The FTC's Big Data Message: Privacy is Fundamental

Jessica Rich, Federal Trade Commission
Email Sender and Provider Coalition (ESPC) Annual Meeting
September 10, 2014

Good morning. I am delighted to be here to talk about the FTC's work on Big Data. My aim today is to provide a brief overview of the FTC's concerns and activities in this area. Then, I'd really like to take your questions.

Big Data is a catch phrase we hear constantly. While definitions of Big Data vary, the term is often used, as I use it here, to refer to the confluence of factors that contribute to the ever-greater combining, analyzing, and use of data in making decisions about individuals and groups of individuals. These factors include the ubiquitous collection of consumer data using the Internet, social media, mobile devices, and sensors; the plummeting cost of storing such data; and the powerful new capabilities to use the data to make inferences and predictions.

Some say Big Data is the answer to save us from disease, from hunger, from environmental collapse, you name it. Others think Big Data will cause the downfall of a free and civilized society, as we are tracked, watched, and analyzed in every aspect of our lives. In reality, various forms of Big Data have been around for some time. For years, information about consumers has been compiled, organized, and used to provide products and services to consumers. What's new is the sheer volume of data produced and collected, and the seemingly infinite uses of this data in ways that provide great benefits to consumers but also have the potential to harm them.

The challenge for us at the FTC – and for all of those who create and use Big Data – is to find ways to ensure that our core consumer protection principles, which have been around for decades, are followed in this era of Big Data.

1. Benefits and Risks of Big Data

Before I describe the FTC's efforts in this area, I'd like to briefly tease out what I view as the benefits and risks of Big Data.

Big Data can of course drive valuable innovation. Smartphones can deliver the latest personalized news stories and connect consumers with friends on social media; home automation systems know when you are leaving work and when to turn on your front-porch light; and ankle monitors track how far you have biked or run so you can reach your fitness goals.

Big Data also can be used to enhance traditional products and services. For example, data brokers can use Big Data analytics to provide granular marketing lists to brick-and-mortar retailers, which provide consumers with more relevant coupons for the goods and services they want, just when they're ready to buy them. Big Data also provides benefits beyond the marketing sphere, such as improving algorithms for determining whether a transaction is fraudulent or assessing what medical treatments are most effective across a large population. .

However, Big Data also raises obvious risks for consumers. First, the enormous volume of data being collected about consumers creates security risks. A main concern, of course, is the storage of sensitive information like Social Security numbers, financial account numbers, and usernames and passwords. If compromised – and we are seeing

data compromises at historic levels *every day* – this data can be used to perpetrate identity theft.

Even without such direct compromises of sensitive information, identity thieves may be able to exploit the sheer volume and richness of consumer profiles available in today's Big Data economy. These profiles can give bad actors a clear picture of consumers' habits over time, thereby enabling them to predict passwords, challenge questions, or other authentication credentials in order to perpetrate identity theft or basic fraud. The compromise of Big Data can also reveal sensitive information about people's homes, their children, and their whereabouts – data consumers do not want revealed about themselves.

Second, data collection is ubiquitous – it's there at every turn, as you surf the web, walk down the street, update your Facebook page, use mapping apps, monitor your exercise and health, and browse the shoe department at your favorite sporting goods store. It's simply impossible to avoid, even if you are a determined consumer that's willing to read privacy policies and opt out wherever you can.

Third, as ubiquitous as Big Data is, it's often quite invisible. Today, many companies that collect consumer data – advertisers, data brokers, all sorts of middlemen – operate behind the scenes, invisible to consumers. This means, again, that it's impossible for consumers to learn about, much less control, the data collection taking place all day long.

Finally, in the absence of clear norms and expectations surrounding the use of Big Data, there are risks that Big Data may be used in unfair, discriminatory, and even illegal

ways. For example, maybe you use a risk management product to screen for fraud and it results in a consumer being denied the opportunity to write checks. Perhaps you reward certain customers with better customer service or shorter wait times than other customers. Maybe you offer different prices or discounts to different consumers, or give out different offers based on income levels. And maybe you only advertise to certain populations or groups of consumers with certain characteristics.

