

Before the
National Telecommunications and Information Administration
Department of Commerce
Washington, DC

In re

Big Data and Consumer Privacy in the
Internet Economy

Docket No. 140514424-4424-01

**COMMENTS OF
COMPUTER & COMMUNICATIONS INDUSTRY ASSOCIATION**

Pursuant to the request for comments issued by the National Telecommunications and Information Administration (NTIA) and published in the Federal Register at 79 Fed. Reg. 32,714 (June 6, 2014), the Computer & Communications Industry Association (CCIA) submits the following comments on the subject of big data and consumer privacy in the Internet economy.

I. Introduction

CCIA represents large, medium and small companies in the high technology products and services sectors, including computer hardware and software, electronic commerce, telecommunications and Internet products and services. CCIA members employ more than 600,000 workers and generate annual revenues in excess of \$465 billion.¹

CCIA commends NTIA for its framing in this request for comments, and the balanced discussion of the benefits of, and potential concerns pertaining to, the many modern information analysis strategies that are collectively referred to as “big data.” Maintaining this balance is very important, so as not to deter or stifle beneficial uses of data. The policy discussion regarding data should be focused on harms that occur from data misuse, rather than the collection or use of

¹ A list of CCIA members is available at <http://www.cciagnet.org/members>.

data itself. The implications of data collection depend greatly upon who it is using the data, under what terms, and for what purpose.

The PCAST Report is consistent with this perspective. Its first recommendation is that “[p]olicy attention should focus more on the actual uses of big data and less on its collection and analysis.”² Similarly, in the Administration’s Digital Government Strategy, one of the three main objectives is to “[u]nlock the power of government data to spur innovation across our Nation and improve the quality of services for the American people.”³ This document started from the important premise that data can be used for innovative purposes that enhance the lives of Americans, from disease control to public transit systems.⁴ It is encouraging to see the Administration continuing from this premise with this inquiry.

II. The Internet economy depends on the use of data.

As the title of this inquiry acknowledges, data is essential to the Internet economy. The value of the global Internet economy is projected to reach \$4.2 trillion within a few years,⁵ and as a recent McKinsey report put it, “it is increasingly the case that much of modern economic activity, innovation, and growth simply couldn’t take place without data.”⁶ Recognizing that the

² Executive Office of the President, President’s Council of Advisors on Science and Technology, *Big Data and Privacy: A Technological Perspective* (May 2014) (hereinafter “the PCAST Report”), at 2, available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.

³ White House, *Digital Government: Building a 21st Century Platform to Better Serve the American People* (May 2012), at 2, available at <http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government-strategy.pdf>.

⁴ *Id.* at 9-10.

⁵ David Dean et al., *The Internet Economy in the G-20: The \$4.2 Trillion Growth Opportunity*, Boston Consulting Grp. (2012), at 3, available at <https://www.bcg.com/documents/file100409.pdf>.

⁶ James Manyika et al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, preface, McKinsey Global Inst. (2011), available at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation.

continued growth of the Internet economy depends on data, regulators should take care not to stifle this growth, and should pursue policies that allow the Internet economy to flourish.⁷

III. Data has many economic and societal benefits.

Both the PCAST Report and the White House Big Data Report highlight the benefits of big data.⁸ Similarly, FTC Chairwoman Edith Ramirez recently spoke about the benefits of big data, starting her speech with a nuanced approach, considering “how we can reap the benefits of big data without falling prey to possible pitfalls.”⁹ She added that:

Big data is now, or soon will become, a tool available to all sectors of the economy. Of course, many uses of big data bring tangible benefits to consumers and businesses alike. And many uses of big data raise no threats to consumer privacy. For example, many firms use big data analytics for purposes that have nothing to do with individuals — forecasting weather and stock and commodity prices; upgrading network security systems; and improving manufacturing supply chains.¹⁰

Chairwoman Ramirez’s remarks highlight some of the potential economic and consumer benefits that applied uses of data can make possible.

⁷ Notably, NTIA’s mission includes “ensuring that the Internet remains an engine for continued innovation and economic growth.” NTIA, <http://www.ntia.doc.gov> (last visited July 29, 2014).

