

MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES



3 4456 0379246 6

ORNL/TM-12651

ornl

**OAK RIDGE
NATIONAL
LABORATORY**

MARTIN MARIETTA

Energy Division

**Using The National Information
Infrastructure For Social Science,
Education, and Informed
Decision Making**

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this
report, send in name with report and
the library will arrange a loan.
UCR-7998 (1-87)

B. E. Tonn

January 1994

MANAGED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P. O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-12651

400
5

Energy Division

**USING THE NATIONAL INFORMATION
INFRASTRUCTURE FOR SOCIAL SCIENCE,
EDUCATION, AND INFORMED DECISION MAKING**

(A White Paper)

Bruce E. Tonn

January 7, 1994

Prepared by
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-84OR21400



3 4456 0379246 6

Table of Contents

1.0 INTRODUCTION	1
2.0 ELEMENTS OF THE NII-SSEDM	5
2.1 DATA COLLECTION	5
2.2 DATABASES	6
2.3 DATA ANALYSIS AND MODELING	7
2.4 POLICY ANALYSIS AND DECISION SUPPORT	8
2.5 EDUCATION	9
2.6 IMPLEMENTATION ISSUES	10
3.0 NII-SSEDM APPLICATION EXAMPLES	13
3.1 ECONOMIC PROSPERITY	13
3.2 HEALTH CARE	14
3.3 CRIME AND VIOLENCE	15
3.4 LIFE-LONG LEARNING	16
4.0 CREATING THE NII-SSEDM	19
4.1 PRE-YEAR 2000 GOALS	19
4.2 POST-YEAR 2000 GOALS	21
5.0 CLOSING REMARKS	23

EXECUTIVE SUMMARY

The United States has aggressively embarked on the challenging task of building a National Information Infrastructure (NII). This infrastructure will have many levels, extending from the building block capital stock that composes the telecommunications system to the multitude of higher tier applications hardware and software tied to this system. This "White Paper" presents a vision for a second and third tier national information infrastructure that focuses exclusively on the needs of social science, education, and decision making (NII-SSEDM).

NII-SSEDM will provide the necessary data, information, and automated decision support and educational tools needed to help this nation solve its most pressing social problems. The proposed system has five components: data collection systems; databases; statistical analysis and modeling tools; policy analysis and decision support tools; and materials and software specially designed for education. This paper contains: a vision statement for each component; comments on progress made on each component as of the early 1990s; and specific recommendations on how to achieve the goals described in the vision statements.

The white paper also discusses how the NII-SSEDM could be used to address four major social concerns: ensuring economic prosperity; health care; reducing crime and violence; and K-12 education. Examples of near-term and mid-term goals (e.g., pre- and post Year 2000) are presented for consideration. Although the development of NII-SSEDM will require a concerted effort by government, the private sector, schools, and numerous other organizations, the success of NII-SSEDM is predicated upon the identification of an institutional "champion" to acquire and husband key resources and provide strong leadership and guidance.

1. INTRODUCTION

In a publication released by the Clinton Administration in the fall of 1993, entitled *The National Information Infrastructure: Agenda for Action*, the opening paragraph states:

"All Americans have a stake in the construction of an advanced National Information Infrastructure (NII), a seamless web of communications networks, computers, databases, and consumer electronics that will put vast amounts of information at users' fingertips. Development of the NII can help unleash an information revolution that will change forever the way people live, work, and interact with each other...."

Following this spirit, this "White Paper" presents a vision for a national information infrastructure to support social science, education, and decision making (NII-SSEDM). Users of the NII-SSEDM will be social scientists, policy analysts, governments, businesses of any size, decision makers (which includes individuals and households as well as those in business, government, etc.), educators, students, and any citizen interested in learning. Thus, the NII-SSEDM is envisioned as a true national resource.

Although few would argue against the goals of the NII-SSEDM, this paper is important because social science and decision making are topics that are often neglected in NII discussions (and subsequent funding allocations). Also, these are fields that can both benefit from the NII (and other computer resources) and contribute to the efficient and equitable development of the NII. This paper is also important because it stresses linkages between education and social science and decision making that can be strengthened through an advanced telecommunications network.

The proposed NII-SSEDM will support social science, public and private sector decision making, and education in the broadest possible sense. Social science encompasses research and analysis aimed at understanding past, present and future human behavior. Thus, social science includes economics, psychology, sociology, political science, anthropology, and geography, as well as policy science, urban planning, transportation, environmental planning, education, and the humanities. In an ideal world, knowledge gained by the social sciences should be used to educate citizens about human nature and how it drives economic, cultural, and political behavior. It can be strongly argued that appreciation of such knowledge is necessary for the effective working of a democratic society.

