sharing expertise, sharing people. One simple way is to move them around a bit. One ought to consider more broadly, for example, the kind of techniques that the CIC has developed in the big ten schools of allowing graduate students to move among the institutions rather freely as required by their programs. Methods of sharing expertise across our institutional barriers are also furthered by the study that the Sedelows are doing – examining institutions and determining what it is that can be shared among them.

The problem of organization seems to me to be a very diffuse one. My view is that organizations of every type, characteristic, and nature that might be desired already exist. The question is whether or not they satisfy the particular needs that exist at a particular moment. I found it very surprising that there are over five hundred interinstitutional organizations of universities. There is one for any activity imaginable. The question is, what is it you want to accomplish and how do you identify the entrepreneur who's going to accomplish it? The combination of answers to these two questions will produce a new organization or make an existing organization do the job. In my own field, we have many organizations and there are new ones developing all the time. We have program exchanges that are operational. We have interinstitutional cooperation in the discipline through our national society and through our interuniversity regular meetings. This doesn't mean, however, that we have satisfactory organizations to do all the jobs required. As an example, the chemists are just completing a one-year study to determine whether there is a need for a national center for computational chemistry. The committee is finally going to conclude that there *is* a need for a center in computation and chemistry for exactly the kinds of reasons that become apparent as you discuss the needs for future development. We *haven't* been standing on the shoulders of the people before us. We do need long-term continuity. There is an opportunity to develop a national and international resource to further develop the uses and utility of chemistry in human welfare. These kinds of needs will lead to an organization that will satisfy those needs. The real question is, can we find the entrepreneurs to coordinate the backing, the support, and the mechanism to make the organization alive and thriving? Another example of an organization badly needed within chemistry, quite apart from the general need of an organization like OCLC, is an interlibrary organization. People are working on this on an individual, ad hoc basis. Almost certainly, the result will be an organization designed to accomplish perceived needs. Organizations do follow the needs of the people rather than precede them.

DR. HOFFERBERT:

The social science workshop did not have quite the clarity of focus some of the other workshops may have had, but we did flow rather directly from the themes which James Davis set forth yesterday. We focused our attention primarily on the output side of the data question: what do we do after we have material, recognizing that there are considerable social science data being produced and also that the nature of social science theory and research practices today are such that one needs a lot of material. One makes up for the absence of theory with search and destroy missions through large quantities of data.

The resources needed is a difficult question to confront. Discussion in the workshops was not unfocused, but did cover much ground. In Political Science, we already have a track record of supply having structured demand. The caricature is: computer as toy, preceding creation of a group of people to play with it. However, there is an intense need on the part of a few people for clearly defined forms of assistance. The track record of which I speak focused somewhat on activities such as those of the Interuniversity Consortium for Political Research. ICPR supplies large quantities of data to many schools. We have some record of what happens to data and how much of a need there is. We also have some record of costs for supplying large amounts of diverse types of data; and we have enough experience in this to recognize that we could not have foreseen our present structure ten, or even four, years ago. The kind of material provided has affected the market, giving rise to the questions: what is the magnitude of the market and, how many different needs can be met?

The themes we touched upon, however, were data transmission, remote access to data, and analytical capacities -- that is, programs (both hardware and software) available from distant points. We also touched upon the problem of inventorying the availability of these resources. How does the social science community learn that machine-readable data sets, for example, exist in a particular substantive area in a particular setting? How, in turn, do they learn how to get? How do they learn what kind of analytical capacities are available elsewhere (software, hardware, networks, etc.)?

It is fairly obvious that questions of sharing resources and organizational structures are intricately intertwined. One of the facts emerging from ICPR's experience, which should be self-evident, is that norms for sharing develop as mechanisms for sharing prove their own success. Certainly this is true in the data area. Ten years ago, we were not able to convince some people to give us data. Today, we can convince these same people because others have given us data in the intervening years. The same thing holds true in the software area. The fact that software is being distributed at minimal cost from places such as the Institute for Social Research is going to encourage the general availability of software for distribution. To the extent that organizational structures are established to facilitate such distribution, the norm itself will accelerate exponentially.

As we move down the schema which Dr. Greenberger has outlined, I am wondering if it might not be helpful to make some distinction between the types of needs that we are talking about, the kinds of things that we are talking about sharing, and the sorts of organizations that we are talking about establishing. As I listened to the conversations yesterday and today, I was increasingly convinced that there are at least three different types of phenomena we are discussing. They overlap a good deal, but there are certain centers of gravity; we have at least a trimodal distribution of needs and interests.

First is the question of data volume. In the social sciences particularly, in political science, history and sociology -- the need is for large amounts of data, transmission and access to vast records of human experience, submitted to relatively simple analytical techniques. There is a second set of needs, however, where the primary interest is in analytical procedures: modeling, simulation, and various forms of open-ended analytical tasks. The need in the second case is more for access to hardware and software rather than access to primary material. The final set of needs, where the primary interest is inventorying, is shared by many persons whose modal need fits one of the first two categories. Inventorying required the development of classifications and the transmission of information about what lies where. The kind of sharing of information currently takes place (increasingly in the social sciences) in the form of episodic newsletters. Assuming that some mechanism of more rapid communication is devised, I would suspect much of the content of that mechanism of communication will be in the form of inventoried material. Thus, I see at least three different kinds of sharing: data, analytic capacities, and information.

We should be disturbed by the indication that there may be a trend toward the commercialization of information and analytical capacities. We are all dealing with the same public agencies who provide us with taxpayers' money or with the money they have been able to donate because they got tax deductions. The kinds of capacities that we are building may indeed be the product of our own genius, but they are also the products of community resources. Market mechanisms in the areas we are discussing are crude analogies at best. Sometimes the hand that is presumably invisible in the free market is eminently visible for the scientific market in the form of the National Science Foundation. The problem is how to define the buyer. Is the buyer the undergraduate who wants access to a particular resource? Is the buyer the research organization of a quasi-governmental or non-profit nature? Is it the collegium defined by the various institutions that we represent?

The consortium with which I am associated made an early determination that may be a lesson for us all since we are still alive and seem to have survived reasonably well. Although it is not without some problems, this early and important decision was that services were going to be provided to people in return for an institutional commitment. Dues are assessed to member universities. Although we don't sell data, we do provide services to people at nonmember institutions on a fee basis. Our basic commitment is to interact with representatives of institutions. The institution provides the commitment, just as it does in the case of our host organization today. This institutional commitment does several things. The fact that a scholar at an institution is able to acquire data or software or training from us as a result of his institution's commitment lowers the barrier to usage and invites trivial usage. Trivial access to data and software is not unlike the trivial wanderer into the stacks of the library who accidentally finds the book that may be useful or, the listening room in the music department where one randomly wanders through the files of records. It is a stimulus to serendipity. Most of us learned early that serendipity does not just happen — skilled people with good resources find things that are valuable. Easy access to data is also a stimulus to the norm of sharing. Because the buyer and the provider in most cases are two different organizations, the problem of the provider's paying to use his own data is avoided.

In the organizational realm, furthermore, we are not talking about just a resource investment, that is, dollars in return for data, training and software. It is really a human investment; there is a person or set of people at each of these institutions with whom we interact. As often as not, our staff interacts with them on a first-name basis. The staff provides tapes, software and, when scholars come to Ann Arbor for the training program in quantitative analysis, they meet, talk, and write together. They are fairly task-oriented, with a set of tasks in common, and around that set of tasks they develop a certain community of interest. The tasks combined with the community of interest and with the sense of mutual involvement in this kind of a sharing apparatus, do a good deal to dilute the harsher outlines of a pure market model. Collectively, we should do all we can to discourage the commercialization of research resources. Certain market analogies would be appropriate in terms of funding, but there's a certain model here, in the ICPR case at least, that deviates significantly from a commercial model. It suggests that one can indeed operate sharing organizations without a straight "fee for service" structure. The absence of a "fee for service" structure may stimulate utilization. It also stimulates some wastage. We sometimes worry about what happens to these data when they are released, but the net effect is beneficial and multiplicative.

DR. SEDELOW:

I should begin by noting that the participants in the language and humanities session represented a wide range of disciplines. As you probably inferred from Walter Sedelow's comments yesterday on our project concerning language research and the computer, such a range seemed particularly appropriate to us. Just to be amusing, we categorized the participants in the session as "humanists and humanitarians." We defined humanitarians to be those people who are nice to humanists. More seriously, we welcomed social scientists who have an interest in language, physical scientists, and computer center directors, as well as people from art, music, language and literature.

We did circumscribe our discussion somewhat, as I assume the other sessions did, by concentrating upon research and teaching which is computer related. Not surprisingly, given the workshop chairmen's interests, there was a strong tendency to define the humanities in terms of languages or symbol systems. We produced a list of resources and services which isn't at all comprehensive, but I will indicate some of the items on the list.

First of all, we need reference materials of all sorts, including information as to what has been done and is being done. Walter indicated yesterday that we included more than 200 pages of bibliographical references in our monograph, simply because we found that neither the informal nor the formal communication networks are satisfactory to reach the researchers who might have things to say to each other in these fields.

In order to use the computer, we need good documentation. Cited under documentation are tutorials, self-teaching programs concerning any computational device the user might wish to employ. Programs and operating systems must be adequately documented.

We need programs, both application programs, and, if you want to think of computer languages as programs, languages better suited to our needs. It is the case in language research -- perhaps more so than in many other fields -- that there hasn't been as much incremental progress as is needed. The programs that have been produced have tended to be very complicated; they've been implemented on a particular machine and on a particular operating system. They haven't been easily replicated, nor even replicated with a great amount of diligence and effort, on other machines. The result is that we've had good starts and the starts haven't led as far as we would like.

Data banks represent another important resource. We need data of all kinds for the range of symbol systems represented by natural languages, programming languages, music, art, etc. We need reference works such as the *Oxford English Dictionary*. We have an edition of *Roget's Thesaurus* and we have an edition of *Webster's Dictionary*; we need more information of that sort.

We need expertise, including not just the very important technical expertise represented by those people who are concerned with making available computer power, but also access to other scholars' work within the disciplines comprehended by our workshop.

Finally, we want computer power, and that includes availability. One of the aspects of the study in which Walter and I have been engaged, is discussion of computer power in terms of availability and reliability. We've talked with many people and attended a conference on computer software development, design, and validation, and we were interested to note that numerical analysts, for example, were very much concerned with reliability. However, the people in our area of concern were focused upon availability. Apparently, it's only after something becomes available that you then begin to brood about whether or not it's reliable.

I wasn't able to elicit much conversation from our workshop concerning the topic of sharing. It isn't that we weren't brought up properly — we were all taught to share. Rather, the situation as to *having* something to share is rather bad. In fact, whenever someone ventured an opinion as to what the current sharing situation might be with reference to some particular program or artifact which might be of interest to various sets of researchers, someone else would quickly interject, "Well, what we need is . . ." It was very hard to pin down the current "sharing situation."

I can make some obvious points about facilities which exist for the sharing of information in the area of interest to our workshop. There are, of course, publications; and our situation is that of many other fields — that there's a serious time lag. For example, people who have written interesting computer-based programs for various aspects of language research find that the publication of the description of their research often doesn't occur for a year and a half or longer. There are newsletters, to be sure, and they appear with a little more speed. But newsletter editors very often either don't know about relevant work which is in progress or has been concluded, or don't have space to report on it. It's especially true that work on language research or symbolic behavior, which is one way of talking about the humanities, really isn't divided up as it is in those often used and often misleading "classical" guides to classification — the college or university catalog. Knowledge, at least insofar as it is relevant to our concerns, isn't cut up as it is in those particular artifacts. There are professional organizations which produce bibliographies but the point I just made applies again. People in linguistics sometimes look at the Modern Language Association bibliography, but often they don't. And those computer scientists who are interested in programming languages or formal descriptions which might be applied to natural languages do not, I am sure, frequently browse through bibliographies published by the Modern Language Association. This point could be made in a more extended fashion by citing the many other examples of undesirable compartmentilization, but I don't think doing so is neccessary.

Other than these topics concerned with information, the state of sharing within the languages and humanities might best be described by employing, as a speaker did yesterday morning, economic metaphors. The condition currently seems to be one of barter and philanthropy. If someone has a program or a data bank, he very often is willing to exchange it with someone else who has something attractive; or, very often one finds either a particularly generous individual or one who has been endowed by the National Science Foundation or some other agency or foundation so that he can afford to dispense largesse in the form of programs or data banks. This form of sharing has taken place, for the most part, on a very informal basis. There is no organization which facilitates such sharing.

A central theme in the workshop on languages and humanities as well as in the study with which Walter and I have been involved (especially in a working conference we held on organization) was centralization versus diffusion. In general, I think that in an area so broadly defined and diffuse as ours, it would be a great mistake to move toward premature closure by insisting upon centralization, by moving toward what many people see as a kind of an inbreeding which would be exceedingly unhealthy, at least at this stage, in the areas represented by our workshop. We have detailed in the report some of the implications of this conclusion and I would be glad to talk more about them if you have questions and if there's time.

In summary I should say that what we do recommend (and I think there was sentiment for this recommendation in the workshop as well) is some realization of a distributed center within some form of network.

DR. SADOWSKY:

We began the economics workshop by enumerating the actors in the economic research and policy arena which fall into three very broad groups: business economists; government economists in federal, state and local governments; and economists in academic and non-profit organizations. These economists tend to move with some frequency both geographically and back and forth between and within the groups. The research arena in economics is fairly decentralized with the exception of some accumulation points such as the Federal government and major universities. These economists rely to a great extent for their work upon data from the real world. Ours is an observational science, and as Sandy Berg pointed out, we rely to a large extent upon observing economic behavior around us. This situation is slowly changing as controlled experimentation in economics begins to take place; as it does, we ought to be in a position to capture data very much like laboratory experiment data. But, at the present time, most of our data is observed in an environment over which we have limited control.

The data can be broken down roughly into "macrodata," of which there's not very much, and "microdata," of which there's a great deal. Economists involved in micro economic studies tend to demand computational support of the type Jim Davis described yesterday. It's interesting to note, however, that although all of the macro economic data available in the world and likely to be used for economic policy purposes can be recorded on about eight 2314 disc packs, most systems for accessing and processing these data are still rudimentary. As a result, poor data often substitute for accurate data or for more reliable or more appropriate data in economic research studies.

Decentralization of teaching and research economists and of the resources they use is a central issue in networking for economics. The promise of networks for us is to provide an organization for this decentralized activity that will diminish replication of effort, increase ease of transferability and communication, and generally allow more efficient use of our resources as a professional community.

