Privacy and Data Security in the Age of Big Data and the Internet of Things
U.S. Federal Trade Commissioner Julie Brill
Delivered at
Washington Governor Jay Inslee's Cyber Security and Privacy Summit
January 5, 2016

Thank you, Alex Alben, for your warm introduction and for inviting me to share my thoughts with this most impressive gathering of lawmakers and leaders from companies and community groups. Protecting consumers' data from unauthorized disclosure and unexpected and inappropriate uses are some of our top priorities at the Federal Trade Commission (FTC), and the challenges of protecting consumers' privacy and security are becoming more pressing as we move further into a world of constantly connected devices and big data analytics. Today's Summit shows that states will continue to play a key role, alongside the FTC, in protecting consumers' privacy and security as our economy and society becomes more connected and data-driven.

We are connecting nearly everything – from cars and buildings to clothing and light bulbs – to the Internet. The pace and scale of these changes is breathtaking. Network equipment manufacturer Cisco reports that there are 25 billion networked devices in the world today and predicts that there will be 50 billion by 2020.[1] These sensors, along with our smartphones, tablets, and computers, generate twice as much data today as they did two years ago, and this trend is expected to continue. Sensors that are so small and efficient that they can power themselves with ambient radio waves are becoming a reality.[2] As a result, data is becoming cheaper to collect and keep, it is coming from an incredibly diverse range of sources – including the physical world around us – and our ability to analyze all of this data is constantly improving.

Let me be clear at the outset: I believe that big data and the Internet of Things have potentially tremendous benefits. Cities can better maintain their infrastructures by developing sophisticated early warning systems for gas and water leaks. Medical researchers can enroll patients in large-scale research projects and collect streams of useful data that, in the past, would have been a mere trickle coming from surveys and patients' own reports.[3] Connected and online courses are making it possible for people all over the world to learn and earn degrees from the

[1] DAVE EVANS, CISCO INTERNET BUS. SOLUTIONS GRP., THE INTERNET OF THINGS: HOW THE NEXT EVOLUTION OF THE INTERNET IS CHANGING EVERYTHING 3 (2011), *available at* http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf. These estimates include all types of connected devices, not just those aimed at the consumer market.

[2] BBC News, Tiny Chip That Powers Itself from Radio Waves (Dec. 8, 2015), *available at* http://www.bbc.com/news/technology-35038430.

[3] *See, e.g.*, Elizabeth Whitman, *Apple ResearchKit: Is New Open-Source Software for Sales or the Greater Good of Health Care*, INTL. BUS. TIMES (Mar. 16, 2015 3:51 PM), *available at* http://www.ibtimes.com/apple-researchkit-new-open-source-software-sales-or-greater-good-health-care-1848612.

world's leading experts and most prestigious institutions.[4] Data is helping governments to better plan and deliver their services.[5] And we are only at the very beginning of these developments.

Some significant risks go along with the potential benefits of connected devices and big data. As we add devices to our homes, classrooms, and clothes, much more sensitive data will be collected. User interfaces on devices will shrink or disappear, making it more difficult for consumers to know when data is being collected, or to exercise any control. In fact, I expect that the Internet itself will soon "disappear" because connectivity will just be part of how things work, as electricity is today.[6]

These developments pose difficult challenges for privacy, security, and fairness in our society. The data from connected devices will be deeply personal, and big data analytics will make the data more readily actionable. Some of these devices will handle deeply sensitive information about our health, our homes, and our families. Some will be linked to our financial accounts, some to our email accounts. And devices themselves will be more closely connected with our actions in the physical world, making data security and device security critically important.

But some fundamental aspects of our world will not change, no matter how connected and data-driven we become. Most importantly, we as individuals will remain roughly the same. We will not suddenly become capable of keeping track of dozens or hundreds of streams of our data, peering into the depths of algorithmic decision-making engines, or spotting security flaws in the countless devices and pieces of software that will surround us. Faced with a world of uncertainty about which devices are safe and whether consumers are getting a fair shake in the big data world, consumers could use some help.

To help consumers navigate and benefit from this complex, uncertain, and exciting world, the Internet of Things and big data analytics need to meet consumers' expectations and earn their trust. Appropriate privacy and security protections, as well as broader assurances that consumers are being treated fairly, are key elements of consumer trust. And these three elements – security, privacy and fairness in the world of big data and the Internet of Things – are what I would like to discuss today.

