TESTIMONY OF DR. GREGORY D. HAGER
JOHNS HOPKINS UNIVERSITY

BEFORE THE HOUSE COMMITTEE ON SCIENCE, SPACE AND TECHNOLOGY
SUBCOMMITTEE ON RESEARCH AND TECHNOLOGY

HEARING ON
A REVIEW OF THE NETWORKING AND INFORMATION TECHNOLOGY RESEARCH AND DEVELOPMENT PROGAM

OCTOBER 28, 2015

I would first like to express my appreciation to Chairwoman Comstock, Ranking Member Lipinski, and the other members of the Subcommittee on Research and Technology for this opportunity to present my perspectives on the Networking and Information Technology Research and Development (NITRD) Program.

As the co-Chair of the NITRD working group of the President's Council of Advisors for Science and Technology (PCAST), I am pleased to share my perspective of the most recent review of the NITRD program with you. That perspective is also informed by my position as the Mandell Bellmore Professor of Computer Science at Johns Hopkins University, as well as my role serving as the Chair of the Computing Research Association's (CRA) Computing Community Consortium (CCC) – a partnership between the National Science Foundation and CRA that serves as a visioning body for the computing research community.

As you are aware, 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 information technology (IT) research and development to ensure continued United States leadership in this important area. This coordination is carried out by the NITRD subcommittee of the National Science and Technology Council (NSTC). By Executive Order, PCAST is charged with periodically reviewing the NITRD program.

To perform this most recent review, PCAST convened a working group consisting of seven experts from academia and industry. I co-chaired this group together with Dr. Susan Graham (a PCAST member, and professor emerita in Computer Science Division of the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley). 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.

My remarks today summarize and offer my perspective on the findings and recommendations of the report. These cover the state of research and development in IT research 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 the National Science and Technology Council (NSTC).

The Evolution of Computing Research

Before I present the recommendations contained in the report, I will first highlight how the field of information technology has evolved over the past three decades, as informed by the work of the working group.

Information technology 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 information technology. Information technology empowers scientific inquiry, space and Earth exploration, teaching and learning, and consumer buying and selling. It informs decision-making, national security, transportation, advanced manufacturing, and protection of the environment. The National Bureau of Labor Statistics projects that more than half of all new jobs in Science, Technology, Engineering and Mathematics (STEM) will be related to information technology.

An important driver has been the exponential advances in computing power – dubbed Moore's law – coupled with Dennard scaling which allows such advances to take place while consuming roughly constant power. As a concrete example, a device with the power of a 1980's supercomputer, which consumed nearly 200 KW of power, now fits in our pocket and consumes less than 10 watts. As a result, a significant fraction of the majority of Americans who own a smart phone are now carrying a 1980's supercomputer in their pocket.

These incredible advances in computing have also reshaped the field of computing itself. When the High Performance Computing Act 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.

Over the subsequent decades the notion of a computer system has come to encompass a much broader range of components – computers, networks, specialized hardware such as graphics accelerators, and the software that runs on them. Large computing systems consist of networks of computers and high-capacity storage systems. Sensors now play a larger role – many computers are now equipped for video, sound, provide location services via GPS chips, velocity information via accelerometers, and bearing via a digital compass. The range of applications that run on these systems has broadened to include many physical devices (from home appliances to cars to surgical robots), web services, and mobile apps.

As a result of these advances, data – its acquisition, transmission, management, and use, have come to play an ever-increasing role in computing research. Users of computing simultaneously grew from a small group of specialists to encompass a much larger and more diverse user community, fueled by enormous growth in the computing industry. As a result, the breadth and role of "applications" grew. This has led to modern computing research to embrace a far broader research agenda increasingly driven by the intimate interactions among computing devices, people, and the physical world.

Along with those transitions, the domains in which IT is used - health, transportation, manufacturing, robotics, societal computing, smart infrastructure, defense, and scientific discovery - have begun to shape IT R&D. These domains create new opportunities for our country – the emergence of autonomous vehicles that will serve an aging population or provide ac-

cess 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. In short, IT research now impacts many of our most pressing national priorities.

PCAST Review of the NITRD Program

In light of this evolving landscape for computing research, there are many important areas of IT research and development that fall under the NITRD program. The PCAST NITRD working group examined previous NITRD reviews (from 2010 and 2013), interviewed experts in a variety of areas, and ultimately chose eight key areas upon which to present findings and recommendations. They are as follows:

Cybersecurity: Concern about the security of computing systems has intensified with the widespread global interconnectedness enabled by the Internet. PCAST found that federal investment in at least five key research and development 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 – as systems are in use, they need ongoing mechanisms for authentication, authorization, data provenance and integrity checks, and powerful tools to automatically detect potential vulnerabilities; 3) systems resilience – improved methods to mitigate the effects of an attack; 4) implementation support – methods to formally express cybersecurity policies 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, so that both technical and non-technical mitigations are possible.

PCAST recommends that:

- 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.

- 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.

- 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: A growing community of IT researchers, primarily with support from NSF and 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. PCAST found that many opportunities for research 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.

PCAST recommends that:

- 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).

- 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.