The primary resources that economists have to share are substantive knowledge and a large supply of data which is geographically decentralized and used by many economists. Part of the substantive knowledge is embodied in computer programs that implement statistical methods and economic and econometric models.

For academically oriented economists, professional incentives result in the dissemination of research results, new knowledge, in the form of publications. Unless a researcher has entreprenurial ambition which he wants to exercise in an academic setting, there is little professional incentive for widespread sharing of either computer programs or primary data. In fact, there are disincentives. Materials developed in the course of one's own research are generally not as well documented as one would like prior to releasing them to others. Thus, even if substantial externalities could be obtained for the whole economic research and policy community by disseminating machine-readable programs or data, the reward structure of the profession tends to discourage it. Certainly the exporting technology that we now have available discourages it also. We hope that networks will substantially decrease both the financial cost and the time costs of the exporting process and lead to more resource sharing in economics than would otherwise be the case.

One other observation touched upon repeatedly in our workshop was that there appears to be a lack of leadership by the data originating agencies, primarily in the Federal government, in making data resources available in an easily usable form. With some exceptions, the role of the Federal government as a data producer is not very different today than it was thirty or forty years ago when the primary method of dissemination of statistics by the government was the production and publication of aggregate tables. This results in multiple distribution of tables, multiple inputs of time series and tables and other data into computers and a large amount of duplicated work. Resources which have alternative uses would not be squandered in duplicating efforts if there were a better mechanism such as a national computer network for originating and sharing access to this primary data resource.

Let me introduce our contribution to the issue of "organizational framework" by repeating a story related in our workshop about the process by which the President gets the monthly unemployment rate. After approximately fifty thousand households have been surveyed, using a reasonable sophisticated panel sample survey, optical scanning devices and a great deal of processing on a Univac 1108 computer, the numbers are finally delivered to the Bureau of Labor Statistics. A person at BLS then telephones a staff member at the Council of Economic Advisors and reports the principal unemployment rate and various other supporting statistics. It is important that the President be notified of the facts before they are released to the general public. A runner is then dispatched from the Council to the Bureau of Labor Statistics, which is located about 7 blocks away, to pick up more complete information too volunious to be relayed over the telephone. This process may repeat, and interacts with analysis performed by members of the Council staff. Then the President is briefed on the magnitude and meaning of these new data. This process takes about an afternoon, a quite timely transfer of information compared to what happens in the research community; yet it's rather shocking compared with the technical alternatives that are currently available to support information systems.

An important point made by this example is that the organization of "additional networking facilities," whatever their implications for economics, should unify the process of providing reliable data to public agencies, private researchers and commercial economists. In addition to supporting the timely distribution of macrodata, networking facilities would be equally useful in disseminating microdata sets and in producing and making available programs which include economic models and techniques for application to the data.

DR. PENNIMAN:

Fred Kilgour asked me to apologize for him. He had to return to Columbus to finish a grant application. Because I'm not with OCLC but with Battelle, my comments should not reflect upon OCLC's opinions.

I think in the workshop yesterday, a couple of points came through pretty loud and clear. One was that networks do exist already. Resources and services are being shared. One good example of that is what Dave McCarn presented concerning the National Library of Medicine. In addition to sharing information, OCLC, for example, shares what I would term analysis. They're sharing the cataloging that goes on in the various libraries. I think also that within the next several months, as Dr. McCarn will attest, there are going to be several new services emerging. It's interesting that these services are going to be tested and will have to be viable in the marketplace. I'd really like to comment more on that when we get to organization, because that's where the comments are most appropriate.

There's no question that the kind of sharing which Fred Kilgour described in OCLC is worthwhile and is very beneficial from a number of standpoints. Ed Parker indicated last night he thought OCLC might be an exception rather than the rule in the sharing of that kind of an information resource. I'm not sure that I agree. Just because OCLC indicates a dollar savings for the participants should not make it an exception. But that brings up a valuable point and here I'll be a little more negative in terms of sharing than some of the other speakers have been.

There are a number of reasons why one might share. First, information being shared might not be worth anything anyway so one doesn't mind an exchange. Secondly, what one gets in return might be of the same value as that given. Historically, thanks to Andrew Carnegie, information is generally considered a free commodity; you shouldn't have to pay for it. While this is true in the public libraries, it also seems to carry over into other kinds of information services. We're seeing a change and as an example of that change, I'd like to describe a typical information-sharing activity as it used to exist at Battelle. Through information analysis centers, Battelle collected specialized information and provided it to technical specialists in an area such as aerospace. We would collect data and information; they would contact us and ask us for a specific analyzed answer to a technical question. We would provide that answer and at the same time say, "Hey, have you got anything that would be worthwhile putting in the file?" Because they were pleased with the answers, they would give us new data. More recently the government agencies funding these information analysis centers have said "We've got to start recovering costs. Start charging for your outputs." Now when somebody calls us and says, "I've got a technical question, what's the answer?", we give him the answer and say, "X dollars, please." Then we say: "Hey, have you got any data I can put in the file so I can sell it back to you later on?"

The necessity of charging for services has begun to limit the kind of sharing that can take place. However, we're going to see more and more of information services moving into the commercial sector. Even government services where charging is being implemented will look more and more like commercial services. The sharing of resources among these various competing commercial groups probably will not be of the type that's being discussed here. What will be shared will be the customers. There is a way that the customer could benefit from this. If the organizations that have to either make it or not make it commercially learn to share techniques for information access, then one customer using a remote terminal could easily access several information services from several different organizations. Everyone would benefit, not only the customer, but also the supplier. I would look for more sharing of access approaches among information services than anything else.

I hope the audience will have a chance to comment later on this because Dave McCarn's probably wanting to disagree.

A few years ago I was deeply involved in an NSF-sponsored study to come up with a plan for a united engineering information system. I don't know whether any of you are familiar with that study or not. We spent many heartbreaking months trying to determine what kind of organization would best serve the engineering community. Because it was NSF sponsored, we had a number of advisory panels and project monitors guiding us along the way. We ended up with four alternatives. The first one looked very much like a planning, coordinating agency that would help set standards and identify needs. This one would have, been funded through NSF or some other government organization, with no outside support or very little outside support. The final or fourth case was at the opposite end of the spectrum and was completely commercially viable. The organizational structures of these two cases were entirely different. If what I said earlier about the best in the marketplace holds true, then the organizational structure of any kind of sharing activity has to have built into it at least some aspects of the marketing structure. If one can be assured of funds from NSF or some other agency for continuing activity, then the organization can be structured very differently. However, the experience of some organizations like ASM in establishing information services indicates that there has to be some kind of marketplace test, and that the organizational structure should reflect the results of that test. ASM had some initial funding through the government. When that aid was suddenly withdrawn ASM had to try to fly completely on its own and didn't make it.

MODERATOR:

We would probably do well to think more about the possible conflicts between commercialization of services and the sharing patterns that have grown up in non-commercial environments.

One point that George Sadowsky made is also food for thought. One of the ways that economists package knowledge these days is in the form of models; models can have a very strong knowledge component. This is a new form of intellectual expression. We all know how to share knowledge through the writing, publication, and distribution of textbooks, but we have not yet developed very good ways of sharing the knowledge that is incorporated in the models. Perhaps networks might provide a way.

In discussing the question, "How do we organize to get on with the job?" the panelists can draw not only on their workshops' thinking into the future, but also on their own experiences in the organizations in which they have been and are involved.

It seems we are out of time. It may be very significant that we have had such an interesting and meaningful discussion without even getting to the question of technology. This obviously is just a start. We have to get into these issues in much more depth and with more care than is possible here, and this is what we intend to do during the EDUCOM General Working Seminars on the National Science Network. The output of the General Working Seminars will be the subject of the EDUCOM Spring Conference next year.

I thank my colleagues on the panel for their very interesting remarks and the members of the workshops for their contributions to these remarks and to the ideas the panelists have brought with them to this session.

Networks and Computer Centers: Cooperation or Conflict?

Chairman: Gerard P. Weeg, Director, Computer Center
University of Iowa

Maurice P. Brown, Director
Office of Computer Coordination
Ontario Council of Universities

E. Rex Krueger, Director
Computation Center
University of Colorado

David O. Harris
Director of Computing Services
University of California, Santa Barbara

Clair G. Maple, Director
Computation Center
Iowa State University

Julian Kateley, Director, Computer Center
Michigan State University

DR. WEEG:

This panel has been convened to discuss the political and economic problems attendant upon the creation and operation of a computer to computer network. We are not discussing the very valuable "star" networks, but rather networks of free standing equal-rights computers. To establish such a net, a number of interpersonal, political, economic, and technical compromises and decisions have had to be made. It is just those compromises, promises, and nuances I hasten to add, which I want our panel to discuss today.

To place each network here represented I have asked each panelist to describe his own local network, realization of a network, etc.; however, I would prefer the lecture not be bogged down with discussion of particular hardware and software utilized by these gentlemen. Therefore, I would prefer to let them dedicate the principal part of their lecture to the problems they have faced in the realm of economics, politics, and interpersonal.

I'm happy to say I have located five existing computer-to-computer networks in the North American hemisphere. There are probably more however, you understand that these are a very well kept secret. Let me now introduce my panelists in the order in which they will speak.

First of all is Dr. Maurice Brown from the Council of Ontario Universities in Toronto. Dr. Brown represents the Ontario network. Secondly, Dr. David O. Harris, on the end, from the University of

California at Santa Barbara. Dr. Harris has been involved with a minimum of three networks; he'll describe one which he is currently involved in and the other which requires involvement no longer. Next, Dr. Rex Krueger, in the middle of the panel, from the University of Colorado at Boulder, who will speak about the Colorado net, which I believe is on the verge. OK? Dr. Clair Maple, on this end, whom I admire a lot, from Iowa State University, Ames, Iowa, will present possibly the most outstanding example of a network which I have ever heard of, namely his and mine. And then second from the end is Dr. Julian Kately from Michigan State University, one of my Ph.D. students. He represents the MERIT network.

I'm concerned as to the best way to handle questions but we will take some questions between panels and open it up to a free-for-all at the end. I would like to ask each gentleman to give his ten to fifteen minute presentation and we'll ask questions right there if you have some. One of the questions, if it doesn't come up, I'll propose is: What the devil good are networks at all, much less how did you solve the problems connected with them?

Well we will be staying seated, at least most of us, during our presentations and I'd like first of all to introduce then Dr. Maurice Brown from the Council of Ontario Universities.

MR. BROWN:

The Council of Ontario Universities is an organization which is responsible to fourteen autonomous universities in the province of Ontario. We are responsible for the development of an intercomputer network called METANET. The program grows out a thrust toward rationalization. The universities collectively spend twenty-one million dollars a year on computing and there are a large number of different computer types in the universities. The network proposal is an outgrowth of a concern by the presidents of the universities, which goes back to 1967 when it was felt that the current mode of financing computers was not appropriate. We began looking at the possibility of large regional star-type centers. This floundered on the basis of politics and we have, as a result, looked at distributed networks as a way to approach the problem. We have, we feel, come up with a technological solution which happens to meet the political basis that exists, namely: distributed centers of computing power and expertise in the individual universities.

We have a number of different levels of involvement in the design of the network. My group, the Office of Computer Coordination, is concerned with the technical design, and we are also responsible for marketing the concept to the universities. I use the word "marketing" advisedly, because to me, the big problem of networks is marketing, not technology. We have a network design group which is a group of computer scientists, electrical engineers, and computing center personnel who are acting as consultants to our group and are helping us channel the design into the appropriate direction. We have a task force on network utilization, which is looking at the applications for the network when it comes about. We have a task force on computer charging which is looking at the question of the currency of exchange: how are services traded between computing centers? As a parallel activity, I am a member of an advisory committee to a project called CANUNET, which may be a future national university network for Canada. This is developing quite well. We expect that there will be a submission to the cabinet of the government of Canada later this fall for a substantial degree of funding which would, in effect, create a network of networks, interconnecting regional networks across the country.

To give you an idea of the relative costs, we expect that the METANET program will cost about 3.6 million, and my own estimate for CANUNET is that it will cost about nine million dollars to develop. The operating costs for METANET will be, as far as I can see, about 1.2 million dollars a year. I'll come back to that later.

Our motivation, in the METANET project, is towards the rationalization of computing resources; a better way of spending that twenty million dollars a year.

There is, at the national level, a serious concern about communications across the country. There is within Canada a very strong element of regionalism. Each province has its own particular way of doing things, its own particular set of interests, its own particular jurisdiction. There is an awareness on the part of most Canadians of the fact of the United States just south of the border. We are also very conscious of the physical fact that north-south communication lines can be easily established. It's quite feasible to think in terms of people in Ontario accessing systems in New York State. It's quite feasible to think in terms of Vancouver people doing work, via communication lines, with organizations in Seattle. In terms of information systems, at the national level, there is a very direct objective developing, which is becoming an imperative: "that we establish east-west lines of communications," and this is the basis for our national program in networks.

In the province of Ontario there is, as I indicated, an implicit commitment in the network plan to, what I would call, "system planning." By that, we mean the coordination of long-range plans for computing in the universities. Clearly, to do this, we have to solve technical problems; we have to solve political problems. The point that has been discussed with the gentleman from MERIT relating to interinstitutional trade is extremely important. We are required to resolve the problem of the currency of exchange; we have to look at the whole question of balance of payments. We have to discuss what is reasonable in terms of protective tariffs.

In terms of where we expect to get the payoff with the network, we see it primarily in the economics of specialization. We feel that it is reasonable to think in terms of one or two centers serving the whole of the province in their particular field of excellence. We feel also that with the network we can in fact achieve economies of scale. We can have, perhaps, fewer but bigger computers accessible to the community at large.

Some of the questions I want to ask are: What are the alternatives to networks? How will the network be controlled? What does the network mean to the individual computing center? Is there going to be increased control from outside? Or does it mean that there is a larger market? And how does the computing center director handle the market, if there is a larger market? What about competition? What does the network mean to the individual user? I submit that it offers him increased choice. I feel it is going to lessen his influence on the local center. It is going to require that he have dollars which he can control, which he can spend for either computing, or travel, or salaries. To me, the long-run problem of the network is how to manage a distributed organization. There is the question of distributed marketing. You would probably agree with me that data service bureaus can do a reasonably good job, in a large part because of their aggressive and positive approach, they take their product and sell it. Universities tend to have a "laissez faire" relationship with the world at large. We are faced in the network environment with "laissez faire" liaison, which is the current practice, versus product endorsement, which means going out and selling what you have. I submit that "laissez faire" liaison means failure. What is required is for computing centers, computing center directors, to realize that the network is a product which they can use. It is a means of delivering their product and it is up to them to go out and sell it. Perhaps they have to hire salesmen.

