REPORT TO THE PRESIDENT AND CONGRESS

DESIGNING A DIGITAL FUTURE: FEDERALLY FUNDED RESEARCH AND DEVELOPMENT IN NETWORKING AND INFORMATION TECHNOLOGY

Executive Office of the President

President's Council of Advisors on Science and Technology

DECEMBER 2010

REPORT TO THE PRESIDENT AND CONGRESS

DESIGNING A DIGITAL FUTURE: FEDERALLY FUNDED RESEARCH AND DEVELOPMENT IN NETWORKING AND INFORMATION TECHNOLOGY

Executive Office of the President

President's Council of Advisors on Science and Technology

DECEMBER 2010

About the President's Council of Advisors on Science and Technology

The President's Council of Advisors on Science and Technology (PCAST) is an advisory group of the nation's leading scientists and engineers, appointed by the President to augment the science and technology advice available to him from inside the White House and from cabinet departments and other Federal agencies. PCAST is consulted about and provides analyses and recommendations concerning a wide range of issues where understandings from the domains of science, technology, and innovation may bear on the policy choices before the President. PCAST is administered by the White House Office of Science and Technology Policy (OSTP).

For more information about PCAST, see http://www.whitehouse.gov/ostp/pcast

The President's Council of Advisors on Science and Technology

Co-Chairs

John P. Holdren
Assistant to the President forScience and Technology
Director, Office of Science and Technology Policy

Eric Lander
President, Broad Institute of Harvard and MIT

Harold Varmus*
President, Memorial Sloan-Kettering Cancer Center

Members

Rosina Bierbaum
Dean, School of Natural Resources and Environment
University of Michigan

Christine Cassel
President and CEO, American Board of Internal Medicine

Christopher Chyba
Professor, Astrophysical Sciences and International Affairs
Director, Program on Science and Global Security
Princeton University

S. James Gates, Jr.
John S. Toll Professor of Physics
Director, Center for String and Particle Theory
University of Maryland

Shirley Ann Jackson
President, Rensselaer Polytechnic Institute

Richard C. Levin
President
Yale University

Chad Mirkin
Rathmann Professor, Chemistry, Materials Science and Engineering, Chemical and Biological Engineering and Medicine
Director, International Institute of Nanotechnology
Northwestern University

Mario Molina
Professor, Chemistry and Biochemistry
University of California, San Diego
Professor, Center for Atmospheric Sciences
Scripps Institution of Oceanography
Director, Mario Molina Center for Energy and Environment, Mexico City

Ernest J. Moniz
Cecil and Ida Green Professor of Physics and Engineering Systems
Director, MIT's Energy Initiative
Massachusetts Institute of Technology

Craig Mundie
Chief Research and Strategy Officer
Microsoft Corporation

** Dr. Varmus resigned from PCAST on July 9, 2010 and subsequently became Director of the National Cancer Institute (NCI).*

Ed Penhoet
Director, Alta Partners
Professor Emeritus of Biochemistry and of Public Health
University of California, Berkeley

William Press
Raymer Professor in Computer Science and Integrative Biology
University of Texas at Austin

Maxine Savitz
Vice President
National Academy of Engineering

Barbara Schaal
Chilton Professor of Biology
Washington University
Vice President, National Academy of Sciences

Eric Schmidt
Chairman and CEO
Google, Inc.

Daniel Schrag
Sturgis Hooper Professor of Geology
Professor, Environmental Science and Engineering
Director, Harvard University-wide Center for Environment
Harvard University

David E. Shaw
Chief Scientist, D.E. Shaw Research
Senior Research Fellow, Center for Computational Biology and Bioinformatics
Columbia University

Ahmed Zewail
Linus Pauling Professor of Chemistry and Physics
Director, Physical Biology Center
California Institute of Technology

Staff

Deborah Stine
Executive Director

Mary Maxon
Deputy Executive Director

Gera Jochum
Policy Analyst

EXECUTIVE OFFICE OF THE PRESIDENT

PRESIDENT'S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY

WASHINGTON, D.C. 20502

President Barack Obama
The White House
Washington, DC 20502

Dear Mr. President,

We are pleased to send you this report, *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology*, prepared by your President's Council of Advisors on Science and Technology (PCAST) acting in its role as the President's Innovation and Technology Advisory Council (PITAC). This report fulfills PCAST's responsibilities under Executive Order 13539 and the High-Performance Computing Act of 1991 (Public Law 102-194) as amended by the Next Generation Internet Research Act of 1998 (Public Law 105-305) and by the America COMPETES Act of 2007 (Public Law 110-69).

The Networking and Information Technology Research and Development (NITRD) Program is the primary mechanism by which the Federal government coordinates its unclassified networking and information technology (NIT) research and development (R&D) investments. Fourteen Federal agencies, including all of the large science and technology agencies, are formal members of the NITRD Program, with many other Federal entities participating in NITRD activities. The program helps ensure that the Nation effectively leverages its strengths, avoids duplication, and increases interoperability in such critical areas as supercomputing, high-speed networking, cybersecurity, software engineering, and information management.

To provide a solid scientific basis for its assessment of NITRD, PCAST appointed an expert 14-member Working Group, which consulted with more than 50 individuals, including government officials, industry representatives, and experts from academia.

PCAST finds that NITRD is well coordinated and that the U.S. computing research community, coupled with a vibrant NIT industry, has made seminal discoveries and advanced new technologies that are helping to meet many societal challenges. Importantly, however, PCAST also finds that a substantial fraction of the NITRD multi-agency spending summary represents spending that supports R&D in other fields, rather than spending on R&D in the field of NIT itself. As a result, the Nation is actually investing far less in NIT R&D than the $4 billion-plus indicated in the Federal budget. To achieve America's priorities and advance key research frontiers to support economic competitiveness in NIT, this report calls for a more accurate accounting of this national investment and recommends additional investments in NIT R&D, including research in networking and information technology for health, energy and transportation, and cyber-infrastructure, among others.

NIT has yielded enormous benefits for the Nation's economic competitiveness, national security, and quality of life. To maintain America's leadership in NIT in an ever more competitive global environment, the Federal Government must be bold in its investments, including funding of high risk/high reward research with the potential to move this essential field in unanticipated directions. PCAST believes that execution of the recommendations in this report will enable us to address critical priorities and challenges in the years ahead.

Sincerely,

John P. Holdren	**Eric Lander**	**Shirley Ann Jackson**	**Eric Schmidt**
PCAST Co-chair	PCAST Co-chair	PITAC Co-chair	PITAC Co-chair

Executive Report

From smartphones to eBook readers to game consoles to personal computers; from corporate datacenters to cloud services to scientific supercomputers; from digital photography and photo editing, to MP3 music players, to streaming media, to GPS navigation; from robot vacuum cleaners in the home, to adaptive cruise control in cars and the real-time control systems in hybrid vehicles, to robot vehicles on and above the battlefield; from the Internet and the World Wide Web to email, search engines, eCommerce, and social networks; from medical imaging, to computer-assisted surgery, to the large-scale data analysis that is enabling evidence-based healthcare and the new biology; from spreadsheets and word processing to revolutions in inventory control, supply chain, and logistics; from the automatic bar-coding of hand-addressed first class mail, to remarkably effective natural language translation, to rapidly improving speech recognition – our world today relies to an astonishing degree on systems, tools, and services that belong to a vast and still growing domain known as Networking and Information Technology (NIT). NIT underpins our national prosperity, health, and security. In recent decades, NIT has boosted U.S. labor productivity more than any other set of forces.

The United States has a proud history of achievement and leadership in NIT. The Federal Government has played an essential role in fostering the advances in NIT that have transformed our world. Steady Federal investment in NIT research over the past 60 years has led to many of the breakthroughs noted above, often a decade or more after the research took place. The Federal investment in NIT research and development is without question one of the best investments our Nation has ever made[1,2,3].

In order to sustain and improve our quality of life, it is crucial that the United States continue to innovate more rapidly and more creatively than other countries in important areas of NIT. Only by continuing to invest in core NIT science and technology will we continue to reap such enormous societal benefits in the decades to come.

Recent technological and societal trends place the further advancement and application of NIT squarely at the center of our Nation's ability to achieve essentially all of our priorities and to address essentially all of our challenges:

- **Advances in NIT are a key driver of economic competitiveness.** They create new markets and increase productivity. For example, an investment in the National Science Foundation's Digital Library Initiative in the 1990's led to Google, a company with a market capitalization of nearly $200 billion[4] that has transformed how we access information.
- **Advances in NIT are crucial to achieving our major national and global priorities in energy and transportation, education and life-long learning, healthcare, and national and homeland security.** NIT will be an indispensable element in buildings that manage their

1. National Academies Press. (1995). *Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure*.
2. National Academies Press. (2003). *Innovation in Information Technology*.
3. President's Information Technology Advisory Committee Report to the President. (1999). *Information Technology Research: Investing in Our Future*.
4. See the Section 12 sidebar "Why We're Able to Google" (page 107).

own energy usage; attention-gripping, personalized methods that reinforce classroom lessons; continuous unobtrusive assistance for people with physical and mental disabilities; and strong resilience to cyber warfare.

- **Advances in NIT accelerate the pace of discovery in nearly all other fields.** The latest NIT tools are helping scientists and engineers to illuminate the progression of Alzheimer's disease, elucidate the nature of combustion, and predict the size of the ozone hole, to cite just a few examples.
- **Advances in NIT are essential to achieving the goals of open government.** Those advances will allow better access to government records, better and more accessible government services, and the ability both to learn from and communicate with the American public more effectively.

Both the science and the practice of NIT have seen dramatic changes during the sixty-year history of the field. The ability of the computing research community, coupled with a vibrant NIT industry, to deliver those changes – to discover and advance new areas of NIT research and development (R&D) that stimulate technological progress and meet societal challenges – has been essential to the Nation's success. There are enormous opportunities for future transformations. To meet the challenge of change, America must continue to make R&D investments in new areas of NIT.

Of course, the Government is not alone in investing in NIT R&D. Industry has made, and continues to make, major contributions. It is important, however, not to equate the very large industry R&D investment in NIT with fundamental research of the kind that is carried out in universities and a small number of industrial research labs. The vast majority of industry R&D in NIT is focused on development – on the engineering of future products and product versions. Few major NIT companies have formal research organizations, and even those that do invest relatively little in research compared to their investment in development activities. Fundamental research with the potential for future transformational application represents a small fraction of overall industry R&D in NIT – a situation that is both appropriate and unlikely to change[5]. For that reason, among others, Federal investment in NIT R&D is and will remain essential.

As a field of inquiry, NIT has a rich intellectual agenda – as rich as that of any other field of science or engineering. In addition, NIT is arguably unique among all fields of science and engineering in the breadth of its impact. Computer science research, carried out to a great extent in America's research universities with funding from Federal agencies such as the National Science Foundation (NSF) and the Defense Advanced Research Projects Agency (DARPA), lies at the heart of our Nation's leadership. It is this research – which ranges from the design of computers and networks to robotics, software, and algorithms – that has repeatedly led to the introduction of entirely new product categories that became multi-billion-dollar industry sectors. The "extraordinarily productive interplay of federally funded university research, federally and privately funded industrial research, and entrepreneurial companies founded and staffed by people who moved back and forth between universities and industry"[6] has been well documented.

5. See Section 12.

6. National Research Council. (1999). *Funding a Revolution: Government Support for Computing Research*. Washington, DC: National Academies Press.

Essentially all unclassified federally funded R&D activities in NIT and related fields fall within the scope of the Networking and Information Technology Research and Development (NITRD) Program. The term "NITRD Program" refers both to the mechanism by which the Federal Government coordinates its unclassified R&D investments in NIT, and to the unclassified Federal NIT R&D portfolio itself. The NITRD member agencies report aggregate NIT R&D investments in excess of $4 billion annually. The largest investments are reported by the National Institutes of Health (NIH) and NSF (roughly $1 billion each), followed by the Office of the Secretary of Defense and the Department of Defense Service research organizations (OSD/DoD), the Department of Energy (DoE), and DARPA (roughly $500 million each)[7]. However, analysis indicates that a substantial fraction of the NITRD crosscut budget (the multi-agency spending summary) represents spending on NIT that supports R&D in other fields, rather than spending on R&D in the field of NIT itself. For example, an expert review of the top 100 awards (by award size) in NIH's NITRD portfolio – totaling nearly $600 million, roughly half of NIH's NITRD crosscut total – concluded that only between 2% and 11% (by dollar value) should be considered NIT R&D[8]. The remainder is spent on various forms of NIT infrastructure that provide essential support for biomedical research, but not on NIT R&D. We have used NIH as an example only because the laudable transparency of its records and reporting allowed such an analysis to be performed. Although other agencies do not report NIT R&D spending in sufficient detail to make the same analysis possible, it seems likely that in many cases a similar confusion in classification of NITRD investment occurs. An important finding of this report is that the Nation is actually investing far less in NIT R&D than is shown in the Federal budget.

In summary, the transformative NIT research that fuels innovation and achievement and strengthens our Nation needs to come from Government investment, yet it is currently difficult to ascertain the magnitude of that investment. Furthermore, going forward, the participating agencies in the NITRD Program must more aggressively embrace the expanding role that advances in NIT play in America's future. A broad spectrum of Federal agencies – those currently participating in NITRD and some which are not yet doing so – must recognize that their abilities to accomplish their missions are inextricably linked to advances in NIT, and must invest in NIT R&D to catalyze the advances that are critical to their missions. Strategic leadership must come from the top – from those within the Federal Government with the authority to implement new strategies.

The PCAST NITRD Program Review Working Group was asked to assess not only the coordination function of NITRD but also the investment portfolio itself. In the remainder of this Executive Report, and in greater detail and breadth within the body of the report, we describe some of the compelling and important scientific and technical problems that must be addressed in order to maintain and strengthen the transformative effect of NIT on the Nation and the world, and we describe some of the essential research that will be needed to solve those problems. A bottom-up analysis of some of the key initiatives that we recommend in this report suggests that an investment of at least $1 billion annually will be needed for new, potentially transformative NIT research. Uncertainty regarding the precise nature of current expenditures makes it difficult to determine how much of this investment can be obtained

7. *Networking and Information Technology Research and Development Supplement to the President's FY 2011 Budget,* (February 2010) (page 21).

8. Analysis conducted for this report by the Science and Technology Policy Institute of the Institute for Defense Analysis. See sidebar, "The NITRD Crosscut Budget Significantly Overstates the Federal Investment in NIT R&D," in Section 10. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stpi-nitrd-9-15-2010.pdf

through repurposing and reprioritization and how much will require new funding. We believe, however, that a lower level of investment in this critically important area could seriously jeopardize America's national security and economic competitiveness.

Recommended Initiatives and Investments in NIT R&D to Achieve America's Priorities and Advance Key NIT Research Frontiers

The Federal Government's investment in NIT R&D dates from the birth of the field more than sixty years ago. NITRD as a coordination effort, though, had its genesis in the High-Performance Computing Act of 1991 – "An Act to provide for a coordinated Federal program to ensure continued United States leadership in high-performance computing." Its scope was broadened by the Next Generation Internet Research Act of 1998, and again by the America COMPETES Act of 2007.

In its early years, NITRD's role was seen as coordinating research in the fundamentals of computing, while the use and advancement of the resulting technology to address our national priorities was left to individual agencies. In recent years, the value and importance of multi-agency coordination in the development and application of NIT to achieve the Nation's priorities has become apparent, and has led to the creation of NITRD Senior Steering Groups in the vital areas of Cyber Security and Information Assurance, and Health IT. NITRD is well-positioned to facilitate similar coordination in NIT for other important national priorities, among them energy and transportation, and education and life-long learning.

The role of NIT in addressing our national priorities, and the NIT research frontiers that contribute to making progress in strengthening our NIT capabilities, raise many important research questions that must be tackled. It is essential that short term needs not crowd out the longer term research that anticipates future needs. It is also essential that some NIT research explore bold, unconventional ideas that would have enormous impact if they could be realized. A recent report from the American Academy of Arts & Sciences[9] describes both the benefits of such transformative research and the mechanisms that can be used to foster it.

The Federal Government must invest in new multi-agency NIT R&D initiatives in areas of particular importance to our national priorities. Such investments should include funding for high risk/high reward research with the potential to move these areas in unanticipated directions. Some of this research will require large project teams and sufficiently long time horizons to allow ambitious goals to be achieved. We see three areas in which such initiatives are particularly timely and important.

9. American Academy of Arts & Sciences. (2008). *ARISE: Advancing Research in Science and Engineering – Investing in Early-Career Scientists and High-Risk High-Reward Research.*

Recommendation [Section 5]: **The Federal Government, under the leadership of NSF and Health and Human Services (HHS), with participation from the Office of the National Coordinator for Health Information Technology (ONC), the Centers for Medicare and Medicaid Services (CMS), the Agency for Healthcare Research and Quality (AHRQ), the National Institute of Standards and Technology (NIST), the Veterans Health Administration (VHA), DoD, and other interested agencies, should invest in a national, long-term, multi-agency research initiative on NIT for health that goes well beyond the current national program to adopt electronic health records.** The initiative should include sponsorship of multi-disciplinary research on three themes:

- to make possible comprehensive lifelong multi-source health records for individuals;
- to enable both professionals and the public to obtain and act on health knowledge from diverse and varied sources as part of an interoperable health IT ecosystem; and
- to provide appropriate information, tools, and assistive technologies that empower individuals to take charge of their own health and healthcare and to reduce its cost.

This program should build on national activities promoting the adoption and meaningful use of electronic health records that are usable by all appropriate organizations; it should complement the shorter-term ONC programs; and it should augment the research investments that the various agencies are currently able to make. In addition to increased attention on using NIT for wellness and for addressing chronic conditions, the departments and agencies mentioned above should continue to investigate novel uses of NIT, such as NIT-assisted surgery, to deliver care for acute conditions. They should continue to pursue advances in the innovative use of NIT, such as sensing and monitoring, to understand the basic biological and psychological mechanisms that underlie disease. And they should continue to address NIT research opportunities that support current and continuing work by HHS and NSF on transformational innovation in healthcare delivery and basic research in health and wellness.

Recommendation [Section 5]: **The Federal Government should invest in a national, long-term, multi-agency, multi-faceted research initiative on NIT for energy and transportation.** As part of that initiative:

- DoE and NSF should be major sponsors of research for achieving dynamic power management in applications ranging from single devices to buildings to the power grid.
- NIST should organize the multi-stakeholder formulation of interoperable standards for real-time control. Interoperability facilitates repeated cycles of innovation by multiple vendors, promoting the development of versatile and robust NIT.
- DoD should continue to be a major sponsor of research on using NIT to achieve low-power systems and devices.
- The Department of Transportation (DoT) should sponsor ambitious NIT research relevant to surface and air transportation.

Current research in the computer simulation of physical systems should be expanded to include the simulation and modeling of proposed energy-saving technologies, as well as advances in the basic techniques of simulation and modeling.

Recommendation [Sections 5 and 7]: **The Federal Government should invest in a national, long-term, multi-agency research initiative on NIT that assures both the security and the robustness of cyber-infrastructure.** NSF and DoD, in collaboration with the Department of Homeland Security (DHS), should aggressively accelerate funding and coordination of fundamental research

- to discover more effective ways to build trustworthy computing and communications systems,
- to continue to develop new NIT defense mechanisms for today's infrastructure, and most importantly,
- to develop fundamentally new approaches for the design of the underlying architecture of our cyber-infrastructure so that it can be made truly resilient to cyber-attack, natural disaster, and inadvertent failure.

Infrastructure to be protected includes the Internet and the national telecommunication system as well as computing systems controlling such national resources as the electric power grid and the financial system. Where fundamental NIT advances are needed to support these initiatives, mission agencies should invest in fundamental research in NIT, either alone or in collaboration with NSF, and should not limit their programs to application-specific research.

Effective use of NIT in increasing our economic competitiveness and achieving our other national priorities depends not only on incorporating innovative NIT into a wide variety of domains, but also on ensuring that the basic science and engineering of NIT remain vibrant and strong. At the time of the High-Performance Computing Act of 1991, the importance of high performance computing and communication (HPCC) to scientific discovery and national security was a major factor underlying the special attention given by Congress to NIT. Although HPCC continues to contribute in important ways to scientific discovery and national security, many other aspects of NIT have now risen to comparable levels of importance. Among these NIT areas are the interactions of people with computing systems and devices, both individually and collectively; the interactions between NIT and the physical world, such as in sensors, imaging, robotic and vision systems, and wearable and mobile devices; large-scale data capture, management and analysis; systems that protect personal privacy and sensitive confidential information, are robust in the face of malfunction, and stand up to cyber-attack; scalable systems and networking (i.e., systems and networks that can be either increased or decreased in complexity, size, generality, and cost); and software creation and evolution. HPCC is but one of many important areas of NIT, and America's prowess in HPCC is but one of many measures of our international competitiveness in NIT.

To achieve our national priorities, and to stimulate the next generation of transformative advances in NIT, we must ensure that the modern and emerging research frontiers are well supported. Investment in those areas must include funding for high risk/high reward research with the potential to move these areas in unanticipated directions.

Recommendation [Section 7]: **The Federal Government must increase investment in those fundamental NIT research frontiers that will accelerate progress across a broad range of priorities.** Among such investments:

- NSF and DARPA, with the participation of other relevant agencies, should invest in a broad, multi-agency research program on the fundamentals of privacy protection and protected disclosure of confidential data. Privacy and confidentiality concerns arise in virtually all uses of NIT.
- NSF, DARPA, and HHS should create a collaborative research program that augments the study of individual human-computer interaction with a comprehensive investigation to understand and advance human-machine and social collaboration and problem-solving in a networked, on-line environment where large numbers of people participate in common activities. Understanding such collective human-NIT interactions is increasingly important for defense, for health, and for the activities of daily life.
- NSF should expand its support for fundamental research in data collection, storage, management, and automated large-scale analysis based on modeling and machine learning. Our ever-increasing use of computers, sensors, and other digital devices is generating huge amounts of digital data, making it a pervasive NIT-enabled asset. In collaboration with NIT researchers, every agency should support research, to apply the best known methods and to develop new approaches and new techniques, to address data-rich problems that arise in its mission domain. Agencies should ensure access to and retention of critical community research data collections.
- NSF and DARPA, in collaboration with those agencies tackling problems whose solution entails instrumenting the physical world – including the Environmental Protection Agency (EPA), DoE, DoT, parts of DoD other than DARPA, NIH, the Department of Agriculture (USDA), and the National Oceanic and Atmospheric Administration (NOAA) – should increase research in advanced domain-specific sensors, integration of NIT into physical systems, and innovative robotics in order to enhance NIT-enabled interaction with the physical world.

At the same time, new investments must not supplant continued investment in important core areas such as high performance computing, scalable systems and networking, software creation and evolution, and algorithms, in which government-funded research is making important progress. Topics of importance within these more established core areas continue to change in response to advances in technologies and applications. High performance computing (HPC) is a case in point. Although HPC plays a critical role in ensuring our national security, our economic competitiveness, and our scientific and technological leadership, the United States must anticipate and adapt to the broadening of its high-end computational needs and changes in the underlying technologies available to address them. Highly influential comparative rankings of the world's fastest supercomputers are for the most part based on metrics relevant to only some of our national priorities, and must not be regarded as the sole measure of our continued leadership in this essential area. Although it is important that we not fall behind in the development and deployment of HPC systems that address pressing current needs, it is equally important that we not allow either the funding allocated to the procurement of large-scale HPC systems, or undue attention to a simplistic measure of competitiveness, to "crowd out" the fundamental research in computer science and engineering that will be required to develop truly transformational next-generation HPC systems. To lay the groundwork for such systems, we will need to undertake a substantial and sustained program of fundamental research on hardware, architectures, algorithms and software with the potential for enabling game-changing advances in high-performance computing.

The Importance of Government Leadership

Many of our recommendations address multiple agencies – sometimes in collaborative roles, sometimes in coordinated roles, and sometimes in addressing different issues within an overall area of need. A successful coordinated attack on the Nation's most challenging and important problems requires focused attention on multi-disciplinary, problem-driven research in NIT. That focus must come from Federal leadership. NITRD is chartered and staffed to *coordinate* multi-agency programs, not to create them. Strategic leadership, when necessary, must come from those with the authority to implement new strategies, namely the Office of Science and Technology Policy (OSTP) and the National Science and Technology Council (NSTC), to which NITRD reports. That leadership must have continuity, breadth and depth, and a focus on NIT.

Both the need for leadership and the need for broad multi-disciplinary research require action on the part of the Federal Government.

Recommendation [Section 11]: **The Federal Government must lead in ensuring that strong multi-agency R&D investments are made in NIT to address important national priorities**:

- OSTP should establish a broad, high-level standing committee of academic scientists, engineers, and industry leaders dedicated to providing sustained strategic advice in NIT.
- The NSTC should lead in defining and promoting the major NIT research initiatives that are required to achieve the most important existing and emerging national priorities.

In addition to ensuring that NIT research in support of the Nation's priorities is conducted and that the results are translated into practice, it is essential that appropriately motivated and educated individuals are available as both researchers and practitioners. All indicators – all historical data and all projections – argue that NIT is the dominant factor in America's science and technology employment, and that the gap between the demand for NIT talent and the supply of that talent is and will remain large. Increasing the number of graduates in NIT fields at all degree levels must be a national priority. Fundamental changes in K-12 education are needed to address this shortage. Here too the Federal Government must take the lead.

Recommendation [Section 9]: **The NSTC's Committee on STEM Education proposed in a recent PCAST report[10] must exercise strong leadership to bring about fundamental changes in K-12 STEM education in the United States, among them the incorporation of computer science as an essential component.**

Improved Effectiveness of NITRD Coordination

Thus far, we have focused primarily on the Federal NIT R&D portfolio and the need for multi-disciplinary collaboration in many areas. We now turn to the government coordination process for those investments.

The NITRD inter-agency coordination mechanism is widely – and we think correctly – viewed as successful and valuable. The collection of NITRD working groups has, over the years, enabled government research managers to become familiar with the activities of their colleagues in other agencies, and to formulate joint programs in areas of mutual interest. Nonetheless, steps can and should be taken to improve the effectiveness of the coordination process.

Recommendation [Section 11]: **The effectiveness of government coordination of NIT R&D should be enhanced**:

- The number of NITRD member agencies should be increased. The duration, management levels, and topic areas of the NITRD coordinating groups should be flexible. Budget reporting categories should be decoupled from the coordinating structure.
- The National Coordination Office (NCO) for NITRD should create a publicly available database of government-funded NIT research, and should provide regular detailed reporting to the Director of OSTP.
- The Office of Management and Budget (OMB) and OSTP should reflect NITRD priorities in their annual Budget Priority Memorandum.

In addition, it is important to recognize the inherent limitations of any such process. In particular, each agency's representatives are charged with advancing that agency's mission, and not with devising a broader national strategy. As recommended previously, the NSTC must provide strategic leadership where necessary.

Continued attention must also be given to stable, evolvable, state-of-the-art shared NIT infrastructure for research, as well as new forms of infrastructure to support new research areas and paradigms. Shared NIT infrastructure – whether computational resources, communication networks, community databases (e.g., PubMed and the Protein Data Bank), or collaboration tools – has become essential to research in virtually all fields. NIT is one such field; NIT infrastructure that supports NIT research is a crucial component of NIT R&D, essential to achieving advancements in networking and information technology, which (among many other benefits) will yield the next generation of NIT infrastructure for all fields.

10. President's Council of Advisors on Science and Technology. (September 2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

The Federal investment currently included in the NITRD crosscut budget includes NIT R&D, NIT infrastructure that supports NIT R&D, and NIT infrastructure that supports R&D in other fields. PubMed and the Protein Data Bank are examples of NIT investments that provide essential shared infrastructure for *biomedical* R&D; they do not represent NIT R&D. Similarly, high-end computing facilities, while essential for many types of research, are for the most part shared NIT infrastructure for physical, biological, and engineering fields other than NIT.

It is appropriate that investments in shared NIT infrastructure for R&D be included within the NITRD Program. However, it is important that investments in NIT that support R&D in other fields be clearly differentiated from investments in NIT R&D. A large portion of the "High End Computing Infrastructure and Applications" budget category, which accounts for roughly $1.5 billion of the $4.3 billion NITRD crosscut total, is attributable to computational infrastructure used to conduct R&D in other fields, and not to NIT R&D or to infrastructure for NIT R&D. In addition, as illustrated earlier by the analysis of the NIH NITRD portfolio, various agencies include in their reports for other NITRD budget categories investments in NIT that support R&D in non-NIT fields. Thus the aggregate NITRD crosscut budget significantly overstates the actual Federal investment in NIT R&D. By leading policymakers to believe that we are spending much more on such activities than is actually the case, this discrepancy contributes to a substantial, systematic underinvestment in an area that is critical to our national and economic security.

Recommendation [Section 11]: **The NCO and OMB should redefine the budget reporting categories to separate NIT infrastructure for R&D in other fields from NIT R&D, and should ensure more accurate reporting of both NIT infrastructure investment and NIT R&D investment.**

In summary: The United States has a proud history of achievement and leadership in NIT that has yielded enormous benefits for our economic competitiveness, our national security, and our quality of life. Execution of recommendations in this report will play an essential role in ensuring the vitality of our Nation's NIT endeavors and enabling us to address our priorities and meet our challenges.

Crosscutting Themes

The five broad themes listed below recur throughout this report, and are of great importance to the future of all Federal agencies:

- Data volumes are growing exponentially. There are many reasons for this growth, including the creation of nearly all data today in digital form, a proliferation of sensors, and new data sources such as high-resolution imagery and video. The collection, management, and analysis of data is a fast-growing concern of NIT research. Automated analysis techniques such as data mining and machine learning facilitate the transformation of data into knowledge, and of knowledge into action. Every Federal agency needs to have a "big data" strategy.
- Engineering large software systems to ensure that they are secure (behaving as expected in the presence of an adversary) and trustworthy (behaving as expected in the absence of an adversary) remains a daunting challenge. The growing complexity of the systems we are building and our increasing societal reliance upon them outpace our ability to reason about them, and to engineer them to be secure and trustworthy.
- As NIT increasingly pervades daily life, systems are storing and processing a greater volume and diversity of private information about individuals. Privacy is a critical issue in all societal applications of NIT – most obviously in areas such as healthcare and electronic commerce, but also in areas such as energy, transportation, and education. Privacy challenges do not and must not require us to forgo the benefits of NIT in addressing national priorities. Rather, we need a practical science of privacy protection, based on fundamental advances in NIT, to provide us with tools we can use to reconcile privacy with progress.
- Interoperable interfaces – the means by which components of the smart grid can talk to each other, for example, or by which electronic health records can be shared and added to by many parties – are an important stimulus to technology innovation and adoption. Optimally, these interfaces would be *open*: anyone may create products that use the interfaces without paying fees; and a public, transparent process is used to establish and revise the standards that define the interfaces.
- The NIT supply chain is vulnerable. The hardware and software components used to build systems are sourced worldwide. We must anticipate and be prepared for various forms of threats to supply, quality, and security.

PCAST NITRD Program Review Working Group

Co-Chairs

David E. Shaw*
Chief Scientist, D. E. Shaw Research
Senior Research Fellow, Center for Computational Biology and Bioinformatics
Columbia University

Edward D. Lazowska
Bill & Melinda Gates Chair in Computer Science & Engineering
Director, eScience Institute
University of Washington

Members

Francine Berman
Vice President for Research
Professor of Computer Science
Rensselaer Polytechnic Institute

Stephen Brobst
Chief Technology Officer
Teradata Corporation

Randal E. Bryant
Dean of the School of Computer Science
Carnegie Mellon University

Mark Dean
IBM Fellow and Vice President
IBM Research

Deborah Estrin
Jon Postel Professor of Computer Science
Director, Center for Embedded Networked Sensing
University of California, Los Angeles

Edward W. Felten
Professor of Computer Science and Public Affairs
Director, Center for Information Technology Policy
Princeton University

Susan L. Graham
Pehong Chen Distinguished Professor of Electrical Engineering and Computer Science Emerita and Professor of the Graduate School
University of California, Berkeley

William Gropp
Paul and Cynthia Saylor Professor of Computer Science
Deputy Director for Research, Institute for Advanced Computing Applications and Technologies
University of Illinois Urbana-Champaign

Anita K. Jones
University Professor Emerita
University of Virginia

* PCAST member

Michael Kearns
Professor of Computer and Information Science
Founding Director, Market and Social Systems Engineering Program
University of Pennsylvania

Paul Kurtz
Managing Partner
Good Harbor Consulting, LLC

Robert F. Sproull
Vice President and Director of Sun Labs
Oracle

Staff

Mary Maxon
Deputy Executive Director
President's Council of Advisors on Science and Technology

Table of Contents

1. Introduction

This report assesses the status and direction of the Federal Networking and Information Technology Research and Development (NITRD) Program. Responsibility for assessment of the NITRD Program, originally assigned to the President's Information Technology Advisory Committee (PITAC), was transferred to the President's Council of Advisors on Science and Technology (PCAST) in 2005.

The phrase "the NITRD Program" has two meanings, both of which are addressed in this report:

- "The NITRD Program" can refer to the mechanism by which the Federal Government coordinates its unclassified research and development (R&D) investments in Networking and Information Technology (NIT). This coordination takes place under the aegis of the NITRD Subcommittee of the National Science and Technology Council's (NSTC) Committee on Technology. The NITRD Program has 14 member agencies, with additional agencies participating in specific NITRD activities. The member agencies report aggregate NIT R&D investments in excess of $4 billion annually, with the largest investments reported by the National Institutes of Health (NIH) and the National Science Foundation (NSF) (roughly $1 billion each), followed by the Office of the Secretary of Defense and the Department of Defense Service research organizations (OSD/DoD), the Department of Energy (DoE), and the Defense Advanced Research Projects Agency (DARPA) (roughly $500 million each)[11].
- "The NITRD Program" also can refer to the unclassified Federal NIT R&D portfolio itself – that is, the ensemble of unclassified research and development efforts in NIT supported by the Federal Government rather than the coordination effort for this NIT R&D.

The Federal Government's investment in NIT R&D dates from the birth of the field more than 60 years ago. NITRD as a coordination effort, though, had its genesis in the High-Performance Computing Act of 1991 – "An Act to provide for a coordinated Federal program to ensure continued United States leadership in high-performance computing." Its scope was broadened by the Next Generation Internet Research Act of 1998, and again by the America COMPETES Act of 2007.

To assist in this assessment, PCAST appointed an expert 14-member Working Group, which consulted with more than 50 individuals and drew upon a number of recent studies.

1.1 The Organization of this Report

Sections 2 and 3 of this report set the stage for the assessment. Section 2 discusses the profound impact of NIT R&D – arguably unique among all fields of science and engineering, and arguably among the best investments that our Nation has made. Section 3 describes a number of recent technological and societal trends that have dramatically broadened and deepened the role of NIT, placing the advancement and application of NIT squarely at the center of our Nation's ability to achieve essentially all of our priorities and to address essentially all of our challenges.

---

11. Networking and Information Technology Research and Development Supplement to the President's FY 2011 Budget, (February 2010) (page 21). http://www.nitrd.gov/pubs/2011supplement/FY11NITRDSupp-FINAL-Web.pdf

Section 4 discusses in some detail the essential role that advances in NIT will play in achieving America's priorities in six important areas: Health, Energy and Transportation, National and Homeland Security, Discovery in Science & Engineering, Education, and Digital Democracy. In each of these areas, we first describe a vision of the future that our Nation must create, and then describe the role that advances in NIT – *true advances*, rather than merely the application of existing NIT systems – will play in creating that future. Where appropriate, we also note beneficial near-term initiatives. Section 5 presents recommendations developed from that discussion.

Section 6 discusses necessary advances in the research frontiers of NIT: NIT and People, NIT and the Physical World, Large-Scale Data Management and Analysis, Trustworthy Systems and Cybersecurity, Scalable Systems and Networking, Software Creation and Evolution, and High Performance Computing. Progress on the NIT research frontier topics is an essential contributor to progress on NIT for national priorities. Section 7 presents recommendations.