Appreciation of human nature must be complemented by wise application of this knowledge in decision making contexts. The United States is a society of over 250

million people that is organized into approximately 87,000 governmental bodies, over 100 million households, and tens of millions of firms, non-profit institutions, and voluntary organizations. Literally, millions of decisions are made every day. Most of these decisions are rather routine, a few may be monumental, but all are important in some respect to those involved. To make good decisions, insights into human nature must be supplemented by: reliable and timely data; information describing the problem and aspects of the surrounding situation; and informative analyses of the consequences of the various decision alternatives. Thus, decision makers can benefit from access to databases, information services, knowledge bases, and decision support tools that can be made available through the NII-SSEDM.

Inherent in the most widely accepted concepts of "citizen" and underlying good decision making is education. Education has utilitarian value (e.g., reading, writing, and arithmetic, vocational training, professional studies, etc). Education also has humanistic value. We all need to understand our heritage and the heritage of others, to be conversant about the great ideas of philosophy, democracy, and literature. With "technical" knowledge, "shared values" and an understanding of others' viewpoints, people can participate knowledgeably in public discourse, and make better decisions to benefit themselves, their employers and other organizations, their communities, and their nation. Thus, the NII-SSEDM, through its emphasis on society and education, logically supports the goals of the National Information Infrastructure.

It should also be noted that the NII-SSEDM can also be seen as being crucial for the success of another Clinton Administration initiative, that of reinventing government. Certainly, social science methods will be needed to evaluate current and prospective government programs. Based on these evaluations, difficult decisions will be made about the future goals, scope, and even existence of government programs. As part of this entire process, many federal program administrators will have to educate themselves on the topic of program evaluation, if not on numerous other topics.

What capabilities should the NII-SSEDM have to support the reinventing government initiative as well as social science, decision making and education in general? The NII-SSEDM should allow: access to important databases (e.g., 1990 Decennial Census) over the Internet and its descendants; the efficient collection of data; and use of sophisticated data analysis and modeling tools. It is also envisioned that the system will support complex economic and social modeling, and provide both public and private sector decision makers with on-line access to data, models, and decision support tools. Lastly, NII-SSEDM should be a user-friendly and economical source of educational resources fall all ages, topics, and types of learning. NII-SSEDM capabilities are discussed in much more detail below.

The goals associated with NII-SSEDM are fundamentally different and more challenging than goals associated with the NII alone or goals associated with other

federal computing initiatives. In the first case, the NII alone can be seen as an initiative promoted by and implemented by a relatively few large public and private organizations. In the second case, federal programs such as the Computing Grand Challenge Program can be pursued rather independently by various research groups around the country. Building the NII-SSEDM, on the other hand, will require both the commitment of large public and private organizations as well as a significant grass roots effort to shape the NII-SSEDM to the needs of an extraordinarily diverse user groups. Also, because the NII-SSEDM requires a range of components—hardware, software, data and materials—a genuine national effort will be needed to integrate hardware, set standards for software, provide access to materials, etc.

The first section of this paper addresses the development of the NII-SSEDM, with specific attention focused on five specific components: data collection systems; databases; data analysis and modeling tools; policy analysis and decision support tools; and materials and software designed for education. Discussions of each component include: a vision statement (i.e., what would be nice to have); a summary of the current state-of-the-world (as of the early 1990s); and recommendations about how to move from here to the preferred future. This section also tackles general implementation issues and especially identifies the need for an institutional "champion" for the NII-SSEDM.

The second section describes topics of national concern and how a national information system architecture that focuses on social science, policy analysis and decision making, and education can help the nation confront these serious problems. The third section presents action items related to organizing this effort and tasks targeted to be accomplished in the near- and mid-term.

2.0 ELEMENTS OF THE NII-SSEDM

The purpose of this section is to set out the functional requirements for the NII-SSEDM. In other words, this section focuses on system components and capabilities. The next section presents examples of how the components and capabilities could be used to tackle specific problems related to social science, decision making, and education.

It needs to be stressed at this point that ideas presented herein for the NII-SSEDM require a national telecommunications infrastructure, in other words the NII itself. This infrastructure will need to have a very high bandwidth to handle multi-media applications, and be widely distributed, eventually to every home, office, school, etc. Access to and use of this technology-base needs to be economical, equitable, and timely. It's technology will be a complex mesh of technology from at least three major industries: phone, cable, and computer. It will feature fiber optics, cellular systems, and satellites. People will be able to be connected to the system through several modes, including standard televisions, standard computers, hybrid "telecomputers", and mobile personal communications systems.