Some of this sounds like business as usual, the way you've always done it. But with the increasing collection of data about consumers from multiple sources, the increasing reliance of businesses on such data to make predictions about consumers, the decreasing ability of consumers to control what's happening, and the limited reach of current laws, the consequences are much greater today. The data could be wrong. And consumers can be forever stuck with labels and in "buckets" that deny them opportunities.

2. The FTC's work in the Big Data area

Some of the concerns raised by Big Data are easier for the FTC to address than others. As you know, we enforce a variety of privacy and consumer protection laws but there is no one law that establishes clear norms across different commercial contexts as to how data can be collected, used, and shared. We're using every tool that we have to protect consumers – enforcement, reports, education, the bully pulpit, etc. Our goal is to protect consumers from harmful practices without undermining the many benefits of Big Data.

A. Law Enforcement

On the law enforcement side, we enforce a variety of laws that have relevance in this area but I'll focus on two in particular. First is the Fair Credit Reporting Act, which sets forth procedures governing some of the most important uses of Big Data – determining whether to give consumers credit, a job, or insurance.

I won't spend too much time discussing our FCRA work because I've been asked to focus on our approach to Big Data for non-FCRA covered entities. But I do want to underscore that this is a huge area of enforcement for us. We've brought over 100 FCRA cases involving over $30 million in civil penalties, including a number of recent cases against data brokers that were covered by law but not complying with its most basic requirements. Among other things, these entities failed to provide reasonable accuracy for the data they provided to employers, failed to follow proper data dispute procedures, and failed to assure that purchasers of the data would use it only for the "permissible purposes" specified by the FCRA.

I also want to emphasize that the FCRA's reach is broader than some of you may realize. You don't have to be TransUnion, Equifax, or Experian to be a consumer reporting agency. A consumer reporting agency is a company that compiles consumer information and provides it to third parties to make employment, credit, housing, insurance, and other similar decisions. If you fall in that description, you may well be a consumer reporting agency, subject to the FCRA's accuracy, dispute, notice, and privacy requirements. Our recent cases made that clear to TeleCheck, Certegy, HireRight,

Spokeo, Filiquarian, Infotrack and Instant Checkmate, some of which paid millions in civil penalties.

The second important Big Data law – and the one that I was asked to focus on – is Section 5 of the FTC Act, which prohibits unfair or deceptive practices. The FTC has used Section 5 for decades to challenge practices involving pyramid schemes, business opportunity scams, cramming, deceptive weight loss and health claims, and many other areas of consumer protection. The basic rules are that companies cannot make deceptive claims about things that matter to consumers, or cause substantial injury to consumers in ways that consumers cannot avoid and in ways that do more harm than good. These same rules apply to privacy, data security, and Big Data.

Now, there isn't some line we draw where "regular" privacy and data security ends and Big Data begins. Many of our recent privacy and data security cases involve the issues and themes that characterize Big Data – the collection of detailed information about consumers, including by companies that are behind the scenes and unknown to consumers; the storage of consumer information in comprehensive databases; and the use of analytics to combine and transform the data and make predictions about consumers. And most of these cases obtained orders that contain relief directly relevant and critical to these issues – greater transparency and choices for consumers, a precious and diminishing commodity in today's Big Data economy; greater attention to privacy-by-design; and stronger security protections for Big Data.

Let me give some examples:

Our cases against data powerhouses Google and Facebook both involved the sharing of data with third parties contrary to claims made when the data was collected and contrary to consumers' expectations about how the data would be used.

In Google, we alleged that the company used data collected for one purpose (to register for, and use Gmail) for another purpose (to enroll consumers in its new Buzz social network), contrary to promises and choices provided to consumers. In Facebook, we alleged that the company claimed that consumers could designate their data as private but then repeatedly shared the data with third-party apps and advertisers. In both cases, there was considerable consumer outcry when it was revealed that the companies were using and collecting data in ways different from what had been promised and different from what users expected. This stuff matters to consumers – when they find out about it.

Another very telling case is our case against Goldenshores Technologies, the maker of a popular flashlight app that allows consumers to use their mobile devices as flashlights. In that case, we alleged that the app promised it would collect data from users' devices for certain internal housekeeping purposes, but failed to disclose that the app automatically transmitted the device's location and device ID to third parties, including mobile ad networks. A flashlight app! Again, there was huge uproar when consumers learned a flashlight app was collecting and sharing sensitive data totally unrelated to its functionality and contrary to its promises.