Section 8 discusses technological and human resource requirements for progress in NIT. With respect to the former, we note that shared NIT infrastructure has become essential to research in virtually all fields, and that NIT infrastructure that supports research in other fields, while a crucial component of R&D in those fields, is not NIT R&D. With respect to the latter, we note that NIT is the dominant factor in America's science and technology employment, and that the gap between the demand for NIT talent and the supply of that talent is large and will continue to be so. Increasing the number of graduates in NIT fields at all degree levels must be a national priority. Section 9 presents recommendations.

Section 10 discusses the strengths and limitations of the NITRD coordination process and structure. NITRD is a successful and valuable coordination mechanism, but there are limits to what it can be expected to achieve. Ongoing strategic advice and leadership via a separate mechanism is necessary. Section 11 presents recommendations.

Section 12 closes the report with a discussion of the complementary roles of federally funded research and industry R&D in NIT. Industry has made, and continues to make, crucial contributions to NIT R&D. It is important, however, not to equate the very large industry R&D investment in NIT with fundamental research of the kind that is carried out in universities and a small number of industrial research labs. Fundamental research with the potential for future transformational application represents a small fraction of overall industry R&D in NIT. We discuss certain well-established principles of economic theory that underlie the need for Federal investment in NIT R&D.

1.2 A Preview of the NITRD Portfolio and the NITRD Coordination Process and Structure

As discussed in Section 10, the NITRD Program currently includes eight Program Component Areas (PCAs)[12]. These PCAs represent NIT R&D budget categories, and map fairly directly onto a set of Interagency Working Groups and Coordinating Groups that carry out much of NITRD's coordination work.

12. http://www.nitrd.gov/subcommittee/program.aspx

The extraordinary payoff from past Federal investments in NIT R&D is discussed in Section 2. However, as discussed in Section 3, the landscape is changing rapidly and dramatically. The NITRD portfolio must change, too. We have chosen to focus this assessment less on NITRD as it is, and more on NITRD as it should be:

- Increased emphasis on advances in NIT necessary to achieve America's priorities, as outlined in Sections 4 and 5.
- A new view of the core of the field, as outlined in Sections 6 and 7.
- The need for larger and more multidisciplinary teams of researchers for longer periods of time, required by both of the above.
- The need for a broad, high-level standing committee of academic scientists, engineers, and industry leaders dedicated to providing sustained strategic advice in NIT, as outlined in Sections 10 and 11.

These changes will require additional resources – some combination of new funds and redirected existing funds – along with additional attention by multiple Federal agencies. Of crucial importance is our finding that the Nation is investing far less in NIT R&D than is shown in the Federal budget. Within the NITRD crosscut budget, there is widespread confusion between spending on NIT that supports R&D in other fields, and spending on R&D in the field of NIT itself. Investments in NIT R&D are investments that are broadly transformational. Section 12 discusses the necessary level of investment further, as well as the essential Federal role.

2. The Impact of Networking and Information Technology

As a field of inquiry, NIT has a rich intellectual agenda – as rich as that of any other field of science or engineering. In addition, NIT is arguably unique among all fields of science and engineering in the breadth of its impact:

Advances in NIT are crucial to achieving our major national and global priorities in areas such as energy and the environment, education and life-long learning, healthcare, and national security. Tackling these challenges requires *advances* in NIT that go well beyond the application of existing systems.

Advances in NIT accelerate the pace of discovery in nearly all other fields. The impact of computer simulations of real-world problems on the physical sciences and engineering has been profound. New paradigms, such as the automated analysis of vast amounts of data now emerging because of dramatic progress in sensors and sensor networks, will revolutionize many more fields.

Advances in NIT are essential to achieving the goals of open government. NIT touches everyone's lives, changing the way we live, work, learn, and communicate. Increasingly widespread use of NIT has important public policy implications, ranging from e-voting and identity management to the nature and global spread of democracy.

Impressive as these examples are, they point to a still larger and more fundamental theme. Unless the Nation's economy continues to thrive, none of our goals in energy, healthcare, education, national security or other crucial areas will be achievable – and the expansion and advancement of NIT are key drivers of America's economic competitiveness.

The enormous economic impact of NIT derives not only from the growth of the NIT industry itself, but to an even greater extent from NIT-enabled productivity gains across the entire economy. The development and application of NIT-related systems, services, tools and methodologies have boosted U.S. labor productivity more than any other set of forces in recent decades. Advances in NIT, deployed pervasively throughout the U.S. economy, have helped U.S. workers become the world's most productive and have enabled the U.S. to remain one of the world's most competitive economies[13]. Advances in NIT are central to achieving the goals set out in the President's *Strategy for American Innovation*[14], which include investing in the building blocks of American innovation, promoting competitive markets that spur productive entrepreneurship, and catalyzing breakthroughs for national priorities.

While the fruits of NIT advances are most evident in the rise of the modern technology sector – now-familiar corporate names such as Apple, Facebook, Google, Intel, Microsoft, and others – the impact in other areas of the economy has been equally dramatic. Companies as diverse as FedEx and Walmart,

---

13. The United States ranks second only to Switzerland in the most recent competitiveness rankings of the World Economic Forum: Global Competitiveness Report, 2009-2010. World Economic Forum, Geneva, Switzerland. (2009).

14. Executive Office of the President. (September 2009). *A Strategy for American Innovation: Driving Towards Sustainable Growth and Quality Jobs.* http://www.whitehouse.gov/assets/documents/SEPT_20_Innovation_Whitepaper_FINAL.pdf

although they provide services that existed long before the current technology boom, have used advances in NIT to revolutionize their industries, boosting operational efficiency and economic output to an unprecedented extent. Small and mid-size companies have also gained new capabilities and efficiencies through the use of NIT. Whether it is access to powerful yet affordable systems that allow virtual prototyping for parts suppliers, point-of-sale systems that allow for precise inventory controls, or simply the availability of sites like Etsy.com that make it easy for communities of artists to reach customers, advances in NIT empower U.S. businesses, augment their competencies, and enable them to compete successfully in an increasingly global economy.

The still-growing economic impact of NIT has been compellingly documented in a number of important studies. The influential "Digital Prosperity" report[15] details the profound economic benefits of NIT in five distinct categories: productivity, employment, efficient markets for goods and services, higher quality goods and services, and innovation producing new goods and services. Analyzing the tremendous recent growth of productivity in the U.S. economy, Harvard economist Dale Jorgenson and colleagues[16] conclude that the use and production of NIT accounted for "roughly two-thirds of the post-1995 step-up in labor productivity growth."[17]

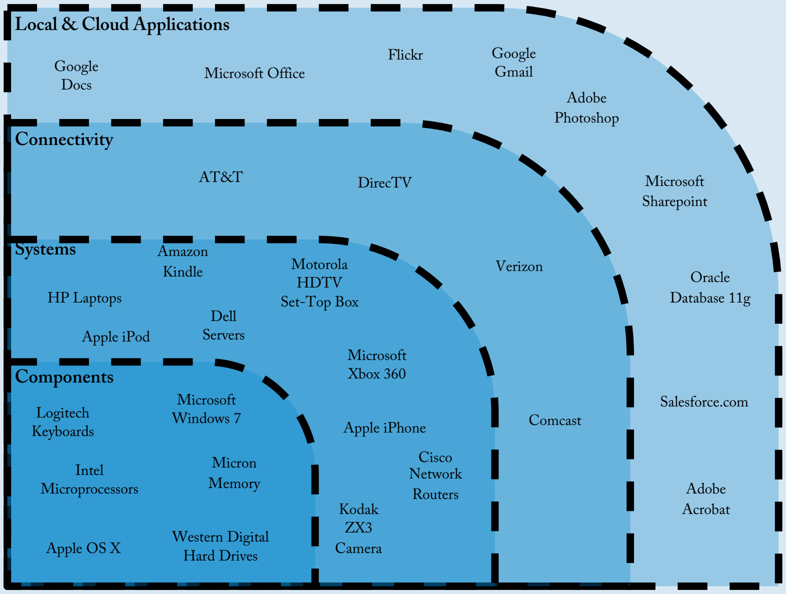
Even the most casual observer of business, industrial, and economic trends over the past several decades can see that the number of companies that benefit from and rely on NIT in crucial ways has grown dramatically. NIT has transformed every scale of organization and every aspect of production: R&D in corporate labs and universities; extraction and processing of raw materials; supply chain management; inventory control; the management, human resource, and marketing functions of an organization; assembly and distribution of end products; customer support; even the social communities that form around a product

15. The Information Technology & Innovation Foundation. (March 2007). *Digital Prosperity: Understanding the Economic Benefits of the Information Technology Revolution.*

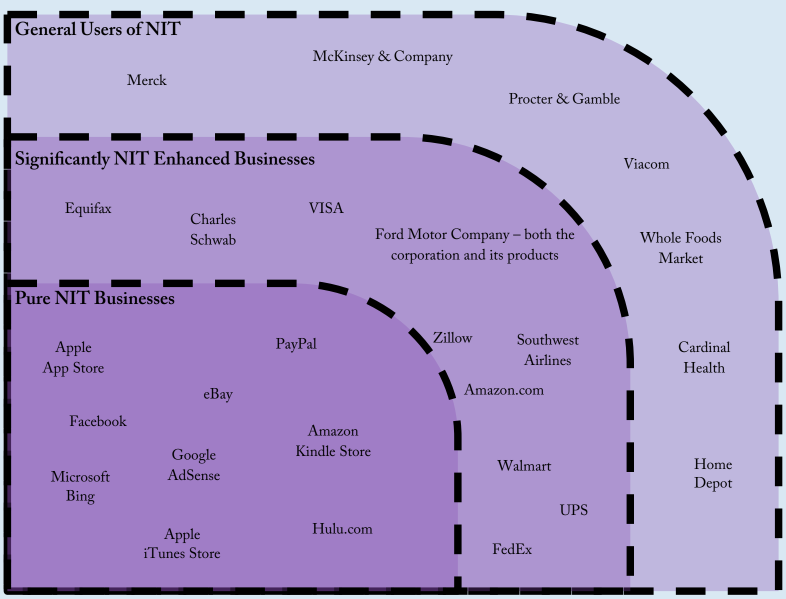
16. Jorgenson, Dale W., Mus S. Ho, and Kevin J. Stiroh. (2005). *Productivity, Volume 3: Information Technology and the American Growth Resurgence.* MIT Press..

17. Oliner, Stephen D., Daniel E. Sichel, and Kevin J. Stiroh. (2007). "Explaining a Productive Decade." Federal Reserve Board.

The Pervasiveness of NIT from the Consumer Perspective



These examples of NIT products and services illustrate their ubiquity and show how advances in particular aspects of NIT influence a broad swath of the tools that we use at work and at home. NIT components, which include hardware and lower-level software, combine to create cohesive, unified NIT products, most of which are not thought of as computers per se. Digital connectivity links systems, enabling transparent choice between local and cloud applications for a wide variety of purposes.



These examples show how companies are embracing NIT at a relentless pace. Pure NIT businesses focus on providing digital services and products. Significantly NIT enhanced businesses use NIT advances to greatly improve their services and products. General users of NIT include consumers and firms that utilize NIT in their daily workflows.

NIT has had equally dramatic effects on the service sector. Consumers today obtain a remarkable array of services through network-connected devices: banking and other financial services; news reporting; travel and entertainment ticketing; on-line video entertainment; and many more. NIT is also essential to efficient service delivery of more traditional kinds: route planning for transportation and logistics; retail inventory control; managing electricity production and distribution; and so on. (See sidebar, "The Pervasiveness of NIT from the Consumer Perspective," preceding page).

For all of the reasons described above – because advances in NIT are crucial to achieving our national priorities, accelerating the pace of discovery in nearly all fields, making open government a reality, and powering America's economic competitiveness – it is essential that America continue to lead the world in NIT.

In assessing how to maintain this leadership, a common fallacy is to overestimate the role of technology development and underestimate the role of fundamental research. In fact, computer science research, carried out to a great extent in America's research universities with funding from Federal agencies such as NSF and DARPA, lies at the heart of our Nation's leadership. This research – ranging from the design of computers and networks to robotics, software, and algorithms – has time and again made possible entirely new product categories that have become billion-dollar industries. The "extraordinarily productive interplay of federally funded university research, federally and privately funded industrial research, and entrepreneurial companies founded and staffed by people who moved back and forth between universities and industry"[18] has been well documented.

Finding: The extraordinary accomplishments of America's NIT research and development efforts are amply evident, and have been authoritatively documented by the National Academies[19,20], the President's Information Technology Advisory Committee (PITAC)[21], and others, with PITAC referring to the "spectacular return" on the Federal research investment. The development and application of NIT-related systems, services, tools, and methodologies have boosted U.S. labor productivity more than any other set of forces in recent decades.

Today, the NITRD Program is the custodian of this Federal research portfolio.

18. National Academies Press. (1999). *Funding a Revolution: Government Support for Computing Research.*

19. National Academies Press. (1995). *Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure.*

20. National Academies Press. (2003). *Innovation in Information Technology.*

21. President's Information Technology Advisory Committee Report to the President. (1999). *Information Technology Research: Investing in Our Future.*

3. Recent Technological and Societal Trends

Progress in NIT continues unabated. The ever-changing technological landscape, coupled with significant societal trends (some induced by technological trends, some evolving independently of them), has dramatically broadened and deepened the role of NIT.

During its first half-century, computing saw four generations of hardware (vacuum tube, transistor, integrated circuit, and microprocessor) and four major classes of systems (mainframe, minicomputer, workstation, and personal computer). Networking has grown from proprietary interfaces and acoustic modems to high-bandwidth system-area, local-area, and wide-area networks. Software has advanced from small programs written in machine language and assembly language to complex systems comprising millions of lines of high-level language code that underlie defense, commerce, and communication.

The impact of networking and information technology during this period has been extensive and profound. As recently as 1995, however, the "umbrella" under which all Federal support for NIT R&D was carried out was the High Performance Computing and Communications Initiative (HPCCI). The National Academies, in an assessment conducted in 1995, observed:

> The HPCCI is the current manifestation of the continuing government research program in information technology, an investment that has been ongoing for more than 50 years. Although it emphasizes research in high-performance computing and communications, the HPCCI now has in its budget nearly all of the federal funding for computing research of any kind. The wisdom of this arrangement is doubtful.[22]

In response to this assessment, and to a 1999 PITAC report[23] that struck a similar chord, the purview of the Nation's NIT R&D portfolio was broadened considerably. Even today, though, this portfolio bears the clear stamp of its high performance computing and communication (HPCC) origins.

Profound technological and societal trends in recent years make it even more essential that this broadening be accelerated. Among these trends are:

- Ubiquitous connectivity at ever-increasing rates; nearly universal reach of the Internet; very inexpensive long-distance communication. Broadband Internet is available to most Americans in their home, library, or school. Long-distance email and voice communications are now nearly free.
- Mobile NIT and location-based services. Smart phones such as the iPhone and Android phones are essentially small networked computers, with a wide variety of software available in app stores. GPS-equipped phones adapt to the user's location, helping the user navigate and find nearby people, places and things.

---

22. National Academies Press. (1995). *Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure.*
23. President's Information Technology Advisory Committee Report to the President. (1999). *Information Technology Research: Investing in Our Future.*

- NIT-driven transformation and convergence of communications, entertainment, journalism, and public discourse, enabling an explosion of online content and service offerings. NIT brings a boundless library of books, newspapers, and video into every home and school, transforming political and social debate, and giving any entrepreneur with a great idea access to a global audience.
- Exponentially increasing volumes of data – from ubiquitous sensors, from higher-bandwidth sensors, from complex simulations, and from the creation of all information in digital form – coupled with algorithms that mine this data for knowledge. The result is striking improvements in predictive capabilities in many areas, science and online advertising being just two. Inexpensive cameras, computers, and networks let people create, edit, and distribute high-quality music and movies, limited only by their talent. Even the smallest businesses can capture, analyze, and visualize detailed data about their operations and revenue.
- Cloud computing, in which users have access through the Internet to shared computing resources, software, and data, makes massive computational and storage capability available to everyone at reasonable cost. The world's most powerful computers, by many important measures, are no longer dedicated to scientific applications. Cloud services such as Gmail and Flickr provide stable, ubiquitous access to users' information, and the same industrial-strength infrastructure is available inexpensively to small entrepreneurs.
- A significant technology flow *up*, from commodity components (processors, graphics processing units, etc.) to high-end systems. High-end computers are increasingly built by aggregating large numbers of ordinary computers, or even video-game consoles.
- The end of performance increases in individual processors, making the use of numerous processors in parallel systems indispensable, as well as the increasing need for system designers to minimize power consumption and heat generation. Improving battery life and energy efficiency can be at least as important as making devices faster.
- The entry into the mainstream of artificial intelligence technologies: human language technology (speech, language, translation), information extraction, machine learning, robotics. Intelligent search helps us find information; language translation lowers cultural barriers; and speech recognition connects the worlds of voice and text.
- The emergence of social computing, communication, and interaction: social networks, crowdsourcing, coordination at a distance. The way people interact has been transformed, the data we have from and about people is transformational, and the ability to crowdsource knowledge creates tremendous new opportunities. People around the world can collaborate to create an encyclopedia; we can "friend" and "follow" our relatives, colleagues, and long-lost acquaintances.
- NIT embedded in physical devices, military technology, and consumer products. Cars use NIT to control antilock brakes, optimize engine performance, decode radio signals, and provide secure keyless entry, among other functions. Smaller devices such as thermostats and coffeemakers rely on NIT to increase functionality and convenience.

- The transformation of commerce by NIT, and the movement online of many economic transactions (from purchasing to banking to voting to health), creating the need for dramatic improvements in privacy and security.
- The rise of cyber-crime, online fraud, and identity theft, and the growing threat of cyber-warfare.
- NIT-driven globalization of production, consumption, workforce, social interaction, innovation; one key aspect of this is the impact on our supply chain, with both hardware and software NIT components sourced worldwide.
- Pressing national and global priorities – such as energy and the environment, education and life-long learning, healthcare, and national security – and the increasingly central role of NIT in addressing these challenges.

The Nation's strategy in NIT R&D must be agile – it must be responsive to these recent technological and societal trends, and to those that will surely follow.

NIT and the Retail Revolution

The past few decades have brought two startling innovations in retailing: the Big Box store and e-Tailing. A Big Box store is a physical establishment that lures consumers by offering a wide variety of goods at low prices. In e-Tailing, consumers shop online and have their purchases delivered to their homes. Unlike traditional retailers, who buy merchandise wholesale and resell it at their physical stores, Big-Box stores and e-Tailers use modern information technology to construct efficient "pipelines" between producers and consumers, removing intermediaries and driving down prices.

When a customer checks out of a Big Box store, point-of-sale terminals capture data on the merchandise just purchased. This information goes upstream to the producers, who use it to adjust their own supply chains in response to demand. The shelf space in a Big Box store becomes the producer's "point of presence," a portal through which the producer sells merchandise directly to consumers. In some cases, the Big Box store never actually owns the merchandise: the producer owns it and sells it directly to the consumer, while the Big Box retailer collects a fee for making the sale possible.

The e-Tail customer goes on-line to find merchandise, makes an order, and pays using a credit card or PayPal. The customer can be confident purchases will arrive quickly because of a revolution in freight logistics called time-definite delivery. In an earlier era, a shipper and the customer were told only that a package would embark on its delivery route at a certain time. In time-definite delivery, the shipper and the customer are told when the package will arrive at its destination. This change came about through the application of modern information technology to package tracking.

Big-Box stores and e-Tailing have revolutionized retailing, a sector of the U.S. economy that's worth more than $4 trillion each year. This revolution could not have happened without modern information technology.

4. The Role of Advances in NIT in Achieving America's Priorities

In this section, we paint the future in six priority areas, assuming that adequate investments are made to ensure America's continued leadership in NIT: Health, Energy and Transportation, National and Homeland Security, Discovery in Science & Engineering, Education, and Digital Democracy. In each of these areas, we first describe a vision of the future that our Nation must create, and then describe the role that advances in NIT – *true advances*, rather than merely the application of existing NIT systems – will play in creating that future. Where appropriate, we also note beneficial near-term initiatives. There are many opportunities for agencies to partner on investments in NIT R&D related to their missions, since research advances in aspects of NIT often will advance the missions of multiple agencies.

Two general findings apply to our evaluation of NIT's role in all six areas:

Finding 1: Recent technological and societal trends place the advancement and application of NIT squarely at the center of our Nation's ability to achieve essentially all of our priorities and to address essentially all of our challenges. America is – and must continue to be – the world leader in NIT. It is crucial that the United States continue to innovate more rapidly and more creatively than other countries in important areas of NIT in order to sustain and improve our quality of life. New initiatives and investments in NIT research and development are required in order to achieve America's priorities and advance key NIT research frontiers.

Finding 2: Federal agencies vary greatly in their appreciation of the dramatically expanded role that advances in NIT – *true advances*, rather than the application of existing NIT systems – play in achieving our Nation's priorities, meeting our challenges, and shaping our world. Some agencies have not yet recognized the extent to which their abilities to accomplish their missions are inextricably linked to advances in NIT.

4.1 NIT for Health

Achieving the Nation's goal of improving people's health and increasing the quality of treatment outcomes while also containing healthcare costs requires fundamental change in our healthcare system. New approaches are needed to manage and prevent chronic diseases, which consume an increasingly large share of healthcare expenditures[24]. Individuals, along with their families and friends, must share with health professionals the responsibility for improving health and treating disease. Improvements in health and quality of life, particularly as we age, depend upon knowledge and information that can drive appropriate actions.

24. For example, see http://healthcarecostmonitor.thehastingscenter.org/kimberlyswartz/projected-costs-of-chronic-diseases/ and http://www.ge.com/visualization/chronic_diseases/index.htm

Advances in NIT will play an essential role in achieving the Nation's health goals. Thus far, NIT has thoroughly infiltrated the business and some administrative aspects of healthcare, but has not come close to fulfilling its potential. The current push for meaningful use of electronic health records, the increasing use of NIT as a surgical tool, the revelation of the structure of the human genome and the biomedical insights that flow from it, and the increasingly broad and pervasive access to health information online are all promising starting points for future advances. The following examples offer glimpses into a new vision of healthcare made possible by innovations in NIT and by innovative approaches to the use of NIT:

- *Comprehensive lifelong individual health records*: Consider a single, patient-centered view of all health-related information pertaining to a person's entire life. The record is complete across time, caregivers, facilities, ailments, and data locations. It includes not only diagnostic and treatment history, but also a genetic profile, mental and psychological characteristics, behaviors, and health-relevant events such as exposures to risks. It includes both explicitly recorded information and information inferred from contextual analysis, perhaps retrieved many years later when relevance has been established. Powerful analyses and tools, tailored to the particular user, extract useful information from the record. Powerful abstraction and data mining tools provide concise summaries of the information required for a given purpose. Views are tailored to the needs of the user. Those tools transcend multiple natural languages and diverse non-textual forms of information. Personal privacy is protected; the trustworthiness of the record is maintained. The contents of this electronic health record do not come from health professionals alone. People contribute information about themselves, or about those they care for. NIT-enabled observation supplies additional data. For example, sensors and other observational tools maintain a robust and continual assessment of an individual's physical and mental state. The assessment incorporates monitoring and sensing of vital signs, chemistry, mobility, and behavior; human observation, including prompted self-reporting; integrated imaging of internal physiological systems; continuing analysis and comparison to detect significant changes; and embedded knowledge of variations from the norm. The results of the assessment are incorporated in the lifelong health record, but might also trigger more immediate action by the individual, clinicians, or family members.
- *Empowerment through health knowledge from myriad sources*: Everything about health, disease, diagnosis, and treatment – scientific, clinical, and experiential – derived from all possible sources in the world, is synthesized and available electronically in an appropriately comprehensible way. Information might come from the biological and medical literature, from social computing, from mining and analysis of aggregated health records, from simulated or actual clinical trials, from public health studies, from epidemiological studies, from genomics, and so on.[25] Information is validated and constantly updated as scientific and medical knowledge progresses. Access can be appropriately tuned to the cognitive and educational characteristics of both professional and non-professional users. Powerful analyses and tools exist that can apply information to individuals, populations, and public health situations. As health knowledge is mapped to individuals

25. An interesting recent example is the use of query data to detect influenza epidemics. See http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en/us/archive/papers/detecting-influenza-epidemics.pdf

by the logical aggregation of their lifelong health information, event monitoring techniques, carried out with the participation of health professionals, detect situations requiring attention and trigger proactive interventions.

- *Tools for informed, personalized healthcare*: Individuals and their extended families are active and responsible participants in their own wellness and healthcare. Easy and comprehensible access to all individual and collective information, along with NIT tools to analyze the information, allows individuals to determine appropriate actions and carry them out. As individuals gain command of their own health knowledge, face-to-face doctor visits might become less frequent. Instead, a variety of multi-modal NIT systems, among them offline communication such as email and online communication such as video conferencing, allow people to have easier and more frequent non-emergency interactions with health professionals. Individuals have ready access to reliable, high-quality "self-help" systems, tailored to their particular needs. Assistive technologies of all kinds are available for disabled, aging, and ill individuals and for their non-professional caregivers. The technologies provide physical assistance – including "smart" prosthetics, auditory and visual aids, and motor aids – and also cognitive assistance, behavioral aids, and monitoring. Robotic devices assist the elderly and the infirm as well as those that care for them.

The role of NIT in achieving this vision. The business side of clinical care already makes wide use of NIT. NIT is the basis for electronic health records, for patient web sites, for image-guided surgery, and for biomedical discovery. We are moving toward a future in which all health information sources, all professional and non-professional care, and most non-pharmacological interventions – although necessarily informed by humans – will be NIT-based.

Interoperable Interfaces and Demonstration Testbeds Drive Innovation and Economic Growth

The impact of NIT on key national priorities, including healthcare, energy, and transportation, will be magnified and accelerated through the use of *well-defined* and *interoperable interfaces*, and *demonstration testbeds*. These are mechanisms that breed unfettered innovation.

An *interface* enables one NIT component to connect to and work with others, whether through a network, by exchanging data, or by executing programs. Examples of widely used interfaces include the Internet communication protocols, the HTML document format, and the Microsoft Windows and Apple iPhone software platforms. These interfaces have all been essential to the development of multi-billion dollar NIT industries: the Internet, the World Wide Web, the personal computer, and smart phones.

Interoperable interfaces allow equipment or software from different vendors to work together or communicate. They allow new, innovative creations to work with older, established services. For example, innovation in Web browsers has been possible in part because new browsers use the established HTML document format and HTTP network protocol, and thus are able to access all existing Web content. Innovation has also proceeded on the other side of the interfaces – in Web servers – and in similar fashion a new server implementation works with old browsers because of the standardized interfaces.

Interoperable interfaces arise and evolve in various ways. In some cases, an *open standards process* defines and modifies a standard interface definition. A public, transparent standards process considers the requirements and contributions of the entire user community, and is designed to ensure that the standard can be used freely and that its use or implementation infringes on no patents and requires no payments of royalties or license fees. The Internet communication protocols and the HTML document format are examples of such interfaces.

Interoperable interfaces can also spring from proprietary contributions. A commercial developer of an interface may seek to expand its use by contributing its definition – and associated intellectual property – to an open standards process. Or it can offer licenses to its intellectual property on reasonable and non-discriminatory terms. Sometimes a consortium is formed to act as steward of an interface – this is the case with the Bluetooth wireless protocol[26]. Commercially successful designs have great value and can accelerate adoption by a wider community if they are available with low barriers to entry.

Innovation requires that standards evolve, and there are ways to define standards so that new or experimental functions can be introduced quickly without interference with interoperability of the established standard. The ultimate aim is to develop an *ecosystem of innovators* – commercial or not – that can rapidly and inexpensively adapt to meet pressing needs.[27] The ecosystem that emerged from the open interfaces of the Internet has led to U.S. domination of innovation in that area.

Two of our highlighted national priorities – health NIT, and energy and transportation – are now at stages where developing interoperable interfaces is critical to creating active innovation ecosystems.

The future of efficient health services requires an interface definition for electronic health data and for mechanisms to allow providers and patients to share data. The system must work as well for individual self-employed physicians as it does for regional healthcare organizations. An interoperable specification will spur diversity and innovation in the creation of software that lets doctors and patients make best use of healthcare data.

The smart grid will depend on interfaces that allow home appliances, electric cars, homeowners, building operators, electricity generators, distribution network operators and many other participants to interact with the grid and make efficient energy choices. Likewise, smart transportation systems will require interfaces that permit transport providers to supply information about their services, highways to report their status, and users to explore alternatives.

Testbeds serve to stimulate innovation and competition by testing interoperable interfaces. They demonstrate that different products and services interoperate as intended and help uncover minor difficulties or inconsistencies in interface definitions before they are deployed for public use. They allow vendors and researchers to demonstrate their solutions in a public forum.

Interoperable interfaces and demonstration testbeds will lead to innovative ecosystems that rapidly develop and deploy new technologies for important national priorities.

26. http://www.bluetooth.com/English/SIG/Pages/default.aspx

27. *Roadmap for Open ICT Ecosystems,* Berkman Center for Internet & Society at Harvard Law School, http://www.apdip.net/resources/policies-legislation/guide/Berkman-Roadmap4OpenICTEcosystems.pdf

The computing and storage infrastructure supporting this vision must provide for interoperability (see the sidebar "Interoperable Interfaces and Demonstration Testbeds Drive Innovation and Economic Growth" on previous two pages). It must provide flexible, reliable, and ubiquitous access, and must allow easy non-disruptive integration of best-practice components. All aspects of the infrastructure – access to information, the ability to customize and to add capabilities, the nature of the abstracting, navigating, and decision support tools, and the incorporation of particular processes – must scale down to the needs of people of modest education as well up to the demands of the most capable professionals. It must be possible to create smaller, simpler systems within the infrastructure to meet the needs of individual self-employed physicians. At the same time, the ability to serve regional and national healthcare organizations by enhancing the infrastructure – substituting, modifying, and adding capabilities – must be readily available. The infrastructure must support national-scale healthcare endeavors, for example comparative effectiveness research, and it must function in disaster and emergency conditions.

The actual capture of medical information, either by the provider or the patient, should be accomplished primarily as a byproduct of the routine use of decision support and health maintenance tools rather than by direct entry. Methods for collecting information should include automated capture of measurements, sensor output, and other information that is generated in digital form; analysis of spoken dialogue; and easy-to-use tools for self-reporting. Capture of information, for instance during a patient-caregiver interaction, must be non-intrusive and minimally invasive. These innovations can greatly improve the value of the time that physicians spend with patients. The "understanding" provided by heavy-duty analyses of all this captured information must lead to both human and automated actions on behalf of individuals and larger groups. It is application-motivated advances in NIT that provide these capabilities.

Specific needed advances in NIT. Significant progress in using the information contained in health records, in the literature, and in observational and real-world data depends on the ability to extract meaning from the information. Without that ability, we cannot capitalize fully on the investment being made in electronic health records. Research is needed in methods for the semantic analysis of natural language (i.e., spoken and written records), for the extraction of semantic information from raw data, and for translating among different terminologies and categorizations used for the same kinds of information. Making use of the potentially huge quantities of health information of all kinds requires novel techniques to abstract higher-level concepts and explanations from lower-level data, to determine relevance in context, and to resolve conflicting information. Fruits of the research into data and knowledge management technologies outlined in the Large-Scale Data Management and Analysis section of this report will need to be put to use for health-related NIT problems.

Given the myriad sources of health knowledge, methods must be devised for the automated and semi-automated exploration of alternative diagnoses and treatments. More generally, methods are needed that can specialize general health information to particular situations (e.g., methods that apply the results of empowerment through health knowledge from myriad sources to the information from comprehensive lifelong individual health records in order to support tools for informed personalized health). Scientifically sound risk analysis techniques are needed to assist human decision makers.

Furthering the specific goals for personalized health requires the creation of broadly usable and affordable technologies for chronic disease prevention, management, and research. Advances are needed in the modeling of interpersonal interaction and in the understanding of human behavior and motivation to allow both physicians and patients to gain maximum advantage from tools that facilitate personalized healthcare. Given the broad availability and relatively low cost of personal computing devices, it will be fruitful to explore novel uses of them, including portable and wearable devices, to achieve personalized healthcare goals. Such research is in its infancy.

For the treatment of chronic conditions, including those due to aging, progress in the discovery and development of advanced assistive technologies is essential. Further advances are needed in image analysis, in new kinds of robotics, in non-intrusive monitoring and response technologies, and in all manner of cognitive assistance.

Using today's NIT. As the recent PCAST Health Information Technology report explains[28], existing NIT technology and understanding can implement location-independent access to all health-related information and analyses that are in digital form, share the information and analyses among organizations, and provide appropriate access control and auditing. That report observes that methods for universal information exchange must be incorporated into current planning for meaningful use of electronic health records. It urges the use of metadata-tagged data elements, the establishment of initial minimal standards for the metadata associated with tagged data elements, the development of a roadmap for more complete standards over time, and the rapid mapping of existing semantic taxonomies into tagged data elements. Demonstration testbeds (see the sidebar on page 15) available to the research community as well as to vendors would allow the creation of demonstration systems that show how all these capabilities can be achieved in a location-independent way, and would provide a plug-and-play environment in which to evaluate new components in context and at scale. To achieve the full benefits of such testbeds, some aspects of such systems will require new policies on the sharing of information while protecting both personal privacy and intellectual property.

4.2 NIT for Energy and Transportation

Through the American Recovery and Reinvestment Act (AARA), the Administration recognized the need for accelerated progress in a number of priority areas, including energy and transportation. (For a review, see *The Recovery Act, Transforming the American Economy through Innovation*[29].) The Nation's priorities for energy and transportation are closely linked: in both areas, the new, more efficient designs and operation that we need cannot be achieved without major contributions from NIT. Advances in NIT can assist in providing the same services at lower cost. For example, a recent National Academies study[30] concludes that even the aggressive deployment of *today's* technologies in the buildings, transportation, and industrial sectors could reduce energy use in 2030 by about 30 percent relative to current projections; full deployment in buildings alone could eliminate the need to construct any new

28. President's Council of Advisors on Science and Technology. (2010). *Realizing the Full Potential of Health Information Technology to Improve Healthcare for Americans: The Path Forward.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-health-it-report.pdf

29. http://www.whitehouse.gov/sites/default/files/uploads/Recovery_Act_Innovation.pdf

30. National Academies Press. (2010). *Overview and Summary of America's Energy Future: Technology and Transformation.*

electricity-generating plants in the United States. Even more importantly, advances in NIT are the key to providing new functionalities and new services – a significant change in the equation.

Although efficient control of energy consumption in the heating and cooling of buildings could realize huge energy savings, consumers and businesses do not at present have the information they need to make worthwhile decisions. How much electricity would they save by shutting off computers at night, or by replacing the kitchen light bulbs with compact fluorescents? What would they save if they lowered the thermostat by 2 degrees? Monthly bills that report total gas and electricity consumption are woefully inadequate for answering such questions, but instrumenting every appliance or fixture in a home or business to provide more precise information would be prohibitively expensive.

One example of how NIT research can help is a University of Washington project that has developed low-cost sensors that consumers can install on the outside of their circuit panels and gas meters – one or two sensors per home or business, rather than one for each appliance and fixture[31]. Using signal processing and machine learning algorithms – fundamental products of computer science research – these sensors can detect and learn the characteristic voltage spikes, water pressure transients, or natural gas sounds made as different lights and appliances turn on and off. Over time, they can accumulate precise information about individual usage patterns and the resulting energy consumption. These sensors, allied with controls that give consumers the easy ability to adjust appliances and fixtures in a centralized way, can put the means to save energy directly into the hands of home and business owners.

