REPORT TO THE PRESIDENT AND CONGRESS ENSURING LEADERSHIP IN FEDERALLY FUNDED RESEARCH AND DEVELOPMENT IN INFORMATION TECHNOLOGY

Executive Office of the President

President's Council of Advisors on Science and Technology

August 2015

REPORT TO THE PRESIDENT AND CONGRESS ENSURING LEADERSHIP IN FEDERALLY FUNDED RESEARCH AND DEVELOPMENT IN INFORMATION TECHNOLOGY

Executive Office of the President

President's Council of Advisors on Science and Technology

August 2015

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 often makes policy recommendations concerning, the full range of issues where understandings from the domains of science, technology, and innovation bear potentially on the policy choices before the President.

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

The President's Council of Advisors on Science and Technology

Co-Chairs

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

Eric S. Lander
President
Broad Institute of Harvard and MIT

Vice Chairs

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

Maxine Savitz
General Manager (retired)
Honeywell

Members

Wanda M. Austin
President and CEO
The Aerospace Corporation

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

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

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

Christine Cassel
President and CEO
National Quality Forum

Mark Gorenberg
Managing Member
Zetta Venture Partners

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

Michael McQuade
Senior Vice President for Science and Technology
United Technologies Corporation

Chad Mirkin
George B. Rathmann Professor of Chemistry
Director, International Institute for Nanotechnology
Northwestern University

Mario Molina
Distinguished Professor, Chemistry and Biochemistry
University of California, San Diego
Professor, Center for Atmospheric Sciences
Scripps Institution of Oceanography

Craig Mundie
President
Mundie Associates

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

Barbara Schaal
Dean of the Faculty of Arts and Sciences
Mary-Dell Chilton Distinguished Professor of Biology
Washington University of St. Louis

Eric Schmidt
Executive Chairman
Google, Inc.

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

Staff

Marjory S. Blumenthal
Executive Director

Jennifer L. Michael
Program Support Specialist

Diana E. Pankevich
AAAS Science & Technology Policy Fellow

Ashley Predith
Assistant Executive Director

PCAST NITRD Working Group

Working Group members participated in the preparation of an initial draft of this report. Those working group members who are not PCAST members are not responsible for, nor necessarily endorse, the final version of this report as modified and approved by PCAST.

Co-Chairs

Susan L. Graham*
Pehong Chen Distinguished Professor
Emerita in Electrical Engineering and
Computer Science
University of California, Berkeley

Gregory D. Hager
Mandell Bellmore Professor
Computer Science
John Hopkins University

Working Group Members

William J. Dally
Willard R. and Inez Kerr Bell Professor
Computer Science and Electrical
Engineering
Stanford University

Michael McQuade*
Senior Vice President for Science and
Technology
United Technologies Corporation

Eric Horvitz
Distinguished Scientist/Managing Director
Microsoft Research, Redmond

Eric Schmidt*
Executive Chairman
Google, Inc.

Sara Kiesler
Hillman Professor of Computer Science and
Human-Computer Interaction
Carnegie-Mellon University

Staff

Marjory S. Blumenthal
Executive Director

Ashley Predith
Assistant Executive Director

Jennifer L. Michael
Program Support Specialist

*Denotes PCAST Member

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

Research and development (R&D) in information technology (IT) has propelled successive waves of commercial success in IT and driven societal changes in ways unimaginable even a few decades ago. Access to digital information has become as important to Americans as access to the electricity system or the transportation system. Mobile technologies and cloud computing now enable Americans to interact with others, with ideas, and with the physical world in important new ways. And, as the United States and the world reap the rewards of greater connectedness, data, and computing power, the Nation has also new responsibilities to keep individuals and institutions safe from those with malicious intent. National investments in research and development are crucial to continued progress in all these areas.

By Executive Order 13539, your President's Council of Advisors on Science and Technology (PCAST) must periodically review the Federal Government's coordinated program of Networking and Information Technology R&D (NITRD). PCAST executes these duties as well as advises you and the Federal Government across all aspects of IT R&D on behalf of the President's Innovation and Technology Advisory Committee (PITAC). For this third evaluation of the NITRD program during your Administration, we convened a working group of seven experts from academia and industry and focused on eight critical research and development areas. The working group spoke with over 60 experts and agency representatives.

Our recommendations reflect the need for a refreshed R&D investment portfolio and coordination process given the pressing concerns of the IT ecosystem. Our recommendations for cybersecurity, for example, call for greater support of research methods that will give rise to systems with end-to-end security by construction. The coordination of the NITRD program has created an important R&D community in the Federal Government; the modernization of the topics, strategies, and working groups that the program prioritizes is essential to its future success.

The United States has been a leader in information technology R&D for decades, but that leadership is not untested. We believe that acting on these recommendations will undergird the Nation's IT R&D system and allow the United States to advance innovative ideas and exceptional talent for the next era of information technology.

Sincerely,

John P. Holdren
PCAST Co-Chair

Eric S. Lander
PCAST Co-Chair

Susan L. Graham
PITAC Co-Chair

Eric S. Schmidt
PITAC Co-Chair

Table of Contents

Executive Summary

Information technology (IT) drives the modern world. Nearly 80 percent of the households in the developed world have access to the Internet, and nearly half of the world is connected. Nearly every device – be it a car, a kitchen appliance, a device on the manufacturing floor, or a child's toy – is enhanced by IT. IT empowers scientific inquiry, space and Earth exploration, teaching and learning, consumer buying and selling, informed decision-making, national security, transportation, advanced manufacturing, and protection of the environment. The Bureau of Labor Statistics projects that more than half of all new science, technology, engineering, and mathematics (STEM) jobs will be related to information technology. It is difficult today to imagine a major economic sector, governmental function, or societal activity that does not directly or indirectly benefit from advances in information technology.

Today's advances rest on a strong base of research and development (R&D) created over many years of government and private investment. Because of these investments, the United States has a vibrant academia-industry-government ecosystem to support research and innovation in IT and to bring the results into practical use. It is essential that the Nation continue to invest in IT R&D and to steward those investments wisely in order to address important national goals in areas such as defense, economic prosperity and inclusion, health and human safety, education, and quality of life.

The High Performance Computing Act of 1991 (Public Law 102-194) established the Networking and Information Technology Research and Development (NITRD) program. The purpose of that Act and subsequent legislation is to coordinate the Federal investment in IT R&D to ensure continued United States leadership. This coordination is carried out by the NITRD Subcommittee of the National Science and Technology Council (NSTC). The coordinating efforts are meant to ensure the Nation takes full advantage of its IT investment and the program ensures a healthy portfolio as the next generation of research and technology evolve.

The legislation mandates a periodic review of the program's activities, progress, and direction. This report captures the output of the most recent review, performed by a working group consisting of members of the President's Council of Advisors for Science and Technology (PCAST) and outside experts from the fields covered by the NITRD program. The working group and PCAST consulted with government agencies, leading academic and national laboratory experts, industry leaders, and other stakeholders to assess the health and evolution of the forefront of research and development in IT. The working group also assessed the operation and effectiveness of the coordination activities carried out by the NITRD Subcommittee. This report summarizes our findings on the Federal Government's NITRD program, examines the state of research and development in IT fields, the preparation of the future IT workforce, and the coordination of IT activities in 18 member agencies and Federal entities by the NITRD Subcommittee of NSTC.

Research and Development in Information Technology

The working group and PCAST members chose to examine closely eight specific areas from a wider of set of topics of which all are critical to the future of IT: cybersecurity, health, Big Data and data-intensive computing, IT and the physical world, privacy protection, cyber-human systems, high-capability computing, and foundational computing research. Singly and in combination, these eight topics are germane to many national priorities. For example, Big Data, IT interactions with the physical world, and high-capability computing are essential contributors to addressing issues within energy and the environment; Big Data and cyber-human systems research play a major role in advances in STEM education; research in topics within IT and the physical world and in cyber-human systems are particularly important for advanced manufacturing.

Several common themes emerged from our review of the eight selected topic areas:

- Every area of IT spans a continuum from basic conceptual foundations; to system building, hypothesis testing, and experimentation; to innovative engineering; to real-world usage via first-mover applications; and finally to translation into common practice. Increased coordination and collaboration between fundamental research programs and mission work will facilitate the translation of fundamental research knowledge to application.
- Many of the important areas of IT R&D require large-scale shared infrastructure for computing and for data. These needs have been highlighted more extensively in earlier reviews, but persistent gaps remain and needs continue to grow.
- A growing number of academic research areas require access to production-quality platforms, large data sets, large-scale infrastructure, or large numbers of representative users – resources that are increasingly developed within the private sector.
- Research in IT is increasingly interdisciplinary, requiring larger and more diverse research teams and researchers who can create and lead multi-disciplinary research teams.
- Academia-government-industry partnerships are increasingly important.
- Researchers face continuing tension when choosing between short-term, problem-solving research and the riskier, more speculative long-term investigations. Funding pressures and publication practices in IT-related disciplines are making it more difficult to sustain the long-term research that is an essential component of a strong and balanced research ecosystem.
- As knowledge in certain research areas grows and the applicability of that area broadens, the demand for workers in that field increases faster than the education system can prepare workers. In some instances, academic research organizations compete with private-sector companies for skilled people, at all levels, including both potential graduate students and current faculty, putting our long-term research capabilities at risk by "eating our seed corn."

The report discusses each of the eight R&D areas in detail, and we summarize those findings here. Our recommendations are listed at the end of the Executive Summary, naming both especially relevant components and the departments in which they are housed when indicating responsibility for an activity.

Cybersecurity

Concern about the security of computing systems has existed for over forty years, and that concern has intensified with the widespread global interconnectedness enabled by the Internet. Cybersecurity attacks can disrupt the normal operation of computing systems (for example, by distributed denial of service), damage systems with computing components (for instance, by altering the computer control of physical devices), or cause the theft of proprietary, secret, or private information. Cybersecurity is often neglected in the early stages of the design and use of computing in new domains. In addition, human behavior, whether deliberate or inadvertent, can threaten the security of computing systems and must be considered as part of more robust cybersecurity solutions.

Federal investment in at least five key R&D areas will improve the foundations of cybersecurity: (1) cybersecurity by design – an understanding of how to construct secure and trustworthy systems; (2) defense against attack – ongoing mechanisms for authentication, authorization, data provenance, and integrity checks, as well as powerful tools to detect potential vulnerabilities automatically, for systems in use; (3) systems resilience – improved methods to mitigate the effects of an attack; (4) implementation support – methods to express cybersecurity policies formally in ways that are understandable both to people and to computers and tools to use them for policy implementation and compliance checking; and (5) better and faster methods for attribution, enabling both technical and non-technical mitigations.

IT and Health

A growing community of IT researchers, primarily with support from National Science Foundation (NSF) and National Institutes of Health (NIH), is actively developing technologies at the frontier of IT and health. These include the use of mobile devices and biometric technologies to support patient monitoring and coaching, new smart devices that augment human physical and intellectual capabilities, and new modeling methods that provide enhanced diagnosis or prediction of disease. Research is needed on new methods to reduce data complexity and provide actionable decision support to health practitioners, patients, and health care administrators. Many opportunities for research, however, are inhibited by significant barriers in gaining access to health data and the lack of standards to ensure inter-operability and promote technology and data exchange. Translation of new technologies into health care settings has lagged other areas where open standards, public data sets, and fewer regulatory hurdles allow for rapid innovation, testing, and evolution of technology that can have tremendous benefits in health care.

Big Data and Data-Intensive Computing

As computation becomes widespread and data generation increases dramatically, Big Data and data-centric computing play a central role in the vitality of the public and private sectors. Research is needed on error analyses and confidence measures for analyzing massive data sets, the determination of causality from data, better understanding neural network models and their construction, on machine learning in ways that consider associated larger decision making pipelines, and tools and methods that enable interactive data visualization and exploration. Attention should be paid to data stewardship to mitigate losses of data and corresponding losses of opportunities for machine learning, inference, and longitudinal studies.

IT and the Physical World

Research and commercial opportunities in systems that couple IT technologies with sensing and actuation – from the Internet of Things to smart infrastructure to robotics – are evolving rapidly. Researchers in robotics, artificial intelligence, cyber physical systems, and related areas will drive innovation in basic IT, as well as the development of robust, reliable autonomous systems that incorporate advanced sensing and sensors. As new products and technologies emerge for IT-enabled sensing and acting in the physical world, it is important to put in place open standards and platforms that encourage sharing of new technologies with and among the research community. There is a need for research focused on human interaction with systems that operate in the physical world, particularly around issues of safety, trust, and predictability of response. Additional research is needed in areas such as physical IT and human interaction, physical IT and robust autonomy, physical IT and sensing, development of hardware and software abstractions for physical IT systems; and trustworthy physical IT systems.

Privacy Protection

At the time of the 2013 NITRD review, privacy research programs were emerging as a focus. Since then, public awareness of digital privacy issues has heightened substantially. Privacy is an important human and societal value, and its protection is increasingly threatened by the growing amounts of online data, the increasing ability to combine and analyze that data, and the absence of adequate safeguards to keep private both original data and information derived from it. Research collaboration is needed among a broad range of computer scientists, government and legal scholars, behavioral and social scientists, and researchers in domains in which IT and human data are increasingly important to inform both the design and modification of computing systems and the drafting of policies and regulations. Technology should be developed so that the burden of privacy protection does not fall on the people being protected. Among the research challenges are understanding and clarifying what is meant by "privacy," automatically tracking the use of all forms of personal data; automating compliance checking; devising methods to use private data without disclosing private information; detecting, signaling, and mitigating information leakage and privacy violations as they occur; and creating mechanisms, frameworks, and tools to enable system builders to construct privacy-preserving systems without themselves being privacy experts.

Cyber-Human Systems

Computing is integral to Americans' work and personal lives and to the aims and processes of organizations, government, and society. The field of cyber-human systems has matured and broadened in scope, encompassing communication and coordination of people, computational systems and methods supported by people, rich collaborations between people and computational systems, and socially intelligent devices and systems, in addition to human use of computational services. Interagency coordination has been effective in some areas of cyber-human systems such as visualization and team science, but it has been sporadic in other areas such as social computing, human-robot interaction, privacy, health informatics, and human learning and education. Many aspects of cyber-human systems are not well understood and warrant research. Among them are computer-based learning as it relates to various socioeconomic groups; social media and networked communication and the effects on cognitive behavior; emerging "smart" consumer products or services and their social influences; human-machine collaboration and complementary problem solving; and development of ways of integrating Big Data analysis and traditional scientific method into new research pedagogies.

High-capability Computing

Since the High Performance Computing Act of 1991, the Nation has strived to develop and exploit extreme-scale computing, networking, and data management. The most capable systems (i.e., the largest, fastest, and most powerful) are critical to national defense, discovery-based research in all fields of endeavor, and commerce. In science and engineering research, high-end systems are scientific instruments much like particle accelerators or telescopes. Advances in the most capable systems lead to advances in smaller and more specialized systems as well. R&D progress has brought the field near to the fundamental physical limits of computer chips and to a state of ever-increasing complexity in software and computational design. New approaches are essential for all aspects of the design of high-capability systems, from hardware to applications programming. Innovations in the energy use of computer systems, programmability, runtime optimization, system software, and software tools are all needed. PCAST applauds the development of a National Strategic Computing Initiative, but emphasizes that it must lead to implementation of a long-term research program for high-capability computing that will also benefit data-intensive applications and IT research at lower demand levels.

Foundational IT Research

All of the paradigm-shifting achievements in information technology that people enjoy today rest on years, and sometimes decades, of basic research. Areas such as advances in computer architecture, domain-specific languages, algorithm development, scalable and reliable software systems, networking, machine learning, artificial intelligence technologies, and more will provide for the next generation of IT. Foundational long-term research in IT is essential for the application areas that build on it and for the future of the Nation's robust IT industry.

