Information Technology Industry Council

Submission to the U.S. Federal Trade Commission
Big Data: A Tool for Inclusion or Exclusion?
Workshop, Project No. P145406

Comments of the Information Technology Industry Council

October 27, 2014

## I. Introduction

The Information Technology Industry Council (ITI) appreciates this opportunity to submit comments to the U.S. Federal Trade Commission (FTC or Commission) in connection with the agency's September 15, 2014 workshop on *Big Data: A Tool for Inclusion or Exclusion* (the "FTC Workshop").

ITI, a U.S.-based global trade association representing 60 of the world's most dynamic and technologically innovative companies, works to advance effective policies that promote privacy and that also enable the technology sector to continue to innovate and develop new products and services.

As discussed below, the FTC Workshop identified areas where the FTC's expertise can move the discussion forward in connection with big data, particularly with regard to enforcing existing laws. In addition, the FTC Workshop highlighted the importance of the recommendations ITI has previously put forward to the administration in connection with big data, in particular those outlined in the ITI submissions to the White House and to the National Telecommunications and Information Administration (NTIA).[1] In these submissions, ITI offered a number of recommendations focusing on three main areas: (a) a responsible use and risk-based approach; (b) accountability mechanisms; and (c) data security and breach notification. We discuss these recommendations below.

[1] These submissions are available at http://www.itic.org/dotAsset/bcae1b74-eb8e-4f01-a02d-7e8aa8bdaf0f.pdf and http://www.itic.org/dotAsset/d/7/d751b77c-6ed0-49e9-8271-585e2dec63fc.pdf, accessed October 27, 2014.

II. Big Data: Moving the Discussion Forward

The capabilities of big data and the significant societal benefits big data analysis can yield (often in real time) are universally recognized. Indeed, many of the FTC Workshop participants discussed the opportunities created by big data. Chairwoman Ramirez remarked, as she opened the FTC Workshop, that big data "has the capacity to save lives, improve education, enhance government services, increase marketplace efficiency, and boost economic productivity."[2]

Big data offers tremendous opportunities in many areas, among them health care (both medical research and delivery of health care), agriculture, energy efficiency, transportation, and education. FTC Workshop participants emphasized that big data supports data analysis that could not previously be done. In areas such as electricity efficiency, vehicle maintenance, and combating government insurance fraud, big data enables powerful capabilities. We further note that big data enables marketers to provide consumers with more tailored offers, and the enormous benefits associated with privacy-protective personalized web experiences and targeted advertising—a practice that subsidizes many free Internet services and contributes significantly to the U.S. economy. Of course, as noted by FTC Workshop speakers, big data can also help empower the underserved, as discussed below.

While recognizing the benefits of big data, some observers have also raised concerns about its use. Specifically, panelists at the FTC Workshop discussed how big data analysis could lead to discriminatory outcomes as more decisions are determined by algorithms and automated processes. ITI agrees that discriminatory outcomes relating to the protection of civil rights should be avoided and that this is an important area to be considered, separate and apart from discussion of any privacy harms. ITI believes that FTC resources devoted to identifying *actual* discrimination harms caused by particular uses of big data would be a worthwhile use of the Commission's resources. We point out that the White House report on big data—prepared by a U.S. government inter-agency working group led by John Podesta and delivered to President

[2] FTC Workshop "*Big Data: A Tool for Inclusion or Exclusion*," September 15, 2014, transcript at 5, accessed October 27, 2014 http://www.ftc.gov/system/files/documents/public_events/313371/bigdata-transcript-9_15_14.pdf.

Obama in May 2014—recommended that the U.S. government's lead civil rights and consumer protection agencies expand their technical expertise to identify such discriminatory practices and outcomes and to develop a plan to investigate and address violations of law.[3] The FTC is well-suited to participate in this examination and to use its authority to enforce existing laws addressing discriminatory practices.