⁸ PCAST Report at x-xi (“The beneficial uses of near-ubiquitous data collection are large, and they fuel an increasingly important set of economic activities. Taken together, these considerations suggest that a policy focus on limiting data collection will not be a broadly applicable or scalable strategy – nor one likely to achieve the right balance between beneficial results and unintended negative consequences (such as inhibiting economic growth).”); Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values* (May 2014) (hereinafter “the White House Big Data Report”), at 3, *available at* http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf (“Big data is creating value for both companies and consumers. The benefits of big data can be felt across a range of sectors, in both large and small firms, as access to data and the tools for processing it are further democratized. . . . Technology companies are using big data to analyze millions of voice samples to deliver more reliable and accurate voice interfaces. Banks are using big data techniques to improve fraud detection. Health care providers are leveraging more detailed data to improve patient treatment. Big data is being used by manufacturers to improve warranty management and equipment monitoring, as well as to optimize the logistics of getting their products to market. Retailers are harnessing a wide range of customer interactions, both online and offline, in order to provide more tailored recommendations and optimal pricing.”).

⁹ FTC Chairwoman Edith Ramirez, *The Privacy Challenges of Big Data: A View from the Lifeguard’s Chair*, Technology Policy Institute Aspen Forum, Aug. 19, 2013, at 1, *available at* http://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf.

¹⁰ *Id.* at 3.

From transportation to education to hospitals, uses of big data have contributed substantially to innovation.¹¹ The PCAST Report lists many specific examples of big data's benefits, from law enforcement to commerce, healthcare to education.¹² IBM alone has demonstrated big data benefits in more than 10 different industries: automotive, banking, consumer products, energy and utilities, government, healthcare, insurance, oil and gas, retail, telecommunications, and travel and transportation.¹³ Other socially beneficial uses of big data range from developing tools to track flu trends,¹⁴ calculate the risks of climate change,¹⁵ and translate languages to aid victims of domestic violence.¹⁶

IV. Privacy frameworks should be use-based, rather than collection-based.

The Consumer Privacy Bill of Rights (CPBR), which is based on the Fair Information Practice Principles (FIPPs), can be reinvigorated to avoid stifling innovative uses of data while protecting consumers. This can be accomplished by focusing on preventing identified harmful uses of data, including those that are discriminatory, while allowing for data innovation that does not harm consumers.

¹¹ *The World's Top 10 Most Innovative Companies in Big Data*, FAST COMPANY, Feb. 10, 2014, available at <http://www.fastcompany.com/most-innovative-companies/2014/industry/big-data>.

¹² PCAST Report at 11-14.

¹³ IBM, *Big Data at the Speed of Business: Big Data in Action*, available at <http://www-01.ibm.com/software/data/bigdata/industry.html> (last visited July 30, 2014).

¹⁴ Thomas M. Lenard & Paul H. Rubin, *The Big Data Revolution: Privacy Considerations*, Technology Policy Institute (Dec. 2013), at 7, available at http://www.techpolicyinstitute.org/files/lenard_rubin_thebigdatarevolutionprivacyconsiderations.pdf (“The poster child for big data is Google Flu. Testing 450 million models, researchers identified 45 search terms that could predict the spread of flu more rapidly than the Centers for Disease Control, which relies on physicians’ reports. By tracking the rate at which the public searched for terms like “flu” and “cough medicine” using Google, an outbreak of influenza could be spotted a week or two ahead of CDC reports. Using data from internet searches for a service such as Google Flu was not and could not be envisioned when these data were collected.”) (citations omitted).

¹⁵ Katherine Noyes, *Big data's biggest challenge: climate change*, FORTUNE, June 23, 2014, available at <http://fortune.com/2014/06/23/big-data-climate-change-map-sea-levels/>.

¹⁶ Michael Hendrix, *How Google Translate Can Help Police Fight Domestic Violence*, U.S. Chamber of Commerce Foundation Blog, Feb. 6, 2014, available at <http://www.uschamberfoundation.org/blog/post/how-google-translate-can-help-police-fight-domestic-violence/34117> (“The use of services like Google Translate to help victims of domestic violence is merely one example of the good that data can do when it is put to work.”).

Using a harms-based framework is critical to allowing the development of new innovations that enhance the lives of consumers and increase public safety. Traditional FIPPs requirements, including concepts like pervasive notice and choice, purpose specification, and data minimization, would cripple innovation, especially serendipitous data innovation. Serendipitous uses of data typically involve unanticipated use of data sets generated for one purpose to solve unrelated problems.¹⁷ Requiring a company to obtain a user’s consent for socially beneficial or non-harmful uses of data would severely curtail innovative reuse of data, as such a requirement for any unanticipated use or unspecified purpose would make notice and choice a constant hurdle to ever-evolving research and development. Similarly, associating collected data with a specified or noticed purpose that requires deletion following the corresponding analysis would also prevent reuse and thereby reduce the frequency of serendipitous outcomes. The aggregate burdens of the traditional FIPPs requirements may so increase the costs of obtaining socially beneficial uses from collected data that would-be innovators might simply not attempt to develop such uses at all.

