April 9, 2020

The Honorable Roger Wicker
Chairman, Senate Committee on
Commerce, Science, and Transportation
512 Dirksen Senate Office Building
Washington, DC 20510

The Honorable Maria Cantwell
Ranking Member, Senate Committee on
Commerce, Science, and Transportation
512 Dirksen Senate Office Building
Washington, DC 20510

Dear Chairman Wicker and Ranking Member Cantwell:

New America's Open Technology Institute (OTI) appreciates the opportunity to submit a statement for the record for the Senate Committee on Commerce, Science, and Transportation's hearing on "Enlisting Big Data in the Fight Against Coronavirus." OTI promotes equitable access to an internet and digital technologies that are open and secure, and views technology not as an end in and of itself, but a means. We support and defend the right to privacy and freedom from surveillance; our technologies are designed for safety and security; our research methods are respectful and sensitive to privacy concerns.

As policymakers turn to big data to combat the impact of the coronavirus, Congress must ensure that robust privacy protections of consumers are upheld and not compromised with the desire to take sweeping action. While every resident is feeling the threats to our health as well as the effects from this pandemic on our economy and way-of-life, marginalized communities are especially feeling the dire consequences.

Mobilizing private industry to leverage the vast troves of data they collect and work hand-in-hand with public health entities holds promise to be a useful tool against this devastating virus. However, especially in this time of pandemic, there must be strict limitations on the data shared with the government to avoid further threats to consumers during these troubled times. Further, companies should only provide government entities with data that is necessary and helpful to public health authorities. During this difficult time, it is even more important to apply modern, sophisticated privacy-enhancing technologies such as differential privacy. Finally, those companies that decide to collect data in an effort to combat the spread of the coronavirus must be held accountable and provide transparency for their actions.

Use Limitations and Privacy Protections Must Empower Consumers

Despite the work thus far of this Committee and others, the United States still lacks a comprehensive law to protect consumer privacy, and existing frameworks have proven

insufficient even in non-crisis times.[1] In the effort to activate private companies, our government must demand limitations on the use and purpose of data collected by companies. Furthermore, data minimization—the practice of reducing the total amount of data collected, used, and stored—must play a prominent role in any public health collaboration between government entities and private companies. Companies should not only minimize the data they collect and retain, but also justify why they collect that data and how they use it.[2] Consumers must have access to any personal data collected, in addition to retaining other data rights such as the right to correct, delete, and port data.[3]

In addition, by centering civil rights in this privacy discussion, we ensure that measures taken in the pursuit of a healthier populace through personal data collection also include robust safeguards to prevent discrimination.[4] Personal data has historically been collected and used by government and private companies to manipulate criminal justice, housing, financial, and health care outcomes to the disproportionate detriment of already marginalized communities.[5] Protections for personal data and public health should and can go hand in hand. We need to provide strict use limitations for personal data, particularly to protect those who are most vulnerable and empower them with rights to their own data.

Government Should Collect Only the Data Necessary and Effective to Combat Coronavirus

As our country struggles to combat the coronavirus crisis, proposals related to large-scale government collection of data from tech and telecom companies may seem tempting. Certainly, a number of governments around the world have increased surveillance of their citizens to fight the pandemic.[6] In China, the government required citizens to use phone software that classifies each person with a color code — red, yellow or green — indicating risk, thereby dictating their quarantines, and meanwhile sending their personal data to the police. Chinese officials have not explained how the system determines these color codes, rendering citizens powerless to challenge.[7] In Israel, the government has been using its cache of citizens' phone location data – collected by intelligence officials for counterterrorism purposes – to track citizens in an effort to