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

IT and the Physical World: Research and commercial opportunities in systems that couple IT technologies with sensing and actuation are seeing the fruits of prior decades of research emerge into the broader view – for example automated driving capabilities, now emerging in the marketplace, originated in research that began over 30 years ago with support from NSF and DARPA. PCAST found that research in: 1) physical IT and human interaction; 2) physical IT and robust autonomy; 3) physical IT and sensing, 4) development of hardware and software abstractions for physical IT systems; and 5) trustworthy physical IT systems is needed to advance this field. Additionally, as new products and technologies are developed for IT-enabled sensing and acting in the physical world, it is important to put in place open standards and platforms that will encourage sharing of new technologies with and among the research community.

PCAST recommends that:

- 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.

- 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.

Cyber-human Systems: Computing is integral to Americans' work and personal lives and to the aims and processes of organizations, government and society. Yet, many aspects of cyber-human systems are not well understood and warrant further 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.

PCAST also found that interagency coordination has been effective in some areas of cyber-human systems such as visualization and team science, but has been only sporadic in other areas such as social computing, human-robot interaction, privacy, health informatics, and human learning and education.

PCAST recommends that:

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

- 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.

Privacy Protection: Privacy is an important human and societal value, and its protection is increasingly threatened by the growing amounts of online data. Advances in privacy research require collaboration among computer scientists, government and legal scholars, and behavioral and social scientists to inform both the design 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: 1) understanding and clarifying what is meant by "privacy", 2) automatically tracking the use of all forms of personal data, automating compliance checking, 3) devising methods to use private data without disclosing private information, 4) detecting, signaling, and mitigating information leakage and privacy violations as they occur, and 5) creating mechanisms, frameworks, and tools to enable system builders to construct privacy-preserving systems without themselves being privacy experts.

PCAST recommends that 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).

High Capability Computing: High capability computing continues to be critical to national defense, to discovery-based research in all fields of scientific endeavor, and to commerce. Advances in R&D have 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. Fundamental 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 recommends that:

- 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.
- 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 the senior steering groups (SSGs).

Big Data and Data-Intensive Computing: As the widespread use of computing grows 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, on the determination of causality from data, on better understanding 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. Attention should be paid to supporting data stewardship to mitigate losses of data and associated opportunities for machine learning, inference, and longitudinal studies.

PCAST recommends that

- 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.

- 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.

Foundational IT Research: All of the paradigm-shifting achievements in information technology that we enjoy today rest on years, and sometimes decades, of foundational IT 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 advances. 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.

PCAST recommends that

- 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.

Education and Training

Education and training directly impact the nation's ability to create new innovations and to translate them into products and services. 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. 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.

PCAST recommends that:

- 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 underrepresented populations.

- 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.

- 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.

NITRD Coordination

The working group reviewed how the 18 member agencies in the NITRD Program invest their contributions to efficiently and effectively expand cross-government research and development. Budgetary Program Component Areas (PCAs) are the investment amounts in technical areas of interest that are tracked for record keeping and budget analysis. OSTB and OMB introduced the initial set of PCAs in 1995, and many of them have remained virtu-

ally unchanged despite the evolution of the scientific and technical fields in IT as described above. PCAST found strong support in OSTP and OMB to establish a process to update the PCAs periodically to ensure PCAs align with current priorities in IT fields. In the report, the working group proposes eight new PCAs beginning in FY 2017 and recommends updating them every 5 or 6 years.

PCAST recommends that:

- 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.

- 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.

The NITRD Program includes multi-agency focus groups ("Groups") that traditionally had been aligned with (PCAs). In addition to 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. This review also 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 and 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.

PCAST recommends that:

- 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.

- 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.

- 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.

Overarching Themes and Observations

The review noted several cross-cutting themes within the many research areas that were reviewed.

The working group noted that an increasing spectrum of areas of IT span the 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. As a result, increased coordination and collaboration between fundamental research programs and mission agencies is an important means to facilitate that translation of fundamental research to practice. Likewise, academia-government-industry partnerships (sometimes termed public-private partnerships) will continue to play an important role. Indeed, 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.

Second, the working group noted that the IT research and education community is under increasing stress. 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. In addition, 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 a growing number of instances, academic research organizations are competing with private-sector companies for skilled people, at all levels, including both potential graduate students and current faculty. This places the nation's long-term research capabilities at risk by "eating our seed corn."

Finally, the working group found that Research in IT is increasingly interdisciplinary, requiring larger and more diverse research teams and researchers who can create and lead multi-disciplinary research teams.

In conclusion, the founding goal of the NITRD program was to coordinate federal investments in computing research to ensure wise investment of taxpayer dollars. PCAST found that the NITRD leadership manages coordination and information-sharing through 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. The growing breadth and impact of computing research further reinforces the need to leverage government investments by engaging the broadest possible spectrum of agencies throughout the federal government. PCAST strongly supports the mission and role of NITRD, and finds that NITRD leadership across the community is essential to lead change so that NITRD continues to achieve the strategic technical or policy vision that is in the best interests of the program and the country.

I would like to again thank the committee for this opportunity to report the findings of the report of the NITRD working group. I was honored to be chosen to co-chair the working group, I fully endorse its findings and recommendations, and I stand ready to help the committee in efforts to advance computing research to enhance the long-term competitiveness of our nation.