Built upon the national telecommunications infrastructure will be the NII-SSEDM. The NII-SSEDM can be conceived as having ***five*** basic elements: data collection systems; databases; data analysis and modeling tools; policy analysis and decision support tools; and on-line materials and software designed for education. The five are related in the following way: data collection techniques can be used to build sophisticated databases; the databases are needed to support data analysis and modeling; and data analyses and model results, in conjunction with decision support tools, are needed to support policy analysis and decision making. Educational activities can incorporate access to databases, data analysis software, models, and decision making software, and even teach people about data and decision making.

This section presents for each component a vision statement, a brief description of current accomplishments, and a development strategy. The section concludes with a discussion of implementation issues

2.1 DATA COLLECTION

Vision Statement: The NII-SSEDM facilitates on-line, real-time, automated data collection when and where data are created (e.g., economic transactions in stores, medical procedures in hospitals). The NII-SSEDM will be accessible to cellular, geosynchronized hand-held devices as well as stationary data collection technology. Accessible over the network will be systems capable of interpreting digital

representations of any sort of data (e.g., handwriting on 19th Century Census forms) or symbolic information collected through knowledge acquisition exercises. Data will be collected in ways to ensure privacy and minimize burden on citizens and will not be collected without proper consent.

Current Status: Automated data collection has made great strides within the social sciences and in other areas of endeavor. In existence are computer-assisted telephone interviewing systems, computer-assisted personal interviewing systems, touch-tone data collection systems, bar code scanners, and hand-held data entry devices. Some progress has been made in the area of handwriting recognition. In most cases, these technologies have been developed as stand-alone systems and are not accessible over the Internet. Point-of-sale data collection systems are spreading quickly in commercial retailing. However, these data are not typically available to people outside of the companies collecting the data and other "real time" data of interest to social scientists and policy analysts are generally not being collected.

Development Strategy: Develop a conceptual roadmap of where important data are being generated and how to collect the data by electronic and networked means. Institute a federally funded RD&D program that addresses but is not limited to the following topics: handwriting recognition; automated data checking and editing; hand-held data entry; computer-assisted interviewing; intelligent survey techniques; privacy issues; security of private data entered into real-time, networked systems; and automated electronic data collection. Assign coordination responsibility to the major federal statistical agencies in partnership with appropriate private sector organizations that specialize in data collection. Encourage entrepreneur, given flexible standards.

2.2 DATABASES

Vision Statement: Users will have on-line, real-time, 24 hour access to a myriad of databases that describe the human condition, past and present. Among the many databases that ought to be made available are: every Decennial Census; Census Tiger Files and other spatial databases; Current Population Surveys; other databases compiled by the major federal statistical agencies; topographic data; environmental data; and data collected by private companies, universities, and other institutions. The infrastructure will provide access to sophisticated record linkage algorithms, and intelligent software capable of assembling custom-made databases. Controls will be implemented to ensure privacy. Interfaces to the databases will be very user-friendly and have intelligence to "surf" the database environment to find data and construct answers to complex queries.

Current Status: Great progress had been made in storing new databases in digital form and some progress has been made in converting old databases to digital form

(e.g., work at the National Center for Supercomputing Applications at the University of Illinois with 19th Century Census data). In addition, the Office of Management and Budget recently announced the federal government's intention of making all its public data more accessible (see Circular A-130). Access to digital statistical databases over networks to download data or submit queries is not common or even possible in most cases, although proposals to begin the development of such capability are beginning to be developed (e.g., by the University of Michigan, Population Studies Center). The Consortium for International Earth Science Information Network (CIESIN) is making progress in putting on-line databases to support the analysis of social science issues associated with global climate change. No systems exist to allow researchers and analysts to easily custom-build complex databases from numerous, disparate, remotely located databases.

Development Strategy: Assign responsibility for the database component of the infrastructure to the major federal statistical agencies (e.g., Census Bureau, Bureau of Labor Statistics) and other major repositories of data (e.g., Smithsonian Institution, Library of Congress). Direct the agencies to solicit input on their plan of action from federal mission agencies (e.g., Dept. of Commerce, Dept. of Labor, Environmental Protection Agency), universities, the private sector, K-12 education institutions, non-profits, etc. Work towards implementing a national database server center model, using existing high performance computer organizations (e.g., NSF Supercomputer Centers, National Laboratories) as potential sites for national database centers. Implement research, development and demonstration (RD&D) programs to support this infrastructure component. Research topics include: database design (especially for parallel architectures); record linkage; data quality assessment; confidentiality; interface design (especially for disabled, illiterate and other people who might have difficulties using standard interfaces); intelligent data retrieval; development of standards to facilitate better database access and custom database creation; and client-server architectures.

