

Before the
Federal Trade Commission
Washington, DC

In re

“Big Data: A Tool for Inclusion or
Exclusion?” Workshop

Project No. P145406

**SUPPLEMENTAL COMMENTS OF
COMPUTER & COMMUNICATIONS INDUSTRY ASSOCIATION**

Pursuant to the request for comments¹ issued by the Federal Trade Commission (“FTC” or “Commission”) prior to the public workshop entitled, “Big Data: A Tool for Inclusion or Exclusion?” (“big data workshop” or “the workshop”), the Computer & Communications Industry Association (“CCIA”) submits the following supplemental comments regarding the Commission’s workshop on 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 hosting the public workshop on the many modern information analysis strategies that are collectively referred to as “big data.” As the field of “big data” expands quickly both in the United States and internationally, the workshop was a timely

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 <https://www.cciagnet.org/members>.

and important opportunity for stakeholders to gather together to examine the many benefits and potential concerns that arise in this context.

Several noteworthy themes developed over the course of the daylong workshop. The most significant involved the policy discussion around whether to regulate both data collection and data use, or only to prohibit data misuse in order to mitigate harms to consumers. It became clear during that workshop that the regulation of data collection would constitute an *ex ante* limitation on the many beneficial data applications that could serve the same communities that were at the center of the workshop—without yielding any material privacy protections for consumers. On the other hand, panelists spoke at length of the potential for data, once collected, to be used for beneficial purposes, particularly for low income and underserved consumers, and underscored the need to consider careful regulation of data misuse without deterring or stifling beneficial uses of data.

II. A use-centric framework that prevents identified harmful uses of collected data would allow for data innovation that does not harm consumers.

The big data workshop confirmed the need for a harms-based approach to managing uses of big data. Such an approach, which CCIA advocated in its comments prior to the workshop, would focus on mitigating harms that occur from data misuse, rather than the collection or use of data itself.³ Numerous other stakeholders presented concurring perspectives,⁴ which highlight that a framework based on promoting responsible uses and avoiding those that are harmful is the best way to enable the social and economic benefits of big data tools. This view is shared by the

³ See Comments of CCIA, Project No. P145406, F.T.C. (2014), *available at* <http://www.ccianet.org/wp-content/uploads/2014/08/CCIA-Comments-FTC-Project-No.-P145406-8-15-2014.pdf>.

⁴ See *generally* Comments of Center for Data Innovation, Project No. P145406, F.T.C. (2014), *available at* http://www.ftc.gov/system/files/documents/public_comments/2014/08/00026-92395.pdf; Comments of The Internet Association, Project No. P145406, F.T.C. (2014), *available at* http://www.ftc.gov/system/files/documents/public_comments/2014/08/00019-92375.pdf; Comments of Internet Commerce Coalition, Project No. P145406, F.T.C. (2014), *available at* http://www.ftc.gov/system/files/documents/public_comments/2014/08/00020-92376.pdf; and Comments of U.S. Chamber of Commerce, Project No. P145406, F.T.C. (2014), *available at* http://www.ftc.gov/system/files/documents/public_comments/2014/08/00021-92389.pdf.

President’s Council of Advisors on Science and Technology (PCAST), which recommended in its report that “[p]olicy attention should focus more on the actual uses of big data and less on its collection and analysis.”⁵

We can observe the efficacy of a harms-based approach in the example of Boston’s Street Bump app, which collects location and accelerometer data from drivers’ smartphones to report potential potholes to municipal authorities. During the workshop panelists acknowledged that, at first, the app collected a disproportionate amount of data about potholes in higher-income neighborhoods, where residents were more likely to own smartphones and download the app.⁶ To a lesser extent, panelists later touched on the subsequent versions of the app that worked to remedy the disparity in data collection.⁷ It is important to recognize that the process of updating of the app illustrates the how a harms-based approach to big data uses can work to avoid discrimination. While the app initially collected data in a way that led to an income-based disparity in city services, developers detected the anomalies and modified the app’s design to mitigate that harm so the app could serve the city’s purposes of providing more consistent and equitable road repair services.⁸ Any big data approach that would restrict the collection of data, for fear that it could potentially be used in a discriminatory or unfair way, would have prevented the development of the Street Bump app in the first place—which would deny the public the broad economic benefits the app’s data now provides.

⁵ Executive Office of the President, President’s Council of Advisors on Science and Technology, *Big Data and Privacy: A Technological Perspective* (May 2014), at 2, available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.

⁶ Transcript of *Big Data: A Tool for Inclusion or Exclusion*, Project No. P145506, F.T.C. at 13, 22, 258 (Sept. 15, 2014) (hereinafter “Workshop Transcript”), available at http://www.ftc.gov/system/files/documents/public_events/313371/bigdata-transcript-9_15_14.pdf.

⁷ *Id.* at 259.

⁸ See Phil Simon, *Potholes and Big Data: Crowdsourcing Our Way to Better Government*, WIRED, Mar. 25, 2014, available at <http://www.wired.com/2014/03/potholes-big-data-crowdsourcing-way-better-government/>.

The Street Bump app’s release cycle also demonstrates how a harms-based approach to big data uses can support the important iterative development process. No developer, consumer, or privacy advocate can consistently and accurately predict the harms that might result from a particular use of big data. Only in practice can actual harms be observed and corrected; aiming to prevent *all* potential harms through *ex ante* collection restrictions will merely limit all possible benefits. In the interest of allowing big data tools to lead to demonstrated public benefits, stakeholders must be understanding of the often complex, iterative development processes that big data tools entail.

