# Information Infrastructure Technology and Applications: Report of the IITA Task Group

High Performance Computing and Communications Information Technology Subcommittee

*Dedicated to the memory of A. Nico Habermann*

CONTENTS

Information Infrastructure Technology and Applications
Report of the IITA Task Group
High Performance Computing, Communications and
Information Technology Subcommittee

Following the introduction of S. 2937, the Information Infrastructure and Technology Act of 1992, S. 4, the National Competitiveness Act of 1993, and H.R. 1757, the National Information Infrastructure Act, it became apparent to the HPCC agencies that the country would be committed to create a National Information Infrastructure and the applications that aid in its initial development.

In response to a suggestion by Dr. Nico Haberman, a task force of the HPCC agencies was formed for the purpose of drafting an initial statement of our endorsement for these ideas and an initial set of descriptions for the agencies' participation. After Dr. Haberman's untimely death, Drs. Mel Ciment, Randy Katz, and Y. T. Chien replaced him, and with other HPCC participants brought to the HPCCIT Subcommittee a final draft of their recommendations (on time) at its regular meeting on December 1, 1993.

Subsequently, many members and agencies have contributed ideas to completion of this initial plan. These have been included in the present document. The document has been accepted by the HPCCIT agencies as our initial plan for the Information Infrastructure Technology and

Applications component of our work. We expect to issue updates to this statement as our plans mature and as experience with the NII accumulates. All the agencies will welcome constructive suggestions on how best we can perfect this charge and accomplish the task.

Donald A. B. Lindberg, M.D.

Chair

HPCCIT Subcommittee

IITA Task Group
Melvyn Ciment (NSF), chair

Randy Katz (ARPA), YT Chien (NSF), co-editors

Sherri de Coronado (NIH), Ernest Daddio (NOAA), Steve Griffin (NSF), Dan Hitchcock (DOE), Paul Hunter (NASA), Selden Stewart (NIST)

Contributors
Howard Bloom (NIST), Roger Callahan (NSA), Steve Cross (ARPA), Norm Glick (NSA), Jerry Linn (NIST), Dan Masys (NIH), Cherie Nichols (NIH), Walter Shackelford (EPA), Doyle Weishar (ARPA), Gio Wiederhold (ARPA), Stephen Squires (ARPA)

## 1.1. Relationship between HPCC and NII

High Performance Computing and Communications (HPCC) and the National Information Infrastructure (NII) are tightly interwoven. The high performance computing and communications technologies developed under the HPCC Program provide much of the base of technologies critically needed for the National Information Infrastructure. The processing elements and servers of the NII will be founded upon the scalable computing technologies being developed under HPCC. The Program's networking component is driving the development of the technology for the high bandwidth data pipelines of the NII, while also examining diverse extensions of the network into radio, broadcast, and cable distribution mechanisms. The software technologies, including those for operating system software as well as compilers and development environments, being developed under HPCC have widespread application to complex, general-purpose distributed systems, not just those identified as "numerically intensive." And many of the Grand Challenge applications being supported under HPCC carry forward into the National Challenges articulated for the NII.

The HPCC Program's High Performance Computing Systems (HPCS) component, focuses on the development of high performance computing elements that can be arrayed and scaled to achieve higher levels of performance. These range from workstations to the most powerful supercomputers. At the heart of today's most powerful massively parallel machines are the same microprocessors to be found in relatively inexpensive desktop workstations and personal computers. And the high speed interconnection technology found in these scalable machines will form the basis of tomorrow's high speed communications network switches and processor cluster interconnections.

It has long been observed that today's highest performance computing technologies are tomorrow's desktop technologies, and today's desktop technologies will find their way into tomorrow's appliances. The same server technology needed to feed high performance computations with high data rate streams of information can be used to disseminate information to American households. For example, several of the emerging system architectures for "video on demand" entertainment services of the NII are based on massively parallel processors coupled to the kind of software- managed mass storage systems that have long been attached to supercomputers. HPCC HPCS activities continue to push forward the leading edge of high performance computing capabilities, while creating system architectures that continue to allow desktop and smaller computers to take advantage of the advancing technology.

The HPCC networking program (NETS) continues to advance networking technology, focusing on issues of interoperability and scalability, while providing new services for multimedia applications and increasingly higher levels of performance. As the distinction between channels, backplanes, and networks continues to blur, the focal point is scalable interconnection technologies. For example, the same components used in massively parallel processor backplanes are now being used as the basis for very high speed network switches for local area networks. In addition, NETS is examining alternative, more ubiquitous "on- ramps" to the emerging high bandwidth backbone of the NII. At the Monte Vista High School in Cupertino, CA, Vice President Gore observed a class of students directly accessing the Internet over the local cable TV system. This technology was developed under HPCC sponsorship.

Much of the activity in the Advanced Software and Technology (ASTA) component of the HPCC Program is directed toward the general issues of developing new programming languages, compilers, reusable module libraries, object management systems, software development environments, and operating systems software for a diverse collection of distributed system and machine architectures. While NII applications will extend the scope of these efforts to new domains, much of the core software technologies directly support the kinds of information-intensive processing that characterizes the National Challenge applications.

Solutions to the Grand Challenges of science and engineering, as undertaken by the HPCC Program, will play key roles in supporting the societal challenges of the NII's National Challenges. Many of the National Challenges will depend upon core computationally-intensive capabilities. For example, design and manufacturing applications depend critically upon the ability to prototype new products and new processes using numerically-intensive techniques.

Similarly, a crisis management system must be able to invoke a weather prediction application as part of its support for disaster planners.

Recognizing the close relationship between HPCC and NII, the President's Science Advisor, Dr. Jack Gibbons, and the Office of Science and Technology Policy (OSTP), have instructed the High Performance Computing, Communications, and Information Technology (HPCCIT) Subcommittee to establish a new program component within the HPCC Program to support the Administration's initiative to build a National Information Infrastructure. This component, Information Infrastructure Technology and Applications (IITA), describes a research and development program to create the technology base underlying a universally accessible National Information Infrastructure (NII) and to use this technology to develop and demonstrate prototype "National Challenge" applications. These National Challenges will help to extend the benefits of HPCC technology to applications with broad impact on American society.

## 1. 2. Role of the Federal Government in Information Infrastructure

Several issues have arisen in the discussion of the National Information Infrastructure since the subject was first raised by then-Senator Al Gore, Jr. during the 1992 Presidential campaign. U.S. industry was at first concerned that the infrastructure would be created and run by the Federal government, counter to the long tradition in the United States of private sector ownership of telecom- munications systems. Over the course of time, the government's role has been clarified: to focus on the crucial issues of information and telecommunications policy, to simultaneously encourage private sector development of new information and communications technologies, and to protect citizens' rights of access and privacy in the electronic age, balanced with law enforcement and national security concerns.

Moreover, the Federal government has a critical role to play to reduce the risk in creating a National Information Infrastructure, while ensuring that the technology will be cost effective and appropriate for developing implementations to the National Challenge applications. Information infrastructure technologies will play a critical role in the Federal government's own plans to re-engineer its work processes. In a recent articulation of the government's role in information infrastructure, Vice President Gore has drawn an analogy between the NII and the first use of telegraphy in a limited government-sponsored demonstration (between Washington and Baltimore), subsequently followed by widespread private sector deployment on a national scale.

The HPCC-sponsored gigabit testbeds offer an attractive model for government- industry-academic research partnerships, and represent a strategic resource upon which to build prototype implementations of the National Challenge applications. Each testbed, cost-shared between government and the private sector, brings together companies from the computer and telecommunications industries, research groups from universities and the National Laboratories, and applications developers from the leading centers of high performance computing. A critical function of the testbeds is to experiment with high risk networking technology while driving the early-stage development of interoperability.