Government Support for Information Technology

For the Nation's worldwide leadership in IT to continue, it is essential that the government lead in the preparation of its people for participation in the IT workforce and that it continue to coordinate its investment in IT R&D.

Education and Training

To satisfy the growing need for IT expertise, education and training are needed at multiple levels, from highly skilled researchers and practitioners to users of conventional IT tools and methods. In order to have a healthy educational pipeline, a major effort is needed to attract young people to IT and to keep them from dropping out of the field. Talented young people need to be enticed to consider future careers in IT. Special efforts are needed to ensure that a large and diverse population of young students enters the pipeline and that the pipeline-leakage is minimized. Strong lifelong-learning programs are needed to help workers to keep up with technological change. Well-prepared teachers are essential to maintaining an educational pipeline. Although it is the States and the private sector that provide most education and training, it is essential that the Federal government lead in designing educational programs, tools, and technologies that enable IT learning and education, as well as continue to support programs at all levels.

NITRD Coordination

PCAST evaluated how the NITRD program coordinates the investments of the 18 member agencies and considered ways to make this cross-government R&D effort more efficient and effective. Investment levels in technical topics areas of interest are tracked for record keeping and budget analysis through Budgetary Program Component Areas (PCAs). The Office of Science and Technology Policy (OSTP) and the Office of Management and Budget (OMB) introduced PCAs in 1995, and many PCAs have remained virtually unchanged despite the evolution of the scientific and technical fields in IT. PCAST found strong support in OSTP and OMB to establish a process to update the PCAs periodically to ensure that PCAs align with current priorities in IT fields. This NITRD review report proposes eight new PCAs beginning in fiscal year (FY) 2017, and it recommends updating them every five or six years.

The NITRD program includes multi-agency focus groups ("Groups") that traditionally had been mapped to PCAs in the annual NITRD Supplement to the President's Budget.[1] In addition to calling for new PCAs, this report highlights the continued need to de-couple the NITRD Groups from the PCAs, so that the NITRD Groups can be modernized independently of the PCAs.

The review also found that NITRD leadership clearly manages coordination and information-sharing in the NITRD Subcommittee and Groups and that the NITRD Subcommittee and the National Coordinating Office (NCO) have established an important community in the Federal government. However, more aggressive leadership across the community is needed to lead change so that NITRD continues to achieve the strategic technical or policy vision that is in the best interests of the program. This review found very little understanding or documentation of the processes by which Groups of all kinds are created, evaluated, and retired. IT has clearly evolved over the last 25 years. New topics emerge and others decline in importance, and the NITRD organization needs to evolve in response. As a result of these findings, PCAST recommends changes to the Groups to help realign the coordination structure. Some Groups might have long lives, while other Groups should stand up and end in as short as six months or a year. Groups should be flexible for change.

In Summary

The United States has been a proud leader in IT for over seven decades and must continue that rich history of achievements and leadership. IT advances thus far have had a tremendous benefit to this country, its citizens, and the world. The implementation of the recommendations presented in this report will play a critical role in ensuring economic competitiveness, securing national security, enhancing quality of life, and improving the overall vitality of the Nation.

[1] Supplement to the President's Budget, https://www.nitrd.gov/Publications/SupplementsAll.aspx, 2016.

Recommendations

Research and Development

Recommendation 1: Cybersecurity

a. The National Science Foundation (NSF) should sponsor broad foundational research on methods to facilitate end-to-end construction of trustworthy systems, particularly for emerging application domains, and on ways to anticipate and defend against attacks, engaging not only computer science but also other engineering disciplines and behavioral and social science.

b. In coordination with this research program, the mission agencies – Department of Defense (DOD), National Security Agency (NSA), Department of Homeland Security (DHS), and Department of Energy (DOE), in particular, but also others – should sponsor both foundational and more applied mission-appropriate investigations of these topics.

c. The research sponsors should work closely with all agencies, including their own, and the private sector to facilitate the translation of the most promising research results into practice.

Recommendation 2: IT and Health

a. NSF, Health and Human Services (HHS), NIH, Defense Advanced Research Projects Agency (DARPA), DOD, and other agencies with responsibility for aspects of health care should continue to support foundational research in health IT. The National Science and Technology Council (NSTC) should continue to support coordination efforts such as the Health Information Technology Research and Development (HITRD) Senior Steering Group (SSG).

b. NSF, HHS, National Institutes of Health (NIH), and National Institute of Standards and Technology (NIST) should develop and nurture open interfaces, standards, and also incentives for promoting the leveraging of electronic health data in data analyses in support of biomedical research and in the delivery of health care.

c. NIH and HHS should create funding mechanisms that will encourage accelerated deployment, testing, and evolution of translational IT systems for clinical use.

Recommendation 3: Big Data & Data-Intensive Computing

a. NSF, in collaboration with mission agencies that collect large amounts of data for R&D, should continue to sponsor research on methods for performing inference, prediction, and other forms of analysis of data to advance all areas of science and engineering, and on methods for the collection, management, preservation, and use of data. Emphasis should be placed on formulating and disseminating methods for representing and propagating error analyses and confidence measures in large-scale data analysis; developing the theory and practice of computational and statistical methods for causal discovery from large data sets; developing deeper understandings of the foundations of neural network models and of systems challenges with scaling up these methods; uses of machine learning to guide decision making; and human understanding of large data sets and the results of their analysis.

b. NITRD, through its Big Data R&D SSG, should work to establish a common set of best practices and support structures for data capture, curation, management, and access. The NITRD Subcommittee of the National Science and Technology Council (NSTC) should encourage uniform adoption of these policies through the NITRD membership.

Recommendation 4: IT and the Physical World

a. NSF and DARPA should lead cross-disciplinary programs that will advance research and development of new approaches to robust autonomy, advance security and reliability of such systems, promote integrative approaches to human interaction, explore new sensing and interface technologies, and incentivize fundamental science on the cognitive and social aspects of interactive physical systems.

b. Mission agencies – particularly DOD, Department of Transportation (DOT), and NIH – and NIST should promote the development of open platforms and sharable infrastructure for research on physical systems within application domains – transportation, agriculture, urban infrastructure, health care, and defense.

Recommendation 5: Privacy

The Office of Science and Technology Policy (OSTP) and NSTC should continue to develop and expand a multi-agency research and development program to advance the science, engineering, policy, and social understanding of privacy protection. Agency participation should include at least NSF; NIH; the units of DOD, NSA, and DHS studying the extensions of cybersecurity R&D to encompass privacy; and other relevant units within HHS, NIST, DOT, and the Department of Education (ED).

Recommendation 6: Cyber-Human Systems

a. NSF should continue to broaden its support for fundamental research on the systems and science of the interplay of people and computing.

b. OSTP and the NITRD Subcommittee of NSTC should establish or strengthen coordination at both higher and lower levels among at least NSF, DOD, DARPA, NIH, and ED. In particular, coordination and support in areas such as social computing, human-robot interaction, privacy, and health-related aspects of human-computer systems should be enhanced.

Recommendation 7: High-Capability Computing for Discovery, Security, and Commerce

a. NSTC should lead an effort by NSF, DOE, DOD, NIH, member agencies of the Intelligence Community, and other relevant Federal agencies to implement a joint initiative for long-term, basic research based on the new National Strategic Computing Initiative aimed at developing fundamentally new approaches to high-capability computing. That research should be sufficiently broad that it encompasses not only modeling and simulation, but also data-intensive and communication-intensive application domains.

b. Under the leadership of OSTP, NSTC and the NCO should establish multi-agency coordination not only at the level of program managers, but also at the higher administrative levels reflected in SSGs.

Recommendation 8: Foundational IT Research

NSF should continue to invest in long-term foundational research in information technology. Other NITRD agencies, including DARPA, IARPA, DOE, and NIH, should support foundational research in those aspects of IT that most affect their missions.

Government Support for Information Technology

Recommendation 9: Education and Training

- **a.** The NITRD Subcommittee, working in partnership with NSF, ED, and the private sector, should create new educational opportunities in IT at all levels, beginning with K-12, to grow the pipeline of skilled workers and identify future innovators and leaders. These programs should incorporate approaches that will engage under-represented populations.
- **b.** As part of that effort, NSF should lead the development and implementation of model programs for pre-college students that attract the most talented young people to study IT. These will be the future innovators and leaders. The program should be designed to address differences in gender, economic status, and cultural background, and to collaborate with industry to provide resources to expand these programs broadly in the U.S. education system.
- **c.** NSF and ED should create programs for training and retraining of workers at all age levels with the goal of providing targeted "on-ramps" for those individuals to develop careers in the IT industry. They should fund research that includes 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 and with external partners to deploy these programs and capture data on performance and outcomes.

Recommendation 10: NITRD Program Component Areas

- **a.** OSTP, NCO, and the NITRD Subcommittee, in collaboration with OMB, should revise the PCAs for the FY 2017 Budget cycle and beyond to reflect both the current nature of IT and the major national priorities in which IT plays a major role. Chapter 4 suggests revisions to the PCAs.
- **b.** Those four stakeholders should create a process to review the PCAs every five to six years and implement proposed modifications. PCAST or its PITAC subcommittee should provide recommendations for changes to the PCAs.

Recommendation 11: NITRD Groups

- **a.** The NITRD Subcommittee, in collaboration with the NSTC and OSTP, should establish specific language specifying what the purpose of each *type* of Group is and what mechanisms should be used to establish, monitor and terminate a Group. They should define a process to create a new Group, set its charter, and specify its correspondences with existing PCAs.
- **b.** The NITRD Subcommittee, in collaboration with NCO and OSTP, should define a process and timeline for periodic review of each Group, with a recommendation for continuation, modification, or sunset. A

process for acting on those recommendations should be defined and executed. Reports on these reviews should be provided for each NITRD Review.

c. Each Group at the Senior Steering level should coordinate a process to publish and publicly discuss periodically a research and coordination plan for its area of interest.

1. Introduction

Information technology (IT) research and development (R&D) has fueled an exploding frontier of new innovations and ideas that shape the world today. Across science, commerce, and society at large, innovations based on fundamental IT research have emerged at an astounding pace for more than half of a century, driving the expansion or reinvention of innumerable existing services and practices and creating entirely new disciplines and industries. Indeed, it is difficult today to imagine a major economic sector, governmental function, or societal activity that does not directly or indirectly rely on advances in IT R&D.

The High Performance Computing Act of 1991 (Public Law 102-194) established the Networking and Information Technology Research and Development (NITRD) program, "To provide for a coordinated Federal program to ensure continued United States leadership in high-performance computing."[2] The Act was amended by the Next Generation Internet Research Act of 1998 (P.L. 105-305) and the America COMPETES Act of 2007 (P.L. 110-69). The legislation calls for a periodic review of the NITRD program.

This review of the NITRD program has three components. First, it provides findings on the progress observed since PCAST's prior reviews in 2010 and 2013. The discussion of the responses to previous recommendations is contained in context in subsequent chapters of the report. Second, it highlights some of the new themes that have gained in importance in recent years and provides findings and recommendations related to those themes. Last, it provides an extensive discussion of, and recommendations for, coordination of the NITRD program within the Federal Government. These recommendations reflect the evolution of the field and are intended to provide the NITRD program with the flexibility to evolve its activities and organization as the field itself evolves.

The remainder of the report is organized as follows. Chapter 2 describes the state of the field of IT,[3] first from the perspective of society, then from a science and technology perspective, and finally from a Federal Government perspective. Chapter 3 discusses the current research agenda – the major research questions, the ways in which the NITRD program is addressing them, and the ways in which it is not. That chapter contains findings and recommendations for IT R&D. Chapter 4 addresses coordination of the NITRD program and contains findings and recommendations concerning the Government's role.

[2] The NITRD program has undergone various name changes. The 1991 legislation established the National High Performance Computing Program, which became the High Performance Computing and Communications (HPCC) Program the following year. In 1997 it became the Computing, Information, and Communications Program (CIC), was renamed the Information Technology Research and Development (IT R&D) Program in 2000, and became the Networking and Information Technology Research and Development (NITRD) program in 2005.

[3] The term "information technology" normally includes networking. In this report, we use "NITRD" to refer to the Federal Government program, but "information technology" or "IT" to refer to the fields of computing and telecommunications, encompassing computer science, computer engineering, bioinformatics, information science, robotics, computational science, and the like.

2. Evolution of Information Technology

The Ever-Increasing Impact of Information Technology

Information technology (IT) drives the modern world. Nearly 80 percent of households in the developed world have access to the Internet, and nearly half of the world is connected.[4] Roughly two-thirds of Americans own a smart phone,[5] and the use of the voice-call feature is an ever-smaller fraction of the function of that device. Almost two billion photos are updated on the Internet every day[6] – a particularly astonishing number when considering that there are "only" about three billion Internet users worldwide.[7] Nearly every device – whether a car, a kitchen appliance, a device on the manufacturing floor, or a child's toy – is now powered by IT. IT is fundamental to the ways in which people understand the world, powering scientific inquiry, space exploration, and knowledge about the physical spaces people inhabit. The Bureau of Labor Statistics projects that more than half of all new STEM jobs will be related to information technology (Figure 1).[8]

The positive outlook for IT must be balanced against the growing wave of vulnerabilities to this ecosystem. The number of active, organized Internet threat groups has grown by a factor of four in the past five years. It is estimated that 95 percent of all networks are compromised in some way.[9] Public disclosure of the theft of private information through security breaches has become an almost weekly occurrence, and the unreported (or undetected) losses are quite possibly far larger. Less obvious, but equally important, are the lost opportunities for innovation due to an inadequate supply of trained IT specialists at all levels.

These trends create enormous new opportunities and challenges for computing research. As the Worldwide Web scales, new ideas and fundamental science are needed to grow bandwidth, reliability, and security of networking and storage technologies. As more of the population interacts with computing, there is a growing need to enhance an understanding of ways to enhance and protect human interaction with IT-enabled systems, and the ways IT-enabled devices enhance human existence. As computing increasingly supports systems in the physical world, research to enhance the reliability, security, and safety of such systems will be essential to

[4] (1) The official World Wide Web Anniversary Site, World Wide Web Consortium (W3C), World Wide Web Foundation "Internet Live Stats," http://www.internetlivestats.com/internet-users/ ; (2) ICT Data and Statistics Division, Telecommunications Development Bureau, International Telecommunications Union "ICT Facts and Figures," http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf, 2014.
[5] Smith, Aaron "U.S. Smartphone Use in 2015," *Pew Research Center, Internet, Science &Tech* http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/, 2015.
[6] Edwards, Jim "Planet Selfie: We're Now Posting A Staggering 1.8 Billion Photos Every Day," *Business Insider*, http://www.businessinsider.com/were-now-posting-a-staggering-18-billion-photos-to-social-media-every-day-2014-5, 2014.
[7] Computing Research Association. The world population is approximately 7.3 billion people. Roughly 5.4 billion people are over the age of 14.
[8] The Bureau of Labor Statistics projections were presented and discussed in: ACM Education Policy Committee, "Rebooting the Pathways to Success: Preparing Students for Computing Workforce Needs in the United States," http://pathways.acm.org/ACM_pathways_report.pdf, 2014.
[9] Meeker, Mary, "Internet Trends 2014- Code Conference," *Business Insider,* http://www.businessinsider.com/mary-meekers-2014-internet-presentation-2014-5#-1, 2014.

ensure public safety and trust in the world that surrounds them. As scientific discovery increasingly exploits the availability of vast quantities of data, research to attain ever-greater levels of performance is imperative.