The Snapchat case is another interesting Section 5 case that implicates Big Data. The very idea behind Snapchat is to respond to consumer concerns about the storage and accessibility of personal data, including photos – a real consumer concern that we saw

very much on display two weeks ago with the huge uproar over the celebrity hacks on Apple devices. Snapchat claimed that photos and video messages sent through its service would *disappear forever* after a period designated by the user. Guess what? The photos didn't disappear, and the case was one of the most reported and tweeted case the FTC has brought yet. Consumers want limits and control over how much of their private data is shared with third parties, and they were outraged when such limits and controls were promised and not delivered.

We also brought a series of cases last year against Aaron's rent-to-own stores for installing invisible tracking software on rented computers that captured consumers' keystrokes, location, account information, and emails, and even took pictures of them in their homes. The cases were all about the invisibility theme I mentioned – unbeknownst to customers, Aaron's installed the software so it could keep track of the computers but didn't tell consumers and captured a whole lot of private information in the process. The company is now under orders for 20 years.

And finally, data security has become even more important in the era of Big Data, as companies collect and store more sensitive data than ever before. With massive breaches in the news, we are reminded of this on almost a daily basis. The FTC has brought over 50 cases against companies that we alleged failed to provide security for sensitive consumer data. The companies have ranged from large, sophisticated entities to mid-sized or smaller entities and include names like Eli Lilly, Microsoft, Guess, Petco, BJ's Warehouse, DSW, TJX, ChoicePoint, LexisNexis, CVS, Rite Aid, Wyndham Worldwide, Twitter, Dave & Busters, and the list goes on. And the information that was

compromised included not just credit and debit card data, but also social security numbers; bank account information; account passwords and personal emails; and consumers' medical conditions and diagnoses, prescriptions, and doctors visited.

One recent example that has particular relevance to Big Data is our first Internet of Things case – against home video monitoring company TRENDnet. In that case, we alleged that the company failed to provide reasonable security for IP cameras used for home security and baby monitoring, resulting in hackers being able to post private video feeds of people's bedrooms and children's rooms on the Internet. It's great that consumers can keep an eye on their homes from work or monitor their babies from a downstairs monitor, but not when criminals can watch too. The company is now under a 20-year order that requires the company to implement better security and provide consumers with free tools to fix the problem.

B. Policy Initiatives

In addition to law enforcement, the FTC has engaged in a number of policy initiatives to explore the challenges raised by Big Data and to encourage and promote stronger protections among industry members. These initiatives include workshops, studies, and reports; testimony before Congress; and consumer and business education. We view this policy work as a critical part of our consumer protection mission.

These initiatives do not necessarily track or reflect current law. Indeed, they often encourage companies to implement "best practices" over and above the law or recommend that Congress pass new or stronger laws to protect consumers and make the

rules clearer for everyone. For example, the Commission has repeatedly and unanimously urged Congress to pass a federal data security and breach notification law.

Many of our policy initiatives have focused – and continue to focus – on Big Data. For example, several years ago, the FTC hosted a series of public roundtables to examine the privacy challenges raised by proliferation of social networks, mobile and connected devices, data brokers, and other tools and building blocks of the Big Data era. Among other things, the workshops examined ways in which companies can amass "little data" from consumers and turn it into comprehensive databases and detailed consumer profiles.

Based on the roundtables, the Commission issued a Privacy Report in 2012 setting forth a general framework for addressing consumer privacy issues, designed to be applicable across multiple contexts. The report urged companies to adopt the following three core approaches to consumer privacy:

- *Privacy-by-Design*. Build in privacy protections at every stage of product development so they are part of the fundamental business model and not overlooked or added later as an afterthought. Such protections should include reasonable collection and storage limits, reasonable security measures, sound retention and disposal practices, and data accuracy.

- *Streamlined Choice*. Give consumers the ability to make decisions about their data at a relevant time and context – which in certain contexts will mean at the time their data is collected, not buried in a privacy policy. At the same time, there's no need to provide choices for data uses that are obvious or consistent with consumers' reasonable understanding and expectations, such as fulfillment or fraud prevention.

- *Greater Transparency*. Make sure that data collection and use practices are transparent to consumers, such as by giving consumers reasonable access to their information and standardizing and improving privacy policies.