The gigabit testbeds will continue to examine the next generation of high speed networking technologies, while also integrating more diverse communications modes like satellite and cable distribution. However, some portion of the gigabit testbeds will be transitioned into a more operational mode (where these capabilities are not duplicated by commercially available services) to support advanced National Challenge applications prototyping teams. These will provide a crucial research tool with which to experiment with new kinds of information services, intelligent interfaces both for users and machines, and new systems development environments. As the HPCC Program expands to encompass new kinds of applications challenges, a critical need will be to nurture interdisciplinary National Challenge applications prototyping teams with access to the network testbeds.

Will government in general, and the HPCC Program in particular, implement the National Challenge applications? In some cases, such as environmental monitoring, the answer is clearly yes. Because of the risks involved or the government's being the only customer, the private sector will not undertake such development efforts. However many of the National Challenges, from health care to education and training to design and manufacturing, will require very substantial investments by the private sector. Government's first role is to be the patient investor underwriting the development of early technologies. Its second role is to foster interoperability at an early stage of development by sponsoring extensive testbed demonstrations with the private sector.

## 2. 1. Introduction

While entertainment applications will drive the connection of homes to the National Information Infrastructure, Vice President Gore has set the additional ambitious challenge of interconnecting all schools and libraries by the turn of the century. But other than "video on demand," what kinds of things will the average American citizen be able to do with this "information infrastructure" when it becomes a reality? This section provides example descriptions of how the infrastructure will be used in the 21st century, by the common citizen as well as scientists, engineers, and business people.

## 2. 2. Electronic Commerce

Business and government experience with online electronic commerce services (for example, the FAST system for electronic parts, a service used by Department of Defense laboratories for ordering components) has demonstrated rather convincingly that electronic commerce can be cost effective and can dramatically reduce the time to purchase items (even very large quantities). The popularity of cable services, such as home shopping channels, has prepared the Nation by providing a limited form of electronic commerce via television distribution, telephone "uplink," and credit cards. In the near future, every home and business could have an information

appliance that combines the capabilities of telephone, television, newspaper, computer, and Internet services such as electronic mail.

To see some of the possibilities, consider the following two scenarios describing how consumers might use electronic commerce by the end of the century:

> A couple realizes that their van has about 45,000 miles on it and it is time to replace the tires. Rather than scanning the ads in the Sunday morning sports page, they ask their computer to retrieve multimedia tire ads. These ads contain useful information about the companies' consumer services (hours, loaner cars, payment plans, etc. ). In this case, the consumer asks the computer to display on a map those companies that have 215X75R tires on sale with 60,000-mile warranties, are located within 5 miles of their home, have hours that extend to at least 8 p.m., and offer shuttle service to and from the Metro. These are displayed on a computer-generated map and the consumer selectively examines the electronic advertisement of each offering. The consumer then issues an electronic bid to one or more of these companies. This in turn would be automatically answered by the tire companies' electronic commerce service. The consumer then picks the best offer (based in part on third party "consumer reports" about tire quality and customer satisfaction), and upon acceptance, activates a process that would reserve (or order) the tire from inventory, schedule an appointment, and reserve other services requested (such as a shuttle to work). Payment could be arranged electronically as negotiated in the bid (for example, in advance, upon performance, or upon inspection by the consumer).
>
> Electronic commerce is an excellent example of "dual-use" capabilities. A military logistics agency increases its performance and decreases its costs, while simplifying its inventory control, by exploiting the services of an electronically available inventory service. This service downloads daily logistics requirements (orders) and predicts demand based on consumer models to determine inventory on-hand requirements. Through electronic brokering capabilities, the inventory provider secures "just in time" inventory at the best cost and provides the material to the logistics agency.

To make electronic commerce readily accessible and frequently used, it must be available, convenient, and inexpensive. Availability will be enabled by services that support the publication, dissemination, and access of multimedia information. Convenience will be enhanced by intelligent services that allow inexperienced users to browse information spaces with a combination of speech and graphics and to delegate tasks associated with brokering to automated agents. Expenses will be mitigated by the large number of consumers and providers that will use electronic commerce and by the savings that they will realize.

## 2. 3. Digital Libraries

The remarkable expansion in the generation and dissemination of digital information in the last decade as a result of the availability of high speed electronic networks has dramatically changed

the nature and role of data archives and traditional libraries. Since the mid-1980s, information sources accessed via the Internet have multiplied rapidly. These include a mixture of data and knowledge sources in all electronically available forms: reference volumes, books, journals, newspapers, national phone directories, sound and voice recordings, images, video clips, scientific data as well as private information services such as stock market reports and private newsletters. These knowledge sources, when connected electronically through a network of networks, are the ingredients of a digital library.

A digital library is literally a library "without walls" -- a knowledge resource open 24 hours a day, accessible via the network. To explore the full benefits of such digital libraries, the challenge for research and development is not merely how to connect everyone and everything together in the network. Rather, it is to achieve economically feasible technologies with which to digitize massive corpora of existing and new information from heterogeneous and distributed sources; then store, search, process and retrieve the information in a user friendly way. Among other things, this will require both fundamental research and development of "intelligent" software and user interfaces; ubiquitous network services for reliable and secure information access; and a new generation of database technologies for managing multimedia, cross-cultural, and interorganizational complexities of knowledge explosion. Equally important are mechanisms to ensure the maintenance (and potential payment for use) of intellectual property. One could image a small fee being charged to each reader of a "fee for access" document published on an open network.

Consider the following scenario. A high school student has been given the assignment to produce a report on folklore of the Middle Ages. The student uses the network to connect to college libraries across the country, and to search the bibliographical indexes for books and articles about the period and the topic she is interested in. Focusing her report on the myths surrounding death in rural England in the early Middle Ages, she filters the enormous amount of information available to a reasonably manageable list of references. She accesses these books directly over the network, viewing them on her computer screen while creating a multimedia report that contains linkages back to the underlying sources. These sources, in turn, have linkages to primary sources, such as an online copy of the famous "Doomsday Book" of English history. She hands her report in to her instructor, who views the report on his own computer. This allows him to check her references directly. For her thorough job, she gets an A on the report.

### 2. 4. Crisis Management

Crisis management involves the use of command, control, communications, and intelligence information systems to support decision makers in anticipating crises, formulating plans, and executing these plans through coordinated responses. Recent relief efforts in the aftermath of Hurricane Andrew and the Los Angeles earthquake are compelling demonstrations of the need for a National Information Infrastructure. While responsible government agencies had developed disaster relief plans in the event of a major disaster, there was not a readily available information infrastructure to communicate the plan, or to coordinate the myriad of relief agencies that flocked to the disaster sites. For example, one conclusion in a post Hurricane Andrew report was the need to have federal information systems that could visually display "areas of need" on

electronic maps that could be shared with the broadcast media. The need for such information systems was further demonstrated during the L.A. quake relief efforts.

Now consider the future. Satellite and mobile communications will ensure "anytime, anywhere" reliable communication infrastructure upon which to base an information infrastructure. Crisis management command and control systems can then be designed and readily tailored to meet the needs of the crisis at hand. Weather center models together with continuously interpreted overhead imagery can more accurately predict crisis locations and the degree of damage. More reliable and faster situation assessment will be accomplished through a combination of intelligent agents, image understanding, and language understanding systems that will assess and aggregate information from a diverse set of situation reports (telephone, radios, etc. ). Decision-aiding models and "lessons learned" histories from similar past crises will aid in refining crisis plans, predicting resource needs, and defining precedent constraints for achieving those needs. For example, in the context of a serious earthquake where heavy cranes are required to help remove building debris, such a service may remind the crisis planners to plan for the insertion of portable power units (the failure to do this in the recent Mexican earthquake resulted in a loss of many lives).