The electric grid must become "smarter," enabling it to accommodate rapidly changing sources and loads without taxing transmission and generation equipment. Homes and businesses equipped with richly NIT-enabled energy monitoring and control systems will be better able to integrate with the smart grid, negotiating energy consumption and prices with the utility company according to predicted usage patterns. The smart grid should signal requests to suppliers and consumers to modify their consumption or production of energy to match instantaneous needs, while insuring that the overall grid remains stable. It should analyze consumption patterns to forecast and plan energy production. Its users must be able to analyze data from the grid to optimize their own energy production and consumption.

NIT advances also have a part to play in exploiting new energy sources that reduce dependency on fossil fuel and have lower greenhouse emissions. Efficient use of such resources will require simulation and modeling during the design stage, along with real-time control of operation and management.

Transportation systems must evolve to be more energy-efficient, and NIT is essential for achieving that goal. Hybrid and all-electric vehicles consume no energy while motionless, but drivers on today's congested roads have little control over when they can move and when they must wait. Real-time measurement and control systems can reduce traffic congestion and the energy waste that accompanies it, and, with the help of forecasting based on historical and current data, can divert drivers to optimal routes for their trips.

A more futuristic vision that has recently come much closer to reality is the possibility of vehicles that use sophisticated NIT to drive themselves. Cars in Google's autonomous vehicle project have now logged over 100,000 miles of driving on California streets and highways, with only occasional human

31. http://ubicomplab.cs.washington.edu/wiki/Projects#Sustainability_Sensing

intervention[32]. This remarkable achievement derives directly from research programs in universities. In 2004, 2005, and 2007, DARPA hosted a series of autonomous vehicle competitions in which researchers outfitted cars and trucks with sensors, sophisticated computer hardware and software systems, and effectors (controls) so that they could drive without any human intervention. These systems drew upon decades of computer science research in robotics and artificial intelligence. The progress displayed over these competitions was remarkable. In the first year, no vehicle successfully traveled more than 7 miles, while in the 2007 competition, the top 5 vehicles – all from university teams – each completed an 80 mile circuit in an urban environment, correctly handling the rules for 4-way stops, merging into traffic from turns, and negotiating parking lots. To start its autonomous vehicle program, Google recruited members of the top two teams from the 2005 and 2007 competitions.

Further advances are required before we can rely on robotic chauffeurs. Still, we can already see this research having practical impact in assisting drivers with intelligent cruise control, automatic emergency braking, and the ability to warn drivers drifting out their lanes. Controls in vehicles that allow denser traffic flows, such as convoying, would make travel more efficient by increasing the capacity of existing highways. They also promise to make our roads safer – they never get tired or drive under the influence of alcohol.

Advances in NIT also can promote the use of public and shared transportation through smart control that customizes routes and schedules in response to instantaneous demand. And ride-sharing can become commonplace as a form of "social networking," an extension of a group formed on the Internet that develops sufficient trust to share a car.

NIT-enabled advances in both energy and transportation technology are essential to the long term goal of making our built environment more energy-efficient. Examples of structures commonly found in the world today that allow more efficient energy use are high-density urban areas that mix residential and commercial development to reduce heating, ventilating, and air conditioning (HVAC) loads and commuting, expanded transit networks that dovetail with dense urban design, and localized electricity generation.

Advances in NIT are essential to achieve an efficient energy future. A recent PCAST report[33] called for "support for technological change" to create a new era of energy innovation. Advances in NIT are central to the achievement this vision. NIT is already deeply embedded in all aspects of today's energy and transportation systems. Simulation and modeling are widely used for planning and designing new energy production technologies, such as new nuclear reactor designs, carbon sequestration, wind turbine design and placement, and biofuel production. Likewise, NIT is used in the design of new technologies that reduce energy consumption, such as new insulation materials, lighter vehicles, and buildings that use passive heating and cooling methods. NIT is also heavily used in the operation of facilities that produce and consume energy: the "sense and act" loop of real-time control, or the "sense, model, and act" cycle used in planning and forecasting both short- and long-term production and consumption.

32. http://www.nytimes.com/2010/10/10/science/10google.html

33. President's Council of Advisors on Science and Technology. (October 2010). *Accelerating the Pace of Change in Energy Technologies through an Integrated Federal Energy Policy.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-energy-tech-report.pdf

Using funds made available by the 2009 ARRA, the Department of Energy has implemented an ambitious "Smart Grid Investment Grant Program" that provides funds to utilities to deploy current-generation digital meters in homes and businesses. This metering equipment represents a significant advance over predecessor equipment, in that load information is periodically communicated "upstream" to the utility.

But today's state of the art cannot satisfy all the needs of an efficient energy future. In some cases, existing information technologies must become more economical and be more widely deployed in the energy and transportation sectors. In other cases, new algorithms, new control schemes, new user interfaces, new communication protocols, and new devices will be required. Additionally, barriers to adoption must be overcome.

- *Integrated building networks*: Improved energy efficiency in heating and cooling systems will require instrumented homes and buildings, with numerous sensors and actuators, linked by robust communications – in some cases wireless – that reliably and securely connect to a controller. Sensors will proliferate – to measure room occupancy, external light, temperature, and humidity – and actuators will control HVAC, lights, window settings, fresh air intake, energy storage (e.g., for solar water arrays), and others. Common, widely accepted communication standards are required, to avoid isolated islands of connectivity characteristic of proprietary products. Standards will drive volume up, cut costs, and increase deployment, leading to further efficiencies.
- *Simple user controls that balance efficiency and comfort*: Instrumented homes and buildings, as described above, will require simple user controls to make complex systems easy to use – allowing residents to add and remove equipment, to diagnose problems, and especially to accommodate their need to balance comfort and efficiency. Making such systems more capable but also more user-friendly will require answers to unsolved user interface problems. Intelligent sensor and machine-learning technology should be used to create systems that automatically configure themselves on installation, that automatically detect and diagnose malfunctioning or inefficient components, and that automatically "learn" and adapt to user preferences.
- *Beyond the "single appliance" model of control*: Many efficiencies will come from coordinating multiple devices: doing the laundry and grocery shopping may require scheduling energy for a washer (and its hot water), a dryer, and one of two electric cars with different amounts of charge in their batteries. Solving the scheduling problem may be straightforward, but devising a simple way for a customer to indicate her needs is not. She may need to share resources with others in the family; she may be willing to pay a higher price today because the laundry is needed for guests visiting tomorrow. There is both enormous need and enormous scope for invention for controls that the public can understand and operate effectively.
- *The smart electric grid*: The control system that achieves the truly "smart grid" has yet to be devised. Can the control be decentralized, to permit equal access to any energy producer or consumer? Is the control like a market, striking deals between producers and consumers who agree on terms? What happens under extreme conditions, when supply cannot be brought online nor load reduced fast enough? How can stability be assured?

- *Security of communications*: The smart electric grid must be rigorously protected against cyber-attack. Perhaps the network communications needs of the grid, with requirements somewhat distinct from the open Internet, will merit a new class of network protocols that require strong authentication for all participants. A second line of defense may also be required, in which not just protocols but energy-related behaviors are monitored as well. For example, a consumer might be prevented from bidding for more energy than his connection to the grid can carry; such action requires verification of the credentials and capabilities of all participants. A wind farm that consistently supplies less power than it bids to deliver may find that its standing with the grid falls; it may even be disconnected altogether. Routine automated scrutiny of transaction logs can be used to detect anomalies and intrusions.
- *Increasing multi-passenger vehicle use*: Ride-sharing and public transportation are effective ways to save transportation energy. New IT systems can make these modes of travel much more desirable and effective by making it easier to find rides at the right time, to find the best public transportation options based on real-time position of buses and trains, and to dispatch multi-passenger taxis and vans more efficiently. Synchronizing the elements of the public transportation network – trains, buses, taxis – can reduce waiting times, e.g., by making sure that bus connections at a hub wait for the arrival of a slightly late train.
- *Transportation logistics efficiency*: In the last twenty years, computerized routing and scheduling has made transportation of goods much more efficient. Many opportunities remain, however, to boost efficiency by consolidating transportation further, especially in "the last mile" of delivering packages and groceries to houses and businesses. Batching package deliveries during peak seasons or when a modestly delayed delivery is acceptable offers further saving.
- *Transportation monitoring and tracking*: Sensors, many of which report readings using wireless infrastructure, monitor and report on condition of palettes of goods. In addition to detecting routing errors, these reports can identify theft or diversion of goods. This is one of many applications for low-cost sensor platforms that communicate to worldwide communications infrastructure either directly or through gateways (e.g., on the roadside or in a truck cab).
- *Improved highway vehicle management*: New technologies are emerging to manage traffic, such as measuring and reporting congestion, metering entrants to a highway, and recording payments for highway tolls, driving in controlled areas (e.g., central London), and parking. Technologies that will lead to greater use of these techniques include low-cost sensors in highways, wireless communications systems, and vehicles equipped with sensors and communications. There are also control and algorithmic problems to be solved, such as how to give advice to drivers seeking to avoid congestion without simply clogging alternate routes. Knowledge of each driver's ultimate destination – something many drivers already signal to their GPS navigators – could inform a computer solution that optimizes each car's route.
- *Real-time driver assistance*: Sensors on roads and cars, in addition to short-range communication among nearby cars, can alert drivers to imminent dangers and even initiate corrective action, such as braking. Short of truly autonomous driving, research may make possible an advanced form of "cruise control" that would increase the number of vehicles traveling a highway without compromising safety. Standards for signaling between cars will play a vital role in these devel-

opments. Insuring correctness of the real-time software that participates in vehicle control will be essential for safety.

- *Continuous improvement*: Data from energy and transportation transactions will be easy to collect. The challenge is to analyze the data and make useful recommendations that improve efficiency without sacrificing convenience. Energy usage data can enable conservation strategies, for example by showing that a person can reduce energy consumption by combining two errands in one or by notifying a smart management system that a house will be unoccupied for 8 hours. Less easy is to offer suggestions in advance, or to use other prognostic methods to reduce energy use.

Shorter-term development for energy and transportation. Many of the innovations mentioned above are already underway, and continued deployment will increase their effectiveness. However, realizing the largest visions for the future of energy and transportation requires the capacity to access a common infrastructure, and thus requires standards that define how different components can connect to the infrastructure. Innovation will accelerate when independently devised components can connect to the shared and standardized infrastructure. Here are some examples:

- *Smart meters*: Smart electric meters are beginning to be deployed, but so far they focused on load sensing and control by the electric utility. To be more effective, this sensing and control should be integrated with equipment in the building – either directly via local communications between the meter and the building or indirectly via Internet communication between the utility and the building. Developing such a system in practical form presents many questions. What are the protocols that a device uses to communicate with the utility, via the smart meter or the Internet? Can a suitably equipped appliance work with any smart meter? Can a customer obtain data recorded by her smart meter? What prevents a utility from running a closed system, inaccessible to customers? Are the protocols sufficiently "future proof" to allow innovative computer applications to act as controllers for multiple appliances, in conjunction with the electric supplier (and its meter)? In other words, do the interfaces to the smart meter support open innovations?
- *Car-to-car signaling*: Automobile manufacturers will be reluctant to start building certain kinds of advanced driver assistance into vehicles until there are agreed standards for communicating with adjacent cars and highway features (signs, signals, etc.).

And of course the smart grid will need a suite of definitions for its infrastructure as well. These standards must embrace interoperability (see the sidebar "Interoperable Interfaces and Testbeds Drive Innovation and Economic Growth" on page 15); they must be designed to allow revisions with backward compatibility so that the infrastructure and devices connected to it can evolve continuously without interrupting service. If integrated systems are to evolve, which they must if optimizations are to govern more than single components, the appropriate communications infrastructures must be in place. It won't do to have a cacophony of incompatible offerings, whether open or proprietary. The sorry state of "home automation" today shows what can happen: proprietary and royalty-bearing protocols are stifling deployments that would otherwise be progressing rapidly because of the falling costs of sensors and computers. And ubiquitous, inexpensive home automation protocols are an essential precondition for energy-saving controls.

4.3 NIT for National and Homeland Security

National and homeland security is the broad priority of protecting the Nation. It involves four key constituencies: the armed forces, the intelligence community, law enforcement, and homeland security. While each constituency is working toward a common goal, their missions are different. They use NIT to support their missions in different ways but in many cases the technology and data they rely upon is common and interconnected. Each constituency should seek the active advancement of science and technology that supports its mission, but also must maintain a common foundation.

Agency Constituencies

- *DoD*: DoD aggressively applies information technology in its operations. For example, it is reputed that, because of NIT, the U.S. military has more remotely piloted air vehicles than manned aircraft. DoD has mature offices that assess future needs and actively seek to develop technology that may serve those needs – organizations such as the office of the Director of Defense Research and Engineering (DDR&E), DARPA, and the Army, Navy and Air Force Offices of Research
- *Intelligence*: The intelligence agencies are fairly effective at employing networking and information technology, and recently have focused on technology that helps analysts to "connect the dots" – that is, to detect patterns of interest in their massive data collections. They have internal science and technology (S&T) offices that seek technology advancement. Work in several areas, such as high performance computing, cryptography, and mathematics, is reputed to be excellent, but much of that work is classified. Except for some selected National Security Agency (NSA) relationships, the intelligence community does not have close ties to the research universities. (The new Intelligence Advanced Research Projects Activity (IARPA) may improve this situation.) There are few avenues by which the intelligence community can get access to rapidly advancing information technology developed within small start-up companies.
- *Law enforcement*: Law enforcement organizations such as the Federal Bureau of Investigation (FBI) rely on NIT to sift through data and to track investigations involving reams of source data ranging from video and photos, biometric samples, fingerprints, interview results, and court filings. The many local police departments that span the Nation are neither funded nor equipped to advance technology, and because there are so many different, independent law-enforcement organizations adopting technology, they often deploy non-standard systems that actually impede communication and information sharing. At the Federal level, the adoption record of law enforcement technology by the FBI is mixed. For example, the FBI may use advanced systems to track money laundering, while at the same time struggling to establish a system to track cases. But it is the FBI and the National Institute of Justice that should be leading the advancement of technology for law enforcement.
- *Homeland security*: Homeland security organizations within the Department of Homeland Security (DHS), such as the Transportation Security Agency (TSA), Customs and Border Protection (CBP), Immigration and Customs Enforcement (ICE) and the U.S. Coast Guard, depend upon NIT to prevent terrorists from boarding aircraft, identify unlawful imports, protect our borders, and

issue employment authorization documents. Since its creation, however, DHS has followed a procurement strategy in technology acquisition, aiming to purchase rather than create the systems it needs. DHS's information and networking R&D program is inadequate, leaving it with no early stage technology investments or effective counterpart to DARPA.

Terrorists and Crooks: Internet-Enabled

Today's terrorists and crooks exploit pervasive 21st century information technologies to work more effectively and dangerously. They are finding new ways to use the Internet to organize, target, destroy and steal.

Worms used to be blunt instruments; now they are capable of surgically targeting specific systems, potentially to reprogram them to cause major damage to the physical systems that they control. Routinely, electronic voting software is penetrated and compromised. Terrorists and crooks use Internet communication to organize themselves.

The incremental adaptations that are often used today to protect cyber systems frequently fail. There are compelling reasons to believe that acceptable cybersecurity will only be achieved using entirely new approaches arising from fundamental computer science research. Research in hardware and software systems and networks, as well as into issues related to trust and privacy, may lead to the development of new architectural designs for systems that can actually be protected.

Research in data mining, data visualization and algorithms can produce new approaches for monitoring and combating threat activity, as well as for analyzing massive streams and reservoirs of data to detect and disrupt adversarial social networks.

Capabilities. The following list outlines a set of capabilities that are required in the future by at least one of the national and homeland security constituencies, and typically by all four. Each of these capabilities relies upon technological superiority in NIT. All depend upon adaptation of processes, tactics, and procedures to create and take advantage of emerging technology, which requires that some part of each organization be deeply involved in driving the development of needed technology. Merely purchasing off-the-shelf NIT products will put the organization many years behind where it could be, and where its adversaries may be.

- *Information superiority*: This capability requires mastery of several related elements: collection, analysis, and distribution of critical information, often on a *real time* basis. It includes near-real-time awareness of the location and activity of friendly, adversary, and neutral forces throughout an area of engagement. It also includes the collection and processing of intelligence regarding the intentions and motivations of adversaries ranging from criminal organizations to spies to military organizations. Each element also requires a seamless, robust command and control network linking all friendly forces that can deliver data (even from remote sensors in space) to where it is needed. Meeting national and homeland security challenges involves the entire spectrum of extreme scale computing – for example, in the analysis of very large intelligence data collections, simulations of physical systems, analysis of advanced weapons designs, and code breaking.

- *Identification and attribution*. All constituency organizations must be able to differentiate between friend, foe, and neutral parties in both physical space and cyberspace. Depending on the situation, the agencies field individuals, teams, and large forces, placing them in dangerous situations and locations. To protect a force – small or large – requires the ability to differentiate between friend, foe, or neutral in sufficient time, with high confidence, and at the requisite range to support decisions on engagement and weapons use. It requires the ability to see all aspects of the environment that affect tactical decision choices. Similar requirements exist in cyberspace. We must be able to identify, authenticate, and trust parties as well as differentiate between neutral parties and adversaries. Identifying the sources of cyber-attacks is of critical importance.
- *Precision engagement*: This is the capability to disable or destroy selected targets – in physical space as well as cyberspace – with precision, while limiting collateral damage. In physical space it relies upon precision-guided munitions, surveillance, precise targeting, and the "sensor-to-shooter" information flow necessary for responsive, timely force application. In cyberspace, the United States needs to have the ability to achieve information superiority by debilitating an adversary's information, information-based processes, information systems, and computer-based networks, while defending its own. Adversary attacks need to be detected and mitigated. This requires active monitoring of our own infrastructure as well as the infrastructure of adversaries. It requires the development of specific software/hardware packages to project an attack when directed, and the ability to assess damage.
- *Infrastructure protection*: Cybersecurity is a critical weakness for government and private systems alike. Of particular concern is protection of all critical infrastructure, including communications networks (civil and governmental), electricity, financial systems, logistics, fuels, water, and emergency services. Critical systems must be robust and resilient.
- *Readiness*: All four constituencies can use information technology to support exercises and training to ensure readiness. Multiple organizations – e.g., multiple military services as well as multiple civil services, such as the FBI and the Drug Enforcement Agency (DEA), and even international police organizations – need to be able to combine forces and act effectively. Readiness depends upon training, in particular on training that relies on enhanced simulation; advanced storytelling and cultural sensitivity awareness; and advanced learning techniques. Use of advanced IT – such as gaming – can help personnel plan, prepare, and assess a wide variety of hostile situations.
- *WMD detection and attribution*: Countering proliferation requires the ability to detect and evaluate the existence of manufacturing activity for weapons of mass destruction (WMD), to locate, identify, and assess the threat posed by fully constituted devices, and to track launched WMDs so that the proper level of counterforce can be exerted in a timely way. Homeland security requires the ability to detect biological agents, chemical agents, and nuclear and radiological material in modest quantities across large urban areas.

The above capabilities are needed for the national security and homeland security organizations to protect the Nation in the coming decades. Every single capability depends upon NIT, and upon advancing the science and technology to improve what can be accomplished.

Research priorities. In what follows we describe selected opportunities in which advancement of NIT can dramatically improve the prowess of our national and homeland security organizations.

- *Trustworthy software and cybersecurity*: It is critical to find better solutions than today's, which are not sufficiently effective. Fundamentally new approaches will only come from fundamental research. The new solutions will, hopefully, increase tolerance of both system failures and deliberate attacks.
- *Comprehensive situation awareness*: In the way that snapshots remind a person about an event, a person, or a situation in the past, what the security constituencies require to detect possibly significant developments are current and historic representations of an area, people, or a virtual space of interest. Histories of changes through time need to be accessible. Time-stamped sensor data of different kinds needs to be maintained and merged to provide the user with just the perspectives and information needed for the current objectives. This capability is required even for activities that do not necessarily involve contested situations, such as routine police protection of an area or Coast Guard monitoring to ensure that fishing fleets obey limits on catches.
- *Machine learning*: Machine learning techniques could be applied to identify patterns that are of interest and to cue a sensor operator. In baggage screening, for instance, there is only a finite number of distinctly different devices that travelers carry, e.g., cameras and laptops. Their signatures could be learned. The screening device software could automatically label parts of the image and then flag what it does not recognize or what looks suspicious, cueing the operator and making better use of the operator's attention and expertise. There is ample training data. This same technology could be applied to better focus the humans who monitor the real-time images generated by remotely piloted vehicles and security perimeter camera suites. Automation can track the mundane, and cue the human to attend to what is interesting, suspicious, and relevant.
- *Data discovery*: The generation capacity of today's sensor and computation devices yields an amount of information that now exceeds human processing ability. The capability to find patterns automatically among vast volumes of data would materially aid all four constituencies. The ability for the software system to be taught by the operator in order to refine old patterns and add new ones would reduce the time analysts spend on the mundane tasks that automation can perform.
- *Usable interfaces to automated systems*: Continued progress is required to ensure that the system evolves to perform better for the human, rather than the human needing to adapt to mitigate the shortfalls of the system.
- *Robotics and cyber-physical systems*: It is early in the era of using robots to perform physical activities such a carrying materiel for a soldier or a police officer or fighting a fire at the hottest

locations. Miniature cyber-systems, or suites of them, can be embedded in physical locations, including in human beings. The leverage that they offer has just begun to be realized.

In many cases the actual application of NIT for the national and homeland security constituencies is classified. That should not impede research. For most topics, there are good sources of unclassified applications and data that can serve as proxies for classified applications. Photo-sharing web sites, commercial satellite images, commercial logistics data collections, and blogs are just a few examples. The intelligence communities are making increased use of open sources, e.g., online newspapers, blogs, and video feeds. Their ability to do so is limited by their ability to process the media automatically. Research using these unclassified sources can have a direct bearing on security, and offers several other advantages. First, it exposes students moving through the research universities to surrogates for the classified problems, training a part of the future workforce of the constituencies. Second, it makes available a community of smart, experienced people, beyond the restricted number of government laboratory workers and contractors, to work out alternative solutions, providing useful competition for the best ideas. Third, it provides unclassified new ideas that can be the basis for new industries to serve the constituencies.

4.4 NIT for Discovery in Science & Engineering

Discovery in science and engineering is the driver for innovation, economic growth, and the advancement of both science and society.

Over the past several decades, computational science – the large-scale simulation of real-world phenomena – has joined theory and experiment as a fundamental tool for discovery in many branches of science and engineering.

Today we are at the dawn of a new revolution in discovery – a revolution that will have even more dramatic and pervasive impact. This revolution is once again being driven by rapid advances in NIT – both hardware and software. The focus of this new revolution is *data* – specifically, the ability to collect and manage data in quantities orders of magnitude greater than ever before, the ability to make this data directly and immediately accessible to a global community, and the ability to use algorithmic approaches to extract meaning from huge volumes of data.

Enormous numbers of tiny but powerful sensors are being deployed to gather data – deployed on the sea floor, in the forest canopy, in gene sequencers, in buildings and bridges, in living organisms (including ourselves!), in telescopes, in point-of-sale terminals, in social networks, in the World Wide Web. These sensors (and, indeed, simulations too) produce huge volumes of data that must be captured, transported, stored, organized, accessed, mined, visualized, and interpreted in order to extract knowledge.

This "computational knowledge extraction" lies at the heart of 21st century discovery. Its fundamental tools include sensors and sensor networks, broadband networks, databases, data mining, machine learning, data visualization, and cluster computing on an enormous scale. Advancing these tools, so that science may advance, is a major challenge for NIT R&D. The world of science is transitioning from data-poor to data-rich, vastly expanding the potential for new breakthroughs, particularly when combined with other NIT-enabled advances, such as simulation and new modes of collaboration.

Data is already accelerating new discovery. For example, a recent and unprecedented cross-sector community collaboration aggregated digital information from PET scans and other sources to illuminate the progression of Alzheimer's disease in the human brain[34]. The same data-driven approach is now being applied to accelerate discovery of the progression of Parkinson's disease. In environmental science, the collection and retention of sensor data, along with the improvement of climate algorithms, have enabled us to predict the size of the ozone hole with increasing accuracy.

The abundance of digital information has already led to the democratization of scientific research and education, leveling the playing field for new research discoveries, and has expanded our ability to collect, combine, and analyze information in real time. Over the next decade, ubiquitous access to raw data, and to the digital information and knowledge that results from its analysis, will drive the creation of smart environments (including power and transportation systems), personalized medicine and greater understanding of disease and treatment, the development of new materials and designs for energy-efficient built environments, and a wealth of new innovations, application areas, and enterprise.

An important path to discovery continues to be the ability to use computation for simulation and modeling. Access to data expands our ability to simulate complex environments. The integration of large-scale data and large-scale simulation provides a powerful new tool, fueling our ability to scale up (to macro scales, greater range, and larger studies) as well as to scale down (to nano levels, smaller granularity, and finer detail). In astrophysics, large-scale simulations, vetted with observations from the world's most powerful telescopes, will advance new understanding of how the universe formed and evolved. On the other end of the scale, experimental data and computational models will allow digital "observations" of molecular dynamics and increased understanding of the microscopic world, precluded from direct observation by scale and by the Heisenberg Uncertainty Principle.

New discoveries in science and engineering depend on continued advances in NIT. These advances are needed to ensure the privacy and effective use of electronic medical records, model the flow of oil, and drive the powerful data analysis needed in virtually every area of research discovery. It is important to understand that advances in NIT include far more than advances in processor design: in most fields of science and engineering, performance increases due to algorithm improvements over the past several decades have dramatically outstripped performance increases due to processor improvements.

Advances in NIT to support discovery in science and engineering. Advances in NIT in a data-rich world will require new approaches. New algorithms will be needed to incorporate multiple sources of data and create new efficiencies; next-generation programming environments and software will need to support interoperability of an unprecedented variety of sensors, devices, data streams, and computers; new models will be required that incorporate contextual and real-time data to more accurately represent complex environments. Key areas for NIT research investment will include:

- Strategies for using data and for transforming data into information and knowledge: data mining, machine learning, data visualization, data analysis.

34. Sharing of Data Leads to Progress on Alzheimer's, *New York Times*, Aug. 13, 2010; http://www.nytimes.com/2010/08/13/health/research/13alzheimer.html?_r=1&ref=the_vanishing_mind

- Systems that support next-generation applications: sensors and sensor networks, robotics, hardware and software particularly suited for data analysis, next-generation scalable and high-performance computer architectures, and wireless and wired networks.
- Algorithms and systems that support policy and social interaction: data privacy and security, crowdsourcing, social networking.
- Reliable systems for accessing and retaining digital information: digital management and preservation systems, long-term storage.

In addition, research and discovery in a data-rich world presupposes adequate enabling NIT infrastructure. Stable, reliable, and economically sustainable NIT infrastructure environments are needed to support data access, use, management and retention. Scalable systems are needed to support a broad spectrum of computationally-enabled applications, and high performance architectures are needed to support applications at extreme scales. Enabling software and systems are needed to create an environment in which the barrier to access is low for innovation and new discovery. To drive American innovation, leadership and competitiveness, critical NIT infrastructure will need to be supported as infrastructure (with a focus on stability, usability, evolvability, and reliability), rather than as research (with a focus on new starts and yet-to-be-investigated concepts). Creation of viable economic models to support, evolve, and sustain NIT infrastructure can provide an opportunity for unprecedented partnership between the public, private, and academic sectors.

A broad range of continuous investment in NIT-enabled discovery is fundamental for U.S. leadership and competitiveness. However, a number of specific investments made now can dramatically accelerate U.S. discovery and innovation. These include:

- Investment in interdisciplinary initiatives that advance both data-oriented computer science and "domain applications" in priority areas (e.g., healthcare and the life sciences, the environment and climate change, energy and smart grids), as well as investment in data-oriented computer science that supports privacy, data security, and other characteristics.
- Investment in the "science" of social networking, crowdsourcing, and other emerging paradigms that exploit extreme-scale usage, scalable systems, and clouds and data centers.
- Investment in NIT data infrastructure that supports broad-based data use, management and retention.

4.5 NIT for Education

One essential, strategic investment that the United States must make to maintain leadership and competitiveness is in the education of the next generation. Today's young people live in an interconnected world in which NIT has democratized access to resources, expertise, and information. A basic understanding of how to use NIT and of the foundational ideas behind NIT should be a fundamental part of every child's education. The ever-expanding role of NIT in our society creates an ever-increasing demand not only for NIT professionals, but also for people who can utilize NIT flexibly and creatively and who can apply NIT "modes of thought" in a wide variety of endeavors. Education and learning expert Marc Prensky writes:[35]

> Power will soon belong to those who can master a variety of expressive human-machine interactions.… I believe the single skill that will, above all others, distinguish a literate person is … the ability to make digital technology do whatever, within the possible, one wants it to do – to bend digital technology to one's needs, purposes, and will, just as in the present we bend words and images.

Clearly schools must give all students access to the world of NIT. Providing physical access to computers and the Internet is but a starting point. The challenge for education is twofold: first, to focus on fluency in NIT, on "computational thinking," and on the fundamental concepts of computer science in order to prepare today's students to be the next generation of leaders and professionals throughout our society; and second, to use NIT technologies to enhance teaching and learning.

It is important to meet these challenges from early childhood education through adulthood. As with other areas of Science, Technology, Engineering, and Mathematics (STEM) education, our educational system in computer science works best at the college and postgraduate level, and needs the greatest improvement at the K-12 level. The first challenge concerns the content of computer science education. It is discussed in Section 8.2 of this report. The second challenge, the use of NIT in the delivery of education for all fields, is the subject of the remainder of this section. In this area, some progress is already evident, but investment in new NIT research can dramatically accelerate the trend, expanding the student experience and enhancing teachers' abilities to guide the process.

NIT to enhance teaching and learning. Many of today's students come to school familiar with the World Wide Web and with simulated environments and games. Our educational system should leverage this familiarity and expand students' knowledge and experience from this baseline. Two transformative contributions that NIT can bring to the provision of education are especially noteworthy: personalized electronic tutoring, and ubiquitous and seamless education ("education everywhere").

The personalized electronic tutor offers education that fits the needs of each individual student. Using virtual models of the learning process, personalized tutors will track individual students as they acquire concepts and skills, and tailor presentations and exercises to build and test the skills each student needs. In this way, a new generation of virtual tutors, like the best human teachers, will adapt to the learner's life circumstances and learning style. Rather than offering a fixed sequence of materials based on age and time-in-grade, personalized tutors will offer a pace and educational experience to meet each learner's needs, and will provide teachers with detailed evaluations of the student's progress.

Virtual tutors have shown particular success in helping students to learn mathematics. For example, when teachers in Oklahoma compared students using Carnegie Learning's Cognitive Tutor (Carnegie Learning is a Carnegie Mellon University startup company) against those using a traditional mathematics textbook, the tutored students "scored higher on a standardized test..., received higher grades, reported greater confidence in their mathematical abilities, and were more likely to believe that mathematics

---

35. Prensky, Marc. (February 2008). "Programming is the New Literacy." *Edutopia* (The George Lucas Educational Foundation), http://www.edutopia.org/literacy-computer-programming

would be useful to them outside of school." A similar study found improvements that "were particularly dramatic for special populations" such as students with limited English proficiency.[36] Analogous approaches are showing early success in creating adaptive games for learning mathematics.[37]

Expanding this early success to other levels of mathematics and to other fields, especially those in the humanities and social sciences requiring subjective knowledge and sophisticated use of language, will require further breakthroughs in natural language processing, concept modeling, and data analysis.

"Education everywhere" makes use of the Internet and other mobile technologies, such as future smart phones, to offer rich interactive experiences in any location, embedding learning in the learner's life and allowing specialized subjects to be taught by observation. Examples might be ecology taught in a wildlife preserve or art taught in a museum. Education everywhere also cuts through barriers of age, geography, culture, language, or socioeconomic status. Learners can form communities, enabled by social technologies, to bring the collective learning experience of classrooms to students who cannot come together in traditional classrooms.

In these and other ways, NIT can transform the organization and delivery of education, just as it has transformed business and government. The Department of Education's (ED) National Educational Technology Plan[38] lays out a vision for the future of technology-enabled learning, and the NSF report on Connecting Learning and Education for a Knowledge Society[39] describes how advances in machine learning and human education reinforce each other. Government must seize this opportunity by laying the R&D foundation for this transformation.

In its recent report on K-12 STEM education[40], PCAST recommended the creation of an Advanced Research Projects Agency for educational technology R&D, called ARPA-ED. It would have the primary goal of "propelling and supporting (i) the development of innovative technologies for learning, teaching and assessment across all subjects and ages, and (ii) the development of effective 'deeply digital' whole-course instructional materials for STEM education that prepare and inspire the next generation of American students." The creation of such an agency could play an important role in realizing the promise of technology for enhancing the effectiveness of education in the United States.

A research agenda for NIT in education. Transforming education requires fundamental advances in NIT. An NIT R&D agenda should include the following components:

- *Data analysis*: Develop new data mining and machine learning methods that analyze data on large numbers of students to assess the efficacy of educational methods and evaluate policy and resource choices, at any scale from an individual school to the country as a whole. Build user models that take information about individual learners and automatically infer students'

36. Ritter, Steven, John R. Anderson, Kenneth R. Koedinger, and Albert Corbett. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review* 14(2), 249-255

37. http://games.cs.washington.edu/Refraction/

38. National Educational Technology Plan 2010. Office of Educational Technology, U.S. Department of Education. *Transforming American Education: Learning Powered by Technology.*

39. National Science Foundation. (January 30 2010). Internal Task Force on Innovation in Learning and Education, *Connecting Learning and Education for a Knowledge Society.*

40. President's Council of Advisors on Science and Technology. (September 2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

"traits and states": what they know, how they learn best, and where they need help. These data analysis methods must be designed to protect student privacy.

- *Social learning*: Develop social networking and social media tools, optimized for use in education, that build communities of learners across boundaries and extend the social learning experience beyond the classroom into students' lives.
- *Games for learning, and immersive environments*: Develop and evaluate "serious games" that combine the engaging experience of electronic games with a serious educational purpose. Create immersive environments that can emulate situations in which students apply what they have learned. Devise tools to make the development of games and environments easier, cheaper, and more practical for teachers.
- *Mobile technologies and learning in the world*: Create tools and methods for integrating education with the world, both in learning-friendly settings such as museums and forests, and in students' lives. These tools, by sensing the students' locations and surroundings, will connect learning opportunities to the students' educational goals and to communities of learners.
- *Human-computer interfaces*: Human-friendly interfaces, important in any use of electronics, are particularly important in education, where users may be young and may not yet be fluent with technology. As part of a broader human-computer interface (HCI) research effort, government should support HCI research that focuses on modeling, coaching, and tutoring student users.

4.6 NIT for Digital Democracy

Information technology is transforming government operations and opening new communication channels between government and citizens. A broad vision going by the name of digital democracy envisions the use of information technologies to improve public discourse, increase dialogue between citizens and government, make government more open and transparent, improve the operation of government, and bring the benefits of technology to everyone. Existing technologies have much to offer, and governments – from national to local levels – have been alert to many of these opportunities. However, unlocking the full benefits of digital democracy will require NIT research to address fundamental challenges.

What digital democracy can deliver for society and government. Digital democracy can transform our society over the next decade or two. In a world where digital democracy is fully realized:

- Government will be more responsive and accountable because citizens can see and measure how it responds to requests and problems and can more fully participate in planning processes. Personal technologies and social computing will enable real-time citizen reporting on civic needs and government service quality. Government will have unprecedented awareness and access to public opinion and public expertise, thanks to new methods of soliciting and gathering input from citizens. These tools will be available to all people, whenever and wherever needed, via Internet and mobile devices.