I'd like to conclude my remarks by mentioning what I view as possible alternatives to networks. The question which I am frequently asked and which I am trying to answer all the time is, "Why a network?" To me there are essentially four different types of alternatives: one, a stand-alone development, whereby we go on as we currently do in that each center strives to get bigger and better. Secondly, we have what I call ad hoc, bilateral arrangements. Then, we have centralized services. Or, we have the option of purchasing services from the market. And finally, we have the network. I would like to discuss very quickly some of the pros and cons for these various approaches. There is an advantage in stand-alone development. There is clearly a minimum of external influence on policy as to how computing is delivered. The disadvantage is that there is a replication of services, in a jurisdiction such as Ontario's. There is generally a surfeit of capacity. Also, economies of scale work against the small centers.

In terms of bilateral ad hoc arrangements, whereby University X negotiates with University Y for the exchange of services, there is the distinct advantage of no compulsion, but I submit that it won't really work because there is limited commitment to the principle of sharing. It's an idea of suboptimization, of suiting yourself and not really helping to solve the larger problem. Another disadvantage is that this kind of "ad hocery" would tend to favor large centers at the expense of the smaller ones.

With respect to centralized services, it's quite clear that economies of scale are realizable. From our own experience of trying this some years ago, it probably won't work, because it tends to reduce autonomy, which is much sought after. It also tends to ignore existing centers of expertise. One might ask, "Why don't we purchase services from the market?" I think there are some things to be said in favor of this. This approach does allow the university to meet demands incrementally. My studies indicate, however, that the economies are dubious as compared to in-house production. If, in fact, you can be assured of a market for the capacity of your system, it's clear that it is more economical to produce the service in-house.

Finally, the network has the advantage of bringing economies of scale and economies of specialization. There is, too, in our particular jurisdiction, the possibility of economies through deferred acquisition. Thus, if a university wishes or plans to go to the next larger system, it can, perhaps, defer that acquisition for a year. There's certainly money in that, from the point of view of the province as a whole. Also, the network brings service independent of distance. This is very important to us since we have a very unbalanced kind of population pattern in the province of Ontario. But most important, there is the potential of rationalization through sharing.

I would like to close by giving a couple of figures. We, in Ontario, are spending $118 per student, per year, on computing. My estimate for the operation of the network — this is independent of the initial development costs which I mentioned — was $3.6 million. The per student cost per year for network operations co[illegible] to $7.05, which is about six percent of the total expenditure per student per year. So, the question we must answer is: "Will the network be worth an extra six percent?"

QUESTION:

Why does it cost the student more money if you have a network? You said that it would cost six percent more.

MR. BROWN:

Yes. The seven dollars comes from communications costs, marketing costs, and central organization for the network. This is a cost over and above the current expenditure of $118 per student.

QUESTION:

What improvement in output do you expect from economies of scale?

MR. BROWN:

This is a dangerous question to answer. A preliminary working paper which we have produced suggests an annual rate of return of twenty-four percent. This is just a suggestion. We need to prove this by getting some good economic forecasters involved, and we hope to have some expert assistance from the Institute of Quantitative Analysis at the University of Toronto and, perhaps, verify twenty-four percent per year or something better.

QUESTION:

Is this network envisioned as an interconnection of what is roughly the existing centers now so that operating systems, and so forth, are determined locally?

MR. BROWN:

Yes. We are constrained, properly, by the principle of autonomy and we are working on the basis of minimum disruption of operating systems in the individual computing centers.

QUESTION:

When you say "we," are you speaking for a separate authority that exists independently of the authority of the separate institutions that are going to be part of this network? Are you a separate organization?

MR. BROWN:

The Council of Universities is a committee formed essentially of the presidents of the universities. We are funded by a levy on the individual university on a per-student basis. There is a secretariat which handles a number of different programs, one of which is Computer Coordination. So we, if you like, are representing the system of universities as a whole, not any particular university.

QUESTION:

You mentioned fourteen different universities that would be a part of the net, and then you speak of one of your economies as an economy of specialization. If each of those fourteen is at this point in the business of providing general purpose, general service computing, just what kind of specialization do you have in mind? I think it is a good idea, but in looking it over in our own case, we don't really see that one of these centers is specializing in something that another is not.

MR. BROWN:

The simplest example is APL. It is provided almost universally in each center and there is clearly a problem in terms of the resources required to support APL. I think a case could be made for providing APL, via a network, at one or two centers, making sure that there is first class support and excellent marketing for the product.

QUESTION:

How many students?

MR. BROWN:

We have an estimated 170,000 students in Ontario.

DR. HARRIS:

The topic of this panel is how to finance computer centers and computer networks. I have been involved with three. The first really wasn't a network because it was a star: a lot of terminals were connected to a central computer at Santa Barbara. It had some of the aspects of a network in terms of the interaction between people that is necessary to make any network work. This involved connections to eight universities around the country, ranging from large to small ones: the University of Washington, Illinois Institute of Technology, The University of Minnesota, the University of Missouri, The University of Pittsburgh, Georgia Tech, Louisiana State, and Florida State. The purpose of the connection was to explore the use of an interactive, on-line graphics system, known as the Culler-Fried system, which is resident at Santa Barbara, in the application toward chemical education. The problems that we experienced were mainly technical. We, at the time, were developing low-cost graphics consoles before they were available commercially, and if anybody has developed consoles, then you realize that it is a hard job. The other major problem that we had in terms of the time that it ran, which was about a year, was the man-to-man communication of the various participants among one another because of the long distances between each of the sites. In terms of the financing, that was fairly straightforward. The National Science Foundation simply paid the bills, with the exception of a modest contribution by each of the participants. The political consequences were rather interesting. Almost invariably, the individuals at each of the sites had to go through a long and involved hassle with the dean, the president, the computer center director, the digital computer committee or someone else before they would let a terminal be connected to an outside source of computing. I think that reflects, or is symptomatic of, the paranoia that every university or campus feels toward competition from the outside. One site actually was so disturbed by the prospects of computing coming in from outside that they quickly installed an on-line system and provided a user with a free terminal and all the free time he required. That was an interesting consequence.

The second network that I've had some experience with is the ARPANET. The ARPANET, as everyone probably knows by now, is a very large-scale, computer-to-computer communications network involving about twenty different nodes. Some of them are strictly military, some are universities. Some of the sites are designated as service sites, some are designated as user sites – some people call them parasites – but it's a network. It is somewhat separated into two pieces: there's a communications subnet which is very well defined and works well indeed. One can, in fact, communicate from one node to another node, independent of the distance, virtually a hundred percent error free. The network goes down very seldom.

The problems that one experiences are really problems associated with trying to understand somebody's system well enough to be able to use it. These problems are being solved. The communication costs are thirty cents a kilopacket transmitted from any point to any other point over the network. A packet is a maximum of approximately a thousand bits. Most cases would average near the maximum, which would be about thirty cents a megabit from point to point across the country. That corresponds to the amount of data you could transmit over a twenty-four hundred bit per second line for about seven minutes. That is quite economical in terms of straight communication costs. The motivation for building the network was to achieve savings in specialized computer services. ARPA has a number of sites which have special needs and, rather than duplicate services at different places, it was thought the network would provide savings on these very expensive specialized software systems. There were other motivations, of course, load sharing is one.

The third network which we are also concerned with is one that probably is similar, at least in terms of problems, to what people in other educational institutions have. This is a project which has National Science Foundation funding to look at the possibility, specifications and so on of a network for the University of California. There are several aspects to that: there is the political, the social, the technical, etc. The questions that are being asked are virtually the same questions that the panelist from Canada brought up; questions like tariffs, how to determine who gets what, systems of exchange — we seem to be in the fourteenth century with respect to computers and our systems of exchange. We still like to barter; we don't like to pay dollars for outside services. Thus, people are bartering among one another — you buy ten thousand dollars of computer time on my machine and I'll buy ten thousand dollars on yours.

The various individual identifiable aspects within the University of California are: academic computing, which is mostly teaching; research computing; administrative computing; hospitals and libraries. There are five areas and there's a task force which has the responsibility for looking into the various aspects of each one of those separate areas.

The University of California has nine campuses spread over seven hundred miles along the coast. There are several different kinds of machines: Berkeley has a CDC-6600; San Francisco, an IBM 360/50; Santa Cruz, a 360/40; Davis has a Burroughs 6500; Santa Barbara has a 360/75; UCLA has a 360/91; Riverside has a 360/50; Irvine has a couple of machines, a Sigma 7, a PDP-10; and San Diego has a Burroughs 6500 as well. So, there are different machines and different skills on campuses which are also different. Santa Barbara, UCLA and San Diego are currently on the ARPA Network. Berkeley is either on or will soon come on the ARPA Network. If one includes all those who have a telephone line to one place or another, and all the data communication costs within the university, we are spending approximately $200,000 a year. Some of that is coming out of individuals' pockets, some is administrative, and so on, but if we analyze that number and consider a network for the University of California, it appears that the annual operating costs for the network communications would be in proximity. Thus, we have strong motivation, from our point of view, to try to make something better.

There is a stronger motivation, however, which is manifested by something that happened recently. Governor Reagan signed a bill which was relatively unnoticed, but somewhat portrays what the future might be. That bill was to consolidate all State of California computing into four major computing centers, including the Department of Motor Vehicles, Department of Human Resources, welfare, budget, payrolls and so on. We are not included as part of it; neither the university nor the state college system, of which we are also independent, are included. However, I think that the time has come when the legislature, or the governor will require the universities (in California and other states) to do something about wasted computing. As a matter of fact, the state college system, which is under more direct control of the legislature, is presently feeling very strong pressure to consolidate their computing in several centers which, to my knowledge, have not been specified. The University of California is traditionally more independent of the legislature and of the governor because we have a regency system and the regents are fairly autonomous. Nevertheless, some of us think that the University of California will be instructed to do something soon. Our interest was, therefore, to examine building a network and possibly build it in our own way. We all have some experience in networks; we know what some of the problems and some of the goals are. Furthermore, if we don't build our own network, we may be very dissatisfied with whatever final network arrangements are made.

Now for some of the questions that have come up in the study. First is, "Will local computer center directors lose their jobs?" I think the answer to that is, "No." If one is on a network using someone else's processor, there is still a need for printers and all the I/O equipment; operators, consultants, and software people are also required. The software people may not be concerned with the next release of the operating system but they will certainly be considering network software. As an example, sixty percent of the computer center budget of Santa Barbara, which is about $900,000 a year, goes for non-hardware related elements, so forty percent is hardware. Perhaps the answer for people who worry about the loss of prestige or identity is to go into the business of building big empty boxes, painted red, and labeled "COMPUTER" on the outside. After installing them in the computer center and letting people walk through to see that they really exist, the computer center director can feel good about it.

Another question is, "What mechanisms should be used to govern a network?" I think the answer to that is some kind of marketplace economy, patterned more after the Common Market than after a totally free market. I think that some of those issues were alluded to earlier. The Common Market subsidizes the French farmer because France thinks it's good to have French farmers. Computer networks would like to take that sort of thing into consideration as well.

One area that greatly concerns me in terms of networks is the possibility that very large computers will come on, CDC 7600s or others, which are unrealistically priced in view of the way universities price theirs. Some of these offer extremely inexpensive time, as $100 an hour for 7600 time, or 91 time, or the equivalent. This situation would draw users from established centers and then, all at once, could dry up as the situation changes politically. But, I think as long as the computer centers that are involved in the network are run properly financially, and as long as everything is priced on a true cost basis, then networks will work.

Finally, the question that one needs to answer in terms of the user is whether or not a network can be friendly. Can you get the same friendly service you can get through your local computer center? I don't know the answer, but, if there are enough sites on a computer network, friendliness will result. The girl who rented me my car today or yesterday was very friendly. The reason she was friendly was because she wanted my $30 or $40. That same market economy works for computer centers as well. If they aren't friendly, the ones that have bad turn-around or bad service will eventually dry up. I think that is good; the user will ultimately benefit from that.

I have some other remarks about how I think computer centers ought to be financed but I'll pass those up for now.

QUESTION:

I'd like to know how Dr. Harris thinks computer centers ought to be financed.

DR. HARRIS:

I can tell you how I think they ought not to be financed. Traditionally computer centers obtained their money from somewhere else; the users did not provide the money. There are many where the administration gives the computer center two or three hundred thousand dollars and tells them to pass it out to the faculty. That's a very artificial way to distribute resources and computer centers inevitably get into trouble working on that basis. I believe that computing in the university ought to be financed the way everything else is financed: the department chairman finds out what the members of his department need. The chairman of the department submits a budget to the dean who considers it in terms of the total college budget. The dean, in turn, then submits his total college budget in competition with other deans. The president or the chancellor finally decides who gets the money. There should be a line item in the budget. Ideally, as a matter of fact, computing money ought to be able to be spent any way the director wants to spend it. He should not be captive to his local center. Now I'm not naive; I know that, in many cases, it has to be captive but that should not be the goal. If the money is captive and, if the computing center is not providing good service, then that money will not be spent. And, if after a couple of years the department chairman notices that he has to give $20,000 back to the administration because he couldn't spend the money, then he is going to start putting pressure on the administration to give him $20,000 he can spend on something else. So, if the funding is done through the user, the computer center will respond to the needs of that user. If it's given to the computer centers, they will do what they please with it.

QUESTION:

I've heard much about ARPA and the fact that there has been a very low level of use on the communications net. I don't know if this refers to the past or present. Is the present opening it up to a large use or is it just going through growing pains or something like that? I would like to get a picture of where ARPA is and where they're going.

DR. HARRIS:

There is isolated use; in other words, one place is using another place and, typically, only that place – or maybe a couple of places. There are some people in the audience who have had experience with the ARPANET as users. One gentleman is right there – Harvey Kriloff from the Chicago Circle Campus of the University of Illinois. I'm sure he'd be willing to say what his experiences are. There's also a man here from the MITRE Corporation who has used the Santa Barbara computing center over the ARPA Network. I think both or either of these gentlemen could answer the question better than I.