Rapidly assembled response teams, consisting of Federal, state, and volunteer organizations, will be able to share and update electronic plans, collaborate during the execution of the plan, and use that plan as a basis for real-time training. Crisis planners would also be able to "collaborate through time" by comparing planned actions with historically relevant plans and through advanced simulation services. The underlying services that are needed include: collaboration and interaction, intelligent agents to monitor communications and provide relevant summaries and advisories, "plug and play" and easily tailorable decision aids, interoperable maps and other electronic communications media, simulation services to predict the effectiveness of the crisis response plan, and access and navigation through historical libraries.

## 2. 5. Advanced Manufacturing

American manufacturers seek to recapture world leadership and respect. The specific technical goals are to exploit lean manufacturing (e.g., greater efficiency and lower cost), flexibility (e.g., variation in production runs to allow for consumer preferences), and agility (e.g., supporting small production runs, rapid retooling, and exploitation of the kinds of electronic commerce services described in Section 2.2).

In addition, companies will be able to band together to jointly manufacture goods. This will require rapid tailoring and composition of shared information services such as inventory control, work scheduling, and product delivery. Potential machine tool vendors and other manufacturing support companies will be willing to provide simulations of new process-control and planning software to enable companies to test before they buy. In addition, software specialty companies will provide access to powerful computer-aided design tools that are currently too expensive for purchase by small companies. This is economically viable because a small company can access both the software and the human expertise that lies behind it, through coordinated on-line consulting services.

## 2. 6. Education and Training

Making HPCC resources a major universal resource for the Nation's education, training, and learning systems for people of all ages, abilities, and locations, is a critical goal of the National Information Infrastructure. It is well known that the most effective learning takes place at the "application" level - that is, "learning by doing. " The National Information Infrastructure will make available a wide variety of simulation-based education and training services that can be accessed through home or industry telecommunication systems. These simulation-based training systems will be tailored to a specific domain and will allow a student to live in that domain. The system will present a variety of skill-based and cognitive-based problems, provide coaching and remedial instruction, and tailor follow-on assignments to address student weaknesses. Fundamental to this vision are authoring tools, digital libraries, associate systems, and the ability to collaborate across space and time.

Consider the concept of a National Virtual Laboratory (NVL), concentrating on university research to enable geographically distributed researchers to share experimental results, collaborate on team research projects, and share coursework. It is envisioned that integrated research teams will share ideas, software, and students - that is, students at many locations may receive credit for NVL-based courses.

## 2. 7. Environmental Monitoring

While sitting at her scientific workstation at the Western Regional Environmental center, a biological oceanographer wants to explore the possibility that observations from an ongoing cruise in Alaska's Shelikof Strait support the hypothesis that a recent El Nino-Southern Oscillation (ENSO) event may be influencing a major pollock fish recruitment off Kodiak Island. During the last three months, data from satellite-linked moored subsurface ocean sensors, polar-orbiting satellites, numerical forecasting models, and coastal sampling stations have indicated that warm tropical water has progressively moved northward along the U. S. and Canadian Pacific coast. Now, real-time data from satellite-linked drifting buoys and current meter arrays, as well as ship surveys of fish egg and larvae show that a major circulation anomaly is occurring in the spawning region off Shelikof Strait. What looked like a high pollock catch year seems to be turning into an economic bust!

Using a mouse, she clicks a satellite icon that downloads the most recent 24-hour set of visible and infrared imagery from NOAA and NASA ocean-observing satellites. After clicking another icon, a new window opens that allows her to extract and compile a time series of the last 30 days of infrared imagery from the NASA satellite database in Pasadena, CA and the NOAA satellite database in Suitland, MD. Using tools available on the NII, she is able to integrate these heterogeneous data sets and to create an animation that clearly shows the progression of warm tropical waters toward the Alaskan coastal region that serves as a safe nursery for pollock larvae. With another icon selection, she overlays several tracks of drifting buoy data and subsurface moored current meter data onto the satellite imagery. These data suggest that the planned cruise track to study the new circulation feature will have to be modified to properly sample the feature.

To plan for a new cruise track, she checks the on-line National Weather Service five-day weather forecast and the Navy's Pacific Ocean Circulation Model forecast for the study region.

In order to put this environmental event into perspective, she queries the National Oceanographic Data center ship observations data base in Washington, DC for the past 50 years of sea surface temperature data. Using a graphical user interface available on the NII that lets scientists intuitively explore and visualize a variety of multi-dimensional data products, she zooms in on the North Pacific and quickly creates a 50-year animation loop of sea surface temperature. In another window of her workstation, she accesses the ENSO and Equatorial Undercurrent oceanographic database at the University of Washington. She creates a new graphical overlay of her data that suggests a long-term, phased relationship between equatorial processes and sea surface temperature in the northern Pacific. She confirms this visual observation by running correlation and coherence analyses using a point-and-click time series package, available on the network, from the Scripps Institution of Oceanography in La Jolla, CA.

2. 8. Health Care

While vacationing in Orlando, a Washingtonian slips and breaks his leg. While being taken to the Emergency Room, the ambulance relays ahead a download request for his universal patient care record. This electronic record contains his complete medical history, but only restricted fields are available to the ER doctors. One crucial field describes the medications he is allergic to.

By the time he arrives, his health maintenance company has already been notified and has approved the procedures for setting his broken leg. All he needs to do is sign an electronic form; all of the other relevant information (medical history, insurance coverage, home address, contact, etc. ) has been automatically filled in.

He is wheeled into the X-ray room, where the X-ray is taken and immediately rendered onto a high resolution display. A copy of the X-ray is automatically forwarded to his primary care physician back in Washington.

Since the fracture is fairly straightforward, he is next moved to the operating room. The room is filled with wall-sized displays, reporting his vital signs and the read- outs of the various medical equipment. After placing the patient under anesthetic, the doctor quickly sets the fracture.

Next step is the recovery room and his hospital room. Wearing a small device on his wrist, the patient's vital signs are relayed to the hospital's information infrastructure, where his "chart" can be viewed by his physician on any of the displays within the hospital or even at the office of his doctor back in Washington.

The next day, he is released, and his hospital and medical bills already have been settled by his insurance company.

## 2. 9. Public Access to Government Information

The government collects an enormous amount of information, such as trade data, environmental data, reports, and white papers, yet it is often difficult for the average citizen to gain access to such data. The National Information Infrastructure will make it much simpler for U. S. citizens to gain access to information they have already paid for.

Consider the following. A farmer contemplates purchasing a new spread of land. He can access a land resources database that gives him information on the mineral content and vegetation to be found on the land he is considering. He can also trace the water rights, to determine how much it will cost to irrigate his new land. He can find out what kind of crops have been grown in the land around his, and whether the land is near any hazards, such as flood plains or toxic waste dumps. This kind of information is available in a variety of different local, state, and Federal government databases. The challenge is integrating the diverse data sources in an unambiguous way while making the access to it as easy to use as possible.

## 3. 1. Overview

Information Infrastructure Technology and Applications (IITA) activities will demonstrate feasible solutions to problems of national importance, such as health care or 21st century manufacturing, using the full potential of the Nation's rapidly evolving high performance communications and information processing capabilities.

