**Comments of the Internet Association in Response to the White House Office of Science and Technology Policy's Government "Big Data" Request for Information**[1]

The Internet Association[2] welcomes the opportunity to submit the following comment in response to the White House Office of Science and Technology Policy's (OSTP) request for information on "big data" to inform its 90-day review. OSTP seeks comments from interested parties to examine how "big data" will affect Americans' daily lives, the relationship between government and citizens, and how the public and private sectors can spur innovation and maximize the opportunities and free flow of information while minimizing the risks to privacy.[3] OSTP has a unique opportunity as part of this 90-day review to educate the public on the benefits of big data to society and how the government and companies are leveraging big data in a privacy-enhancing way consistent with the robust and effective privacy regime that exists in the United States (U.S.).

The current U.S. policy framework is critically important to the continued growth of the Internet ecosystem. It provides a framework by which companies respect and promote the privacy of the people who use their services, while allowing for technological advancements benefitting individual users and society. We encourage the Administration to highlight the

[1] The Internet Association comments electronically submitted at bigdata@ostp.gov.

[2] The Internet Association represents the world's leading Internet companies including: Airbnb, Amazon, AOL, eBay, Expedia, Facebook, Gilt, Google, IAC, LinkedIn, Lyft, Monster Worldwide, Netflix, Practice Fusion, Rackspace, reddit, Salesforce.com, SurveyMonkey, TripAdvisor, Twitter, Uber Technologies, Inc., Yelp, Yahoo!, and Zynga.

[3] White House Office of Science and Technology Request for Information, 79 Fed. Reg. 12251. (Mar. 4, 2014).



1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

advantages and successes of the existing framework in its 90-day report, just as it did in releasing the Consumer Privacy Bill of Rights in 2012.[4]

Finally, the Internet Association encourages the federal government to devote research and development towards determining solutions to make more government data available, fund research towards emerging technologies as well as support research to determine methods to educate consumers on "big data" practices.

## I. The White House Office of Science and Technology Policy should dedicate efforts to further exploration of methods and policies to effectuate government surveillance reform.

The Administration's review of this issue comes at a challenging time for the Internet industry. The ongoing revelations concerning the nature and scope of government surveillance programs create the potential for diminished user trust and confidence in Internet services. It is critical, therefore, that the Administration and Congress focus on policy solutions that are directly responsive to concerns that have surfaced in light of these revelations.

As it explores the policy frameworks under which "big data" issues should be examined, we urge the Administration to be cognizant that government surveillance and commercial privacy are separate and distinct issues. Given that Internet companies aim to provide transparency, choice, and control to consumers, efforts to conflate these issues are

[4] The White House, Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy at i (Feb. 2012), *available at* http://www.whitehouse.gov/sites/default/files/privacy-final.pdf.



1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

counterproductive, particularly given how little transparency citizens currently have when it comes to government surveillance.

As discussed in greater detail in Section III, our current public policy framework provides an effective and balanced approach in allowing for this innovation while protecting consumers. The success of existing laws and frameworks in guiding the private sector does not justify a need to initiate wholesale changes in how we approach these policies. Rather, we urge the Administration to continue its effort to address concerns that emerged in the aftermath of the NSA revelations. There have been real consequences from the disclosure of current surveillance practices, particularly outside the U.S. and the Administration should act swiftly to prevent declining confidence in U.S.-based Internet companies. Reforming surveillance law will help to rebuild user trust and to maintain the U.S.'s competitive edge in technological innovation. In the near term, the Administration can advance this debate by supporting the following reforms:

- **First, the Administration should endorse legislation pending in Congress that would update the Electronic Communications Privacy Act to** require governmental entities to obtain a warrant before they can compel online companies to disclose the content of users' communications. The Internet Association - along with more than 100 companies, trade associations, and civil society organizations - supports legislation pending in the House (H.R. 1852) and Senate (S. 607) that would update ECPA in this manner. Over 100,000 citizens have petitioned the White House to support this update to ECPA without exceptions that would whittle away at the bright-line, warrant-for-content rule that these bills would create. This update will provide certainty to both service providers and citizens that their content stored online will receive the same Fourth Amendment protections as their offline information.

- **Second, the Administration should build on the emerging consensus that U.S. law should prohibit the bulk collection of communications metadata to support comprehensive surveillance reform.** We understand from recent reports that the Administration intends to end the existing bulk telephony metadata collection program in favor of a legal regime that is narrowly tailored and subject to greater judicial oversight. The general thrust of Congressional legislation is consistent with this approach, and we believe it is critical that the bulk collection of Internet metadata be


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

encompassed within Congressional legislation. More remains to be done, and we urge the Administration to continue engaging Internet industry stakeholders on policy prescriptions that would form the basis for comprehensive government surveillance reform.

We encourage the Administration to lead the world in making reforms that ensure government surveillance efforts are clearly restricted by law, proportionate to the risks, transparent and subject to independent oversight. In so doing, the Internet Association recommends the Reform Government Surveillance principles.[5] By addressing these serious gaps in federal law, the government will demonstrate that it takes seriously its responsibility to protect the privacy of millions of Americans.