DR. KRILOFF:

Our experience has been, just simply as a user, to try and see what some of the capabilities of the networks are. We found, in many cases, it is extremely difficult, by the normal channels, to get information about what is available through the network. Most of the non-standard systems which are the reason for using the network were essentially undocumented and the only way of finding out about it was to track down the person who was responsible for it, get him on the telephone and spend half an hour talking to him and then go back to the terminal and iterate this process a number of times. The system itself is fairly well defined from a technical standpoint. From the software standpoint, there is still a whole range of wide open questions that much discussion is being done about but little full implementation; for example, the thing you see here with the MERIT network. They've been talking a lot longer about intercomputer communication and they don't have anywhere near the facilities. Even though MERIT is only starting this kind of intercomputer communications, there's a long way to go. ARPA started earlier and seems to have gone not quite as far.

QUESTION:

It isn't clear how far you have progressed in the University of California network. Is it operational? Do you have a timetable?

DR. HARRIS:

There is computer communication back and forth between campuses via the ARPANET but that is not the University of California network. On November 10, a group of people will get together in a hotel room for twenty-four or forty-eight hours straight to write a final report, compiling all the answers as they see them after almost a year-long study of what the political, social, and economic specifications ought to be for the University of California network. That will be a final report to the National Science Foundation and, as such, will be public information as well.

QUESTION:

Does the communications cost that you quoted earlier assume that the available lines are fully utilized?

DR. HARRIS:

No, that is the least cost; that's what you have to pay the telephone company to have them there all the time. Which one are you asking about, the thirty cents per kilobit? That's ARPA's number. That is what it costs, they figure, when the network gets loaded to the point where it's really a fully operating network.

QUESTION:

You said earlier that you can see the circumstances where the state would want to get together and account for the computer money. What do you see the relationship to be between state data processing and university data processing? What problems do you see there?

DR. HARRIS:

The State of California is sufficiently large so that there are no more economies of scale to be obtained by combining the Department of Motor Vehicles, the Highway Department, and others with the University of California. I think the State College System, with 200,000 or so students, the University of California system with 110,000 or 115,000 students, and the State itself, can all have independent data processing. Each of them is large enough so that there is really no motivation for marrying them. There is really no common interest. Now, maybe in a smaller state with a lesser population than twenty million people, it would make sense. The State of Oregon is, for example, looking at that.

DR. MAPLE:

I believe that all of you have a fairly accurate concept of an academic computer center; and to be sure that we are all thinking about the same thing when I mention the word "network," I will define a computer network to be a set of autonomous independent computer systems which are interconnected to permit sharing of resources among them. The resources to be shared are programs and data and the ability to access and execute these items on any one of the computers in the network. The fundamental concept here is the ability for each computer to make every local resource available to any computer in the network in such a manner that programs and data which are available to local users can be used remotely without degradation. Currently a computer center, with a few notable exceptions, is forced to recreate all of the application programs and data files which may have been established elsewhere. Quite often this is very expensive in both manpower and computer time and is very restrictive in the transferability of capabilities from one center to another. A successful computer network would reduce many of these problems of resource sharing and this is the goal of our network. The discussion which follows will give you some idea of some of the problems we have met, how we have dealt with them and some of the problems we expect to be facing in the future.

There are three state universities in the State of Iowa with a combined enrollment of 50,000 faced with the classic problem of expanding needs for computers in their academic, research and administrative areas with insufficient resources, both financial and human, to supply the required facilities. About four years ago discussions aimed at the solution of this problem led to the so called "cornfield" computer concept. Basically, this concept as originally envisioned, said that we should obtain a large computer and locate it somewhat centrally to the universities, and pool our financial and human resources to get the required computer capability from a single computation center, a la TUCC. Needless to say, with the recession and the accompanying down turn in the availability of financial support for the universities and their computation centers, these plans never got off the drawing board. However, since the University of Iowa is on a semester plan and Iowa State is on the Quarter plan, we saw the possibility of sharing the resources of these two centers to our mutual advantage. Consequently about two years ago we started discussions which led to a plan to tie our two IBM 360/65's together. Our initial plans were predicated on the thought that since our systems are so similar, we should be able to establish a connection between the two CPU's and ultimately be able to load level between these two systems. For well over a year now, we have had our systems linked together via a 4800 baud telpac line. In order to make this connection functional, there were certain compromises that had to be made in the systems area. To indicate the nature of these adjustments, I will give a few examples. Neither SUI nor ISU were using the standard generic names for classes of devices and naturally we had names different from each other. We had to agree on a set of standard names for all common devices. We also had to come to an agreement as to which space would be available for user scratch files at both locations and settle the question as to how the user would obtain the available user permanent file space.

In order that identical results be obtained on both machines, the PROCLIB's on both systems had to be made to agree and the ordering of DD statements had to be agreed upon in order that the overrides would work properly.

It is clear that with the numerous system changes that have been made over an extended period of time, the inclusion of additional computers into the network becomes more difficult and indeed impossible unless they are willing to conform to the conventions already in use in our systems. This point touches upon a political question concerning the use of computers in the state of Iowa. About two years ago, the State Legislature created a general services department whose duties among others included the "Establishing, supervising, and maintaining a system of centralized electronic data processing, including a data processing service center for the benefit of the state agencies in need of data processing services." At the time this legislation was passed, the Board of Regents institutions were exempt from inclusion under this act for reasons which relate to the extensive magnitude of the computer activities at the institutions, the fact that the equipment is being utilized essentially at capacity and the extensive cost of conversion of very complex software. The largest activity at the universities is associated with academic computer activity which requires a continuing program of software development. This in turn requires systems analysts at each educational institution in order to provide adequate services to academic personnel who are preparing their own applications programs. Thus, it did not appear practical to have the software people in a centralized group prepare programs for academic personnel because of the intimate knowledge of the particular academic discipline required. For example, specialized knowledge in mathematical programming is available at Iowa State and the social sciences archives are available at the University of Iowa.

The thrust to form a Board of Regents Computer Network was based in part on a desire to be able to share some of the unique capabilities which currently exist on our individual campuses as well as to find out if we can effectively level the loads over the two computer systems. I am also sure that there is an attitude on both centers, whether justified or not, that the university personnel better understand the computer techniques required to effectively serve the academic and research needs of our campuses. There is also a considerable interest on our campuses to provide help in the computer field to the several small colleges, junior colleges and high schools on a scale considerable larger than that now is being provided by the University of Iowa through their small college network; and it is thought that this could be more effectively done in a network containing more than a single resource node.

One of the questions which quickly arises in a multi-node network is concerned with the exchange charge rates to be used for the computing done at one resource node for the benefit of another. Usually, even though the computer systems at two nodes are practically identical, the local charging algorithms may be quite different due to the different local environments. We found ourselves in this situation and have developed an exchange rate which we consider very practical and worth mentioning. The computing charges at ISU are based on the systems management facilities records completely while those at SUI also use some records gathered by HASP and even though our charges are fairly comparable they are derived from different algorithms. Now one objective at both our centers is to be able to make the charges for a given program be as nearly reproducible as possible regardless of which computer it is run on. To accomplish this, we append the appropriate records to the output of each program when it is returned to originating node and these records are used to compute the charges according to the local algorithm. Thus a user on either campus is charged approximately the same amount for similar runs of the same program regardless of which computer is used. This leaves only the question as to how much one center charges another for services rendered. In each of our centers, personnel costs represent approximately fifty percent of our total expenses and we consider that approximately the same amount of personnel is contributed by each to support our network activities. It is probably also true that the level of our payrolls would not change if the link were not in existence. Thus we have agreed that each center will support its own personnel and not enter that expense into the exchange rate. We then average the local rates based upon this provision. As a practical expedient, our exchange rate is established by taking fifty percent of the average of the local rates of the two centers. Thus the wholesale rate between the two centers is somewhere in the vicinity of one half of a local rate. Under these circumstances, if one computer has some computing capacity available for sale to the other node, that capacity can be sold and at least some income will be realized instead of letting that resource go unused. From the viewpoint of the node using this excess capacity of the other node, this can provide the capacity needed to carry the campus through a peak load period without having to increase the local computing facilities. In order to encourage use of the link, our two centers have exchanged a credit of $20,000 for each to use at the other center. If this credit is not used during the fiscal year, it is lost. If one center uses more than $20,000 at the other facility, it pays for this usage at the established exchange rate.

This arrangement is most equitable when the two nodes involved are approximately the same size, but this algorithm clearly could be modified to fit other circumstances as long as the imbalance between the capacities and expenses of the nodes involved are not too severe.

The Computer Center at the University of Iowa has run an MVT – HASP system for a considerable time and ISU has run a straight OS/MVT system. This presented a problem in communications which was solved by ISU inserting a mini-HASP into our system so that we could use the HASP-to-HASP communications facilities originally developed by TUCC. This of course required that some of our systems personnel become thoroughly familiar with HASP in order to make the changes required for our environment, and required that our operations staff establish appropriate procedures to support the communications without degradation of our local OS system. The development of our communications link required that we establish a workable relationship among the systems personnel of the two centers, as equitable division of the necessary developmental and maintenance work, and common period of test time. This type of activity requires a large number of phone conversations and frequent joint meetings of sub-groups from the two centers. It is my opinion that our ability to coordinate activities was enhanced due to the fact that we are only 125 miles apart. Had we been 1500 miles apart, the required coordination would have been much more difficult.

With changes of the above nature being made it was very important that we establish some means of notifying the users of anticipated changes in order to minimize the inconvenience of any such change. The basic elements of user support are supplied in essentially three different ways. We are using our individual newsletters and bulletins to publish this type of information. This in turn is backed-up by consultative types of activities which involves diagnostic help for a knowledgeable user, information about what is available in our system and how to get at it, and basic support in the implementation and debugging of user programs. The third area of user support is in the provision of certain physical elements such as keypunches, disk space, catalogued procedures and the like.

In the near future we must make a decision as to how far we want to carry the integration of our individual centers' activities. We currently both have our own users' manual, newsletters and similar publications. It may be that we will want to merge these types of activities into network publication, or perhaps supplement them by additional publications.

At some time in the near future we need to decide upon a mutually agreeable division of the applications packages each center will support. Undoubtedly, certain applications packages will be supported at both centers whereas others may be of such a nature that we can afford to support them only at one site and require users from outlying campuses to use those facilities remotely.

DR. WEEG:

I have one question which I think is very significant. How much use have we got now after this past year of real hard effort?

DR. MAPLE:

Last month we used five hundred nineteen dollars worth of computing at the exchange rate at your place.

DR. WEEG:

I think that's a significant fact. Here we've got two guys who really want to make this work, we've got two systems which are nearly identical, and we've got all the wires set up. It's been going for a year and we know how to make it run; yet we're buying only five hundred dollars of time back and forth. Even if a thousand dollars were spent, think about that when you think about national nets.

QUESTION:

How much different are your charging rates?

DR. WEEG:

We're converging, I think.

DR. MAPLE:
Not radically different.

QUESTION:
Five percent?

DR. MAPLE:
Probably. We recently changed our charging by lowering our rates so we are almost down to the University of Iowa's rates.

DR. WEEG:
We're just changing our rates now in order to come into almost equivalence.

DR. HARRIS:
I'd like to make one other comment in terms of dollars. The University of California at San Diego with a Burroughs 6500 is on the ARPANET. I understand that they get twenty-five percent of their income from the network.

DR. WEEG:
I anticipate that's what's going to happen in any national net. Some nodes will become supernodes and other nodes will be total users, or parasites.

DR. HARRIS:
Bill Kehl at UCLA, I believe, draws about four hundred thousand dollars a year from the ARPANET.

QUESTION:
Is the rate for exchange of computer resources between the University of Iowa and Iowa State University defined by signed contract or strictly a gentleman's agreement between you two?

DR. MAPLE:
Whenever you deal with Gerry, you get a signed contract, which he signs in disappearing ink.

DR. WEEG:
What we really have is a series of memos back and forth. I suppose the most recent memo is the one which is the most binding.

DR. MAPLE:
Seriously, there are no formal contracts. However, there are statements of agreement. My boss, the vice president, for example, is aware that this is going on. I'm sure that Gerry's boss is aware of this too.

QUESTION:

Has there been any consideration given to the impact that selling time at wholesale rates might have on your government-sponsored research? You're not giving those rates to the government.

DR. MAPLE:

Yes. I depend very heavily upon the AEC for contract computing. It turns out that their rate is considerably lower than our network rate because they buy on a percentage basis, fifty percent of the computing, and pay accordingly. They've got a bargain still.

QUESTION:

Is there other government-sponsored research on your campus?

DR. WEEG:

Maybe for Clair but not for me. Our federal users pay the full price. We make the argument, and we've made the argument successfully stick with the auditors so far, that our users on our campus are paying the same price for equivalent services. If a federal user accesses my computer for one job he'll pay a price for it. If he uses Clair's computer for the same job, he'll still pay the same price for it. I conclude that he's not getting short shrift.

QUESTION:

It would be nice for us at Stanford to run a line up to that 7600 at Lawrence Berkeley Laboratory, get our computing done at one hundred dollars an hour, and sell it at our rates. Maybe we could swing the auditors into ignoring this situation as long as it's pretty much in doubt.

DR. WEEG:

Well, you've got to make a justification; we do make the justification. As a user of computer services I want my computer center director to buy me time the cheapest place he can buy it. Even if it means that they haul our machines away.

DR. MAPLE:

What do you think will happen when you buy fifty thousand dollars of computer time each month rather than five hundred dollars? Do you think the auditors will look at the problem in quite the same way?

DR. WEEG:

Certainly – you've got contracts to the tune of a half million dollars. I've got contracts to the tune of close to half a million dollars in federal business. The amounts are significant so they're looking hard. I've had week-long audits and they come away saying yes, it's all right.

DR. HARRIS:

What I suspect is that they're magnitude conscious.

DR. WEEG:

I understand that: that's right, but I'm saying there is already a significant magnitude, significant at least as far as the State of Iowa is concerned. Five hundred dollars goes a long way for us, buys lots of pigs.

QUESTION:

If the argument for going through this process is load sharing, do you visualize that there will be much individual use of the Iowa State machine? Will ISU do a lot of the Iowa work and vice versa?

DR. MAPLE:

Yes, I do. For example, SUI is moving in January. Gerry may be running quite a few programs in Ames during the time his machine is down moving to another location. During our end of quarter activities we're always very busy and we like to dump a lot of stuff over there on him, when that happens not to coincide with the end of his semester.

DR. KATELEY:

The memo that I have from Gerry Weeg says that I am supposed "to try to identify and address in an intelligent way some of the difficult political and financial questions that seem to be hindering the development of networks." He further says that this means "that we ought to discuss what kind of political, interpersonal, and financial problems have been faced and which remain to be faced." Well, I thought I would tell you about three things: first, about the history of MICIS and MERIT; second, about the early problems; and third, about some of the existing problems.