The High Performance Computing and Communications (HPCC) Program will produce enabling technologies critical to developing the National Information Infrastructure. With the incorporation of IITA activities, the HPCC Program will advance intelligent system interfaces, real environments augmented with virtual environments, image understanding, language and speech understanding, intelligent agents aiding humans "in the loop," and next generation data and object bases for electronic libraries and commerce. This will be coupled with a vigorous program of testbed experimentation that will ensure the continued leadership of the United States in critical information processing technologies.

IITA efforts are designed to strengthen the HPCC technology base, broaden the markets for these technologies, and accelerate industry development of the NII. Federal HPCC agencies will work closely with industry and academia in pursuit of these objectives. This will be accomplished in part by accelerating the development of readily-accessible, widely used, large-scale applications with significant economic and social benefit. The HPCC Program's original focus of enhancing the Nation's computing and communications capabilities is thus extended to address a broader set of technologies and applications that have an immediate and direct impact on critical information capabilities affecting every individual in the Nation.

The development of such applications will be predicated on (1) creating the underlying scalable computing technologies for advanced communication services over diverse bitways, effective

partitioning of applications across elements of the infrastructure, and other applications support services that can adapt to the capabilities of the available infrastructure; and (2) creating and inserting an intelligent service layer that will significantly broaden the base of computer information providers, developers, and consumers while reducing the existing barriers to accessing, developing, and using advanced computer services and applications. In parallel with these activities, a more effective software development paradigm and technology base will be developed. This will be founded on the principles of composition rather than construction, solid architectures rather than ad hoc styles, and more direct user involvement in all stages of the software life cycle. The entire technology base developed in this program, including services and software, will be leveraged across the National Challenges, leading to significant economies of scale in the development costs.

The underlying elements of the IITA program component consist of the following four broad topic areas:

*Information Infrastructure Services*

These are the collection of services provided to applications developers and end- users that implement a layered architecture of increasing levels of intelligence and sophistication on top of the communications bitways. Services provide a universally available, network-aware, adaptive interface upon which to construct the National Challenge applications, spanning communications-based services at the low end, to intelligent information processing services at the high end. These services include network support for ubiquitous access, resource discovery in a complex distributed network environment, and intelligent support services that can negotiate and adapt to the service quality needs of the application. Information Infrastructure Services also include system software and services that implement pervasive privacy, security and trust mechanisms for the information infrastructure, persistent object bases with which to build large-scale data repositories, reliable computing technologies to support the mission-critical nature of the infrastructure, and defensive software organized to protect the infrastructure from intrusion and attack.

*Systems Development and Support Environments*

This subprogram area consists of the enabling technologies to develop and support large, complex information systems that exploit a national-scale information infrastructure. Fundamental to the subprogram is the use of that infrastructure in the software development and support process. Virtual organizations consisting of end-users, contractors, and management will synergistically work together to develop software systems that are easy to use, that can be adapted through use to fit human needs and changing requirements, and that enhance end-user productivity, all in spite of the complexity of the underlying infrastructure. To achieve these goals, the program will focus on software architectures, component prototyping, software composition, libraries of reusable and reliable software modules, end-user tailoring, intelligent documentation and on-line help, machine learning, and scalable compiler and interpreter technology.

*Intelligent Interfaces*

Many of the National Challenge applications require complex interfacing with humans or intelligent control systems and sensors. In addition, these applications must be able to understand their environment and to react to them. This program area consists of high-level, network-capable applications building blocks for real- time planning and control, image processing and understanding, human language technology, extensive use of intelligent computer-based "agents," and support technologies for more effective human-computer interaction.

*National Challenges*

These are large-scale, distributed applications of high social and economic impact that have been identified as containing an extensive information-processing component and which could benefit greatly by building an underlying information infrastructure. Specific applications include Digital Libraries, Manufacturing, Education and Training, Health Care (Trauma Care, Medical Information Infrastructure), Environmental Monitoring, Crisis Management, and Government Information Delivery.

These elements are described in more detail in the following subsections.

## 3. 2. Information Infrastructure Services

Services provide the underlying network-capable building blocks upon which the National Challenges can be constructed. They will form the basis of the ubiquitous "Information Web" of the 21st century. Services bridge the gap between the communications bitways and the application-specific software components that implement the National Challenges.

- Universal Network Services: Extensions to the existing Internet technology base will provide more widespread use by a much larger number of users. These include techniques for improved ease-of-use, "plug and play" network interoperation, remote maintenance, exploitation of new "last mile" technologies like cable TV and wireless, management of hybrid/asymmetric network bandwidth, guaranteed quality of service for continuous media streams, and scale-up of network capabilities to dramatically larger numbers of users.
- Integration and Translation Services: These services support the migration of existing data files, databases, libraries, and programs to new, better integrated models of computing, such as object-oriented systems. They also provide mechanisms to support continued access to older "legacy" forms of data as the models evolve. Included are services for data format translation and interchange as well as tools to translate the access portions of existing programs. Techniques include "wrappers" which surround existing elements with new interfaces, integration frameworks which define application-specific common interfaces and data formats, and "mediators" which extend generic translation capabilities with domain knowledge-based computations, permitting abstraction and fusion of data.

- System Software Services: These include operating system services to support complex, distributed, time-sensitive, and bandwidth-sensitive applications such as the National Challenges. The services support the distribution of processing across processing nodes within the network, the partitioning of the application logic among heterogeneous nodes based on their specialized capabilities or considerations of asymmetric or limited interconnection bandwidth, guaranteed real-time response to applications for continuous media streams, and storage, retrieval and I/O capabilities suitable for delivering large volumes of data to very large numbers of users. Techniques include persistent storage, programming language support, and file systems.
- Data and Knowledge Management Services: These services include extensions to existing database management technology for combining knowledge and expertise with data. These include methods for tracking the ways in which information has been transformed. Techniques include distributed databases, mechanisms for search, discovery, dissemination, and interchange, aggregating base data and programmed methods into "objects," and support for persistent object stores incorporating data, rules, multimedia, and computation.
- Information Security Services: These services provide support for the protection of the security of information, enhanced privacy and confidentiality for users of the infrastructure, the protection of intellectual property rights, and the authentication of information sources within the infrastructure. Techniques include privacy-enhanced mail, methods of encryption and key-escrow, and digital signatures. Also included are techniques for protecting the infrastructure, such as authorization mechanisms and firewalls against intrusion attacks such as worms, viruses, and trojan horses.
- Reliable Computing and Communications Services: These include system software services for non-stop, highly reliable computer and communications systems that can operate for 7 days a week and 24 hours a day. The techniques include mechanisms for fast system restart such as process shadowing, reliable distributed transaction commit protocols, and event and data redo logging to keep data consistent and up-to-date in the face of system failures.

### 3. 3. System Development and Support Environments

These will provide the network-based software development tools and environments needed to build the advanced user interfaces and the information- intensive National Challenges themselves.

- Rapid System Prototyping: These consist of software tools and methods that enable evolvable software. These consist of the tools and methods that enable the incremental integration and cost effective evolution of software systems. Technologies include tools and languages that facilitate end-user specification, architecture design and analysis, component re-use and prototyping; testing and on-line configuration management tools; and, tools to support the integration and interoperation of heterogeneous software systems.
- Distributed Simulation and Synthetic Environments: These software development environments provide the specialized underlying support mechanisms for the creation of

synthetic worlds, which can integrate real as well as virtual objects, in terms of both their visual as well as computational descriptions. Methods include distributed simulation algorithms, geometric models and data structures, tools for scene description, creation, and animation, and integration of geometric and computational models of behavior into an integrated system description.