More recently, last November, we held a public workshop to explore the phenomenon known as the Internet of Things. This growing form of data collection occurs via sensors placed in everyday devices that connect to the Internet. Participants discussed the privacy issues raised by the Internet of Things – including the increased ubiquity and invisibility of data collection, the challenges of providing notice and choice in this environment, and what incentives exist for designing products with privacy and security in mind. We plan to issue a report on the workshop in the coming months.

This spring, we also hosted a series of seminars to examine the privacy implications of three growing business models that enable companies to gather Big Data about consumers: (1) mobile device tracking of consumers in retail stores (2) predictive score modeling, through which companies can determine consumers' likely response to product and service offers, and (3) websites, apps, and devices that consumers use to manage their health data, most of which are not covered by HIPAA. We plan to issue reports on these seminars too.

Our activities on the Big Data front also include the May 2014 release of a report on data brokers. The report was the result of an in-depth study of nine companies representing a cross-section of the industry. The report discussed how these data brokers acquire and store billions of data elements on nearly every U.S. consumer and develop detailed profiles for sale to other companies for use in marketing, fraud prevention, and people search.

And the report discussed how data brokers don't just collect and share raw data, but also develop inferences about people and put them into categories – for example,

Urban Scramble and Mobile Mixers, which characterize low income, minority consumers; Thrifty Elders; Financially Challenged; Bible Lifestyle; Leans Left; and many other such categories. Like many of the Big Data practices I have been discussing today, virtually all of this happens behind the scenes, without consumers having any idea, let alone control over it.

To create more visibility and control for consumers – it's their data after all – the report encouraged Congress to enact legislation to create a centralized mechanism, such as an Internet portal, where data brokers would describe their data practices and provide links to information and choices about how it is used.

The report also recommended that the legislation require consumer-facing entities, such as retailers, to provide prominent notice to consumers when they share information with data brokers, along with choices about such sharing.

Now, all of initiatives I have been discussing this morning – not just the Data Broker Report – highlight Big Data practices and trends that would come as a big surprise to many consumers. I think we can all agree that consumers generally have no idea they are being tracked as they walk through stores or workout at home, or that they are labeled with various marketing scores or categories that refer to their incomes, health, religion, and education.

Sometimes, the data is used in ways that have clear benefits to consumers and businesses. For example, the data may be used to predict whether a certain transaction will result in fraud, whether a consumer will repay a loan, whether sending a catalog to a

certain address will result in an in-store or online purchase, and whether particular consumers will welcome or ignore a coupon or advertisement sent to them.

At the same time, there are potential harms associated with this type of behind-the-scenes labeling and categorization of consumers. Some consumers may find it troubling to receive advertisements based on their inclusion in various categories, especially when the ads relate to their health conditions. They may be troubled to know that these categories are used to determine what prices they may get, as opposed to their neighbors or someone across town. And they may find it even more troubling if a company used these labels and categories to decide whether to provide them credit, insurance, employment, or other benefits. For example, an insurance company could use the fact that a consumer is a biker, a smoker, or a person interested in diabetes information to infer that the consumer poses an insurance risk. And there's the related question of whether companies are using *unregulated* scores and labels in ways that mimic data uses subject to the protections of the FCRA or the Equal Credit Opportunity Act.

In just a few days, on September 15th, we will host a workshop to examine these and similar concerns raised by Big Data. Entitled "Big Data: A Tool for Inclusion or Exclusion?," the workshop will expand on what we have learned to date and explore how the categorization of consumers may be both creating and limiting opportunities for consumers, with a focus on low income and underserved consumers. Participants will discuss, among other things: the ways in which organizations use Big Data to classify consumers; the benefits and concerns these practices raise; how existing laws apply to such practices and if are there gaps in the legal framework; and whether companies are

appropriately assessing the impact of big data practices on low income and underserved populations. Please come or listen in – we're webcasting too.

3. Conclusion

This was just a brief overview of the FTC's concerns and initiatives regarding Big Data. As you can probably tell, my central message is – you don't need to leave consumer protections behind and you shouldn't. The basic principles we laid out in our privacy report – privacy-by-design, and providing transparency and choices to consumer – still apply. Thank you again for having me. I'm happy to answer questions.