## II. The ability to leverage datasets in a privacy-protective manner has allowed for continued innovation and important advancements within the Internet industry. The government should devote research and development to generate additional innovative solutions and to educate consumers on government and commercial "big data" practices.

The Internet industry uses large datasets to enable activities that promote the public good and facilitate the creation of beneficial products and services, consistent with robust, existing legal frameworks that safeguard against harms arising from the misuse of personal data. For instance, cloud service providers' housing of datasets allows scientists and researchers to address societal changes relating to cancer research and climate change. Additionally, these datasets allows Internet companies to ensure the efficient operation of their platforms. The following examples illustrate these benefits:

- **Cloud services for scalable and reusable analytics.** With the increasing complexity of genomic research, scientists are using Amazon Web Services (AWS) as a platform to store and capture large volumes of scientific data in a cost-effective and timely manner.

[5] *See* Reform Government Survhttps://www.reformgovernmentsurveillance.com/ (last visited Mar. 31, 2014).



1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

AWS allows researchers to build scalable and reusable analytical tools. For instance, AWS hosts data for the 1000 Genomes Project, an international public-private effort that seeks to build the most detailed map of human genetic variation available. The project has grown to 200 terabytes of genomic data including DNA sequenced from more than 1,700 individuals that researchers can now access on AWS for use in disease research. The samples for the 1000 Genome Project are collected via informed consent and are mostly anonymous. AWS permits anyone with an account to access vast amounts of data to use in research to gain further insights into human health and diseases.

- **Improved communication on efficient and secure platforms.** It is common for a company to randomly distribute information across many servers. For example, to make its infrastructure more efficient, Facebook uses data analysis to intelligently distribute information by mapping data storage based on an understanding of how people communicate with their friends. This analysis protects privacy by relying on aggregated review of communication patterns at scale, and it not only promotes energy efficient infrastructure but also helps people communicate more reliably.

- **Improving users' daily lives.** Google Now helps Google users better manage their lives. With the affirmative consent of users that avail themselves of the service, Google Now uses information in the background to bring users the information they want, when they want it - showing the weather as you start your day, finding the best route to your next event to avoid traffic, telling you a flight is delayed or checking your favorite sports team's score. To do this, Google integrates information from the user with a number of back-end data sets, like maps, flight schedules, calendars, emails, weather, and public transit. Google Now also surfaces AMBER alerts, weather warnings, and other public alerts. Taking these many kinds of data - operational data, process data, statistical data, aggregated data, linguistic data, ethnographic data, and metadata - and linking them together to do something useful is one of the challenges of "big data," and provides significant value to Google users.

- **Spam filtering.** Also based on data analytics, Internet companies are able to enhance their ability to analyze message traffic to prevent spam, while leveraging de-identification technologies that reduce unwanted communications for all users without the need to maintain or share identifiable information about individual's communications.

These beneficial uses of big data in the Internet space are only the beginning. More work must be done in order for our society to gain the full benefits that can be achieved in areas of public health, economic growth, education, and social research from the analysis of large data sets. As a key component of its "big data" review, the Administration should commit to devoting substantial resources towards research and development aimed at unlocking the societal


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

benefits of large datasets, in both the private and public sectors. The Administration's current Open Data Initiative, intended to solve complex problems ranging from consumer to environmental issues, is a commendable first step at unleashing government data to fuel scientific discovery and spur innovative growth. For example, the Administration's recent launch of the National Oceanic and Atmospheric Administration (NOAA) open data website[6] will undoubtedly promote both of these goals by making publicly available scores of datasets containing valuable environmental data. Scientists will be able to harness this data to assess environmental risks, and start-ups will utilize it to provide critical products and services to consumers. While this initiative continuously seeks to make positive contributions, more needs to be done. The Administration should increase its efforts to make more government data readily accessible, which could create significant societal and economic benefits, including job creation. And, it should do this in a way that illustrates how large data sets can be used for research in a privacy-preserving manner.

Additionally, the National Science Foundation and other research funding should focus on areas of emerging research such as: (a) how to analyze big data effectively, (b) how to de-identify large data sets (e.g., privacy enhancing technologies), (c) how to build accountable systems, and (d) how to safely release data for research purposes. This research is just beginning, and it is critical to our long-term ability to attain big data benefits. The Administration should prioritize support for these areas of emerging research before seeking to circumscribe cutting

[6] *See* Press Release, NOAA, *NOAA announces RFI to unleash power of 'big data' Agency calls upon American companies to help solve 'big data' problem, available at* http://www.noaanews.noaa.gov/stories2014/20140224_bigdata.html (Feb. 24, 2014).


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

edge research in the private sector.