- Problem Solving and System Design Environments: These environments provide the techniques that support the software and system design process through the use of automated tools, with particular emphasis on maintaining flexibility and tailorability in tool configurations to enable organizations to tailor their support environments to their needs. Examples include efficient algorithms for searching huge planning spaces, more powerful and expressive representations of plans, operators, goals, and constraints, and the incorporation of efficient methods to facilitate scheduling and resource allocation. The effects of uncertainty must be taken into account as well as the effects of goal interactions.
- Software Libraries and Composition Support: These software tools and methods support the development of common architectures and interfaces to increase the potential for reusability across multiple underlying models of computation, the diversity of programming languages in use, and the varying degree of assurance provided by software components. Important elements of this area include the development of the underlying methodology, data structures, data distribution concepts, operating system interfaces, synchronization features, language extensions, and other technology to enable the construction of scalable library frameworks.
- Collaboration and Group Software: These tools provide support for group cooperative work environments that span time as well as space. Methods include shared writing surfaces and "live boards," version and configuration management, support for process and task management, capture of design history and rationale, electronic multimedia design notebooks, network-based video conferencing support, document exchange, and agents serving as intermediaries to repositories of relevant multimedia information. The technology should be developed to make it possible to join conferences in progress and to be automatically brought up to date by assistants (agents) with memory.

### 3. 4. Intelligent Interfaces

Advanced user interfaces will bridge the gap between users and the future National Information Infrastructure. A wide range of new technologies that adapt to human senses and abilities must be developed to provide more effective human-machine communications. The IITA program must achieve a high level user interface to satisfy the many different needs and preferences of vast numbers of citizens who interact with the NII.

- Human-Computer Interface: This subelement supports work in a broad range of technologies and their integration to allow humans and computers to interact effectively, efficiently, and naturally. Work in this area includes development of technologies for speech recognition and generation; graphical user interfaces that allow rapid browsing of large quantities of data; user- sensitive interfaces that customize and present information for particular levels of understanding; language corpora for experimental research; and human-machine interaction via touch, facial expression, gesture, etc. The new IITA

emphasis is on integration, real-time performance, and demonstration of these new communication modalities in multimedia, multi-sensory environments.

- Heterogeneous Database Interfaces: This subelement supports development of methodologies to integrate heterogeneously structured databases composed of multi-formatted data. To support NII information dissemination, a user will be able to issue a query to be broadcast to the appropriate databases. A timely response can then be returned and translated into the context of the user's query. Multi-formatted data may range from ASCII text to numerical time series, to multi-dimensional measurements, to time series of digital imagery, etc. Also of critical importance is the integration of metadata with the data and its accessibility across heterogeneous databases.
- Image Processing and Computer Vision: This subelement supports work in making images, graphics, and other visual information a more useful modality of human-computer communication. Research areas include all aspects of theory, models, algorithms, architectures, and experimental systems from low-level image processing to high level computer vision. Methodologies of pattern recognition will be further developed to allow automated extraction of information from large databases; in particular, digital image data. The new IITA emphasis is on integration, scalability, and demonstration of easy access and usability of visual information in real-world problems.
- User-centered Design Tools/Systems: This subelement consists of work in models and methodologies leading to interactive tools and software systems for design and other user-centered activities. User-friendly tools that combine data-driven and knowledge-based capabilities is one of the areas for new research. The new IITA emphasis is on supporting the development of ubiquitous, easy-to-use, and highly effective interactive tools.
- Virtual Reality and Telepresence: This subelement supports work that will provide tools and methods for creating synthetic (virtual) environments to allow real-time, interactive human participation in the computing and communication loop. Such interaction may be through sensors, effectors, and other computational resources. The IITA focus is creating shared virtual environments which can be accessed and manipulated by many users at a distance in support of National Challenge application areas.

## 3. 5. National Challenges

National Challenges are information-intensive applications that have broad and direct impact on the Nation's well-being and competitiveness. The IITA program activity will support the use and integration of component technologies developed in other parts of HPCC to seek solutions for such applications. It will run the gamut of projects and programs ranging from concept demonstrations and experimental testbeds to the actual delivery of application systems in specific domains.

- Digital Libraries: This subelement includes work in both technologies and applications which will lead to significant advances in the generation, storage and use of digital information of different kinds across high speed networks. A digital library is a knowledge center without walls, open 24 hours a day and accessible by way of a

network. Research areas range from advanced mass storage, on-line capture of multimedia data, intelligent filtering, knowledge navigation, effective user interfaces, system integration, and prototype and technology demonstration.

- Electronic Commerce: Electronic commerce integrates communications, data management, and security services, to allow business applications within different organizations to automatically interchange information. Communications services transfer the information from the originator to the recipient. Data management services define the interchange format of the information. Security services authenticate the source of information, verify the integrity of the information received by the recipient, prevent disclosure of the information to unauthorized users, and verify that the information was received by the intended recipient. Electronic commerce applies and integrates these infrastructure services to support business and commercial applications including financial transactions such as electronic bidding, ordering and payments, and exchange of digital product specifications and design data.
- Advanced Manufacturing: This subelement supports work in advancing manufacturing technologies through the use of HPCC capabilities in design, production, planning and quality control, marketing and user services. Research areas include concurrent engineering, protocols for electronic exchange of product data, and virtual design technologies. The IITA focus will be on testbeds and technology demonstration on the Internet.
- Education, Training, and Lifelong Learning: The subelement supports work in making high performance computing and communications capabilities a major universal resource for the Nation's education, training, and learning systems at all levels and in all environments. Areas of work include innovations in providing network access and conducting pilot projects that demonstrate computing and communications technologies for improving learning/training that can be scaled up.
- Health Care: This subelement supports work in developing the concepts and technologies for applications of HPCC in the health sector. Such applications make use of high speed networking capabilities for linking health care resources of all kinds to support research and shared care delivery. It also supports novel use of emerging technologies, e.g., visualization, virtual reality, 3D imaging, and multimedia databases. The IITA focus will be on the use of networking and other HPCC capabilities in medicine and health care delivery.
- Crisis Management: Crisis management refers to the use of command, control, communications, and intelligence information systems to support decision makers in anticipating threats, formulating plans, and executing these plans through coordinated response. This subelement will support work in crisis management that deals with a variety of large-scale, time- critical, resource-limited problems, including nuclear monitoring, disaster operations, and riot control. The IITA focus will be the use and demonstration of ubiquitous information services and other HPCC technologies in dealing with crisis management.
- Energy Management: This subelement supports work in managing the Nation's energy resources -- supply and demand -- in an increasingly complex global environment. The IITA focus will be the use of HPCC assets and capabilities to demonstrate ubiquitous benefits (e.g., energy savings) to the public from effective energy management.

- Government Information Delivery: The subelement supports work in implementing new programs to develop and apply high performance computing and high speed networking technologies to vastly improve public access to information generated by Federal, state, and local governments. Work includes projects to connect agency depository libraries and other sources of government information to the Internet to enable public access; and to demonstrate, test, and evaluate technologies to increase such access and effective use of government information. The IITA focus in the near term will be to provide cross agency technology demonstrations and in the long run to develop information delivery systems tailored to agencies.
- Environmental Monitoring: Environmental monitoring relates to our ability to observe, understand, and predict changes in the environment, both natural and man-made, on scales from seconds to millennia and from local to global. Efforts in this IITA National Challenge area will focus on applying high performance computing to advance environmental understanding and predictive capabilities, and on integrating multi-agency environmental observing systems (land-based, ocean-based, or satellite-borne) and historical databases using high speed networks. The NII will tie together environmental observing systems, computational resources, historical databases, libraries, scientists, forecasters, and all potential consumers of environmental information to support research, transportation, commerce, education, policy making, emergency preparedness, recreation, and day-to-day activities. This will provide ready access to the best possible, most up-to-date and comprehensive information needed to address the range of environmental questions, whether they be as mundane as "Will it rain this afternoon?" or as profound as "Is the ozone hole intensifying?"
- Other Category: NSF Discipline Network centers. This subelement of NSF National Challenge activities focuses on the establishment of discipline applications centers in selected areas of the physical, biological, and geological sciences. It also includes the development of a network of discipline-specific networks. The new IITA emphasis is the use of the high speed network and high performance computers to enhance collaborative research among discipline scientists in both academia and industry.