[1] Claire Park, *How "Notice and Consent" Fails to Protect Our Privacy*, Open Technology Institute (March 23, 2020), https://www.newamerica.org/oti/blog/how-notice-and-consent-fails-to-protect-our-privacy/
[2] Eric Null, Becky Chao,Sharon Bradford Franklin, WC Docket No. FTC-2018-0098 (May 31, 2019), https://newamericadotorg.s3.amazonaws.com/documents/Comments_of_New_Americas_OTI_Consumer_Privacy.pdf
[3] *Id.*
[4] *Id.*
[5] Becky Chao, Eric Null, Brandi Collins-Dexter, Claire Park, *Centering Civil Rights in the Privacy Debate*, Open Technology Institute (August 14, 2019), https://www.newamerica.org/oti/reports/centering-civil-rights-privacy-debate/for-marginalized-communities-the-stakes-are-high
[6] Natasha Singer and Choe Sang-hun, *As Coronavirus Surveillance Escalates, Personal Privacy Plummets*, The New York Times (March 23, 2020), https://www.nytimes.com/2020/03/23/technology/coronavirus-surveillance-tracking-privacy.html
[7] Paul Mozer, Raymond Zhong, and Aaron Krolik, *In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags*, The New York Times (March 1, 2020), https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html

enforce social distancing and warn people about potential coronavirus exposure.[8] Both of these are troubling examples of governments using big data that is not tailored to efficacy and need, without any transparency or privacy protections, and as a means of unfettered social control.

Indeed, *some* data may be useful in helping stop the spread of coronavirus in the United States as well. But, as government officials weigh such proposals and consider what data to collect from companies and use in this battle, they must carefully consider whether such data will be actually useful to public health authorities, and limit any collection to *only* the necessary and efficacious data. There is a significant interest in using data for the purpose of tracking the *disease* and thereby allocating resources, but this type of effort should be distinguished from efforts to use location data to track *specific individuals*.[9]

Aggregated and anonymized location data has potential to assist government officials in the interest of tracking the *disease* and appropriately allocating lifesaving resources—a task the government is currently struggling to do effectively. Aggregate data can provide public health officials with the mapping information that they so desperately need in order to both direct resources to hospitals, and also to enforce social distancing orders. Smart thermometer data, for instance, may be useful in the aggregate to provide COVID-19 heat maps indicating where the virus is spreading next. Likewise, aggregate location data can show trends over time and demonstrate whether too many people are gathering in certain public spaces. Approaches through which the government receives aggregate data from companies of these and other types could therefore be effective in stopping the spread, while also privacy-protective.

However, government officials should not gather location data to track specific individuals; it is important to recognize that current technology generally cannot provide the level of granularity required to reliably inform officials of whether people are complying with social distancing guidance or to conduct epidemiological contact tracing. Cell site location information (CSLI), or the location records generated by mobile carriers based on phones connecting to cell towers, is not precise enough to allow assessments of whether particular people were closer than the CDC-recommended six-foot distance from one another, and while GPS data is more precise, GPS data only works outside, and only when opted into.[10] Wi-Fi network data and Bluetooth data are other options, but are less than ideal because they are not ubiquitous enough to track exposure (and additionally, Wi-Fi network data is unlikely to be

[8] David M. Halbfinger, Isabel Kershner, and Ronan Bergman, *To Track Coronavirus, Israel Moves to Tap Secret Trove of Cellphone Data*, The New York Times (March 18, 2020), https://www.nytimes.com/2020/03/16/world/middleeast/israel-coronavirus-cellphone-tracking.html

[9] Sharon Bradford Franklin, *The Right and Wrong Ways to Use Location Data in the Pandemic*, Slate (April 8, 2020), https://slate.com/technology/2020/04/coronavirus-location-data-heat-maps-privacy.html

[10] Susan Landau, *Location Surveillance to Counter COVID-19: Efficacy Is What Matters*, Lawfare (March 25, 2020), https://www.lawfareblog.com/location-surveillance-counter-covid-19-efficacy-what-matters; Adam Schwartz and Andrew Crocker, *Governments Haven't Shown Location Surveillance Would Help Contain COVID-19*, Electronic Frontier Foundation (March 23, 2020), https://www.eff.org/deeplinks/2020/03/governments-havent-shown-location-surveillance-would-help-contain-covid-19

granular enough as well).[11] Ultimately, there is simply not enough likelihood that a certain phone—and therefore its owner—can be located with six-foot precision at a given time.[12] Accordingly, this privacy-intrusive method of tracking individuals through mobile phone location data would not be effective and is not warranted for either enforcing social distancing guidelines, or disease-related contact tracing.