The FTC's Role in Protecting Consumers' Privacy and Data Security

Before I get to that discussion, let me first describe the role that the FTC plays in privacy, data security, and consumer protection in general. We are the nation's leading consumer protection agency, and we share competition enforcement with the Department of Justice.

---

[4] *See, e.g.*, Madeleine Parker and Sarah Rockwood, UC Berkeley Offers New Online Data Science Master's Degree, The Daily Californian (last updated June 24, 2014), *available at* http://www.dailycal.org/2014/06/24/uc-berkeley-offers-new-online-data-science-masters-degree/.

[5] *See* Ben Casselman, *Big Government is Getting in the Way of Big Data*, FIVETHIRTYEIGHT ECONOMICS (Mar. 9, 2015), *available at* http://fivethirtyeight.com/features/big-government-is-getting-in-the-way-of-big-data/.

[6] Chris Matyszczyck, *The Internet Will Vanish, Says Google's Eric Schmidt*, CNET (Jan. 22, 2015, 6:00 PM), *available at* http://www.cnet.com/news/the-internet-will-vanish-says-googles-schmidt/.

Eighty years ago, Congress gave the FTC authority to protect consumers from a broad range of "unfair or deceptive acts or practices."[7] Under this authority, the FTC has brought nearly 100 privacy and data security enforcement actions.

The flexibility and breadth of our authority to obtain remedies that protect consumers has allowed us to keep up with rapid changes in technology. For example, we have brought actions against companies for allegedly collecting information inappropriately from consumers' mobile devices,[8] making unwarranted intrusions into private spaces,[9] exposing health and other sensitive information, exposing previously confidential information about individuals' networks of friends and acquaintances,[10] and providing sensitive information to third parties who in turn victimize consumers.[11]

In addition to these privacy and data security enforcement actions, we have also brought hundreds of cases vindicating consumers' rights under more specific laws that protect sensitive information about children,[12] financial information,[13] medical data,[14] and information used to make decisions about consumers' credit, insurance, employment and housing.[15]

The FTC also maintains a busy policy docket. At the beginning of this year, we published a report on the Internet of Things, which emphasizes the importance of data and device security as well as the applicability of established privacy principles to connected devices.[16] Before that, we published a detailed study of the data broker industry,[17] which was in the big data business long before the words "big data" became part of our policy lexicon. We also held

---

[7] 15 U.S.C. § 45(a).

[8] *See, e.g.*, Goldenshores Techs. LLC C-4466 (F.T.C. Mar. 31, 2014) (decision and order), *available at* https://www.ftc.gov/system/files/documents/cases/140409goldenshoresdo.pdf.

[9] *See* FTC, Press Release, Aaron's Rent-To-Own Chain Settles FTC Charges That It Enabled Computer Spying by Franchisees (Oct. 22, 2013), *available at* https://www.ftc.gov/news-events/press-releases/2013/10/aarons-rent-own-chain-settles-ftc-charges-it-enabled-computer.

[10] *See* Facebook, Inc., C-4365 (F.T.C. July 27, 2012) (decision and order), *available at* https://www.ftc.gov/sites/default/files/documents/cases/2011/10/111024googlebuzzdo.pdf/

[11] FTC v. Sitesearch Corp., d/b/a LeapLab (D. Az. Dec. 23, 2014) (complaint), *available at* http://www.ftc.gov/systems/files/documents/cases/141223leaplabcmpt.pdf.

[12] *See* Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501-06.

[13] 15 U.S.C. §§ 6801-09.

[14] Health Insurance Portability and Accountability Act, Pub. L. No.104-191, 110 Stat. 1936 (1996) (codified in scattered sections of 18, 26, 29, and 42 U.S.C.).

[15] 15 U.S.C. § 1681 *et seq.*

[16] *See generally* FTC, INTERNET OF THINGS: PRIVACY & SECURITY IN A CONNECTED WORLD 1-4 (2015) (staff report), *available at* https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf (discussing views of workshop participants) [IOT REPORT].