Socially beneficial innovations based on consumer data are far from a new phenomenon; Ford introduced the seat belt in the 1950s after analyzing data from crashes.¹⁸ More recently, technology has increased the possibilities for unanticipated uses of existing data to achieve new insights. A potential cost associated with over-regulating the use of data is inhibiting these future serendipitous uses of data. A number of such uses have been undertaken by government agencies, and many have helped to improve healthcare, including research about causes of

¹⁷ The PCAST Report also described this phenomenon as “new, non-obvious, unexpectedly powerful uses of data.” PCAST Report at 38.

¹⁸ *The Fog of War: Transcript*, available at http://www.errolmorris.com/film/fow_transcript.html (former Ford President Robert McNamara explains how Ford introduced seat belts after analyzing data from thousands of deaths and millions of injuries).

cancer, side effects of drugs, and other health risks.¹⁹ For example, the Danish Cancer Society was able to study whether cell phone use increased the risk of cancer by using both Denmark’s national registry of cancer patients and cell phone subscriber data.²⁰ Yet when Danish wireless carriers were compiling this subscriber information, they could not have anticipated that it would have some valuable application in improving public health. If that subscriber data were restricted solely to uses relating to improving network reliability and customer billing, for example, these gains might not have been achievable. Restricting data collection will foreclose incredible opportunities such as these to improve healthcare, security, and public safety—not to mention to contribute to the growth of the Internet economy.

Harm-based limitations on data use can also effectively address concerns about discrimination that may be enabled by big data technologies. While the White House Big Data Report warns that “big data could enable new forms of discrimination and predatory practices,”²¹ the authors also understood that information gleaned from big data analysis can also be used to fight discrimination or help underserved populations.²² A framework based on restricting collection of data that might lead to discrimination would foreclose the opportunity for such socially beneficial uses, which might be reliant on the exact same types of data.

For example, Pigeonly, a startup that reduces costs of external communication for prison inmates, is based around a 50-state prisoner database.²³ Inmates that communicate more

¹⁹ Lenard & Rubin, *The Big Data Revolution: Privacy Considerations*, *supra* note 14, at 6 (“The examples of the serendipitous use of data are numerous. . . . The FDA used Kaiser Permanente’s database of 1.4 million patients to show that the arthritis drug Vioxx increased the risk of heart attacks and strokes. The Centers for Disease Control combine airline records, disease reports, and demographic data to track health risks.”) (citations omitted).

²⁰ *Id.*

²¹ White House Big Data Report at 53.

²² *Id.*

²³ Marco della Cava, *Pigeonly’s CEO helps prison inmates*, USA TODAY, Aug. 2, 2014, available at <http://www.usatoday.com/story/money/business/2014/08/02/pigeonly-frederick-hutson-tech-start-up-prison/13324873/> (describing Pigeonly’s services that deliver low-cost hard copy prints of photos to inmates and provide cheaper local numbers for inmates’ families).

frequently with the outside world while incarcerated are more successful at acclimating upon release, which reduces recidivism rates.²⁴ While inmate location and incarceration data can be used to discriminate against prisoners upon release, Pigeonly demonstrates that the same information can also be used to aid an underserved population in a way that benefits society. This socially beneficial use would not be feasible if inmate data availability were restricted, but a framework that accounts for the potential benefits and identified harms of a particular use of data would weigh in its favor. Frameworks that incorporate nuanced and flexible contextual analyses of the harms of particular uses of collected data would effectively account for the risk of discrimination, while still encouraging uses that are novel and beneficial.

Accordingly, limitations on data collection and use should be grounded in preventing demonstrable harms. Regulation should, wherever possible, focus on harms emanating from the misuse of data, not on the collection or use of data itself. As James Cooper put it in comments to the Office of Science and Technology Policy (OSTP) several months ago, “[b]efore relying on intangible harms as a justification for restrictions on big data, however, policy makers should have a firm grasp on their variance and magnitude.”²⁵

V. Existing and emerging privacy enhancing technologies can mitigate privacy risks to individuals while preserving the benefits of big data.

