Before the
Federal Trade Commission
Washington, DC

In re

"Big Data: A Tool for Inclusion or
Exclusion?" Workshop

Project No. P145406

COMMENTS OF
COMPUTER & COMMUNICATIONS INDUSTRY ASSOCIATION

Pursuant to the request for comments[1] issued by the Federal Trade Commission (FTC)

prior to the public workshop entitled, "Big Data: A Tool for Inclusion or Exclusion?" the

Computer & Communications Industry Association (CCIA) submits the following comments on

the subject of the effects of big data on low income and underserved consumers.

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.[2]

CCIA commends the Commission 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

---

[1] *Notice of Workshop and Request for Public Comments*, Project No. P145406, *available at*
https://ftcpublic.commentworks.com/ftc/bigdataworkshop/.
[2] A list of CCIA members is available at http://www.ccianet.org/members.

discussion regarding data should be focused on mitigating harms that occur from data misuse, rather than the collection or use of data itself. The implications of data collection depend greatly upon who it is using the data, under what terms, and for what purpose.

The report issued by the President's Council of Advisors on Science and Technology (PCAST) 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."[3] 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."[4] 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.[5] The Commission's focus on potential harmful and beneficial uses of data is encouraging because it is consistent with the PCAST Report's primary recommendation.

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

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,[6] and as a recent McKinsey report put it, "it is increasingly the case that much of modern economic activity, innovation, and growth simply

---

[3] 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.

[4] 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.

[5] *Id*. at 9-10.

[6] 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.

couldn't take place without data."[7] Recognizing that the continued growth of the Internet

economy depends on data innovation, regulators should take care not to stifle this growth, and

should pursue policies that allow the Internet economy to flourish.

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, 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] 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.

[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.   Responsibly used big data tools are socially beneficial and can help combat discrimination and aid underserved populations.

Chairwoman Ramirez envisions that the Commission's upcoming workshop "will examine the potentially positive and negative effects of big data on low income and underserved populations."[17] The most effective means of reaping the social and economic benefits of big data while protecting consumers from potential negative effects is the development of responsible use

---

[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.").

[17] Press Release, FTC, *FTC to Examine Effects of Big Data on Low Income and Underserved Consumers at September Workshop* (Apr. 11, 2014), *available at* http://www.ftc.gov/news-events/press-releases/2014/04/ftc-examine-effects-big-data-low-income-underserved-consumers.

---

frameworks. A use-centric privacy framework would focus on preventing identified harmful

uses of collected data, including those that are discriminatory, while allowing for data innovation

that does not harm consumers.

Using a harms-based framework is critical to allowing the development of new, socially

beneficial innovations that enhance the lives of consumers and increase public safety. Such

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 kinds of future serendipitous

uses of data.

Serendipitous uses of data typically involve unanticipated use of data sets generated for

one purpose to solve unrelated problems.[19] A number of such uses have been undertaken by

government agencies, and many have helped to improve healthcare, including research about

causes of cancer, side effects of drugs, and other health risks.[20] 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.[21] 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

---

[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).

[19] The PCAST Report also described this phenomenon as "new, non-obvious, unexpectedly powerful uses of data." PCAST Report at 38.

[20] Lenard & Rubin, *The Big Data Revolution: Privacy Considerations, supra* note 15, 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).

[21] *Id.*

---

customer billing, for example, these gains might not have been achievable. Restricting data collection or limiting use to previously anticipated research 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.

A responsible use framework for big data that restricts identified harmful uses can also effectively address concerns about the negative effects of big data technologies on low income and underserved consumers. While the White House Big Data Report warns "big data could enable new forms of discrimination and predatory practices,"[22] the authors also understood information gleaned from big data analysis can also be used to fight discrimination or help underserved populations.[23] 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.

Big data has been regularly used to aid underserved populations and to combat traditional forms of discrimination. The clearest example of the latter comes from the NYPD's meticulous tracking of demographic data from its "stop-and-frisk" program. Analysis of the collected data confirmed anecdotal evidence that the program had a racially disproportionate effect,[24] and the lessons from stop-and-frisk can be used to shape future programs to ensure that they are fairly targeted. Analysis of large public data sets can also help improve health outcomes in underserved groups. The current lack of data on LGBT communities makes it difficult to accurately gauge health disparities and tailor programs.[25] In response, various New York state

---

[22] White House Big Data Report at 53.
[23] *Id.*
[24] Maurice Mitchell, *Big Data Can Be Used for Good in the Community*, N.Y. TIMES, Aug. 6, 2014, *available at* http://www.nytimes.com/roomfordebate/2014/08/06/is-big-data-spreading-inequality/big-data-can-be-used-for-good-in-the-community.
[25] Robert Graham *et al.*, *The Health of Lesbian, Gay, Bisexual, and Transgender People Building a Foundation for Better Understanding*, Institute of Medicine of the National Academies (Mar. 2011), at 1-2, *available at*

agencies have set about collecting data that might have once been used to discriminate or

stigmatize the LGBT population to now improve the provision of public health services.[26] Big

data is likewise used to enable civic engagement in underrepresented communities: the New

York State Civic Engagement Table merges publicly available voter and public housing data to

increase agency in the least politically active developments.[27] Similar use of unrelated public

employment data and social media allows third parties to develop tools that allow more effective

employment recruiting of populations underrepresented in the workplace or particular fields.[28]

In fact, entire business models are being built around the idea of using big data tools to

aid underserved groups for broad social benefit. For example, Pigeonly, a startup that reduces

costs of external communication for prison inmates, is centered on a 50-state prisoner database.[29]

Inmates that communicate more frequently with the outside world while incarcerated are more

successful at acclimating upon release, which reduces recidivism rates.[30] 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 was 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

---

http://www.iom.edu/~/media/Files/Report%20Files/2011/The-Health-of-Lesbian-Gay-Bisexual-and-Transgender-People/LGBT%20Health%202011%20Report%20Brief.pdf.

[26] Laura Nahmias, *State agencies launch LGBT data-collection effort*, CAPITAL NEW YORK, July 24, 2014, *available at* http://www.capitalnewyork.com/article/albany/2014/07/8549536/state-agencies-launch-lgbt-data-collection-effort.

[27] Mitchell, *Big Data Can Be Used for Good in the Community*, *supra* note 24.

[28] Jon Bischke, *Announcing Entelo Diversity*, Entelo (Apr. 30, 2014), *available at* http://blog.entelo.com/announcing-entelo-diversity.

[29] 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).

[30] *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.").

---

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."[31]

V.   Conclusion

CCIA encourages the study of this important subject by the Commission. 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 15, 2014   Respectfully submitted,

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

---

[31] 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. *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.").