2.3 DATA ANALYSIS AND MODELING

Vision Statement: The NII-SSEDM will allow users to assemble custom-made databases, submit them for analysis using the most appropriate statistical procedures, process the data on the most appropriate computers, and evaluate results using preferred numerical or graphical modalities (e.g., 3-D visualization). The NII-SSEDM will also allow users to assemble inputs for sophisticated models, which reside on appropriate computers, and inspect outputs in real-time using preferred numerical or graphical modalities (e.g., virtual reality). Lastly, users will have the ability to assemble meta-models (i.e., collections of models) to conduct integrated analyses of important problems, such as global climate change.

Current Status: Social scientists and policy analysts have had access to statistical software for a number of years. The software first resided on mainframes, has recently migrated to microcomputers and workstations, and is beginning to migrate up to supercomputers and parallel processors. In general, though, users' data and software reside on the same computer, either on their desk-top or in their institution's computer center. It is rare for users to access data from a remote site, and ship the data to another, computationally appropriate remote site for analysis. A similar story holds true for models. There are countless models running on all types of computers, even high performance computers. However, only rarely do those who were not intimately involved in building the models get to exercise others' models, remotely or otherwise. The potential of using the Internet to run an integrated set of models is not being realized.

Development Strategy: Assign responsibility for this component to a select consortium of institutions heavily engaged in data analysis and modeling (e.g., mission agencies such as Dept. of Labor, non-profit think tanks such as Rand, and research universities). Integrate client-server models for large scale database analysis with the database site concept. Develop a national model access system, again integrated with the database site concept. Implement an RD&D program, whose research topics would include, but not be limited to: data visualization; parallelization of statistical codes and models; integrated micro-macro economic-energy-environmental models; artificial intelligence models of human behavior; and custom database design and development facilities.

2.4 POLICY ANALYSIS AND DECISION SUPPORT

Vision Statement: The NII-SSEDM will allow users to have answers to questions in real-time, whether the question requires a simple database query, a series of queries, statistical analysis of one or more databases, exercising of one or more models, or any combination of these activities. Answers will be provided in the preferred modality. The NII-SSEDM will also facilitate collaborative decision making, creative thought, and structured decision making. It is envisioned that this component of the NII-SSEDM will be particularly valuable to government and widely used by business, households, etc.

Current Status: Policy analysts and decision makers are not being well served at the present time. They are presented with the results of statistical analyses and model runs, but, by and large they do not have access to software on their desk-tops that can assist them in making decisions. There are a growing number of decision analytic-based microcomputer systems on the market and groupware is becoming more widespread. However, decision analytic software and groupware have not been synthesized with each other or with the ability to access numerous remote databases

and models over a national network. Nothing exists that approaches the vision presented above.

Development Strategy: Jointly assign responsibility for this component to: a select group of mission agencies that depend most highly upon timely, accurate, and informative policy analyses (e.g., Dept. of State, Dept. of Defense); a group composed of representatives from the private sector; and a group representing the needs of families and individual decision makers. Create a small set of policy analysis/decision making support sites, where master software will reside that provides to these people the kinds of support envisioned above. Institute a RD&D program focusing on, but not limited to, the following topics: query software; knowbots; virtual reality and visualization; data mining and machine learning; intelligent problem solving; groupware; and collaborative decision making. Special attention should be paid to developing interfaces that are easily used by families and individual decision makers to make every day decisions, such as those regarding employment, major purchases, insurance, health care, personal safety, etc.

2.5 EDUCATION

Vision Statement: The NII-SSEDM facilitates human learning, regardless of age and ability. Intelligent software quickly develops individualized instructional strategies, depending on the subject matter, the learner(s), and setting (e.g., distance learning, on-the-job training). Instruction is available on almost every conceivable subject. Intelligent software builds upon software developed under the other components of the NII-SSEDM to provide access to the widest range of educational materials. The NII-SSEDM uses graphical and other sensory techniques to communicate with illiterate, vision impaired, and people who possess other learning disabilities. The NII-SSEDM is developed in such a way to allow entrepreneurs, school teachers, businesses, researchers, museums, and others to contribute materials, software, and other resources to a national, on-line educational system.

Current Status: In the past few years, use of information technology for educational purposes has grown rapidly. Availability of highly sophisticated multi-media systems is leading the way. The Global Jukebox by Alan Lomax of Hunter College is a wonderful example. More schools are tying into the Internet and other networks, and are building computer-based materials into their curriculums. However, most educational software is microcomputer-based, unintelligent, and unnetworked, and many schools that do have computer resources do not understand how to best apply them for teaching. Businesses and other large organizations are implementing more computer-based training, but often this type of training still takes place in traditional classroom settings. The nation has only scratched the surface with respect to the potential for using information technology for education.