MICIS is a committee that started meeting early in 1966. The membership includes Wayne State, University of Michigan, and Michigan State University. Allan Smith, Vice President for Academic Affairs at the University of Michigan, has been at most every meeting; the Vice President of Research Development and Advanced Graduate Studies at Michigan State, Milton Muelder, has been at most every meeting, as has the Assistant Provost of Wayne State, a man by the name of Bob Hubbard. In addition a variety of faculty people have come and gone. However, there has been a core of people attending monthly MICIS meetings since 1966. The purpose of this committee is identical to the purpose of EDUCOM, and that is to use technology to help provide better education at lower cost. We have had a few problems with MICIS itself. For example, one of the worst arguments I was ever in at a meeting occurred at a MICIS meeting. This occurred at a time about two years after we had been meeting. If we had not had this history of meeting with one another, that probably would have been the end of our meetings. There have been pressures from the outside that have been just as intense that would have broken up this group had we not developed the rapport we have.

MERIT was incorporated in November of 1966 by the three universities for the purpose of accepting funds and distributi , f.nds for the three universities, again pretty much for the same purpose as MICIS. It is a non-profit organization incorporated in the State of Michigan. The members of the Board of Directors of MERIT are the same three individuals that I mentioned. In addition, MERIT employs Bert Herzog as Director. Bert and his staff are located in Ann Arbor. Their most immediate purpose is to direct and promote the use of the MERIT network. The MERIT network activities got started about July, 1968, and has used 1.5 million dollars since then. The sources were NSF, the State of Michigan, and the three institutions, the last mainly contributed services. The monies that were available were expended between the period July 1969 – June 1972 to develop the network.

Some of the early problems I had in connection with the development of the network I will describe next. I went to many, many meetings having to do with the selection of a communications computer, this computer to be the same at each of the three institutions. People from the three institutions had various technical requirements of this computer that made it impossible to go out and buy a computer that would do the job. This may sound like a technical problem; I don't think it is. I did not understand why the technical requirements were such as they were, but each institution was allowed to say what they wanted in a communications computer. We had no mechanism for telling some one institution they were overspecifying the interface, for example.

The second problem I had was that Wayne State was made the fiscal agent. I understand the bookkeeping system at Michigan State reasonably well. The bookkeeping systems at Wayne and Michigan are just enough different that I do not understand their systems. As a result, I often did not understand what the fiscal agent was doing. I had the same sort of problem with Bert Herzog, that is, not understanding his budget and reporting processes because they were, of course, designed along the lines of the University of Michigan system.

There was and still is a serious problem of jealousies among the staffs of the three institutions. My people objected to many of the decisions that were made by the MERIT staff as undermining them or shortchanging them, and it was my job to try to explain to them why we had to do this. This still is an unsatisfactory situation.

There was and still is within MICIS and the related MERIT activity an attempt to make all things equal and fair among the three institutions to the point where many of us do not say some of the things that we should in our meetings.

At the present time, a number of problems exist with respect to the network. The main problem is the uncertainty as to the use of the network as a budget matter. Or, where does the money come from to use the network? A related problem is that of unequal flows. I think that the flows will be unequal, and furthermore I am inclined to think the flow is going to be in dollars away from the MSU system. Another related problem is that I have to support the use of the network on the MSU campus. For example, I provide consultants and systems programmers for the use of the MSU system. We hold workshops to try and teach people how to use the MSU system. We hold seminars for the same purpose. We publish a variety of memos and other documents to explain to users how to use the MSU system. The same sort of support activity *is* required on the MSU campus for the use of the network and does not exist. A related problem has to do with credits. The MSU consultants serve three purposes. One is to help the user, the second is to collect information about bugs in the system, and the third is to give users a dollar credit when they have a run that is bad because it is the fault of the MSU system. I won't give credits for bugs in the network software or for problems in the system at the University of Michigan or Wayne State that one of my users encounters. The problem of unequal flows bothers me a lot because I cannot afford, out of the money I have available to me, to properly support network use.

The MERIT network has had an impact on the procedures that we follow very faithfully with respect to the integration of new software into the MSU system, and with respect to operational procedures for putting new software into production use. Installation of network software into our system has required us to temporarily ignore some of our procedures. That's both good and bad. In making this exception to our procedures for MERIT we have reexamined those procedures. It takes a very long time for us to make changes to the MSU operating system and to do the notification and reporting that is involved. Well, the result of MERIT's pressure on me was that we thought back and remembered the reasons we do things the way we do and have decided that we will keep doing them that way.

I would like to take a minute to explain the way that I think a user on my campus will see the MERIT network service. Each college at MSU is required to pay for twelve and a half percent of their total computer use. This twelve and a half percent comes from the college, and it comes from whatever sources they have available; for example, it could come from supplies and services accounts. They could buy computer service instead of buying pencils. It could also come from research grants. We do not know where it comes from, but the point is we get twelve and a half percent from each college. We have been using this scheme for three or four years. Well, the result is that the users of Michigan State do not expect to pay a hundred percent for anything and in particular, they do not expect to pay a hundred percent for the use of the network. I am not willing to use the money I have to subsidize the use of the network. So we have worked out the following arrangement: the flow in, in dollars, to Michigan State will be used to subsidize the use of the network. And we will watch the unbalance in flow: when I am in the hole fifteen thousand dollars, we will change the arrangement I am about to describe. The user at Michigan State on a college by college basis will pay for twenty-five percent of his network use. In other words, seventy-five percent of a college's network use is subsidized. If I see that the deficit that I am running with the network runs to five thousand dollars, the pay required will go to fifty percent. And, of course, when it reaches fifteen thousand dollars, the pay that the user will have to provide will be one hundred percent – no subsidy for use of the network. I hope that the flows are equal, of course; if the flows are equal, I will have an additional income item of twenty-five percent of the flow. I would argue that that would about pay for my cost of providing service to the network.

Let me close with the following, and unfortunately Clair took my line here. The question is cooperation or conflict, and the answer is, of course, cooperation *and* conflict. The conflict is exquisite, and I have great expectations for the cooperation. I have done things for MERIT that from my point of view are wrong, for example, things that have the effect of offending my people. And this is a serious problem for me. On the other hand, the challenge presented to us by Allan in his talk at lunch yesterday is clear. It is my responsibility, I suppose, to try to convey this challenge to my people and tell them about the grand future for the networks and how we are going to save a lot of money. They are very skeptical about that. I suppose I do not do a good job of delivering that message.

DR. KRUEGER:

Institutions of higher education are established and maintained for the preservation, creation and dissemination of knowledge. In achieving these goals, computers have been introduced as a service entity to support the programs of *instruction*, *research* and *administration*. The institutions have also provided computing services to the public.

Computers have been utilized to meet the identified goals in institutions of higher education in Colorado since 1962. In 1969, CHESS (Colorado Higher Education System Sharing) was instituted as a concept to promote sharing of computing system resources among institutions of higher education. Currently, eighteen computers are operating to provide centralized services to their institutions. At the time of implementation of the CHESS concept, service was oriented around a decentralized type of network.

Recently, networking has resulted in sharing of applications software by the instruction and research programs. Administrative resource sharing has also been promoted by restricted hardware networking. Software sharing, which has been limited in the adminstrative area, is also now slowly progressing.

From a hardware networking point of view, there are currently twenty-three remote batch terminals utilized in two star networks operational in the state. This batch service is provided by Colorado State University to its two campuses and to the community colleges utilizing the COTIE network. The University of Colorado provides academic computer services to its four campuses and to Metropolitan State College. In addition, the university provides resources, in the best interests of the state, to serve the Colorado Division of Highways' scientific computer needs. These networks are operated as auxiliary enterprises, charging real dollars for all services.[1] Interactive network service is provided by the University of Colorado and the Colorado School of Mines, and again, all usage is charged on a real dollar basis. In March of 1972 a new computing center was created, the Denver Regional Computing Center (DRC), which is managed by the Colorado Commission on Higher Education, the agency responsible for providing overall state coordination within the institutions of higher education. Following establishment of the DRC, network evaluation activities have increased substantially, looking toward creating a computer to computer network, somewhat like MERIT, involving Colorado State University, the University of Colorado, and the Denver Regional Center.

During the current fiscal year, it is estimated that enrollment in higher education in the state will be approximately one hundred thousand. Of these, approximately sixty thousand students are registered in one of the three universities. Based on data provided by the institutions, actual class headcount of students using computer resources in fiscal 1972 was 48,155. Planning for fiscal year 1973, based on instructional programs justification, calls for a student user class headcount of 66,015. State support of ADP during this period is just over four million dollars; and I might add that there was no increase for this current fiscal year from the preceding year. Approximately three million dollars, or thirty dollars per enrolled student, of this sum is expended in administrative ADP support. The balance is available for instruction, that is ten dollars per student enrolled or, if you want to count users, fifty dollars per student user.

Coordination and control of ADP in higher education in Colorado is vested, based on legislation, with two state agencies. The Colorado Commission on Higher Education is responsible for coordination of activities of all institutions. ADP funds for the institutions are appropriated directly to the Commission. The second agency involved is the Department of Administration, which is responsible for control of all state ADP, including that in higher education.

Through institutional cooperation resulting from funding constraints and, of course, encouragement from the cognizant state agencies, resource sharing is continually increasing. Currently, a Commission appointed committee of institutional representatives is involved in evaluating and coordinating institutional ADP and associated FY74 budget requests. Evolving from this activity will be a state-wide ADP resource sharing plan for higher education. By "resource sharing" we think in terms of systems. Resource sharing involves hardware, software, and people networks. These plans address the needs of programs supported by ADP services within the institutions as well as services provided in the public interest. Of particular interest relative to this latter point is a proposed cooperative project between the Commission on Higher Education and the Department of Education to develop the so-called HITIE computer to computer network to provide services to state high schools. Institutions of higher education that will be involved in this project include Colorado State University and the University of Colorado. Of course, this will interconnect to the state-wide higher education network mentioned earlier. Although the state-wide plan is not yet complete, the following reflect the general direction of the resource sharing planning to date. I will list these as objectives.

In meeting the continuing institutional goals, the unifying concept of CHESS is endorsed; that is, the sharing of institutional ADP resources including personnel, hardware and software. In supporting the general concept of resource sharing, specific objectives have been established, identified by program served. It should be noted that these goals influence both short- and long-range planning.

Instruction. Each student should have the opportunity to obtain a general understanding of computers, their organization, applications, limitations, and impact on society. Each student should have the opportunity to study the computer and apply the computer in the manner and to the extent pertinent to his course of study. Adequate computer resources (hardware, software, personnel and instructional programs) should be made available to meet the first two objectives and to maintain the level of education commensurate with the standards existing nationally, utilizing netw rks to achieve this objective as necessary. Student and faculty members should have access to computer resources appropriate to his course of study and discipline, regardless of his institutional affiliation.

Research. Use of the computer as a research tool to create new knowledge is a necessary function of higher education. Faculty should have access to computer resources appropriate to his research area.

Administration. Through coordinated institutional activity, administrative data processing should meet state-wide needs as follows:

- Continue and improve support of operations and management function.
- Continue and expand development of data base oriented planning models, e.g., CAMPUS.
- Continue to evaluate and implement existing applications both within the state and available from other sources including CAUSE and WICHE-NCHEMS.
- Continue development of standard data elements.
- Establish evaluation process for implementing new applications.

Public Service. Cooperative efforts in ADP with federal, state, and local agencies should be encouraged so as to supplement each institution's educational activities and be of direct benefit to the citizens of the state. The HITIE project, which I referred to earlier, is specifically endorsed. Cooperative efforts in ADP with industry should be encouraged when oriented toward benefiting the institutional goals in instruction, research and administration.

In summary, state-wide ADP network planning in Colorado is currently based on existing capabilities, funding limitations and controls. Resource sharing has progressed to provide computer access when economically feasible. Usage of the networks has been on a real dollar charge basis with overall control vested in the Colorado Commission on Higher Education. Completion of the state-wide computer to computer network plan will most likely recognize and continue these attributes.

REFERENCE

[1] Computing in Higher Education 1971: Successes and Prospects (EDUCOM, Princeton, N.J.) 1972.

supplements

CONFERENCE PARTICIPANTS

Leon Adkison
Assistant Professor of Information Systems
Virginia Commonwealth University
901 West Franklin Street
Richmond, Virginia 23220

John P. Allen
Chief, Technical Services
Washington State University Library
120 Holland
Pullman, Washington 99163

Ernest J. Anastasio
Assistant Director (ODAR)
Educational Testing Service
Princeton, New Jersey 08540

Ronald Anderson
University of Minnesota
Minneapolis, Minnesota 55455

Rosemarie Areten
EDUCOM
Princeton New Jersey 08540

Gene Arnold
Administrative Dean
Delta College
University Center, Michigan 48710

William F. Atchison
Director, Computer Science Center
University of Maryland
College Park, Maryland 20792

Robert L. Ashenhurst
Director, Institute for Computer Research
University of Chicago
5640 South Ellis Avenue
Chicago, Illinois 60637

Mary Jill Ault
Center for Research on Learning and Teaching
109 East Madison Street
University of Michigan
Ann Arbor, Michigan 48104

Eric Aupperle
Project Leader
MERIT Computer Network Project
University of Michigan
Ann Arbor, Michigan 48104

George Badger Jr.
Associate Director, Computer Center
University of Pittsburgh
Pittsburgh, Pennsylvania 15213

F. W. Bakker
Professor
Michigan State University
East Lansing, Michigan 48823

W. P. Barley
IBM Corporation
10401 Fernwood Road
Bethesda, Maryland 20034

Michael Barnett
Director, Research and Development
The H. W. Wilson Company
950 University Avenue
Bronx, New York 10452

James J. Batter
Director
University of North Carolina
Computation Center
39 Phillips Hall
Chapel Hill, North Carolina 27514

John H. Beaman
Professor
Michigan State University
East Lansing, Michigan 48823

Virginia Beauchamp
Michigan State University
East Lansing, Michigan 48823

Stoughton Bell
Director, Computing Center
University of New Mexico
Albuquerque, New Mexico 87106

Sanford Berg
Economics Department
University of Florida
Gainesville, Florida 32601

D. L. Bernitt
Pennsylvania State University
105 Computer Building
University Park, Pennsylvania 16802

R. E. Bishop

Robert Blanc
Technology B216
National Bureau of Standards
Washington, D.C. 20234

Ronald Blum
Director, CONDUIT
Duke University
Durham, North Carolina 27706

Lorraine Borman
Manager of Information Services
Northwestern University
2129 Sheridan Road
Evanston, Illinois 60201