[17] FTC, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY (2014), *available at* http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf [DATA BROKER REPORT].

public workshops on cutting edge topics like consumer generated health information, so-called alternative consumer scores,[18] and the potential for big data analytics to be used in ways that discriminate against consumers.[19] And, just last month, we held a public workshop on cross-device tracking, which refers to companies' efforts to correlate a consumer's activities as she moves from smartphone to tablet to desktop computer.[20]

Of course, the FTC does not do this work alone. Other federal regulators have a role in privacy and data security with respect to health care providers and hospitals,[21] banks and depository institutions,[22] and common carriers.[23][FN] States also play a vital and active role in advancing consumer privacy and data security protections. Last year, approximately 60 new privacy laws were passed at the state level in the U.S. State privacy laws range from limiting employers' ability to view their employees' social network accounts[24] and prohibiting employers and insurers from using information about certain medical conditions,[25] to requiring companies to notify consumers when they suffer a security breach involving personal information.[26] And the FTC and states work closely together on privacy, data security,[27] and a range of other consumer protection issues.[28]

---

[18] *See* FTC, Spring Privacy Series: Alternative Scoring Products (Mar. 19, 2014), *available at* https://www.ftc.gov/news-events/events-calendar/2014/03/spring-privacy-series-alternative-scoring-products.

[19] *See* FTC, Big Data: A Tool for Inclusion or Exclusion? (Sept. 15, 2014), *available at* https://www.ftc.gov/news-events/events-calendar/2014/09/big-data-tool-inclusion-or-exclusion.

[20] *See* FTC, Cross Device Tracking: An FTC Workshop (Nov. 16, 2015), *available at* https://www.ftc.gov/news-events/events-calendar/2015/11/cross-device-tracking.

[21] *See* Dept. of Health & Human Svcs., HIPAA Compliance and Enforcement, *available at* http://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/index.html (last visited Jan. 4, 2016).

[22] *See* FDIC, Privacy Choice, *available at* https://www.fdic.gov/consumers/assistance/protection/privacy/privacychoices/ (last updated July 28, 2014) (describing roles of different agencies with responsibilities for enforcing privacy laws against banks and other financial institutions).

[23] *See* FCC, Customer Privacy, *available at* https://www.fcc.gov/general/customer-privacy (describing FCC's role in enforcing privacy protections under the Communications Act and FCC rules) (last visited Jan. 4, 2016).

[24] *See* Nat'l Conf. of State Legislatures, Employer Access to Social Media Usernames and Passwords, available at http://www.ncsl.org/research/telecommunications-and-information-technology/employer-access-to-social-media-passwords-2013.aspx (last updated Nov. 18, 2014) (noting that in 2014, at least 28 states had introduced social media and employment legislation or had such legislation pending).

[25] *See, e.g.*, Privacy Rights Clearinghouse, *California Medical Privacy Fact Sheet C5: Employment and Your Medical Privacy*, *available at* https://www.privacyrights.org/content/employment-and-your-medical-privacy (last updated July 2012).

[26] *See* Nat'l Conf. of State Legislatures, Security Breach Notification Laws (Jan. 12, 2015), *available at* http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx (collecting references to more than 45 state laws).

[27] *See, e.g.*, FTC, Press Release, LifeLock Will Pay $12 Million to Settle Charges by the FTC and 35 States That Identity Theft Prevention and Data Security Claims Were False (Mar. 9, 2010), *available at* https://www.ftc.gov/news-events/press-releases/2010/03/lifelock-will-pay-12-million-settle-charges-ftc-35-states (stating that LifeLock agreed to pay $11 million to the FTC and $1 million to a group of 35 state attorneys general).

[28] *See, e.g.*, FTC, FTC and Ten State Attorneys General Take Action Against Political Survey Robocallers Pitching Cruise Line Vacations to the Bahamas (Mar. 4, 2015), *available at* https://www.ftc.gov/news-events/press-

Device and Data Security

Security is one of the biggest challenges we encounter with the Internet of Things and big data. And because these connected devices are linked to the physical world, *device* security also is a top concern. Unfortunately, there is some evidence that security vulnerabilities are rampant in the Internet of Things. A 2014 study by Hewlett-Packard found that 90 percent of connected devices are collecting personal information, and 70 percent of them are transmitting this data without encryption.[29] Part of the reason may be economic. Traditional consumer goods manufacturers that are now entering the Internet of Things market may not have spent decades thinking about how to secure their products and services from hackers in the way that traditional technology firms have. For these companies, adding security expertise may be particularly costly. But many connected devices will be inexpensive and essentially disposable. If a vulnerability is discovered on such a device, will such manufacturers have the appropriate economic incentive notify consumers, let alone patch the vulnerability?[30]