There was robust discussion at the FTC Workshop about existing U.S. laws that address discriminatory practices, including the Fair Credit Reporting Act, the Equal Credit Opportunity Act, Americans with Disabilities Act, the Age Discrimination and Employment Act, the Genetic Information Non-Discrimination Act, and employment laws under Title VII of the Civil Rights Act of 1964. The FTC Act, which prohibits unfair or deceptive practices, was also cited as an enforcement tool that could be utilized. We further note that self-regulatory codes of conduct currently in use prohibit certain discriminatory uses of data, and we urge that self-regulatory mechanisms be encouraged.[4] Companies that represent they abide by such a code of conduct but then violate its requirements could become subject to an action by the FTC for engaging in a deceptive practice in violation of Section 5 of the FTC Act.

The FTC Workshop discussion surrounding the current legal framework that prohibits discrimination highlighted the need for a comprehensive examination of the current foundation of privacy and anti-discrimination protections currently afforded to consumers, including privacy laws related to health, financial information, children, and credit, and anti-discrimination laws in the areas of employment, education, housing, and credit worthiness. Workshop participants pointed out that these issues are not brand new, and that any discussion of what ought to be done going forward must include a full analysis of the current legal landscape. Such an examination by the FTC could determine whether any legal or regulatory gaps actually exist and, if so, their

[3] Big Data: Seizing Opportunities, Preserving Values," *Executive Office of the President*, May 2014, accessed October 27, 2014, http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf at 60.

[4] For example, codes of conduct developed by the Digital Advertising Alliance (DAA) include restrictions on the use of data for eligibility purposes in connection with employment, credit, health care, and insurance. See: "Self-Regulatory Principles for Multi-Site Data," *Digital Advertising Alliance*, November 2011, accessed October 27, 2014, http://www.aboutads.info/msdprinciples and "Self-Regulatory Principles for Multi-Site Data," *Digital Advertising Alliance*, July 2009, accessed October 27, 2014, http://www.aboutads.info/msdprinciples.

scope. If gaps in the law are identified, it should be carefully considered whether they are unique to big data, or whether they exist regardless of the technology used. Any policy proposals that would address privacy and discrimination harms should be "technology neutral."

We also note that there was discussion at the FTC Workshop about how big data can help solve social problems impacting vulnerable populations, such as health disparities, education disparities, and disproportionate consumer impacts. It was also noted at the FTC Workshop that big data can be utilized to identify discrimination—and be further used to avoid it. For example, reference was made to a series of case studies presented by the Future of Privacy Forum and the Anti-Defamation League.[5] These case studies illustrate how big data can actually be used to protect and empower vulnerable groups. For example, one case study showed how data analytics is utilized to further workforce diversity.

In addition, there was discussion at the FTC Workshop about how underserved communities could be at risk from being *under*represented in big data—and that policymakers should consider social and economic inequalities that could result from a dearth of data in certain communities. Reference was made to a recent paper—*The Rise of Data Poverty in America*—where the author referred to such inequalities as a "data divide."[6] The paper's author, who also appeared as a panelist at the FTC Workshop, points out that if data is not collected in certain communities, "data deserts" could emerge where there is a lack of access to high-quality data that can be used for social and economic benefits. The author points out that data has always had an important impact on communities and discusses how census data is used to apportion congressional seats—and that inaccuracies lead to clear negative impacts. Further, the author points out that failure to capture data about underserved communities could make certain big data benefits unavailable to those very communities that could benefit the most, particularly in areas such as education, financial services, and health care.

As the FTC continues its work in the big data area, we urge the agency to take into

[5] Future of Privacy Forum and Anti-Defamation League, "Big Data: A Tool for Fighting Discrimination and Empowering Groups" accessed October 27, 2014, http://www.futureofprivacy.org/wp-content/uploads/Big-Data-A-Tool-for-Fighting-Discrimination-and-Empowering-Groups-Report1.pdf.