Development Strategy: Assign responsibility for this component to organizations: engaged in educational activities (e.g., Dept. of Education, National Science Foundation, universities, K-12 institutions); possessing valuable expertise in technology (e.g., national laboratories, universities); and representatives from the private sector. Adopt the following principle as the guiding strategy: "Universal-network access first; duplicated and distributed resources second." Implement a broad based RD&D program that includes but is not limited to the following topics: intelligent computer aided instruction; database/network interfaces for K-12 and disabled persons; automated curriculum development support; using technology to integrate classrooms with real world activities; virtual reality and holographic image generation; and educational material database design and access.

2.6 IMPLEMENTATION ISSUES

Vision: The NII-SSEDM is developed in a coordinated fashion through the combined efforts of the federal, state, and local governments, universities, the private sector, and non-profit institutions, with a special emphasis on balancing top-down planning with user driven and bottom-up needs and activities. Resources on the network are made available in an economically efficient but equitable fashion. Intellectual property, confidentiality, security, and other important issues are handled in an ethical fashion. Competition is encouraged when appropriate. Much if not most of the management activities required for the NII-SSEDM are facilitated by the NII-SSEDM itself.

Current Status: Nothing remotely resembling this vision exists in the United States at the current time. People who would most benefit from the NII-SSEDM—social scientists, decision makers, educators, students—have not enjoyed the opportunity of a large-scale, coordinated effort at technology development analogous to the Apollo Program, the Human Genome Project, the Intelligent Vehicle Highway System, the Stealth Bomber, or the Hubble Telescope. It can be argued that in past years, technology was not available to accomplish much in to create the NII-SSEDM. However, this is no longer the case.

Development Strategy: Implementing an Apollo-like program for the social sciences, decision making, and education requires strong leadership. Therefore, the most important implementation goal is to find a "Champion" for this initiative. The champion should, in all likelihood, be a federal agency that possesses both the mission needed to justify its involvement with creating the NII-SSEDM and the staff with the technical and organizational background to handle the task. The champion needs to work closely with other federal and government agencies. It is recommended that something similar to the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET) Committee, which is overseen by the Office

of Science and Technology Policy (OSTP), be created to oversee the development of the NII-SSEDM. The "Council" would organize efforts in the five areas mentioned above and work to achieve budgetary goals needed to meet the expectations listed below.

It is also important that both the champion and any coordinating bodies work closely with the private sector and others on the development of the NII-SSEDM. Under the "let a thousand flowers bloom" strategy, most of the real work in developing software and materials will not be done by the government but by others. The champion needs to establish the vision and the government needs to establish the groundrules and mediate disputes. Others will take these cues and work in various ways to make the NII-SSEDM a reality.

3.0 NII-SSEDM APPLICATION EXAMPLES

Unfortunately, there is no lack of pressing social problems to attack with the aid of a national information system infrastructure. The four topics discussed in this section—economic prosperity, health care, crime and violence, and life-long education—are serious problems in their own right, but their selection for discussion should not indicate that topics not discussed are viewed as being less important. The four topics were chosen to cover the spectrum of social problems, social science disciplines, and high performance computing and telecommunications applications.

Each discussion of each application area has three parts. The first part presents a general discussion of the problem and the promise of the NII-SSEDM providing value to its solution. The second part focuses on a specific NII-SSEDM application area within the broader application area. These discussions are meant to be illustrative only. The ideas presented herein would need much additional thought before one could evaluate their potential value and begin to implement them. The third part discusses how the specific application relates to social science, education, and decision making.

3.1 ECONOMIC PROSPERITY

Economic prosperity is a central concern for this nation as we approach the 21st Century. Challenges to economic prosperity include: job creation and security; global economic competition; an aging workforce; lagging productivity; and skittish and/or insolvent financial institutions. The NII-SSEDM could provide access to socioeconomic databases of all kinds. The system could facilitate simple database queries, provide on-line access to databases and computing resources needed by social scientists developing and exercising complex socioeconomic models, and provide decision makers with the latest model forecasts and even with the ability to run models remotely. The system may have special 'visualization' servers to portray the results of various data analyses and simulations. The system may also provide the capability needed to improve the collection of timely economic data related to gross national product, balance of trade, price changes, and employment. Lastly, the system may even host numerous 'virtual' markets that could involve 'products' as diverse as household-generated municipal solid waste to independently produced electric power.

With respect to jobs, one could envision a national jobs database and decision support system as part of the NII-SSEDM. The database could have two components: one containing information on current jobs (e.g., from occupational classification databases and current job openings); and the other containing personal information on employment histories, skills, etc. The first component would be accessible to the

public at large. Records in the second component would only be accessible to those who created the records and in an anonymous fashion to others. The personal employment "accounts" could include information on current and past employment, references, resumes, and other personal information which people may want to store electronically and which could be made available to interested parties.

