



October 31, 2014

Secretary  
Federal Trade Commission  
Room H-113 (Annex X)  
600 Pennsylvania Avenue, NW  
Washington, DC 20580

**Re: “Big Data: A Tool for Inclusion or Exclusion? Workshop, Project No. P145406,”**

I hereby submit the attached comments on behalf of the Marketing Research Association (MRA) regarding the FTC’s Big Data workshop on September 15, 2014.

Sincerely,

Howard Fienberg  
Director of Government Affairs  
Marketing Research Association (MRA)



## Introduction

On September 15, 2014, the FTC workshop “Big Data: A Tool for Inclusion or Exclusion?” focused on the potential discriminatory effects of segmentation and analysis in large data sets. Participants covered a lot of benefits from Big Data, as well as plenty of potential scenarios with nebulous negative effects.

As FTC Chairman Edith Ramirez said in opening the workshop, “advances in computational and statistical methods mean that this mass of information can be examined to identify correlations, make predictions, draw inferences, and glean new insight.”

She further observed that, “the proliferation of connected devices; the plummeting cost of collecting, storing, and processing information; and the ability of data brokers and others to combine offline and online data mean that companies can accumulate virtually unlimited amounts of consumer information and store it indefinitely.” In the process, Big Data is growing in its capacity to “reinforce disadvantages faced by low-income and underserved communities. As businesses segment consumers to determine what products are marketed to them, the prices they are charged, and the level of customer service they receive, the worry is that existing disparities will be exacerbated.”

Big Data frequently serves as sort of a cypher. The term is widely-referenced but poorly understood, and people in the public policy community see in it a variety of hopes and fears.

The Marketing Research Association (MRA) aims to: (1) explain how Big Data fits into the survey, opinion and marketing research space that we represent; (2) address the legitimate use of segmentation; (3) explore the challenges facing the notice and choice model and the need for a use-based model; (4) consider the appropriateness of consumer data access and

correction; (5) look at the promise of data deidentification; and (6) specifically identify and counter Big Data harms.

MRA wants to ensure that Big Data does not become a pejorative term, such as the FTC and privacy activists have made of the term “data broker.”

### **Big Data analysis is part of survey, opinion and marketing research**

MRA is a non-profit national membership association representing the survey, opinion and marketing research profession.<sup>1</sup> MRA promotes, advocates for, and protects the integrity of the research profession, and works to improve research participation and quality.

Survey, opinion and marketing research is the scientific process of gathering, measuring and analyzing opinion and behavior. Researchers traditionally use statistically-balanced samples<sup>2</sup> to determine the public’s opinion and behavior regarding products, services, issues, candidates and other topics. Such information is used to develop new products, improve services, and inform policy. Analysis of massive data sets plays a growing role in the research business, both on its own and in conjunction with more traditional research methodologies.

Survey, opinion and marketing research is sharply distinguished from commercial activities, like marketing, advertising and sales. Research focuses on understanding groups (“segments”), not individuals. No sales, promotional or marketing efforts are involved in bona fide research.<sup>3</sup> In fact, MRA and other research associations prohibit and attempt to combat sales or fundraising

---

<sup>1</sup> The research profession is a multi-billion dollar worldwide industry, comprised of pollsters and government, public opinion, academic and goods and services researchers, whose members range from large multinational corporations and small businesses to academic institutes, non-profit organizations and government agencies.

<sup>2</sup> A “sample” is a subset of a population from which data is collected to be used in estimating parameters of the total population.

<sup>3</sup> MRA, in consultation with the broader research profession, has developed a legal definition of bona fide survey, opinion and marketing research, which implicitly includes Big Data analytical research: “the collection and analysis of data regarding opinions, needs, awareness, knowledge, views, experiences and behaviors of a population, through the development and administration of surveys, interviews, focus groups, polls, observation, or other research methodologies, in which no sales, promotional or marketing efforts are involved and through which there is no attempt to influence a participant’s attitudes or behavior.”

under the guise of research (referred to as “sugging” and “frugging”), “push polls,”<sup>4</sup> and any attempts to influence or alter the attitudes or behavior of research participants as a part of the research process.<sup>5</sup> Quite to the contrary, professional research has as its mission the true and accurate assessment of opinion and sentiment in order to help individuals, companies, organizations and governments design products, services and policies that meet the needs of, and appeal to, the public.