A harms-based framework on the use of data would encourage companies to seek out ways to minimize the risk of harm. As a result, such a framework would encourage companies

²⁴ *Id.* (“A successful re-entry is always linked to how well an inmate kept in touch with the outside world. To the extent that a company (like Pigeonly) can mitigate the harsh and stressful world of prison and give people that sense of self through contact, that is very positive.”).

²⁵ James C. Cooper, *Comment, Office of Science and Technology Policy Big Data Study*, George Mason University School of Law Law & Economics Center, OSTP, Mar. 31, 2014, at 2, available at [http://masonlec.org/site/rte_uploads/files/Cooper%20OSTP%20Big%20Data%20Comment%20\(FINAL%203.31\).pdf](http://masonlec.org/site/rte_uploads/files/Cooper%20OSTP%20Big%20Data%20Comment%20(FINAL%203.31).pdf). See also CCIA, *In re Mobile Security Project*, FTC, May 30, 2014, at 5, available at <http://www.ccianet.org/wp-content/uploads/2014/06/CCIA-Mobile-Security-Project-Cmts-May-2014.pdf> (“Evidence-based policy must be favored over government intervention in any nascent industry based on merely speculative harms.”).

to implement existing best-practices for managing consumer data, and also spur them to develop new tools that are more responsive to varying consumer perspectives than prescriptive standards.

Accordingly, analyzing a use of a consumer data set should account for existing harms-reducing privacy controls used by companies, such as data de-identification. De-identified data, which has been made not reasonably linkable to a particular consumer or device, retains much of its analytic value and can still lead to beneficial uses or serendipitous outcomes. The FTC's 2012 report finalizing its Privacy Framework explained how data can be properly de-identified, a process which requires an enforceable public commitment by companies to not re-identify data, and mandates contractual restrictions for downstream recipients of de-identified data to preempt attempted re-identification.²⁶ Proper de-identification of data effectively eliminates the risk of harm from collected data, and is thus excluded from the scope of FTC's Privacy Framework.²⁷ Use of de-identified data accompanied by the FTC's recommended safeguards is a practice that would be encouraged in a framework that limits identifiable harms and promotes uses that mitigate risks.

The harms-based approach to uses of consumer data also promotes the development of responsive privacy controls, as evidenced by the expansion of the marketplace for tools centered on consumer privacy. Ephemeral messaging apps like Snapchat, which are designed to “delete by default,” are growing in popularity because they allow consumers agency in sharing their information.²⁸ Similarly, the growing range of startups that enable consumers to manage and

²⁶ FTC, *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers* (March 2012), at 21-22 available at <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.

²⁷ *Id.* (implying that efforts to de-identify data in accordance with the FTC's requirements conform with the FTC's principle of “privacy by design”).

²⁸ Parmy Olson, *Delete By Default: Why More Snapchat-Like Messaging Is On Its Way*, FORBES, Nov. 22, 2013, <http://www.forbes.com/sites/parmyolson/2013/11/22/delete-by-default-why-more-snapchat-like-messaging-is-on-its-way/>.

benefit from the use of their personal data through storage and selective sharing tools gives companies the incentive to ensure that they limit harmful uses of aggregate data.²⁹ A privacy framework focused on harms reduction will ensure that organizations work to use consumer data in innovative and beneficial ways.

VI. Conclusion

CCIA encourages the study of this important subject by the Administration. Given the importance of data use in the Internet economy, and the economic benefits that can result from current and future applications of data, regulators should avoid stifling innovation in the public and private sectors with overly prescriptive measures.

August 5, 2014

Respectfully submitted,

Ali Sternburg
Public Policy & Regulatory Counsel
Dan O'Connor
VP, Public Policy
Bijan Madhani
Public Policy & Regulatory Counsel
Computer & Communications Industry
Association
900 Seventeenth Street NW, 11th Floor
Washington, D.C. 20006
(202) 783-0070

²⁹ Shane Green & Ann Cavoukian, *A Rising Tide of Data, Partnered With Privacy By Design, Will Lift All Boats*, Disruptive Competition Project, Apr. 23, 2013, at <http://www.project-disco.org/privacy/042313-a-rising-tide-of-data-partnered-with-privacy-by-design-will-lift-all-boats/> (highlighting a trend of new private and public services focused on the sharing of consumer and public data).