The national job database would be up-dated in real-time as employers and workers see fit. The decision support component could assist workers develop career plans, assess and evaluate their skills, and search for new jobs. The data analysis and modeling components could assist employers with workforce planning, training program development, and other employment issues. The system could also be used to disburse unemployment benefits, following the "One-Stop Shopping" concept. Lastly, the system could contain an educational component to teach people how to search for jobs, and why it is important to develop work-related/career plans.

Certainly, such a system would have the potential to help workers in the United States find and maintain jobs and develop satisfying work-lives. The system would also provide to social scientists a wealth of up-to-date cross sectional and time series data about employment, which could be analysed and used to develop models also available as part of the NII-SSEDM. Government employment offices could use the data, data analyses, and models as well as the policy analysis and decision support software to assess employment training programs, to study effects of government programs on reducing structural unemployment, and to evaluate unemployment benefit programs.

3.2 HEALTH CARE

Economic prosperity must be accompanied by improvements in the nation's health care system. Health care is more than the successful application of medical science. The problem has economic roots, e.g., the costs of drugs and insurance. The problem also has political roots, as nearly all solutions are the target of influential special interests. The problem has organizational roots—e.g., how should a hospital be run?—; sociopsychological roots—e.g., how can people of all types and ages be convinced to protect themselves from sexually-transmitted diseases, especially HIV?—; and geographical/urban planning aspects—e.g., how is one's health related to where one lives and works?—.

With respect to health care, a most important NII-SSEDM-related application would be a national medical database on symptoms, diagnoses, treatments, and outcomes, as well as on costs and other factors important in health care. The database would be created and up-dated in real-time as patients are initially examined, tests are given, diagnoses are rendered, treatments are prescribed, and outcomes are evaluated.

Similar to the jobs database mentioned above, the health database would have components open to the public as well as records made only accessible to treating physicians and their patients. Also, people may use the national medical database to conveniently store and access their personal medical histories on the system.

Data analysis and modeling software could be used to assess the cost effectiveness of procedures and treatments and be used as the source of new medical knowledge. This new knowledge, as well as existing knowledge in the forms of text books, expert systems, etc., could be provided on-line in several forms to assist teaching and medical practice, as well as to provide people with access to basic medical knowledge (e.g., through expert systems). The system could facilitate distance-diagnosis through the transmittal in real-time of medical images. The system could also improve the efficient handling of donated organs for transplant as well as provide decision support tools and on-line, "human" support to doctors to help them through ethically difficult treatment decisions. Lastly, the system could host a national health care transactions system that would be used by individuals to track their health care benefits and costs and options.

As with a jobs database, a health database would allow social scientists to study relationships between health and numerous variables (e.g., behavior, occupation, place of residence) and the cost of health care. The system could support difficult policy decisions regarding national health care, health insurance regulation, and other health-related laws. The system could also be an educational resource for schools (e.g., through simple statistical queries, access to materials, etc.).

3.3 CRIME AND VIOLENCE

Possibly the greatest challenge facing humanity in the post Cold War world is, ironically, peaceful coexistence, among nations, among communities, and between individuals. The world suffers from widespread ethnic and religious strife, civil wars, international terrorism, rising prejudice, rising levels of urban violence, social intolerance, violence against women, and family violence. To facilitate human evolution toward more peaceful forms of organization, values, and beliefs, every aspect of social science knowledge must be brought to bear and done so in a conceptually integrated fashion. The NII-SSEDM could assist in this goal.

In this country, progress is being made in developing a national information system to support law enforcement. National, state, and local law enforcement agencies are beginning to pool and share data and databases. An NII-SSEDM application could take these efforts much further to create valid and useful longitudinal databases on crime and violence. A national database could be created and updated in realtime to include: initial data from the scene and/or about the crime; progress on cases;

evidence; indictments and verdicts, with supporting reasons; information on admitted and convicted criminals, pooled by case and over time; punishment and/or other court mandated restitution; and special information surrounding the context of the crime (e.g., characteristics of location, community) and the perpetrator (e.g., education, exposure to media influences). The national database could also include information on victims as well as information collected as part of special longitudinal studies which could be merged with data collected as part of law enforcement and judicial activities.

Such a national database would provide a wealth of longitudinal information not now available to social scientists. The database could be combined with a breathtaking array of data on demographics, economics, voting behavior, education, environmental degradation, health, energy, transportation, etc. to provide the basis for fundamentally new insights about crime and violence in American society. Decision makers and policy analysts could use the tools provided by NII-SSEDM to evaluate prison programs, law enforcement strategies, sentencing guidelines, and many other questions, and even conduct on-line mediation activities.. The system could be a source of information to the public about crime and violence.