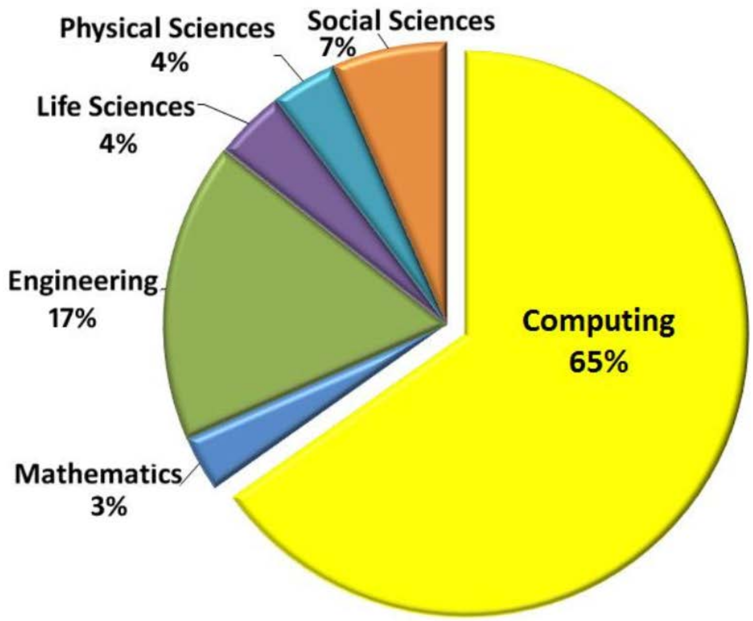

Figure 1. Projected distribution of growth in newly-created STEM job openings by field from 2012 to 2022. Jobs data are calculated from the Bureau of Labor Statistics Employment Projections 2012-2022 (http://www.bls.gov/emp). Image courtesy of the Association for Computing Machinery.

It is important to note that all of these trends are taking place in a world that, because it is ever more interconnected, is also increasingly globally competitive. The United States no longer leads in many important areas of IT. By some measures, the Nation no longer has the fastest supercomputer. There are no longer any major industrial robotics companies in the United States. There is rapidly growing competition to Google, Facebook, and Twitter in Asia. Anyone, anywhere in the world can access the best teachers in the world by taking courses online, access incredible resources through cloud computing vendors, and deploy a new system using "app stores." National defense, economic well-being, and quality of life depends on continued, vigilant, and wise investment in long-term IT research.

Continuing Research Progress in Information Technology

When the High Performance Computing and Communications (HPCC) program (now the NITRD program) was introduced in 1991, much of computing research, particularly at the high end, focused on advances in computing systems themselves. Computing research was organized around three large areas – computing hardware and systems, software, and networking. Software systems research encompassed such topics as operating systems, programming languages, and software engineering. Software research also included uses of computing systems in areas such as graphics, artificial intelligence (AI), simulation, databases, and the like, but that research was often in one general "applications" bucket.

As the field continued to move forward, the division between hardware and software weakened. Researchers came to consider computing systems, including not only conventional computers but also more specialized devices such as robots, as integrated hardware/software systems. Software became so pervasive a component that it was viewed increasingly in the context of its functionality, not as an abstract entity of its own.

The role of networking changed as well. Envisioned originally as a means to connect computers, computer networks increasingly connected people, first through the spread of email, and then, as personal computers and the World Wide Web entered the scene, as a vehicle for shared information. The roles of the human users changed from specialists to a more diverse user community. The modes of interaction broadened from terminal to workstation to personal computer and from black-and-white text to richer visual and audible outputs.

The enormous expansion in computing capability, in the form of increased computing speed, storage capacity and network interconnectedness, together with enhanced collection of observational data in digital form, led to greater importance for gathering, storing, and analyzing data. Research focusing on data became a major theme in many computing research organizations. The overall research agenda changed from networking and computing to networking and information technology as the role of "applications" grew.

In the third major transition, as computing has become more mobile, cheaper, and more highly commoditized, and as wireless networking technologies have become ubiquitous, there has been an evolution toward a far more pervasive and intimate interaction among computing, people, and the physical world (Figure 2).

There is now a vibrant and growing commercial ecosystem around mobile devices carried by people, and a growing discussion of "cyber-human systems" that consist of large-scale network devices that allow people to communicate, collaborate, and carry out activities of normal life through online platforms. A similar evolution is occurring in the expanding sphere of "smart systems" that operate in the physical world. In the 1990s, high-performance computers modeled the physical world. Now computers also sense the physical world and control physical systems.

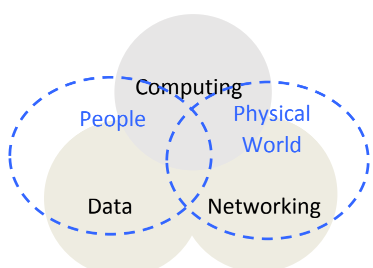

Figure 2. The integration of IT components with each other and with people and the physical world

As IT devices have gotten smaller, IT systems have gotten bigger. The emerging "supercomputers" that motivated the NITRD program in 1991 have now become complex computing systems that are approaching exascale performance (10^{18} operations per second). A large commercial data center may have tens of thousands of servers managing petabytes of data.

Along with those transitions, the domains in which IT is used have become preeminent in IT R&D (Figure 3). The integration of IT with people and the physical world is central to achievements in security, privacy, health, transportation, manufacturing, robotics, societal computing, smart infrastructure, and scientific discovery, addressing many of the most important national priorities. Some of those domains, notably scientific discovery and security, have been IT-enabled for a long time. Indeed they were the genesis of the HPCC program. Other domains, among them cybersecurity, trusted systems, and privacy, have increased in priority as a consequence of new capabilities enabled by IT. And many domains represent new opportunities – the emergence of autonomous vehicles that will serve an aging population or provide access to hostile territory, the opportunity to understand medical conditions and treatments at a greater depth than ever before, the ability to manufacture products efficiently and competitively in the United States, and many more. Advances in many of the same technologies that led to the HPCC program nearly 25 years ago – hardware and computer architecture, software, and networking – remain important in enabling those domains.

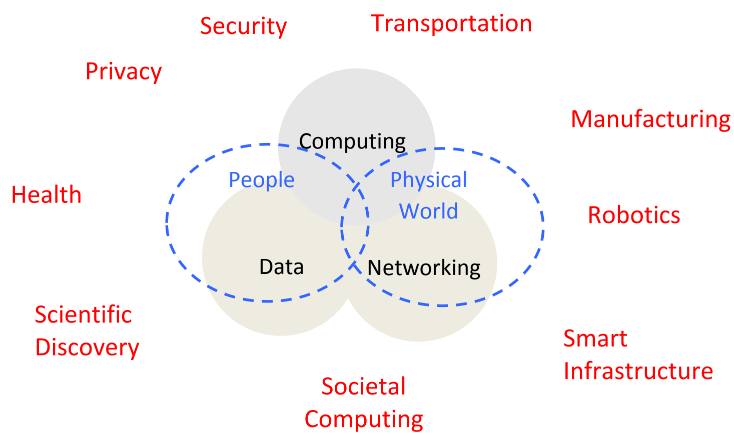

Figure 3. The integration of IT components, people, and the physical world in the context of proliferating uses

The Importance of Sustaining Government Investments in IT R&D

The enormous progress in IT-enabled domains has resulted from both incremental and disruptive innovations, virtually all of which drew on a substantial knowledge base built over many years of research and experience. The National Research Council (NRC) has documented the importance of long-term research in creating many of the most successful IT-based industries.[10] Reports from the NRC's Computer Science and Telecommunications Board (CSTB) have also illustrated that years of research are sometimes needed for a novel idea to mature to the point that it is commercially viable.

The U.S. ecosystem of academia-industry-government collaboration in R&D is the envy of the world, and many countries seek to emulate the basis of its success. In that ecosystem, particularly as it applies to IT, all three partners play a role. It is the nature of the commercial private sector that its focus is on bringing products to market, creating markets, and making money. To achieve those goals, its investments in R&D tend to be on short-term research and advanced development. Even companies with well-respected research divisions often turn to colleagues in academia to discover the science and technology results that are important for future products. The role of academia is to advance knowledge and to educate people. Advancing knowledge is a long-term endeavor, in which studying topics for the sake of a deeper understanding is coupled with discovery that is in some sense useful. Academia helps to transfer that knowledge to the private sector when students graduate and take jobs in industry.

[10]Computer Science and Telecommunications Board; "Continuing Innovation in Information Technology," *National Academies Press,* http://www.nap.edu/catalog/13427/continuing-innovation-in-information-technology 2012.

What is the government's role in this ecosystem? The government is a consumer of research. For example, many of the highest capability systems enabled by over 70 years of R&D are owned by the Federal Government. The Federal Government also articulates challenges of national importance and sets priorities for addressing those problems. And the Federal Government invests in finding solutions to those challenges. Given the relatively short-term focus of industry and the opportunity for relatively long-term investigation in academia, together with the preparation of an educated workforce, Federal support of academic research is a crucial part of the ecosystem.

Talented and skilled people are one of the most important resources in maintaining and growing national successes in the IT research enterprise. Academia, industry, and government all draw from that talent pool. The success of the ecosystem depends on the success of all three sectors in attracting those people. In academia, attracting and retaining talented faculty researchers, in particular, is facilitated by adequate and stable research funding, access to the infrastructure needed for research, and help in handling growth in numbers of students seeking IT education and training. Government success depends on attracting and retaining strong leadership in managing research funding programs, which is facilitated by making those jobs attractive to potential and current employees. The success of the entire ecosystem depends on growing the workforce by motivating and preparing more people to enter it.

3. Research and Development in Information Technology

Introduction

The Nation invests in information technology R&D in order to develop the capabilities necessary to tackle problems of national importance, such as defense, economic prosperity and inclusion, health and human safety, education, and quality of life. Recent NITRD reviews have examined many topics that advance the Nation's collective expertise in IT and its ability to use advances in IT to achieve America's priorities. This review highlights eight areas of information technology R&D that are of major importance at the present time. Two of those areas, health and cybersecurity, are among our most important national priorities. Two of them, Big Data and IT interaction with the physical world, encompass recent Government initiatives to highlighting cross-society opportunities; two others, privacy and cyber-human systems, exemplify the deepening integration of people with information technology. The final two, high-capability computing and foundational IT research, reflect the technology base on which all of the uses of IT depend.

Many of the Nation's priorities can benefit from some or all of these eight areas. For example, Big Data, IT interactions with the physical world, and high-capability computing are essential contributors to addressing issues involving energy and the environment. Big Data and cyber-human systems research play a major role in advances in STEM education. Research in topics within IT and the physical world and in cyber-human systems are particularly important for advanced manufacturing.

Cybersecurity and privacy are often combined in discussions of IT, but they are separated in this report (although NITRD coordination of those topics takes place largely within the same organizational structures). Concern about the security of computing systems has existed for over 40 years. That concern has intensified with the widespread global interconnectedness enabled by the Internet. More recently, there has been increasing awareness of the privacy concerns raised by the abundance of personal information in digital form and the growth in all forms of online interaction. Cybersecurity protects personal information from theft, but it also protects information, services, and systems that have little to do with privacy. Threats to privacy come not only from information leakage, but also from inappropriate use of both protected and publically available information in electronic form. Cybersecurity and privacy protection have different but overlapping research agendas.

Some areas that were singled out in earlier NITRD reviews are now incorporated within one of the eight topics described herein. For example, software is a crucial component in any use of information technology and no topic can be treated without consideration of its software aspects, but this report does not include a separate section on software. The coverage of topics in this report is selective, not comprehensive – the decision to not highlight any particular topic in this review does not signal a lack of importance or opportunity to do more in that area.

In preparing this review and consulting with leading experts in the various fields, several common themes were noted:

- Every area of IT spans a continuum from basic conceptual foundations to system-building, hypothesis-testing, and experimentation; to innovative engineering; to real-world usage via first-mover applications; and finally to translation into common practice. The NITRD program primarily supports the earlier stages of that continuum, but without attention to translation to practice the Nation runs the risk of delaying the benefits of its R&D investments and the iterative cycle of feedback from real-world engagement and experiences. Increased coordination and collaboration between fundamental research programs and mission agencies is an important means to facilitate that translation to practice.

- Many of the important areas of IT R&D require large-scale shared infrastructure for computing and for data. Infrastructure needs are discussed in some of the subsequent sections – these needs have been highlighted more extensively in earlier reviews, but persistent gaps remain.

- A growing number of academic research areas require access to production-quality platforms, large data sets, large-scale infrastructure, or large numbers of representative users – resources that are increasingly developed within the private sector. If researchers have only limited access to these resources, they will be forced to replicate versions of commercially available products or services in order to do representative research. Limited funding to create such resources deters some researchers from those areas and can lead to "toy" experiments that limit scaling up or working in naturalistic settings.[11] Academia-industry partnerships sometimes occur, but they may not be possible when the research is long-term or counter-normative.

- Research in IT is increasingly interdisciplinary, requiring researchers who can create and lead effective multi-disciplinary research teams. Technological advances have made such research collaboration easier but not necessarily more efficient or more effective.[12] New approaches to training that balance cross-disciplinary training with depth and rigor in individual fields are needed.

- Academia-government-industry partnerships (sometimes termed public-private partnerships) continue to evolve. At their best, each partner provides something unique – industry offers insight into real-world problems, access to infrastructure, and some funding; government provides financial and community support for research; and academia contributes foundational research, synergistic education, and a cadre of expertise.

- There is continuing tension between short-term, problem-solving research and the riskier, more speculative long-term investigations. Funding pressures and publication pressures are making it more difficult to sustain the long-term research that is an essential component of a strong research ecosystem.

- As knowledge in certain areas grows and the applicability of such areas broadens, workforce needs increase faster than the supply of the educated workers can increase. Those pressures are particularly

---

[11] See, for example, Bernstein, M. S., Ackerman, M. S., Chi, E. H., & Miller, R. "The Trouble with Social Computing Systems Research," *In CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11). ACM, New York, NY, USA, 389-398.* http://doi.acm.org/10.1145/1979742.1979618, 2011.

[12] Board on Behavioral, Cognitive, and Sensory Sciences, "Enhancing the Effectiveness of Team Science" *The National Academies Press* http://sites.nationalacademies.org/DBASSE/BBCSS/Enhancing_Effectiveness_of_Team_Science/index.htm, 2015.

> strong in data analytics and cybersecurity. In some instances, academic research organizations compete with private-sector companies for skilled people, at all levels, including both potential graduate students and current faculty, putting our long-term research capabilities at risk by "eating our seed corn."

In the remainder of this chapter, we review our findings and recommendations for IT research.

Cybersecurity

Both the frequency and the severity of attacks on computer systems continue to increase. Cyberattacks can disrupt the normal operation of computing systems (for example, by distributed denial of service), damage systems with computing components (for instance, by altering the computer control of physical devices), or compromise proprietary, secret, or private information. Such attacks threaten national security, economic well-being, and quality of life.

The U.S. government is well aware of the importance of protecting against cyberattacks and has various protection programs (aimed, in particular, at governmental and critical-infrastructure systems). In addition, multiple Federal agencies invest in cybersecurity R&D. The NITRD program includes two interagency groups, the Cyber Security and Information Assurance (CSIA) SSG and the CSIA interagency coordinating group. In 2011, the NITRD Subcommittee of the National Science and Technology Council prepared a strategic plan for cybersecurity R&D.[13] The plan is currently being revised.[14] There is communication and, as appropriate, coordination on cybersecurity between NITRD and another NSTC component, the Special Cyber Operations Research and Engineering (SCORE) subcommittee, which addresses the classified research needs of U.S. national security and intelligence agencies, such as the National Security Agency.[15] In 2013, PCAST reported more broadly on the issues surrounding cybersecurity.[16]

Discussions of the Federal investment in cybersecurity R&D appear in both the 2010 and 2013 NITRD reviews. The 2010 report recommended that "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" and listed some particular areas of focus. The 2013 report observed that cybersecurity R&D continued to be important and that, while noticeable progress had been made on interagency coordination since 2010, the topic remained a critical focal point. The report said, "Continued emphasis and even greater coordination is recommended."

