March 31, 2014

Ms. Nicole Wong
Deputy U.S. Chief Technology Officer
Attn: Big Data Study, Office of Science and Technology Policy
Eisenhower Executive Office Building
1650 Pennsylvania Ave. NW
Washington, DC 20502

**Re: Big Data RFI**

Dear Ms. Wong,

Unleashing the power of big data is a key priority for both government and industry, and it is in the interests of the US economy to establish a policy framework that promotes innovation and protects the interests of all stakeholders associated with or affected by the use of the data.

BSA | The Software Alliance ("BSA") appreciates the opportunity to present its views in response to the Request for Information (RFI) on the Obama Administration's comprehensive review of the ways in which ''big data'' will affect how Americans live and work. Our comments focus on the benefits of big data as well as the implications of collecting, analyzing and using such data for privacy, the economy, and public policy.

BSA is the leading global advocate for the software industry. It is an association of nearly 100 world-class companies that invest billions of dollars annually to create innovative software solutions that make enterprises more productive, competitive, and secure.[1] Through international government relations and educational activities, BSA helps build trust and confidence in digital networks and in the new technologies that drive the global economy.

[1] *BSA | The Software Alliance (www.bsa.org) is the leading advocate for the global software industry before governments and in the international marketplace. Its members are among the world's most innovative companies, creating software solutions that spark the economy and improve modern life. With headquarters in Washington, DC, and operations in more than 60 countries around the world, BSA pioneers compliance programs that promote legal software use and advocates for public policies that foster technology innovation and drive growth in the digital economy.*

*BSA's members include: Adobe, Apple, ANSYS, Autodesk, AVG, Bentley Systems, CA Technologies, CNC/Mastercam, Dell, IBM, Intel, Intuit, McAfee, Microsoft, Minitab, Oracle, PTC, Rockwell Automation, Rosetta Stone, Siemens PLM, Symantec, Tekla, and The MathWorks.*

20 F Street, NW, Suite 800
Washington, DC 20001
P 202-872-5500
W bsa.org

The role of information is evolving. BSA's response to this RFI reflects our member companies' evolving experience in developing the products and tools that are used to improve the reliability, security and trustworthiness of big data collection, storage and analytics.

Governments, companies and consumers across the world rely on data-driven products and services to derive actionable insights from the ever-expanding pool of digital information. These insights and information help governments better serve their citizens at lower cost; they help businesses grow and expand in unprecedented ways; and, they help consumers better their lives – from improving their personal health and fitness to increasing the efficiency of their commercial interactions with new and established enterprises.

As a threshold matter, BSA would draw attention to the RFI's attempt to define "big data" as "datasets so large, diverse, and/or complex, that conventional technologies cannot adequately capture, store, or analyze them." While we recognize that such a definition allows the Administration to focus on issues and concerns that stand in the way of tremendous potential problems, such a definition provides too narrow a perspective for the dramatic societal and economic gains that an examination of the true scope of big data affords.

We believe that a key feature of big data is that through the application of complex analytics we can identify patterns that can be converted into reliable predictions. These predictive outputs can be used to advance science, improve health, and save energy as well as for a variety of commercial undertakings. Such benefits already are all around us, and to suggest that big data lies beyond our current technological capabilities would be to deny the economic and personal value already being derived from big data analysis.

BSA does appreciate the Obama Administration's focus in the RFI: the opportunity to advance the policy discussion and the big data solutions for the most difficult cases. By examining the impediments here, BSA hopes the Administration can continue to promote interests and advances across the board.

Big data already is having major positive impacts across all aspects of the economy. Developments such as predictive analytics have helped traditional manufacturing companies save millions of dollars in testing and other costs. At the same time, the wealth of new data streams has helped spawn entirely new business lines and revolutionized existing industry models. By way of potential impact, one estimate finds that increased big data analytics could increase annual GDP in retailing and manufacturing by up to $325 million even as it produces cost savings of as much as $285 billion in health care and government services.[2]

In the research realm, big data has the potential to unlock tremendous societal advances ranging from medicine to energy efficiency and the environment. In the health care field, for example, scientists' ability to analyze the immense amounts of data that are increasingly available will help explain disease processes and could identify new treatments and cures.

[2] Game Changers: Five Opportunities for US Growth and Renewal, McKinsey & Co. (July 2013), available at: http://www.mckinsey.com/insights/americas/us_game_changers.