Big Data presents an important opportunity for research, but also for respondents. While response rates have declined for research studies in the U.S. in the last couple of decades, and busy consumers are less likely to want to actively participate, the analytical possibilities of Big Data present the chance to do in-depth long-term research of consumer attitude and behavior without ever actually directly disrupting consumers’ valuable time and attention.

### **Is Segmentation Good or Bad**

The FTC’s continued focus on the practice of segmentation indicates that a key aspect of research is under misguided scrutiny.

According to Ramirez, discrimination “is what Big Data does in the commercial sphere – analyzes vast amounts of information to differentiate among us at lightning speed through a complex and opaque process.” That is a generally accurate description of segmentation. However, the Chairman went farther, asking if segmentation is “unfair, biased, or even illegal discrimination” in terms of offering different rates and prices or eligibility for products and

---

<sup>4</sup> “Push polls’ - Deceptive Advocacy/Persuasion Under the Guise of Legitimate Polling.” <http://www.marketingresearch.org/push-polls-deceptive-advocacy/persuasion-under-the-guise-of-legitimate-polling>

<sup>5</sup> For instance, in the *MRA Code of Marketing Research Standards*: “7. Ensure that respondent information collected during any study will not be used for sales, solicitations, push polling or any other non-research purpose. Commingling research with sales or advocacy undermines the integrity of the research process and deters respondent cooperation. In addition, the possibility of harm from data sharing – such as health insurance companies adjusting an individual’s costs based on information disclosed about their health behaviors or financial companies denying someone credit based on their propensity for online shopping – are the focus of growing public debate about Big Data and data brokers. Respondents should be assured that information shared in a study will only be used for research.” <http://www.marketingresearch.org/code>

services based on racial, ethnic, income or other status, and if there are possible ways to “level the playing field.”

Ramirez explained at the workshop that some data brokers “create segments or clusters of consumers with high concentrations of minorities or low income individuals.” While Ramirez admits that “there may be legitimate reasons why businesses would want to sort consumers in this fashion,” such segmentation could be used for what she calls “discrimination by algorithm,” and what the White House Big Data report has referred to as “digital redlining.”

The “legitimate reasons” for segmentation include dividing consumers into groups in order to better understand them. The application of that understanding to uses of which the FTC may not approve is a completely different matter – if so, the FTC would do well to articulate the policy changes necessary to combat such specific uses, instead of denigrating through generalizations a key part of research.

Much of this issue relates back to the FTC report on data brokers,<sup>6</sup> which was referenced throughout the workshop. The FTC specifically identified “marketing analytics” activities in their report as falling within their definition of data brokers. Many “marketing analytics” products, such as the measurement of online and mobile ad effectiveness, are services that many researchers, their suppliers, and other business partners, now routinely offer.

According to the FTC data brokers report, “In addition to using raw data, data brokers often aggregate and analyze it to make inferences about specific consumers. For example, they may categorize a consumer as an expectant parent, a car enthusiast, interested in diabetes, a discount shopper, and more likely to be interested in brand medications than generic.” The FTC said that the data brokers the agency investigated develop segments by: (1) combining data elements to create a list of consumers who have similar characteristics; and/or (2) developing complex models to predict behaviors.

---

<sup>6</sup> <http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>

The line separating marketing analytics from marketing research can at times be gray. The FTC, in the data brokers report, crossed it.

- The agency observed that five of the data brokers they investigated provided “analytics for marketing purposes, as a way to predict consumers’ likely behavior. Among other things, the analytics products offered by some of the data brokers enable a client to more accurately target consumers for an advertising campaign, refine product and campaign messages, and gain insights and information about consumer attitudes and preferences. For example, some data brokers will analyze their client’s customer data and advise a client regarding the type of media channel to use to advertise a particular product or brand (e.g., online, newspapers, television) and where the advertisements should be shown (e.g., Florida or California). As part of this analysis, the data brokers can also help their clients model the expected outcomes of various marketing tactics, thus allowing the clients to better advertise their products to consumers. For example, a data broker might be able to predict whether advertising a new product exclusively through Twitter will yield the desired outcome.”
- Some data brokers also can “evaluate the impact of an advertising campaign after it has run. These analytic products are generally based on algorithms that consider hundreds or thousands of data elements, including historical data provided by the client and data that the broker gathers” from various sources.

Both descriptions of Big Data analytics are also textbook marketing research cases. Similarly, the FTC has painted data brokers with so broad a brush as to include most marketing research companies under such a designation.