A. There is scant evidence that big data uses cause actual harms to low income and underserved consumers.

Beyond the focus on the growing pains of the Boston Street Bump app, the workshop failed, unfortunately, to present strong evidence of concrete harms that big data uses pose to low income and underserved consumers—and whether current regulatory regimes are incapable of addressing any concrete harms. Presenters frequently speculated as to what might happen when large data sets are combined to build predictive models of consumer behavior or to classify groups of individuals for business purposes,⁹ but did not offer examples of those activities actually harming at-risk communities or individuals. Policy cannot be developed in a vacuum—consumer protection regulation must protect consumers from substantive dangers backed up by real-world evidence. In the absence of examples of such big data harms, it does not make sense to develop a regulatory framework based on preventing data collection, which would preclude the very real benefits of big data uses.

The lack of evidence of harm also demonstrates the need for more collaborative work in exploring the big data landscape. The White House Big Data Report emphasized that a range of

⁹ See Workshop Transcript at 32-33 (describing hypothetical inferences about consumers that could be made by predictive big data models); *id.* at 88-89 (detailing potential harms resulting from the classification of individuals by predictive models).

agencies involved in civil rights and consumer protection should be involved in assessing the potential for discrimination from big data uses.¹⁰ The Commission’s workshop had a participant from just one other federal agency, the Equal Employment Opportunity Commission, and its stakeholder representation was limited to privacy and industry advocates and a single representative from the civil rights community.¹¹ Representatives from across the federal government and subject-matter experts in civil rights, ethics, economics, and data science should contribute to a full-assessment of the benefits and concrete discriminatory risks of big data tools.

B. Disclosure of the means of big data analysis is neither helpful nor necessary.

Given the dearth of concrete harms presented by the use of big data tools, there is little reason to mandate that developers of big data tools, under the banner of “algorithmic transparency,” disclose the technological means through which they conduct their analysis. The algorithms, statistical tools, and practices employed to perform analysis of large data sets are regularly proprietary and are often trade secrets—as such there must be compelling reasons for their disclosure to regulators beyond speculation about potential harms from their use.

In addition, disclosure of the particulars of the tools used in big data analyses would be unhelpful to those examining them. Data analysis algorithms are complex formulas that evolve and learn quickly and require a large number of variable inputs.¹² Attempting to determine in a vacuum whether a particular snapshot of an algorithm is harmful to consumers is not illustrative of the actual use cases, and may not even reflect the version ultimately implemented. Big data

¹⁰ Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values* (May 2014), at 65, available at http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

¹¹ See *Big Data: A Tool for Inclusion or Exclusion*, FTC (Sep. 15, 2014), available at <http://www.ftc.gov/news-events/events-calendar/2014/09/big-data-tool-inclusion-or-exclusion>.

¹² See Workshop Transcript at 53-54 (“And it’s not only just not visible to outsiders, it’s often not visible to the actors themselves as they’re trying to do a lot of the predictive analytics that they’re working on. Right. We’re working with complex learning algorithms . . .”). See also *Algorithms*, GOOGLE | INSIDE SEARCH, at <http://www.google.com/insidesearch/howsearchworks/algorithms.html> (noting that Google’s search algorithms rely on over 200 unique signals to respond to user queries, and that over 665 improvements to search were rolled out to users over 2012).

tools are reliant on real-world context: as demonstrated by the revised Street Bump app, it is only in the wild that their beneficial and harmful uses can be observed.

III. Ensuring that there is available data about underserved and low-income consumers is the best means of preventing discriminatory impact in big data use.

Rather than encouraging unfair treatment of underserved and low-income consumers, prevailing evidence has shown that big data is regularly used to aid underserved populations and combat traditional forms of discrimination.¹³ In fact, in many instances it is a distinct lack of information that leads to disparities in benefits for these communities.¹⁴ Public and private organizations that use big data can provide improved outcomes in health, education, and economics, but such services can only be tailored to the needs of underserved consumers if there is available information about those groups. As Nicol Turner-Lee of the Minority Media and Telecommunications Council noted during the workshop, vulnerable minority populations will be left out of the positive uses of big data if they are unable to leave data footprints.¹⁵

For example, many low-income individuals have traditional credit histories that paint an incomplete or negative picture of their financial situations, and thus often find it difficult to obtain credit.¹⁶ To serve these consumers—who would otherwise be left out of credit markets—one lending startup offers them the opportunity to produce other types of meaningful data points that have not been traditionally captured, and then uses that data to make better-informed predictions about borrowers' creditworthiness for higher value loans.¹⁷ To encourage widespread access to positive types of big data use, including more accurate credit evaluation for

¹³ See, e.g., *Big Data: A Tool for Fighting Discrimination and Empowering Groups*, FUTURE OF PRIVACY FORUM (2014), available at <http://www.futureofprivacy.org/wp-content/uploads/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-Report1.pdf>.

¹⁴ See generally Daniel Castro, *The Rise of Data Poverty in America*, CENTER FOR DATA INNOVATION (2014), available at <http://www2.datainnovation.org/2014-data-poverty.pdf>.

¹⁵ Workshop Transcript at 100-102.

¹⁶ Jake Rosenberg, *Extending Credit To Sub-Prime Customers Through Big Data*, N.Y. TIMES, Aug. 6 2014, available at <http://www.nytimes.com/roomfordebate/2014/08/06/is-big-data-spreading-inequality/extending-credit-to-sub-prime-customers-through-big-data>.

¹⁷ *Id.*

underbanked consumers, regulators should avoid restricting the collection of data about underserved populations and instead seek to educate consumers about the importance of making data available to responsible users.

IV. Conclusion

CCIA was encouraged by the outcomes of the Commission's workshop on big data. Given the economic and social benefits that can result from current and future applications of data and the widely-held desire that those benefits be available to as many consumers as possible, regulators should encourage the adoption of frameworks that both promote innovative uses of big data and avoid identifiable harms.

October 31, 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
bmadhani@ccianet.org