- The Internet will become a forum for substantive public collaboration and deliberation, thanks to new social technologies that foster positive interactions and are not polluted by spam and flame wars. All citizens will be able to make effective use of government data, thanks to breakthrough tools supporting access, analysis, and visualization for non-experts. Archivists, historians, journalists, and the public will have better and more convenient access to government records, including information previously available only in paper form.
- Digital tools will foster institutional innovation within government, making government services and processes faster, better, and cheaper. Government employees will have vastly better access to information and expertise, inside and outside government. Information and knowledge will be able to flow where they are needed, even as agencies maintain their distinct missions and cultures. Government regulation will be more effective and overregulation will be reduced because regulators have better information about where their actions are effective, and about the costs of overregulation. The availability of excellent, flexible tools and well-documented best practices will bring these benefits of digital democracy to all levels of government, including state and local.

A Picture is Worth a Thousand Numbers

Data visualizations can illuminate trends in large and complex data sets. Visualizations have typically required significant time and expertise to create. But new tools allow anyone to create useful visualizations, broadening participation in public debate by letting ordinary citizens explore government datasets and create persuasive arguments from them.

Two examples of recent highly usable data visualization tools are Tableau[41] (the product of a Stanford University startup located in Seattle WA) and Many Eyes[42] (a publicly available experiment by IBM Research). Many Eyes, as of this writing, features citizen-created visualizations of carbon emissions by G-20 vs. non-G-20 countries, sources of Pakistan flood aid, and teacher starting salaries by state, among many others. These visualizations are built using data-exploration tools on the site, from data sets provided by sources such as the Census Bureau, the Bureau of Labor Statistics, and the Federal Reserve. Visualizations can be embedded in external web sites, so they can enter the public conversation.

As open government initiatives put more data online, the value of further advances in easy-to-use visualization tools will only increase. What was once available only to experts will now be open to any citizen.

- Publication of data will become a valuable public policy tool. By opening data sets to the public, government will foster transparency and competition, improving outcomes in the market and in public life.
- Election processes will be secure, convenient, inclusive, and strongly error-resistant, thanks to judicious use of NIT in voting and tabulation.
- Access to government and access to justice will be improved for *all* citizens.

41. http://www.tableausoftware.com/
42. http://www-958.ibm.com/software/data/cognos/manyeyes/

NIT research to support digital democracy. Experience teaches that digital democracy requires much more than simply deploying technology. Success depends on judicious use of technology, on management and consensus building, and on building a culture of continual and iterative institutional innovation inside and outside government. At present, much effort is spent wrestling with the technical details of data formats, system architectures, and technology management.

What may be less obvious is that the practical difficulties of deploying transformative technology are often manifestations of deep technical challenges that NIT research can address. An NIT R&D program aimed at these challenges can open new opportunities for digital democracy that are not possible today. Some examples of these research opportunities are:

- *Privacy and security*: Methods for putting datasets online without compromising citizen privacy. Methods for verifying, without compromising privacy, that an online commenter is a real person, or is not the same person using a different identity. Effective methods for users to verify the authenticity of information from government datasets.
- *Social computing*: Understanding how to make online interaction engaging for citizens. Mechanism design for public deliberation: understanding how to structure and (semi-)automatically manage discussion forums to keep the discussion on-point and productive. Helping citizens find and help each other. Measuring public opinion and sentiment in ways that resist gaming and strategic behavior. Specializing social media and crowdsourcing technologies for use in government.
- *E-voting*: Methods for voting from remote locations, e.g., for overseas military personnel, that fully address inherent security risks. Making voting more convenient without compromising security. Protecting the secrecy of the ballot in an environment of ubiquitous, high-resolution cameras and sensors.
- *Computer-aided text analysis and processing*: Methods for government employees to sift through large sets of comments and discussion to find relevant information. Methods for automatically summarizing discussions and extracting the most important arguments and facts.
- *Data analysis tools*: Tools to aid export of legacy datasets, making them available to the public with whatever documentation exists, as well as porting of data, which connotes greater effort to make the data more easily understandable, standardize formats, create thorough explanatory documentation, and so on. Tools for analyzing the format and semantics of unstructured or poorly documented data sets. Tools for finding likely errors in diverse datasets. Processes for managing errors, including error handling and metrics for responsiveness.
- *Data tools for non-experts*: Tools to let non-experts effectively access, analyze, and visualize government data, and to publish the results of their analyses. (See the sidebar "A Picture is Worth a Thousand Numbers on the previous page").

Shorter-term development for digital democracy. There are many things the Federal Government can do, and is doing, to advance digital democracy. The Office of Science and Technology Policy (OSTP) Open Government Initiative, which includes the opening of government datasets at Data.gov and the publication of government spending data, is a valuable first step. Progress toward opening data and documents at OSTP, the National Archives and Records Administration (NARA), and elsewhere, deserve continued effort and attention.

The Community Health Data Initiative (CHDI)[43] is just one example where OSTP and the Department of Health and Human Services (HHS) are taking a leadership role in targeting NIT innovation toward today's information needs in community and public health. The effort engages private and public players and promotes engagement by a broad range of organizations, from county health departments to patient advocacy groups to social media startups. This strong beginning could be further enhanced by leveraging already available technical capabilities in new ways in a form of technical and social innovation. For example, moving beyond the integration of existing institutional data sets, the CHDI could create a channel for citizen participation and dissemination of citizen-generated data using mobile, social, and traditional web media. Similar efforts throughout government deserve continued support.

Success in digital democracy requires participation from both government and the public (including companies and non-profits). Some things can only be done by government, e.g., publication of government-held data as in the Data.gov initiative and the USAspending.gov dataset. Other things are best done by non-government actors, e.g., organizing political discussions related to government data and activities. In some areas, such as defining data-driven metrics for quality of healthcare, both government and private parties may have things to offer. It is important to enable innovation both inside and outside of government.

Much can be learned from the impact of NIT on business. Initially, organizations transfer their legacy paper-based processes into the electronic realm. This offers relatively modest benefits. Over time, processes are re-engineered to truly take advantage of what the digital world can offer. This is a slow process – often taking a decade or more – but it unlocks the big benefits of going digital. This will surely be the pattern in government, with the greatest benefits of digital democracy only becoming evident over time.

The difficulty of changing government processes presents additional challenges. But getting this transformation right offers huge benefits: by making government more efficient, more responsive, and more transparent, we can make it more effective at addressing the challenges of the 21st century.

43. http://www.hhs.gov/open/plan/opengovernmentplan/initiatives/initiative.html

5. Recommendations: Initiatives in NIT R&D to Achieve America's Priorities

The Federal Government must invest in new multi-agency NIT R&D initiatives in areas of particular importance to our national priorities. This section contains specific recommendations regarding NIT R&D for Health, for Energy, for Transportation, for National and Homeland Security, for Education, and for Digital Democracy. Section 7 contains recommendations regarding NIT R&D in various research frontiers of the NIT field – frontiers that are described in Section 6. Although no specific recommendations are made in the current section regarding NIT R&D for Discovery in Science & Engineering, advances in many of those research frontiers are crucial for progress in this national priority, as are the infrastructure improvements addressed in Sections 8 and 9. More generally, advances in the various NIT research frontier topics are essential for progress in *all* the national priorities.

Addressing the recommendations in this section will require the formation of multi-disciplinary collaborations, sometimes between basic researchers and domain experts. Two principles must be kept in mind. First, it is important that short term needs not "crowd out" the longer term research that anticipates future needs. Some of the research must explore bold unconventional ideas that would have enormous impact if they could be realized. Second, some of the collaborative research will require large project teams, drawn from communities that have not worked together before. These large projects must have sufficiently long time horizons to allow ambitious goals to be achieved.

NIT for Health

Improving people's health while containing healthcare costs is an important national priority. Advances in NIT are essential to the realization of improved health and quality of life for people at every stage of life at an affordable cost.

The Federal Government has recognized the importance of NIT for health and has embarked on an aggressive program to institutionalize electronic health records. The Strategic Health IT Advanced Research Projects (SHARP) program, currently funded by ONC, is excellent, but it addresses relatively short-term problems. Other agencies have initiated promising longer-term programs (e.g., NSF's Smart Health and Wellbeing program, the National Library of Medicine's (NLM) health data standards and telemedicine projects, NIST's Healthcare Infrastructure Integration projects) but larger investments in these and comparable efforts are needed.

Recommendation: **The Federal Government, under the leadership of NSF and HHS, with participation from ONC, the Centers for Medicare and Medicaid Services (CMS), the Agency for Healthcare Research and Quality (AHRQ), NIST, the Veterans Health Administration (VHA), DoD, and other interested agencies, should invest in a national, long-term, multi-agency research initiative on NIT for health that goes well beyond the current national program to adopt electronic health records.** The initiative should include sponsorship of multi-disciplinary research on three themes:

- to make possible comprehensive lifelong multi-source health records for individuals;
- to enable both professionals and the public to obtain and act on health knowledge from diverse and varied sources as part of an interoperable health IT ecosystem; and
- to provide appropriate information, tools, and assistive technologies that empower individuals to take charge of their own health and healthcare and to reduce its cost.

This program should build on national activities promoting the adoption and meaningful use of electronic health records that are usable by all appropriate organizations; it should complement the shorter-term ONC programs; and it should augment the research investments that the various agencies are currently able to make. In addition to increased attention on using NIT for wellness and for addressing chronic conditions, the departments and agencies mentioned above should continue to investigate novel uses of NIT, such as NIT-assisted surgery, to deliver care for acute conditions. They should continue to pursue advances in the innovative use of NIT, such as sensing and monitoring, to understand the basic biological and psychological mechanisms that underlie disease. And they should continue to address NIT research opportunities that support current and continuing work by HHS and NSF on transformational innovation in healthcare delivery and basic research in health and wellness. The funding levels and project durations must be sufficient to foster substantive collaborations between NIT researchers and clinical experts.

NIT for Energy and Transportation

Realizing the potential of the smart energy grid requires long-term research, to develop not only new electricity-processing elements of the grid (sources, controls, and especially storage) but also the information technologies to control them. Other important research directions with the broad goal of reaping energy savings include exploration of innovative ways to use the current grid more efficiently and continuing efforts to exploit information technologies that make operations of all kinds more efficient.

Recommendation: DoE should identify key information-related technologies for the future grid that may not be developed by utilities, and should build a long-term research program that focuses on them. Examples are robust control algorithms for a highly dynamic grid, extreme resistance to cyber-attack or disruption, advanced grid monitoring and diagnostic techniques, and data-mining techniques designed to continuously improve grid efficiencies. This research will need to acknowledge key constraints of the country's electrical landscape, such as the private ownership of most generation and distribution assets, together with a need for public-private partnerships for operation and evolution.

Recommendation: In developing the Quadrennial Energy Review recommended recently by PCAST[44], DoE should integrate research, development, demonstration, and deployment (RDD&D) of NIT-enabled energy technologies as one of its elements.

Recommendation: DoE and NSF should lead an R&D effort across multiple NITRD agencies, including NIST, DoD, DoT, and the Federal Aviation Administration (FAA), to use NIT to measure and achieve optimum performance of operating equipment with a view to obtaining energy efficiencies. The research should encompass new sensors, controls, and algorithms for optimal control. Especially important is better control of building heating and cooling, where better NIT can turn large opportunities for efficiency into real energy savings. A particular opportunity involves commercial buildings, where automated control systems can be much more sophisticated. End consumers (individual citizens and commercial enterprises) should be empowered to make choices in their energy consumption. To do so they need near-real-time knowledge of their consumption and information about pricing of energy over time. Economic research should be done to create models for differential pricing that enable the end-consumer to both save energy and reduce costs.

Recommendation: NIST should use its convening power, in conjunction with DoE, to create and impel a structure for promoting interoperable standards for real-time control of the grid and the energy consuming systems that feed from it. This effort should build on the experience of the Internet Engineering Task Force (IETF), whose emphasis on "rough consensus and working code" – pragmatism, interoperability, the interworking of multiple transmission technologies, and foresight achieved by significant engagement of the research community – has fostered the Internet's ability to "innovate at the edges." Particularly important are the requirements for communication privacy, security, and robustness. It will also be essential to develop a framework that smoothly evolves from today's grid, with its limited opportunities for dynamic control, to a much more efficient and dynamic future grid. These standards should ensure that open innovation flourishes to improve efficiencies in electricity consumption.

Recommendation: DoE should increase its use of high-performance computing capabilities in support of research in energy-saving technologies such as new materials, new electro-chemical processes (e.g., for batteries), new bio-fuel processes, and new applications of combinations of technologies (e.g., electric hybrid systems).

Creative use of new NIT tools can make U.S. transportation systems more efficient, more flexible, and safer. NIT-enabled improvements in transportation could also lead to significant declines in transportation energy use and emissions. The greatest benefits would result from embedding sensors in vehicles and vehicle components that would allow NIT to optimize the performance of individual vehicles and of integrated transportation systems as a whole.

Recommendation: DoT should upgrade its research program, taking responsibility for managing an ambitious program of technical research as well as economic and policy analysis. This might be done by greatly expanding the Research and Innovative Technology Administration in DoT. Included in the program should be creation of multidisciplinary university-based research centers with long-term funding

44. President's Council of Advisors on Science and Technology. (October 2010). *Accelerating the Pace of Change in Energy Technologies through an Integrated Federal Energy Policy*. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-energy-tech-report.pdf

to explore surface transportation technologies, including long-term synergies with urban design. The effort should include innovations that present real-time information to transportation users, giving them the ability to make convenient and personalized decisions that optimize their transportation choices.

Recommendation: Under the leadership of NSTC, the Federal Government should build an integrated R&D program, coordinated by NITRD and involving DoT, NIST, DoE, DoD, and the Department of Housing and Urban Development (HUD) that furthers a broad range of transportation missions.

Recommendation: NIST should use its convening power to create and impel a structure for promoting interoperable standards for real-time transportation information. It should build on the experience of the IETF, whose emphasis on "rough consensus and working code" – pragmatism, interoperability, the interworking of multiple transmission technologies, and foresight achieved by significant engagement of the research community – has fostered the Internet's ability to "innovate at the edges."

The above recommendations are condensed and summarized in the Executive Report as follows:

Recommendation: **The Federal Government should invest in a national, long-term, multi-agency, multi-faceted research initiative on NIT for energy and transportation.** As part of that initiative:

- DoE and NSF should be major sponsors of research for achieving dynamic power management in applications ranging from single devices to buildings to the power grid.
- NIST should organize the multi-stakeholder formulation of interoperable standards for real-time control. Interoperability facilitates repeated cycles of innovation by multiple vendors, promoting the development of versatile and robust NIT.
- DoD should continue to be a major sponsor of research on using NIT to achieve low-power systems and devices.
- DoT should sponsor ambitious NIT research relevant to surface and air transportation.

NIT for National and Homeland Security

NIT has three important roles in national and homeland security. First, in order for government agencies charged with protecting the Nation to act in a situation, they must collect, analyze, and distribute critical information in a timely way. To do that, they must have superior information, both historic (large data collections) and current. Second, the NIT infrastructure that these agencies use to collect and analyze information must be stable and resilient in the face of natural disaster and cyber-attack. Third, many elements of our Nation's critical infrastructure – the financial system, the electric grid, the air traffic control system, etc. – rely on NIT; these systems too must be stable and resilient. NIT is a key enabler for security, but cybersecurity is the Achilles heel. Fundamental NIT advances are needed.

Recommendation: Section 7 of this report includes recommendations for fundamental research in two areas, Large-Scale Data Management and Analysis, and Trustworthy Systems and Cybersecurity, that are important enablers for national and homeland security. The recommendation that there be research to explore fundamentally new approaches to cybersecurity is particularly important for protecting our

Nation. DoD and DHS should work closely with the managers of those programs and with key researchers, both to ensure that the problems on which they work include those of crucial importance for national security and to incorporate the most promising research results into all stages of system development.

Recommendation: DoD and DHS should sponsor research to:

- find ways to apply as appropriate different standards of reliability and trustworthiness, so that life- and mission-critical systems can be scrutinized with the utmost care, while more relaxed standards can be applied elsewhere;
- develop techniques for recording and analyzing the provenance of hardware and software components, as well as techniques for forensics, so that when systems falter or fail, analysts can determine the problem and upgrade the system for the future;
- continue to develop new defense mechanisms for today's infrastructure as well as methods to reduce the likelihood of inadvertent human action that contributes to security breaches.

We observe that the Nation's cybersecurity leaders have not yet fully engaged the academic research community. Technology is moving rapidly and tomorrow's Internet will have or can have different properties than today's.

Recommendation: Intelligence, defense, and administration cybersecurity leaders should have frequent interchange with the research community that is inventing tomorrow's Internet, not simply those experts whose field of vision is limited to protecting the Internet of today. We recommend establishing a standing advisory committee that is structured for frequent, substantive and direct interaction among these groups.

These recommendations, along with other recommendations from Section 7, are condensed and summarized in the Executive Report as follows:

Recommendation: The Federal Government should invest in a national, long-term, multi-agency research initiative on NIT that assures both the security and the robustness of cyber-infrastructure. NSF and DoD, in collaboration DHS, should aggressively accelerate funding and coordination of fundamental research

- to discover more effective ways to build trustworthy computing and communications systems,
- to continue to develop new NIT defense mechanisms for today's infrastructure, and most importantly,
- to develop fundamentally new approaches for the design of the underlying architecture of our cyber-infrastructure so that it can be made truly resilient to cyber-attack, natural disaster, and inadvertent failure.

NIT for Education

Realization of an information-rich economy based on extensive use of NIT will require an increasingly well educated workforce. That, in turn, will require better education of children and young adults, and continuing education of the workforce throughout their careers. The potential for advances in NIT to greatly enhance teaching and learning has barely been tapped.

Recommendation: ED, in collaboration with NSF, should provide robust and diversified support for fundamental NIT R&D that will lay the foundation for educational technologies such as personalized electronic tutors, serious games and interactive environments for education, and mobile and social education technologies. The support for NIT-based education should extend from pre-school settings to lifelong learning.

Recommendation: ED, in collaboration with NSF, should have a long-term program to evaluate promising technology coming out of the research community in trials that include large numbers of sites and participants. Technology that proves its worth should be transferred into the schools. This program will require evolution of curricula and school processes and procedures.

In its recent report on K-12 STEM education[45], PCAST recommended the creation of an Advanced Research Projects Agency, ARPA-ED, for educational technology R&D. The above recommendations fit naturally within the charter of such an agency.

NIT for Digital Democracy

NIT provides an opportunity to enhance in practical and constructive ways the flow of information about government and the ability of citizens to participate in government. The research recommendations in Section 7 concerning Privacy and Confidentiality, NIT and People, and Large-Scale Data Management and Analysis will lead to important advances in our ability to enable all our citizens to engage in digital democracy. Today's technologies and tomorrow's advances must be put into practice.

Recommendation: NSTC should lead a multi-agency effort to define infrastructure, tools, and best practices that will increase the opportunities for digital democracy at all levels of government. Both the NIT research community and representatives of the public at large should participate in the planning process. The plan should have an emphasize on using the results of fundamental research in NIT to enable more efficient government and to improve the quantity and quality of information and ideas flowing into, out of, and within government. It should create pathways for fundamental research to be explored and evaluated on national testbeds, and for high impact approaches to be translated into practice.

---

45. President's Council of Advisors on Science and Technology. (September 2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

6. NIT Research Frontiers

At the time of the High-Performance Computing Act of 1991, the importance of HPCC to scientific discovery and national security was the major factor justifying special attention to NIT. Today, many other aspects of NIT have risen to comparable levels of importance. HPCC is now but one of many important areas of NIT, and America's prowess in HPCC is but one of many measures of our international competitiveness in NIT.

> **Finding:** The discipline of networking and information technology has both broadened and deepened in the roughly 60 years of its history. Although high performance computing and communication continues to contribute in important ways to scientific discovery and national security, many other aspects of NIT have now risen to comparable levels of importance.

In this section, we map the research frontiers in NIT, identify key opportunities in each of the frontier areas, and relate them to the national priorities discussed in Section 4.

6.1 NIT and People

One of the most striking features of the revolution enabled by the Internet and the World Wide Web over the last two decades is the extent to which it has been fueled by the contributions of millions of users, the vast majority of whom have little or no technological or programming prowess. It is easy to forget how fundamental this phenomenon was and still is. The Internet was the domain of scientists and the military for decades before the World Wide Web overlay made it possible for anyone to easily access and publish information. For search engines to offer any value, someone had to provide content interesting enough that others would want to seek it out. Those "someones" were the users themselves. More recently, companies such as Facebook and Twitter have further evolved and shaped user contributions, structuring them in novel ways and transforming them into more personalized, local, and social content.

The willingness of users to contribute freely to public NIT systems and artifacts in interesting and powerful ways has taken another leap forward in recent years, in the form of the phenomenon we refer to as "social computing" (which includes "crowdsourcing," "peer production," and "social media"). In many recent systems, users are not merely contributing information about themselves and their interests or expertise, but are actively participating in much broader collective causes. The influential World Wide Web site Wikipedia, a user-created modern online encyclopedia of general knowledge (and much else), is perhaps the canonical example.

Researchers in many disciplines, however, are increasingly taking note of a new phenomenon, in which surprising numbers of non-specialists are showing an appetite and aptitude for participating in an astonishing variety of collective, distributed tasks, from the mundane to the sophisticated. Examples that have yielded genuine scientific insights include labeling digital images according to their content[46] (which in turn can be used to train statistical models for image recognition), playing an online game that contributes to protein design and protein structure prediction[47], classifying galaxies in astronomical images[48], labeling of news stories and other documents by sentiment and tone, and many others.

Users participating in these diverse examples have a wide range of motivations and interests. People providing Wikipedia entries and folding proteins, for example, are acting quite intentionally, and expending skill and a significant time commitment to do so. But people doing character recognition for reCaptcha or labeling images as part of a game are making contributions as a side effect of their normal activities. And many people rating their Netflix rentals or shopping at Amazon may be unaware that the World Wide Web sites are capturing information derived from their actions. But all of these users, directly or indirectly, contribute to the creation of information systems and artifacts of tremendous collective power.

These developments have led to a distinct sense among scientists and technologists that we are in the earliest stages of a new era in human-machine interaction, in which advanced NIT methods coordinate the distributed but concerted efforts of human users on collective tasks of commercial, scientific, societal, political, and military importance. The questions raised thus far by such applications are already deep and numerous. What is the best way of organizing distributed users in the service of any given goal, task, or problem? Why do so many people voluntarily contribute to a collective task, and what incentives, economical or otherwise, will encourage them further? How can we harness distributed human input to perform tasks that are not easily parallelizable but require careful coordination among subtasks? How can humans and computer systems collaborate on tasks, utilizing the greatest potential of each?

"Crowdsourcing" and "peer production" have to date largely been applied in the service of practical goals such as protein folding or image labeling. But the same underlying phenomena also herald a new era in economics and the social sciences. The intersection of NIT with these disciplines is in its infancy. But the Internet and the World Wide Web provide the opportunity for large-scale experimental studies that could not have been imagined just a few years ago. Crowdsourcing efforts and the empirical documentation of online social network structure have demonstrated that even uncontrolled environments and studies can provide valuable scientific insights in the investigation of both traditional sociological questions and more recent ones, such as the diffusion of influence through social networks. Among the many revolutionary capabilities provided by the World Wide Web are those of scale – both in the number of people that can participate in studies, and in the number of studies that can be conducted. There is also tremendous promise in using NIT-enabled studies involving large numbers of people to shed new light on areas such as privacy and security, where behavioral considerations are paramount yet poorly understood.

This is just the beginning of the new field of collective intelligence, in which modern technology yields new understanding of collective human behavior and new methods for problem-solving in complex systems and networks. We now summarize some of the most important general research directions in these topics.

46. http://images.google.com/imagelabeler/

47. http://fold.it/portal/ See also "People power: Networks of human minds are taking citizen science to a new level." *Nature* 466, (August 2010). http://www.nature.com/news/2010/100804/pdf/466685a.pdf

48. http://www.galaxyzoo.org/

- *Creating a science of social computing*: The successful examples of crowdsourcing to date are just that – examples. Scientists and technologists don't yet know how to take the lessons of one success or failure and apply it to another problem. So far, in other words, we have ad hoc solutions without any underlying theory or engineering principles. An ambitious goal for a science of social computing would be a "crowdsourcing compiler," which would view people, computer, and network ensembles as a single architecture and enable designers to write in a high-level language that would automatically compile systems that knew what parts of a problem to outsource to peer production, how to organize the human contributions, how to incentivize them, and so on. But well short of this perhaps overly ambitious goal remain many basic questions whose answers could have great societal benefit. Online collaboration offers great opportunities across many scales, from a local club or Girl Scout troop up to large transnational groups. As with a number of World Wide Web technologies that have become standard commodities, these social computing systems could be leveraged by small mom-and-pop operations and large multi-national corporations alike. In health-related social media applications, for example, the same underlying technology can be used to solicit data and deliver messaging to many millions of participants about broad health concerns such as diabetes, but could also serve communities of thousands with an interest in much rarer conditions such as ALS (cf. PatientsLikeMe[49]).
- *NIT-enabled sociology*: As mentioned above, NIT-enabled sociology is not merely a matter of using the Internet and World Wide Web for investigating questions of traditional sociology on large scales (though it holds great promise there as well), but is particularly important for NIT itself, since so many important questions today involve the interaction of technology with large numbers of people. One "grand works" project would be the creation of a national infrastructure for large-scale, controlled Web-based human-subject studies. "Large-scale" in this context means not only studies with large populations, but studies involving large numbers and repetitions of trials on small populations or even individual subjects (something that is onerous and expensive for in-person, offline human subject studies). While some behavioral researchers are already starting to use services like Amazon Mechanical Turk to recruit subjects, there has also been some discussion of the kind of shared platform we envision, complete with subject panels, potentially detailed demographic and cross-sectional data, adherence to institutional review board considerations, and so on.
- *A New HCI for NIT and People*: The long-standing and important interdisciplinary field of HCI has traditionally focused on a single person interacting with a single computer or system. The importance of this work increases as computer-based systems become more ubiquitous and as the population of computer users expands from the technically knowledgeable to embrace an ever-larger segment of the general public. In addition, there is a new challenge: *populations* – sometimes very large ones –interacting both with machines and each other via systems that are distributed in space and time, and that often have a common collective goal, as in Wikipedia, Facebook, and crowdsourcing systems. These new types and scales of human-machine systems will present important new HCI research challenges, such as understanding patterns of adoption

49. http://www.patientslikeme.com/

or non-adoption of tools and services, the management of diverse and sometimes competing incentives, the determination of appropriate compensation of various kinds, and the detection and control of free-riding and other socially undesirable human or system behavior.

6.2 NIT and the Physical World

The Internet is expanding from the virtual world to embrace the physical world. We can access weather information almost anywhere around the globe. Smartphone users can set their home security systems no matter where they are. iRobot shipped roughly 1 million Roomba vacuum cleaning robots during the past year. This revolution will only accelerate in the coming decades. By embedding instrumentation in our buildings, vehicles, and factories, we can transform them from static, inefficient edifices to adaptive infrastructures with just-in-time resource consumption that accommodates growth in the face of pressure to reduce fossil fuel consumption. By developing rich ecological observing systems, we can create accurate high-resolution models that support forecasting and management of increasingly stressed watersheds and ecosystems. By taking advantage of increasingly miniaturized and powerful sensors and radios embedded in mobile, wearable, and residential devices, we can create a healthcare system that helps people prevent and manage chronic and acute diseases in their own everyday contexts, not just on occasional visits to clinical facilities. And by combining smart sensing with actuation, we can bring the power of robotics to a broader range of transportation, medical, and manufacturing settings.

Already, we swim in a sea of sensors. From our phones and our cars, from our increasingly instrumented homes and offices, from health monitors and environmental sensors, streams of new data constantly emerge. We can transform our interactions with the physical world by creating intelligent, adaptive, and highly personalized and private systems that continually capture and analyze a broad array of data about us and about the world around us. So far, though, we are vastly under-utilizing the available data. We need to devise systems that analyze all this data, both in real-time and retrospectively, to create a coherent picture along with meaningful feedback, that can help us navigate the modern world while also protecting our privacy.

To achieve this vision we need numerous NIT advances in data fusion and inference techniques that operate over diverse and noisy raw data streams in combination with historical and contextual data. A rich set of basic data processing techniques, including computer vision techniques as well as systems for understanding speech and natural language, will be needed to deal with the diversity of data types. Then, to extract useful and personally relevant events and insights from the cacophony of processed, but still disparate, data streams, we need advances in data fusion, causal analysis of multivariate data, and data mining. Finally, it will take further advances to create these systems in ways that are accountable and usable, preserve privacy, but also make maximum use of secure, robust and increasingly cloud-based data stores. These challenges present fundamental NIT research opportunities in the development of much-needed miniaturized transducers, data fusion in human-cyber-physical systems, autonomous actuation, and open architectures.

Physical world transducers. For countless applications, smart transducers capable of "digitizing" the physical world are a core need. Examples are wearable devices that can accurately report when the wearer falls, embedded devices that monitor and control the operation of precision engines, or small imagers, on the scale of plant roots, that would constitute subterranean observatories. The development of smart, miniaturized, low-power, self-calibrating, inexpensive, and adaptive instrumentation has been and continues to be critical to the advancement of physically-coupled systems in all challenge domains we discuss in this report.

As evidenced by heterogeneous observing systems such as the National Ecological Observatory Network, (NEON), the Ameriflux network, and the Ocean Observatories Initiative, (OOI), the ecological, climate, and ocean science communities have made data-intensive monitoring, modeling and forecasting of full watershed ecosystems a key priority. These full-system, geographically widespread observatories rely on heterogeneous instrumentation arrays, working in concert with remote sensing assets, region-scale instrumentation, surveys, mobile data sources, and models. Sensors are the elements of these observatories that actually couple to the physical world; advances in transducers are thus crucial to advances in observatories and observatory science.

One of the cross-cutting advances in this area is the development of sensor arrays based on microelectromechanical systems (MEMS), such as those found in high-resolution projectors and in biological applications (lab on a chip, a device that integrates multiple laboratory functions on a single integrated circuit of only millimeters to a few square centimeters in size). While many design advances have been achieved in the laboratory, relatively few of these advances have made their way into deployable devices because of a lack of coordinated funding for translational work. Sizable capital investment is needed to move published research results from the laboratory into deployable products. Moreover, each sensor type, particularly chemical and biological sensors, tends to be highly specific to a particular chemical or biological species. Consequently, many technically feasible transducers with potentially high scientific impact are never produced, because they do not have a large enough market demand to justify the investment needed for complete product design and manufacturing. Of particular importance is advancing the availability of accurate, low-cost, miniaturized deployable chemical and biological sensors for use in soil and water monitoring, the need for which is shared across a wide array of agencies.

Processing, correlation, and coordination across heterogeneous human-cyber-physical system data streams. A second core technical challenge is automatic extraction of information from dramatically diverse inputs – including continuous digital streams generated by embedded physical and chemical sensors, processed features extracted from high-resolution images, and highly subjective yet highly perceptive human inputs emerging from a wide range of social media. In addition to the classic challenges of signal processing and data fusion from sensor arrays, this plethora of input sources presents unprecedented heterogeneity in their temporal and spatial scales, as well as in their accuracy, coverage, and completeness. Moreover, these inputs are increasingly part of control loops that must make decisions and take actions in close to real time. The meaning of the information will depend on the contextual metadata about what was measured, how, and when. Methods need to be developed to catalog, search, and summarize the data, so that it can also be used outside of the particular vertically integrated system in which it is acquired.

NITRD investments in Cyber-Physical Systems have contributed to R&D advances in several aspects of this area, but far broader progress is needed to realize the application of these advances in our national priority areas. For example, personalized measurement is key to understanding, managing, and preventing chronic diseases, which present one of the greatest challenges to our Nation's health and healthcare system. The advent of personalized measurement systems depends on advances in NIT to create robust, usable, trustable, and affordable instrumentation of health-related exposures, habits, and status. Meaningful personal health monitoring requires a synthesis of diverse inputs from wearable and mobile devices, along with environmental measurements. Some of these devices automatically or passively capture data, while others support patients' self-reporting. The outputs of these systems will inform diagnosis and treatment plans, adjustment of medication, and patient support, as well as population and community health research. Broader advances in cyber-physical systems will also revolutionize personalized assistive instruments for patients with more acute needs, by creating adaptive systems that can be personalized to the status and context of the user. Examples range from wheelchairs and prosthetic limbs to systems that compensate for visual, auditory, and memory impairments, for example by leveraging rich contextual knowledge fused from diverse sensor and data sources (from physiological measures to GPS to maps).

A second example of the critical role of broadly capable cyber-physical systems is in crisis response, which may emerge as one of this decade's most important national security concerns. Crisis response requires the synthesis of highly diverse and noisy inputs to support decision and action. Knowing the situation on the ground, coordinating logistics and triage, supporting the needs of individuals to find family, friends, and other sources of support – all these capabilities are enabled by rapidly deployable distributed and mobile sensing and communication systems that are designed to incorporate remote and *in situ* sensing, broad human input, participation, and visibility.

Autonomous actuation. Robotics and vision-based NIT is the engine for systems that not only observe the physical world but also navigate and manipulate it. The demand for autonomous and semi-autonomous mobile systems will grow tremendously across challenge domains – from seafloor discovery to intelligent transportation systems to adaptive prostheses and robotic surgery. Also relevant here are defense applications using unmanned aerial vehicles, (UAV), from the battlefield to search and rescue.

Careful consideration and nurturing of R&D in robotics and automated systems can reap dramatic benefits in productivity and growth in the manufacturing sector of the U.S. economy. Over the last several decades, the plummeting cost of computation and physical sensors has enabled robotic automation in manufacturing to expand from very high-end applications (such as military systems and aircraft) and repetitive routine actions (such as automobile assembly) to much wider adoption at the low end of manufacturing, and to less routine, more adaptive and responsive applications made possible by increased computational "intelligence."

However, significant advances are needed to reduce the need for fine-grained control of robots and other automated devices by people, and to increase the ability of robots to work in hybrid human-cyber-physical systems. Advances are needed in visual object recognition in uncontrolled, obstructed, and dynamic environments, in the development of new sensors and materials to enable more flexible orientation, navigation, manipulation, and interaction, and perhaps most importantly, in new learning algorithms for creating robotics with the adaptive intelligent behavior it needs to operate as part of our everyday environments. Ultimately, greater adaptability and lower costs will make it possible for people to "teach" or "program" robots to perform simple manufacturing tasks (in much the same way that office workers learned to "program" or codify their knowledge using tools such as Microsoft Excel).

Open architectures. Application of NIT to the physical world will benefit greatly from the adaptation of open systems architectures that promote scaling and robustness. Open architectures also facilitate rapid decentralized innovation by encouraging the use of modular and interoperable components.

Smart grid system architecture, for example, will support adaptive, distributed resource management and control (including energy storage) at levels of aggregation that produce efficient system results. This architecture would be capable of iterative innovation, driven by comprehensive instrumentation that generates rich analytics about itself. Open architecture in smart transportation will require NIT advances that allow sensors and other devices to be embedded at every scale of the system – from individual vehicles down to vehicle components, and from passengers to operators of the system. Such a transportation system can become efficient through multi-timescale adaptation (from efficient engines to collision avoidance to congestion-based routing and coordination of shared vehicles and micro-public transport), and self-describing, enabling the designers, operators, and users of the system to contribute to both local and global efficiencies.

Our healthcare system is also greatly in need of more open, more modular components that promote data liquidity. Moving away from monolithic systems that cannot interoperate with other systems will foster agility and innovation that is now markedly absent from health information technology.