---

[13] National Science and Technology Council, Executive Office of the President, "Trustworthy Cyberspace: Strategic Plan for the Federal Cybersecurity Research and Development Program," https://www.whitehouse.gov/sites/default/files/microsites/ostp/fed_cybersecurity_rd_strategic_plan_2011.pdf, 2011.

[14] In addition to the intrinsic value of updating the 2011 plan, the Cybersecurity Enhancement Act of 2014 calls for NITRD to provide Congress with a Federal cybersecurity R&D strategic plan by mid-December 2015.

[15] Whereas the NITRD subcommittee falls under the NSTC Committee on Technology, the SCORE subcommittee falls under the Committee on Homeland and National Security.

[16] President's Council of Advisors on Science and Technology, Executive Office of the President, "Report to the President: Immediate Opportunities for Strengthening the Nation's Cybersecurity," "http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_cybersecurity_nov-2013.pdf, 2013.

Reviews of the literature and discussions with experts lead to a variety of findings:

Insecure systems are easy targets for terrorists, strategic adversaries, and economic competitors, with a wide range of damaging consequences. Current research in cybersecurity encompasses both work on identifying potential security threats (e.g., by attacking existing systems to discover weaknesses) and work to develop more secure systems. The latter includes system designs that provide new or enhanced functionality using a growing set of tools and techniques, as well as some fundamental work on new methods for encrypting data and computing with encrypted data. It is still the case, however, that many new systems are designed and launched with inadequate attention given to cybersecurity. Reducing code vulnerabilities from the outset should have as high a priority as eliminating functionality bugs or adding features.

Finding: Improving the security of computing systems continues to be essential to the nation's safety, economic growth, and overall wellbeing.

The need for cybersecurity becomes particularly acute as new generations of IT-enabled physical systems are developed. For example, vulnerabilities are reported to exist in aspects of roadway safety and communicating vehicles, smart buildings, emergency management, and medical devices.[17]

Finding: Cybersecurity has often been neglected in the early stages of the design and use of computing in new domains.

Additionally, recent widely publicized incidents illustrate that even well engineered systems are prone to infiltration due to the lack of attention to the most basic of best practices in managing computing systems. For example, system administrators fail to install updates, while users fail to use strong passwords or log out of publicly accessible computers.

Finding: Human behavior, whether deliberate or inadvertent, can threaten the security of computing systems and must be considered as part of more robust cybersecurity solutions.

Following are some research areas that PCAST suggests for emphasis in the next few years. The list is illustrative but not complete. There are two kinds of topics. One kind deals with deepening the understanding of what constitutes a trustworthy or secure system – what properties it should have and how a system with those properties could be constructed. The second set of topics deals with existing imperfect systems and human fallibility. These topics address questions of detection of attack, defenses against attack, and improvements to robustness and resilience of existing systems. In many instances, research in cyber-human systems (discussed in a later section) can play a key role and should be addressed as part of the research.

- Cybersecurity by design – an understanding of how to construct secure and trustworthy systems. That understanding will require the ability to express the properties that a trustworthy system must have while taking the context of its use into account; the techniques to analyze the trustworthiness of

[17] (1) Miller, C., and Valasek, C, "A Survey of Remote Automotive Attack Surfaces," http://www.scribd.com/doc/236073361/Survey-of-Remote-Attack-Surfaces 2014; (2)"Security and the Internet of Things" - *Computer Science Zone*, http://www.computersciencezone.org/security-internet-of-things/ 2014.
(2) Kramer, D, B.; Baker, M.; Ransford, B.; Molina-Markham, A.; Stewart, Q.; Fu, K.; Reynolds, M, R. "Security and Privacy Qualities of Medical Devices: An Analysis of FDA Postmarket Surveillance" *PLOS ONE*, http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0040200, 2012.

components and entire systems; the ability to consider systems as a whole, from hardware, to communication, to systems software, to applications, to human interaction; and methods for composing secure components and preserving trustworthiness.

- Defense against attack – as systems are in use, they need ongoing mechanisms for authentication, authorization, data provenance-tracking, and integrity checks, as well as powerful tools to automatically detect potential vulnerabilities. Cryptography continues to be an important component of such defense, but it must be complemented by effective systems designs. Defenses must encompass both external and internal attacks.
- Systems resilience – improved methods to mitigate the effects of an attack. Examples include tools and methods that continuously monitor and measure systems for degraded or compromised performance, methods that provide resilience by allowing a system to operate reliably in the presence of failures at any level of the system stack, and analyses that identify and disable possibly compromised system components as soon as the compromises occur. Although cataloging attacks is valuable, and imagining them can be instructive for workforce training, it is learning about the vulnerabilities exposed by an attack that should be the focus of research efforts.
- Implementation support – Methods to express cybersecurity policies formally in ways that are understandable both to people and to computers. Associated tools and methods to help creators of novel systems in new application domains to integrate cybersecurity from the beginning, without themselves being cybersecurity experts. Software tools to determine compliance with policies are also needed.
- Better and faster methods for attribution (for identifying the source of an attack), so that both technical and non-technical mitigations are possible.

Recommendation 1: Cybersecurity

a. NSF should sponsor broad foundational research on methods to facilitate end-to-end construction of trustworthy systems, particularly for emerging application domains, and on ways to anticipate and defend against attacks, engaging not only computer science but also other engineering disciplines and behavioral and social science.

b. In coordination with this research program, the mission agencies – DOD, DHS, NSA, and DOE, in particular, but also others – should sponsor both foundational and more applied mission-appropriate investigations of these topics.

c. The research sponsors should work closely with all agencies, including their own, and the private sector to facilitate the translation of the most promising research results into practice.

IT and Health

The 2010 and 2013 PCAST NITRD reports both highlighted the importance of IT to human health and health care. The 2010 PCAST NITRD report called for harnessing IT to enable comprehensive lifelong records, new kinds of health-related empowerment through provision of knowledge, and new tools for informed personalized health care and called out opportunities for greater coordination among agencies. The 2013 PCAST NITRD report found that significant progress had been made, including coordination by the NITRD Health Information Technology Research and Development (HITRD) SSG, but that more research is needed.

The opportunities for high-impact research and development continue to grow and evolve, and there are signs of increasing Federal investment in this area, as well as a growing community of researchers and practitioners. Of particular note are President Obama's January 2015 announcement of the Precision Medicine Initiative, the appointment by NIH of a Director of Data Science, the launching of the NIH "Big Data to Knowledge" (BD2K) program, and the continued support and evolution of the NSF-led Smart and Connected Health program. These investments are complemented by investments by DARPA and DOD on military health-related programs and funding from cyber-physical systems, the National Robotics Initiative, NSF Frontiers, and other Federal programs that include projects with medical applications as their focus. The former HITRD SSG enhanced the visibility and communication across agencies in this important area.

Finding: There is a growing community of IT researchers who are actively developing technologies at the frontier of IT and health care, primarily with support from NSF and NIH.

While progress has been made through Federal incentive programs on introducing electronic health records into hospitals over the last eight years, much of the data collected by these systems remains disjoint and untapped for machine learning and inference. Existing data resources for research, such as the large amounts of data collected and stored electronically by the Department of Veterans Affairs (VA) hospital system, are generally difficult for researchers to gain access to for studies. Some impediments are based in the Health Insurance Portability and Accountability Act (HIPAA) privacy constraints, others in procedural mechanisms for providing data access to external researchers (e.g., non-VA personnel). Creating well-constructed toolkits and frameworks that would allow institutions to create data sets and support distributed processing of data without the need for explicit sharing would enable advances in computing and data analytics to be applied to health in new ways. Currently, there are few vehicles to advance these objectives, leaving every institution to chart its own course.

The potential for leveraging the data assets being stored in electronic records is immense, including opportunities for large-scale research spanning multiple systems, and for real-time hospital-centric health care delivery. There is the promise of weaving together heterogeneous data sets, including genomics, imaging, physiological, family history, interventional care, and post-treatment activities to learn about health and wellness, and to guide decisions about clinical care and biomedical research. Development of research platforms that enable machine learning from hospital data, and methods for protecting patient identity when using patient data for studies would stimulate research on machine learning, inference, and decision support in health care.

Finding: Barriers in gaining access to health data for research purposes and insufficient data interoperability inhibit many opportunities for health IT research.

Another opportunity comes from marrying new, emerging mobile-health applications with traditional patient records and treatment. Despite enthusiasm for possibilities for harnessing data from mobile devices, little is known about how basic biometrics such as movement, galvanic skin response, and heart-rate level and rate variability, at both the individual and population level, might influence diagnosis, treatment, or the study of outcomes. More basic research is needed at the intersection of IT and health care research about the potential for harnessing readily available data from wearable sensors in screening, diagnosis, and management of illness. Other opportunities include developing better links between traditional electronic health record systems and devices and cloud services that capture data. Supplying actionable information at the right place and time promises to improve the patient and caregiver experience greatly. Applications for the chronically ill may enhance adherence to treatment regimes, reduce hospitalizations and readmissions, and facilitate new studies of disease and treatment.

Finding: New forms of data and new IT platforms are becoming available, but few standards exist that ensure interoperability and promote technology and data exchange. Health care innovation is also limited by the ability to deploy, test, and evolve new technologies rapidly. Most hospital infrastructures and workflows do not readily admit the introduction of new technology for the purposes of testing and evaluation. Many of these innovations do not fit naturally into either clinical or basic science methodologies, making it difficult to acquire the resources necessary to translate work into clinical settings. Absent such testing and evaluation, private investment is difficult to attract. As a result, many promising technologies that would quickly advance in non-health care settings are left unrealized.

Finding: Translation of new technologies into health care settings has lagged other areas where open standards, public data sets, and fewer regulatory hurdles allow for rapid innovation, testing, and evolution of technology.

The following are research areas that PCAST suggests should be emphasized in the coming years:

- The development of new systematic science approaches that will support study of treatments and outcomes of individual and groups of patients as complements and as replacements to randomized controlled trials. As the number of experimental treatment options grows, and the precision with which identifying relevant patient populations increases, the ability to make more accurate and informed hypotheses will increase the speed and cost-effectiveness with which new diagnostic and therapeutic tools are developed.
- The exploitation of advances in mobile and biometric technologies to create new approaches to patient monitoring, outpatient care, and patient-provider interaction. Such systems have immense potential for advancing care in particular for the chronically ill or infirm.
- The development of new technologies and methods for enhancing care as well as relating the performance of care to patient outcomes. For example, advanced surgical robots, smart interventional systems, or patient-focused and provider-focused sensing technologies can fuel new systems and workflow models that can inform the "learning health care system" discussed in recent reports from the Institute of Medicine.[18]

[18] Learning Health System Series, *Institute of Medicine* http://www.nap.edu/catalog/13301/the-learning-health-system-series, 2007-2011.

- Technologies, platforms, and programs to acquire, curate, and share data while incorporating strong protections on the inappropriate use of personal data and adhering to regulations for privacy and human-subjects research. Current approaches to health care innovation are limited by the size and quality of data available for identifying problems, creating solutions, and assessing the results of new innovations.
- Methods to reduce data complexity and provide actionable decision support to health practitioners, patients, and health care administrators. Ultimately, health care relies on skilled and trained individuals to make the correct decisions that govern the care of an individual patient. As the wealth of data relevant to a patient grows to include biometrics, genomics, and increasingly rich records of treatments and outcomes, computational tools will be needed to extract meaningful and relevant information needed to make appropriate patient-specific decisions.

Recommendation 2: IT and Health

a. NSF, HHS, NIH, DARPA, DOD, and other agencies with responsibility for aspects of health care should continue to support foundational research in health IT. NSTC should continue to support coordination efforts such as the HITRD SSG.

b. NSF, HHS, NIH, and NIST should develop and nurture open interfaces, standards, and also incentives for promoting the leveraging of electronic health data in data analyses in support of biomedical research and in the delivery of health care.

c. NIH and HHS should create funding mechanisms that will encourage accelerated deployment, testing, and evolution of translational IT systems for clinical use.

Big Data and Data-Intensive Computing

Large data collections and data-centric computing play a central role in the vitality of the public and private sectors and will continue to be critical in science and engineering, commerce, health care, government services, and national security. For example, advances in machine learning procedures and practices over the last five years have led to significant improvements in core technologies, such as speech recognition, face recognition, translation, and image interpretation that now appear in commercial services. Such advances also have led to compelling demonstrations and prototype fielded systems with remarkable new capabilities, including self-driving cars, real-time speech-to-speech translation, and automated image-captioning systems. As described in the IT and Health section, analysis of large and diverse data sets promises to play an important role in understanding health and in personalized health care. Advancing scientific principles and applications, nurturing a competent workforce, and leveraging the creativity and complementarity of multiple agencies via coordinated activities and programs in the realm of Big Data and data-intensive computing are critical to the health and competitiveness of the Nation and to the quality of life of people throughout the world.

The 2010 PCAST NITRD report recommended that, "NSF should expand its support for fundamental research in data collection, storage, management, and automated large-scale analysis based on modeling and machine learning. The ever-increasing use of computers, sensors, and other digital devices is generating huge amounts of digital data, making it a pervasive IT-enabled asset. In collaboration with IT 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."

The response to that recommendation was strongly positive. In March 2012, OSTP announced a Big Data Initiative, and many Federal agencies are now participating in R&D on data collection, machine learning, and other aspects of data analytics such as visualization. Among them are NSF, NIH, DOD, DARPA, DOE (including the DOE-funded Institute of Scalable Data Management, Analysis and Visualization (SDAV)), and United States Geological Survey (USGS). In 2012, NSF and NIH issued a joint initiative in Core Techniques and Technologies for Advancing Big Data Science and Engineering (BIGDATA). The November 2013 OSTP-NITRD meeting on Data to Knowledge to Action highlighted multiple research projects funded by both Federal Government agencies and public-private partnerships. Earlier in 2015, NSF announced the upcoming formation of a national network of Big Data Regional Innovation Hubs, centering on the fostering of "cross-sector collaborations and partnerships" around data, with input and involvement of stakeholders from academia, industry, government, and non-profit sectors. Such efforts are still in their formative phases and need to be monitored and evaluated over the next several years.

Finding: The Federal government has been responsive to previous recommendations concerning large-scale data collection and analysis.

There are growing opportunities for researchers to share important massive data sets such as those collected in biomedicine, astronomy, and biology. Beyond capture of data sets, it is important to maintain large-scale data sets over time and to provide continued easy access to them. Storing and then making large datasets available over extended periods of time can be a costly endeavor. In addition, the holders of large data sets might move on at the end of a time-limited project, leaving uncertainty about the successor custodian and the disposition of the data resources. OSTP has provided leadership in highlighting the issue and encouraging both government agencies and the private sector to participate in such "data stewardship." Investigators funded by NSF and sometimes other Federal agencies are expected to submit data-management plans for long-term preservation and access and can request funding for that purpose. Researchers continue to find access to some data they need difficult or impossible, even with strong protections for privacy and security. More must be done to incentivize people and institutions to share data sets and to promulgate best practices for providing maintenance and continued access to those data.

Box 1. Data Analytics and Machine Learning

Data analytics are a collection of methods centered on the discovery and analysis of patterns, structure, or correlations in large collections of data. Data analyses provide visualizations, predict outcomes, confirm or disconfirm hypotheses, identify causality, and, more generally, provide insights and guide decisions. Mathematical analyses employing statistics play a central role in data analytics. The results of data analyses include the development of predictive models that forecast outcomes in healthcare, identify artifacts such as faces from images, discern meaning from sequences of words, recognize speech from audio streams of utterances, and create classifications such as all the people who will likely live to 90 years of age.

Predictive analytics is based on the use of sets of examples to build predictive models that can classify new instances into classes of interest. Statistical procedures for predictive analytics that are scaled to large numbers of variables and data sets are often referred to as *machine learning*. Machine learning includes several different classes of methodology, including *supervised, unsupervised,* and *reinforcement* learning. A set of procedures referred to as *active learning* refers to methods for extending data sets by directing the collection and use of new data.