Maurice P. Brown
Director, Office of Computer
Coordination
Council of Ontario Universities
102 Bloor Street West
Toronto 181, Ontario, Canada

Russell W. Burris
Associate Professor
University of Minnesota
400 Ford Hall
Minneapolis, Minnesota 55455

W. Richard Burrows
Associate Director
Administrative Systems and Computation
Services
University of New Hampshire
Room 201, Thompson Hall
Durham, New Hampshire 03824

Coleman H. Burton
Associate Director, Computer Center
University of Missouri
Columbia, Missouri 65201

Peter Calingaert
Professor
University of North Carolina
New West 215D
Chapel Hill, North Carolina 27514

Hardy Carroll
Assistant Professor
School of Librarianship
Western Michigan University
Kalamazoo, Michigan 49001

Henry Chauncey
President
EDUCOM
Princeton, New Jersey 08540

Robert Chenhall
Coordinating Office
University of Arkansas Museum
Fayetteville, Arkansas 72701

John W. Childs
Assistant Dean College of
Education
Wayne State University
Detroit, Michigan 48202

John A. Clark
Acting Director
Office of Computer Planning
Department of Higher Education
225 West State Street
Trenton, New Jersey 08625

Nancy Clark
Project Manager
Virginia Polytechnic Institute
and State University
Computing Center
Blacksburg, Virginia 24061

Ronald Code
Stanford University
Stanford, California 94305

Mary Frances Collins
Assistant Director
University Library
State University of New York at Alban
Albany, New York 12203

Ronald W. Collins
Chemical Department
Eastern Michigan University
Ypsilanti, Michigan 48197

Jack Cooper
Director, Computer Services
University of South Carolina
Columbia, South Carolina 29208

Nancy Covert
Oakland University
Rochester, Michigan 48063

Ronald D. Crain
Visiting Lecturer
University of Kansas
Department of Chemistry
Lawrence, Kansas 66044

Thomas A. Croil
T. A. Croil Associates Ltd.
54 Daneswood Road
Toronto, Ontario, Canada

John Daniels
Operations Manager
Shared Educational Computer Systems Inc.
P. O. Box 204
Poughkeepsie, New York 12602

Alvin L. Dasen
Staff Analyst
State of Michigan House Fiscal Agency
Room 119, Capitol Building
Lansing, Michigan 48902

Robert Davies
Ithaca College
Ithaca, New York 14850

James Davis
Director, NORC
University of Chicago
6030 South Ellis Avenue
Chicago, Illinois 60637

Ralph Deal
Professor of Chemistry
Kalamazoo College
Kalamazoo, Michigan 49001

Joseph Denk
Curriculum Coordinator
North Carolina Educational Computing Service
Box 12175
Research Triangle Park, North Carolina 27709

Gerald W. Dennis
AVTECH
Purdue University
Lafayette, Indiana 47907

Dixon R. Doll

Robin N. Downes
Assistant Director
University of Michigan Library
Ann Arbor, Michigan 48104

Joselyn Druschel
Assistant Chief Technical Services
Washington State University
120 Holland
Pullman, Washington 99163

Howard Duffield
MITRE Corporation
McLean, Virginia 22101

Brenda Dunlap
EDUCOM
Princeton, New Jersey 08540

Catherine Dunnagan
CONDUIT
Duke University
Durham, North Carolina 27706

William B. Durant, Jr.
Executive Officer for the Faculty of Arts and Sciences
Dartmouth College
201 Wentworth Hall
Hanover, New Hampshire 03755

Phil Dylhoff
Telecommunications Manager
Computer Center
University of Iowa
Iowa City, Iowa 52240

Carol Easthope
Systems Analyst
LMIS Project
MERIT Computer Network
University of Michigan
Ann Arbor, Michigan 48104

Harry A. Eick
Professor of Chemistry
Michigan State University
East Lansing, Michigan 48823

Mark Eisner
National Bureau of Economic Research
545 Technology Square
Cambridge, Massachusetts 02139

Robert Enriss
Budget Analyst
State of Michigan
Executive Office
Bureau of Programs and Budget
Lansing, Michigan 48902

Keith Evans
Director of Planning and Institutional Research
Portland State University
P. O. Box 751
Portland, Oregon 97207

David Freeman
Rutgers University
New Brunswick, New Jersey 08903

Joyce Friedman
Professor
University of Michigan
2076 Frieze Building
Ann Arbor, Michigan 48104

Martell Gee
University of Dayton
Dayton, Ohio 45409

Ray Geitka
Oakland University
Rochester, Michigan 48063

Robert J. Gettelfinger
Director, Computer Center
Educational Testing Service
Princeton, New Jersey 08540

Edvard Glaaser
Research Economist
Denver Research Institute
Industrial Economics Division
2135 East Wesley
Denver, Colorado 80210

Edmond Goings, Director
Instructional Computing Center
Eastern Michigan University
Ypsilanti, Michigan 48197

James M. Gradolph
Staff Associate
Illinois Educational Consortium for Computer Services
160 ERL
University of Illinois
Urbana, Illinois 61801

E. E. Graziano
Assistant Librarian
University of California
Santa Barbara, California 93106

Martin Greenberger
Professor of Computer Science
The Johns Hopkins University
34th and Charles Street
Baltimore, Maryland 21218

Stuart Grout
Director of Academic Planning
Boston University
147 Bay State Road
Boston, Massachusetts 02215

Richard E. Grove
Director, University Computer Center
Virginia Commonwealth University
1015 Floyd Avenue, Box 174
Richmond, Virginia 23220

Dale J. Hall
Director, Computer Systems Development
Indiana University
Computers and Information Systems
317 East Second Street
Bloomington, Indiana 47401

Michael A. Hall
Director of Computer Services
Lawrence University
Appleton, Wisconsin 54911

Randall Hansen
Manager, Education Industry
Control Data Corporation
Box O
Minneapolis, Minnesota 55440

Rodney B. Harrington
Animal Science Department
Purdue University
Lafayette, Indiana 47907

David O. Harris
Acting Director
Computer Systems Laboratory
University of California
Santa Barbara, California 93106

Paul R. Hart
University of Toledo
Computer Services
Toledo, Ohio 43606

E. C. Hertzler
Biology
University of Michigan
Dearborn, Michigan 48120

Bertram Herzog
Director
MERIT Computer Network
University of Michigan
Ann Arbor, Michigan 48104

Robert C. Heterick, Jr.
Director
Virginia Polytechnic Institute
and State University
Computing Center
Blacksburg, Virginia 24061

Clayton Highum
Director of Libraries
Illinois Wesleyan University
Bloomington, Illinois 61701

Laurence G. Hoard
The University of Michigan
Ann Arbor, Michigan 48104

Richard I. Hofferbert
Executive Director
Inter-University Consortium for
Political Research
P. O. Box 1248
Ann Arbor, Michigan 48106

Michael Holcombe
Delta College
University Center, Michigan 48710

Mary W. Hook
Senior Staff Economist
Council of Economic Advisors
Old Executive Office Building
Washington, D. C. 20506

David Howard
Delta College
University Center, Michigan 48710

Helen Ianni
Wayne State University
5524 Williamson
Dearborn, Michigan 48126

John Iseli
MITRE Corporation
McLean, Virginia 22101

Frederick H. Jackson
Director, Committee on
Institutional Cooperation
1603 Orrington Avenue
Evanston, Illinois 60201

W. Carl Jackson
Dean of University Libraries
130 Bryan Hall
Indiana University
Bloomington, Indiana 47405

David P. Jasper
Consultant for Business Development
Control Data Corporation
8100 34th Street South
Minneapolis, Minnesota 55440

Lawrence L. Jesky
Director of Computer Programming Courses
Seton Hall College
Greensburg, Pennsylvania 15601

James W. Johnson
Curriculum Development Manager
University of Iowa
Iowa City, Iowa 52240

Jeremy E. Johnson
Computing and Data Processing Services
University of Maine
Orono, Maine 04473

Whitney L. Johnson
Director of Management Information
Services
Northern Michigan University
Marquette, Michigan 49855

T. S. Johnston
Director, Computer Center
Michigan Technological University
Houghton, Michigan 49931

Julian [illegible]teley
Director, Computer Laboratory
Michigan State University
East Lansing, Michigan 48823

Donald L. Katz
University of Michigan
Chemical Engineering
Ann Arbor, Michigan 48104

H. Eugene Kessler
Executive Director and Treasurer
EDUCOM
Princeton, New Jersey 08540

Frederick Kilgour
Director, Ohio College Library
Center
1314 Kinnear Road
Columbus, Ohio 43210

Ethel B. Kinney
EDUCOM
Princeton, New Jersey 08540

John C. Klensin
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Joretta Klepfer
Faculty Programmer Consultant
University of North Carolina at
Greensboro
Computer Center
Greensboro, North Carolina 27408

John G. Kolp
Social Science Curriculum Coordinator
University of Iowa
Iowa City, Iowa 52240

Robert R. Korfhage
Director and Professor
Computer Science/Operations
Research Center
Southern Methodist University
Dallas, Texas 75222

Harvey Krilott
Assistant Professor
University of Illinois
Computer Center
Chicago, Illinois 60680

John Kronebusch
University of Pittsburgh
5526 Howe Street
Pittsburgh, Pennsylvania 15232

E. Rex Krueger
Director, Computer Center
University of Colorado
Boulder, Colorado 80302

Thomas E. Kurtz, Director
Kiewit Computation Center
Dartmouth College
Hanover, New Hampshire 03755

Warren Lackstrom
Computer Center
National Bureau of Research
Cambridge, Massachusetts 02139

Carolyn P. Landis
Secretary
EDUCOM
Princeton, New Jersey 08540

Dave L. Landom
Computer Center Director
Black Hills State College
Spearfish, South Dakota 57783

Alfred J. Lata
Lecturer
University of Kansas
Department of Chemistry
Lawrence, Kansas 66044

Kurt Lauckner
Associate Professor of
Mathematics
Eastern Michigan University
Ypsilanti, Michigan 48197

Hans E. Lee
Department of Computer Science
Michigan State University
East Lansing, Michigan 48823

Sheldon G. Levy
Associate Professor Psychology
Wayne State University
Detroit, Michigan 48105

Miles A. Libbey, Director
Research Center for Library and
Information Science
Graduate Library School
Indiana University
Bloomington, Indiana 47401

Peter Lykos
Office of Computing Activities
National Science Foundation
1800 G Street, N.W.
Washington, D. C. 20550

Ron McAdams
University of Dayton
Dayton, Ohio 45409

David McBlain
Austin College
Sherman, Texas 75090

Davis B. McCarn
Acting Associate Director for
Science Communications and
Computer Engineering Service
National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 20014

Donald H. McClain
Assistant for CAI
University of Iowa
Iowa City, Iowa 52240

Kenneth McCool
Pennsylvania State University
University Park, Pennsylvania 16802

Forrest McGill
708 Pauline No. 2
University of Michigan
Ann Arbor, Michigan 48104

James McKenney
Professor of Business Administration
Harvard Graduate School of Business
Administration
Soldiers Field
Boston, Massachusetts 02163

D. S. Macey
Office of Computer Coordination
Council of Ontario Universities
102 Bloor Street West
Toronto 181, Ontario, Canada

Dana Main
Department of Psychology
University of Michigan
Ann Arbor, Michigan 48104

Clair Maple
Director of Computer Center
Iowa State University
Ames, Iowa 50010

Dille , E. Mapother
Director
Computing Services Office
University of Illinois
Urbana, Illinois 61801

Gregory Marks, Head
Computer Support, Center for
Political Studies
Institute for Social Research
University of Michigan
Ann Arbor, Michigan 48104

Joseph E. Matar
Associate Provost
Marquette University
615 North 11th Street, Room 201
Milwaukee, Wisconsin 53233

Ignatius G. Mattingly
Professor of Linguistics
University of Connecticut
Haskins Labs
270 Crown Street
New Haven, Connecticut 06511

Jack R. Meagher, Director
Computer Center
Western Michigan University
Kalamazoo, Michigan 49001

John P. Menard
Associate Director
Computer Science Center
University of Maryland
College Park, Maryland 20742

Thomas E. Michels
Assistant to the Director for
Technical Affairs
University of Nebraska Computer
Network
419 Administration Building
Lincoln, Nebraska 68508

E. P. Miles, Jr.
Professor of Mathematics
Florida State University
Tallahassee, Florida 32306

Ronald F. Miller
Director, NELINET
New England Board of Higher
Education
40 Grove Street
Wellesley, Massachusetts 02181

Gordon W. Mills
Education Specialist
Burroughs Corporation
Burroughs Place
Detroit, Michigan 48232

Robert A. Mills
Assistant Professor of Mathematics
Eastern Michigan University
Ypsilanti, Michigan 48197

Don Mittleman
Director, Computer Center
Oberlin College
Oberlin, Ohio 44074

K. Leon Montgomery
Assistant Professor
IDIS
Room 408, LIS Building
University of Pittsburgh
Pittsburgh, Pennsylvania 15213

Roland H. Moody
Director of Libraries
Northeastern University
360 Huntington Avenue
Boston, Massachusetts 02115

Michael Moreno
AVTECH
Purdue University
Lafayette, Indiana 47907

Gilbert G. Moser
Educational Information Services, Inc.
Rutgers University
Hill Center for the Mathematical
Sciences
New Brunswick, New Jersey 08903

Wayne A. Muth
Director, Computer Center
West Virginia University
Morgantown, West Virginia 26506

Albrecht J. Neumann
Systems Development Division
National Bureau of Standards
Building 225, Room A264
Gaithersburg, Maryland 20760

Gordon L. Nordby
University of Michigan
Biological Chemistry
Medical School
Ann Arbor, Michigan 48104

C. E. Nordman
Department of Chemistry
University of Michigan
Ann Arbor, Michigan 48104

Stanley Norris
Assistant to the Vice President
The University of Calgary
Calgary, Alberta, Canada

Edward J. O'Donnell
Chief Scientific Applications Branch
Director, Management Information Systems
Army Materiel Command
Washington, D.C. 20350

K. Okashimo
Director
Institute of Computer Science
University of Guelph
Guelph, Ontario, Canada

Eric V. Ottervik
Vice Provost
Lehigh University
Bethlehem, Pennsylvania 18015

Edwin B. Parker
Institute for Communication Research
Stanford University
Stanford, California 94305

Ed Paulette
MERIT Computer Network
University of Michigan
109 East Madison
Ann Arbor, Michigan 48104

Ben Paulson
Director of Data Processing
Delta College
University Center, Michigan 48710