Of course, companies *should* disclose how they will protect consumers' data, and those disclosures must be truthful and not misleading. There is a long line of FTC enforcement actions against companies that failed to meet this standard.[31] And the rather arcane nature of data security does not excuse the failure of companies to apply fixes for well-known vulnerabilities or take other reasonable steps to protect consumers' data or devices.[32] The FTC is also encouraging developers to go beyond the legal requirements based on Section 5, and adopt security measures that create stronger protections for consumers.

---

releases/2015/03/ftc-ten-state-attorneys-general-take-action-against-political; FTC, FTC and Federal, State and Local Law Enforcement Partners Announce Nationwide Crackdown Against Abusive Debt Collectors (Nov. 4, 2015), *available at* https://www.ftc.gov/news-events/press-releases/2015/11/ftc-federal-state-local-law-enforcement-partners-announce (announcing 30 new actions targeting illegal debt collection practices, including joint actions brought by the FTC and Illinois and New York).

[29] Hewlett-Packard, *Internet of Things Research Study* 2 (July 2014), *available at* http://h20195.www2.hp.com/V2/GetDocument.aspx?docname=4AA5-4759ENW&cc=us&lc=en.

[30] *See* IOT REPORT, *supra* note 16, at 13-14.

[31] For early cases, see *In re GeoCities, Inc.*, 127 F.T.C. 94 (1999) (consent order) (settling charges that website had misrepresented the purposes for which it was collecting personally identifiable information from children and adults); *FTC v. Toysmart.com, LLC*, No. 00-11341-RGS, 2000 WL 34016434 (D. Mass. July 21, 2000) (consent order) (challenging website's attempts to sell children's personal information, despite a promise in its privacy policy that such information would never be disclosed). For a more recent case, see, e.g., Snapchat, Inc., No. C-4501 (F.T.C. Dec. 23, 2014), (decision and order), *available at* http://www.ftc.gov/system/files/documents/cases/141231snapchatdo.pdf.

[32] See, e.g., TRENDNet, Inc., No. C-4426 (F.T.C. Feb. 7, 2014), at ¶ 8 (complaint), *available at* http://www.ftc.gov/system/files/documents/cases/140207trendnetcmpt.pdf; GMR Transcription Servs., No. C-4482 (F.T.C. Aug.14, 2014) (consent order), *available at* http://www.ftc.gov/system/files/documents/cases/140821gmrdo.pdf; HTC America, Inc., C-4406 (F.T.C. June 25, 2013) (decision and order), *available at* http://www.ftc.gov/sites/default/files/documents/cases/2013/07/130702htcdo.pdf.

Still, security vulnerabilities may be hidden deep in the code that runs an app or device. A vulnerability may not become apparent until a device is connected to an environment for which it wasn't designed, or perhaps until consumers use a device or service in unexpected ways.

All of these factors point to the need to take an "all hands on deck" approach to data security, with security researchers playing an important role in bringing security flaws to light. Researchers have found vulnerabilities in systems ranging from electronic voting systems[33] to connected cars[34] to online learning platforms.[35] This kind of research continues to raise difficult questions about how and when to disclose vulnerabilities to the developer of the product or service.[36] It is not always the kind of news that companies – or law enforcement agencies – want to hear, and some researchers have been prosecuted for their activities.[37] In October, the FTC criticized a proposal in Congress that would have made certain kinds of security research on connected cars illegal, on the ground that this provision would cut off a useful source of information about security vulnerabilities that could affect consumers' physical safety.[38] Fortunately, many companies see the value of learning about vulnerabilities in their products, and many are willing to pay "bug bounties" to be the first to be told of a vulnerability in one of their own products.

Still, security research is difficult and uncertain. Even when researchers have access to source code, they may have a hard time identifying errors.[39] Fortunately, security experts have

---

[33] *See, e.g.*, Calif. Sec. of State, Top-to-Bottom Review, *available at* http://www.sos.ca.gov/elections/voting-systems/oversight/top-bottom-review/ (last visited Jan. 4, 2016) (describing security, accessibility, and documentation reviews of electronic voting systems and linking to reports); Ariel J. Feldman, J. Alex Halderman, and Edward W. Felten, Security Analysis of the Diebold AccuVote-TS Voting Machine (Sept. 13, 2006), *available at* https://citp.princeton.edu/research/voting/.