A supervised learning system often starts with a data set of cases, each of which contains observational variables (evidence) and a ground-truth label. Such pairs are partitioned into training and validation sets. Predictive models are developed by using the training data and then tested for accuracy with the “held out” validation data. The accuracy of predictions on the validation data serves as an estimate of the predictive power of the model when it is used in real-world applications to classify or predict on previously unseen cases.

A second paradigm in machine-learning research is unsupervised learning. In unsupervised learning, methods such as clustering are used to discover patterns and labels without having access to signals about ground truth. Beyond purely supervised and unsupervised methods, a mixture of approaches can be used. In semi-supervised learning, initial analyses with labels about ground truth are used as seeds for further refinement. For example, in some approaches, successful predictions based on supervised methods can be used iteratively to tag unlabeled cases, thereby strengthening the ability to predict. For example, a language translator takes a phrase in the first language and produces an equivalent phrase in the second language by analyzing the collection of translations known to the system and then adding the new translation to that collection.

A third machine-learning paradigm is reinforcement learning, which draws on ideas from control theory. In reinforcement learning (used, for example, for robot control or game playing), the training data comprise a model of the problem environment and predictions are made based on maximizing reward (the quality of the next action) using algorithmic approaches such as approximate dynamic programming and Monte-Carlo methods.

Recent advances in machine learning include supervised learning with the use of convolutional neural networks (referred to as “deep-learning”) centering on the use of a hierarchical representation. The methodology has been harnessed with considerable success for such problems as language translation, speech understanding, and recognition of objects in still and moving images.

The successful use of machine learning is sensitive to the quantity of data and computation; large amounts of data and high-capability computing systems may be required for effective solutions. The quality of an analysis also rests on the artistry of the design of the machine learning method that is used, especially the selection of learning procedures and the parameters that are used in the predictions.

Finding: Inadequate programs for data stewardship are leading to losses of data and associated opportunities for machine learning, inference, and longitudinal studies. Individual investigators and program are not well incentivized or resourced to manage or share data, and there are few programs that provide support or guidance for data stewardship.

Multiple research directions have come to the fore as important in the advance of machine learning and in the broader realm of data analytics. Key themes include the following research challenges and opportunities:

- Effective uses of large-scale data analyses require methods for representing and propagating error analyses and confidence measures in large-scale data analysis, especially when the data is derived from heterogeneous data sets. Additionally, analysis of high-dimensional data sets can be expected to yield false discoveries at some rate and methods for characterizing and checking for false discoveries can be critical for high-quality inferences and decision support. These areas of research, which pose important challenges, are often underappreciated, but they are essential in many high-value applications such as health care, national defense, or climate modeling.
- Discovery of new causal associations, influences, and structures sits at the heart of science. Identifying causal dependencies (as opposed to correlation) from analyses of large data sets would have enormous practical impact (for example, reducing the need for costly randomized controlled trials in medicine or providing complementary analyses that reduces the scope and costs of trials). Despite promising recent progress, this topic remains one of the most challenging areas of machine learning and data-intensive computing. We urge more study of this important topic, especially with regard to practical, real-world tools and practices.
- The renaissance and fast-paced adoption over the last five years of a family of methods referred to as neural network methods (e.g., convolutional neural nets, “deep learning”) has framed new questions about opportunities and challenges with these approaches. Enthusiasm for applications of neural network methods has outpaced understanding of the foundations for the jumps seen in discriminatory power. Other R&D opportunities include studies at the intersection of machine learning and distributed systems and networking infrastructures aimed at identifying systems that best support large-scale learning of neural networks and methods to combine different types of information (e.g., language and vision) in the same neural network models.
- The march of machine learning into real-world applications has surfaced questions about the role of machine learning within larger pipelines of information gathering and decision-making, including opportunities to fold considerations of the utility structure of application domains into the learning procedures. There are opportunities to perform “active learning,” the study of how machine-learned models can be used to guide the collection of additional data based on current data available and predictive goals to guide data collection. Such methods include techniques for automated experimentation.
- As data sets and appropriate inferences grow in complexity, the need for tools to visualize and explore data interactively – making the processes and results of machine learning more understandable – is growing in importance. Research is needed to introduce transparency into both the process of learning models and the uses of models for making inferences. Research and development of new, flexible, and easily customized tools for “end users” of data visualization are an important complement to research on data analytics. Tools that combine human intelligence and abilities with advanced computation have the potential to advance science, engineering, and

private-sector applications in ways that neither alone could. Further research is needed to develop them.

Finding: More research is needed on error analyses and confidence measures for data analysis, on the determination of causality from data, on better understanding of neural network models and their construction, on widening studies of machine learning to consider the larger decision making pipelines that they support, and on tools and methods that enable interactive data visualization and exploration.

In light of these findings, PCAST makes the following recommendations:

Recommendation 3: Big Data & Data-Intensive Computing

a. NSF, in collaboration with mission agencies that collect large amounts of data for R&D, should continue to sponsor research on methods for performing inference, prediction, and other forms of analysis of data to advance all areas of science and engineering, and on methods for the collection, management, preservation, and use of data. Emphasis should be placed on formulating and disseminating methods for representing and propagating error analyses and confidence measures in large-scale data analysis; developing the theory and practice of computational and statistical methods for causal discovery from large data sets; developing deeper understandings of the foundations of neural network models and systems challenges with scaling up these methods; uses of machine learning to guide decision making; and human understanding of large data sets and the results of their analysis.

b. NITRD, through its Big Data R&D SSG, should work to establish a common set of best practices and support structures for data capture, curation, management, and access. The NITRD Subcommittee of NSTC should encourage uniform adoption of these policies through the NITRD membership.

IT and the Physical World

Computing is exploding into the physical world. Compared to even a few years ago, the commercial interest and opportunities in systems that couple IT technologies with sensing and actuation, from the Internet of Things (IOT) to smart infrastructure to robotics, have evolved enormously. Commercial products already let people observe and control their homes remotely, enable cars to park themselves, and allow factory workers to teach a robot to perform tasks on the shop floor. Future systems will provide assistance to the elderly, function as aerial mobile communications platforms, and provide autonomous personal transportation. Rapid technology advances in remote sensing, powerful mobile devices, and cloud-powered analytics have set the stage for this trend, but they will also be reshaped by it.

Historically, support for IT and the physical world has focused on experimental intelligent systems – primarily through artificial intelligence and robotics research – and more theoretically driven investigations of high-reliability and embedded physical systems, now commonly referred to as cyber-physical systems (CPS). More recently, a new class of systems comprising many networked components, IOT, has evolved within the commercial sector. For historical reasons, researchers in these areas have formed multiple communities, each with their own conferences, journals, and professional societies. While different in character and intellectual focus, all of these communities share the common goal of advancing the capabilities of IT to observe, understand, engage with, and manipulate the physical world.

Finding: Researchers in robotics, artificial intelligence, cyber physical systems, and related research communities will contribute to the important advances anticipated in the various aspects of computing and the physical world, and will drive research and innovation in basic IT.

The 2013 PCAST NITRD report observed that substantial progress on prior recommendations has been made in this area. In particular, the National Robotics Initiative (NRI), a cross-agency initiative announced in 2011, and NSF's investments in cyber-physical systems (CPS) research have strengthened research in IT for the physical world. Although the Cyber-Physical Systems senior steering group and the High Confidence Software and Systems coordinating group provide coordination in CPS, there is currently no coordinating group for robotics in NITRD. Experts with whom we spoke observed that the NRI and CPS cover only part of the research agenda for IT-enabled physical systems – NRI focuses primarily on systems that collaborate or interact with humans, and CPS research is focused on developing formal techniques and models for systems. Research on autonomous intelligent systems (e.g., autonomous under-water vehicles) and on large-scale networked systems (e.g., IOT-type systems) has lagged in comparison. DARPA and DOD investments in these areas appear to be constant or even waning, but current budget reporting methods do not provide a means to verify these beliefs.

Finding: There is a continuing need to develop the basic science that will enable robust, reliable autonomous systems that incorporate advanced sensing and sensor feedback. Continued investment in IT technologies for sensing and acting in the physical world is important to maintain world leadership in these areas.

Finding: Federal investments in computing related to the physical world are difficult to track because existing Program Component Areas do not capture this activity.

The rapidly evolving landscape of this area presages disruptive change that will affect the associated research communities. In particular, basic research in physical IT R&D has historically included *both* exploratory technology development as well as fundamental research to advance the frontier of possible technology innovations. We can anticipate that industry and venture capital will begin to invest in exploring technology innovations in this field, which – in turn -- will increase the number of available platform technologies upon which research can build. This has happened in cloud computing or cyber-human systems, for example. Research on these new platforms will demand new and synergistic collaboration between industry and academia.

Finding: As new products and technologies develop for IT-enabled sensing and acting in the physical world, it is important to put in place open standards and platforms that will encourage activities that share the new technologies with the research community and among its researchers.

Another emerging aspect of IT in the physical world is the increased relevance of cyber-human systems research. Many of the most compelling opportunities for commercial advancement in physical IT will address important societal needs: more efficient transportation, more effective health care, more productive agriculture, more agile manufacturing, and overall improvement in quality of life in the home and workplace. These applications will motivate new research in cyber-human systems, safety, trust, and privacy in the physical world. We note that these advances will, in turn, expose new opportunities for basic IT research on resilient and reliable hardware and software systems, real-time sensor data analysis and control methods, new interface methods and technologies, and information security.

Finding: There is a need for research focused on human interaction with systems that operate in the physical world, particularly around issues of safety, trust, and predictability of response.

The following are research areas that PCAST suggests should be emphasized in the coming years:

- Physical IT and human interaction – Physical interaction with computationally enabled physical systems must evolve beyond keyboard, mouse, or touch-screen. Interaction will involve speech, gesture, and physical contact. The growing intimacy and variety of interaction will dictate research on usability, efficacy, trust, and other cognitive and socio-technical aspects of systems operating in the physical world.
- Physical IT and robust autonomy – Rapid advances in sensing technologies, both software and hardware, have created enormous data archives in cloud-based storage and computing systems plus concomitant rapid advances in machine learning on large static data sets. Sensing and action related to physical world interaction – manipulation of materials or tissue, reaction to the movements of humans or pets, or interpretation of human action – remain challenging, unsolved problems. Work that supports robust, reliable, and effective autonomy through sensing and interaction is essential for future growth in this area.
- Physical IT and sensing – Algorithms for sensor interpretation are advancing at a rapid pace. Sensing needs of IT systems differ, however, from those of (for example) an information-retrieval application. Physical IT requires short response times, high reliability, and often depends on highly application- or task-specific information extraction. Often multiple modalities – vision, ranging, touch, and force – must be integrated. Together, these form an important sub-area within physical IT.
- Development of hardware and software abstractions for physical IT systems – New software, hardware, and theoretical computing abstractions will be needed to model, develop, and deploy systems with the same velocity of innovation that today's web-based systems are deployed. Many of these future systems, e.g., automated driving, will be likely to require very high dependability, but must achieve it in far less controlled settings than current high-reliability systems, such as today's aircraft or medical devices.
- Trustworthy physical IT systems – The opportunities for unexpected behavior due to misuse, abuse, attack, or accident leading to physical damage, harm, or loss of life will grow rapidly. There will be a strong need for research that enhances reliability, resilience, and safety through formal and empirical validation and verification throughout the entire hardware and software stack.

Recommendation 4: IT and the Physical World

a. NSF and DARPA should lead cross-disciplinary programs that will advance research and development of new approaches to robust autonomy; advance security and reliability of such systems; promote integrative approaches to human interaction; explore new sensing and interface technologies; and incentivize fundamental science on the cognitive and social aspects of interactive physical systems.

b. Mission agencies – particularly DOD, DOT, and NIH – and NIST should promote the development of open platforms and sharable infrastructure for research on physical systems within application domains – transportation, agriculture, urban infrastructure, health care, and defense.

Privacy Protection

In the 2010 NITRD report, PCAST recommended that, "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." At the time of the 2013 NITRD report, privacy-research programs were emerging, but there was still no significant multi-agency attention to this topic. The 2013 PCAST NITRD report recommended that, "NSTC should create a multi-agency collaborative effort led by NSF, the Department of Health and Human Services (HHS), and DARPA to develop the scientific and engineering foundations of privacy R&D. NITRD should coordinate across government agencies to develop deployable technologies and inform policy decisions."

Since the last NITRD review, public awareness of digital privacy issues has been heightened substantially. In response to a request by the President, PCAST reported on the technology aspects of privacy protection in 2014.[19] At the request of OSTP, the National Coordination Office (NCO) and the CSIA R&D Senior Steering Group, in consultation with the research community,[20] are developing a National Privacy Research Strategy to guide Federal R&D investments in the context of both interagency collaboration and interactions with the private sector. The topic of privacy protection is also receiving increased attention from the research community.

Finding: Privacy is an important human and societal value. Its protection is increasingly threatened by the growing amounts of online data, the increasing ability to combine and analyze that data, and the absence of adequate safeguards to keep private both original data and information derived from it.

The research needs and opportunities in the realm of protecting personal privacy in a digital age are both diverse and challenging. Privacy entails both human values and personal information. Data that an individual willingly shares with certain people in particular roles – for instance a physician, a spouse, or even a salesperson – she might not share with anyone else. Information that would be shared willingly at a certain stage in one's life might be regarded as secret at a later date. The same individual data that might be withheld from other individuals might be readily shared in an aggregated collection in which those data would not be associated with the individual. The aggregated data, particularly as accumulated over time for longitudinal analysis, can be invaluable for purposes such as research on health and disease, understanding of social and cultural behaviors, predictions of future resource needs, preventing crime or attack, and many more. The inappropriate or unwanted disclosure of personal information can cause irreparable harm, both to individuals and to groups.

Simultaneously, the abundance of available personal data continues to increase, and the technological power to collect and analyze that data continues to grow in scope and power. The PCAST report on privacy discusses this phenomenon in considerable detail, describing the potential threats to privacy posed by large-scale data collection and retention, data fusion, increasingly sophisticated data analysis, and the possibilities for misuse.

---

[19] PCAST, Executive Office of the President, "Report to the President, Big Data and Privacy: A Technological Perspective," https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf, 2014.

[20] Cranor, L.; Rabin, T; Shmatikov, V.; Vadhan, S.; Weitzner, D.; "Towards a Privacy Research Roadmap for Computing Community," *Computing Community Consortium,* http://www.cra.org/ccc/files/docs/white-papers/CCC%20National%20Privacy%20Research%20Strategy.pdf, 2015.

The present working group has also spoken with various experts who did not participate in the PCAST privacy study and has reviewed other subsequent documents.

Finding: Research on privacy protection must draw on a breadth of expertise. Collaboration is needed among a broad range of computer scientists, government and legal scholars, behavioral and social scientists, and researchers in domains in which IT and human data are increasingly important (e.g., health, transportation, education, public safety, "smart" infrastructure), together with the practitioners in those domains.

Finding: An understanding of the technologies for privacy protection should inform the design and modification of computing systems and the drafting of policies and regulations. Realizable privacy protection must be incorporated from the outset, and both systems and regulatory frameworks must incorporate advances in privacy protection automatically. Tools are needed to help both system builders and policy makers.

Finding: Technology should be developed so that the primary burden of privacy protection does not fall on the people being protected and dependency on limiting collection and anonymizing personal information is reduced. Those mechanisms are known to be inadequate forms of protection. Technology that protects the *use* of personal information is a much stronger form of protection than attempting to inhibit data collection or identifiability.