6.3 Large-Scale Data Management and Analysis

Collecting, storing, preserving, managing, analyzing, and sharing exponentially increasing quantities of data present a variety of significant NIT challenges that research must address. A vast range of data sources, from webcams and weblog posts to telescopes and supercomputer simulations, is flooding our world with enormous amounts of data in many different forms. These data are stored in many different formats and many different environments, from computer hard drives to large-scale data warehouses. Numerous unresolved issues arise in attempts to ensure that all of this data maintains its integrity and availability, both now and long into the future. Making use of the data presents another set of challenges. Advanced machine-learning algorithms enable sophisticated analyses of data sets, leading to breakthroughs in science, medicine, commerce, and national security (see sidebar "Extracting Worldly Knowledge from the World Wide Web", on page 50 for one example). Graphical visualization and other methods allow people to gain valuable insights from large collections of data. To gain maximum advantage from data sets that continue to grow larger and more complex, new techniques are needed in both these areas. There is also a critical need for better techniques for sharing both data and the results of data analyses while respecting the privacy rights of individuals and the needs for government and corporate confidentiality. Effective use of data will be critical to meeting every one of this report's technical priorities. Below we describe the fundamental elements of working with data and some of the key R&D challenges and coordination needs they engender.

Extracting Worldly Knowledge from the World Wide Web

Although the field of machine learning has made enormous progress in recent years, the scope for further research remains almost unlimited.

People learn throughout their lifetimes, and at an accelerating pace – accumulated knowledge facilitates further learning. For example, to understand the difference between the sentences "Clarissa went to the store in her car" and "Clarissa went to the store in her neighborhood," we need to know that a car is a vehicle while a neighborhood is a place. Various computer science research projects over the years have attempted to manually create "knowledge bases" containing essential information of this sort, but the task has proved impractical.

Instead, the Never Ending Language Learning (NELL) project at Carnegie Mellon University is extracting these facts from the World Wide Web[50,51]. Starting with a small set of categories and examples, NELL seeks statistical reinforcements for the facts it has already learned, and uses what it knows to extract new facts. For example, when it sees the statement "Michael McGinn is the Mayor of Seattle," it can infer that Michael McGinn is a person and that Seattle is a city. When it later sees "Dallas is bigger than Seattle," it can infer that Dallas is also a city. NELL uses a collection of 1 billion pages extracted from the World Wide Web by means of a system that Google provided to NSF-sponsored researchers. Each iteration of its analysis requires around four hours on a massive computer system made available to university researchers by Yahoo!®. Funding for the project has come from DARPA and Google. As of October 2010, NELL had extracted nearly 500,000 facts with an estimated accuracy of at least 87%. Other researchers are starting to use NELL's knowledge base to improve natural language understanding programs and to create better World Wide Web search engines.

We use two additional practical examples to illustrate both the promises and the challenges of large-scale data collection and analysis. First, consider a consortium of cancer research institutes creating and managing a data repository based on data collected from millions of patients. These data include copies of medical histories in textual and spoken form, x-ray and MRI images, and test results from biopsies and genetic microarrays. Assembling and managing all of these data is a monumental task, but it could lead to much a deeper understanding of disease processes and the way different medications affect different populations. As a second example, consider the case of the FBI collecting data from many different sources in many different forms: video surveillance data, intercepted email and telephone calls, law enforcement records, and even online information, such as web pages, videos, and weblog posts. Among the masses of information are small traces indicating the activities of crime rings and terrorist organizations. This data must be collected in ways that preserve individual rights against unreasonable searches and seizures, but by analyzing this information, the FBI may be able to uncover and disrupt major threats to the safety and security of U.S. citizens. In both these examples, management and preservation of the data to ensure its future accessibility is an important element of our ability to gain new insights and understanding.

50. http://rtw.ml.cmu.edu/
51. http://www.nytimes.com/2010/10/05/science/05compute.html

Research challenges in data collection, storage, and management. It is estimated that around 1.2 zettabytes (1.2 billion terabytes) of digital data are generated worldwide each year by numerous devices in numerous forms: remote sensors, online retail transactions, text documents, email messages, web posts, camera and video images, computers running large-scale simulations, and scientific instruments such as particle accelerators and telescopes.[52] The core technology for data storage, especially magnetic disks, has progressed rapidly, enabling government, research, and corporate organizations to create massive data warehouses that can store of much of the data as fast as it is created. But storing raw data is just a small part of the larger issues of creating and maintaining data repositories. We must view data repositories as archives requiring long-term stewardship based on sustainable economic models. Important issues include:

- *Representations*: How to adopt and evolve standards for important categories of information. These representations must allow different companies and organizations to create software tools that generate, manipulate, and analyze societally important data. Left on its own, the software industry is likely to create a number of incompatible, proprietary standards that become obsolete. (Consider, for example, the case of word-processing formats and the fact that the Federal Government still mandates using WordPerfect format for official documents, long after most organizations have transitioned to other software.)
- *Detecting and correcting errors or inaccuracies in the data*: Although various forms of outlier detection have been developed and applied, these methods need to be more sophisticated and comprehensive when applied to data sets of societal importance.
- *Support for data management policies*: Systems to support data privacy and access limitations, retention requirements, requirements for mechanisms to reduce the risk of data loss or damage, and other aspects of increasingly stringent data policy and regulatory requirements.
- *Data provenance*: Tracking how, where, and when data are created and modified. This is an important and often overlooked aspect of data stewardship.
- *Data integrity*: Ensuring that data are not corrupted either accidentally or maliciously.
- *Data storage engineering*: Ensuring reliability, reducing power consumption, incorporating new technology. Management of data across multiple storage technologies and multiple hierarchies, and with replication across multiple geographic locations. Continued research is required to adapt to changing technology (e.g., nonvolatile RAM), performance requirements, and the need to provide consistent views of data worldwide.
- *Development of sustainable economic models*: Necessary for supporting data access and preservation over the long term, especially beyond the durations of typical research grants.

Many of these requirements appear in our cancer research institute example. Interoperability for electronic health records, considering both current and future needs and technology, are of critical national importance both to realize the promise of better healthcare and to enable the use of patient data for medical research. As for recording and tracking data provenance, suppose it was determined that an

52. http://gigaom.com/2010/05/04/we-cant-squeeze-the-data-tsunami-through-tiny-pipes/

automated blood analysis instrument had been giving faulty readings over a one-month period. It should be possible to identify all patients who might need to be retested and all scientific analyses that may have been tainted by faulty data. It is crucial to maintain original data for medical and scientific research to enable validation of results and to support longitudinal studies. In addition, regulations such as the Health Insurance Portability and Accountability Act (HIPAA) create many requirements for data access and retention. Similarly, the FBI must collect and manage data in ways that satisfy rules of evidence. It must keep track of data provenance so that it can later carry out deeper investigations of critical information and use the information when prosecuting legal cases. Errors either in the initial data or due to subsequent corruptions could have devastating effects on innocent people, as well as on the ability of the FBI to carry out its mission.

Research challenges in data analysis. Increasingly sophisticated methods of data mining and machine learning allow us to extract more and more useful insights from many data sources. Prominent recent examples include search engines, automated language translation, customer recommendation systems, and credit card fraud detection. This is an area of very active research with ever-increasing capabilities, but also with increasingly high expectations. Important issues include:

- *Systems*: Engineering computer systems that can perform complex processing of data on very large scales. Internet-based industries are building computer systems of unprecedented size to house and analyze their data and to serve millions of customers worldwide. Comparable systems could also provide powerful capabilities for scientific research, for making government data available to citizens, and for national security.
- *Algorithms*: Developing more sophisticated machine learning techniques, especially ones that apply to very large data sets. Machine learning is still in its infancy, and we can reliably predict that great strides will be made in creating algorithms with new capabilities that can scale to handle the very large data sets being generated now and in the future.
- *Programming*: Computational models and languages suited for expressing data analysis algorithms that map onto large-scale, parallel systems. Recently developed proprietary and open-source programming tools have demonstrated greatly enhanced scalability and programmer productivity. These tools and models must be extended and refined to handle wider classes of applications and to make them easier for non-specialists to use.
- *Cross-media information extraction*: Understanding speech, images, video, and unstructured data; translating speech and text to other languages. Although these topics have been the subject of decades of research, new data-driven approaches promise to be much more effective.
- *Information fusion*: Analysis that combines multiple data sources in multiple different forms. Many important insights can be gained by analyzing different representations of a single event or phenomenon.

Using our cancer research institute as an example, we can imagine important medical breakthroughs being made by systematically analyzing imaging data (x-ray, MRI) along with patient histories, including automatic transcription of dictations from patients and caregivers. These can lead to new diagnostic regimens that are far more efficient and effective than today's methods, which rely heavily on the experience and judgment of medical specialists. Dealing with the data from millions of patients will require computing capabilities far beyond those currently being used in medical research, and will require collaborations between medical researchers and a wide range of computer scientists and engineers. Interoperability will permit the analysis of much larger patient populations.

For the case of the FBI, we can see that the activities of a criminal or terrorist organization will show up in many different forms. The patterns of communication between different individuals, via phones and email, can be analyzed as a social network, revealing its command structure and ways that it can most effectively be disrupted. It would be possible to track the movements of individuals through the locations of their phone calls, their use of public transportation and credit cards, and from video surveillance. Creating a comprehensive picture of these activities from the many different forms of data requires much higher levels of automation and sophistication than exists today.

Research challenges in controlled and effective data sharing. Sharing different forms and different aspects of data offers important societal and organizational benefits. However, numerous instances in which researchers have been able to breach the confidentiality of supposedly anonymized data sets demonstrate the difficulty of sharing data in the face of increasingly sophisticated analytic techniques. Some key issues include:

- *Models and algorithms for controlled data sharing*: Current statistically-based anonymization methods provide no real guarantees for data privacy. For example, such methods often assume that all data come from a single dataset, whereas many breaches of privacy result from correlating multiple data sources. The recently devised differential privacy model, on the other hand, considers the existence of additional data sources. Developing and applying algorithms based on such models is crucial to taking full advantage of the wealth of data sources available to society.
- *Data presentation and visualization*: Making complex data understandable to people, specialist and non-specialists alike. This involves determining what information to feature, and how. Visualizations of cancer tumors, Internet traffic, weather patterns, sociological data, etc., are important to facilitate new insights critical to understanding.

Again, we can see how these issues arise for our cancer research institute example. Although medical record privacy is considered very important, current policies and regulations are a patchwork of poorly specified and ineffective rules. With healthcare data, as with other data sets concerning individuals and organizations, we must greatly improve our ability to derive and share useful insights from data while preserving privacy and confidentiality. Otherwise, there is a major risk that either private data will be disclosed, or that we will have to impose such strict controls that useful information cannot be shared.

For the case of the FBI, there are many instances where data must be shared with other organizations – local law enforcement agencies, the security services of other countries, and even the public – but this must be done in ways that protect the rights of individuals and do not accidentally leak information about covert sources and methods. Current methods of classifying information, unfortunately, are not sufficiently reliable in the face of sophisticated data analysis methods. They are also far too labor intensive.

6.4 Trustworthy Systems and Cybersecurity

Information technology is a key enabler for all of the national priorities described in this report – education, health, energy and transportation, national and homeland security, discovery, digital democracy, and economic competitiveness. Cybersecurity is the Achilles heel of this enabler. In public presentations, leaders of the intelligence community have described wholesale theft of information, inability to protect networks, the ease of access for attackers who seek to implant attack mechanisms in U.S. systems.

The design of current systems followed an evolutionary path in which security measures were layered onto fundamentally insecure architectures. Such systems give attackers a large advantage over defenders. There have been enormous expenditures, at various levels of sophistication, on short-term defensive technologies – that is, countering threats as they arise. This work is essential, and it is heroic. It will not, however, get us ahead of the problem. What is required is a dramatically increased focus on fundamental research in trustworthy systems and cybersecurity that will shift the balance of power from attacker to defender.

We define trustworthy systems as those that do what users expect, and nothing more. Here, systems may be anything from pacemakers to electric grid control systems to distributed command and control systems in factories and for emergency responders. These systems can fail from hardware malfunction or software error. They can also be deliberately attacked. We use the term cybersecurity to denote protection of information and property from theft or corruption in the face of attacks, while allowing the information and property to remain accessible and productive to its intended users. There must be security policies, including privacy policies, that define the desired level of protection.

Dramatically increased focus is necessary on advancing the art and practice of designing and implementing trustworthy systems which act only as users expect them to act, even in the face of failures. In addition, the NITRD agencies should fund an aggressive basic research program to derive first principles and fundamental building blocks of trustworthiness and security – units of functionality that can be implemented in hardware, software or both. Using these principles and building blocks, researchers need to develop the knowledge to relate classes of attacks, defenses, and security policies, so that outcomes can be predicted. The objective is to create a basis for designing and building more secure networks and more secure systems. This "clean slate" fundamental research should span most aspects of cybersecurity.

Most cybersecurity today is based on perimeter defense. As has been often noted, perimeter defense techniques (walls, moats, locks on doors, the Maginot Line, firewalls) are inherently flawed. Pervasive solutions that are effective against both internal and external attackers need to be found. New solutions are needed so that a sub-community of users can operate with trust in each other, even while untrustworthy users share part of the same virtual space. That is the motivation for the "clean slate" research.

In parallel with gaining a better fundamental understanding, the broader R&D community needs to continue to pursue refinement of the techniques used today for cybersecurity and in the broader area of trustworthy hardware/software systems. Defenses need to be stiffened, and placed in depth, while hoping that substantive improvements can be found.

The following is a list of key R&D topics, illustrating the breadth and depth of the challenge.

- *Advancing trustworthy system characterization*: We must be able to clearly specify the desired properties of systems and then characterize how well a hardware or software design realizes these properties. Relevant properties include not just the nominal functionality, but the ability of the system to operate in the face of failures (including hardware failures, communication errors, and software defects) and attacks. We must also be able to track the provenance of components and their integration in order to assess the trustworthiness of the design and manufacturing process. We must be able to apply different standards of reliability and trustworthiness as appropriate, so that life- and mission-critical systems can be scrutinized with the utmost care, while more relaxed standards can be applied to other classes of systems.
- *Understanding and improving the social dimensions of trustworthy systems and cybersecurity*: Many security breaches stem from human shortcomings, such as poor password selection and vulnerability to social engineering attacks. Similarly, many system failures are characterized as "operator error." These vulnerabilities call for behavioral research in support of privacy and security and in systems designed to respect the cognitive abilities and limitations of their users. Conversely, we can make use of the "wisdom of crowds" by having security measures that rely on crowdsourcing to detect cyber-attacks, spam, and social engineering attacks.
- *Creating foundations for cybersecurity*: This includes starting with a clean slate, formulating a framework by which secure hardware, software, and networking building blocks can be constructed and composed. Fundamental mechanisms for authentication, authorization, and trust management must be designed, and there continues to be a need for fundamental research in cryptography, based on both traditional and quantum technology. Cybersecurity for infrastructure and process control systems must be based on sound foundations, rather than the current approach of applying industry best practices.
- *Formulating the definition and application of security and privacy policies*: Current policies, such as HIPAA, are a conglomeration of English-language documents that leave many details vague and that contain possible conflicts, making it difficult to design software systems that are truly compliant. Actionable policies must have a clear logical basis and be expressed in terms of languages that are understandable to both people and computers. Better methods of evaluating policies and the compliance of systems to these policies must be developed. Policies must be formulated in ways that are consistent with the capabilities and limitations of NIT, and the implications of new policies on the NIT industry and research communities must be considered.
- *Improving methods to detect and mitigate security attacks*: These methods must deal with attacks of many different forms and from many different sources, such as insider threats, large-scale attempts to compromise or overwhelm the Internet, and attacks on critical industries or infrastructures. Mitigation methods must include ways to operate in the face of attacks as well as forensic mechanisms to identify attack sources both during and following an attack.
- *Developing methods for the implementation of a "survivable core" of essential cyber-infrastructure*: Research in NIT-related areas relevant to critical infrastructure protection should not focus exclusively on defending the complex information systems currently in routine use. We must also develop methods suitable for implementing a small, rigorously isolated set of very basic

capabilities that can be relied upon with a high degree of confidence to provide truly essential NIT-based services on a temporary basis in the event that we are unable to prevent, for example, a catastrophically damaging cyber-attack.

6.5 Scalable Systems and Networking

Computer technology over the past 60 years has seen an astonishing triumph of scaling, producing systems that have grown in capability far beyond their early incarnations. Moore's Law, the familiar observation that the semiconductor industry can double the number of transistors on a single chip every 18-24 months, has held true for 45 years; the computing power of microprocessors has grown by six orders of magnitude over that time.[53] Similarly, the number of Internet hosts worldwide has grown in 40 years from four to four hundred million, while magnetic disk drives have increased in capacity by over seven orders of magnitude. A central tenet of modern computer science is that both hardware and software systems must be designed from the outset to be highly scalable – for example, by using data formats that allow for future growth, using protocols that eliminate bottlenecks as computations become larger, using well defined interfaces to permit connection to other software modules, and using algorithms that remain efficient over a wide range of input values and sizes, numbers of resources, and numbers of simultaneous users.

Success in all these aspects of scaling has been built on a foundation of sustained research. Semiconductor manufacturers, as well as the vendors who supply their fabrication equipment, have invested huge sums in techniques to make transistors smaller and wafers cleaner, and to ensure that circuits remain robust as they become smaller. An entire research field and industry has arisen to provide computer-aided tools for the design of billion-transistor chips. Disk and networking technologies have also required fundamental research to achieve and deploy scaling increases. Software systems and applications have had to change in fundamental ways to take advantage of scaling – for example, systems must behave reliably in large-scale distributed environments where component failures are inevitable, and applications must be designed to exploit large-scale parallelism to deliver improved performance. As an example, over the past 30 years, Microsoft operating systems have grown from 4,000 to nearly 100 million lines of source code[54]. Scaling will continue to provide future growth for NIT, but only if research can invent the techniques that will make it possible.

Most of the focus on scaling has been in one dimension, *up* – creating larger and faster processors, networks, and storage systems. This has been a crucial element in making systems that can solve more complex problems and communicate and process increasing amounts of data. But scaling has two other important dimensions:

- *Scaling down*: Creating systems that are smaller, more portable, and more affordable has led to an explosion of consumer products (mobile phones, cameras, MP3 players), and to easy access by everyone – individuals, small organizations, and large companies – to powerful NIT resources.

53. The Intel 4004 integrated circuit of 1971 had 2,300 transistors; Intel's next generation microprocessor will have over 1 billion transistors. See http://www.intel.com/about/companyinfo/museum/exhibits/4004/facts.htm and http://www.anandtech.com/show/3916/intel-demos-sandy-bridge-shows-off-video-transcode-engine

54. Swedin, E.G. & Ferro, D.L. (2005). Computers: The Life Story of a Technology. Greenwood Press, Westport, CT. http://www.forbes.com/forbes/2007/0226/050.html

By lowering the cost of entry in NIT, downward scaling has enabled small groups of researchers and entrepreneurs to innovate and experiment outside of mainstream efforts. Scaling down applies not only to physical characteristics, but also to the feature set and complexity of systems and applications.

- *Scaling out*: Embedding information technology everywhere and connecting everything via networks – creating what is called the "Internet of Everything" – has led to wide-scale sensor and control networks, as well as networked services for commerce, social interactions, communication, and computing. These ubiquitous services are often provided in a form known as "cloud computing," reflecting the fact that most customers no longer know or care about the location and configuration of the computing infrastructure.

The three dimensions of scaling – *up*, *down*, and *out* – are intimately connected. For example, reducing electricity consumption by microprocessors pays dividends for all kinds of scaling. Low-power microprocessors can run longer under battery power in consumer devices, but they can also be assembled by the thousands to form large-scale data centers and supercomputers. And supplying consumers with low-cost IT resources promotes more widespread access to the Internet of Everything.

Scaling issues for networked environments. The Internet is a marvel of engineering design. It has accommodated truly astonishing growth in scale and diversity over its 40-year history. Aspects of its design, though, are showing the strain. Real-time, safety-critical communications of the sort needed by the military or by crisis response teams provide one illustration of the gap between today's networking and our future needs. Today's networks lack some crucial properties: guaranteed operation despite equipment or transmission failures, extreme robustness against cyber-attack, dependable real-time transmission of critical data, rapid reconfiguration of communication resources, and others. Similar gaps come to light with the steady growth in networked systems that control our transportation and energy infrastructure, our national security, and our financial services. These and other shortcomings of our current networked systems are the object of research programs that promote "clean slate" design. Such research can lead to fundamentally better approaches, but governments, companies, and standards organizations will need to provide considerable and sustained effort and leadership to incorporate these ideas into the global Internet structure. Following are some of the concerns that need to be addressed:

- *Adapting network architectures to meet future needs*: Networks are proliferating in many different forms, going beyond Internet-style networks to include the local networks connecting processors within a data center, the on-chip networks connecting processors within a chip, and the networks connecting large sets of sensor nodes. They are being called on to perform more sophisticated services, such as routing among mobile users and supporting peer-to-peer data sharing. Current implementations provide such services by adding more layers to the system architecture and increasing the software complexity. Research programs such as the NSF-sponsored GENI, the Global Environment for Network Innovations[55], and Future Internet Architecture[56] programs are exploring fundamental changes to overall Internet system archi-

55. http://www.geni.net/
56. http://nets-fia.net/

tecture that will provide these services more efficiently. Similar efforts are required for other classes of networks.

- *Rethinking wireless spectrum management*: Radio frequency spectrum is an inherently limited resource. Existing static allocations lead to inefficient usage and provide limited space for emerging wireless services. Well-crafted public policy can foster innovation through incentives that lead to the development of spectrally efficient wireless systems. Fundamental research is required in improved coding algorithms, energy efficient circuits, and more efficient and secure routing techniques. Models of spectrum management should be developed with which to analyze policy alternatives. Experimental research infrastructure is needed to test theoretical innovations and to stimulate the recognition of as-yet unarticulated needs.

Scaling issues for emerging computing architectures. Improvement in the performance of single microprocessors has slowed. For that reason, processor technology is increasingly shifting to "many-core" architectures, where a single chip contains multiple independent processing elements, or "cores," working in concert with one another. Performance gains will come from the ability to program those chips and to construct systems containing thousands or even millions of cores on multiple chips connected by high-speed communication networks. Even individual devices as small as mobile phones and sensor nodes will become complex, internally networked systems. Harnessing the computing power of many-core systems both large and small will require a fundamental rethinking of how systems are structured and how application programs are written. Furthermore, the circuit technology that makes many-core systems possible will be much more susceptible to transient errors and failures, creating an urgent need for fault-tolerant design methodologies that build reliable systems from unreliable components. Research progress on the following issues is needed to enable current and future use of the emerging architectures.

- *Constructing and programming million-core machines*: Totally new system architectures, algorithms, programming models, and programming languages will be required to enable programmers to identify parallelism within their applications and to get a large collection of processors to work together in parallel while coordinating their activities and exchanging data with one another.

- *Radically reducing power requirements*: The power consumption of individual processing elements has become the most challenging design factor not only for small-scale systems operating on batteries, but also for very large scale systems, where the number of processors in a data center or supercomputing facility depends mainly on how many megawatts of power are available. Radically improving energy efficiency requires fundamental changes in many aspects of a system, including circuit technology, power-supply and cooling system design, and the software controlling the scheduling and mapping of processing and storage resources.

Privacy, security, and robustness challenges induced by scaling. The Nation's NIT infrastructure contains many layers of scalable systems forming a complex and interconnected set of resources. Much of the server infrastructure is organized into data centers, each one a single facility comprising thousands of machines. On the client side are millions of devices: desktop and laptop computers, smart mobile phones, and an increasing number of network-connected sensors and controllers, such as web cameras and intelligent thermostats in offices, homes, and out in the world. Servers and clients connect to each other via the Internet, which is itself a complex and layered interconnection of routing and communications infrastructure.

The design of this entire NIT infrastructure came about in a decentralized and evolutionary way. Some elements (e.g., the Internet protocols) were carefully engineered through a deliberative and open process. Some major elements (e.g., the Microsoft Windows platform) were developed by a single company. Other aspects arose through the efforts of many companies and organizations operating both competitively and cooperatively. This "bottom-up" approach tapped into the entrepreneurial spirit to generate a rich set of capabilities providing valuable services. At the same time, the decentralized process grew in the context of early government investment in university R&D that seeded coherent development of both the architecture of the Internet and World Wide Web and their open governance through the IETF and the World Wide Web Consortium (W3C).

The manner in which all this NIT infrastructure has grown has created obstacles to further progress, because some of the core decisions made in designing the original underlying system layers make it difficult to support key attributes required by today's and tomorrow's applications. The original design did not envision widespread use by untrusted parties, nor the profound dependence of modern society on the services supported by this infrastructure. Following are some of the concerns that need to be addressed:

- *Improving the security and privacy preservation capabilities of networked systems*: Security and privacy must start with authentication mechanisms that can reliably identify end users and reliably guarantee that data will be stored, communicated, and processed without corruption and with appropriate tracking for auditing and provenance identification. Such guarantees become more challenging as services become richer and interactions between service providers become more complex. At the same time, we must ensure that data is managed in ways consistent with privacy requirements, including, for example, the rights of citizens to have access to public information without fear of government monitoring or interference.
- *Creating systems that are robust in the face of both intentional and accidental disruptions*: System architectures at all levels must incorporate mechanisms to monitor their own operations, detect and diagnose failures and anomalies, and automatically adjust to provide the best level of service permitted by the available resources. Current systems are prone to catastrophic failure, for example when a number of machines are updated simultaneously with defective software. The advent of many-core systems greatly increases the need for fault tolerance mechanisms.
- *Providing stronger semantics for specifying performance and availability*: The current "best effort" service model has proved inadequate for managing some critical resources. Stronger semantics and a richer set of models are needed to quantitatively express performance and availability requirements for given services. Improved mechanisms are needed to create reliable systems services out of unreliable components, as is increasingly done in hardware systems.

The challenge of scaling down software. Scaling down becomes necessary both when the platform on which software executes has limited resources, as in the case of sensors or other very small devices, or when it is desirable that a software system be adopted by a different user base, with different requirements and a need for lower cost. In order to scale down, it may be necessary to simplify or eliminate some features. The choice of those design changes depends both on a clear understanding of requirements and constraints, and on a software architecture amenable to those changes.

An important example of scaling down is deployment of software systems to small businesses and organizations. For example, the way that healthcare software is built and deployed today is completely incompatible with the small physician marketplace. Software for the healthcare industry is chiefly designed for large providers (complex systems, sophisticated capability, expensive to install and maintain, etc.). Software as a Service (SaaS) models delivered via public cloud offer great potential for addressing this problem. Making the cloud work, however, especially in healthcare, requires major progress in software customization, together with all the other challenges described above that arise when working with a large networked infrastructure.

Improving operational efficiency. Setting up and maintaining current systems, including even individual personal computers, requires far too much effort and expertise. Burdensome system administration reduces the productivity gains NIT can provide and reduces the usability of NIT systems to large segments of the population. The administrative burden grows ever greater as systems become more complex and as the need to safeguard systems against attack increases. As we consider deploying networked sensors everywhere – throughout our homes, offices, and roadways – we must dramatically reduce the human effort required to operate these systems. Specific sources of improvement include:

- *Creating self-configuring and self-tuning devices*: A newly deployed device should automatically locate available computing and communication resources and negotiate how it will connect to and cooperate with them. During system operation, devices should automatically optimize their operations by continually monitoring their own activity and the environment.
- *Creating a richer set of service models and capabilities*: Many forms of cloud services are already appearing and developing. These operations must be more clearly defined and standardized so that services provided by different vendors can interoperate. The current patchwork of proprietary services means that customers run the risk of their data or operations being "trapped" within a single vendor's system.

6.6 Software Creation and Evolution

The role of software and the complexity of software systems continue to grow rapidly. NIT, of which software is an intrinsic part, appears in more aspects of our lives, not just in the form of computers, but as all manner of novel digital devices. Many devices that once relied on a great deal of customized hardware now consist of relatively standardized microprocessor platforms customized by software. Our expectations for functionality increase relentlessly.

Since the early days of computing, the United States has led the world in creating and deploying innovative software. Leadership in software is important for our economy, our security, and our quality of life. Yet despite enormous advances in our ability to build and maintain extraordinarily large and complex software systems, new and novel demands challenge our abilities. There is a continuing need for new software of diverse kinds, often at an increased scale and complexity. Because of the world's dependence on software-based systems and infrastructure, improvements continue to be needed in the security and trustworthiness of software systems, in our ability to adapt, evolve, and maintain existing software, and in our ability to integrate software with new computer hardware, new devices, and new kinds of interactions with people. The other sections of this report discuss the *uses* of software that require research and innovation. This section describes research that is needed to produce and maintain innovative software, to understand software behavior and properties, and to discern the characteristics underlying the design of certain kinds of software systems.

Improving software production. Software is created, maintained, and modified by people using languages, libraries, and software tools. Most commonly, people work in teams whose composition changes over time. Both the technological and the human aspects of software production require additional research:

The choices of language, libraries, and tools are sometimes dictated by economics, training, and compatibility concerns. Continuing research is needed to improve the available languages, libraries, and tools, in order to enhance software productivity and quality by giving developers a wider range of choices. The focus should be on improving productivity and quality for higher level applications development rather than for lower level systems software. To ensure major advances, we need fundamental long-term research that is not tied to immediate commercialization.

The human side of software production is itself an important area of study. The collaborative nature of the work, as well as the complexity of the product, require that structure be imposed on the process, that design and development choices be explicit, and that individuals be able to understand what others have done, both when software is created and long after. Many methodologies have been proposed to facilitate the collaboration and the transfer of understanding, but the principles that underlie the design of such methodologies are not well understood. Examples such as agile development, extreme programming, open source systems, design reviews, and virtual teams merit further study.

Some people have dramatically better than average software design and programming skills, and development teams can be structured to take advantage of those skills. However, the needs for software production swamp the availability of such heroic programmers. Continuing research is needed into approaches that will make it easier to create software for particular purposes. Research into ways of empowering domain experts or end-users (people whose expertise is in the purposeful use of the software) to create or customize software for themselves has not received sufficient attention.

Determining and ensuring software properties. Software should do what it is supposed to do, and not do anything unexpected or bad. Among the desirable properties of good software are correctness, customizability, high performance, low resource utilization, usability, and robustness. Properties that guard against undesirable behavior are trustworthiness, security, privacy-protection, and fault-tolerance. Research is needed into methods to achieve those properties and to preserve them in the face of changes in the software itself or in the hardware and devices with which the software interacts. Research is also needed into how, for a given software system, it can be determined whether the system has the desired properties. The methods need to scale up and out to very large and complex systems, since those are often the ones on which we are most dependent.

Improving Software Quality: "No Silver Bullet"

Frederick P. Brooks, a leading software engineer, once observed that "there is no single development, in either technology or management technique, which by itself promises even one order of magnitude improvement within a decade in productivity, in reliability, in simplicity."[57] "No Silver Bullet," in other words. True. However, *very significant progress has been made in software quality* by designing tools that draw upon a foundation of deep research. Here we profile the progress Microsoft Corporation has made in the past decade.

Ten years ago, software developers at Microsoft relied exclusively upon the same three software development tools that they and their colleagues at universities and other companies had used for decades: an editor for writing code, a compiler for translating the code into a form that computers could execute, and a debugger for examining and controlling a running program to find and understand program errors ("bugs").

These tools are still fundamental to software development. But as Microsoft strived to improve the reliability of software that was becoming part of the fabric of society, these tolls proved increasingly inadequate. The explosive growth of the Internet exposed another problem: malicious individuals were discovering software defects and exploiting them to quickly infect large numbers of online computers.

Meanwhile, computer science researchers were exploring new ways to build reliable software, ranging from languages with stronger safety properties to tools that reduced the incidence of "bad code." In the late 1990's, Microsoft Research started two related efforts. Amitabh Srivastava's Programmer Productivity Research Center built a number of very successful ad hoc defect detection tools, beginning with the acquisition of the tool PREfix from the startup Intrinsa. PREfix used a large collection of heuristics to search code for patterns indicative of coding errors. It was successful in finding large numbers of simple software bugs and is widely believed to have improved Microsoft's software quality. Unfortunately, no systematic studies were performed to analyze the improvement, but PREfix found one eighth of the bugs in Windows Server 2003.

Concurrently, Jim Larus started the Software Productivity Tools (SPT) group to develop systematic defect detection tools based on computer science research. The first problem was to develop scalable program analysis, which could efficiently analyze millions of lines of code. SPT's work led to scalable alias and value flow analysis algorithms that could understand complex relationships in millions of lines of code. These techniques were heavily used in later Microsoft tools, particularly for finding buffer overruns as part of the company's major security push. A second line of work for SPT was finding better ways to identify bugs in device drivers, complex, low-level parts of an operating system that are particularly difficult to write correctly. This research extended a technique called software model checking, which combined several ideas from hardware verification with ideas from program analysis, to create a new way of finding software defects – a highly effective application of formal methods to software.

Around the same time, Wolfram Schulte's Fundamentals of Software Engineering research group developed a series of innovative testing tools based on high-level specifications or models of program behavior. One of the tools, Spec Explorer, was used to precisely document the Windows application program interfaces as part of Microsoft's settlement with the European Union.

57. Brooks, Frederick P., Jr. (April 1987) . "No Silver Bullet: Essence and Accidents of Software Engineering." *Computer 20*, 4, pp. 10-19.

Attacks such as Code Red and BLASTER exposed the fundamental vulnerability of programming languages such as C and C++ to buffer overrun attacks and led to the development of tools to find these defects. These tools were built on an extensible framework called PREfast, which came out of earlier work in Microsoft Research. The tools required the annotation of function interfaces in millions of lines of code header files, a process that was largely automated using scalable program analysis techniques from Microsoft Research. The tools were incorporated into Microsoft Visual Studio and made available to all Windows developers, both inside and outside the company.

Microsoft's successful development and deployment of sophisticated program development and analysis tools built on several decades of research in programming languages, static analysis, and formal methods at universities worldwide. Within the United States, funding for this research, which necessarily preceded an understanding of how it would be incorporated into industrial-strength tools, came largely from DARPA, NSF, and the Semiconductor Research Corporation.

There are many facets to these problems and many approaches to their solution. Following is a sampling of some of the most important challenges:

- Provenance is the origin and derivation of a piece of software. Since systems are often constructed from components created or modified by unknown and perhaps untrusted suppliers, techniques are needed to determine and preserve the provenance of software.
- Static analysis is the determination of software properties by analyzing the program text. Simple measures such as lines of text or counts of the numbers of significant operations are sometimes used to estimate the development effort or complexity of the software. Analyses of the possible sequences of execution steps are sometimes used to estimate correctness or safety properties. Dynamic analysis is the determination of properties by observing the execution of the software. Simple measures include the number of instructions executed, the amount of storage used, or the branches taken. Neither of these methods is good enough to ensure the adequacy of the properties they aim to strengthen. More research is needed into the efficacy of such measures, and also into what software performance metrics would be both feasible to discover and more informative. Analysis of parallel and asynchronous execution also merits further study. Properties such as usability defy formal analysis at the present time.
- How to achieve desired properties in combination with each other is not well understood. Does privacy protection jeopardize security? Must performance be traded against fault tolerance? Understanding these issues by combining formal methods, static analysis, and dynamic analysis is in its infancy, and needs further research.
- Insuring correctness, reliability, security, and so on during development is better than retrofitting those properties later. Formal methods of development and analysis, software testing, and validation methods show some promise in addressing these concerns, but research is needed to do better.
- One way of improving software is to study and learn from existing deployed software systems. Evaluative research to understand why systems that were developed with the best of current

practice nevertheless fail to achieve all the desired properties would suggest where attention needs to be placed. Barriers to such studies need to be lowered.