[34] *See, e.g.*, Andy Greenberg, *Hackers Remotely Kill a Jeep on the Highway – With Me in It*, WIRED (July 21,2015 6:00 a.m.), *available at* http://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/.

[35] *See* Jonathan Mayer, A Funny Thing Happened on the Way to Coursera (Sept. 4, 2014), *available at* http://webpolicy.org/2014/09/04/a-funny-thing-happened-on-the-way-to-coursera/; Coursera, Response to Reported Vulnerability in Instructor Access to Learner Data (Sept. 5, 2014 3:02 a.m.), *available at* http://blog.coursera.org/post/96686805237/response-to-reported-vulnerability-in-instructor.

[36] *See* Angela Simpson, Deputy Asst. Sec. for Communications and Information, NTIA, Enhancing the Digital Economy Through Collaboration on Vulnerability Research Disclosure (July 9, 2015), *available at* https://www.ntia.doc.gov/blog/2015/enhancing-digital-economy-through-collaboration-vulnerability-research-disclosure.

[37] *See* David Kravets, *Appeals Court Reverses Hacker/Troll "weev" Conviction and Sentence*, ARSTECHNICA (Apr. 14, 2014 11:49 a.m. EDT), *available at* http://arstechnica.com/tech-policy/2014/04/appeals-court-reverses-hackertroll-weev-conviction-and-sentence/ (describing prosecution and ultimate reversal of conviction for conspiracy to violate the Computer Fraud and Abuse Act in the course of demonstrating how an online registration process exposed consumers' personal information).

[38] *See* FTC, Prepared Statement on "Examining Ways to Improve Vehicle and Roadway Safety" Before the Subcommittee on Commerce, Manufacturing, and Trade of the Committee on Energy and Commerce, United States House of Representatives 5-6 (Oct. 21, 2015), *available at* https://www.ftc.gov/public-statements/2015/10/prepared-statement-federal-trade-commission-examining-ways-improve-vehicle.

[39] *See, e.g.*, Ka-Ping Yee, Building Reliable Voting Machine Software 148 (2007) (PhD thesis), *available at* http://zesty.ca/pubs/yee-phd.pdf. Yee introduced three errors into a 100-line section of code for a system that he was developing. He told a group of computer security experts that this section contained errors (though not how

devised many other ways to test components and systems more quickly and on a larger scale. No single method is capable of finding every vulnerability, but different methods, used in combination, can provide more assurance that systems are trustworthy.

To bring all of this back to consumers and their trust in companies and their services, giving consumers timely information about vulnerabilities, how to fix them, and making it easy to do so are some of the best ways to maintain trust. We will be discussing all of these issues at the FTC's upcoming "Start with Security" event, right here in Seattle, on February 9th. Co-sponsored by the University of Washington's Tech Policy Lab and its School of Law Technology, Law and Public Policy Clinic, our one-day conference will provide companies – particularly small companies – with practical advice and recommended strategies for implementing effective data security.[40]

Privacy

So let me now turn to privacy in the era of big data and the Internet of Things. Consumers want to know – and should be able easily to find out – what information companies are collecting, where they're sending it, and how they're using it. This kind of information is important to consumers' decisions about whether to use digital products and services in the first place. The FTC came to this realization early in the history of the commercial Internet.[41]

But many companies, including data brokers, ad networks, and analytics firms, operate in the background with consumer data, and their activities can significantly affect consumers. While consumers might benefit from the activities of some of these "behind the scenes" operators – by receiving more relevant advertising, for example – consumers should have choices about where their data ends up and how it is used.

Consumer choice is enhanced by giving consumers "just in time" information, at key moments when it is most relevant to them, such as when they are deciding to download an app or make purchases on a connected device. But companies should also help consumers navigate the complex ecosystem of data, devices, and big data analytics operating behind the scenes, so that consumers understand the practices that can affect them, and exercise choices about the practices.

---

many). In a total of 20 person-hours devoted to these 100 lines of code, the security experts found two of the three bugs – and not the one that Yee considered the most difficult to find.