Such data can help identify illnesses – even before symptoms develop. It will help with early cancer detection and lead to much more effective treatments.[3]

It is important to note that this process is playing out against the backdrop of an international conversation about data privacy and international surveillance. All efforts should be made to avoid conflating government and commercial interests in data. Ensuring trust in the full spectrum of the big data environment is key to its success, and it requires that we fully separate the topics in order to avoid missing the promise of "big data."

**BSA Responses to the RFI's Questions to the Public**

**(1) What are the public policy implications of the collection, storage, analysis, and use of big data? For example, do the current U.S. policy framework and privacy proposals for protecting consumer privacy and government use of data adequately address issues raised by big data analytics?**

BSA appreciates the attention that the Obama Administration has given to data-related policy in the United States, particularly the work on the 2010 Commerce Department green paper,[4] the 2012 Federal Trade Commission white paper,[5] and the Administration's 2012 Privacy Bill of Rights.[6] In a sign of how quickly the big data economy is developing, none of these recent reports fully assesses the policy implications of the collection, storage, analysis and use of big data. The current White House effort can help close those gaps.

The Administration should ensure that any policy proposals related to big data carefully account for the varying types of data that will fall under any new policy umbrella. The rising tide of information that big data innovations is based on comes in a range of types and from a multitude of sources, including public data, sensor data, transaction information and demographic information. The type and source of this data matters because different types of data implicate a range of different policy concerns.

In recent years, much of the debate about privacy has focused on the ways data is collected and associated notice and consent rules. We believe this is a necessary element but does not fully fit the realities. In most instances, concerns arise when data is used in ways that threaten or cause harm to the individual. Thus, we believe privacy policies should specifically take into account the risk of harm that the exposure or misuse of the relevant data represents. Accordingly, policies should be tailored to account for such risks in context-specific circumstances. The most sensitive data must

[3] The Impact of Big Data on Medical Research, Sanford-Burnham Science Blog, Sanford | Burnham Medical Research Institute (June 25, 2013), available at: http://beaker.sanfordburnham.org/2013/06/the-impact-of-big-data-on-medical-research/#sthash.bUhWRLKu.dpuf.
[4] Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework (December 2010), available at: http://www.ntia.doc.gov/files/ntia/publications/iptf_privacy_greenpaper_12162010.pdf.
[5] Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers (March 2012), available at: http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.
[6] Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy, (February 2012), available at: http://www.commerce.gov/sites/default/files/documents/2012/february/privacy-final.pdf

be accorded the highest protections. For example, individualized health information in the care of a patient's personal physician cannot be treated in the same manner as weather information that is the product of federally funded meteorological research.

Even with within the narrow categories of information, the sensitivity of data can vary – ranging in the health care space, for example, from information gathered in long-term public health assessments to that in an oncologist's diagnostic files. Taking a risk-based approach that accounts for such variations is particularly important in the evolving big data environment, where the traditional focus on notice and consent fails to account for the ever-increasing size and scope of data sets.

In addition, as the Administration considers updating policy frameworks it should also take into account the steps that can be taken to protect data – and the underlying individuals that data represents. Already BSA member companies build in privacy protections to their systems from the point of inception. This practice of "privacy by design" ensures that companies thoroughly incorporate privacy protections into their products and services.

Such efforts begin with the development and use of adequate privacy and security standards to establish a responsible benchmark for business practices. By putting the appropriate emphasis on security, for example, companies can reduce the risk that data is improperly accessed or disseminated. In addition, by using anonymization, de-identification, and encryption tools a company can further minimize the impact of any breach.

**(2) What types of uses of big data could measurably improve outcomes or productivity with further government action, funding, or research? What types of uses of big data raise the most public policy concerns? Are there specific sectors or types of uses that should receive more government and/or public attention?**

The Administration already has taken positive steps on big data by increasing access to the government's existing data sets and encouraging additional focus on areas of great societal impact. The recently announced initiative to share climate data with the goal of helping the public better understand risks on coastal flooding serves as one example. BSA would encourage continued similar efforts aimed to produce the greatest potential societal gains, including around such areas as health, public safety, and education.

Government should not focus its limited resources on areas where consumers can more directly shape the emerging big data environment. For example, systems built on user-based preferences for personalization and self-selected preferences in the commercial space should be allowed to mature and evolve to fit consumer needs. Finally, the lowest priority should be given to business analytics where no individuals are exposed to risk or harm.