The FTC used the workshop to reiterate the agency’s call for legislation targeting data brokers. Based on the scope and definitions in the FTC’s report, and legislation introduced so far, like the “Data Broker Accountability and Transparency Act” (DATA Act) (S. 2025) from Senator Jay Rockefeller (D-WV), research companies would likely be required to provide access and correction rights to research respondents, potentially upending the research ecosystem.

## **Notice, Choice, Use**

Notice and choice have been the pillars of our approach to consumer privacy for a long time. However, while some research indicates that consumers, on average, are concerned about their privacy, notifying consumers of every minute thing happening with their data could work counter to the interest of actually keeping consumers well informed, since over-notification and excessively lengthy privacy policies already may be causing consumers to stop paying close attention to their own privacy needs and wants and growing more careless in how they handle their own data.

Traditional research is focused heavily on consumer notice and consent, but social media and Big Data make that traditional model significantly harder to apply. While an approach of “just-in-time” notification holds some promise in the Big Data context, getting consent for Big Data analysis may not make sense in many cases.

Most importantly, policymakers should instead consider a use-based approach to consumer data privacy that draws sharper lines between the research uses of Big Data and its implementation for other purposes, whether for marketing, identification/verification, or the determination of eligibility. The traditional model of prior notice and consent is often impractical or impossible in a Big Data context, because Big Data allows for the discovery of insights unimagined or unknown at the original time and place of data collection.

Researchers and others in the private sector should strive to achieve workable opt out mechanisms and greater transparency for consumers. Companies like Axiom are already delivering on that.<sup>7</sup> However, MRA cautions strongly against policymakers requiring specific mechanisms and approaches to transparency. A one-size-fits-all approach will not work well for such policies when applied to research, and will stifle the private sector’s ability to innovate in delivering greater consumer transparency and control.

In the meantime, we need to look more closely at shifting regulation towards the uses of data – different protections and requirements for data privacy, depending on the uses of that data.

---

<sup>7</sup> <https://aboutthedata.com/>

Data collected, used and shared strictly for bona fide research<sup>8</sup> should be held to a different standard than ordinary commercial use, which will differ from purely transactional use. Those uses all should be treated differently than data used for determining a consumer's eligibility for things like health insurance, credit or a mortgage, or data used to prosecute crimes or prevent terrorism.

### **Consumer Access and Control**

The question of whether or not consumers should be given access and control over data may be even more complicated. The demand of access to consumer data, a key part of the FTC's data broker report, may make sense in contexts of eligibility, where such data (particularly if inaccurate) could adversely impact a consumer's credit rating, personal or professional reputation, or in the likelihood of becoming a victim of identity theft or fraud. However, none of these conditions should reasonably be assumed to apply to research data.

The cost of access and correction could potentially be quite onerous, especially for smaller research companies and organizations, given a potential deluge of frivolous or pointless inquiries. Since the research process is interested in broad groups, not individuals, compiling and tracking individual consumer data, by the individual, would require complex and expensive procedures and infrastructure not currently in use. Moreover, such tracking could lead to a much greater threat of harm from data leakage and empower the kind of consumer tracking that causes some privacy concerns.

The ability of companies to authenticate the identity of consumers requesting access is another potential problem. That kind of authentication would require collecting and checking even more data, which runs counter to the FTC's interest in data minimization and limited data retention. Plus, necessary authentication procedures and processes would add to the cost in money and time on the part of research organizations and possibly that of research participants.

---

<sup>8</sup> As noted earlier, research purposes, unlike most of these other purposes, involve data about individuals only for the purpose of understanding broader population segments and demographic groups.

MRA supports the concept of a “sliding scale” for access and correction responsibilities in order to reconcile the vague consumer benefits with the expected costs. We propose that the availability and extent of access should depend on the data actually being susceptible to use for criminal or fraudulent purposes. As MRA has repeatedly stressed, purpose and use matter.

### **Data Deidentification**

Multiple workshop speakers mentioned the potential promise of deidentification to allow greater consumer protection alongside increased use of Big Data.

Anonymizing or de-identifying personal information whenever possible, by aggregating or pseudonymizing it, is a sensible principle, and one that MRA and the research profession support. In fact, carving out a safe harbor from many privacy regulations for data that has been protected in this way makes a lot of sense, but we are concerned with the specific definition of how and to what extent to do it. There is an ongoing debate in the academic and policy arenas on whether or not data can ever be fully de-identified or anonymized. If it cannot, then any piece of data could ultimately be personally identifiable.