[40] FTC, Start with Security – Seattle, *available at* https://www.ftc.gov/news-events/events-calendar/2016/02/start-security-seattle (last visited Jan. 7, 2016).

[41] *See, e.g.*, *In re GeoCities, Inc.*, 127 F.T.C. 94 (1999) (consent order), *available at* https://www.ftc.gov/enforcement/cases-proceedings/982-3015/geocities (settling charges that website had misrepresented the purposes for which it was collecting personally identifiable information from children and adults); *FTC v. Toysmart.com, LLC*, No. 00-11341-RGS, 2000 WL 34016434 (D. Mass. July 21, 2000) (consent order), *available at* https://www.ftc.gov/enforcement/cases-proceedings/x000075/toysmartcom-llc-toysmartcom-inc (challenging website's attempts to sell children's personal information, despite a promise in its privacy policy that such information would never be disclosed).

Many companies and organizations understand this important connection between transparency, privacy, and trust. But being transparent in this data-intensive age is challenging. With the Internet of Things, many connected devices do not have a user interface to present information to consumers about data collection. Devices are becoming more numerous, adding to the mountain of information that companies present to consumers in privacy policies. As devices become integrated into homes and other physical spaces, there are also questions around who should receive disclosures about data collection and use practices. How will the consumers who buy a device – and the innocent bystanders around them – know when a device is recording images or audio? And there are other questions, like how can consumers choose to avoid having their data collected? For how long will their data be kept by the companies who are collecting it? And how will these companies keep the data secure?

Companies that provide connected devices should recognize that providing transparency will require some creative thinking. Visual and auditory cues, and immersive apps and websites should be employed to describe to consumers, in a meaningful and relatively simple way, the nature of the information being collected. The same signals should be used to provide consumers with choices about whether any of this information can be used by entities or persons who fall outside the context in which the consumer is employing the device, and in which the consumer expects her information to remain private.

Another promising tool for providing information to consumers, as well as allowing them to exercise meaningful choices, is the "command center" that companies are now developing to run multiple household connected devices.[42] The driving force here is convenience, but these command centers could also provide an opportunity for consumers to understand the information their devices are generating, and to control where that information goes. After all, if you can have a centralized interface to program your garage door, thermostat, television, refrigerator, and who knows what else, you ought to be able to use that same interface to make meaningful choices about the data your devices will collect and where they will send it.

Fairness and Consumer Trust

Finally, let me turn to the issue of fairness in big data analytics. Data brokers are a good place to start this discussion. Data brokers are companies that assemble individual profiles on consumers by collecting information from far-flung sources, but typically do not interact with consumers themselves. Through these profiles, consumers can end up in marketing segments drawn along lines of race, ethnicity, financial status, health conditions, and other sensitive characteristics. With all of this data, and the inferences that data brokers can draw from it, they put consumers into categories or segments that had labels such as "Financially Challenged", "Bible Lifestyle",[43] "Diabetes Interest"[44], "Metro Parents" (which are lists of single parents who are "primarily high school or vocationally educated" and are handling the "stresses of urban life on a small budget"), and "Timeless Traditions" (a list of immigrants who "speak[] some English,

[42] *See* Don Clark, *The Race to Build Command Centers for Smart Homes*, WALL ST. J. (Jan. 5, 2015), *available at* http://www.wsj.com/articles/the-race-to-build-command-centers-for-smart-homes-1420399511.

[43] DATA BROKER REPORT, *supra* note 17, at 20 n.52, 21.

[44] *Id.* at 46, 55.

but generally prefer[] Spanish").[45] These practices were recently brought to light by the FTC and others.[46] Once again, while this kind information can be used for relatively benign purposes, or even in ways that will enhance financial inclusion, this kind of information has also been used to harm vulnerable consumers.

Even big data analytics projects within a company, using its own data, could create questions about fairness. For example, a company might analyze its own data in an effort to identify "good" versus "troublesome" customers. But what if this analysis ends up tracking individuals along racial or ethnic lines? A *Harvard Business Review* article argues that this kind of result isn't just possible, but inevitable.[47]

Many big data-driven decisions are based on some kind of score – a number generated by an algorithm that gives some indication of what a consumer is likely to be interested in or how she is likely to behave. A familiar example is the credit score. Credit scores are basically predictions of how likely a consumer is to repay a debt. The higher the score, the lower the credit risk. In their early days, credit scores were used strictly for credit decisions – whether you would qualify for a mortgage, for example, and the interest rate that you would be offered. Over time, the use of these same credit scores expanded to other major decisions about consumers, such as whether a prospective employer would extend a job offer to an applicant, or an insurance company would charge a higher premium on auto or homeowners insurance.