Bibliography

President William J. Clinton, Vice President Albert Gore, Jr., Technology for America's Economic Growth, A New Direction to Build Economic Strength, February 22, 1993.

Computer Systems Policy Project, Perspectives on the National Information Infrastructure: CSPP's Vision and Recommendations for Action, January 12, 1993.

Al Gore, Jr., "Infrastructure for the Global Village," Scientific American, (September 1991), V. 265, N. 3, pp. 150-153.

Information Infrastructure Task Force, The National Information Infrastructure: Agenda for Action, September 15, 1993.

Information Technology Association of America, "National Information Infrastructure: Industry and Government Roles," July 1993.

Institute for Information Studies, A National Information Network: Changing Our Lives in the 21st Century, Aspen Institute, Queenstown, MD, 1992.

Brian Kahin, "Information Technology and Information Infrastructure," in Empowering Technology: Implementing a U. S. Strategy, L. M. Branscomb, ed. , MIT Press, Cambridge, MA, 1993.

Ed Krol, The Whole Internet User's Guide and Catalog, O'Reilly & Associates, Inc. , Sebastopol, CA, 1992.

J. Motiwalla, M. Yap, L. H. Ngoh, "Building the Intelligent Island," IEEE Communications Magazine, V. 31, N. 10, (October 1993), pp. 28-34.

National Computer Board of Singapore, A Vision of an Intelligent Island: The IT2000 Report, March 1992.

Office of Science and Technology Policy, FCCSET Initiatives in the FY 1994 Budget, April 8, 1993.

## A. 1. Advanced Research Projects Agency

The Advanced Research Projects Agency (ARPA) is the lead agency for developing advanced dual-use technologies (i. e. , those which are equally applicable to commercial and defense systems) in information processing, electronics, optics, and materials in support of the Department of Defense's missions. Within the High Performance Computing and Communications Program (HPCC), ARPA develops the underlying computing and communications technologies needed by the program. The focus is on creating a scalable technology base of interoperating workstations, networks, parallel computing systems, massive storage, systems software, and development tools and environments. This identical technology base, enriched by an intelligent, adaptive layer of application support services and with support for a more diverse collection of bitway technologies, will underlie the development of the National Information Infrastructure.

The Department of Defense has unique mission responsibilities that influence its interests in the National Challenge applications. These are reflected in ARPA's emphasis on the development of NII capabilities in support of Crisis Management, Training, Health Care, and Manufacturing. The applications are described in more detail below.

*Crisis Management:* This is the use of command, control, communications, and intelligence information systems to support decision makers in anticipating threats, formulating plans, and

executing such plans through a coordinated response. Operation Desert Storm is one form of a large-scale crisis; the recent floods along the Mississippi, the Los Angeles riots, Hurricane Andrew, and the Loma Prieta earthquake also could benefit from a comprehensive information-centered approach to command and control. The underlying systems are of direct interest to the National Guard as well as law enforcement and public safety agencies.

*Education and Training:* The major thrust of ARPA's activities is in the area of simulation-based training, in particular, distributed simulations made possible by network technology. In addition, Defense has considerable experience in constructing and using (hardware) simulators of complex systems. Nevertheless, software-based simulators and "virtual realities" will dominate in future training systems.

*Health Care:* ARPA's interests are particularly focused on exploiting information technology to improve the survival of trauma victims. The key to survival is rapid response. This means determining the most capable care provider who can get to the scene most rapidly with the right equipment and instruments. In addition, ARPA is interested in developing technologies that make available remote medical expertise where it is needed and where it can be used in the critical first few minutes of injury. Thus, there is considerable interest in telemedicine.

*Design and Manufacturing:* Defense depends on an advanced design and manufacturing base from which its specialized systems can be implemented and constructed. Given the new demands for affordable military systems, information technology must be exploited to make possible low volume manufacturing that is also low(er) cost. To this end, ARPA has been focusing on the underlying infrastructure elements to make communications and coordination along the producer-supplier food chain more efficient.

## A. 2. Department of Energy

The Department of Energy (DOE) has unique and extensive capabilities that could provide major contributions to the establishment of a National Information Infrastructure. Some of the Nation's premier high performance computing systems, applications, and networking capabilities have been developed, perfected, and routinely utilized by DOE scientists and engineers in pursuit of the department's missions. Many departmental facilities are routinely accessed and used for gathering information on energy production and use, for performing research and education programs and projects, and for collaborating with industry and academia in manufacturing and other technologies.

The proposed DOE IITA program combines research and development of information infrastructure technologies with mission related applications testbeds, which will be scalable nationally, innovative, information driven, relevant to industrial competitiveness, and performed in conjunction with U.S. industry. This program has three objectives: (1) to improve industry access to the DOE scientists and engineers to improve the flow of technology from the department's national laboratories and production facilities into the private sector; (2) to assist

the department in maintaining its core competencies, especially its traditionally strong information technology expertise, in the DOE national laboratories; and 3) to increase the internal effectiveness of departmental activities.

DOE is in the forefront in the development and use of information technologies, such as high performance computing, high speed networking, data storage and databases, and other information services and system integration technology. It has pioneered the use of computing and networking tools to enable distributed and/or remote collaboration. Its diverse missions to conduct forefront research and development in many energy related fields and to design and manufacture nuclear weapons have mandated that the department stay at the leading edge in these information infrastructure technologies. This strength makes the department's information activities relevant to that of many U.S. industrial sectors.

The benefits to DOE from this program are threefold. First, it will allow a more cost- effective approach to carrying out the missions of the department during a period of constrained budgets. Second, by providing industry an open connection to this infrastructure, it will facilitate collaborations between industrial scientists and engineers and their counterparts in DOE laboratories and production facilities. Third, it will provide testbeds and demonstrations of information infrastructure technology and applications that are relevant to the National Information Infrastructure.

In summary, the DOE's weapons mission and its energy research mission have required that it develop enormous strengths in these technological areas over the last four decades. To accommodate future environmental concerns, these technologies become even more important for future success. It will now be the department's challenge to apply these strengths to the new national priorities.

### A. 3. Environmental Protection Agency

The Environmental Protection Agency must involve a broad range of interests and technological levels in its mission to improve the condition of the Nation's environment. Access to environmental information, collaboration among widely- distributed environmental scientists and decision makers within the Federal Government and States, and an informed public are essential elements in the combination of regulatory approaches and public action that are necessary to address the environmental problems of the twenty-first century. EPA requires a broad range of technologies and applications in order to synergize the diverse disciplines necessary to study the environment and to produce effective policies.

EPA envisions collaboration with other Federal Agencies such as NASA, NOAA, and DOE to create digital libraries of environmental monitoring data for the purpose of determining the health and potential for change in the Earth's environmental systems. Likewise, EPA envisions access to computational models of environmental systems by partners in the states to collaborate in predicting the effects of potential regulations or existing environmental conditions.