As MRA pointed out in our reply comments to the FTC regarding the Compete settlement in 2012,<sup>9</sup> we would be wary of the FTC stepping into such an intense ongoing debate. The current FTC approach as outlined in the FTC’s 2012 privacy report sensibly approaches deidentification from a more strategic level. The agency considers data to not be “reasonably linkable” if “a company (1) takes reasonable measures to ensure that the data is de-identified; (2) publicly commits not to try to re-identify the data; and (3) contractually prohibits downstream recipients from trying to re-identify the data.”<sup>10</sup> The FTC then has the power and authority to support good data protection and punish violations without having to artificially end the deidentification debate.

---

<sup>9</sup> (#452: In the Matter of Compete, Inc.; FTC File No. 102 3155) MRA comments in reply to EPIC’s comments: [http://www.marketingresearch.org/sites/mra.cms.memberfuse.com/mra/files/pdfs/MRA-response\\_to\\_EPIC\\_comments-Compete\\_case-12-21-12.pdf](http://www.marketingresearch.org/sites/mra.cms.memberfuse.com/mra/files/pdfs/MRA-response_to_EPIC_comments-Compete_case-12-21-12.pdf)

<sup>10</sup> Page iv: <http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>

## Identifying and countering harm

Identifying (and especially, quantifying) the harm arising in data privacy often proves elusive for both privacy activists and regulatory authorities. We need to look more closely at the harms from Big Data and weigh them against the benefits. The harms identified to date have often been ethereal. Presenters at the FTC workshop pointed to alternative credit scoring or insurance-eligibility mechanisms, but none of them could show evidence of how the discriminatory use of such alternative scores or mechanisms would fall outside of existing discrimination, housing, financial, insurance or credit report laws, regulations and enforcement.

Ultimately, we have not seen much to distinguish the supposed negative discriminatory effects of Big Data to such effects from long-standing commercial practices, like club memberships or frequent flyer designations. The reward and reinforcement of customer loyalty is nothing new, but Big Data may allow it to be applied more efficiently. As a *Forbes* article described such price discrimination, “In a traditional bazaar a seller might charge a well-dressed buyer twice as much as another based on visual clues or accents. Big data allows for a far more scientific approach to selling at different prices, depending on an individual’s willingness to pay.”<sup>11</sup>

## Conclusion

Making corporate or public policy based on hypothetical scenarios involving no measureable consumer harm is a sure-fire way to strangle data-driven innovation in the research world and the broader marketplace. Prohibiting or rigidly restricting the use of Big Data would be detrimental to the continued progress of data-enabled convenience and community.

Rather than simply assuming negative discriminatory impact from data analysis and segmentation, MRA recommends a practical step: a rigorous long-term economic study of the issue by the FTC. If the net value of Big Data for low-income or vulnerable segments of the American population isn’t shown to be there, and if Big Data’s discriminatory effects are inordinately burdensome, then it might be worth considering policymaking options (whether by

---

<sup>11</sup> “Different Customers, Different Prices, Thanks To Big Data.” by Adam Tammer. *Forbes*. April 14, 2014. <http://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/>

government entities like the FTC or self-regulation) to counter such effects. However, as any researcher would say, we need better data before we can make responsible decisions.

In the meantime, although we differ on some of the details, MRA broadly supports the FTC's call for a baseline federal privacy law<sup>12</sup> and a national data security law.<sup>13</sup>

MRA wishes to work with the FTC to identify and counter real harms against consumers. We believe that continued robust enforcement by the FTC of existing privacy statutes, combined with the in-depth economic study we've proposed, is the best course of action for the agency.

---

<sup>12</sup> MRA's vision of a baseline privacy law is closer to Congressman Cliff Stearns' Consumer Privacy Protection Act (H.R. 1528 in 2011) than to Congressman Bobby Rush's Best Practices Act (H.R. 611 in 2011) or Senator John Kerry's Consumer Privacy Bill of Rights Act (S. 799 in 2011).

<sup>13</sup> While MRA sees many things to like in Congressional data security bills, we're opposed to some potential provisions like private rights of action (found in Senator Blumenthal's Personal Data Protection and Breach Accountability Act (S. 1995)) or unfettered FTC authority to define personal information so broadly as to include all sorts of mundane research data (as would have been the case with Congresswoman Mary Bono-Mack's 2011 SAFE Data Act (H.R. 2577), before its amendment in subcommittee - <http://www.marketingresearch.org/news/2011/07/21/mras-amendments-approved-by-house-subcommittee-before-passing-the-safe-data-act> ).