Lastly, although many of the discussions at workshops convened as a part of this "big data" review have focused on data collected by websites, the reality is that "big data" is collected in many contexts beyond websites, and its collection, analysis, and use does not depend on the existence of the Internet. The Internet of Things, mobile devices, wearable computing, and many other sectors and platforms are also involved in collecting data and performing large-scale analytics, and the retail sector has analyzed "big data" since before the commercial Internet existed. One key outcome of the Administration's work on "big data" should be an effort to educate people about the full range of "big data" practices – particularly in sectors that, unlike the Internet, may not be at the top of consumers' minds when thinking about privacy. In this regard, the government should focus research funding on efforts to pioneer methods that organizations, including government agencies, can use to help consumers gain meaningful understanding of how their data is collected and used. For example, the government can support usability studies aimed at discovering the best methods to inform people about the life cycle of their data, or how they can exercise control when organizations share their data with third parties.

## III. There is nothing dramatically new that would suggest a wholesale move away from our existing framework for regulating data, particularly given the breadth and effectiveness of federal and state enforcement as compared to regulatory regimes in other jurisdictions.

The U.S.'s flexible, multi-layered privacy regime is capable of responding forcefully to remedy violations of consumers' privacy, while permitting businesses that engage in privacy-protective practices to flourish. Under the current U.S. regime, organizations that engage in harmful information practices, in the big data context or otherwise, are subject to a wide range of


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

laws and regulations at both the federal and state levels. At the federal level, privacy laws protect information in the financial, insurance, educational, telecommunications, credit and health sectors, as well as information about children. At the state level, privacy laws cover these areas and more: employee data, spam, event data recorders, phishing and spyware. The U.S. also requires robust information security and data breach notification, which is the front line in preserving and protecting information privacy.

Beyond these sector-specific laws, the Federal Trade Commission Act and equivalent laws at the state level broadly prohibit "unfair or deceptive" acts or practices, and authorize enforcement actions by regulators. A defining feature of the U.S. commercial privacy regime is that it is calibrated to respond to the greatest public concerns. The Federal Trade Commission (FTC) and state attorneys general have acted on consumers' concerns about identity theft and data breaches by taking swift action to punish bad actors in the ecosystem and protect consumers who have been harmed. The FTC and state attorneys general are sensitive to emerging privacy and consumer protection concerns ranging from deceptive health claims to mobile tracking, the "Internet of Things," and "data brokers." Through panels, reports, investigations, consent decrees, and consumer education initiatives, U.S. regulators and law enforcement officials have proven remarkably adept at protecting consumer privacy in a balanced and agile manner that focuses administrative resources on the worst harms while allowing industry to consumers offer innovative products and services.

Nearly two years ago, the White House proposed a Consumer Privacy Bill of Rights, a framework that was intended to capture common privacy principles in a *comprehensive* way. The White House lauded the strength of the U.S. privacy regime when it released the Consumer Privacy Bill of Rights:


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

> "the consumer data privacy framework in the United States is, in fact, strong. This framework rests on fundamental privacy values, flexible and adaptable common law protections and consumer protection statutes, Federal Trade Commission (FTC) enforcement, and policy development that involve a broad array of stakeholders. This framework has encouraged not only social and economic innovations based on the Internet but also vibrant discussions of how to protect privacy in a networked society involving civil society, industry, academia, and the government."[7]

Since the release of the Consumer Privacy Bill of Rights, we have seen the continued and sustained success of privacy regulation at the federal and state levels, as well as the first successes arising from multi-stakeholder efforts to build new sectoral privacy improvements that are legally enforceable but would not have been feasible through legislative means.

We urge the White House to recognize in its 90-day report that our current legal framework in the U.S. can robustly address commercial data practices and is a flexible model for the continued growth of the innovation economy. Numerous jurisdictions are considering measures that would restrict the free flow of data, including data localization requirements and restrictive privacy provisions. Such measures would impede economic growth and even extinguish the promise of big data benefits. The Safe Harbor Framework— an important mechanism for U.S. companies that transfer data from Europe to the U.S.— is in danger of being scaled back or even suspended. Trade negotiations, in particular the Transatlantic Trade and Investment Partnership (T-TIP) with the EU have been adversely impacted by both the NSA revelations and a perception that the U.S. privacy regime is not as privacy-protective as the EU model.

The Internet industry appreciates the Administration's commitment to promoting the continued global competitiveness of U.S. businesses, and consistent with that commitment we urge the White House to uphold its responsibility to protect the economic interests of U.S.

---

[7] The White House, Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy at i (Feb. 2012), *available at* http://www.whitehouse.gov/sites/default/files/privacy-final.pdf.



1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org

industry by accurately characterizing the U.S. commercial privacy regime as fully protective of consumer privacy.

## IV. Conclusion

The Internet Association is pleased to provide input in response to OSTP's request for information on "big data." As the current U.S. public framework is balanced and effectively achieves its goal of protecting consumers while allowing for continued innovation, the system is not in need of wholesale changes. We hope that government surveillance reform remains a top priority for the Administration, and resources are devoted to making government data available and exploring emerging areas of research.

Respectfully submitted,

/s/Michael Beckerman
Michael Beckerman
President & CEO
The Internet Association

March 31, 2014


1100 H Street NW, Washington, DC 20005
www.InternetAssociation.org