Even when using aggregate data, however, there must be strong privacy protection measures in place, and assurances that the data is sufficiently representative. First, the aggregation of such data should be done by the relevant companies before being shared with the government, by knowledgeable data scientists who, as discussed further below, are applying protections beyond basic anonymization.[13] Further, any new initiatives to provide data to the government to combat COVID-19 must be limited to the duration of this pandemic. Additionally, approaches relying on phone or other smart device data, even in the aggregate, must take into account that such data is not likely representative of the overall population. Significantly, the population without smartphones is largely made up of lower income people and seniors;[14] some of the populations that are at the highest risk during this crisis.

Use Modern Privacy-Enhancing Technologies

In recent years, it has become clear that traditional data privacy techniques of anonymization and aggregation are not as privacy protecting as had been assumed. There have been numerous cases of re-identifications of information about specific individuals from data that had been thought to have been effectively anonymized. Recent advances have been made in developing privacy-enhancing technologies and techniques that should be implemented as the government looks to big data solutions to address the pandemic. These include employing differential privacy to ensure that data are adequately protected from being able to identify individuals.[15] Differential privacy, a way of injecting precise amounts of statistical noise into results drawn from datasets, protects aggregated data with mathematical guarantees of privacy

[11] Susan Landau, *Location Surveillance to Counter COVID-19: Efficacy Is What Matters*, Lawfare (March 25, 2020), https://www.lawfareblog.com/location-surveillance-counter-covid-19-efficacy-what-matters
[12] Adam Schwartz and Andrew Crocker, *Governments Haven't Shown Location Surveillance Would Help Contain COVID-19*, Electronic Frontier Foundation (March 23, 2020), https://www.eff.org/deeplinks/2020/03/governments-havent-shown-location-surveillance-would-help-contain-covid-19
[13] Chris Sadler, *Protecting Privacy in Data Releases*, Open Technology Institute (February 24, 2020), https://www.newamerica.org/oti/reports/primer-disclosure-limitation/
[14] Monica Anderson and Madhumitha Kumar, *Digital divide persists even as lower-income Americans make gains in tech adoption*, The Pew Research Center (May 7, 2019), https://www.pewresearch.org/fact-tank/2019/05/07/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption/; Monica Anderson and Andrew Perrin, *Technology Use Among Seniors*, The Pew Research Center (May 17, 2017), https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/
[15] Chris Sadler, *Protecting Privacy in Data Releases*, Open Technology Institute (February 24, 2020), https://www.newamerica.org/oti/reports/primer-disclosure-limitation/

within certain bounds. It can greatly mitigate re-identification of individuals' data[16] and is being increasingly adopted by both commercial and government sectors for privacy protection. For example, differential privacy is currently being used by Google for COVID-19 Community Mobility Reports[17] and for the 2020 Census.[18]

Companies Must Be Held Accountable

Though it is laudable that private entities want to share their tools and data with the government and public health experts to help address the COVID-19 crisis, companies must be transparent about what information they are collecting, as well as the method and format in which they share that data with government entities. As discussed above, private-public partnerships in this area can introduce significant benefits, if companies provide aggregate data to allow public health officials to map the disease and where response efforts are needed, while considering the disparate impacts of this pandemic on the poor.[19]

However, efforts to harness big data can also severely threaten privacy and civil liberties, as has been the case with some other public-private partnerships, and therefore require the utmost transparency. For example, Amazon's partnerships with local police departments through its subsidiary Ring, which manufactures a range of smart home security products that incorporate motion-detecting cameras and facial recognition technology, has increased the threat of surveillance against already marginalized and overpoliced communities like undocumented immigrants and Black activists, fomented distrust between neighbors, and threatens our basic freedom to walk in public without being continuously tracked and watched.[20]

There is much opportunity here for companies to aid the effort to slow the pandemic and reduce the number of infections. However, they must be transparent in showing how they are