[6] Daniel Castro, "The Rise of Data Poverty in America," The Center for Data Innovation, September 10, 2014, accessed October 27, 2014, http://www2.datainnovation.org/2014-data-poverty.pdf.

account the benefits of big data, how big data can also be used to *avoid* negative outcomes that have been theorized as being facilitated by big data, and how underserved communities may actually suffer from not being adequately represented in big data sets. We also note the importance of recognizing that the data ecosystem continues to evolve. Finally, we note, as did panelists at the FTC Workshop, as well as Commissioner Brill, that consumer trust is critical. Consumer trust is paramount to ensuring continued economic growth and innovative new products and services. ITI member companies recognize that customers entrust them with their information and count on ITI member companies not only to continue to provide innovative products and services but also to continue to be transparent and fair in connection with their information.

## III. Recommendations

ITI's specific recommendations with regard to big data focus on three main areas: (a) a responsible use and risk-based approach; (b) accountability mechanisms; and (c) data security and breach notification.

### A. *Responsible Use and Risk-Based Approach to Privacy*

#### (i) Responsible Use of Big Data

As discussed at the FTC Workshop, large-scale data analytics allows for insights that are unexpected or were previously unknowable. This capability relies on vast amounts of data. Accordingly, as discussed at the FTC Workshop, it is often more appropriate to focus greater attention on *how data is used* and less on its collection. Data, in itself, is not necessarily good or bad, but it can be used for a range of purposes. In the big data environment, shifting the focus to the responsible use of data rather than its collection enables beneficial uses of data, while at the same time requires that privacy and other harms be considered in connection with any potential data use. ITI supports such an approach.

#### (ii) Risk Assessment and Mitigation

A responsible use framework based on risk-based assessment and mitigation is well suited to the big data environment; it requires organizations to thoroughly consider the distinct privacy and discrimination issues involved in decisions about whether data should be used for a given purpose.

A responsible use framework requires an organization to implement robust procedures and mechanisms to determine, based on the risks involved, which uses of data should be pursued and which should not. Thus, a use-based framework involves examining the potential risks of a particular data use. A risk-based analysis would be based on a common set of factors, such as the type of data being analyzed and used, how the data was amassed, the public interest in the use of the data, the consumer benefits of the use, the security measures in place, whether the data is shared with third parties, and the potential harmful impact to individuals resulting from the use. This risk assessment would serve not only to determine whether a particular use or analysis is appropriate, but also to identify how privacy-protecting or anti-discrimination safeguards might be implemented to mitigate any risks. By assessing the privacy or discrimination risks at the outset (through what might be thought of as a privacy or discrimination risk assessment), data scientists could identify how to derive the maximum benefit from the data while minimizing the risks to individuals. The development of a framework for privacy or discrimination risk management can be an effective mechanism to address the challenges posed by big data.

(iii) The Role of De-identification

As noted above, the type of data being analyzed and used is one factor to be considered in determining the appropriateness of a data use. For example, whether data is de-identified will be a consideration in a company's overall risk-based assessment and mitigation strategy. If the data an organization plans to analyze is de-identified, any potential privacy risk is lessened. If the data is not presently de-identified, a company may choose to de-identify the data in order to mitigate potential privacy risks. While concerns have been raised about the potential to re-identify data—and the question of the value of de-identification was raised at the FTC Workshop—recent research indicates that the "real world" risk of re-identification may be far lower than expected.[7] We note that examples of re-identification often involve datasets that were publicly accessible, enabling robust efforts to defeat identification. When datasets are kept confidential, the risk of re-identifying the data is remote.

[7] Daniel Castro and Ann Cavoukian, "Big Data and Innovation, Setting the Record Straight: De-Identification Does Work," *Office of the Information and Privacy Commissioner, Ontario, Canada and Information Technology and Innovation Foundation*, June 2014, accessed October 27, 2014, http://www2.itif.org/2014-big-data-deidentification.pdf.