There are many research challenges to be addressed to improve privacy protection in current and future computing-enabled information systems. Some of those challenges are:

- Understanding and clarifying what is meant by "privacy." There are many concepts of privacy; among them are some grounded in history and law, some based on potential harms, some based on rights to ownership of personal data.
- Automatically tracking the use of all forms of personal data, starting from their origination, and including data fusion and information derived from that data by other means such as analytics. The techniques must be sufficiently powerful to encompass multiple copies and distributed data.
- Specifying privacy protection policies and privacy laws formally, so that privacy preserving services can be automated. In addition, such specifications can be used in creating methods to track and monitor policy and legal compliance that are not instance-specific. Note that a privacy policy might be associated either with the data subject or with the data user.
- Devising methods to use private data without disclosing private information. The comparative strengths and weaknesses of different approaches such as data analytics performed with encrypted sources or methods that selectively remove information or inject noise need to be better understood, and new approaches need to be created.
- Devising methods that minimize collection of data and sensing in accordance with the purposes of their use.
- Detecting, signaling, and mitigating information leakage and privacy violations as they occur.
- Creating mechanisms, frameworks, and tools to enable system builders to construct privacy-preserving systems without themselves being privacy experts.

Recommendation 5: Privacy

OSTP and NSTC should continue to develop and expand a multi-agency research and development program to advance the science, engineering, policy, and social understanding of privacy protection. Agency participation should include at least NSF; NIH; the units of DOD, NSA, and DHS studying the extensions of cybersecurity R&D to encompass privacy; and other relevant units within HHS, NIST, DOT, and ED.

Cyber-human Systems

Computing is integral to people's work and lives, and to the aims and processes of organizations and society. Past NITRD reports emphasized that humans are fundamentally social and connected to others through personal, cultural, and national ties. The 2010 PCAST NITRD review recommended that "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. " It was suggested that the program include a science of social computing, a strengthening of collaborative computing and peer-to-peer production systems, and a research platform for computational social science research. The 2013 PCAST NITRD review observed that while all three agencies had increased their activities in various aspects of social computing, there was still little or no interagency coordination.

The field of research referred to in those reports as "social computing" has broadened and matured to encompass a larger set of topics that integrate computing and people. Today, domains of cyber-human systems include computational systems that support communication and coordination of individuals, groups, and organizations (topics include social media, collaboration systems, online training and virtual reality for education, and peer lending and financial services), computational systems and methods supported by people (topics include prediction markets, reputation systems, human-in-the-loop automation, and crowdsourcing), and socially intelligent devices and systems (topics include algorithmic management, human-robot interaction, knowledge sharing and decision-support systems, sensor networks, brain-machine interfaces, and software for customization and personalization). Recent advances in cyber-human systems have enabled new forms of human organization, such as friend-sourcing, peer lending, and human augmentation, labels that have now entered everyday language.

The science of cyber-human systems has advanced significantly. For instance, recent scientific studies have tested theories of structure and incentives in crowdsourcing, have conducted basic research on attention and habituation to better support crowd workers, have probed foundations of team effectiveness, have explored ideal models for pooling the intellect of people and machines in "complementary computing" solutions, and have revisited, from a computational perspective, long-standing ideas from the social sciences, such as models of "social loafing" and engagement. The growth of the importance of this area is reflected in new or expanded programs in human-computer interaction, connection science, design, health informatics, human robot interaction, and social network research at major research universities and in college curricula.

Finding: The field of cyber-human systems has matured and broadened in scope, encompassing communication and coordination of people, computational systems and methods supported by people, and socially intelligent devices and systems, in addition to human use of computational services.

NSF has done a good job of supporting and coordinating social computing and the larger domain of cyber-human systems. DOD and DARPA have focused on research in security and terrorism, topics related to their mission. NIH has made progress in supporting initiatives in team science, but appears to do little research on understanding and assessing the use of technology in medicine.

Given the research advances in online learning, schooling and training, collaborative learning platforms, neuroscience advances in understanding learning and development, and possibilities of diagnostic (as compared to evaluative) testing, one would expect strong interest from ED. But this research seems not to be coordinated with ED investments in evaluating school reform, promulgating best practices and diversity, and educating teachers for tomorrow's education challenges.

Finding: Interagency coordination has been effective in some areas (e.g., in leading new efforts in visualization, team science, and security education) but coordination in other areas such as social computing, human-robot interaction, privacy, health informatics, and human learning and education has been sporadic and highly dependent on the initiative of individuals.

There are many aspects of cyber-human systems that are not well understood and warrant research. Some illustrative examples include the following:

- Education – A major challenge for computer-based learning systems is to build on principles of learning that have solid evidence of effectiveness — principles such as optimized scheduling and pacing of the flow of concepts, feature focusing, visual-verbal integration, worked examples, corrective self-explanation, and personalization.[21] Also, it is important to understand the interactions of technology with race, gender, and socioeconomic status.
- Workforce and Society – Technological change is making possible many new approaches to economic development, for example: enabling small farmers in Africa to order and report delays in fertilizer delivery, community health care workers to report and track immunizations, and better ways to aid in disaster recovery by using robots in search and rescue. Technological change is also having little-understood indirect consequences, both positive and negative, for the overall economy, for communities, for government institutions and companies, and for workers. Many of these effects were unanticipated a decade ago, and will require continuing investments in fundamental and applied interdisciplinary research to understand patterns and develop new tools for workforce development, education, and well-being.
- Communication – Social media and networked communications have given rise to greatly expanded peer-to-peer communications and new, more informal and spontaneous organizations. These new forms have come with some new risks and problems, including free-riding, inappropriate behavior,

---

[21] (1)Koediner, K. R., Booth, J. L., & Klahr, D. "Instructional complexity and the science to constrain it," *Educational Forum, Science,* Vol 342, 2013; (2)Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. "A systematic literature review of empirical evidence on computer games and serious games," *Computers & Education* http://www.sciencedirect.com/science/article/pii/S0360131512000619, 2012.

herding, bullying, conflict, and transmission of rumor and poor advice. A major research challenge is to learn how we can discourage these negative behaviors and mitigate their risks so that people and groups online can have reasoned argument and make decisions based on reliable information. Research will need to be conducted at all levels, from interface to architecture.

- Behavioral economics and social influence – Emerging "smart" consumer products and services build on dynamic patterns of human behavior and employ powerful mechanisms of social influence enabled by computation. For instance, networked homes, cities, and businesses that enable communication and track people's activities could increase efficiency, and reduce waste. A major challenge is to understand how to benefit from these advances and mitigate misuse or overly intrusive applications.
- Human-machine collaboration and complementarity – A rising area of scholarship and engineering in cyber-human systems focuses on challenges and opportunities to combine the complementary skills of people and computational systems in jointly solving challenging tasks. This work includes the need to understand better the role of people in taking shared responsibility in the control of semi-autonomous machines, such as automobiles and surgical robotic systems. It includes efforts in citizen science that employ computational optimization and intelligence in joint work with people (e.g., FoldIT, Eyewire, CrowdSynth, etc.). This topic is of particular interest in light of recent advances in display and interaction technologies, such as the Oculus Rift or Microsoft Hololens – both powerful new interaction platforms.
- Scientific method – New computational methods, such as machine learning and network analyses of Big Data are giving us the capability of understanding people and society in new ways. There are many unsolved problems that can be tackled using these new approaches: inequality, recessions, the spread of disease and rumor, and violent crime. A major challenge is that Big Data analyses are essentially correlational and do not demonstrate causation or generalizability. Interventions may require investment in scientific controlled trials and longitudinal analyses. Researchers need to develop ways of integrating Big Data analysis and the traditional scientific method, two very different ways of conducting research.

Recommendation 6: Cyber-human Systems

a. NSF should continue to broaden its support for fundamental research on the systems and science of the interplay of people and computing.

b. OSTP and the NITRD Subcommittee of NSTC should establish or strengthen coordination at both higher and lower levels among at least NSF, DOD, DARPA, NIH, and ED. In particular, coordination and support in areas such as social computing, human-robot interaction, privacy, and health-related aspects of human-computer systems should be enhanced.

High-capability Computing for Discovery, Security, and Commerce

Since the High-Performance Computing Act of 1991 launched what became the NITRD program, the nation has strived to develop and exploit extreme-scale computing, networking, and data management. The most capable systems (largest, fastest, and most powerful) are critical to national defense, to discovery-based research in all fields of endeavor, and to commerce. In science and engineering research, high-end systems are scientific instruments much like particle accelerators or telescopes. Advances in the most-capable systems lead to advances in smaller and more specialized systems as well.

The 2010 PCAST NITRD review contains an extensive discussion of high performance computing, emphasizing the forthcoming research challenges for continuing to advance these computing platforms. The report recommended that "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) efficiently analyzing vast quantities of both numerical and non-numerical data, (b) handling problems requiring real-time response, and (c) accelerating new applications." The 2013 NITRD review found little progress in creating such an interagency coordinated research program and recommended that "NSTC should lead an effort by NSF, DOE, DOD, member agencies of the Intelligence Community, and other relevant Federal agencies to design and implement a joint initiative for long-term, basic research aimed at developing fundamentally new approaches to high-performance computing."

PCAST is pleased to observe that OSTP and NSTC have responded positively to the 2013 recommendation and that an Executive Order describing a National Strategic Computing Initiative (NSCI) was issued in late July 2015.[22] Long-term research is a key component of that initiative. As the 2013 NITRD review observes,[23] the Federal Government is both the primary enabler and the primary user of these high-capability systems. That creates pressure to find short-term solutions to technical challenges so that improved systems can be deployed and important national priorities that depend on high-capability computing can be addressed. Yet it is essential to discover innovative ways to meet the difficult long-term challenges, so that we will have solutions when we need them. It is also important to recognize that as methods for scientific discovery have evolved, the role of massive amounts of data has increased; as the capability of these systems has increased, the use of scaled-down systems to complement that large ones has also increased.

Finding: The White House has responded to the need for a long-term plan by creating an NSCI.

The advances that have resulted in ever-higher capability to date now allow to push the boundaries of fundamental physical limitations and confronting ever-increasing complexity. Those constraints affect systems no matter the balance in a given system design among computation, data quantity and access, and networking. The research challenges now are to compensate for the physical limitations and master the complexity, while also supporting the creation of systems that balance compute-intensive, communication-intensive, and data-intensive demands as they arise. Resilience and reliability are also important for systems of this size and complexity, but the technologies to address those issues are of less concern, since adequate solutions exist.

Increases in hardware performance have benefit from two phenomena – Moore's law, the decrease in transistor size and consequent increase in the number of transistors per chip, and Dennard scaling, the constancy of power

[22] Executive Order Creating a National Strategic Computing Initiative, https://www.whitehouse.gov/the-press-office/2015/07/29/executive-order-creating-national-strategic-computing-initiative, July 29, 2015.
[23] PCAST, Executive Office of the President, "Report to the President and Congress, Designing A Digital Future: Federally Funded Research and Development in Networking and Information Technology," https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd2013.pdf, 2013.

use in proportion to area. The performance scaling due to miniaturization coupled with energy efficiency is fast reaching its physical limits.

System designers have used increased chip density to improve system performance by increasing the degree of parallelism (coordinated simultaneous execution) in a variety of ways. As chip performance has increased, memory bandwidth (the rate at which memory is read or written) has also improved but at a slower rate. System designers have compensated by incorporating deep memory hierarchies, and by building heterogeneous systems in which high-performance accelerators like graphics processing units (GPUs) are available for particular aspects, creating systems with large numbers of network-connected components. As examples, four of the top ten machines on the July 2015 Top 500 list[24] are based on accelerators; four of those machines have custom interconnection networks. In general, peak performance requires 10^8 (100,000,000)-way parallelism.

Finding: Hardware strategies are reaching fundamental miniaturization and power limitations; computer architectures are using increasingly complex memory hierarchies, increased heterogeneity, and growing degrees of parallelism.

These design characteristics have made high-capability systems very complex and extraordinarily hard to program. That software complexity extends to smaller systems as well. In order to achieve high performance, both systems and application-level software has largely been tailored to particular systems, even when the design concepts were more broadly shared, increasing cost, reducing portability, and inhibiting change. Greater complexity of the simulation, modeling, and inferential analysis problems for which these systems are used increases the software challenge even more. At present, a considerable software burden falls on the applications programmer – not only to find computational solutions for the problems being tackled, but also to express those solutions in a way that provides excellent end-to-end performance on a given system. Creating the needed software requires both deep knowledge of the problem to be solved and deep knowledge of capability computing.

Finding: Software for high-capability computing and for the scaled-down systems used for some problems is extraordinarily complex throughout the software stack. Insufficient progress has been made in addressing that complexity.

This situation cannot continue. To address the programming and energy gaps described above large changes are required in how we build and program high-capability systems. Fundamental new approaches are needed - incremental research extending current approaches is unlikely to suffice. For example, a project that improves the efficiency of a conventional CPU by 25 percent is not even going to match the efficiency of an accelerator, let alone deliver the next order of magnitude improvement in efficiency. Similarly, projects that make small improvements in conventional programming systems are not going to close the large and growing programmability gap. Approaches to programming are needed that both make it easier to create understandable and correct problem solutions and require less expertise in the complexities of high-capability computing technologies.

[24] The TOP500 list, which is updated twice per year, ranks the most powerful general-purpose computers worldwide according to their performance on the LINPACK benchmark for solving a dense system of linear equations. http://www.top500.org/lists/2015/06/, 2015.

Finding: Increases in computing performance can no longer rely on incremental improvements, but require new innovations throughout the hardware and software stack.

PCAST has identified several areas of innovation needed to improve both the performance and usability of future high-performance computing systems:

- Energy – A large and growing fraction of energy is consumed transporting bits from one location to another at each level in a computer system. At the hardware level, energy-efficient signaling systems are a promising area for future research. In the area of architecture, the bulk of the energy and time consumed in a modern computer is spent on *overhead;* comparatively little energy is spent on the actual *payload* arithmetic operations and essential data movement. A promising area for architecture research is to understand how to eliminate much of the overhead energy without sacrificing single-thread performance. Consideration of fundamentally new technologies must be part of the research. Among the more speculative ideas being pursued are quantum computing, the use of superconducting circuits, and neuromorphic (biologically-inspired) computing.
- Programmability – Addressing the programmability gap requires moving from a regime where application programmers manually map their codes to individual target machines – exposing just enough parallelism to fill the machine, carefully blocking data over the memory hierarchy, and so forth – to one where programmers write “target-independent codes” that are mapped to specific target machines using semi-automated tools. The problem of developing target-independent programming systems is much easier than the problem of automatically parallelizing codes (which is largely considered to be intractable), and focusing on semi-automated mappings provides a path toward practical tools that take best advantage of programmer time and experience. Research will necessarily include the development of notations – programming languages – that express the application program in a manner that captures all available parallelism and an abstract notion of data locality. Languages should support the application domain: the considerations that govern efficient execution for data-intensive computing are different from the needs for expressing scientific simulations, which are different from the needs of high-volume, high-throughput genome processing.
- Runtime optimization – A complete set of tools for programming high-capability systems should consist of a suite of compile-time and run-time components that incorporate both the knowledge of efficient mapping strategies and powerful analysis of both the program and the target system. Some of the components may embed knowledge of the application domain. The development of “auto-tuners” that automatically and efficiently search the mapping parameter space, at compile-time or at runtime, to find the optimum mapping for a particular target is an important component. Algorithms to improve energy efficiency, such as algorithms to reduce data movement, are increasingly important. Monitoring and visualization tools that enable the programmer to discover performance bottlenecks so she can guide the mapping tools are another important part of the solution.
- System software and software tools – Rethinking programming notations and mapping technologies must be accompanied by a rethinking of the system-software layers, so that the software services that transcend application domains are shared. Decisions about the design of lower levels of the software stack also interact with decisions about computer architecture. Additional software challenges include approaches to migrating old codes to new hardware/software systems and better methods for reasoning about domain-specific software properties.

Finding: Fundamental new approaches are needed for all aspects of the design of high-capability systems, from hardware to applications programming.