A. 4. National Aeronautics and Space Administration

NASA's five Strategic Enterprises (Mission to Planet Earth, Scientific Research, Aeronautics, Human Exploration, and Development of Space and Space Technology) will all contribute to the establishment of a U.S. National Information Infrastructure, but the first three are especially related to developing an NII and thus become the drivers for NASA's IITA program. Under the Mission to Planet Earth enterprise, NASA has the responsibility for developing and operating the systems for the Earth Observing System and its Data and Information System (EOSDIS). A primary goal of the NASA IITA program will be to extend the outreach of the NASA EOSDIS program to the underserved communities in the U.S. commercial, educational, and public interest sectors.

NASA's Scientific Research enterprise includes missions to understand the genesis and evolution of the solar system, the galaxy, and the universe, and to conduct unmanned exploration of planetary bodies. This enterprise is also generating tremendous amounts of data and information that NASA's IITA program will strive to deliver to interested parties such as the US educational community that NASA currently cannot well serve.

NASA's role as the lead agency for the development and demonstration of fundamental aeronautical technologies for both the commercial and defense U.S. aerospace industry naturally positions NASA to work in partnership with U.S. aerospace companies in the development of an NII to serve their needs. NASA's IITA program will demonstrate the potential for government-industry partnerships in cooperative development efforts within a National Information Infrastructure.

Finally, NASA has substantial programs in educational outreach that will accelerate the utilization of an NII by the education community. A prime example of this is the Spacelink repository on the Internet for teachers and students that currently has almost 30,000 active user accounts. The impact and outreach of resources such as Spacelink will be enhanced under NASA's IITA program.

Major NASA activities under the HPCC IITA component are listed below:

*Digital Libraries:* Jointly support basic research in Digital Libraries with NSF and ARPA through an NSF lead research solicitation. Establish cooperative agreements with partners in industry, academia, and the non-profit sectors for the development and demonstration of Digital Library Technology for providing public access to databases of remote sensing data and images.

*Environmental Monitoring:* Establish cooperative agreements through a competitive solicitation with developers and underserved users for new and innovative applications of remote sensing data that will lead to new commercial and public services founded on remote sensing data within the U.S. NII.

*Design and Manufacturing:* Establish cooperative agreements with the U.S. aerospace industry to jointly develop the NII design and manufacturing services of critical interest this industry.

*Education, Training, and Lifelong Learning:* Expand the NASA center-based programs for outreach in computational science and computer networking for K-12 education. Initiate a NASA Research and Cooperative Agreement opportunity for utilizing NASA assets and NII services and technologies for the betterment of K-12 math, science, and engineering education.

A. 5. National Institutes of Health

The National Institutes of Health (NIH) is the steward of biomedical research for the United States, with the mission of pursuing knowledge in biomedical science and applying that scientifc knowledge to improve human health. The effectiveness of the NIH depends on its ability to foster fundamental, innovative and valuable research and to communicate and implement new knowledge expeditiously.

To this end, the NIH's interests and activities in the IITA component of the High Performance Computing and Communications Program (HPCC) are focused on exploiting developments in information technology as well as the emerging NII to facilitate: improvements in medical technology; group cooperative work environments through software, interfaces and networking; high performance, high speed acquisition, storage and dissemination of health-related patient and research data; better coordination of large scale complex research programs; and high performance communications to accesss high performance computing.

The National Library of Medicine (NLM) IITA program supports research and development to create testbed networks for linking health care facilities, systems for visualizing human anatomy, medical applications of virtual reality technology, telemedicine methods and systems, and database technology for storing, accessing, and transmitting patients' medical records while protecting the accuracy and privacy of those records. This research and development effort is supported via the mechanism of a Broad Agency Announcement.

IITA initiatives in the National center for Research Resources emphasize simulation, image analysis, virtual reality, group software development and science education. In the area of simulation, major goals are to enhance computational capabilities and project the course of epidemics such as AIDS. Image analysis goals are in real-time medical imaging. Virtual reality includes development of tools to expand basic and clinical application of technological methods and instrumentation. Group software development goals are to support group cooperative work environments and use of networks. Science education will emphasize the development of interactive learning tools to increase public and student understanding of science. NCRR programs support these initiatives through both investigator initiated resource centers and research projects.

The Division of Computer Research and Technology (DCRT) is developing new methods of medical imaging for basic research, clinical research, and health care delivery in collaboration with the Radiology and Nuclear Medicine Departments of the NIH Clinical center and the Radiation Oncology Branch of the National Cancer Institute. This work includes the development of scalable parallel algorithms for medical image processing and radiation treatment planning, the deployment of powerful workstations for medical image visualization, and the deployment of high speed networks for the transfer of images between these visualization workstations and high performance parallel computer systems.

The National Cancer Institute (NCI) activities in IITA include an intramural project to develop network based multimedia conferencing in which all users participate in a shared session with the goal of facilitating cooperative research, as well as several projects aimed at increasing the availability of PDQ (Physician Data Query) and other cancer related information over multiple dissemination platforms for the medical and cancer patient communities. These projects involve the development of new interfaces to PDQ through SBIR contracts, use of Internet and Gopher to facilitate access, Spanish language translations through Internet and development of a multimedia CD-I platform for disseminating health information. NCI cooperates with NLM and DCRT in telemedicine and several other IITA areas.

A. 6. National Institute of Standards and Technology

The National Institute of Standards and Technology (NIST) has a program in high performance computing and communications with the following objectives: (1) accelerate the development and deployment of high-performance computing and networking technologies required for the National Information Infrastructure and (2) apply and test these technologies in a manufacturing environment. Building on its leadership in research, standards development, measurement technology and technology transfer in both information technology and manufacturing, NIST will collaborate with industry and other agencies to deploy and utilize a test facility to apply information technologies more quickly to manufacturing.

Program elements include research, development and standards activities related to Electronic Commerce, Distributed Multimedia Environments, Human-machine Interface, and systems integration for advanced manufacturing applications. Electronic Commerce enables the electronic exchange of business transactions and will greatly improve the ability to quickly and effectively develop new intellectual property and to bring new ideas and technologies to the worldwide marketplace through improved manufacturing techniques. Electronic data interchange will be extended to deal with new types of data, and interfaces will be developed with other standards to describe, access, and update data. The Systems Integration for Manufacturing Applications program elements will develop, refine, test, and transfer advanced, computer-integrated, electronically-networked manufacturing technologies and associated applications using HPCC technologies, with emphasis on facilitating the use of standards in the areas of flexible computer integrated manufacturing and product data exchange. The program will

support industrial environments including mechanical part production, apparel, electronics, construction, and chemical processing.

An Advanced Manufacturing Systems and Networking Testbed (AMSANT) will be established at NIST to enable research and development into advanced manufacturing computer systems and networking. In addition the AMSANT will (1) test high performance computer and networking hardware and software to determine their suitability for use within the U.S. manufacturing community; (2) assist industry in the development and implementation of voluntary consensus standards; and (3) serve as a demonstration site that industrial technology suppliers and users can use to identify and overcome technical and other barriers leading to the successful and cost-effective implementation of these systems. NIST will support research and development of advanced manufacturing systems and networking, and will make results of the program available to American industry through workshops, training materials, accessible electronic data repositories and pre-commercial prototype systems that can be installed by potential vendors in user facilities for test and evaluation. A Standards Development Environment will support the development of standards that must underlie the AMSANT facility. A major standards-based data exchange effort will be aimed at flexible, computer- integrated manufacturing to alleviate the lack of data exchange capability among computer aided design, computer-aided process, and computer-aided manufacturing activities.

