U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON ENERGY AND COMMERCE

November 27, 2017

TO: Members, Subcommittee on Communication and Technology and the Subcommittee on Digital Commerce and Consumer Protection

FROM: Committee Majority Staff

RE: Hearing entitled "Algorithms: How Companies' Decisions About Data and Content Impact Consumers"

I. INTRODUCTION

The Subcommittee on Communications and Technology and the Subcommittee on Digital Commerce and Consumer Protection will hold a hearing on Wednesday, November 29, 2017, at 10:00 a.m. in 2123 Rayburn House Office Building. The hearing is entitled "Algorithms: How Companies' Decisions About Data and Content Impact Consumers."

II. WITNESSES

- Catherine Tucker, Sloane Distinguished Professor of Management Science and Professor of Marketing, MIT Sloane School of Management;
- Omri Ben-Shahar, Leo and Eileen Herzel Professor of Law, University of Chicago Law School;
- Kate Klonick, Resident Fellow, Information Society Project, Yale Law School;
- Michael Kearns, Professor and National Center Chair, Department of Computer and Information Science, University of Pennsylvania;
- Laura Moy, Deputy Director, Georgetown Law Center on Privacy and Technology; and,
- Frank Pasquale, Professor of Law, University of Maryland.

III. BACKGROUND

A. Consumer Protection

Recent high-profile incidents such as the unauthorized access of consumer credit information at the credit reporting agency Equifax have raised questions regarding the security of consumers' most sensitive personal information when it is contained in databases that the consumer does not control, and over which the consumer can exert little influence. The

Committee has previously examined the Equifax data breach and ways to secure consumer credit data in prior hearings.[1]

Additionally, revelations of the use of Internet social media sites to attempt to influence voting decisions have highlighted the vast trove of personal data that these sites make available to advertisers and other third parties,[2] and how that data can be used to develop detailed profiles of individual Americans.[3]

The migration of services and products on-line presents challenges to existing modes of regulation designed for the physical world. Maintaining consumer protections and ensuring transparency in such an environment are factors in providing consumers the tools for informed decision making.

B. Personalization Online

In the late 1990s, information technologist Doug Laney developed the term "infonomics" to describe an emerging understanding of the value of information and ways that businesses should account for and monetize information as an asset.[4] By 2011, the benefits of using refined data about individual consumers was becoming more widely recognized. As noted by the McKinsey Global Institute:

> Retailing is an obvious place for data-driven customization because the volume and quality of data available from Internet purchases, social-network conversations, and, more recently, location-specific smartphone interactions have mushroomed. But other sectors, too, can benefit from new applications of data, along with the growing sophistication of analytical tools for dividing customers into more revealing microsegments.[5]

By 2014, the practice of using highly personalized data points from multiple sources began to become widely available to companies seeking to develop more sophisticated profiles of consumers. As commenter Gurbaksh Chahal noted, "[a]ll of these tiny interactions send an overwhelming amount of valuable information to marketers in real-time — browsing behavior,

---

[1] *Oversight of the Equifax Data Breach: Answers for Consumers: Hearing Before the H. Comm. on Energy & Commerce* (Oct. 3, 2017). Rebroadcast available at https://energycommerce.house.gov/hearings/oversight-equifax-data-breach-answers-consumers/; *Securing Consumers' Credit Data in the Age of Digital Commerce: Hearing Before the H. Comm. on Energy and Commerce* (Nov.1, 2017). Rebroadcast available at https://energycommerce.house.gov/hearings/securing-consumers-credit-data-age-digital-commerce/.

[2] *See, e.g.*, Craig Timberg & Elizabeth Dwoskin, *Russian Content on Facebook, Google and Twitter Reached Far More Users Than Companies First Disclosed, Congressional Testimony Says*, Wash. Post, Oct. 30, 2017.

[3] *See, e.g.*, Samatha Shorey & Philip Howard, *Automation, Big Data and Politics: A Research Review*, Int'l J. Comm., 10, 24 (May 2016), http://ijoc.org/index.php/ijoc/article/view/6233/1812.

[4] For an outline of the valuation of information *see* Doug Laney, *Infonomics: The Economics of Information and Principles of Information Asset Management*, presented at The Fifth MIT Information Quality Industry Symposium, (July 13-15, 2011).