3.4 LIFE-LONG LEARNING

Learning is the linchpin of society. It is necessary for economic prosperity, essential for the operation of an effective and efficient health care system, and a key element underlying the peaceful coexistence of peoples. Education is life-long, spanning traditional classroom activities, distance learning, and retraining. Information technology has the potential to bring to teachers and students diverse subject matter in subjects as traditional as reading, writing, and mathematics and as advanced as astronomy, global climate change, and genetics. Technology can also be fashioned to provide intelligent instruction, find and coalesce materials, and facilitate spatially disperse learning communities. If these challenges were not enough, it must be realized that education has the potential to involve almost every U.S. citizen, most in more than one way.

K-12 education can be served by information technology in numerous ways. For example, through NII-SSEDM, a national educational assessment data system could be developed. The database could contain information collected in real-time on student performance and accomplishment by grade and subject matter, and accompanying data on textbooks used, teaching methods used, class size, homework assignments, class environment, etc. Again, the database could have public and private components. The latter refers to an electronic portfolio maintained by and for individual students that would only be accessible to the students, teachers, and parents and only by request otherwise.

The NII-SSEDM would also bring into the classroom resources from the Smithsonian Institution museums, Library of Congress, and any number of other places. The database could allow educational researchers, teachers, and administrators to study the effectiveness of teaching techniques, and overall levels of accomplishment. These findings, in turn, could support decision about school reform, funding, and strategic planning.

4.0 CREATING THE NII-SSEDM

To turn the vision of the NII-SSEDM into reality, it is necessary to develop a strategic plan and then establish specific, accomplishable goals that naturally flow from the plan. This paper has already addressed the topic of strategy: the NII-SSEDM is conceived as having five major components; a champion for the NII-SSEDM must be found; federal agencies, universities, other research institutions, and the private sector need to work together on the various technical challenges; and the development of the NII-SSEDM must be tied to important social problems (e.g., jobs, health care, crime and violence, and K-12 education). This section suggests specific goals whose accomplishment will help create the NII-SSEDM. It should be kept in mind that these goals are only initial ideas designed to contribute to more substantive discussions about the NII-SSEDM.

4.1 PRE-YEAR 2000 GOALS

- Incorporate into existing or create new federal research programs to address research topics necessary for the NII-SSEDM, such as those mentioned associated with the five technical areas discussed above.
- Work with a small set of large companies in a variety of industries (e.g., retail, phone service, travel) to make available for research purposes point of purchase databases.
- Prototype a system to allow rudimentary point of generation data collection with respect to all four application areas mentioned above: jobs, health care, crime, and K-12 education.
- Establish one to three national database server centers to provide access to numerous and major databases and computational resources as appropriate.
- Use existing technology to make several very large federal statistical databases available over the Internet (e.g., 1990 Decennial Census, Current Population Surveys taken in the 1990s).
- Make available as freeware user friendly, PC-based, Internet-oriented interface to the databases, record linkage software, and other resources accessible at the national database server centers and elsewhere.

□ Make available as freeware software to allow disabled individuals, people for whom English is a second language, and illiterate citizens to query important databases and other information system resources.

□ Develop prototype systems to allow those engaged in statistical analysis to ship their databases over the Internet to the most appropriate processing sites, as defined by computational resources, statistical software, and display software.

□ Create the jobs, health care, crime, and K-12 education databases mentioned above, which includes provisions for individuals to establish their own record keeping systems and to ensure the privacy of these personal records.

□ Establish one to three policy analysis/decision support resource centers, effectively networked to the database server centers.

□ Make accessible from these resource centers over the Internet a suite of economic and sociopolitical models possessing graphical and visualization capabilities for use by researchers and public and private sector analysts.

□ Make available as freeware user friendly interface system to these models.

□ Develop decision support systems, as mentioned above, in the areas of employment, health care, safety, and education.

□ Prototype a system that uses the Internet to facilitate mediation and collaborative decision making.

□ Make accessible over the Internet to the educational community resources from two or three Smithsonian Institution collections (e.g., National Zoo, Naturalist Center, American Museum of Art).

□ Identify enthusiastic champion and firmly establish the organizational structure needed to accomplish the tasks needed to create the NII-SSEDM.

□ Develop a detailed blueprint of organizational responsibilities, near- to mid-term funding and tasks, and RD&D activities that will compose the initiative. The blueprint should be consistent with goals for the first tier of the nation's information infrastructure and other activities under the broad rubric of the National Information Infrastructure Initiative.