A. 7. National Oceanic and Atmospheric Administration

The National Oceanic and Atmospheric Administration's (NOAA) mission is to observe, understand, and predict changes in the environment. The information products that result from carrying out this mission are distributed locally, nationally, and internationally and impact nearly every sector of society. These include warnings of severe weather, navigation charts, regional long-range weather outlooks, fishery stock assessments, flood forecasts, stratospheric ozone depletion analyses, the national geodetic reference grid, coastal tidal predictions, and assessments of greenhouse gases effects on global climate. NOAA information products find critical use in many aspects of commerce, science, agriculture, law, education, and everyday life. No building is constructed, no ship sails, no crop is planted, no roadbed is laid, and no power plant is planned without referencing NOAA environmental observations and information products.

Historically NOAA has made large leaps forward in fulfilling its mission through the application of technological advances to providing more accurate and complete environmental assessments and more timely distribution of information products to the Nation. This includes early application of supercomputer technology to operational weather prediction, using Earth-observing satellites to monitor the oceans and atmosphere, and the application of telecommunication technologies to global environmental data acquisiton and product distribution. NOAA finds itself once again at the threshold of such an opportunity to make a quantum leap forward in the quality and timeliness of its environmental information delivery to the Nation. The IITA component of the HPCC Program represents a means by which NOAA, along with the other participating agencies, will parlay the national investment in computing and

communications technology development into significant and measurable benefit to the national well-being.

With IITA enabling technologies, NOAA can make literally thousands of environmental observations from across the Nation and the globe available instantly to the analyst or forecaster; the agency's environmental observing systems, whether land-based, ocean-based, or satellite-borne, can be integrated with observing systems of other agencies or other nations into a virtual national or global environmental monitoring system; its historical databases of environmental observations and information distributed across the Nation at scores of sites can be seamlessly integrated along with other-agency environmental databases into a virtual national database and placed at the fingertips of the researcher, policy maker, lawyer, businessman, educator, or any other user; national policy with regard to sustainable development can be made with the most comprehensive, up-to-date, and best environmental information available, integrated with economic and human-dimension information; the geometrically growing volume of environmental data and information can be effectively and automatically quality- controlled, integrated, synthesized, archived, and retrieved.

A. 8. National Science Foundation

## The Changing Context of NII

The proposed National Information Infrastructure promises a radical transfor- mation of the way Americans at all levels generate, transmit, access, and use information. Whether it is a schoolchild writing a report on black holes or loss of biodiversity, the librarian who assists the child in assembling the information from a digital library, the astrophysicist or ecologist who generates the primary data or the child's parents and teachers communicating about the homework assignment, all Americans may use NII in their daily lives and will be enriched by that use.

## NSF'S Unique Strengths and Assets

Achieving the goals of the NII requires not only new technologies but also life-long training in and exposure to the methods for taking full advantage of NII. NSF is uniquely qualified to lead the research efforts in creating the technologies and to provide the educational foundations on which NII will be built.

NSF's unique qualifications derive from several different factors, all of which must be integrated in order to achieve the NII vision. First, NSF supports scientific and engineering research across the spectrum of disciplines. Second, NSF has a record of actually building the infrastructure for bringing people with a wide spectrum of interests and expertise together through networking and other facilities. Third, NSF is involved with all aspects of the educational enterprise, ranging from graduate education to K-12, and ranging from teacher training to systemic change. NSF's support of diverse communities enabled it to build an effective communication link, NSFNET, between the research and education community at large and the foci of research enterprises in the various NSF centers.

## Ingredients of NII Implementation

The development and implementation of an NII is not a matter of technology alone. It involves: technology based on knowledge creation, development of expertise of how to use new technology in a meaningful and efficient manner, what tools and interfaces are needed for specific application domains, policies for public use of provided services, general socio-economic aspects of introduction and application of infrastructure services and, last but not least, training of the potential users for whom these services are beneficial.

Using the NII will not be similar to using a telephone system that can be used without much training. It requires sharp minds to develop practical access to information, but also helping people how to utilize it creatively. It is more in the style of setting up an aquarium or learning to play an instrument. It requires familiarity with the technology and understanding of its potential application which may vary by individual and purpose. Hence training comes first and then a host of applications emerge for individuals, from the very young to senior citizens, to acquire information, learn from it and to connect with others. Information plays a vital role in every aspect of life and work. Those who are information "literate" will not just survive in this new information environment but will prosper while participating in advancing the frontiers of knowledge.

## NSF'S Contribution

NSF's support for and implementation of the NII is largely coordinated through the HPCC Program, including the new IITA component. NSF has unique experience in building partnerships between the inventors, producers, educators, and users of NII technologies. The variety of users NSF can bring to participate in prototype development and use is without equal. Research conduct is changing radically. Researchers and educators will interact with information in realtime and from remote locations. What can be done today on site can be done remotely over the whole world tomorrow. NSF's leadership role in applying advanced computing and networking technologies in education and training for all citizens will enhance the deployment of the NII and new ways of conducting research and education.

A.9. National Security Agency

Information security is a primary mission of the National Security Agency. In pursuing this mission NSA develops security technology and security products. Network security applications present significant challenges for the National Information Infrastructure. NSA will initiate programs that address authentication, wireless interoperability and real-time operating aspects of network security. NSA will develop authentication techniques (e.g., biometric) for network access control and will create a mechanism for establishing standards of security and interoperability in emerging wireless PCs and cellular networks. NSA and NIST will integrate Smartcard and biometric technology to permit authentication of a person to a network. In collaboration with NIST and hardware and software manufacturers, NSA will develop

technology to integrate real time and data communications in a secure multimedia multi-terminal network environment.

Over the past few years there has been a significant research program funded by ARPA, NSA and other government agencies to develop and test automatic machine translation algorithms. While this research program has been constrained to a limited source of documents and a limited set of languages, results so far have been very promising. However a follow-on program is needed to transfer the results of this research into operational use. NSA will sponsor work to extend the applicability of the best language translation algorithms to more languages and more general domains; to improve the computational efficiency of those algorithms; to port those algorithms to networked workstations; and to develop good human-machine interfaces to allow easy control and operation of the system.

For textual information, there are ongoing research programs for document retrieval by topic, for data extraction and for machine translation. For several years, ARPA, NSA and other agencies have conducted and sponsored research programs to develop algorithms for large vocabulary, continuous speech recognition. A follow-on to this research program is needed to further improve the recognition algorithms and to build a prototype speech recognition system and a system capable of processing continuous speech dictation of arbitrary text. NSA will sponsor work to extend the applicability of the best large vocabulary continuous speech recognition systems to vocabularies with sizes up to 50,000 words and to languages other than English; to improve the computational efficiency of those algorithms; to port those algorithms to networked workstations; and to develop effective human- machine interfaces to allow easy training, testing and general use of the system. The goal of the program is to deliver a usable prototype system for taking dictation on arbitrary topics using continuous speech input.

A major effort needs to be initiated for development of efficient and reliable text summarization technology. Text summarization will combine existing text generation systems with a new understanding of how to identify key points of information in a text to reduce the volume of text an analyst needs to review. Prototype development for text summarization and relevance feedback from users is a near-term goal of the program.

The seamless and transparent access to multiple heterogeneous databases with intelligent, integrated run-time reasoning is the major component of a large knowledge base effort. NSA will develop a prototype environment of the future where the end-user of the data, the application builder, and the data administrator all see, not a collection of relatively unintelligible, difficult-to-access databases, but an integrated information space in terms directly meaningful and accessible to them.

In support of digital library applications, NSA will develop improved interfaces to information resources, including enhancements to WAIS.