[5] Brad Brown, Michael Chui, & James Manyika, *Are You Ready for the Era of 'Big Data'?*, McKinsey Q. (October 2011), http://ce.uoregon.edu/aim/DataDrivenOrgs2013/AreYouReadyforBigData.pdf.

social media interactions, mobile device usage, geo-location, click-through rates, online purchase patterns, and the list goes on."[6]

In addition to the availability of large amounts of personal data, marketers began to develop more and more sophisticated algorithms to help categorize and monetize the newly available information. The development of sophisticated data analytical tools would be a crucial development for the creation of highly individualized advertising campaigns. Chahal predicted that "[s]mart marketers will crack the big data code and use these exclusive findings to beat out their competitor's one customer at a time by making informed decisions on the right course of action to take, at the right time with the right message over the right channel."[7]

Social media platforms provide marketers with the ability to utilize their expansive collection of highly categorized and segmented customer profiles for purposes of targeted advertising. Additionally, third party apps that partner with these sites to allow a user access to the app through a social network login likewise create the ability to gather personal data.[8] By accessing an app utilizing the social network's login credential, the social network obtains information that it can then use to further refine its model of the user, including by tracking the user across multiple third party sites.[9]

Utilizing "cookies," websites can follow a user as he or she visits multiple websites. Cookies can be placed by a particular site a user is visiting, or by a third party that has partnered with the original website to obtain user data. Device fingerprinting can be used to identify a user through the method he or she is accessing the Internet. In addition, mobile apps routinely seek permission to access information on a user's mobile device including photos, call information and web browsing. These and other methods allow websites and developers to gather and use information which is often stored in a proprietary database.[10]

The use of this data allows advertisers and others to deliver highly personalized information to consumers, which can help inform them of their options. This in turn can provide consumers with greater opportunities to compare prices and services across a broader universe of providers.[11] However, some critics contend that the ability of some Internet companies to obtain highly personalized information about individuals, categorize and profile individuals using that information, and profit from selling those profiles to third parties actually limits choice and

---

[6] Gurbaksh Chahal, *2014: The Year Marketer's Big Data Gets Real,* WIRED, (Dec. 2013), https://www.wired.com/insights/2013/12/2014-year-marketers-big-data-gets-real/.
[7] *Id*.
[8] Paul Bischoff, *Facebook, Twitter, Google+, or LinkedIn … Which Should You Log in With?*, Comparitech, (Jan. 13, 2016) https://www.comparitech.com/blog/vpn-privacy/facebook-twitter-google-or-linkedin-which-should-you-log-in-with/.
[9] Baratunde Thurston, *Why Using Facebook, Google, And Twitter to Log Into Apps is a Problem,* FAST COMPANY (Apr. 13, 2015), https://www.fastcompany.com/3044280/the-ghosts-of-app-permissions-past.
[10] For more information see the FTC's Consumer Information Website Online Tracking, available at https://www.consumer.ftc.gov/articles/0042-online-tracking
[11] *See, e.g.*, Lora Kolodny, *How Consumers Can Use Big Data,* Wall St. J. (Mar. 23, 2014).

creates opportunities for unscrupulous behavior.[12] The Federal Trade Commission staff reaffirmed how transparency, choice, and privacy-by-design and other contextual principles guide its enforcement work to protect consumers.[13]

C. Content Shaping and Delivery

Further concerns have arisen as to how content is shaped over private social media platforms, which have emerged as an essential channel for free speech and democratic participation. In this context, understanding the flow of information from data collectors back toward users is as important as the understanding how data is collected and protected. Platforms that provide free services often are able to do so through the monetization of user data for targeted information or other mechanisms that can affect not just one user's online experience, but also that of the individuals that a user is connected with, as well as others that may share similar geographic or other statistical traits. And given the platforms' increasingly prominent governance function as content moderators, their practices in this regard have increasing implications for free speech in our society.

Content is sometimes filtered or prioritized using proprietary algorithms, or other intellectual property, that are not subject to transparency requirements and thus not well understood by the public.[14] There are concerns about intentional or unintentional bias being built into these machine-based decision-makers during their development.[15] Moreover, many controversial decisions regarding content moderation are made not by algorithms, but by employees enforcing or developing internal guidelines, which may or may not be public. In the context of concerns about the diversity of the employees responsible for making these decisions, questions of bias, influence, and control are magnified.[16]

As this Committee further reviews how content reaches consumers over the country's Internet infrastructure, there is a clear nexus in reviewing how private platforms are managing and implementing their power to police users' ability to communicate with others. As these platforms continue to self-regulate and moderate content, they can effectively censor speech with little oversight due to their own First Amendment protections. Witnesses will touch upon how

---

[12] *See, e.g.*, Franklin Foer, *World Without Mind: The Existential Threat of Big Tech,* Penguin Books (2017); Adam Alter, *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked,* Penguin Books (2017).