Peter J. Paulson
Director
New York State Library
State Education Department
Albany, New York 12224

Philip I. Pavlik
Associate Professor Chemistry
Northern Michigan University
Marquette, Michigan 49855

Glenn W. Peardon
Head, University Systems
University of Saskatchewan
Room 3022
Saskatoon, Saskatchewan, Canada

Robert R. Pearson
Director of Computer Services
Marquette University
615 North Eleventh Street
Milwaukee, Wisconsin 53233

W. David Penniman
Associate Chief, Basis-70
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201

Daniel R. Pfoutz
Head, Science and Technology Department
Carnegie Library of Pittsburgh
4400 Forbes Avenue
Pittsburgh, Pennsylvania 15213

Doug Phaneuf
Associate Director, Computing Center
University of Michigan
Flint, Michigan 48503

Jesse H. Poore, Jr.
Director, Computing Center
Florida State University
Tallahassee, Florida 32306

James R. Potter
Interim Executive [illegible]
Indiana Higher Ed[illegible]
Telecommunica[illegible]
Indiana University
1100 West Michigan Street
Indianapolis, Indiana 46202

Charles B. Ray
Director
California Institute of Technology
Booth Computing Center
Pasadena, California 91109

Jerry L. Ray
Academic Coordinator
University of Nebraska Computer Network
Omaha Facility
P. O. Box 688
Omaha, Nebraska 68101

Gene Raymer
Wayne State University
Detroit, Michigan 48202

Martin D. Robbins
Senior Research Economist
Denver Research Institute
Industrial Economics Division
2135 East Wesley
Denver, Colorado 80210

Louis P. A. Robichaud
Directeur, Centre de Traitement
de l'Information
Universite Laval
Laval, Quebec, Canada

Stewart Robinovitz
MERIT Computer Network
University of Michigan
Ann Arbor, Michigan 48104

Robert J. Robinson
Director, Computing Center
State University of New York
at Albany
1400 Washington Avenue
Albany, New York 12222

George H. Roehm
Management Systems Executive
Management Sciences Group
Executive Office of the Governor
Lewis Cass Building, 1st Floor
Lansing, Michigan 48913

Monique R. Rona
Assistant Director, Computer Center
University of Washington
160 Computer Center
Seattle, Washington 98195

Leo Roomets
Coordinator Computer Services
Massachusetts State College System
65 Franklin Street
Boston, Massachusetts 02110

Mitch Roth
Michigan State University
East Lansing, Michigan 48823

Howard Rovelstad
Director of Libraries
University of Maryland
College Park, Maryland 20742

Harry Rowell
Associate Director
Computation Center
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Martin Rubin
Staff Scientist
Human Resources Research Organization
300 North Washington
Alexandria, Virginia 22314

Walter G. Rudd
Assistant Professor
Computer Science Department, LSUBR
205 Nicholson Hall
Louisiana State University
Baton Rouge, Louisiana 70803

Frank J. Ruggiero
Department Head
Data Base Systems
Educational Testing Service
Princeton, New Jersey 08540

George Sadowski
The Urban Institute
2100 M Street, N.W.
Washington, D. C. 20037

Gerald St. Amand, Instructor
Michigan State University
College of Business
418 Eppley Center
East Lansing, Michigan 48823

Frances Saupe

Regina Scarborough
EDUCOM
Princeton, New Jersey 08540

Hal B. Schell
Associate Director of University
Libraries
Southern Methodist University
Dallas, Texas 75222

Ed Schneider
Department of Political Science
Interuniversity Consortium for
Political Research
University of Michigan
Ann Arbor, Michigan 48104

Sally Y. Sedelow
Professor, Computer Science and
Linguistics
The University of Kansas
Lawrence, Kansas 66044

Walter A. Sedelow, Jr.
Department of Computer Science
1 Strong Hall
The University of Kansas
Lawrence, Kansas 66044

R. G. Selfridge
Director, Computer Center
University of Florida
Gainesville, Florida 32601

Harold T. Shapiro
Professor of Economics
University of Michigan
Ann Arbor, Michigan 48104

Harvey A. Shapiro
Oakland University
School of Economics and Management
Rochester, Michigan 48063

Charles Shomper
Regional Computer Center Director
University of Iowa
Iowa City, Iowa 52240

Harrison Shull
Vice Chancellor for Research and Development
Indiana University
Memorial Hall East
Bloomington, Indiana 47401

Donald R. Shurtleff
Director of Computing
University of Missouri
Columbia, Missouri 65201

F. D. Simpkin
Director, Computer Services
York University
4700 Keele Street
Downsview, Ontario, Canada

Robert R. Simpson
Systems Coordinator
The University of Michigan
3071 Administration Building
Ann Arbor, Michigan 48104

S. P. Singh
Department of Chemical Engineering
The University of Michigan
Ann Arbor, Michigan 48104

John E. Skelton
Research Associate, Mathematics
Denver Research Institute-Math Division
John Greene Hall
Denver, Colorado 80210

Vladimir Slamecka
Director, School of Information and Computer Science
Georgia Institute of Technology
225 North Avenue, N.W.
Atlanta, Georgia 30332

Allan F. Smith
Vice President for Academic Affairs
University of Michigan
Ann Arbor, Michigan 48104

Anderson Q. Smith
Director, Data Processing
University of Alabama
University, Alabama 35486

Gary R. Smith
Professor of Education
Wayne State University
Detroit, Michigan 48202

James L. Smith
Assistant Director
Ball State University
Computer Center
Muncie, Indiana 47304

Edward A. Smolsky
Head, Systems Development, Computer Center
Western Michigan University
Kalamazoo, Michigan 49001

Richard L. Snyder
Director of Libraries
Drexel University
32nd and Chestnut Streets
Philadelphia, Pennsylvania 19104

Robin E. Spock
Director, Instructional Systems Development
City University of New York
535 East 80th Street
New York, New York 10021

E. F. Staiano
Director, Freas-Rooke Computing Center
Bucknell University
Lewisburg, Pennsylvania 17837

Donald F. Stanat
Associate Professor
Kalamazoo College
Kalamazoo, Michigan 49001

John M. Steele
Purdue University
Lafayette, Indiana 47907

Einar Stefferud
Computer Management Consultant
Einar Stefferud & Associates
3002 Midvale Avenue
Los Angeles, California 90034

John C. Storlie
Director, Computer Center
University of Wisconsin
La Crosse, Wisconsin 54601

Jon C. Strauss
Director, Computing Facilities
Washington University
St. Louis, Missouri 63130

Len Swanson
Associate Program Director
EDUCOM
Princeton, New Jersey 08540

Bruce L. Taylor
Director, Systems Division
Educational Testing Service
Princeton, New Jersey 08540

Robert C. Taylor
University of Michigan
Ann Arbor, Michigan 48104

Dave Thomas
Department of Library Sciences
The University of Michigan
Ann Arbor, Michigan 48104

Duane Thomas
The University of Michigan
Ann Arbor, Michigan 48104

Karl C. Thomas
Director, Computer Center
Marshall University
Huntington, West Virginia 25701

Larry E. Travis
Director, Academic Computing Center
University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

Robert Tuite
AVTECH
Purdue University
Lafayette, Indiana 47907

Doreen Utmsa
Computerworld
797 Washington Street
Newton, Massachusetts 02160

R. E. Utman
Library Systems Analyst
Princeton University Library
P. O. Box 200
Princeton, New Jersey 08540

David Vance
President
Museum Computer Network Inc.
c/o Museum of Modern Art
11 West 53rd Street
New York City, New York 10019

Frank Verbrugge
Director, University Computer
Services
University of Minnesota
Minneapolis, Minnesota 55455

Stephen Vozella
Systems Analyst
Massachusetts State College
System
65 Franklin Street
Boston, Massachusetts 02110

Frederick H. Wagman
Director, University Library
University of Michigan
Ann Arbor, Michigan 48104

Stephen V. F. Waite
Faculty Applications Resource Advisor
Dartmouth College
Kiewit Computation Center
Hanover, New Hampshire 03755

Harold Wakefield
Assistant Vice Chancellor
State University of New York
9 Thurlow Terrace
Albany, New York 12201

Victor L. Wallace
Department of Computer Science
New West
University of North Carolina
Chapel Hill, North Carolina 27514

R. Brian Walsh
Director of Computing Center
University of Notre Dame
Notre Dame, Indiana 46556

Roger Walton
IBM Canada Ltd.
Toronto Dominion Center
Toronto, Ontario, Canada

Peter J. Wanger, Jr.
Oberlin College
Oberlin, Ohio 44074

Glenn O. Ware
Statistician
University of Georgia
Computer Center
Athens, Georgia 30601

Gerard P. Weeg
Director, Computer Center
The University of Iowa
Iowa City, Iowa 52240

James Wegryn
Fiscal Analyst
Legislature State of Michigan
Capitol Building
Lansing, Michigan 48902

Ed Weiss
Director Research Program
Office of Science Information Services
National Science Foundation
1800 G Street, N.W.
Washington, D. C. 20550

John Wenzel
Instructor, Data Processing
Wayne County Community College
Detroit, Michigan 48201

Franklin W. Westervelt, Director
Computer and Data Processing Center
Wayne State University
Detroit, Michigan 48202

Raymond A. Wiesen
Massachusetts Institute of Technology
455 Old Billerica Road
Bedford, Massachusetts 01730

Paul J. Wolfe
University of Kansas
Computation Center
P. O. Drawer 2007
Lawrence, Kansas 66044

Seymour J. Wolfson
Wayne State University
Detroit, Michigan 48202

Gordon R. Wood
English Department
Southern Illinois University
Edwardsville, Illinois 62234

Helen A. Wright
Associate Director
Computing Center
University of Michigan
Flint, Michigan 48503

Charles Wrigley
Professor
Michigan State University
East Lansing, Michigan 48823

Joe B. Wyatt
Director, Office for Information
Technology
Harvard University
1730 Cambridge Street
Cambridge, Massachusetts 02138

Eugene P. Young
Assistant Vice President for
Academic Affairs
Rutgers, The State University
New Brunswick, New Jersey 08903

Kenneth C. Young
Oakland University
Rochester, Michigan 48063

William A. Young
Director of New Learning Media
Association of Universities and
Colleges of Canada
151 Slater Street, KIP-5N1
Ottawa, Ontario, Canada

Warren T. Zemke
Assistant Professor of Chemistry
Wartburg College
Waverly, Iowa 50677

Karl L. Zinn
The Center for Research on
Learning and Teaching
The University of Michigan
Ann Arbor, Michigan 48104

Timothy Zorka
University of Michigan
23945 Fordson
Dearborn, Michigan 48124

Appendix A

List of Archives Participating in *S S Data* Newsletter Project

Project TALENT Data Bank
AMERICAN INSTITUTES FOR RESEARCH
Palo Alto, California

BEHAVIORAL SCIENCES LABORATORY
University of Cincinnati

BELGIAN ARCHIVES FOR THE SOCIAL SCIENCES

BUREAU OF APPLIED SOCIAL RESEARCH
Columbia University

CENTER FOR COMPARATIVE POLITICAL RESEARCH
State University of New York, Binghamton

DATA & PROGRAM LIBRARY SERVICE
University of Wisconsin

EUROPEAN CONSORTIUM FOR POLITICAL RESEARCH

INSTITUTE FOR BEHAVIORAL RESEARCH
York University, Ontario, Canada

INTERNATIONAL DATA LIBRARY & REFERENCE SERVICE
Survey Research Center
University of California, Berkeley

INTERSOCIETAL INFORMATION SERVICE
Northwestern University

SOCIAL DATA EXCHANGE ASSOCIATION
Providence, Rhode Island

SOCIAL SCIENCE DATA ARCHIVE
Urbana, Illinois

INTERUNIVERSITY CONSORTIUM FOR
POLITICAL RESEARCH
Ann Arbor, Michigan

LATIN AMERICAN DATA BANK
University of Florida, Gainesville

NATIONAL DUALABS, INC.
Rosslyn, Virginia

NATIONAL OPINION RESEARCH CENTER
University of Chicago

POLIMETRICS LABORATORY
Ohio State University

POLITICAL DATA ARCHIVE
Michigan State University

POLITICAL SCIENCE LABORATORY AND
DATA ARCHIVE
Indiana University

REGIONAL SOCIAL SCIENCE DATA ARCHIVE
University of Iowa

ROPER PUBLIC OPINION RESEARCH CENTER
Williams College

SOCIAL SCIENCE DATA CENTER
University of Pennsylvania

SOCIAL SCIENCE INFORMATION CENTER
University of Pittsburgh

SOCIAL SCIENCE DATA ARCHIVES
Carleton University, Ottawa

SOCIAL SCIENCE DATA CENTER
University of Connecticut

SOCIAL SCIENCE USER SERVICE
Princeton University Computer Center

SOCIOMEDICAL RESEARCH ARCHIVES
Columbia University School of
Public Health

For further information, contact:

John G. Kolp, Editor
s s data
321A Schaeffer Hall
University of Iowa
Iowa City, Iowa 52240

Appendix B

A Survey of Communications Activities in Economics

Innovative activities in the field of communication and information dissemination have not been lacking in Economics, particularly with regard to distilling the literature. *The Journal of Economic Abstracts* proved successful in bringing together abstracts of articles, and has been expanded to include review articles in the *Journal of Economic Literature*. Earlier, the Rockefeller Foundation noted the problem of fragmented literature and sponsored a series of survey articles which appeared in the *American Economic Review* and *Economic Journal*.

Activities involving the use of computer technology are becoming more common, but is it here that expertise and cooperative planning become essential. No single society or institution can assume responsibility for maintaining and modernizing communications channels, but together the major scientific societies and research institutions can explore how to use the computer to reduce inefficient lags and leakages. Currently, a number of independent projects are attempting to bring computer technology to bear on communications problems.

Administrative information systems have been candidates for computerization in recent years. The American Marketing Association has computerized its mailing list and uses the system for handbooks; and the 1970 Handbook for the American Statistical Association was computer generated. Also, the Econometric Society has been computerizing the administrative and information dissemination functions of societies. Programs have been developed to produce mailing labels and billings, and under an NSF grant, a free form entry system was developed to facilitate the publication of handbooks and bibliographies. SIPPS (System of Information Processing for Professional Societies) is currently used by the National Bureau of Economic Research and the Econometric Society and is being applied to other organizations.