4.2 POST-YEAR 2000 GOALS

- Make fully functional national job, health-care, crime, and K-12 databases, as well as other important databases, that are continuously up-dated via point of generation data collection technology and are easy to access for personal use, statistical analysis and modeling activities.
- Implement full-scale national economic point of purchase system (e.g., transactions from retailers could be randomly selected for inclusion in a database).
- Fully establish a national system of database server centers, with appropriate computational resources to allow access to and manipulation of sets of very large databases and real-time generation and viewing of complex visual outputs.
- Bring the great proportion of the nation's statistical databases on-line (e.g., by making electronically available historical data), along with materials that succinctly and precisely describe these databases in terms of quality.
- Make available as freeware software to custom-build databases for analysis.
- Fully establish a national system of policy analysis and decision support centers.
- Make accessible over the "Network" a primary health care expert system to allow citizens to better practice preventive health care and in-home, basic health care as well as fully functional employment/career decision support and expert systems.
- Establish prototype policy analysis/decision support systems for executives in the federal government. Further enhance the mediation and collaborative decision support systems.
- Make accessible over the "Network" for educational purposes substantial portions of the Smithsonian's collections, as well as those maintained by the Library of Congress, the National Library of Medicine, the National Oceanographic and Atmospheric Administration, the U.S. Geological Survey, and other agencies. Independently produced multi-media resources should also be made accessible.
- Implement a demonstration project that fully "wires" an entire school district for access to the "Network", integrates accessible resources comprehensively into the curricula, and provides students and teachers with highly portable, networked computer tools that allows the development of "electronic portfolios".
- Implement a sophisticated, intelligent, network-based "virtual" system to manage the growth and evolution of the NII-SSEDM.

5.0 CLOSING REMARKS

The time has come to begin to implement the vision of a national information infrastructure to support social science, public and private decision making, and education, herein termed the NII-SSEDM. The ideas presented above encompass a large number of organizational and technical challenges. However, the basic computer and telecommunications technology exists today to initiate a national program to serve the current and future generations of Americans.

One participant who commented on a draft of this paper remarked: "Why is it that while all of our problems are ultimately social, we are so much more receptive to research in the physical sciences?" The point is that social scientists, along with educators and those interested in improving government, feel that their contributions to society have been undervalued and underfunded, and that it is time to accept the larger challenge of assisting society help itself solve its major problems. Technology can serve as a nexus to bring together these communities and empower people to develop new, creative, and effective solutions to national and international problems.

The ideas contained in this white paper encompass every agency of government that conducts policy analysis, every organization that makes decisions, and every institution and individual engaged in health care educational activities. The moral costs in terms of jobs not created, continued reductions in the quality of life, increasing health care costs, substandard education, etc. associated with not pursuing a national information infrastructure program for the social sciences, public and private decision making, and education cannot be overstated.

The author is a research at Oak Ridge National Laboratory and current president of the Social Science Computing Association. This paper benefitted from the contributions of numerous individuals. Initial ideas for the paper were developed during the 1993 Conference on Computing for the Social Sciences, held at the National Center for Supercomputing Applications, University of Illinois, May 18-21. Contributions and comments were provided on drafts of this paper by Albert Anderson, Mike Bronzini, Fred Conrad, Sherwood Dowling, Joan Combs Durso, Steve Elliott, Bruce Fogarty, Rick Goeltz, Kerry Hake, Robert Hammond, Bruce Johnson, Mike Kuliasha, Alan Lomax, Don MacGregor, and Roberta Miller.

INTERNAL DISTRIBUTION

- | | | | |
|-------|----------------|--------|----------------------------|
| 1. | M. S. Bronzini | 47. | Central Research Library |
| 2. | J. B. Cannon | 48. | Document Reference Section |
| 3. | M. A. Kuliasha | 49-50. | Laboratory Records |
| 4. | D. E. Reichle | 51. | Laboratory Records - RC |
| 5. | R. B. Shelton | 52. | ORNL Patent Office |
| 6-46. | B. E. Tonn | | |

EXTERNAL DISTRIBUTION

- 53. Dr. Helen M. Ingram, Director, Udall Center for Studies in Public Policy, The University of Arizona, 803/811 East First Street, Tucson, AZ 85719.
- 54. Mr. Calvin D. MacCracken, President, Calmac Manufacturing Corporation, 101 West Sheffield Avenue, Englewood, New Jersey 07631.
- 55. Ms. Jacqueline B. Shrago, Director, Office of Technology Transfer, 405 Kirkland Hall, Vanderbilt University, Nashville, Tennessee 37240.
- 56. Dr. Martin Williams, Professor, Department of Economics, Northern Illinois University, DeKalb, Illinois 60115.
- 57. Office of Assistant Manager for Energy Research and Development, DOE/ORO, P.O. Box 2001, Oak Ridge, Tennessee 37831-8600.
- 58-60. Office of Scientific and Technical Information, U.S. Department of Energy, P.O. Box 62, Oak Ridge, Tennessee 37831.