[13] *See, e.g.*, Comment of the Staff of the Bureau of Consumer Protection of the Federal Trade Commission before the Federal Communications Commission, *In the Matter of Protecting the Privacy of Customers of Broadband and Other Telecommunications Services*, WC Docket No. 16-106, p. 16 (May 26, 2016), https://www.ftc.gov/system/files/documents/advocacy_documents/comment-staff-bureau-consumer-protection-federal-trade-commission-federal-communications-commission/160527fcccomment.pdf.

[14] Promises made by companies under the Federal Trade Commission's jurisdiction about their data collection and use practices are subject to compliance with Section 5 of the FTC Act's prohibition against unfair or deceptive acts or practices.

[15] https://techcrunch.com/2017/04/30/algorithmic-accountability/.

[16] https://qz.com/1047453/a-google-employees-viral-anti-diversity-memo-shows-americas-political-divide-has-spread-to-silicon-valley/.

content regulation practices impact consumers both online and offline, who decides the appropriateness and priority of what is being shared, and how those decisions are being made.

D. Disclosures to Consumers: Regulations and Privacy Policies

Regulation of Use of Consumer Information on the Internet. At the Federal level regulation of the gathering and use of information obtained through the Internet is largely accomplished through the regulation of various industry sectors, rather than via a general, Internet-wide rule.[17] For instance, the healthcare industry and the financial services sectors are both subject to proscriptive regulations regarding the use and distribution of personal information that apply equally to data collected on and off the Internet.[18] One area under Federal law that does impose blanket policies with respect to the use of personal information gathered on the Internet is the Children's Online Privacy Protection Act, which requires operators of websites or online services directed to children under 13 years of age, and operators of other websites or online services that have actual knowledge that they are collecting personal information online from a child under 13 years of age, to fulfill certain obligations.[19] These include the requirement that Internet operators provide notice to parents and obtain verifiable parental consent prior to collecting, using, or disclosing personal information from children under 13 years of age; keep secure the information they collect from children; and bars operators from conditioning children's participation in activities on the collection of more personal information than is reasonably necessary to participate in such activities.[20]

In a 1998 report the Federal Trade Commission (FTC) issued the Fair Information Practices Principles, which summarized privacy protections common to governments, guidelines and model codes (which in most cases predated the Internet).[21] The FTC has brought over 150 enforcement actions regarding online privacy issues, including spam, social networking, behavioral advertising, pretexting, spyware, peer-to-peer file sharing, and mobile.[22] These actions have been taken under the FTC's general authority to protect consumers from "unfair or deceptive acts or practices in or affecting commerce. . . ."[23]

Several States have implemented their own laws regarding the use of personal information gathered online.[24] These include the requirement to post a conspicuous privacy policy on websites or online services, and that website operators disclose in a privacy policy how

---

[17] Clark D. Asay, *Consumer Information Privacy and the Problem(s) of Third-Party Disclosures*, 11 Nw. J. Tech. & Intell. Prop. 321 (2013), http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss5/3.
[18] *See* Health Insurance Portability and Accountability Act (42 U.S.C. § 1320d–6); Fair Credit Reporting Act (15 U.S.C. § 1681 et seq); Gramm-Leach-Bliley Act (15 U.S. Code §§ 6801-6809).
[19] Children's Online Privacy Protection Act of 1998 (15 U.S.C. §§ 6501–6505); *see* Matt Richtel & Miguel Helft, *Facebook Users Who Are Under Age Raise Concerns*, N.Y. Times, Mar. 11, 2011.
[20] 16 CFR 312 *et. seq*.
[21] https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-report-congress/priv-23a.pdf.
[22] *See* https://www.ftc.gov/reports/privacy-data-security-update-2016.
[23] Federal Trade Commission Act Sec. 5 (15 U.S.C. § 45).
[24] For a general discussion see National Conference of State Legislatures, State Laws Related to Internet Privacy, available at http://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-related-to-internet-privacy.aspx.

they respond to a web browser "Do Not Track" signal (California); the requirement that an operator gathering personally identifiable information about State residents have a privacy policy conspicuously available on its Internet website, online or cloud computing service, online application, or mobile application (Delaware); and the requirement that operators of Internet websites or online services that collect personally identifiable information from residents of the state notify consumers about how that information is used (Nevada).[25]