Improving the design of certain classes of software. Software is sometimes regarded as a generic category, but certain classes of software present particular challenges and opportunities. Research is needed to identify the unique characteristics of those classes and to address the problems that arise. For example:

- There is an increasing variety of software systems that interact with the physical world, through sensors, imaging, robots, or computers embedded in mechanical devices. Innovative approaches are needed to design the software for those heterogeneous systems.
- Systems that interact with people, through screens and keyboards, hand-held devices, or other technologies, also require novel software that accommodates the particular cognitive, visual, verbal, auditory, and motor characteristics of people, as well as their behavior.
- Highly distributed (cloud) computing and computing that exploits both fine-grained and coarse-grained parallel execution require different algorithms, different software structures, different development approaches, and different analyses than more local and sequential kinds of software. More generally, different kinds of software would benefit from new computational models or abstractions – high-level declarative and rule-based designs, reactive event-driven models, or data-driven approaches.
- The advent of multi-core chips as conventional system components has made parallelism pervasive at every level of software, ranging from simultaneous execution of instructions, to multiple cores within a single chip, to networked many-core systems with multiple levels of parallelism and both shared and non-shared hardware memory access. It is difficult to reason about parallel execution when data is logically shared among processing elements and when the time to access that shared data can vary. Research is needed in the design of programming languages, algorithms, analysis tools, and systems software that make it easier for software developers to create, modify, and maintain programs and systems that take full advantage of parallel computing without requiring undue effort on aspects that could be left to well-designed tools and standardized forms of expression. Since multi-core chips are already on the market, there is considerable pressure to find solutions quickly. Although those efforts are important, they should not supplant long-term research to find approaches that will also accommodate future technology changes in the use of parallelism.

Programming issues for scalable systems. Networked systems necessarily involve the simultaneous operation of components that are subject to both transient and long-term failures of processing and communications – an environment that greatly adds to the complexity of creating, testing, and maintaining software. The emergence of networks of increasingly heterogeneous devices and operating environments exacerbates these challenges. Current methods to deal with these issues require considerable amounts of human labor, and aim to achieve reliability largely through regular downloading of software patches. Specific requirements include:

- *Creating programming models and languages that work across systems of different scales and systems with heterogeneous components*: Ideally, functionally identical code should be able to run on a single small device, in parallel on one thousand processors in a data center, or even throughout an array of geographically distributed processing resources, subject to possible limitations of the available processing and communications resources.

- *Raising the level of abstraction in application development*: Current software must both implement the high-level functionality of the application and also provide low-level management of processing, storage, and communications resources. The application developer must have a more abstract view of these resources if we hope to create application programs that meet the required goals for scaling, robustness, and adaptability to heterogeneous environments.

Open interfaces and open source. The importance of *interoperable* and *open interfaces* has been emphasized in a sidebar on page 15, and specifically noted in relation to areas such as smart transportation, smart grid, and electronic health records.

Open interfaces should not be confused with *open source software*, another important trend, which denotes software whose source code is available to others, sometimes without charge. This software is created by volunteers or by companies who wish to make certain software widely available and enable others to read it or change it. Although there are thousands of open source software components, open source thrives chiefly in two areas: (1) implementations of extremely popular software such as operating systems (Linux) or office suites (OpenOffice); and (2) implementations of extremely specialized software for which there is little or no commercial market. Many software packages that support scientific research, such as modeling natural processes or analyzing experimental data, fall into this category. Another example is the Hadoop framework for analyzing very large data sets: its open source code base is maintained by individual researchers and by programmers in companies that use the framework in their business.

6.7 High Performance Computing

High performance computing (HPC) encompasses the design of algorithms, software systems, and computer hardware to deliver the computing power needed to tackle the most computationally challenging problems, which are often highly:

- *compute-intensive*, requiring the exploitation of massively parallel computation involving a very large number of processing elements;
- *communication-intensive*, requiring the high-speed transfer of data among processing elements;

and, in many cases,

- *data-intensive*, involving the high-speed manipulation of very large quantities of data.

HPC has played a key role in addressing a number of the national priorities described in Section 4; its remarkable contributions have been well chronicled by the National Academies[58] and by the President's Information Technology Advisory Committee[59]. The highest performance computers – supercomputers

58. National Academies Press. (2004). *Getting up to Speed: The Future of Supercomputing*.
59. Report to the President, President's Information Technology Advisory Committee. (June 2005). *Computational*

– have been used, for example, to understand in detail how epidemics develop, to model traffic congestion in order to develop emergency evacuation scenarios, to study how the economic environment influences entrepreneurial behavior, to model climate change over long periods, and to learn how to make wind turbines quieter.

Of course, many aspects of our national priorities do not require supercomputing, and some problems that formerly required supercomputing can now be addressed without it, due to progress in algorithms and in midscale and desktop computer system performance. It is thus appropriate to review our Nation's HPC R&D needs in light of new technologies and technical challenges, the radical transformation of the world's cyber-landscape in recent years, and the future needs that HPC will ultimately be called upon to satisfy.

Breaking the Speed Limit

In recent decades, HPC has brought unthinkable power to bear on a wide range of problems that have proven critical to our national interest. Over the coming decade, however, designers of the world's fastest computer systems, both in our country and in others, will confront unprecedented technical obstacles that will not be overcome without fundamental conceptual changes. As the amount of computing power embedded within a single chip continues to increase, progressively shrinking circuit elements will begin to approach atomic-scale dimensions. Data transfers between one chip and another will encounter inherent limitations imposed by the speed of light. Conventional chips will generate so much heat that even powerful, miniaturized refrigerators will be unable to keep them from quickly frying.

Yet we simply cannot afford to accept these limitations, or to ride in the tailwind of other countries that are determined to overcome them. By strengthening our national defense and sifting through an astronomical amount of intelligence data to "connect the dots," HPC helps ensure the physical safety of America's population. By enhancing the competitiveness of the products and services we offer within the global economy, HPC helps provide high-wage jobs for American workers. By enabling transformative advances in science and technology, HPC helps maintain our historical leadership for future generations of Americans.

Our long-term competitiveness within the field of high-performance computing will require a substantial and sustained investment in basic research involving a number of subdisciplines of computer science and engineering, along with a willingness to underwrite the inescapable cost of technological failures in order to achieve major, game-changing advances. Will tomorrow's high-performance systems incorporate optical, molecular, or quantum computing? How will we program and manage systems whose speed is derived from millions of tiny computers? How will we deal with databases incorporating the equivalent of 10 billion bytes of data for every person on earth?

We do not know. We must figure it out.

The evolving goals of HPC. Although HPC has played an important role in a number of fields, the use of supercomputers is most commonly thought of in connection with scientific and engineering applications. Over the past several decades, simulation has become the third pillar of science, complementing

Science: Ensuring America's Competitiveness.

theory and experiment in cases where the theory is unknown (or too difficult to solve analytically), or where experiments are too difficult (or even impossible), too costly, or too dangerous to perform. Some of the most important breakthroughs achieved through simulation and other forms of scientific computing have been enabled by America's historical leadership in the development and deployment of HPC technology. These advances have in turn played a central role in addressing our Nation's needs and priorities.

In today's environment, however, the notion of "high performance" must assume a broader meaning, encompassing not only the traditional metric of floating-point operations per second (FLOPS), but also the ability to efficiently manipulate vast and rapidly increasing quantities of both numerical and non-numerical data, to handle problems requiring real-time response, and to accelerate many applications that were either non-existent or far less important at the time of NITRD's creation under the High Performance Computing Act of 1991.

Competition within the international community to develop what are typically described as the world's most powerful supercomputers has been based to a large extent on a single metric that, while relevant to *certain* HPC applications, increasingly fails to reflect the broad range of capabilities our Nation needs in the area of high performance computing. This metric, which measures the number of FLOPS executed on a single, classical benchmark involving the solution of a dense system of linear equations, is used to compile the rankings shown on the widely followed Top500[60] list. As of June 2010, three of the top ten positions on this list were occupied by machines in the hands of foreign countries, and in October 2010, China announced the completion of a machine that has now claimed first place. But the goal of our investment in HPC should be to solve computational problems that address our current national priorities, and this one-dimensional benchmark measures only one of the capabilities relevant to those priorities. For data-intensive applications involving the rapid execution of graph operations, for example, the forthcoming Graph500[61] benchmark will be more relevant, while the most salient performance-related characteristics associated with other important applications may be difficult to reduce to *any* single figure of merit.

While it would be imprudent to allow ourselves to fall significantly behind our peers with respect to scientific performance benchmarks that have demonstrable practical significance, a single-minded focus on maintaining clear superiority in terms of FLOPS count is probably not in our national interest. Engaging in such an "arms race" could be very costly, and could divert resources away from basic research aimed at developing the fundamentally new approaches to HPC that could ultimately allow us to "leapfrog" other nations, maintaining the position of unrivaled leadership that America has historically enjoyed in high performance computing.

The evolving dimensions of HPC. If Top500 rankings can no longer be viewed as a definitive measure of a country's high performance computing capabilities, what goals should our nation be setting for fundamental research in HPC systems, and what criteria should be used in allocating funding for such research? Given the natural inclination to quantify the relative performance of competitors in any race, there is a temptation to replace the traditional FLOPS-based metric with another fixed, purely quantita-

60. http://www.top500.org/
61. http://www.graph500.org/

tive metric (or perhaps two or three such metrics) that policymakers can use on an ongoing basis to rank America's competitive position in HPC relative to those of other countries. This approach, however, is subject to several pitfalls that could both impair our ability to maintain our historical leadership in the field of high-performance computing and increase the level of expenditures required to even remain competitive.

First, it is no longer feasible to capture what is important about high-performance computing as a whole using one (or even a small number of) fixed, quantitative metrics, as a result of:

- the progressive broadening of our nation's requirements in the area of high-performance computing;
- the consequent "splintering" of the set of computational tasks required to satisfy these requirements;
- a wide range of substantial advances in the various technologies available to perform such computational tasks;
- significant changes in the "bottlenecks" and "rate-limiting steps" that constrain many high-performance applications as a result of different rates of improvement in different technological parameters.

In addition, transformative advances of the sort that may allow us to leapfrog the competition associated with current-generation HPC systems may well involve unanticipated breakthroughs, in unanticipated dimensions, whose relevance to important national objectives may not be evident in advance. As in many areas of research in science and engineering, such approaches may in some cases involve redefining the problem itself, discovering methods that achieve *different* objectives which ultimately turn out to be instrumental in addressing national priorities related to national security and/or economic competitiveness. In order to capture the potential benefits of such game-changing advances, a substantial portion of our nation's HPC research portfolio should thus be allocated to high-risk, high-return research with very broadly and flexibly defined objectives. As in many fields of research, the objectives for individual projects falling within this component of our research portfolio should in many cases be proposed by the *investigator* as part of his or her application for funding, then evaluated by the program manager and/or peer reviewers on a case-by-case basis.

This is not to say that well-defined goals and metrics have no place in the nation's HPC research portfolio. On the contrary, such measures can often serve as powerful drivers not only for incremental progress, but in some cases, for fundamental advances as well. This type of research may be particularly appropriate in the context of:

- Research programs initiated by specific agencies to address one or more specific applications relevant to mission-related national needs;
- Focused research initiatives that aim to develop particular types of HPC systems or to increase performance along specific dimensions, including (though by no means limited to)
    - time-to-completion for specific computationally-intensive applications;
    - turnaround time for short runs;

- inter-processor communication latency;
- performance metrics related to data analytics;
- performance for graph operations on large databases;
- inter-processor communication bandwidth;
- memory bandwidth at various levels within the system;
- power consumption;
- cooling capacity;
- mean time between failures;
- uptime percentage;
- real-time response;
- time required for programming;
- software reliability.

The evolving ecology of HPC. It is important not to equate "computational science" with "supercomputing." Advances in science and engineering are enabled by computational resources at all levels of what is often referred to as the "Branscomb pyramid,"[62] which extends from individual desktop machines through small clusters to the largest supercomputers. The nature of these levels has been evolving as a result of important changes in commercially available components and in the overall computational landscape. For example:

- Even desktop computers are now available with powerful multicore processor chips that provide parallel computing;
- Graphics processing units (GPUs) originally designed for gaming applications are now being used to accelerate many scientific computing applications;
- The enormous web service datacenters operated by companies such as Amazon.com, Google, and Microsoft offer higher aggregate performance, by some important measures, than the fastest scientific supercomputers.

The availability of such computational resources must be taken into consideration when weighing the Nation's investments in computational infrastructure at all levels of the Branscomb pyramid, and in the portfolio of research activities that will be required to fully exploit these various resources.

The continuing challenges of HPC. Over the past few decades, HPC systems have benefited from rapid and sustained advances in the performance of commodity computing hardware, including processors, memory, and data storage systems. In particular, the processing units of today's fastest supercomputers, unlike those of their early predecessors, are often based in large part on commercially available processor chips (typically, multicore chips designed for server applications). GPUs are also now being incorporated

62. NSF Blue Ribbon Panel on High Performance Computing. (August 1993). *From Desktop to Teraflop: Exploiting the U.S. Lead in High Performance Computing.*

in some high performance systems, exploiting their high on-chip arithmetic density (particularly relative to power consumption) to accelerate the computational power of individual processing nodes.

Both the architecture and the capabilities of a typical massively parallel, high-end supercomputer, however, differ in important ways from those of a conventional cluster based on the same commodity processing components. One of the most important distinctions relates to the way in which those processing nodes are interconnected.

Many critically important problems cannot be decomposed into subproblems that can be executed more-or-less independently on a large number of processors without a great deal of inter-processor communication to exchange data. Depending on the intrinsic communication requirements of the algorithm, on the number of processing elements in the system on which it is to be executed, and on various other parameters, the time required for computation may in some cases be dominated by the time required for communication, rendering processor speed largely irrelevant. Such communication bottlenecks may be attributable to:

- *bandwidth constraints*, which are associated with the *amount* of data that must be communicated among the various processors; and/or
- *latency constraints*, which originate from the fixed delay incurred when *initiating* a (direct or indirect) data transfer from one node to another.

By contrast with conventional clusters of modest to moderate size, massively parallel supercomputers often contain specialized interconnection networks designed specifically to achieve high bandwidth and low latency. Such networks are often relatively costly, but in many cases are capable of supporting a far larger number of processing elements before overall system performance "tops out," allowing scientists and engineers to tackle important problems that would otherwise lie far beyond the reach of computational methods. While the fastest machines in some cases differ in other ways from lower-performance clusters, high-bandwidth, low-latency communication is now a defining characteristic of HPC.

Progress in Algorithms Beats Moore's Law

Everyone knows Moore's Law – a prediction made in 1965 by Intel co-founder Gordon Moore that the density of transistors in integrated circuits would continue to double every 1 to 2 years.

Fewer people appreciate the extraordinary innovation that is needed to translate increased transistor density into improved system performance. This effort requires new approaches to integrated circuit design, and new supporting design tools, that allow the design of integrated circuits with hundreds of millions or even billions of transistors, compared to the tens of thousands that were the norm 30 years ago. It requires new processor architectures that take advantage of these transistors, and new system architectures that take advantage of these processors. It requires new approaches for the system software, programming languages, and applications that run on top of this hardware. All of this is the work of computer scientists and computer engineers.

Even more remarkable – and even less widely understood – is that in many areas, *performance gains due to improvements in algorithms have vastly exceeded even the dramatic performance gains due to increased processor speed*.

The algorithms that we use today for speech recognition, for natural language translation, for chess playing, for logistics planning, have evolved remarkably in the past decade. It's difficult to quantify the improvement, though, because it is as much in the realm of quality as of execution time.

In the field of numerical algorithms, however, the improvement can be quantified. Here is just one example, provided by Professor Martin Grötschel of Konrad-Zuse-Zentrum für Informationstechnik Berlin. Grötschel, an expert in optimization, observes that a benchmark production planning model solved using linear programming would have taken 82 years to solve in 1988, using the computers and the linear programming algorithms of the day. Fifteen years later – in 2003 – this same model could be solved in roughly 1 minute, an improvement by a factor of roughly 43 million. Of this, a factor of roughly 1,000 was due to increased processor speed, whereas a factor of roughly 43,000 was due to improvements in algorithms! Grötschel also cites an algorithmic improvement of roughly 30,000 for mixed integer programming between 1991 and 2008.

The design and analysis of algorithms, and the study of the inherent computational complexity of problems, are fundamental subfields of computer science.

As noted above, powerful new tools are now available to scientists and engineers at *all* levels of the Branscomb pyramid, and large scale, high performance computing facilities should by no means be regarded as the *only* infrastructural resources relevant to scientific computing. At the same time, it is clear that high-end computer systems, which provide a distinct and irreplaceable set of capabilities, will continue to be of critical importance to the Nation's needs and priorities for many years to come. Ongoing and anticipated changes in multiple dimensions make it essential to examine the ways in which the nature and focus of HPC and HPC research will have to evolve in order to most effectively address these needs.

Changes and challenges. The next generation of HPC systems will encounter a number of formidable technical challenges. Failure to address these challenges will seriously handicap our Nation's ability to tackle many significant problems and opportunities. Ironically, many of these challenges arise from the enormous progress we have made over the years in developing and applying new information technologies. From a bottom-up perspective, integrated circuits (ICs) are fast approaching physical limits that will alter the rhythm of technological advances to which we have become accustomed over a period of decades. From a top-down viewpoint, we now find ourselves within striking distance of computationally enabled scientific and engineering breakthroughs that could have a major impact on our lives. We are also facing an embarrassment of riches in the form of a flow of data so overwhelming as to have been unimaginable 20 years ago, the effective utilization of which will require fundamental changes in the way we think about high performance computing.

Beginning at the bottom, we note that the number of processor cores and the amount of logic and memory on each processor chip continue to increase significantly from one generation of technology to the next. Such chips are thus progressively drawing more power, dissipating more heat, and requiring more cooling to function. This trend has already brought back water cooling for some HPC systems, and even the most effective currently employed cooling techniques will soon prove inadequate to fully exploit the computational potential associated with increasing circuit density. In addition, the number of physical connections that each chip can make with the outside world has been increasing more slowly than the number of devices on a chip, thus limiting the rate at which data can be transferred among processors, or between processors and memory. As the processing power of each chip continues to grow, the bandwidth and latency limitations of the interprocessor communication network are also becoming increasingly significant bottlenecks. Progressive increases in the degree of parallelism exploited in HPC systems also pose challenges for the detection and handling of errors.

Many of these limitations may be addressable through the discovery of new parallel algorithms designed to take advantage of the potential power of continuing advances in the underlying technology while avoiding some of the limitations and bottlenecks they expose. The study of parallel algorithms, however, presents special challenges, and has not yet reached the same maturity as its sequential counterpart. The utilization of HPC resources has also been hampered by a relative dearth of systems software and of tools for monitoring and optimizing performance.

Another significant challenge is the design of high performance architectures and algorithms for non-traditional forms of high performance computing, such as systems designed for the efficient performance of various data-intensive computational tasks, including those involving non-numerical data. Such applications pose a unique set of problems and design tradeoffs, and will become progressively more important over time.

Research priorities. It is of course important that we not jeopardize America's national security, economic competitiveness, or other vital interests by falling behind in the near-term development and deployment of HPC systems that address critical and continuing needs. It is equally important, however, that we balance our investments in America's present and future requirements for high performance computing, and that we not allow the procurement of current-generation machines to "crowd out" the fundamental research in computer science and engineering that will be required to develop *next-generation* HPC technologies. To lay the groundwork for such next-generation systems, we need to conduct basic research in hardware, in hardware/software systems, in algorithms, and in both systems software and applications software.

Hardware research must include novel IC designs incorporating a large number of on-chip processor cores; novel intra-chip communication architectures; system-level interconnection networks with high link bandwidth, high bisection bandwidth, and low latency; and IC and chip packaging technologies offering high input/output (I/O) bandwidth.

Advances that combine hardware and software considerations are needed for the design of reliable massively parallel computer systems; for design aspects of HPC systems in which hardware design choices and the ability to write software that takes full advantage of the hardware are interdependent; and for "special purpose" machines designed to achieve high performance on specific classes of algorithms, applications, and data structures.

New methods are needed for both hardware and software design that reduce power needs. Both hardware and software advances are needed for data-intensive computing, including non-numerical applications such as graph operations. In particular, research is needed in algorithms, software systems, and architectures capable of handling extreme-scale data systems (as much as 10^{21} bytes) and data analytics.

Other important software-related research topics include latency-tolerant architectures and algorithms; programming models and languages for massively parallel machines; systems software for massively parallel systems, including operating systems, file systems and data stores; performance and correctness debuggers; system management tools; development environments; and tools and techniques for modeling and tuning the performance of large-scale systems (including computational, data, and networking resources) to optimize current applications and guide the development of new hardware and software architectures.

Preparing for the demise of current technologies. Even if significant progress can be made in addressing such problems as chip I/O, system-level communication, and heat dissipation in HPC systems, fundamental physical limitations will ultimately be encountered as traditional integrated circuit technologies begin to approach atomic-scale feature widths. In the absence of basic research into alternative technologies, improvements in the performance of HPC systems cannot be expected to continue indefinitely at their current pace. Such research will by its nature be risky, and it is difficult to predict in advance which approaches might prove fruitful. A non-exhaustive list of potentially transformative technologies that could be included within a diversified research portfolio, however, might include three-dimensional integration, carbon nanotubes and graphene nanoribbons, optical computing and interconnect, memristor-based technologies, quantum computing, molecular computing and storage, and reliable technologies based on unreliable devices.

Extending the reach of HPC. HPC has transformed many areas of science and engineering, but its potential has yet to be fully realized within certain other application areas, and within certain sectors, types of organizations, and categories of users. To address the obstacles that have limited the adoption and effective utilization of HPC, steps should be taken to:

- Eliminate technical barriers to moving applications from laptop to cloud to high-end systems (programming models, system software, algorithms);

- Ensure that there is an adequate high-speed networking infrastructure to permit the use of HPC systems from anywhere in the United States;
- Support research on new architectures and software that lower the capital, operating, programming, maintenance, and administrative costs incurred by organizations engaging in high performance computing;
- Foster the development of parallel versions of both open-source and commercial application codes that are currently in widespread use on sequential machines;
- Implement mechanisms to make expertise on HPC hardware and software available to smaller public- and private-sector organizations that do not presently have in-house expertise in these areas.

7. Recommendations: Investments in the NIT Research Frontiers

Advances in NIT rest on a broad and deep foundation of more than 60 years of fundamental research. That foundation, which is divided for convenience into a collection of core areas, continues to evolve as changes in technologies and new uses of NIT stimulate new breakthroughs and deeper understanding. In order to make progress in the uses of NIT, continuing research in core areas is essential.

The following summary recommendation, which appears in the Executive Report, highlights the most important elements of the more detailed recommendations that appear later in this section. We note again the importance of high risk/high reward research with the potential to move these areas in unanticipated directions.

Recommendation: The Federal Government must increase investment in those fundamental NIT research frontiers that will accelerate progress across a broad range of priorities. Among such investments:

- NSF and DARPA, with the participation of other relevant agencies, should invest in a broad, multi-agency research program on the fundamentals of privacy protection and protected disclosure of confidential data. Privacy and confidentiality concerns arise in virtually all uses of NIT.
- NSF, DARPA, and HHS should create a collaborative research program that augments the study of individual human-computer interaction with a comprehensive investigation to understand and advance human-machine and social collaboration and problem-solving in a networked, on-line environment where large numbers of people participate in common activities. Understanding such collective human-NIT interactions is increasingly important for defense, for health, and for the activities of daily life.
- NSF should expand its support for fundamental research in data collection, storage, management, and automated large-scale analysis based on modeling and machine learning. Our ever-increasing use of computers, sensors, and other digital devices is generating huge amounts of digital data, making it a pervasive NIT-enabled asset. In collaboration with NIT researchers, every agency should support research, to apply the best known methods and to develop new approaches and new techniques to address data-rich problems that arise in its mission domain. Agencies should ensure access to and retention of critical community research data collections.
- NSF and DARPA, in collaboration with those agencies tackling problems whose solution entails instrumenting the physical world – including the Environmental Protection Agency (EPA), DoE, DoT, other parts of DoD, NIH, the Department of Agriculture (USDA), and the National Oceanic and Atmospheric Administration (NOAA) – should increase research in advanced domain-specific sensors, integration of NIT into physical systems, and innovative robotics in order to enhance NIT-enabled interaction with the physical world.

The recommendations that follow describe some areas in which *additional* investment is needed in order to realize the advances that our national priorities require, and other areas in which attention must be directed to particular challenges. These investments should *supplement* ongoing research in more established areas such as algorithms and computer graphics that are not called out explicitly in this report.

Recommendation: New investments must not supplant continued investment in important core areas in which government-funded research is advancing. Continued attention must also be given to sustained high-quality shared research infrastructure, including new forms of infrastructure to support new research areas and paradigms.

Privacy and Confidentiality

Preserving personal privacy is a critical need that pervades NIT. Our democratic society puts a high value of protection of personal privacy, while the protection of corporate information is a key element of competitiveness. Controls on data disclosure are essential to protecting the safety of individuals and of our Nation. All NITRD agencies should evaluate how considerations of privacy and confidentiality will affect the deployment and use of the technologies emerging from their NIT R&D efforts. In some cases the issues are not well understood. The Federal Government needs to support R&D to better understand, address, and ameliorate the privacy and confidentiality issues that are identified.

Recommendation: NSF and DARPA, with the participation of other relevant agencies, should invest in a broad multi-agency research program on the fundamentals of privacy protection and protected disclosure of confidential data. The program should address at least the following important issues:

- developing methods that allow agents – that is, individuals or software acting in well-defined and suitably constrained roles – to perform analytics on large datasets while preserving privacy and confidentiality;
- creating and investigating formal models of privacy that combine concepts from statistics and computer science; these models should be characterized in terms of what guarantees they provide, what adversaries they can withstand, and what forms of sharing they permit;
- understanding the consequences for privacy and confidentiality of technology trends such as large-scale data gathering, analytics, correlations of multiple sources, machine learning, and ubiquitous sensors, as well as of protective regimes such as cybersecurity;
- devising methods that give individuals knowledge of what data about them is held and appropriate control over the use of that data;
- exploring the privacy-preserving design of human-centered systems that create and use information about people, in financial, medical, demographic, and residential domains;
- creating ways to educate users about, and protect users against, actions they might take that inadvertently compromise privacy;
- using the fruits of research into privacy and confidentiality protection to enable privacy- and confidentiality-related policies to be stated in application-relevant terms, and enforced in

application-specific ways; one example is coordination with agencies including NIH, HHS, the Department of Commerce, and the Department of Justice (DoJ) to ensure that policies and regulations concerning medical records, census data, and other socially important datasets are based on sound scientific principles.

The Ubiquitous Role of Privacy

Online privacy is already a significant issue for many Americans – witness the debate over privacy on social networks and the rise of identity theft. Technology trends can only raise the stakes.

Privacy challenges clearly arise for electronic medical records, but in fact they come up in all of the national priority areas. The smart grid will save energy by instrumenting, analyzing, and optimizing power usage within a home – but these actions convey information about activities within the home. Smart transport will reduce congestion and save energy by optimizing the movement of individual vehicles and tracking people's travel needs – but these details convey information about personal activities. Personalized education will use data about a student's education history and progress to offer the best instruction – but this, too, is sensitive information.

We cannot afford to forgo the benefits of NIT in addressing national priorities. Rather, we need fundamental advances in NIT – a practical science of privacy protection – to give us the tools to reconcile privacy with progress.

Privacy challenges arise whenever we want to allow access to information for some purposes but not for others. A public health researcher may want to search for subtle trends hidden in a large collection of patients' health records, but we shouldn't allow the researcher to learn the details of any specific patient's record. Ideally, we could scrub or anonymize data before giving it to the researcher, but doing this safely, without scrubbing away the very trends the researcher is seeking, is difficult in theory and risky in practice.

Privacy challenges are complicated by problems of inference and side information. Revealing a fact implicitly reveals everything that can be inferred from it. Revealing a patient's prescriptions, for example, could implicitly reveal the patient's medical conditions. Inferred facts lead to further inferences; one seemingly innocuous fact can trigger a cascade of inferences. It is difficult to catalog and control all methods of drawing inferences. Worse yet, an analyst can combine revealed facts with all of the side information available from other sources, so that our concerns about the possible extent of inferences must take into account all the side information that might be available. These are difficult problems, but there is hope that fundamental NIT research can address them.

Understanding how to reconcile privacy with the application of NIT to national goals, both in general and with respect to specific goals, will improve Americans' privacy while enabling ever more beneficial application of NIT.

NIT and People

The modes and the ease with which people interact with computers have improved as richer forms of interaction and better understanding of human capabilities have informed the design of interactive systems. The advent of widely available networking and the introduction of digital consumer products have further empowered people. We are now experiencing another spurt of growth – into the realms of social computing and media, NIT-enabled social science, and collective interaction.

Recommendation: NSF, DARPA, and NIH should create a research program that augments the study of individual human-computer interaction with a comprehensive investigation to understand and advance human-machine collaboration and problem-solving in a networked, online environment. The program should:

- create a science of social computing that, for example, gives insight into how to organize human contributions, how to incentivize participants, and how to design generic social-computing frameworks that could be used by different organizations for diverse purposes;
- foster research that pushes the field beyond the current examples of crowd-sourcing;
- encourage theoretical, algorithmic and engineering foundations that guide the design of peer-production systems (in which large groups of individuals, sometimes tens or even hundreds of thousands, collaborate online) for a wide variety of tasks;
- design novel mission-specific uses of collaborative computing;
- create shared privacy-preserving research platforms to enable researchers in computational social science to share and exchange experimental designs, behavioral experimental data, and human subject panels and subjects. For example, a promising application area for such experimental research is the study of human decision-making regarding security and privacy issues, so as to inform technology and design considerations in those areas.

NIT and the Physical World

The 2007 PCAST assessment[63] of the NITRD Program spurred a new emphasis on Cyber-Physical Systems. New and valuable research activities were launched, and important programs and collaborations were fostered involving NSF, DARPA, and industry. We recommend expanding and deepening these efforts in three particular areas: sensor development, robotics, and open architectures.

Recommendation: NSF, in collaboration with those agencies tackling problems whose solutions entail instrumenting the physical world – including EPA, DoE, DoT, DoD, NIH, USDA, and NOAA – should conduct research to design, fabricate, and test sensors that are problem-domain specific and that are cheaper, smaller, better packaged, lower powered, and more autonomous than those available today. These advances would lead to new applications and new markets that would be sustained in the long term through traditional commercial activities.

---

63. President's Council of Advisors on Science and Technology. (August 2007). *Leadership Under Challenge: Information Technology R&D in a Competitive World.* An Assessment of the Federal Networking and Information Technology R&D Program. http://www.nitrd.gov/pcast/reports/PCAST-NIT-FINAL.pdf

Recommendation: DARPA, NSF, NIH, and DoE should continue to sponsor research on large-scale modular robotics and computer vision and should collaborate to increase the rate of innovation, usability, and scalability of autonomous actuation in environmental, medical, manufacturing and defense contexts.

Recommendation: Interoperability is essential for enabling NIT to be broadly embedded in the physical world. Early adoption of both will create an open and fair playing field for Federal agencies to see competition among their suppliers, and for commercial products built by different companies to operate compatibly. If the United States creates the open architectures and standards, U.S. industry will gain an early advantage. NIST, in consultation with NSF, should lead an interagency effort to build consensus and fund reference implementations.

Large-Scale Data Management and Analysis

Virtually every Federal agency has opportunities to capitalize on the analysis of large volumes of data to further its mission. The data might be human-generated in electronic form, it might be obtained from sensors or other observational tools, it might be derived from computer simulations, or it might be the by-product of other kinds of NIT applications. As data becomes increasingly abundant, fundamental research to advance our expertise in collecting, analyzing, understanding and using that data becomes increasingly urgent. In addition to the recommendations below, controlled data sharing and data privacy are critical issues; these topics are addressed in the "Privacy and Confidentiality" recommendations above.

Recommendation: NSF should expand its support for fundamental research in data collection, storage, management, and analysis. Its programs should address topics such as:

- contextual metadata for data obtained from sensors in the physical world;
- information derived from cross-correlation;
- information fusion algorithms that combine data from diverse sources, differing scales, differing types and metadata;
- long-term preservation;
- data provenance and integrity;
- inference from incomplete and uncertain data;
- deep analysis of the information contained in the data;
- abstraction, summarization, and visualization of complex data and information.

Recommendation: Every agency should engage in R&D to apply the best existing methods and to develop new approaches and new techniques to address data-rich problems in its mission domain. Collaboration between NIT researchers and domain experts is essential to this work.

Recommendation: Under the leadership of NIST, processes and policies should be established for the NITRD agencies to publish real-world primary data sources such as detailed logs of events or sensor readings, in order to facilitate research into techniques for mining and abstracting those sorts of data.

The private sector should be encouraged to participate as well. Examples include a "click stream" of users trying to find data on a government web site, detailed readings of water flows and levels, or video surveillance of busy highways. Care must be taken to release data in a privacy-preserving way based on sound scientific principles.

Trustworthy Systems and Cybersecurity

The trustworthiness and security of NIT systems are characteristics that transcend the uses to which these systems are put. All systems must be secure against unintended behaviors, against unauthorized access, and against threats to their availability and integrity.

Recommendation: NSF and DARPA should aggressively accelerate their initiatives to fund and coordinate fundamental research to find more effective ways to build trustworthy systems and to assure cybersecurity. These initiatives should include programs to:

- Advance the art and the practice of designing and implementing trustworthy systems, which act only as users expect them to act even in the face of failures. Develop methods to analyze the trustworthiness of designs and implementations. The research should focus explicitly on systems critical to society;
- Develop fundamentally new "clean slate" designs and outside-the-box approaches that will provide a new basis for assuring the security of systems and data. These new approaches should provide a basis for relating classes of attacks, specific (possibly new) defense designs, and security policies, so that more careful reasoning and analysis can be performed;
- Develop methods suitable for implementing a small, rigorously isolated set of very basic capabilities that can be relied upon with a high degree of confidence to provide truly essential NIT-based services in the event of, for example, a catastrophically damaging cyber-attack.

Scalable Systems and Networking

Research on the design and implementation of scalable systems has a long history in computer science. Both NSF and DARPA have supported relevant research for many years. This is a critical core activity for computer science research, and continued investment is required to keep up with changing application needs and technology capabilities. Continued progress will draw on expertise ranging across many different system layers – an effort that only collaborative, multidisciplinary teams can provide. The overall system environment today consists of many layers under the control of different companies, governments, and standards organizations. Support for the fundamental advances the Nation needs demands higher levels of coordination between these entities than now exists.

Recommendation: NSF, DARPA, and other organizations should continue their leadership and funding of core research into scalable systems in order to ensure that networked systems will adapt to the ever-changing needs of applications, to the capabilities engendered by new technology, and to evolving needs for security and privacy.

Recommendation: To foster an innovative ecosystem in NIT, government agencies concerned with networked systems operations, including DoD, NSF, and the Federal Communications Commission (FCC), must continue to encourage and invest in open systems development. They must coordinate with standards organizations (IETF, W3C) and the NIT industry to ensure that standards pertaining to the different interfaces within and among the layers of the networked systems environment are defined and kept up to date.

Recommendation: In the area of wireless systems, NSF, FCC, and the National Telecommunications and Information Administration (NTIA) should partner to create, sustain, and promote the use of a nationwide infrastructure for spectrum monitoring that cuts across commercial, public safety and DoD applications. NSF, DHS, and NTIA should partner to create programs that promote innovative use of public safety frequencies. NSF, DHS, and DARPA should jointly articulate the synergies among their individual needs and programs in wireless spectrum management.

Software Creation and Evolution

Over the past several decades, software research has made major advances in our ability to create increasingly large, complex, and critical software systems. The continuing emergence of new challenges requires a steady stream of new advances. Investment in software research must be sustained.

Recommendation: NSF, DARPA, and other organizations that need software tailored to their mission requirements should continue their leadership and funding of core research in methods to improve the design, development, modification, and maintenance of all varieties of software. That research should address language design, tools, analysis methods, methods for collaborative design and development, and techniques that provide security and robustness. Attention must be given to system design and programming for scalability, paradigms for parallelism at multiple levels of granularity, software for heterogeneous systems involving interaction with the physical world, and software for systems that incorporate human interaction. Long term evaluative research is required to determine which tools and techniques yield sustainable improvement in software creation.