Finding: Advances in high-capability computing will also benefit data-intensive systems and IT at lower demand levels, since energy-reduction and better software development are universally important.

Although progress has been made toward this goal, our first recommendation remains essentially unchanged from the previous review.

Recommendation 7: High-capability Computing for Discovery, Security, and Commerce

- **a.** NSTC should lead an effort by NSF, DOE, DOD, NIH, member agencies of the Intelligence Community, and other relevant Federal agencies to implement a joint initiative for long-term, basic research based on the new National Strategic Computing Initiative aimed at developing fundamentally new approaches to high-capability computing. That research should be sufficiently broad that it encompasses not only modeling and simulation but also data-intensive and communication-intensive application domains.
- **b.** Under the leadership of OSTP, NSTC and the NCO should establish multi-agency coordination not only at the level of program managers but also at the higher administrative levels reflected in senior steering groups.

Foundational IT Research

Other sections of this chapter have singled out certain areas of IT research and development that have been highlighted in recent NITRD reviews or that seem worthy of particular attention at the present time. But all achievements in IT rest on the base of longstanding research.[25] It is easy to look back and pick out examples of how inventive "blue-sky" research has created paradigm-shifting changes – the Internet, graphics and animation, or web search. What is important to keep in mind is that these were the result of years, and sometimes decades, of basic research that came to fruition in sometimes-unexpected ways.

The cadence of innovation continues today – there are undoubtedly innovations being explored right now that will be the next "big thing." Given the entire spectrum from theoretical computer science and basic principles of hardware and software design to applications, there is a broad base of research that is necessary to support advances in IT R&D. For example:

- Advances in computer architecture, building on advances in device electronics, continue to provide the computational hardware platforms for computing systems. Recent focus has been on lowering power demands, on miniaturization through increased chip density, and on increasing parallelism and memory access at multiple levels.[26]

[25] Computer Science and Telecommunications Board, "Continuing Innovation in Information Technology," *National Academies Press,* http://www.nap.edu/catalog/13427/continuing-innovation-in-information-technology, 2012.
[26] Computing Community Consortium, "21st Century Computer Architecture: A community white paper," http://www.cra.org/ccc/files/docs/init/21stcenturyarchitecturewhitepaper.pdf, 2012.

- Innovations in domain-specific languages, new capabilities in program analysis, and advances in methods for formal specification and verification enhance the ability to create improvements to software development, maintenance, and reliability.[27]
- Fundamental advances in algorithms that operate on streaming data sets make it possible to dynamically observe and react to torrents of data that would overwhelm traditional approaches.
- Deeper understanding of the fundamentals of algorithms, complexity, approximation, and systems analysis drive progress in every aspect of information technology.
- Progress in the fundamentals of software systems enables the increasingly large and complex systems we use. Topics include operating systems, data management, networking (see sidebar), distributed and concurrent systems, and so forth.
- New systems concepts provide scalable, reliable storage and recall of data sets of unimaginable scope.
- Advances in distributed systems support cloud computing and resilient data availability.
- New algorithms for machine learning that provide improved results and which scale to ever-larger data sets enable impressive advances in speech recognition, object recognition in images, and interpretation of medical data.
- New advances in computer graphics and visualization make it possible to discover new patterns or associations in data, or to visualize ever more complex physical phenomena under realistic conditions and in real time
- Major improvements in computer vision drive progress in robotics; advances in face recognition facilitate search; progress in speech recognition and natural language processing enable human communication and comprehension.

Although this kind of research is less visible to the public and to researchers in other fields, it is essential to progress in IT. It is important that the foundational research have a major long-term emphasis, incorporating both broad and novel approaches that will sustain computing in the future, not just solve the problems of the near term. It is equally important that funding be balanced between projects that involve a single investigator or small group and projects that bring together larger, cross-disciplinary groups; likewise, it is essential that funding be set aside to support early career researchers, who are often the source of the most innovative and "out of the box" ideas.

[27] Ball, T. and Zorn, B, "Teach foundational language principles," *Communications, ACM* 58, 5, 30-31. http://cacm.acm.org/magazines/2015/5/186023-teach-foundational-language-principles/fulltext, 2015.

Box 2. New Directions for Networking Research

A computer network, or simply a network, is a collection of computers and other hardware components interconnected by communication channels that allow sharing of resources and information. The Internet is an example of a system that uses networking technology to provide a vast quantity of resources and services. Important research topics in networking have included the communications protocols, network topology, (wired or wireless) transmission media (the physical layer), performance (quality of service, resilience), and devices and algorithms that move data from a source, onto and through the communications media, and into the destination. Modern networks transmit data in discrete packets that are generated at the source and reassembled at the destination.

Traditionally, "closed" vendor-produced devices such as network interfaces, routers and switches have controlled networks. More recently, there has been considerable research, experimentation and deployment of general – purpose commodity processors as network components, enabling programmable dynamic network control, diversity of network architectures, layering of network topologies, and network support for particular classes of applications such as streaming, cloud computing or large collections of intermittently-available diverse devices (Internet of Things).

The exploration of the many opportunities that this new networking flexibility provides, and the new challenges it presents to traditional network properties (speed, resilience, reliability, ubiquity, for example) are in their infancy.

Finding: Foundational long-term research in information technology is essential for the application areas that build on it and for the future of the Nation's robust IT industry.

Recommendation 8: Foundational IT Research

NSF should continue to invest in long-term foundational research in information technology. Other NITRD agencies, including DARPA, IARPA, DOE, and NIH, should support foundational research in those aspects of IT that most affect their missions.

4. Government Support for Information Technology

Introduction and Operating Structure

The NITRD program is the coordinated investment by multiple government agencies in IT. The program operates under the NITRD Subcommittee of the National Science and Technology Council's Committee on Technology. The Subcommittee provides overall coordination of the program and is made up of representatives of the eighteen member agencies investing in IT. Additional agencies participate in some NITRD activities.

The National Coordination Office (NCO) for NITRD, working closely with the Office of Science and Technology Policy (OSTP) and the Office of Management and Budget (OMB), supports planning, facilitating, and reporting for the NITRD program. Many external groups provide input to the NCO and the NITRD program, including the Computing Community Consortium (CCC), the Computer Science and Telecommunications Board (CSTB) of the National Research Council, the professional societies in IT areas, and the President's Information Technology Advisory Committee (PITAC) subcommittee of the President's Council of Advisory for Science and Technology (PCAST) that periodically conducts a review of the NITRD program.

Education and Training

As illustrated in Figure 1, more than 65 percent of the new jobs in STEM in the next seven years will be in computing, continuing the demand for IT workers that has been described in earlier PCAST NITRD reviews. The 2013 PCAST NITRD report recommended that the National Science and Technology Council (NSTC) Committee on STEM education and the various government agencies lead in educating more children and adults in IT, through programs in schools, skills training programs, and high-quality continuing education programs. That need continues.[28]

Education and training are needed at multiple levels, from highly skilled researchers and practitioners, to users of conventional IT tools and methods.[29] Well-prepared teachers are essential for encouraging young people into

[28](1)Johnson, J., *Google for Education,* "The Computer Science Pipeline and Diversity: Part 1 – How did we get here?" http://googleforeducation.blogspot.com/2015/07/the-computer-science-pipeline-and.html, 2015; (2) Johnson, J., *Google for Education,* "The Computer Science Pipeline and Diversity: Part 2 – some positive signs, and looking towards the future." http://googleforeducation.blogspot.com/2015/07/the-computer-science-pipeline-and_9.html, 2015.

[29] (1) A September 2014 PCAST letter report discusses information technology training and employment matching. (See: PCAST, Executive Office of the President "Information Technology for Targeting Job-Skills Training and Matching Talent to Jobs" https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/PCAST_worforce_edIT_Oct-2014.pdf , 2014; (2) Also, the President's TechHire Initiative launched in March 2015 engages communities and the private sector to provide both continuing education and training and employment opportunities. It aims, in part, to ensure that newly trained people actually enter the IT workforce.

IT fields and for maintaining an educational pipeline.[30] Special efforts are needed to ensure that a large and diverse population of young students enters the pipeline, and that the pipeline-leakage is minimized.

Finding: The demand for skilled IT workers continues to grow.

Finding: Diversity in computing-related fields continues to be a significant issue. This reduces the Nation's ability to fulfill the demand of the IT industry and limits the economic opportunities for those populations.

In addition to supporting programs to engage K-12 students in STEM activities generally,[31] NSF has created new programs to teach middle and high school students computing principles, including professional development for teachers. Among them are a funded collaboration with the College Board to create a new advanced placement course, the STEM+Computing partnerships to incorporate information technology in K-12 STEM education, collaborations with the College Board to create a new advanced placement course, and collaborations with code.org, Project Lead the Way, and other groups to scale and deploy K-12 programs broadly through Web-based platforms and through collaborations with local school districts.[32] These deployments have also created data on many learners that can now be used to support the study of the effectiveness of approaches and the quality of outcomes that will inform new innovations on educational programs. This connection of education with research on its effectiveness is a very effective model for leveraging Federal investments in IT education; it should be fostered and grown, subject to the kinds of privacy protection described above.

[30] [30] Such teachers are particularly scarce in rural areas, which may require additional creativity to reach their students. NSF's CS10K initiative aims to get 10,000 computer science teachers in 10,000 high schools, https://cs10kcommunity.org/.
[31] A notable example is the Innovative Technology Experiences for Students and Teachers (http://stelar.edc.org/about-nsf-and-itest).
[32] National Science Foundation, "College Board and NSF expand Partnership to Bring Computer Science Classes to High Schools across the U.S." Press Release 15-060, http://www.nsf.gov/news/news_summ.jsp?cntn_id=135335&org=NSF&from=news, 2015.

Finding: NSF programs demonstrate that collaboration with organizations outside of government can create opportunities to scale and deploy new educational approaches, to broaden the pipeline of information-technology workers and to create new forms of data to support further research on educational best-practices.

Recommendation 9: Education and Training

a. The NITRD Subcommittee, working in partnership with NSF, ED, and the private sector, should create new educational opportunities in IT at all levels, beginning with K-12, to grow the pipeline of skilled workers and identify future innovators and leaders. These programs should incorporate approaches that will engage under-represented populations.

b. As part of that effort, NSF should lead the development and implementation of model programs for pre-college students that attract the most talented young people to study IT. These will be the future innovators and leaders. The program should be designed to address differences in gender, economic status, and cultural background and to collaborate with industry to provide resources to expand these programs broadly in the U.S. education system.

c. NSF and ED should create programs for training and retraining of workers at all age levels with the goal of providing targeted "on-ramps" for those individuals to develop careers in the IT industry. They should fund research that includes 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 and with external partners to deploy these programs and capture data on performance and outcomes.

Program Component Areas

The overall government investment in information technology research and development is tabulated in the annual NITRD budget.[33] Investments are mapped to the Program Component Areas (PCAs) used to report elements of NITRD spending requests as part of a crosscut budget report that summarizes NITRD spending in each Agency's budget.

Current Program Component Areas

PCAs were first introduced in 1995. All but one of the current PCAs were created in 2001; cyber security and information assurance (CSIA) was added in 2007. The current PCAs are:

- High end computing infrastructure and applications (HEC I&A)
- High end computing R&D (HEC R&D)

[33]The Networking and Information Technology Research and Development Program, "FY2016 Supplement to the President's Budget," https://www.nitrd.gov/pubs/2016supplement/FY2016NITRDSupplement.pdf, 2015.

- Cyber security and information assurance (CSIA)
- Human computer interaction and information management (HCI&IM)
- Large scale networking (LSN)
- Software design and productivity (SDP)
- High confidence software and systems (HCSS)
- Social, economic, and workforce implications of IT and IT workforce development (SEW)

Findings on PCAs

OMB tabulates the Program Component Areas in the annual cross-agency budget report. The areas have remained virtually unchanged despite the evolution of IT in ways impossible to envision in 1991. In their present form, the PCAs do not provide a useful breakdown of planned investment in the areas of national priority. For example, the United States has a National Robotics Initiative, but no visibility into its investment in robotics R&D.

The 2010 PCAST NITRD report recommended that the NCO and OMB modernize the NITRD PCAs to reflect the current nature of the field. The 2013 PCAST NITRD report recommended that "OSTP, with guidance from PCAST, should develop a combination of quantitative and qualitative methods to assess the adequacy and appropriateness of government investments in NIT R&D." Modernized PCAs would be an important contributor to clarifying the Nation's investments. Yet the PCAs were not changed.

In our conversations with NITRD participants, the most common response to questions about changing PCAs was that OMB provided strong guidance to leave PCAs unchanged so that annual budget mapping could continue and historical trends could be analyzed. Intriguingly, in our discussion with OMB and OSTP budget coordination representatives, we found no evidence for this position. On the contrary, there is strong support to establish a process to review and update the PCAs periodically to ensure they track to current priorities.

Proposed NITRD Program Component Areas

Recommendations for updated Program Component Areas, beginning in FY 2017, are shown below. In addition to suggested explanatory sentences to be used by OMB, a rationale is given for each PCA.[34]

- Large-scale data management and analysis (Big Data or LSDMA).
    - Large-scale data management and analysis to develop the ability to analyze and extract knowledge and insights from large, diverse, and disparate data sets.
    - *Explanation for change:* This PCA is an expansion of the former information management portion of HCI&IM (which was originally about databases) to reflect the important issues arising from massive data sets and the significant advances in data analytics. This topic is an additional program focus area in the 2016 NITRD budget report supplement. The descriptive text above is adapted from language used in that document.

[34] The style of the current budget description is of the form X to do Y. We have maintained that style, even though we find it awkward.

- Robotics and intelligent systems (RIS).
    - Robotics and intelligent systems to advance physical and computational agents that complement, augment, enhance, or emulate human physical capabilities or human intelligence.
    - *Explanation for change:* This PCA provides visibility for robotics and the supporting technologies drawn from artificial intelligence in the NITRD structure. No natural place for this topic exists in the previous PCAs, despite its enormous advances and importance, as exemplified by the National Robotics Initiative and recent advances in computer vision, machine learning, and their applications.
- Computing-enabled networked physical systems (CNPS).
    - Computing-enabled networked physical systems to advance the science and engineering of systems that incorporate distributed, communicating sensors and actuators.
    - *Explanation for change*: This PCA incorporates cyber-physical systems, which is currently part of the HCSS PCA, and is an additional program focus area in the FY2016 NITRD budget report supplement. It also includes networked systems of diverse sensors and actuators, sometimes described as the "Internet-of Things" that is the basis for "smart" physical infrastructures such as buildings, transportation systems, and environmental monitoring systems.
- Cyber security and information assurance (CSIA).
    - Cyber security and information assurance to protect computer-based systems from actions that compromise or threaten to compromise the authentication, availability, integrity, or confidentiality of these systems and/or the information they contain.
    - *Explanation for change:* This PCA and budget description are unchanged.
- Computing-enabled human interaction, communication, and augmentation (CHuman).
    - Computing-enabled human interaction and communication to enhance the modes, richness, and effectiveness of interchange among individuals and computing-enabled devices and among large numbers of individuals. Computing-enabled augmentation to advance the use of computing to enhance human capabilities and to provide improved learning, education, and training in all fields.
    - *Explanation for change:* There are now a broad variety of cyber-human systems that have in common the "human-in-the-loop." The revision decouples these topics from the preparation of NIT workers (currently in SEW), and from information management (currently in HCI&IM).
- IT foundational research and innovation (IT-FRI).
    - IT foundational research and innovation to discover fundamental knowledge and methods that enable advances in the science and engineering of IT systems and applications.
    - *Explanation for change:* As IT advances, traditional boundaries among sub-disciplines shift, but the need for basic research across the full discipline continues. Instead of singling out particular topics such as software development or networking research, the category is broadened to incorporate not only computer architecture, software, and networking, but also topics such as

basic research in cryptography, algorithms, artificial intelligence, machine learning, language understanding, and machine perception.