Privacy Policies. Regardless of any legal requirement, many websites and applications contain privacy policies and general disclosures regarding their use of customer information. These often include information on how information is collected, used, and shared, and include consents for such activity from the user. For instance, Twitter's privacy policy states that "[w]hen using any of our Services you consent to the collection, transfer, storage, disclosure, and use of your information as described in this Privacy Policy."[26] Google notes that "[w]e collect information about the services that you use and how you use them, like when you watch a video on YouTube, visit a website that uses our advertising services, or view and interact with our ads and content." This information includes unique device identifiers, log information such as phone numbers and time, date and duration of calls, location information and data stored locally on a user's device.[27] Amazon discloses that:

> Examples of the information we collect and analyze include the Internet protocol (IP) address used to connect your computer to the Internet; login; e-mail address; password; computer and connection information such as browser type, version, and time zone setting, browser plug-in types and versions, operating system, and platform; purchase history, which we sometimes aggregate with similar information from other customers to create features like Top Sellers; the full Uniform Resource Locator (URL) clickstream to, through, and from our Web site, including date and time; cookie number; products you viewed or searched for; and the phone number you used to call our 800 number.[28]

Facebook states that "[w]e receive information about you and your activities on and off Facebook from third-party partners, such as information from a partner when we jointly offer services or from an advertiser about your experiences or interactions with them."[29]

Website and app operators collect this data for a wide variety of purposes largely related to enhancing the accuracy of advertising. One commenter noted that "Big Data and data mining helps businesses understand who is most likely to respond to new marketing campaigns and the best way to reach them. This not only helps control the costs of advertising but increases conversion rates from hit-or-miss income to steady streams of cash flow. In addition to direct advertising data mining can help retailers manage the arrangement of shelf space so that items

---

[25] *See* Calif. Bus. & Prof. Code § 22575 and Calif. Bus. & Prof. Code § 22575-22578; Del. Code Tit. 6 § 205C; NRS 603 *et. seq*.
[26] https://twitter.com/en/privacy.
[27] https://www.google.com/policies/privacy/.
[28] https://www.amazon.com/gp/help/customer/display.html?nodeId=468496.
[29] https://www.facebook.com/about/privacy.

frequently bought together are placed optimally to maximize customer experience, increase transactions and boost ticket averages."[30]

Effectiveness of Privacy Policy Disclosure. Despite the ubiquity of disclosures and Internet company privacy policies, there is a lack of consensus as to whether such policies actually increase protection for consumers. While some scholarship maintains that disclosure policies empower individuals to control the level of information they are comfortable sharing ("effective privacy notices serve an important function in addressing risk issues related to e-commerce"),[31] others purport that greater disclosure can actually increase the propensity of providers of personal information to inaccurately gauge the risks involved.[32] Other scholars question the usefulness of mandated disclosure policies altogether ("Although mandated disclosure addresses a real problem and rests on a plausible assumption, it chronically fails to accomplish its purpose. Even where it seems to succeed, its costs in money, effort, and time generally swamp its benefits. And mandated disclosure has unintended and undesirable consequences, like driving out better regulation and hurting the people it purports to help.")[33]

IV. ISSUES

The following issues may be examined at the hearing:

- How is personal information about consumers collected through the Internet, and how do companies use that information?
- How do companies make decisions about content that consumers see online?
- How effective are current policies and communications with consumers regarding the collection and use of personal data?

V. STAFF CONTACTS

If you have any questions regarding this hearing, please contact Melissa Froelich, Robin Colwell, Gregory Zerzan, Tim Kurth, or Paul Jackson of the Committee staff at (202) 225-2927.

---

[30] *Big Data: Why Do Companies Collect and Store Personal Data,* LE VPN (May 26, 2017), https://www.le-vpn.com/why-companies-collect-big-data/.

[31] George Milne & Mary Culnan, *Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices*, 18 J. of Interactive Marketing (2004), https://doi.org/10.1002/dir.20009.

[32] Laura Brandimarte, Alessandro Acquisti & George Loewenstein, *Misplaced Confidences: Privacy and the Control Paradox*, 4 Soc. Psychol. & Personality Sci. (Aug. 9, 2012), https://doi.org/10.1177/1948550612455931. *See also* Mark Gazaleh, *Online Trust and Perceived Utility for Consumers of Web Privacy Statements,* University of Westminster Business School (2008), http://www.academia.edu/31116853/Online_trust_and_perceived_utility_for_consumers_of_web_privacy_statements.

[33] Omri Ben-Shahar & Carl Schneider, *The Failure of Mandated Disclosure,* 159 U. Pa. L. Rev. 647 (2010), https://www.law.upenn.edu/.../BenShaharSchneider159U.Pa.L.Rev.647(2011).pdf.