We further note that new techniques continue to improve our ability to de-identify data—and technological research into more effective de-identification methods should be encouraged. In addition, policies that encourage de-identification should be pursued. For example, in its 2012 privacy report, the FTC stated that the agency's privacy framework applies to data that is reasonably linkable to a specific consumer, computer, or device.[8] Thus, data that is not "reasonably linkable" would not be subject to the requirements of the framework. The FTC Privacy Report outlines the steps it expects companies to take to render information not reasonably linkable. Organizations, as part of their risk mitigation techniques, can develop processes to de-identify data where appropriate, and processes to prevent re-identification. We emphasize that de-identification may serve as one tool among many that an organization deploys to mitigate privacy risks raised by the use of big data.

B. *Accountability Mechanisms*

In the big data context, and particularly in developing a responsible use and risk-based approach to privacy and discrimination, companies should hold themselves accountable for decisions they make about the processing, management, and protection of data. Several FTC Workshop participants cited accountability as a critical element of building consumer trust. Also, in the FTC Privacy Report, the Commission, recognizing the importance of accountability, pointed out that "companies should implement accountability mechanisms and conduct regular privacy risk assessments to ensure that privacy issues are addressed throughout an organization."[9] Accountability is also one of the principles of the Consumer Privacy Bill of Rights, embodied in the 2012 White House Report on privacy,[10] and a principle of fair information practices articulated in several international privacy instruments, including the Organisation for Economic Cooperation and Development's Privacy Guidelines (1980), and the

[8] "Protecting Consumer Privacy in an Era of Rapid Change, Recommendations for Businesses and Policymakers," (the "FTC Privacy Report") *Federal Trade Commission*, March 2012, accessed October 27, 2014, http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

[9] FTC Privacy Report at 30.

[10] "Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy," *The White House*, February 2012, accessed October 27, 2014, http://www.whitehouse.gov/sites/default/files/privacy-final.pdf.

Asia Pacific Economic Cooperation's Privacy Framework (2005). Both of these instruments state that entities "should be accountable for complying with measures that give effect" to the principles in the applicable instrument.

Accountability requires that organizations implement programs and processes that foster compliance with their commitments. Further, accountability requires that organizations be able to describe how the programs and processes do in fact meet such commitments. Thus, an important component of accountability is how an organization evaluates its processes and programs. The tools that a company uses to evaluate its processes will depend on various factors, including the size, complexity, and nature of an organization's business. Assessments can include internal or external audits, as well as other systems for ongoing oversight, assurance reviews, and verification.

Robust accountability is particularly important in the context of big data and a responsible use framework and risk-based approach to privacy where there may be a lesser reliance on certain fair information practice principles (FIPPs). FIPPs such as "purpose specification" and "use limitation" may limit the possibility of analyzing data in a manner that may not have been contemplated at the time it was collected. And utilizing data in new ways is integral to big data and the benefits it can provide. A responsible use framework and risk-based approach to privacy and discrimination is a practical alternative to a traditional FIPPs framework approach—a responsible use framework and risk-based approach requires organizations to thoroughly consider the distinct privacy and discrimination issues involved in decisions about whether data should be used for a given purpose.

C. *Data security and breach notification*

As recognized at the FTC Workshop, the importance of data security is paramount in the big data context where large amounts of data are collected and processed, often in real time. In addition, certain decisions are made—based on data—about the functioning of devices that impact all facets of our lives. These "devices" include vehicles, alarm systems, medical devices, and countless other ICT-enabled products. We note that the measures that an organization employs to secure data will depend on a number of factors, including the nature of the data and its sensitivity, as well as the size and complexity of the organization.

In addition, the large amount of data that can be amassed in the big data context

highlights the need for federal data breach notification legislation to replace the current patchwork of state breach notification laws. ITI supports a federal standard that would require data breach notification when the unauthorized acquisition of sensitive personal data could result in a significant risk of financial harm or identify theft. Such legislation should preempt existing states laws, be technology neutral, and set forth reasonable time periods for breach notification.

* * *

ITI appreciates the opportunity to submit these comments to the FTC. If you have any questions about these comments, please contact Yael Weinman, VP, Global Privacy Policy and General Counsel, Information Technology Industry Council, at 202-626-5751, yweinman@itic.org.