**(3) What technological trends or key technologies will affect the collection, storage, analysis and use of big data? Are there particularly promising technologies or new practices for safeguarding privacy while enabling effective uses of big data?**

Big data technologies are enabling and spurring a range of products and services to benefit consumers and governments. Data-rich mapping services already have turned our phones into navigation devices that can be leveraged for restaurant reviews or

guidance for homebuyers. Crime statistics are being used by law enforcement agencies for "predictive policing" that helps reduce crime even before it happens. Facial recognition technologies are being developed that could help improve security and public safety and help tailor customer services.

Big data developments will help create new products and services in innumerable sectors. The benefits of many of these advances are obvious. And while such developments also pose potential privacy and security challenges, the same technology that delivers these advances can help provide solutions to address such concerns. By anonymizing or de-identifying the underlying data, technology can diminish or even extinguish many threats. New encryption technologies enable this, as do methods of data analysis that limit the exposure of the data to interception.

BSA supports the recent efforts of the National Institute of Standards and Technology (NIST) to convene stakeholders to discuss the potential for privacy engineering to contribute to the development of effective and repeatable privacy protections.[7] In doing so NIST and industry can work together to develop a basis for the development of technical standards and best practices for the protection of individual privacy and civil liberties.

**(5) What issues are raised by the use of big data across jurisdictions, such as the adequacy of current international laws, regulations, or norms?**

Like the Internet, data does not "respect" political borders. Nor should we restrict the innovations enabled by data analysis. At the same time, we must respect that privacy norms vary in different markets, and policymakers and industry should work together to bridge the differences.

As the United States develops policies with respect to big data, we strongly recommend that any such policies should include considerations of the impact of the policy on cross-border data flows and big data uses. The full rewards of big data will best be achieved by enabling the broadest possible economies of scale and ensuring that the benefits of big data accrue to the widest possible global audience.

Recognizing that there is a fine line between protecting privacy and creating unnecessary burdens to the flow of information, the 21 countries of the Asia-Pacific Economic Cooperation (APEC) established the Data Privacy Pathfinder in 2007.[8] The goal of the Pathfinder is to "achieve accountable cross-border flow of personal information" by developing and implementing a system of cross-border privacy rules (CBPRs).

Those rules are consistent with the APEC Privacy Framework, which was developed in 2004 and aims to: 1) improve information sharing among government agencies and regulators; 2) facilitate the safe transfer of information between economies; 3) establish a common set of privacy principles; 4) encourage the use of electronic data as a means to enhance and expand business; and 5) provide technical assistance to those economies that have yet to address privacy from a regulatory or policy perspective.

[7] NIST Privacy Engineering Workshop (March 6, 2014). More information available at: http://www.nist.gov/itl/csd/privacy-engineering-workshop.cfm
[8] Asia-Pacific Economic Cooperation (APEC), APEC Data Privacy Pathfinder, at http://www.apec.org/About-Us/About-APEC/Fact-Sheets/APEC-Privacy-Framework.aspx.

This year, a joint effort between the US, APEC and the European Union created a tool to build even more-expansive bridges. A coalition of officials from the Federal Trade Commission, APEC and the European Union (EU) created the "Referential" a tool that maps the APEC CBPR's to the EU's Binding Corporate Rules.[9] This new document is intended to serve as a practical reference guide for companies that seek certification under both the APEC and EU systems. Both of these tools can serve as models for the creation of policy frameworks for enabling continued big data innovation.

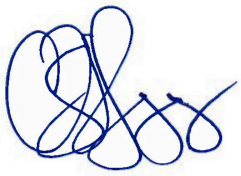
Looking more broadly, BSA believes that trade rules should be developed to enable and ensure the free flow of digital trade generally and the growth of big data services specifically. This entails covering innovative services in trade agreements, keeping borders open to the free flow of data and preventing mandates on where servers or other computing infrastructure must be located.

As the United States examines its policies on big data it should redouble its existing efforts on policy areas that impinge on cross-border data flows, and the government should look for new ways to improve international alignment on privacy and other data-related policy areas.

***Conclusion***

BSA appreciates this opportunity to comment on the Administration's Big Data RFI. We would be delighted to discuss these comments with the Office of Science and Technology Policy or to answer any questions about them.

Sincerely,



Christopher Hopfensperger
Director, Policy

[9] Asia-Pacific Economic Cooperation (APEC), "Joint work between experts from the Article 29 Working Party and from APEC Economies, on a referential for requirements for Binding Corporate Rules submitted to national Data Protection Authorities in the EU and Cross Border Privacy Rules submitted to APEC CBPR Accountability Agents," at http://www.apec.org/~/media/Files/Groups/ECSG/20140307_Referential-BCR-CBPR-reqs.pdf.