We know a great deal about what information is in our credit reports and what our traditional credit scores are for a simple reason. Congress has required some transparency in credit scoring. In 2003, Congress instructed the FTC and the Federal Reserve to study whether one type of popular credit score used for auto insurance employed factors that serve as proxies for race, gender, or other traits that could give rise to unlawful discrimination.[48] In addition, Congress required credit bureaus to make consumers' credit reports available to them for free,[49] and credit scores are increasingly becoming available for free to consumers.[50]

---

[45] *Id.* at 20 n.52.

[46] *See, e.g.*, CBS News, 60 Minutes; The Data Brokers: Selling Your Personal Information (last updated Aug. 24, 2014), *available at* http://www.cbsnews.com/news/data-brokers-selling-personal-information-60-minutes/; World Privacy Forum, *The Scoring of America* (2014), *available at* http://www.worldprivacyforum.org/2014/04/wpf-report-the-scoring-of-america-how-secret-consumer-scores-threaten-your-privacy-and-your-future/; U.S. Senate Committee on Commerce, Science, and Transportation, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes* (Dec. 2013) (staff report), *available at* http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=0D2B3642-6221-4888-A631-08F2F255B577.

[47] *See* Michael Schrage, *Big Data's Dangerous New Era of Discrimination*, HARVARD BUSINESS REVIEW BLOG NETWORK (Jan. 29, 2014, 8:00 a.m.), *available at* http://blogs.hbr.org/2014/01/big-datas-dangerous-new-era-of-discrimination/.

[48] *See* Fair and Accurate Credit Transactions Act of 2003 § 215, Pub. L. 108-159 (Nov. 5, 2003), *available at* https://www.congress.gov/bill/108th-congress/house-bill/2622/text ["FACTA"].

[49] *See* FACTA, *supra* note 48, at § 211 (codified at 15 U.S.C. § 1681j(a)).

[50] *See* Annamaria Andriotis, Millions of Consumers to Gain Access to Credit Scores, WALL ST. J. TOTAL RETURN (Oct. 17, 2014 2:38 PM), *available at* http://blogs.wsj.com/totalreturn/2014/10/17/millions-of-consumers-

These transparency requirements and practices have been good for consumers, credit bureaus, and companies that rely on credit scores to make business decisions.[51] With the transparency provided by free credit reports and, increasingly, scores, consumers can more effectively exercise their rights to dispute and correct inaccurate information. And the thorough analysis of one critical type of credit score by the FTC and Federal Reserve made users more confident that this score was not discriminatory.

Today, we're seeing a proliferation of other types of scores being used to make eligibility determinations covered by the Fair Credit Reporting Act.[52] While these scores are subject to the same obligations of access, accuracy, security, and other requirements imposed by the FCRA, they haven't yet been subject to the same kind of scrutiny that Congress and the federal agencies brought to bear on traditional credit scores.[53] The use of new sources of information, including information that goes beyond traditional credit files, to score consumers raises fresh questions about whether these alternative scores may have disparate impacts along racial, ethnic, or other lines that the law protects.

Unfortunately, it's not realistic to rely on the approach that the FTC took – to understand one type of score used for auto insurance – to gain an understanding of the full spectrum of scoring models used today. It took the FTC nearly four years to conduct its study. The FTC – and all other federal agencies for that matter – simply do not have the capacity to study every score out there. This approach simply will not scale.

Moreover, scoring algorithms and other forms of big data analytics rely on statistical models and data system designs that few on the outside understand in detail. And even if we on the outside could peer into the hundreds of scoring algorithms that could potentially affect consumers, what would we learn? We might learn which features of a data set are used in a given algorithm, and what weight a company attaches to them. These details might be so abstract, and so rapidly changing, that they do not tell government, consumers or other

---

to-gain-access-to-credit-scores/ (reporting that some credit card issuers are reporting consumers' credit scores on their monthly statements).