[16] For further discussion of differential privacy and limiting disclosure, see *LA Department of Transportation Must Address Serious Privacy Threats Posed by Collection of Highly Detailed Scooter and Bike Location Data*, The Open Technology Institute Press Release (April 4, 2019), https://www.newamerica.org/oti/press-releases/la-department-transportation-must-address-serious-privacy-threats-posed-collection-highly-detailed-scooter-and-bike-location-data/

[17] *See how your community is moving around differently due to COVID-19*, Google (Last updated April 2, 2020), https://www.google.com/covid19/mobility/

[18] *Disclosure Avoidance and the 2020 Census*, United States Census Bureau (March 27, 2020), https://www.census.gov/about/policies/privacy/statistical_safeguards/disclosure-avoidance-2020-census.html

[19] Donald G. McNeil Jr., *Restrictions Are Slowing Coronavirus Infections, New Data Suggest*, The New York Times, https://www.nytimes.com/2020/03/30/health/coronavirus-restrictions-fevers.html; Jennifer Valentino-DeVries, Denise Lu, and Gabriel J.X. Dance, *Location Data Says It All: Staying at Home During Coronavirus Is a Luxury*, The New York Times (April 3, 2020), https://www.nytimes.com/interactive/2020/04/03/us/coronavirus-stay-home-rich-poor.html

[20] Jon Schuppe, *Amazon is developing high-tech surveillance tools for an eager customer: America's police*, NBC News (August 8, 2019), https://www.newamerica.org/oti/in-the-news/amazon-developing-high-tech-surveillance-tools-eager-customer-americas-police/; American Civil Liberties Union (ACLU), *Letter from Nationwide Coalition to Amazon CEO Jeff Bezos Regarding Rekognition*, ACLU (June 18, 2018), https://www.aclu.org/letter-nationwide-coalition-amazon-ceo-jeff-bezos-regarding-rekognition; Caroline Haskins, *How Amazon and the Cops Set Up an Elaborate Sting Operation That Accomplished Nothing*, Vice (July 1, 2019), https://www.vice.com/en_us/article/43jmnq/how-amazon-and-the-cops-set-up-elaborate-sting-operation-that-accomplished-nothing

collecting and sharing personal data from individuals, and be held accountable for any practices that may do more harm than good. Particularly because the United States lacks comprehensive data privacy legislation, it is critical that companies provide transparency to the public about how their data is being used. More importantly, this accountability must be in place even after the pandemic has subsided since data collection can have long-term consequences. Companies and the government alike should recognize the sensitive and private nature of the data to be used in these public health efforts, and provide transparency to build trust with their users and policymakers. One method of ensuring accountability is through transparency reporting.[21]

Conclusion

The full extent of consequences from the coronavirus pandemic are of course still unknown; however, while we will need to sacrifice many things, consumer privacy does not need to be one of them. Enlisting big data to take on the challenges arising from this virus can allow innovation to tackle a complex and timely issue. Yet, in the pursuit of improving public health measures, companies collecting massive data sets must limit the nature and extent of data they collect and share with the government, be transparent in their efforts, and ensure that such data will never be used for discriminatory purposes. Accountability will be the key to any public-private partnership. Fortunately, there are tools available to help us address the challenges of this tumultuous time, including modern privacy protecting measures such as differential privacy. For communities already facing the health threats of COVID-19 and hurting economically as a result of this virus, we must employ safeguards to protect privacy and prevent discrimination, and ensure that we do not compound the threats they face.

We are happy to discuss any of the above points further with you or your staff. Thank you again for the opportunity to provide this submission for the hearing record.

Sincerely,

Sharon Bradford Franklin

Sharon Bradford Franklin
Policy Director
New America's Open Technology Institute

[21] Spandana Singh and Kevin Bankston, *The Transparency Reporting Toolkit: Content Takedown Reporting*, The Open Technology Institute (October 2018), https://d1y8sb8igg2f8e.cloudfront.net/documents/The_Transparency_Reporting_Toolkit_Content_Takedown_Reporting_2018-10-24_125414_1.pdf