- Enabling-R&D for high-capability IT systems (EHCS).
  - Enabling-R&D for high-capability IT systems advances hardware, systems software, algorithms, and software development methods for high-capability systems.
  - Explanation for change: Broadening of HEC R&D to incorporate support for extreme-scale data analytics, the technical goals of the forthcoming National Strategic Computing Initiative, and the high-capability IT needs of the biological, medical, and social sciences.
- Large-scale research infrastructure (computing, network, storage) (LSRI).
  - Large-scale research infrastructure (computing, network, storage) to provide shared hardware and software to support research that requires high performance computing, large shared information sources, open software, and large non-production IT platforms and test beds.
  - *Explanation for change:* Investments in shared research infrastructure are separated from investments in the R&D they support. Changes in technology make it increasingly pointless and difficult to separate computing, networking, and storage infrastructure. This PCA incorporates HEC I&A, the infrastructure aspects of LSN, and both data center and cloud infrastructure.

Recommendation 10: NITRD Program Component Areas

a. OSTP, NCO, and the NITRD Subcommittee, in collaboration with OMB, should revise the PCAs for the FY2017 Budget cycle and beyond to reflect both the current nature of IT and the major national priorities in which IT plays a major role. An earlier section of this Chapter suggests revisions to the PCAs.

b. Those four stakeholders should create a process to review the PCAs every five to six years and to implement proposed modifications. PCAST or its PITAC subcommittee should provide recommendations for changes to the PCAs.

Groups

Under the NCO's guidance, coordination of cross-agency activities is managed by the NITRD Subcommittee and by a collection of focal groups ("Groups"), arranged around specific topics within IT. Each PCA historically had a corresponding Interagency Working Group (IWG) or Coordinating Group (CG). The groups provide a forum for agency representatives to exchange information about the programs within or sponsored by their agencies and to learn about advances in their fields.

In addition, starting in 2008, the NITRD Subcommittee introduced Senior Steering Groups (SSGs), each in response to a specific government initiative. More recently, Community of Practice (COP) groups were introduced to facilitate the use of advanced IT within government agencies.

Current NITRD Groups

The NITRD structure includes the following current groups.

Senior Steering Groups (SSG)

- Big Data R&D (BD)
- Cyber-Physical Systems R&D (CPS)
- Cybersecurity and Information Assurance R&D (CSIA R&D)
- Wireless Spectrum R&D (WSRD)

Interagency Working Groups (IWG)

- Cyber Security and Information Assurance (CSIA)
- High End Computing (HEC)

Coordinating Groups (CG)

- Human Computer Interaction and Information Management (HCI&IM)
- High Confidence Software and Systems (HCSS)
- Large Scale Networking (LSN)
- Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW)
- Software Design and Productivity (SDP)
- Video and Image Analytics (VIA)

Community of Practice (COP)

- Faster Administration of Science and Technology Education and Research (FASTER)
- Health Information Technology R&D (HITRD)[35]

Until last year when the Video and Image Analytics (VIA) CG began, the IWGs and CGs and the PCAs mapped one-to-one with each other. Over the years, the set of NITRD Groups has become increasingly disconnected from the changes in IT described in Chapter 2. Group members have dealt with the mismatch in a variety of ways – sometimes dividing into subgroups, sometimes changing the nominal topic of the group to a timelier topic. The 2010 PCAST NITRD Review report recommended specifically that the structure of the IWGs/CGs decouple from the PCAs so that the NITRD groups could be modernized independently of the PCAs. With the exception of VIA, that change has not been implemented.

OSTP and NCO introduced the first SSG in 2008 to augment the Cyber Security and Information Assurance IWG with a group of higher-ranking agency representatives. As the 2013 PCAST NITRD review report explained,

> *The first SSG, also named CSIA, was formed in response to the 2008 Presidential Comprehensive National Cybersecurity Initiative (CNCI). It was joined in 2011 by the SSG for Health Information Technology R&D, chartered at the time that HHS ONC was developing a national strategy for health IT. The Wireless Spectrum Research and Development (WSRD) SSG was formed to advance*

[35] This group was formerly a SSG.

the goals of the June 28, 2010 Presidential Memorandum: "Unleashing the Wireless Broadband Revolution."[36] *The Big Data SSG was formed in 2011 as OSTP was developing a National Big Data Initiative. An SSG for Cyber Physical Systems (CPS) was created in 2012.* [37]

The 2010 report recommended NITRD "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." The 2013 report recommended that the use of SSGs be continued.

The SSGs provide higher-management attention to high priority areas in the Government's investment in networking and information technology and provide a focal point for administration and congressional initiatives in NIT research and development. The 2010 and 2013 PCAST NITRD reports make clear the intention that SSG members have high rank within their organizations, with sufficient level of budget authority to drive their organization's contribution to strategic initiatives.

Findings on Groups

The purpose of CGs, the IWGs, the SSGs, and the recently initiated COPs is to coordinate NITRD activities that cross agency boundaries. Many kinds of coordination are important. Among them:

- cross-agency information exchange about planned or existing agency programs
- sharing of expertise for peer-review or other forms of evaluation
- learning and discussing scientific and technical advances in a topic
- building on ongoing efforts in other agencies (thereby reducing duplication)
- planning and executing new cross-agency programs
- planning and organizing major national R&D initiatives

We generally find the model of coordination and management instantiated through the NITRD Subcommittee, the NITRD Groups, and the NCO to be appropriate and comparable to other Federal cross-cut initiatives. The NITRD Subcommittee and the NCO have built and continue to support an important community that learns and shares information across the Federal agencies about research and development in the NIT ecosystem and about activities in the agencies. The National Coordination Office ably executes its duties under the guidance of the NCO Director.

We further find that a mechanism for providing input from the research community, as exemplified by the Computing Community Consortium, is a valuable voice within the IT ecosystem; something that would be helpful to initiatives in other domains.

Given the program's dedicated support and broad community, how does its leadership advance technical and policy directions for research and development for the future? A significant concern in our evaluation is that while the leadership clearly manages coordination and information-sharing, it does not appropriately lead

[36] The White House, Office of the Press Secretary, "Presidential Memorandum: Unleashing the Wireless Broadband Revolution," www.whitehouse.gov/the-press-office/presidential-memorandum-unleashing-wireless-broadband-revolution, 2010.

[37] PCAST, Executive office of the President, "Report to the President and Congress, Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology," pg. 19-20. https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd2013.pdf, 2013.

change so that NITRD achieves a strategic technical or policy vision agreed upon by the Subcommittee co-chairs, NSTC Committee on Technology, and OSTP that is in the best interests of the Nation.

We look at the challenges here from two points of view. First, our conversations with many Group participants revealed highly varied levels of engagement and productivity across the overall structure. We ascribe some of this variation to the aging of the topics and the de-prioritization of some topics within the overall IT community and to the introduction of new topics by some of the Groups. We also note, as has been recognized in the 2010 and 2013 PCAST NITRD reports,[38] that the Groups function best when senior level agency representatives participate and have sufficient responsibility to effect the actions proposed by the Group.

Second, and perhaps more important, our review found very little understanding or documentation of the processes by which Groups of all kinds are created, evaluated, and retired. IT has clearly evolved over the last twenty-five years. New topics emerge and others decline in importance; the NITRD organization needs to evolve in response. In some ways, the creation of the SSGs over the last seven years is one response to emerging themes. Some Groups that were clearly relevant at one point in time, however, be they IWGs, CGs, or SSGs, become lower priority, either because the primary work of the research community has progressed to where the amount of government-sponsored research is diminished, or because the topic has been supplanted or absorbed by newly emerging science and research. We see little evidence of ongoing discussions to examine such trends and to draw down various once-prominent themes.

Suggested Changes to Groups

The purposes of the NITRD program are best served if the Group process has a large degree of flexibility. It should be possible to have Groups at multiple levels of management responsibility and with varying durations. While one would expect that every PCA and every major initiative without its own PCA should have some form of coordination, PCAs and Groups need not have one-to-one correspondence. As an illustration, three Groups now coordinate various aspects of the CSIA PCA – CSIA (an IWG), VIA (a CG), and CSIA (an SSG). That coordination appears to be working well.

In many important areas, coordination efforts appear to be missing. There is no Group with budget authority that coordinates the present HEC I&A – we suggest that there should be such an SSG, perhaps co-chaired by the DOE ASCR Director of Operations and the NSF CISE Director of Advanced Cyberinfrastructure. No NITRD Group coordinates the National Robotic Initiative, or aspects of robotics such as human-robot interaction or autonomous robots. There is no Group for "Big Data." Broad interagency coordination of activities concerning aspects of IT and People is also absent.

Because of the pervasiveness of IT, some initiatives cross not only agency boundaries but also divisions of OSTP. Although the NITRD Subcommittee reports to the NSTC Committee on Technology, which is chaired by the head of the Technology Division, NITRD activities may touch on the interests of other OSTP divisions, whose staff are encouraged to work with NITRD as appropriate, even to the extent of helping to lead a group. There is already good history of National Security and International Affairs division staff working with NITRD on cybersecurity, for example.

---

[38] (1) PCAST, Executive Office of the President, Report to the President And Congress, Designing A Digital Future: Federally Funded Research and Development in Networking and Information Technology," https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf, 2010; (2)PCAST, Executive Office of the President, "Report to the President And Congress, Designing A Digital Future: Federally Funded Research and Development in Networking and Information Technology," https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd2013.pdf, 2013.

Despite the need for various kinds of coordination, the number of Groups cannot grow without bound. In addition to mechanisms to create Groups, there need to be ways to sunset or reconstitute Groups. It is the responsibility of the NITRD Subcommittee and the NCO to ensure that the existing Groups have strong leadership and vibrant agendas.

Recommendation 11: NITRD Groups

a. The NITRD Subcommittee, in collaboration with NSTC and OSTP, should establish specific language specifying what the purpose of each *type* of Group is and what mechanisms should be used to establish, monitor and terminate a Group. They should define a process to create a new Group, to set its charter, and to specify its correspondences with existing PCAs.

b. The NITRD Subcommittee, in collaboration with NCO and OSTP, should define a process and timeline for periodic review of each Group, with a recommendation for continuation, modification or sunset. A process for acting on those recommendations should be defined and executed. Reports on these reviews should be provided for each NITRD Review.

c. Each Group at the Senior Steering level should coordinate a process to publish and publicly discuss periodically a research and coordination plan for its area of interest.

Acknowledgments

Vincent Aleven
Associate Professor
Human-Computer Interaction Institute
Carnegie Mellon University

Seth Andrews
Senior Advisor
Office of the Chief Technology
Office of Science and Technology Policy

Christopher Atkeson
Professor, RI/HCII
Robotics Institute
Carnegie Mellon University

Tali Bar-Shalom
Program Examiner
Energy Branch
Office of Management and Budget

Francine Berman
Edward P. Hamilton Distinguished
Professor in Computer Science
Rensselaer Polytechnic Institute

Michael Bernstein
Associate Professor
Department of Computer Science
Stanford University

Jeffery Bigham
Associate Professor
Human-Computer Interaction Institute
Carnegie Mellon University

Randy Bryant
Assistant Director
Information Technology R&D
Office Of Science and Technology Policy

David Corman
Program Director
Cyber-Physical Systems
National Science Foundation

Jonathon N. Cummings
Associate Professor
The Fuqua School of Business
Duke University

Ann Drobnis
Director
Computing Community Consortium
Computer Research Association

Deborah Estrin
Professor of Computer Science
Joan & Irwin Jacobs Technion-Cornell
Institute
Professor of Public Health
Weill Cornell Medical College

Joseph Evans
Deane E. Ackers Distinguished Professor
Department of Electrical Engineering and
Computer Science
University of Kansas

Kevin Fu
Associate Professor
Department of Electrical Engineering and
Computer Science
University of Michigan

Susan Fussell
Associate Professor
Department of Information Science
Cornell University

Viktoria Gisladottir
PCAST intern
University of California-San Diego
Office of Science and Technology Policy

Saul Martirena Gonzalez
Assistant Director
Physical Sciences
Office Of Science and Technology Policy

Sol Greenspan
Program Director
Software Engineering and Languages Program
National Science Foundation

Suzi Iacono
Office Head (Acting)
Office of Integrative Activities
National Science Foundation

Farnam Jahanian
Vice President
Research
Carnegie Mellon University

Michael I. Jordan
Pehong Chen Distinguished Professor
Department of Electrical Engineering and Computer Science
University of California, Berkeley

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

Daniel Kaufman
Director
Information Innovation Office
Defense Advanced Research Projects Agency

Jim Kirby
Software Engineering Researcher
Center for high Assurance Computer Systems
Naval Research Laboratory

Kei Koizumi
Assistant Director
Federal Research and Development
Office of Science and Technology Policy

Robert Kraut
Herbert A. Simon Professor
Human-Computer Interactions institute
Carnegie Mellon University

Vijay Kumar
Nemirovsky Family Dean
Department of Electrical Engineering and Computer Science
University of Pennsylvania

Jim Kurose
Assistant Director
Directorate for Computer & Information Science & Engineering
National Science Foundation

Susan Landau
Professor
Department of Social Science and Policy Studies
Worcester Polytechnic Institute

Ed Lazowska
Bill & Melinda Gates Chair
Computer Science & Engineering
University of Washington

Peter Lee
Corporate Vice President
Microsoft Research

Tsengdar Lee
Program Manager
High End Computing
National Aeronautics and Space Administration

Robert W. Leland
Vice President Science & Technology
Chief Technology Office
Sandia National Laboratories

Celinda Marsh
Program Examiner
Science and Space branch
Office of Management and Budget

William "Brad" Martin
Research Advisor
High Confidence Software and Systems
National Securities Agency

Keith Marzullo
Director
Networking and Information Technology Research and Development
Office of Science and Technology Policy

Maja Materic
Professor and Chan Soon-Shiong Chair
Vice Dean for Research
University of Southern California

Douglas Maughan
CSD Director
Homeland Security Advanced Research Projects Agency
Department of Homeland Security

Sendhil Mullainathan
Professor
Economics
Harvard University

Klara Nahrstedt
Ralph and Catherine Fisher Full Professor
Computer Science Department
Director of Coordinated Science Laboratory
University of Illinois, Urbana Champaign

William Newhouse
Program Lead
National initiative for Cybersecurity Education
National Institute of standards and technology

Vern Paxson
Professor
Computer Science
University of California, Berkeley

Alex "Sandy" Pentland
Toshiba Professor of Media Arts and Sciences
MIT Media Lab

Tim Polk
Assistant Director
Cybersecurity
Office of Science and Technology Policy

Carlton Reeves
PCAST intern
University of Wisconsin-Milwaukee
Office of Science and Technology Policy

Jennifer Rexford
Gordon Y.S. Wu Professor
Department of Computer Science
Princeton University

Shankar Sastry
Dean
College of Engineering
University of California, Berkeley

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

Scott Shenker
Professor
Department of Electrical Engineering and Computer Science
University of California, Berkley

Darren Smith
Program Manager
High performance computing and communications Division
National Oceanic and Atmospheric Administration

Sylvia Spengler
Program Director
Division of Information and Intelligent Systems
National Science Foundation

Ram Sriram
Chief
Software and Systems Division
National Institute of Standards and Technology

George Strawn
Director
Federal Networking and Information Technology Research and Development National Coordination Office

Alex Szalay
Alumni Centennial Professor
Department of Physics and Astronomy
John Hopkins University

David Wagner
Professor
Computer Science
University of California, Berkeley

Albert Wavering
Chief
Intelligent Systems Division
National Institute of Standards and Technology

Katherine Yelick
Professor of Computer Science
Associate Laboratory Director
University of California, Berkeley

President's Council of Advisors on Science and Technology (PCAST)

www.whitehouse.gov/ostp/pcast