High Performance Computing

HPC is increasingly important for research in many areas of science and engineering. It is essential to national security, and is a major tool in addressing other important national priorities. In order to maintain its historical leadership in the design and effective utilization of HPC, the United States must anticipate and adapt to the broadening of its high-end computational needs and to changes in the underlying technologies available to address them. The primary focus must be on advances that will address important national needs, and not on the relative ranking of each country's fastest supercomputer on the Top500[64] list.

Recommendation: NSF, DARPA, and DoE should invest in a coordinated program of basic research on architectures, algorithms and software for next-generation HPC systems. Such research should not be limited to the acceleration of traditional applications, but should include work on systems capable of (a)

64. http://www.top500.org/

efficiently analyzing vast quantities of both numerical and non-numerical data, (b) handling problems requiring real-time response, and (c) accelerating new applications. Specific areas of investigation should include:

- Novel system architectures for massively parallel computing
- High-bandwidth, low-latency processor interconnection networks
- Reliability and fault-tolerance in massively parallel computer systems
- Hardware and software design techniques for the dramatic reduction of power consumption
- Data-intensive computing, including non-numerical applications
- Programming models and languages for massively parallel machines
- Systems software for massively parallel systems
- Improved approaches for system management

In addition to designing next-generation systems, significant effort must be devoted to R&D focused on extracting the greatest possible scientific benefit from current leading-edge systems.

8. Technological and Human Resource Requirements

This section addresses two forms of infrastructure: technological and human.

Shared NIT infrastructure – be it computational resources, communication networks, community databases (e.g. PubMed and the Protein Data Bank), or collaboration tools – has become essential to research in virtually all fields. Equally important are new forms of infrastructure that support new research areas and paradigms. It is appropriate that this infrastructure be included under the NITRD Program. However, it is important to distinguish between NIT infrastructure that supports NIT research and NIT infrastructure that supports research in other fields. NIT infrastructure that supports NIT research is a crucial component of NIT R&D – it is essential to achieving advancements in NIT, which (among many other benefits) will yield the next generation of NIT infrastructure for all fields. In contrast, NIT infrastructure that supports research in other fields is a crucial component of R&D in those fields, *but it is not NIT R&D*. The importance of NIT to the Nation's future requires that investment in NIT R&D is accurately known and distinguished from investments in NIT infrastructure that serve other purposes.

The ever-expanding role of NIT in our society creates an ever-increasing demand for NIT professionals. All indicators – all historical data and all projections – argue that NIT is the dominant factor in America's S&T employment, and that the gap between the demand for NIT talent and the supply of that talent is large and will continue to be so. If we are to bridge that gap, increasing the number of graduates in NIT fields at all degree levels must be a national priority. Along with the need for more NIT professionals, there is also a rapidly growing demand for individuals who can utilize NIT flexibly and creatively and who can apply NIT "modes of thought" in a wide variety of endeavors. Meeting each of these needs will require fundamental changes in K-12 STEM education.

8.1 Hardware, Software, and Data Infrastructure

Finding: Shared NIT infrastructure – be it computational resources, communication networks, community databases, or collaboration tools – has become essential to research in virtually all fields.

The ability to effectively conduct NIT-enabled research and education presupposes reliable access to data and information, stable computational platforms to execute applications, continuous network access, and dependable software – i.e., a supporting information technology *infrastructure*.

Infrastructure is the foundation upon which we operate. Just as water and electrical power enable us to function in the physical world, reliable and stable data systems, computers, and networks enable us to function in the digital world.

The best infrastructure is utterly unremarkable – it supports other efforts without distracting from them. We may not notice that our lights stay on continuously or that our Google home page is there every time we call it up, but the high reliability of underlying infrastructure is an essential foundation that enables us to do other things.

The infrastructure needed for most research purposes is of small to medium scale, relatively affordable, and distributed among users. However, some research inquiry requires large, necessarily shared, infrastructure. For example, PubMed (an online biomedical research library run by NIH) and the Protein Data Bank (an online archive of protein and other molecular structures, run by a university consortium) support millions of users performing research and practice in biology and the life sciences. Both are digital infrastructure – community data bases supported by user-friendly software, reliable storage, efficient search algorithms, and stable servers. As another example, the national supercomputer Centers support groundbreaking research in a wide range of science and engineering disciplines, from molecular biology to the evolution of the universe – research that requires large-scale computation and storage, accessible over a reliable, high-speed network.

Because so many users depend on NIT infrastructure, and because infrastructure typically becomes more valuable to users over time as they become more proficient with it, predictable access, resource reliability, and cost-effectiveness for users are important metrics of infrastructure success. For this reason, funding for NIT infrastructure should run on longer time-frames than it does now, so as to maximize leverage of the resource by the user community. In addition, funding should support evolutionary pathways to next-generation infrastructure to enable users to maximize research productivity. Continuity is particularly important in the software arena, where broadly used software systems and tools must evolve as next-generation hardware platforms are deployed. This model is followed in many sciences and by many agencies, where powerful, shared resources (such as accelerators in physics or telescopes in astronomy) are made available to their communities and funded and staffed to provide long-term, reliable, and cost-effective infrastructure.

Good infrastructure doesn't just happen. It requires R&D to ensure that the large-scale NIT infrastructure meets the Nation's needs. For example, standards are needed to ensure that data remains as accessible in twenty years as it is now. To meet the demonstrated need for extreme scale systems, innovative approaches must be developed, and these are likely to require a long-term commitment to NIT R&D. New software approaches are needed to support dramatic increases in scale. This is not just an engineering problem, solvable using known concepts, but requires the discovery and development of new concepts and tools with which to build robust, reliable systems. Only a relatively modest proportion of NIT researchers working to advance large-scale infrastructure actually need large-scale infrastructure for their research. In some cases they can work with large-scale infrastructure that is shared by researchers in other disciplines.

As science and technology, as well as research and practice, become increasingly data-driven, the importance of shared infrastructure that supports data access, management, use, and preservation grows ever greater. The digital data used by research communities is both expensive to acquire (generating the information stored in the Protein Data Bank cost over $80 billion in research funding[65]), and difficult to

65. Testimony of Helen Berman, Director of the Protein Data Bank, to the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. The interim report of the Blue Ribbon Task Force: http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf

replace (e.g., the lost NASA high-resolution video of the first moon walk[66]). Research projects on smaller scales frequently produce valuable data, which in many cases should be preserved for periods far in excess of typical research grant duration. If we are to compete in the global science and technology arena, we must put into place viable and economically sustainable models for funding NIT *infrastructure*, distinct from the models used for funding NIT *research*.

The shared infrastructure we need for America's competitiveness is not the province of the Federal Government alone. Today, the commercial sector is partnering with NSF to provide scalable cloud platforms, data centers, and high performance computers to support research. University libraries and public archives, seeking to reinvent themselves in the digital age, are beginning to discuss a role as stewards of the vast and growing deluge of digital research data. Although the Federal Government must bear the ultimate responsibility for ensuring the preservation of data from Federally funded research, as well as massive Federal "collections of collections" such as Data.gov, it cannot afford to build and maintain the necessary capacity to store all critical data. Partnership with university libraries and other repositories may therefore be a promising strategy.

It is important to distinguish between NIT infrastructure that supports NIT research and NIT infrastructure that supports research in other fields. NIT infrastructure that supports NIT research is a crucial component of NIT R&D – it is essential to achieving advancements in NIT, which (among many other benefits) will yield the next generation of NIT infrastructure for all fields. NIT infrastructure that supports research in other fields is a crucial component of R&D in those fields, *but it is not NIT R&D*. PubMed and the Protein Data Bank are examples: they are essential NIT investments in biomedical R&D, but they are not NIT R&D. It is appropriate to include all of these NIT infrastructure investments as part of NITRD, but they must be properly distinguished. The importance of NIT to the Nation's future requires that we have an accurate estimate of our actual investment in NIT R&D.

8.2 Education and Human Resources

Finding: All indicators – all historical data, and all projections – argue that NIT is the dominant factor in America's science and technology employment, and that the gap between the demand for NIT talent and the supply of that talent is and will remain large. Increasing the number of graduates in NIT fields at all degree levels must be a national priority. Fundamental changes in K-12 education are needed to address this shortage.

NIT workforce: Robust demand, limited supply. NIT workforce supply and demand has been the subject of a significant amount of thoughtful analysis and of a far greater amount of editorializing in the press. Two recent thoughtful analyses are Chapter 4 of the 2009 National Academies report *Assessing the Impacts of Changes in the Information Technology R&D Ecosystem*[67] and a 2009 report prepared for

66. http://www.washingtonpost.com/wp-dyn/content/article/2007/01/30/AR2007013002065.html

67. National Academies Press. (2009). *Assessing the Impacts of Changes in the Information Technology R&D Ecosystem: Retaining Leadership in an Increasingly Global Environment.*

NITRD by SRI[68]. It is clear from the available data that NIT workers are a significant part of the Nation's workforce and a dominant part of the Nation's S&T workforce. It is also clear that the Nation is not granting enough degrees in NIT fields to fill the available jobs.

Estimates of the number of employees in NIT occupations in the United States differ due to definitional variations. A reasonable but conservative estimate is 3.8 million. On the U.S. Bureau of Economic Analysis (BEA) definition of NIT, there were 3.79 million full time equivalent employees in NIT occupations in the United States in 2008 (the number rises to 3.99 million if part time employees are included).[69] Filtering U.S. Bureau of Labor Statistics (BLS) data[70] on BLS job codes corresponding to NIT occupations yields an estimate of 3.81 million workers. At the high end, applying the Organization for Economic Cooperation and Development (OECD) definition of NIT[71] to BEA data yields an estimate of 5.84 million full time equivalent employees.

NIT workers represent a majority of the Nation's S&T workforce. Data from BLS show that NIT occupations have comprised between 52% and 58% of all S&T occupations from 1998 to 2008.[72] To understand the significance of this statement, it is important to recognize that S&T occupations include all engineering, life science, physical science, and social science occupations.[73]

The most authoritative employment demand projections are made semi-annually by the BLS. The most recent projections, issued in November 2009, cover the ten-year period 2008-2018.[74] That report states (p. 85):

> Computer and mathematical occupations are expected to add 785,700 new jobs from 2008 to 2018, and, as a group, they will grow more than twice as fast as the average for all occupations in the economy, according to projections. It is anticipated that computer specialists will account for the vast majority of this growth, increasing by 762,700 jobs. Demand for computer specialists will be driven by the continuing need for businesses, government agencies, and other organizations to adopt the latest technologies… New computer specialist jobs will arise in almost every industry…

Figure 8.2-1 shows BLS projections for job growth in five major S&T categories between 2008 and 2018. For each field, "New Jobs" represents expansion while "New Jobs + Replacements" represents total available positions. The dominance of NIT ("Computer specialists") is clear.

---

68. SRI International. (May 29, 2009). *Networking and Information Technology Workforce Study: Final Report.* http://www.nitrd.gov/About/NIT_Workforce_Final_Report_5_29_09.pdf

69. U.S. Bureau of Economic Analysis. (2009). "Bureau of Economic Analysis, Gross Domestic Product by Industry." http://www.bea.gov/industry/xls/GDPbyInd_VA_NAICS_1998-2009.xls

70. U.S. Bureau of Labor Statistics. (2009). "May 2009 National Occupational Employment and Wage Estimates United States." http://www.bls.gov/oes/current/oes_nat.htm

71. Organisation for Economic Cooperation and Development. (2008). (page 3). *OECD Information Technology Outlook 2008.*

72. U.S. Bureau of Economic Analysis. (2009). "Bureau of Economic Analysis, Gross Domestic Product by Industry." . http://www.bea.gov/industry/xls/GDPbyInd_VA_NAICS_1998-2009.xls

73. BLS codes 15-0000 through 19-0000.

74. "Occupational employment projections to 2018." (November 2009). Monthly Labor Review, U.S. Bureau of Labor Statistics.

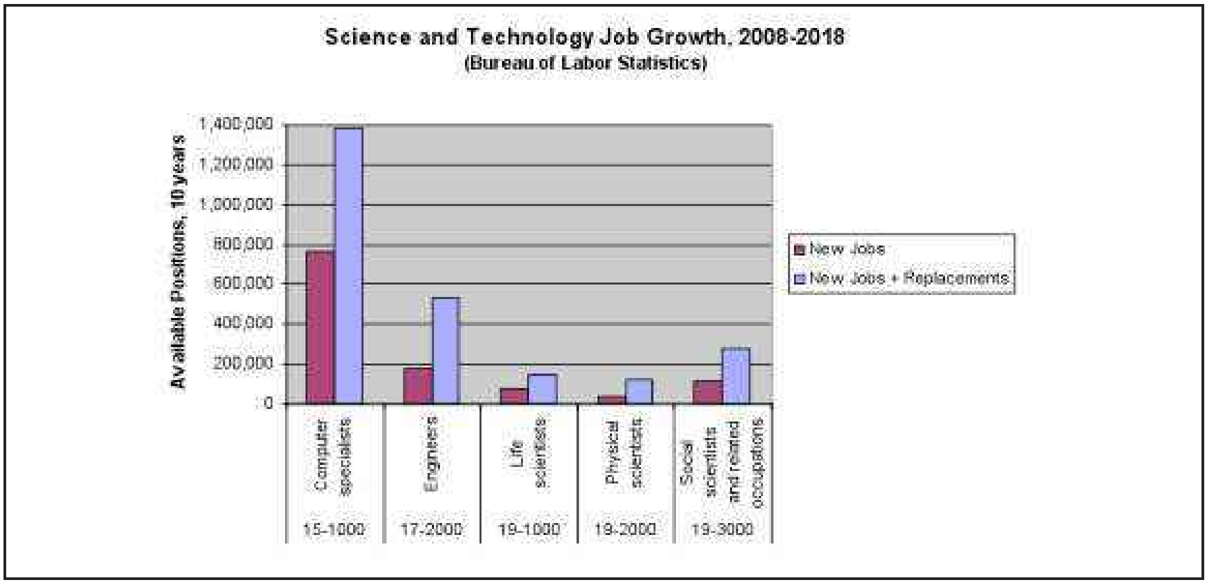


Figure 8.2.1

To summarize the demand side of the equation: While there will be inevitable variations in demand for every field, the long-term prospects for employment in NIT occupations in the United States are exceedingly strong. All other S&T fields pale by comparison.

The supply side of the equation is a cause of great concern. Figure 8.2-2 shows NSF statistics for degrees granted in 2006 in the S&T fields corresponding to the five BLS categories.[75] Degrees in psychology and the social sciences dominate, with degrees in other S&T fields lagging far behind.

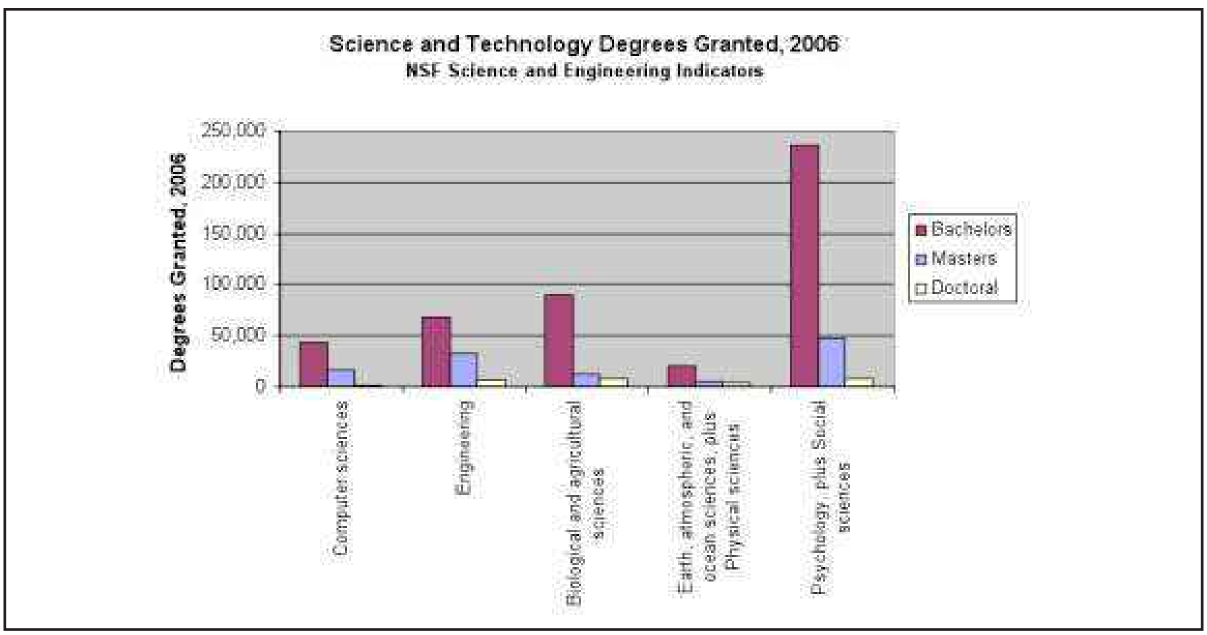


Figure 8.2.2

75. National Science Foundation, Division of Science Resources Statistics. (October 2008). Science and Engineering Statistics, S&E Degrees. (1966-2006). http://www.nsf.gov/statistics/nsf08321/content.cfm?pub_id=3785&id=2

It is difficult to make direct field-by-field comparisons between available jobs and degrees granted: Individuals with degrees in many fields work in NIT occupations, and individuals with NIT degrees work in many fields. With this important caveat, Figure 8.2-3 compares annualized job openings (BLS data) to total annual degrees granted (NSF data) for the five S&T fields reported in the previous two figures. Recent ACT data show an even greater disparity between declared student interest in computing majors and job growth predictions.

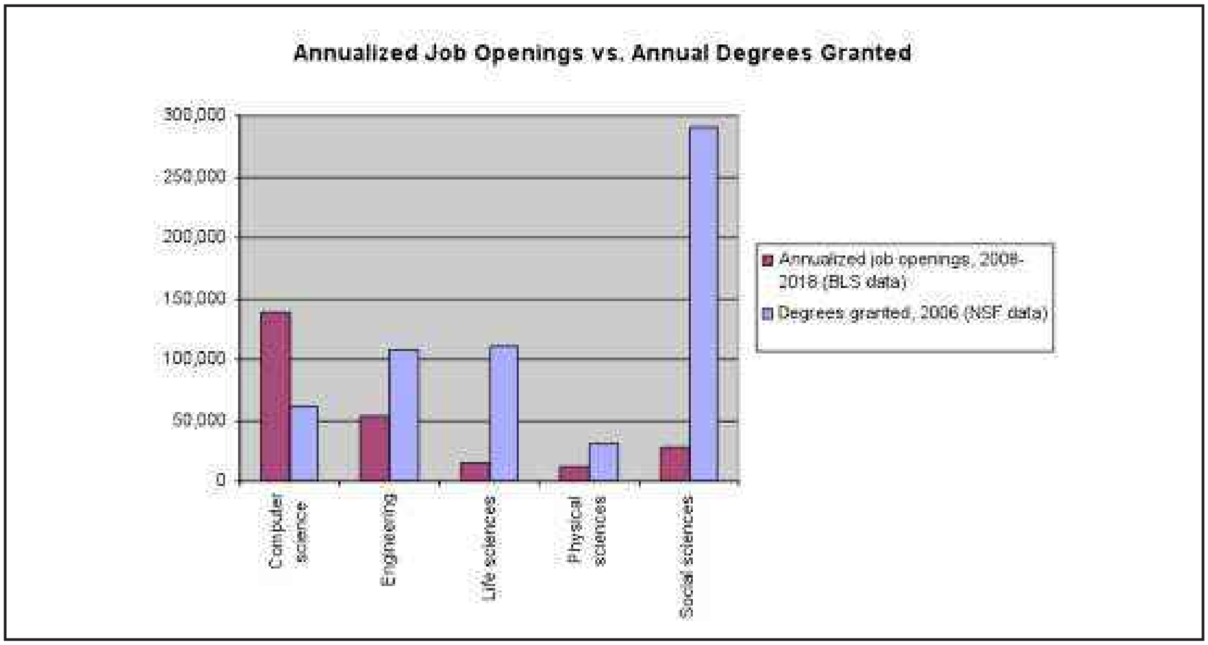


Figure 8.2.3

Understanding information technology – fluency, computational thinking, computer science. NIT pervades modern life. Every citizen – not just the NIT professional – needs to be fluent with information technology. The various dimensions of "NIT fluency" were the subject of a landmark 1999 National Academies study[76] that has stood the test of time remarkably well.

Fluency obviously involves a set of skills, such as using a word processor or spreadsheet, using the Internet to find information and resources, and using a database system to set up and access information. But fluency also involves a set of concepts and capabilities that have little to do directly with the use of a computer, but rather have to do with "computational thinking."[77] Basic concepts of computational thinking include abstraction, modeling, algorithmic thinking, algorithmic efficiency and analysis, stepwise fault isolation, and universality. Basic capabilities include algorithmic expression, managing complexity, and evaluating information. The skills, concepts, and capabilities of NIT are illustrated in the sidebar, taken from the National Academies study.

To illustrate the impact of computational thinking, consider two fields, linguistics and biology. Linguistics was transformed in the 1960s by the introduction of formal grammars, due to Chomsky and others – classic computational thinking. Biology was transformed beginning at about the same time, but extending

76. National Academies Press. (1999). *Being Fluent with Information Technology.*
77. Jeannette M. Wing. (March 2006). "Computational Thinking." *Communications of the ACM 49*,3, pp. 33-35.

over a longer period, by Watson and Crick's discovery that the human genome was a biochemically-implemented digital code; this discovery, and countless subsequent technological and conceptual breakthroughs, transformed biology into an information science.

The term "computer science" has traditionally meant the academic specialization that prepares NIT professionals. But this old characterization needs to be extended. Just as learning mathematics includes everything from children learning to count to post-docs studying algebraic topology, so learning computer science should be understood as gaining a full spectrum of skills from the elements of fluency to the most advanced graduate concepts. The recent PCAST report on Science, Technology, Engineering, and Mathematics (STEM) education also argues for a deeper understanding of computer science[78]:

> Computer-related courses should aim not just for technological literacy, which includes such utilitarian skills as keyboarding and the use of commercial software packages and the Internet, but for a deeper understanding of the essential concepts, methods and wide-ranging applications of computer science. Students should gain hands-on exposure to the process of algorithmic thinking and its realization in the form of a computer program, to the use of computational techniques for real-world problem solving, and to such pervasive computational themes as modeling and abstraction, modularity and reusability, computational efficiency, testing and debugging, and the management of complexity.

If Americans are to acquire proficiency in all levels of computing, their education must begin when they are children. Fluency with NIT skills, concepts, and capabilities; facility in computational thinking; and an understanding of the basic concepts of computer science must be an essential part of K-12 STEM education.

Advancing STEM education. Many previous studies of the NIT workforce[79,80] have emphasized the importance of visas in addressing America's NIT workforce gap. Between 40% and 50% of the 214,271 H-1B petitions approved in Fiscal Year 2009 were for workers in computer-related occupations.[81]

This need for imported NIT expertise persists. However, the Nation must be more aggressive in pursuing a long-term solution by developing the necessary expertise in the American populace. It is our view that such a solution must begin with dramatic enhancements to K-12 STEM education. Since K-12 STEM education is the focus of the PCAST report cited above, we will not discuss it further here, except for a brief elaboration on the role of computer science in STEM.

Today, K-12 education largely ignores computer science. Most high school computing courses, teach only basic literacy – the use of word processors, spreadsheets, etc. Those courses are typically taught

78. President's Council of Advisors on Science and Technology. (September 2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future.* http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

79. National Academies Press. (2009). *Assessing the Impacts of Changes in the Information Technology R&D Ecosystem: Retaining Leadership in an Increasingly Global Environment.*

80. President's Council of Advisors on Science and Technology. (August 2007). *Leadership Under Challenge: Information Technology R&D in a Competitive World.* An Assessment of the Federal Networking and Information Technology R&D Program. http://www.nitrd.gov/pcast/reports/PCAST-NIT-FINAL.pdf

81. U.S. Department of Homeland Security. (April 15, 2010). *Characteristics of H-1B Specialty Occupation Workers: Fiscal Year 2009 Annual Report.*

on the Career and Technical Education (CTE) track, and thus are not attractive to most college-bound students. Even schools that do provide academic computing courses appropriate for college preparatory students generally offer them as electives; few states count computing as a math or science requirement. Nationwide, fewer than 8% of our schools offer Advanced Placement (AP) Computer Science. Few K-12 computing teachers have any formal background in computing. To build a competitive, global workforce we will need to do much better than this.

All students – but most certainly all STEM students – will need a firm grounding in computing. They will need a level of sophistication that goes well beyond merely being able to use computational systems and devices. They will need to be able to "bend" computation to their purposes. To reach this level, all high school students must have the opportunity to take rigorous, academic computing classes. Computing must thoroughly infuse the K-12 curriculum, so that students have an opportunity to understand the role of computing in solving problems in many disciplines. As an example, students often use off-the-shelf simulations in biology or physics classes, but the experience would be much richer if it went beyond just filling in numbers and seeing the results. Students would be enriched by an understanding of what a simulation is: What is an abstraction? How are simulation models developed? How are they tested and validated?

To teach computing either as a separate discipline or as content infused across the curriculum, we will need better prepared teachers. As a Nation we must undertake an aggressive, and perhaps unprecedented, effort to prepare teachers who can effectively teach computing. We should support all of this with state-of-the-art online learning and social communities both for the teachers and their students.

9. Recommendations: Technological and Human Resources

Hardware, Software, and Data Infrastructure

NIT infrastructure – be it computational resources, communication networks, community databases, or collaboration tools – has become essential to research in virtually all fields. Although some infrastructure is acquired and managed exclusively by individual research projects, many research fields benefit from access to large-scale shared infrastructure – shared because of the considerable expense of acquiring and maintaining it, because of the long-term need, or because of the desire that multiple researchers use a common base. High-end computing systems, such as those made available the NSF Supercomputer Centers, and collections of curated data, such as PubMed and the Protein Data Bank, are examples of large-scale shared NIT infrastructure. High-end computing infrastructure provided by the NSF Centers and similar facilities has a long history, but shared infrastructure that supports data access, management, use, and preservation is in its early stages. The health of the Nation's research enterprise depends on sustained and reliable infrastructure of both kinds.

An important observation is that virtual or physical computing centers that provide infrastructure services for general R&D can often boost their value by hosting some NIT research activities as well. Some NIT research that serves to advance the technology underlying the infrastructure can be conducted using that infrastructure without disruption to other users. In addition, some science and engineering research using computer centers may confer extra value by stress-testing the infrastructure and providing a cadre of skilled consultants to help other users.

Recommendation: With NSF taking the lead, the NITRD agencies should develop an improved framework for the development and support of shared large-scale research infrastructure with the following properties:

- Proposed NIT infrastructure projects should be evaluated not only on whether they satisfy a demonstrated need, but also on their adaptability, reliability, adoptability, stability, size of user base, capability, and other appropriate metrics of infrastructure success, as well as on their plans for sustaining the infrastructure over time.
- Shared large-scale infrastructure is best managed with robust rather than minimal levels of support. To that end, the organizations that develop and manage large scale computation and data infrastructure should be constituted so that researchers working on problems for different agencies can perform their work at these common centers. (Occasionally, mission agency constraints will preclude such sharing, for example when data is classified.)
- In budgetary summaries, large-scale infrastructure costs for NIT resources devoted to R&D in areas other than NIT (e.g., in physics or medicine) should be clearly designated as infrastructure for those disciplines rather than mislabeled as NIT R&D. NIT R&D should be explicitly called out in budget summaries, as is R&D for other user communities.

- Plans and practices should be defined and implemented to manage the curation and preservation of long-lived and large data sets, so that data infrastructures survive beyond the lifetime and the boundaries of the projects that generated them.

Recommendation: NITRD should initiate a proactive approach for supporting data-driven research and the preservation of research data. We propose that NITRD work with each agency to designate critical data collections important to their communities and "best of breed" repositories to foster sustainable data infrastructure. This data infrastructure should follow best practices in curation and community standards, offer broad access for the research community, and ensure the sustainability of community data needed for new discovery. Programs should exploit the capabilities of the private sector, university libraries, government repositories, and other facilities that can support best practices and sustainable business models for broad access to long-lived digital information.

Education and Human Resources

The ever-expanding role of NIT in our society creates an ever-increasing demand not only for NIT professionals, but also for individuals who can utilize NIT flexibly and creatively and who can apply NIT "modes of thought" in a wide variety of endeavors. The Nation must take concrete steps to ensure that the American people have the education and skills to meet that demand.

Recommendation: The NSTC's Committee on STEM Education proposed in a recent PCAST report[82] must exercise strong leadership to bring about fundamental changes in K-12 STEM education in the United States, among them the incorporation of computer science as an essential component.

Research is needed to inform the necessary changes to STEM education. That research must address both curriculum content and understanding of the motivations and incentives that will encourage students to seek and persevere in STEM education.

Recommendation: NSF and ED should fund research to determine an age-appropriate progression of concepts for STEM education in computer science that generates strong skills in fluency, computational thinking, and the science and engineering aspects of computer science. That research should include the creation and assessment of the best ways to enable students to learn those concepts. The agencies should work with the academic community to determine and continuously update the appropriate concepts.[83]

Recommendation: NSF and ED should fund research to analyze why people do or do not choose to become computing professionals and why students of all ages, from childhood to post-graduate, do or do not choose to study computer science. That research should identify factors that inhibit greater participation in NIT and should propose and evaluate remedies.

82. President's Council of Advisors on Science and Technology. (September 2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future*. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf

83. *A Model Curriculum for K-12 Computer Science: Final Report of the ACM K-12 Task Force Curriculum Committee.* Computer Science Teachers Association and Association for Computing Machinery, 2006.

10. Strengths and Limitations of the NITRD Coordination Process and Structure

The NITRD Program is the primary mechanism by which the Federal Government coordinates its unclassified NIT R&D investments. Fourteen Federal agencies, including all of the large science and technology agencies, are formal members of the NITRD Program. Many other Federal organizations also participate in NITRD activities. The National Coordination Office (NCO) for NIT R&D supports the planning, budget, and assessment activities of the NITRD Program.

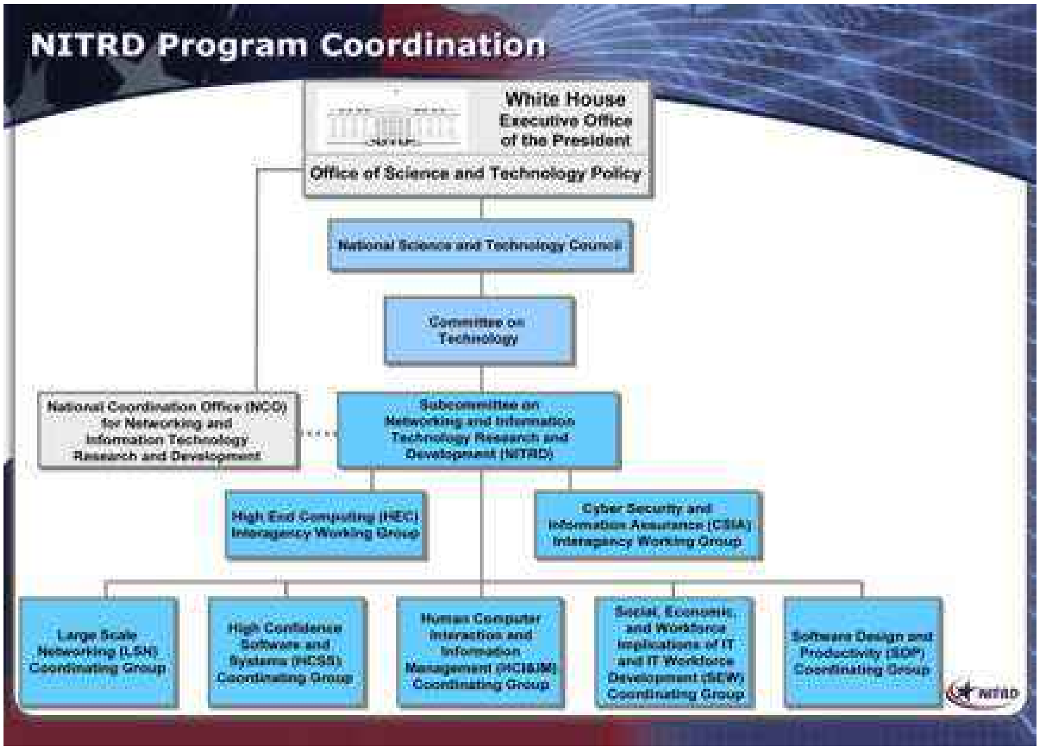
The NITRD Program operates under the aegis of the NITRD Subcommittee of the NSTC Committee on Technology. The Subcommittee, comprising representatives from each of NITRD's 14 member agencies, provides overall coordination for NITRD activities.

The NITRD Program currently includes eight Program Component Areas (PCAs)[84]. These PCAs represent NIT R&D budget categories, and map fairly directly onto a set of Interagency Working Groups and Coordinating Groups that carry out much of NITRD's work. The organization chart (next page), taken from the NITRD web site[85], illustrates these structures and relationships.

The NITRD Interagency Working Groups and Coordinating Groups are largely populated by individuals at the program manager level. The co-chairs of these groups stressed repeatedly to the PCAST NITRD Program Review Working Group the great benefit of having these individuals meet to exchange views and plans. Recent coordination in Cyber-Physical Systems – a response to a recommendation of the 2007 PCAST assessment of the NITRD Program[86] – is a specific example where relationships established through the NITRD process paid substantial dividends.

At the same time, members of the Interagency Working Groups and Coordinating Groups – and even many of the individuals who represent their agencies on the NITRD Subcommittee of the NSTC – typically do not have the seniority to make agency-level decisions and commitments. In fact, many are volunteers, not formally appointed by their agency heads and not formally evaluated on their NITRD participation. The NITRD Subcommittee itself meets only three times per year, each time for less than a day. Most of the agencies participating in NITRD (NSF being a notable exception) naturally concentrate on their own mission focus, and do not feel direct responsibility for the overall health of the Nation's NIT R&D enterprise. Finally, regardless of the level of agency representation, commitments can be difficult to implement because the various NITRD agencies report to different Office of Management and Budget (OMB) examiners and different Congressional appropriations subcommittees.

84. http://www.nitrd.gov/subcommittee/program.aspx
85. http://www.nitrd.gov/SubCommittee/NITRD-Org-Chart_121608.pdf
86. President's Council of Advisors on Science and Technology. (August 2007). *Leadership Under Challenge: Information Technology R&D in a Competitive World.* An Assessment of the Federal Networking and Information Technology R&D Program. http://www.nitrd.gov/pcast/reports/PCAST-NIT-FINAL.pdf



The NITRD coordination function has been strengthened recently by the creation of Senior Steering Groups (SSGs) in two areas: Cybersecurity and Information Assurance, and Health IT. Because these SSGs are focused, ad hoc (vs. standing), and cover areas that have emerged as clear priorities for multiple agencies, they are attracting agency representatives with decision-making authority. NIT for Sustainability (energy, environment, climate change, transportation) and NIT for Education and Life-long Learning are additional areas that might well also warrant such focused attention.

Finding: The NITRD inter-agency coordination mechanism is widely – and we think correctly – viewed as successful and valuable.

There are limits, however, to what the NITRD coordination process can be expected to achieve. A recent report on cybersecurity from the Government Accountability Office (GAO) stated: "While officials within OSTP's Subcommittee on Networking and Information Technology Research and Development … have played a facilitator role in coordinating cybersecurity R&D efforts within the Federal Government, they have not led agencies in a strategic direction."[87] Information gathering for the GAO report was concurrent with the creation of the Cybersecurity SSG, so the SSG's impact had not yet been felt. Nonetheless, it is our view that the sort of strategic leadership envisioned by the GAO is not a reasonable expectation for NITRD even in the best of circumstances. NITRD has few carrots to offer, and few sticks to employ.