Handbooks with biographical listings of the membership help one locate addresses of individuals with similar research interests. A logical extension in the age of the computer is the creation of computer files with both biographical and bibliographical information. Economists regularly find themselves answering questionnaires about their background and research for NSF studies of scientific manpower and for individual societies. Surely, if the stock of required information could be entered into computer files, the scientific community would have a valuable research tool. Furthermore, the societies could update and use the data base for daily administrative requirements. Perhaps within a decade the professional labor market will be improved through the use of a system with machine-readable vitae. Colleges, government, or businesses could search for a person whose interests and talents best matched their needs. Book review editors might search for appropriate reviewers, the sponsors of conferences could make meetings more all-inclusive, and researchers would find retrospective searches of the literature much more efficient.

Computerized bibliographies are currently being developed in conjunction with several projects. The American Agricultural Economics Association has established an Agricultural Economics Documentation Center in conjunction with the National Agricultural Library, and input and retrieval continues on a regular basis. The computerized bibliography is being used to produce a publication which can bring researchers up to date on working papers appearing in the field of agricultural economics. Independently, R. Hicks has developed a computerized bibliographic system for the field of public finance.[1] Also, it should be noted that listings of dissertation titles provide useful pointers to new research. University Microfilms has a computerized retrieval system that would complement future information systems in the discipline.

The American Economic Association is currently exploring the feasibility of applying the computer system developed by the Econometric Society to the production of the quarterly *Journal of Economic Literature*, thus the annual *Index to Economic Literature* could be derived automatically from the *JEL* data base. The *JEL* is central to any computerized information system in Economics, because if author, title, journal, and classification number were entered into a master file, that file could then be sorted alphabetically (by author) or numerically (by classification number), eliminating much of the costly hand operation. As a by-product of the printed volume, the Association would have a machine-readable file that could be searched on-line. Given present technological and cost trends, a computer based information retrieval system will be feasible soon. Physics, Chemistry, and Psychology are well on their ways to such systems. However, searches for articles with particular keywords or print-outs of abstracts classified under the AEA indexing scheme are probably five years away, given the absence of resources and institutional commitment for developmental work.

This survey indicates that Economics should begin to identify problems of compatibility and possibilities for merging current efforts to prevent the duplication of efforts. At present, different organizations are caught in a common syndrome: delivering products before systems are truly operational.

Current efforts lack funds to take the broader view of the work. For example, if the alerting service for agricultural economics proves useful, other fields in economics may be able to utilize the system. We also need user studies which would give us better information about the impact of innovations on communications channels. For example, what should be the role of specialized information centers and should their development be encouraged on an *ad hoc* or systemic basis? Unless consideration is given to such problems as network compatibility, Economics could find itself with several duplicative systems — none of which adequately meets the communications needs of the profession, and none of which is easily integrated with the others.

REFERENCE

[1] Robert J. Hess, "Programmed Bibliography in Public Finance," *Proceedings of the Sixty-second Annual Conference on Taxation*, Stanley J. Bowers, ed., Columbus, Ohio, 1970, pp. 593-601.

Appendix C

Networks — Whose Panacea?

Hans Lee
Computer Science & Sociology
Michigan State University

The purpose of this paper is to place the consideration of computer networks in a broader context so as to facilitate the emergency of constructive criticism. The discussion assumes the supremacy of the value position that whatever we do as individuals, citizens, teachers, political decision makers, or scientists must be directed towards the betterment of the human qualities in mankind. Further, it is assumed that from the standpoint of wisdom, it is imperative to place a consideration of networks and disciplines in a broad context that ultimately will be oriented toward this goal.

Perhaps the most fundamental mistake that could be made is to restrict the question to, "Given computer networks, what do we do with them?" Yet, this question seems to be motivating the emphasis upon funding programs for network development. Instead of responding to this question, we should first consider a number of other questions: As members of an academic discipline, what type of resources do we desire and for what purposes? What are the costs involved with the utilization of a particular resource? How are these resources related to the improvement of our understanding? In what ways may using particular resources distort research objectives and outcome?

The perspective from which these ideas emerge represents a combination of various viewpoints. It includes the sociological framework in which large complex institutions and organizations are studied; the social psychological interests in interpersonal relations and relationships among men and between man and machine; as well as economic and political aspects of decision making. In addition, the issues in the philosophy of science, but especially those emerging from the philosophy of social science, permeate the considerations underlying the thoughts developed in this paper.

Ultimately, most of the financial resources, as well as other forces, pushing the rapid development of computer technology emanates from the federal government of the United States of America. Probably a majority of the applications which initially motivated computer development by the federal government were applications either in the military or in recordkeeping. These same two areas of application were, no doubt, also contributory to the orientation of computer development toward very large computer systems. It is therefore not surprising to find that a substantial proportion of the funds supporting computer development and application are emanating from the advanced projects program of the Department of Defense. It is especially intriguing to see that the first fairly successful nationwide computer network is the ARPA network sponsored by the defense projects division of the Department of Defense.

As a general principle for the development of large systems, it would seem reasonable to maintain that it is better to permit the pressure for the development of the system to arise naturally from its users, rather than to force the development upon the users from higher bureaucratic levels. (This seems to be an even more sensible design philosophy if the designers are concerned for bettering the human qualities in mankind.) Consequently, it is appropriate to inquire whether the demand for networks is arising to satisfy user's needs, or is the policy emphasis upon networks a bureaucratic decision being imposed upon the community? It seems quite plausible that one motivation for the National Science Foundation's pushing the development of its own network must be the competition emanating from the existence of the ARPA network. Another factor possibly motivating the National Science Foundation's computer network development program seems directly attributable to the initial policies of NSF in the deployment of computers in universities. In the interests of getting computers quickly established in university environments, the National Science Foundation either completely or partially supported the costs of computers in several large universities. Complicating the deployment of computers in universities was the early policy of the federal government in its accounting practices which insisted that federally supported research contracts to universities not be charged any more than the lowest rate charged any other computer user. University administrators and faculty obviously also supported the position that it would be ideal to obtain financial support from external agencies for funding their computer development programs. Until quite recently, however, funding computer development programs meant either the funding of the purchase or the funding of the leasing of a computer. Because there were no widely available computer utilities to supply computer services, administrators of universities tended to think of the purchase of a computer as a "one-shot" arrangement. Further, both because of the nature of the computer industry as well as the insatiable desire of faculty, students, and computer operations personnel, for more equipment and more computing resources, administrators seldom fully appreciated the fact that the cost of the computer hardware itself was only a small proportion of the total charges associated with offering high quality computer services to the university communities. Even today, many university administrators do not appreciate the dynamic nature of the development of both computing hardware and all supporting services. Presumably, the National Science Foundation has grown concerned over the policy of funding computing development in universities. One would imagine that it would not be an attractive proposition to anticipate that the National Science Foundation itself would have to continually fund new computers to replace obsolete ones, as well as to fund computers for colleges and universities without their own computers. Another factor would also appear to be contributory to the direction of allocation of the funds of the National Science Foundation. This other factor is that it probably is considerably easier to justify to Congress the expenditures for equipment than it is to justify expenditures made for research programs whose only output are ideas in the form of final reports. This factor presumably subtly influences the direction of the allocation of funds within the National Science Foundation to various programs.

Of particular significance to the humanities, however, is the fact that many forms of research in humanities has been assigned to the National Endowment for the Humanities, whereas the development of computer networks is primarily associated with the Office of Computing Activities within the National Science Foundation. One consequence of separating these agencies is that the humanities which use computer services may find it very difficult to influence the decision making about the kinds and allocations of such services.

Obviously, the federal government and its associated agencies have been responsible for the development of very large computer systems. It may be that agencies and institutions responsible for the development and operation of large systems are suffering from an embarrassment of riches. Perhaps they have a lion by the tail. To justify past resource allocations (both time and money), it is necessary to demonstrate productive utilization of such systems. This, in turn, requires an increasing demand for computing services. As the amount of computation done by a computer increases by a factor of four with the size of the computer, the so far very rapid increase in the size of computers from computer generation to computer generation implies the necessity of creating more demand for computing services.

We must not, however, forget the social-psychological level of analysis. Here it is important to recognize the tremendously interesting problems that exist both in the design of hardware and the development of extensive, complex software systems. Unfortunately, little research is being conducted on the human as a user of computer services. Perhaps the key to the understanding of the critical position taken in this paper is contained in trying to indicate who are the "users" of the computers or networks.

From the perspective of the author of this paper, many policies regarding users seem to be unsatisfactory for the reason that the policies are not designed for the individual consumer. Instead, policies seem to be designed for the convenience of the staff who manage and operate the computational services. The situation appears similar to the hospital. Sociologists have been pointing out that many hospital policies are designed for the convenience of the management and staff, but have not been designed from the perspective of the needs of the patient as a human being. So, maybe just as hospital policy is oriented around the physicians and nursing staff as "users," so are computer centers' policies oriented around systems programmers and operational staff. Consequently, henceforth the word "user" refers to the human user "at the bottom" (and for many academic disciplines, the user has not used a computer, but rather is a possible or potential user). Thus, the word "user" will not, in general, refer to members of the design, operation, or maintenance staff of computer centers.

Consequently, conclusions must be based upon more informal observations. For example, it seems that some users are motivated to obtain computer services from a computer network because they are dissatisfied with the quality of service available from their own computer facility. As another example, consider computer usage in social sciences. Just as there has been a tendency for social scientists to include tests of statistical significance as an imprimatur, there now appears to be a tendency for them to use the fact of computer analysis of data in the same uncritical way. Somehow, the phrase, "the data were analyzed by computer program XXX," seems to enhance the quality of the content of the reported research.

In reflecting upon the emergence of computers, and now the emergence of computer networks, I have found that it is helpful to think of the following politically unreasonable but intellectually exciting idea of asking the prior question, "If I had X amount of resources, how would I allocate them?" There are a number of points that should be made in attempting to deal with this alternative question. These are briefly indicated as follows:

1. We must ask the prior question of whether computing service is necessary for achieving our objectives. For example, in sociology it would appear that before students are trained in the use of statistical programs, they should have a prior understanding of the statistical procedures, the philosophy of science, and the sociological perspective. Otherwise, the ease with which students can learn to submit data to canned programs can only result in meaningless output.

2. In moving toward computer use, the first step frequently is to adapt old methods to the computer. Yet, this may be a fundamental mistake. In social sciences, for example, in some cases there would appear to be no need to use only a sample of cases, but rather to examine exhaustively all cases. As the ideas and strategies for research under such circumstances have not yet been examined in any great detail, one should concentrate on the development of appropriate research strategies before quickly moving toward computer analysis.

3. There seems to be an interesting carry-over from the emphasis upon hardware visible in the area of applying computers to the educational process itself. (All too frequently hardware development dictates software development rather than vice versa.) Even in the development of software programs to assist computer-assisted instruction, the assumption is made that we understand how to teach, and therefore, know how to use the computer to teach. In particular, then, one should suggest that prior to the allocation of money for the development of computer systems, a very large portion of monies should be spent on studying how to teach well and how individuals learn best. Although such research probably would utilize computing services, the research would be more of a social psychological nature than the so-far traditional emphasis on hardware and software. Also

evident in the CAI field is a tendency to hand down policies rather than to permit policies to be developed as a response to users' demands.

The overall picture of the development of the humanities, social and hard sciences via computers is somewhat discouraging. The present institutional arrangements seem to force computers upon the user. Further, it seems that computer services are thought of as primarily providing hardware and minimal software, most of which is oriented either toward simple bookkeeping but on a large scale, or else toward numeric computation. Further, the present policies of funding agencies seem to be furthering the gap between the hard sciences and the soft sciences and humanities. There seems to be no way to close this gap unless policies are implemented which allocate greater expenditures on the arts, humanities, and social sciences, considerably curtailing and slowing the rate of development of hard sciences.

Even if the discussion is restricted to computer services, there still seems to be a number of special things that should be done:

There needs to be considerably more attention toward nonnumeric computation utilizing large data bases. In some typical realistic sociological applications, the problem of interest may require a longitudinal analysis of "messy" data, i.e., incomplete data, varying sizes of subgroups, etc. Such applications require a much more thorough consideration of software development for nonnumeric computation. Analyses using such data bases require sensitive interaction between researcher and the data facilitated by computation. The quality of this interaction could be significantly enhanced by sophisticated (in the underlying details of implementation) interactive graphics systems that, to the user, appeared reasonably straightforward. Designers for such systems should study the various interactive design systems developed in the automobile industry.

Although many people have different answers to the various sub-problems it seems that in the initial phases at least, more attention should be given to operating systems which facilitate the intermixing of various language processors in one job. But this should be done in an environment which permits direct observation of the process of utilizing computers by the individual scholar. That is, in developing improved software systems, research should be done directly on man-machine interaction. Thus, a new type of anthropologist-social psychologist is needed who would observe the activities of the researcher, and at the same time correlate these with the history of actual machine computations. To make a little bit more concrete the two ideas just mentioned (a more flexible operating system and the student of man-machine interaction), I would like to describe what I would like to be able to do with a computer.

I am interested in inference behavior by man and by machine. I would like to study this by developing a computer program which appears to converse with an individual by a teletype utilizing natural language. I would like to take the input statement of the subject and parse it by SNOBOL into a cononical form which would then be passed to a complicated dictionary lookup, probably written utilizing some of the features of a FORTRAN/SLIP system, especially the file manipulation capabilities of FORTRAN and the list processing features of SLIP. The output of this second task would then be handed to a graph-theoretic-graph-processing language locally available (HINT) to compute the "meaning" of the input by graph processing. The result would be referred to the inference making portion of the program, most likely also written in HINT. After the inference making, the program would prepare a response, probably in HINT. Then the response would have to be prepared for output probably by going through a FORTRAN/SLIP system to find appropriate vocabulary items and thence, to the SNOBOL system to construct the appropriate English sentences. Finally, the operating system, as well as each of the language processors, should be programmable in an interactive environment, such as available within the CAI author's language, PLANIT.

If it were possible for researchers to utilize a computer with an operating system that permitted jobs to be done in a variety of languages, and if such research projects were themselves an object of study, then we would have a better information base on which to design newer and more appropriate language processors and operating systems.

In summary then, there seems to be a serious lopsidedness in the emphasis upon computer networks. The policy making bodies seem to be prematurely forcing the rapid development of computer networks rather than permitting user demand to arise naturally. Also, the network development programs seem to act as if the major costs of supplying computer services are hardware and software costs, with emphasis upon hardware. Yet, the major costs of computing services are human costs: documentation, consulting, system maintenance, and last, but not least, human users who have good ideas. Perhaps the entire thesis of this paper is that the network development programs fail to view computers in the context of information sciences and fail to be user oriented.