[51] *See* BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, REPORT TO CONGRESS ON CREDIT SCORING AND ITS EFFECTS ON THE AVAILABILITY AND AFFORDABILITY OF CREDIT S-1 (Aug. 2007), *available at* http://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf ("The large savings in cost and time that have accompanied the use of credit scoring are generally believed to have increased access to credit, promoted competition, and improved market efficiency.") [FRB, CREDIT SCORING REPORT].

[52] *See generally* FTC, Transcript of Spring Privacy Series: Alternative Scoring Products (Mar. 19, 2014), *available at* http://www.ftc.gov/system/files/documents/public_events/182261/alternative-scoring-products_final-transcript.pdf; Pam Dixon and Robert Gellman, The Scoring of America: How Secret Consumer Scores Threaten Your Privacy and Your Future (Apr. 2014), *available at* http://www.worldprivacyforum.org/wp-content/uploads/2014/04/WPF_Scoring_of_America_April2014_fs.pdf.

[53] *See generally* FRB, CREDIT SCORING REPORT, *supra* note 51; FTC, CREDIT-BASED INSURANCE SCORES: IMPACTS ON CONSUMERS OF AUTOMOBILE INSURANCE 62-73 (July 2007), *available at* https://www.ftc.gov/sites/default/files/documents/reports/credit-based-insurance-scores-impacts-consumers-automobile-insurance-report-congress-federal-trade/p044804facta_report_credit-based_insurance_scores.pdf.

concerned stakeholders much at all about what really matters – which is how the algorithms are actually used and whether they have discriminatory or other inappropriate effects.

This suggests that testing the *effects* of big data analytics may be a promising way to go. Doing this kind of analysis from the outside is difficult. Researchers have done some proof-of-concept studies, but they required considerable work and involved efforts to tackle some cutting-edge research questions.[54]

This means that companies using scoring models should themselves do more to determine whether their own data analytics result in unfair, unethical, or discriminatory effects on consumers.

In addition to scrutinizing their own practices, companies should do much more to inform consumers of what is happening with their data. Companies can get creative with user interfaces to provide consumers with more meaningful, usable access to their data. This will serve two purposes: meaningful usable access for consumers will help address questions about the role that big data analytics plays in the marketplace and whether consumers are being treated fairly; and it will provide a helpful check on potentially troublesome data practices. As Louis Brandeis famously said, "sunlight is said to be the best of disinfectants."[55]

Technologists have a key role to play, too. They have the skills to make data access tools that are easy for consumers to use, and they have the technical insights that are necessary to determine whether specific analytics practices pose risks of excluding, or otherwise placing at a disadvantage, groups defined according to sensitive traits. I am hopeful that companies will give technologists, including designers and user interface experts, the support and resources needed to tackle these critically important challenges.

* * * * * *

Protecting privacy and security and assuring fairness in the complex and rapidly changing digital world around us is a critical, yet complex task. Basic consumer protection principles can provide a compass to guide us through this complexity. Consumers want reasonable assurances that companies are keeping the data collected about them, as well as their connected devices themselves, secure. Consumers want to know what data companies are collecting about them, how companies are using it, and how they can exercise some control over the collection and use of their data. And consumers want to know that they are not being treated unfairly or in a discriminatory manner. For now, the rapid changes in big data analytics and the Internet of Things have made it difficult to meet some of these expectations in practice. The key point, however, is that these are the enduring expectations of consumers, rather than relics of a simpler world. Meeting these expectations and, more broadly, ensuring consumer trust is the key to

[54] *See, e.g.*, Amit Datta, Michael Carl Tschantz, and Anupam Datta, *Automated Experiments on Ad Settings: A Tale of Opacity, Choice, and Discrimination*, *in* PROC. ON PRIVACY ENHANCING TECHS. (PETS) 2015, *available at* http://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf; Latanya Sweeney, *Discrimination in Online Ad Delivery*, 56 COMMS. OF THE ACM, No. 5, at 44-54 (2013), *available at* http://cacm.acm.org/magazines/2013/5/163753-discrimination-in-online-ad-delivery/fulltext.

realizing the full potential of this highly connected, data-driven world. And all of you here – government officials, representatives of industry, and civil society leaders – have important roles to play in this endeavor.

Thank you.