87. United States Government Accountability Office GAO-10-466. (June 2010). *CYBERSECURITY: Key Challenges Need to Be Addressed to Improve Research and Development.* http://www.gao.gov/new.items/d10466.pdf

Finding: The potential contributions of advances in NIT to several national priorities span multiple agencies. A successful coordinated attack on the Nation's most challenging and important problems requires focused attention on multi-disciplinary, problem-driven research in NIT. That focus must come from Federal leadership. NITRD is chartered and staffed to coordinate multi-agency programs. Strategic leadership, when necessary, must come from those with the authority to implement new strategies, namely OSTP and NSTC, to which NITRD reports. That leadership must have continuity, breadth and depth, and a focus on NIT.

In the following section, we recommend that the Federal Government establish a high-level standing committee that focuses on a national strategic vision for NIT, and explain the rationale behind this recommendation. We have given careful thought to suggesting yet another advisory committee. We do so because of the broad impact and profound importance of NIT for the United States, which creates an urgent national need for continuing attention to the sustained high-level strategic direction of NIT.

We now turn our attention to NITRD budget reporting. Decisions regarding what investments to report as part of the NITRD crosscut, and under what PCAs to report them, are left to each agency, with no oversight by the NCO or the NITRD Subcommittee. As a result, there is enormous variability in what is included. For some NITRD agencies, such as NSF, the amount included in the NITRD crosscut is a reasonably accurate representation of the agency's investment in NIT R&D. For other agencies, there can be significant discrepancies. These discrepancies arise, for the most part, due to confusion between true NIT R&D and NIT that supports R&D in other fields. The latter is legitimately part of an agency's R&D portfolio – often a crucially important part. But it is not NIT R&D. (For more details, see the sidebar, "The NITRD Crosscut Budget Significantly Overstates the Federal Investment in NIT R&D" on page 96).

Finding: The Nation is actually investing far less in NIT R&D than is shown in the Federal budget. A substantial fraction of the NITRD crosscut budget represents spending on NIT that supports R&D in other fields, rather than spending on R&D in the field of NIT itself.

The broadening role of NIT makes it important to broaden correspondingly the set of agencies involved in NITRD. It is also important to broaden the perspectives of the individuals who represent their agencies – the HPCC origin of NITRD is still evident in the balance of interests of participating individuals.

All agencies should be aware that NITRD is the resource for NIT-related issues, and that achieving agency missions may require significant advances in NIT (that is, R&D) that go beyond the application of existing technology[88]. We understand that there has been some move to create parallel coordination efforts to answer these concerns, but we urge that this path be resisted: issues such as large-scale data management and analysis, NIT workforce, cybersecurity education, and cyber-enabled education fall naturally within the purview of NITRD, which is already providing successful Federal inter-agency coordination.

88. The inclusion of an R&D component in the National Broadband Plan is a good example of the latter. http://www.broadband.gov/plan/

NITRD response to the 2007 PCAST NITRD review. The NITRD Subcommittee Co-Chairs, Drs. George Strawn and Jeannette Wing, were asked[89] to report progress toward and impediments to implementation of the 17 recommendations in the 2007 PCAST NITRD review[90]. Drs. Strawn and Wing indicated that efforts were made to implement all recommendations, with five of them described as achieved[91]. Efforts to increase agency membership and engagement, to improve network capacity within the Federal Government, and to streamline processes from discovery to commercialization were highlighted as successes. In most cases where recommendations were partially achieved, insufficient information was provided regarding impediments to implementation to allow the Working Group to ascertain whether significant barriers exist that limited NITRD's ability to implement the recommendations.

The Draft NITRD 2010 Strategic Plan. The draft NITRD 2010 Strategic Plan[92] articulates several directions that are consistent with those suggested in this report. The challenge is to translate these directions into reality, and to put ongoing strategic advice and leadership into place that will ensure greater agility on the part of NITRD in the future.

The NITRD Crosscut Budget Significantly Overstates the Actual Federal Investment in NIT R&D

The aggregate NITRD crosscut budget[93] – currently in excess of $4 billion – significantly overstates the actual Federal investment in NIT R&D.

Most obviously, a large portion of the "High End Computing Infrastructure and Applications" budget category, which accounts for roughly $1.5 billion of the $4.3 billion NITRD total, is attributable to computational infrastructure used to conduct R&D in other fields, and not to NIT R&D or to infrastructure for NIT R&D.

Beyond this, however, various agencies include in their reports for other NITRD budget categories investments in NIT that support R&D in non-NIT fields. The laudable transparency of NIH's NITRD grant reporting allowed an expert in NIT at the Science and Technology Policy Institute (STPI) to review the abstracts of the top 100 awards (by award size) in NIH's 2009 NITRD crosscut[94] for actual NIT R&D content[95]. STPI developed a 14-part coding scheme to categorize the 100 projects.[96] The analysis showed that of the

89. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nco-nitrd.pdf

90. President's Council of Advisors on Science and Technology. (August 2007). *Leadership Under Challenge: Information Technology R&D in a Competitive World.* An Assessment of the Federal Networking and Information Technology R&D Program. http://www.nitrd.gov/pcast/reports/PCAST-NIT-FINAL.pdf

91. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nco-nitrd.pdf

92. http://www.nitrd.gov/DraftStrategicPlan/NITRDstratplan_Public_Comment.pdf

93. *Networking and Information Technology Research and Development Supplement to the President's FY 2011 Budget,* (February 2010) (page 21). http://www.nitrd.gov/pubs/2011supplement/FY11NITRDSupp-FINAL-Web.pdf

94. *As identified by NIH's RCDC (Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC) – Project Listing by Category – Non-ARRA NIT R&D FY 2009 Funding)* and RePORT system (http://report.nih.gov/rcdc/categories/ProjectSearch.aspx?FY=2009&ARRA=N&DCat=Networking+and+Information+Technology+R+and+D)

95. R&D was defined per OMB's "Character Classification (Schedule C)" (http://www.whitehouse.gov/sites/default/files/omb/assets/a11_current_year/s84.pdf), expanded to include "physical assets" (hardware and software systems) used for R&D.

96. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stpi-nitrd-9-15-2010.pdf

95 projects with available abstracts, only 4 appear to be making a contribution to actual NIT R&D, such as novel computer science methods, novel simulation methods, or novel system design and computing methods. Another 14 projects could be considered "borderline" although they seem to focus on NIT development alone. The remaining 77 of the 95 projects with available abstracts appear to have no NIT R&D content. (For example, they may involve NIT infrastructure to support biomedical or biochemical R&D, but with no novel NIT R&D contribution.)

NIT R&D per OMB's Definition	Number of Awards	Percent of Awards	Dollar Value of Awards	Percent of Dollar Value
Yes	4	4%	$10,882,505	2%
Borderline	14	14%	$52,108,659	9%
No	77	77%	$497,208,700	86%
No abstract	5	5%	$14,722,586	3%
Total	100		$574,992,450	

In terms of dollar value, these 100 projects totaled $575 million, roughly half of NIH's 2009 NITRD cross-cut. Of this, approximately $500 million was allocated to projects that STPI judged to have no NIT R&D content, approximately $52 million went to "borderline" projects that seem to focus on NIT development alone, and approximately $11 million went to projects that are making a contribution to actual NIT R&D. (Approximately $15 million was allocated to 5 projects that could not be assessed because abstracts were not available.)

In short, STPI concluded that 86% of these awards, by dollar value, have no NIT R&D content; 3% could not be assessed, 9% were judged borderline, and 2% were judged to be making a contribution to actual NIT R&D[97]. Although other agencies do not report NIT R&D spending in sufficient detail to make the same analysis possible, it seems likely that in many cases a similar confusion in classification of NITRD investment occurs.

These findings highlight both the increasing ubiquity of NIT infrastructure for conducting R&D in many fields and the difference between this infrastructure and actual NIT R&D – work that makes a novel contribution to NIT.

97. There was one extremely large award – $250,355,440. The other 99 awards ranged from $13,935,921 down to $1,299,732. Even if one ignores the one large award (which, obviously, fell into the "No" grouping), more than 75% (by dollar value) of the 99 remaining top 100 NIH NITRD awards were judged by STPI to have no NIT R&D content, and only 3.3% fell into the "Yes" grouping.

11. Recommendations: NITRD Coordination Process and Structure

Coordination of NIT research and infrastructure by the Federal Government has enjoyed considerable success and is cited as an example for other areas of research. Increasing the recognition of NITRD participants and the flexibility of NITRD coordination, along with broadening the reach of coordination and improving its reporting mechanisms, can make the coordination process still more effective in adapting to changes in NIT and in important national priorities.

Recommendation: To strengthen the memberships of the NITRD Subcommittee (the leadership group), the Working Groups, the Coordinating Groups, and the Senior Steering Groups, heads of NITRD agencies should:

- Appoint to the Subcommittee individuals with significant influence and decision-making authority, and with a balanced and comprehensive view of the role of NIT advances in fulfilling the missions of their agencies;
- Make NITRD activities an explicit part of job duties and performance assessments for all appointees to the Subcommittee, Working Groups, Coordinating Groups, and Senior Steering Groups.
- Charge the NITRD subcommittee with the responsibility to ensure that its meetings are devoted to strategic discussions.

NITRD should be the single Federal organization responsible for coordinating NIT R&D, not one of several. The NCO should increase agency participation in NITRD by communicating to the agencies that achieving their missions requires advances in NIT R&D and that NITRD participation is the avenue for sharing responsibility for those advances.

Recommendation: NITRD should be the sole coordinating body for NIT R&D. To facilitate that coordination, NCO should:

- Include agencies engaged in significant R&D in more than a few of the "NIT Research Frontiers" as full participants in NITRD. The agency heads should ensure that their agencies participate and pay their NITRD operating cost assessments;
- Enable agencies with NITRD-related interests to be included in appropriate NITRD Groups without requiring full participation in NITRD;
- Heighten agency awareness that NITRD is the resource for NIT-related issues, and that achieving agency missions may require *advances* in NIT R&D that go beyond the application of existing technology; and
- Make greater use of mechanisms such as the SSGs to attract agency representatives with decision-making authority in response to specific cross-agency priorities, as has been done recently with Cybersecurity and with Health IT.

As areas of special importance come and go, the NITRD structure must adapt. NITRD must be able easily to start new coordinating groups, to form groups of limited duration, and to have groups representing different levels of management, from program officers all the way to agency heads or heads of major units within an agency. The recently established SSGs are a good first step. To achieve that enhanced flexibility, it must be possible to constitute and define coordinating groups independently of the PCAs, so that the set of groups can change more frequently, while the PCA-based budget reporting remains more stable. Although some stability is needed in the definition of the NITRD PCAs, they too must be redefined periodically to reflect the current nature of the field. To better serve its participating agencies, NCO should increase its sharing of information.

Recommendation: The Federal Government should strengthen the power of NITRD coordination to promote national priorities and important computer science research frontiers by making the following changes:

- The NCO should give the Interagency Working Groups and Coordinating Groups more flexibility by decoupling them from the PCAs, while PCA-based budget reporting remains more stable.
- The NCO and OMB should modernize the NITRD PCAs to reflect the current nature of the field, along the lines suggested in the "NIT Research Frontiers" and "Hardware, Software, and Data Infrastructure" sections of this report.
- The NITRD priorities should be referenced in the annual OMB/OSTP Budget Priority Memorandum.
- The NCO should maintain a single publicly accessible database of all NITRD awards and expenditures, containing data uploaded from NITRD agency databases.
- The NCO should provide to the Director of OSTP regular and detailed reports from the NITRD Subcommittee on NITRD issues, strategy, implementation, and budget.

The following summary recommendation, which appears in the Executive Report, highlights the most important elements of the more detailed recommendations above.

Recommendation: The effectiveness of government coordination of NIT R&D should be enhanced:

- The number of NITRD member agencies should be increased. The duration, management levels, and topic areas of the NITRD coordinating groups should be flexible. Budget reporting categories should be decoupled from the coordinating structure.
- The NCO for NITRD should create a publicly available database of government-funded NIT research, and should provide regular detailed reporting to the Director of OSTP.
- OMB and OSTP should reflect NITRD priorities in their annual Budget Priority Memorandum.

We have noted previously the importance of differentiating investments in NIT infrastructure for other fields from investments in NIT R&D:

> **Recommendation: The NCO and OMB should redefine the budget reporting categories to separate NIT infrastructure for R&D in other fields from NIT R&D, and should ensure more accurate reporting of both NIT infrastructure investment and NIT R&D investment.**

Even if the above changes are made, there are still limits to what the NITRD coordination process can do. In light of the broad impact of NIT in the modern world and of its profound importance for the United States, a sustained high-level standing committee is needed to advise the Federal Government on both long-term and shorter-term strategy for NIT. That committee should include a mix of academic experts from computer science and allied fields and industrial leaders in NIT. The rationale for a committee *dedicated to NIT* is the need for focus, so that important issues get timely and in-depth attention that combines scientific and technical considerations, policy considerations, and economic considerations. The motivation for a *standing committee* is the need for continuous attention, so that advice can be predictive rather than reactive; the committee must identify emerging issues early and incorporate them in a sustained strategic vision for the Nation's strength in NIT. The standing committee must be *sufficiently large* that its members have the sufficient breadth and depth of knowledge and experience together with the shared context that they can provide a continuously evolving strategic vision.

> **Recommendation: The Federal Government must lead in ensuring that strong multi-agency R&D investments are made in NIT to address important national priorities**.
>
> - OSTP should establish a broad, high-level standing committee of academic scientists, engineers, and industry leaders dedicated to providing sustained strategic advice in NIT.
> - The NSTC should lead in defining and promoting the major NIT research initiatives that are required to achieve the most important existing and emerging national priorities.

12. The Role of Federal Investment in NIT R&D

In the past 15 years, the National Academies has published roughly a dozen studies of America's NIT R&D ecosystem, beginning with a highly influential 1995 report assessing the High Performance Computing and Communications Initiative[98], and including most recently a 2009 study of the impact of globalization and other recent forces[99]. The conclusions of many of these studies were summarized in the 2003 National Academies report *Innovation in Information Technology*[100]. Chapter 1 of that report provides an authoritative overview. The short Summary that precedes it, referring to the National Academies' Computer Science and Telecommunications Board (CSTB), states:

Here are the most important themes from CSTB's studies of innovation in IT:

- **The results of research**
  - America's international leadership in IT – leadership that is vital to the nation – springs from a deep tradition of research.
  - The unanticipated results of research are often as important as the anticipated results…
  - The interaction of research ideas multiplies their impact…
- **Research as a partnership**
  - The success of the IT research enterprise reflects a complex partnership among government, industry, and universities.
  - The federal government has had and will continue to have an essential role in sponsoring fundamental research in IT – largely university-based – because it does what industry does not and cannot do. Industrial and governmental investments in research reflect different motivations, resulting in differences in style, focus, and time horizon.
  - Companies have little incentive to invest significantly in activities whose benefits will spread quickly to their rivals. Fundamental research often falls into this category. By contrast, the vast majority of corporate R&D addresses product and process development.
  - Government funding for research has leveraged the effective decision making of visionary program managers and program office directors from the research community, empowering them to take risks in designing programs and selecting grantees. Government sponsorship of research especially in universities also helps to develop the IT talent used by industry, universities, and other parts of the economy.

98. National Academies Press. (1995). *Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure.*
99. National Academies Press. (2009). *Assessing the Impacts of Changes in the Information Technology R&D Ecosystem: Retaining Leadership in an Increasingly Global Environment.*
100. National Academies Press. (2003). *Innovation in Information Technology.*

- The economic payoff of research
  - Past returns on federal investments in IT research have been extraordinary for both U.S. society and the U.S. economy. The transformative effects of IT grow as innovations build on one another and as user know-how compounds. Priming that pump for tomorrow is today's challenge.
  - When companies create products using the ideas and workforce that result from federally sponsored research, they repay the nation in jobs, tax revenues, productivity increases, and world leadership.

The themes highlighted above underlie two recurring and overarching recommendations evident in the eight CSTB reports cited:

Recommendation 1: The federal government should continue to boost funding levels for fundamental information technology research, commensurate with the growing scope of research challenges. It should ensure that the major funding agencies, especially the National Science Foundation and the Defense Advanced Research Projects Agency, have strong and sustained programs for computing and communications research that are broad in scope and independent of any special initiatives that might divert resources from broadly based basic research.

Recommendation 2: The government should continue to maintain the special qualities of federal IT research support, ensuring that it complements industrial research and development in emphasis, duration, and scale.

From *Innovation in Information Technology*, National Academies Press, 2003.

Many of these themes have been explored earlier in this report. Here, we amplify just a few additional aspects to emphasize the crucial role of continuing and vigorous Federal support for NIT R&D.

12.1 The Critical Role of Federal Investment

In assessing how to maintain America's leadership in networking and information technology, a common fallacy is to overestimate the role of technology development and to underestimate the role of fundamental research. In fact, computer science research, carried out to a great extent in America's research universities with funding from Federal agencies such as DARPA and NSF, lies at the heart of this leadership.

The complementary roles of federally funded research and industry R&D merit further emphasis. Industry has made, and continues to make, crucial contributions to NIT R&D. The "extraordinarily productive interplay of federally funded university research, federally and privately funded industrial research, and entrepreneurial companies founded and staffed by people who moved back and forth between universities and industry"[101] has been well documented. It is important, however, not to equate the very large industry R&D investment in NIT with fundamental research of the kind that is carried out in universities and a small number of industrial research labs. The vast majority of industry R&D in NIT is focused on development – on the engineering of future products and product versions. Few major NIT

101. National Academies Press. (1999). *Funding a Revolution: Government Support for Computing Research.*

companies have formal research organizations, and even those that do invest relatively little in research compared to their investment in development activities. Fundamental research with the potential for future game-changing applications is a small fraction of overall industry R&D in NIT.

The Research Component of Industry R&D in NIT

Industry has made, and continues to make, major contributions to NIT R&D. It is important, however, not to equate the very large industry R&D investment in NIT with fundamental research of the kind that is carried out in universities and in a small number of industrial research labs. Appropriately, the vast majority of industry R&D in NIT is focused on development – on the engineering of future products and product versions. Few major NIT companies have formal research organizations, and even those that do invest relatively little in research compared to their investment in development activities.

To illustrate this, an expert in NIT at the Science and Technology Policy Institute (STPI) compared the total worldwide number of R&D employees of IBM and Microsoft to the worldwide number of employees in the research organizations of those companies.[102] (IBM and Microsoft are widely regarded as leaders in their levels of investment in fundamental research, compared to other U.S. NIT companies.) The table below summarizes these findings:

	IBM	Microsoft
Total research personnel worldwide	3,000	930
Total R&D personnel worldwide	40,000	36,000
Percentage of R&D personnel engaged in research	7.5%	2.5%

The research investments of IBM and Microsoft are significant by any measure – whether by comparison of those of other U.S. NIT companies, or in terms of the contributions to the field that they have yielded, or in terms of the actual dollars invested[103]. But they represent a very small proportion of the overall R&D investments of these companies.

The vast majority of industry R&D in NIT is focused on development. Fundamental research with the potential for future game-changing applications represents a small fraction of overall industry R&D in NIT.

Economic theory[104],[105] provides a clear explanation for the reluctance of industry to invest in fundamental NIT research to the extent that would be optimal for the Nation as a whole. Consider the case of a single firm that must decide how much to invest in such research. In order to maximize its profits, the company

102. The worldwide total of R&D employees was obtained via 10-K filings (for Microsoft) and personal communication with senior staff (for IBM). The worldwide number of employees in the research organization was obtained via personal communication with senior staff in each case. See http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stpi-nitrd-9-15-2010.pdf

103. According to their 10-K filings, IBM and Microsoft invested $6 billion and $9 billion in R&D during 2009, respectively. Multiplying these totals by the personnel ratios in the table yields estimates of $450 million and $235 million invested in research, respectively. IBM and Microsoft also make substantial investments in extramural research at universities.

104. Nelson, Richard. (1959). "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 67.

105. Arrow, Kenneth. (1962). "Economic welfare and the allocation of resources for invention," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, R. R. Nelson, ed., Princeton, NJ: Princeton University Press.

should invest only up to the point at which the cost of further research would exceed the incremental returns it might expect such research to generate for its shareholders.

One of the defining characteristics of fundamental research, however, is its breadth of applicability. When such research proves successful, it often drives innovations in a wider range of products, services, and target markets than are relevant to any one company. If the cost of conducting such broadly applicable research were amortized over all U.S. firms that might profit from its results, the optimal level of total investment would be significantly higher than the optimal level for any one firm, acting alone. In the absence of some enforceable mechanism for cost-sharing, however, each of the potential beneficiaries would be expected to systematically under-invest in fundamental research relative to what would be optimal for them as a group. Fundamental research in NIT, in particular, often yields advances whose impact is felt across an unusually wide range of applications and markets, increasing the magnitude of this predicted under-investment.

This problem is exacerbated when two or more companies compete within the same market. If our hypothetical firm anticipates that some portion of any economic value generated by its own fundamental research will be conferred on its competitors, it will tend to invest even less in such research. The patent system is designed to address such issues (at least in part) by granting a period of exclusive use to firms whose research leads to economically valuable innovations. Patents designed to protect computer software and systems, however, are often easier to work around than those in many other industries. Once the basic idea behind a given algorithm or computational approach has been publicly disclosed as part of the patent process, a competitor will often find it possible to design a system that draws on the same underlying idea to accomplish the same objective without infringing on the issued patent.[106]

Even if our company retains its findings and innovations as trade secrets, it may be difficult for it to capture the full economic value generated by its fundamental research.[107] When groundbreaking NIT research leads to qualitatively new types of IT-based systems, applications software, or IT-enabled services, the most valuable competitive information may be simply the demonstration that a market for these products or services exists. That is, once a competitor knows that it is technically and economically feasible to address a specific (and perhaps previously unrecognized) market with a new type of IT-based solution, it may in some cases be relatively straightforward either to reverse engineer the design or to address the same market needs in a different way. The very act of commercially exploiting the results of its NIT-related research may thus result in a transfer of value from the innovating company to its competitors. The anticipation of such "leakage" might be expected to reduce the investment that the innovator would be willing to make in fundamental research.

106. Samuelson, Pamela, and Scotchmer, Suzanne. (2002). "The Law and Economics of Reverse Engineering." *Yale Law Journal* 111,7, pp. 1499–1663.

107. Teece, David. (1986). "Profiting from technological innovation." *Research Policy* 15, 6, pp. 285–305.

Why We're Able to Google

Our ability to access information any time, any place through Bing, Google, and similar services has quickly become an integral part of modern life. A recent report from the National Academies[108] profiles Google as "An Example of Growing from Research to Global Brand:"

> *Larry Page and his Google cofounder Sergey Brin were research assistants at Stanford contributing to the National Science Foundation's Digital Library Initiative. Search was a natural component of this effort. World Wide Web search was not new. But Page and Brin had a new idea for improving search quality: the PageRank algorithm that weights World Wide Web page importance by the number and importance of other World Wide Web pages that link to it.*
>
> *In 1998, Google handled 10,000 search queries per day from a "server farm" located in the dormitory room of Larry Page, a computer science graduate student at Stanford University. Today, Google has 20,000 employees, diverse products, annual revenues of $25 billion, a market capitalization of nearly $200 billion, and is a verb. Google's story illustrates the critical nature of university research for start-ups and the huge difference that individuals make in the trajectory of a start-up.*

The PageRank patent – held by Stanford and licensed to Google – was Google's original "secret sauce." PageRank built upon earlier research at several other universities, and is only one of many research foundations on which Google's success is built. Google's scalable computing infrastructure is another major contributor to the company's success. The challenge in building scalable World Wide Web services is to create reliable systems out of huge collections of components. It's impossible to build hardware on this scale that does not suffer failures. If a single disk drive has a mean-time-between-failure of 3 years, a system of 100,000 disk drives will experience a failure about every 15 minutes! So software algorithms must be used to create an illusion of perfect reliability. At the heart of Google's solution to this problem lies the Paxos algorithm, invented at the DEC Systems Research Center more than 20 years ago building on previous work at MIT.

The list could go on and on. Success on Google's scale is built upon much more than "two bright young people in a garage."

Such problems are particularly pronounced in the case of:

- *Smaller companies* whose market shares are not large enough to capture a substantial fraction of the profits arising from their research (as might be the case, say, for a large firm that dominates its target market);
- *Software products*, which are often especially vulnerable to the leakage of research-derived innovations as a result of
  - Special difficulties in protecting such intellectual property (discussed above)
  - The easy migration to other firms (especially within certain geographic "clusters of innovation") of employees whose skills are often highly transferable to other application areas;

108. National Academies Press. (2009). (Figures updated.) *Assessing the Impacts of Changes in the Information Technology R&D Ecosystem: Retaining Leadership in an Increasingly Global Environment.*

- *High-risk research*, the successful results from which would, if appropriated by competitors, allow them to reap the economic benefits of major advances without underwriting the substantial costs associated with unproductive lines of investigation.

For these reasons, among others, it is unrealistic to expect the private sector alone to invest in fundamental NIT research to the extent that would be economically optimal for U.S. companies as a group, or for their employees, customers, investors, and other stakeholders. Although the more narrowly focused *later* stages of product-oriented R&D may be driven largely (though still not exclusively) by industry, we cannot rely on the private sector to ensure the long-term economic security of our Nation as a whole. Federal funding of fundamental research in NIT is essential if America is to maintain its economic competitiveness and standard of living within an increasingly integrated global economy in which the technological capabilities of a number of countries are now expanding at a faster rate (in relative terms) than our own.

As discussed in Section 4, NIT plays a central role not only in preserving the Nation's economic competitiveness, but also in achieving the great majority of its most important *non-economic* objectives. Industry, however, has no natural incentives to conduct the fundamental research that will be required to achieve many of these objectives. If we are to harness NIT's immense potential to address critically important priorities in such areas as national security, energy, health, and education over the coming years, sustained Federal funding is essential.

Finding: The vast majority of industry R&D in NIT is focused on development – on the engineering of future products and product versions – and not on fundamental research. Private-sector R&D, while important, is (appropriately) driven by economic incentives that preclude its serving as a substitute for the sustained Federal funding of fundamental research in NIT.

12.2 The Incremental Investment Implied by this Report

The investments that our Nation has made in NIT R&D are among the best investments that our Nation has made. As discussed in Section 3 of this report, the NIT research landscape is changing rapidly and dramatically. The NITRD portfolio must change, too. We have chosen to focus this assessment less on NITRD as it is, and more on NITRD as it should be:

- Increased emphasis on advances in NIT necessary to achieve America's priorities, as outlined in Sections 4 and 5;
- A new view of the core of the field, as outlined in Sections 6 and 7;
- The need for larger and more multidisciplinary teams of researchers for longer periods of time, required by both of the above.

These changes will require additional resources – some combination of new funds and redirected existing funds – along with additional attention by multiple Federal agencies.

Of crucial importance is our finding that the Nation is actually investing far less in NIT R&D than is shown in the Federal budget. Our analysis indicates that a substantial fraction of the NITRD crosscut budget represents spending on NIT that supports R&D in other fields, rather than spending on R&D in the field of NIT itself.

A bottom-up analysis of some of the key initiatives that we recommend in this report suggests that an investment of at least $1 billion annually will be needed for new, potentially transformative NIT research. We believe that a lower level of investment in this critically important area could seriously jeopardize America's national security and economic competitiveness. Uncertainty regarding the precise nature of current expenditures makes it difficult to determine how much of this investment can be obtained through repurposing and reprioritization and how much will require new funding. It will be an early responsibility of the standing advisory committee recommended in Section 11, working with the NCO, to resolve this uncertainty, as well as to provide a detailed assessment of the investment requirements of the various initiatives.

By the time of the next assessment of the NITRD Program, it is essential that the NITRD coordination process be strengthened, that greater insight be available regarding the precise nature of NITRD Program expenditures, and that mechanisms be in place to provide sustained strategic advice and leadership – all of which are described in Sections 10 and 11. It is clear, though, that additional investment, focused as indicated in this report, is essential.

Appendix A: Expert Input Into the PCAST NITRD Review

PCAST is grateful for the input of the experts listed below. Listing here does not imply endorsement of this report or its recommendations.

Thomas E. Anderson
Robert E. Dinning Professor
University of Washington

Robert D. Atkinson
President
The Information Technology & Innovation Foundation

Forest Baskett
General Partner
NEA

Kris Berger
National Director of Community Programs
New Leaders for New Schools

Guy Blelloch
Professor
Carnegie Mellon University

Rodney Brooks
Panasonic Professor of Robotics
Massachusetts Institute of Technology

Karen Cator
Director, Office of Educational Technology
Department of Education

Peter Chen
Arthur F. Thurnau Professor
University of Michigan

Aneesh Chopra
Chief Technology Officer
Assistant to the President
Office of Science and Technology Policy

Deborah Crawford
Deputy Assistant Director
Computer and Information Science and Engineering Directorate
National Science Foundation

David Culler
Professor
University of California, Berkeley

Michael Dahlin
Professor
University of Texas, Austin

Carol Diamond
Managing Director
Markle Foundation Healthcare

Thom Dunning
Director
National Center for Supercomputing Applications
University of Illinois, Urbana-Champaign

William Feiereisen
Director of High Performance Computing
Lockheed Martin Information Systems and Global Solutions

Charles Friedman
Chief Scientific Officer
Office of the National Coordinator for Health Information Technology
Department of Health and Human Services

Erwin Gianchandani
Director, Computing Community Consortium
Computing Research Association

Marcy Gallo
Professional Assistant
House of Representatives Committee on Science and Technology

Steven Gribble
Associate Professor
University of Washington

Christopher Greer
Assistant Director for Information Technology R&D
Office of Science and Technology Policy

John Hennessy
President
Stanford University

Peter Harsha
Director of Government Affairs
Computing Research Association

Daniel Hitchcock
Computer Scientist
Department of Energy

John Holdren
Assistant to the President
Director, Office of Science and Technology Policy

Suzanne Iacono
Program Director
National Science Foundation

Sairah Ijaz
Government Accountability Office

Tom Kalil
Deputy Director for Policy
Office of Science and Technology Policy

Henry Kelly
Principal Deputy Assistant Secretary
Office of Energy Efficiency and Renewable Energy
Department of Energy

John Leslie King
Vice Provost for Strategy, and W.W. Bishop Professor of Information
University of Michigan

James Kirby
Center for High Assurance Computer Systems
Naval Research Laboratory

Chris Kemerer
David M. Roderick Professor of Information Systems, and Professor of Business Administration
Katz Graduate School of Business
University of Pittsburgh

Janet Kolodner
Regents' Professor
School of Interactive Computing
Georgia Institute of Technology, and
Program Officer
National Science Foundation

James Larus
Director of Research and Strategy
eXtreme Computing Group
Microsoft Research

Peter Lee
Director, Transformational Convergence Technology Office
Defense Advanced Research Projects Agency

Michael Marron
Associate Director
National Center for Research Resources
National Institutes of Health

Brad Martin
Senior Computer Scientist
National Security Agency

Douglas Maughan
Program Manager
Department of Homeland Security

Andrew McLaughlin
Deputy Chief Technology Officer
Office of Science and Technology Policy

Patricia Muoio
Science and Technology Lead for Cyber
Office of the Director of National Intelligence

Beth Noveck
Deputy Chief Technology Officer for Open Government
Office of Science and Technology Policy

Shannin G. O'Neill
Government Accountability Office

Toby Sanders
Independent Consultant in Technology Education Strategy

Fred B. Schneider
Samuel B. Eckert Professor of Computer Science
Cornell University

Dahlia Sokolov
Staff Director
Research and Education Subcommittee
House of Representatives Committee on Science and Technology

Sylvia Spengler
Program Director
National Science Foundation

Peter Steenkiste
Professor
Carnegie Mellon University

George Strawn
Director, National Coordination Office
NITRD

John Trustman
Former CIO
Aetna Health Plans

Jeannette Wing
President's Professor of Computer Science and Department Head,
Computer Science Department
Carnegie Mellon University

Beverly Park Woolf
Research Professor
University of Massachusetts Amherst

Appendix B: Acknowledgements

PCAST wishes to express gratitude to the following individuals who contributed in various ways to the preparation of this report:

Samuel Blazek
Research Associate
Science and Technology Policy Institute

David Bray
Research Staff Member
Science and Technology Policy Institute

Jennifer Chen
Research Associate
Science and Technology Policy Institute

Chris Greer
Assistant Director for Information Technology R&D
Office of Science and Technology Policy

Nayanee Gupta
Research Staff Member
Science and Technology Policy Institute

Seth Jonas
Research Staff Member
Science and Technology Policy Institute

Bhavya Lal
Senior Research Staff Member
Science and Technology Policy Institute

David Lindley
Writer

Jennifer McGrady
D. E. Shaw Research

Mario Nunez
Research Associate
Science and Technology Policy Institute

Mark Shankar
Student Volunteer
PCAST

Edward Shyu
Research Associate
Science and Technology Policy Institute

Appendix C: Abbreviations Used in This Report

AHRQ	Agency for Healthcare Research and Quality
AP	Advanced Placement
ARRA	American Recovery and Reinvestment Act
BEA	U.S. Bureau of Economic Analysis
BLS	U.S. Bureau of Labor Statistics
CBP	Customs and Border Protection
CHDI	Community Health Data Initiative
CMS	Centers for Medicare and Medicaid Services
CTE	Career and Technical Education
CSTB	Computer Science and Telecommunications Board (NAS)
DARPA	Defense Advanced Research Projects Agency
DDR&E	Director of Defense Research and Engineering
DEA	Drug Enforcement Agency
DHS	U.S. Department of Homeland Security
DoD	U.S. Department of Defense
DoE	U.S. Department of Energy
DoJ	U.S. Department of Justice
DoT	U.S. Department of Transportation
ED	U.S. Department of Education
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FBI	Federal Bureau of Investigation
FCC	Federal Communications Commission
FLOPS	Floating point operations per second
GAO	Government Accountability Office
GPU	Graphics Processing Unit
HCI	Human Computer Interaction

HHS	U.S. Department of Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act
HPC	High Performance Computing
HPCC	High Performance Computing and Communication
HPCCI	High Performance Computing and Communications Initiative
HUD	Department of Housing and Urban Development
HVAC	Heating, Ventilating, and Air Conditioning
IARPA	Intelligence Advanced Research Projects Activity
I/O	input/output
IC	Integrated Circuit
ICE	Immigration and Customs Enforcement
IETF	Internet Engineering Task Force
MEMS	Microelectromechanical Systems
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NCO	National Coordination Office
NEON	National Ecological Observatory Network
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NIT	Networking and Information Technology
NITRD	Networking and Information Technology Research and Development
NLM	National Library of Medicine
NOAA	National Oceanic and Atmospheric Administration
NSA	National Security Agency
NSF	National Science Foundation
NSTC	National Science and Technology Council
NTIA	National Telecommunications and Information Administration
OECD	Organisation for Economic Development and Cooperation
OMB	Office of Management and Budget
ONC	Office of the National Coordinator for Health Information Technology

OOI	Ocean Observatories Initiative
OSD	Office of the Secretary of Defense
OSTP	Office of Science and Technology Policy
PCA	Program Component Area
PCAST	Presidents' Council of Advisors on Science and Technology
PITAC	President's Information Technology Advisory Committee
R&D	Research and Development
S&T	Science and Technology
SHARP	Strategic Health IT Advanced Projects
SSG	Senior Steering Group
STEM	Science, Technology, Engineering, and Mathematics
TSA	Transportation Security Agency
UAV	Unmanned aerial vehicle
USDA	U.S. Department of Agriculture
VHA	Veterans Health Administration
W3C	World Wide Web Consortium
WMD	Weapons of Mass Destruction

President’s Council of Advisors on Science and Technology

http://www.whitehouse.gov/ostp/pcast