Susan Landau[1]

Testimony Before the
US Senate
Committee on Judiciary
Subcommittee on Privacy, Technology, and the Law
Hearing on Protecting Americans' Privacy Information from Hostile Foreign Powers

September 14, 2022

[1] Bridge Professor of Cyber Security and Policy, The Fletcher School and School of Engineering, Department of Computer Science, Tufts University, 160 Packard Ave., Medford, MA 02155. susan.landau@tufts.edu. Affiliation for identification purposes only.

Thank you for the opportunity to offer testimony about protecting Americans' private information from hostile foreign powers.

For more than thirty years, my research and scholarship has focused on security and privacy of communications systems, largely on encryption policy and surveillance, but also on privacy risks. My work has often focused on public policy issues; in this vein, I have testified before Congress previously, as well as having served on study committees of the National Academies of Science, Engineering, and Medicine, the Carnegie Endowment for International Peace, and other organizations.

I am currently the Bridge Professor of Cyber Security and Policy at The Fletcher School and the School of Engineering, Department of Computer Science at Tufts University, where I teach and do research in cybersecurity, national security, law, and policy; I also direct our MS degree in Cybersecurity and Public Policy. Much of my work focuses on communications security and privacy. Previous to my time at Tufts University, I held positions as Professor of Cybersecurity Policy at Worcester Polytechnic Institute, Senior Staff Privacy Analyst at Google, and Senior Staff Engineer and Distinguished Engineer at Sun Microsystems. I have also held academic positions at the University of Massachusetts, Amherst and at Wesleyan University. I hold a PhD in applied mathematics from MIT, an MS from Cornell University, and a BA from Princeton University.

In my testimony I will:

- Describe how the transformations of our communications system over the last three decades has led to a remarkable ability to track individuals' behaviors and actions in great detail and the role of the online ad industry in this tracking;
- Discuss who might have access to this private information; and
- Present potential protections of this data, increasing citizens' privacy and the nation's security. I also recommend that the protections in the privacy bill currently being considered, *American Data Privacy and Protection Act,* be augmented by protections for communications metadata and telemetry information. Currently consumers have little to no control over the collection or use of this remarkably revelatory information, which can disclose personal relationships, and provide precise user location and other highly personal information; the lack of protection on this data must be changed for Americans' safety and security.

## The Online Tracking Ecosystem

*Massive Change in the Availability of Personal Data*

Digitization, Internet communications, and mobile devices have led to a situation in which we leave tracks about virtually all our activities as we go about our lives. The first cause of that is digitization, which has made the ability to search records remarkably easy and fast.

In the mid 1990s, Assistant U.S. Attorney Patrick Fitzgerald investigated the 1993 World Trade Center bombing using Call Detail Records—the records of who called whom when—in an attempt to connect actions of various of the suspects. This effort required hundreds of hours; digitization and dropping costs in storage and search means that such searches can now be done in seconds. This is one of many aspects in the increasing ability to track user actions.

The second cause resulted from cellphones. To connect calls, the phones connect with a cell site, and Cell Site Location Information (CSLI) automatically gives telephone service providers information about a user's locations whenever a user's phone is on. Later technologies, including GPS and other methods, provide far more precise location information. While GPS data only provides information about the location of users outdoors, other techniques can locate users to precise locations within buildings—and therefore locate them as to which office they are visiting, for example.[2] The third cause came about from the shift to Internet communications. Metadata from IP communications is richer than that of phone calls, and thus can be even more revelatory. The fourth change arose from the fact that smartphones, which are both computers and phones, provide software and device telemetry, streaming data that provides measurements on device functioning and user activity.

There are many ways to infer private information about a user using telemetry. For example, it is possible using the accelerometer, gyroscope, and magnetometer data to learn a user's location even if the app is not permitted to collect GPS information. Over the last decade and a half, industry has obtained a number of patents using telemetry to determine user location, proximity of other users over a period of time (and thus enable the app to introduce the user to "someone you may know"), and other types of private information. As someone who rode New York City subways as an adolescent, the last thing I'd have wanted is an app that starts introducing me to strangers who happen to ride the same car I did.

Consumers have control over the information they input to a search engine, a mapping application, a photo album. But consumers do not control the communications metadata that accompanies that transaction, that is, the phone number and date and time of call in the case of phone calls, the IP address in the case of Internet communications. Indeed, few consumers even know what information travels in that Internet "packet header" (the addressing and other information that is needed for an Internet communication to reach its destination). This information not only allows the communication, such as email, to reach its recipient, it also allows the delivery of content a consumer requests, such as a webpage or video. This communications metadata can also provide quite personal information about the consumer, but few consumers are in a position to know this. And short of not accessing a web page or using an app, the consumer has no ability to prevent the dissemination of this data—nor the ability to control its use of the data beyond the anticipated purpose of delivering requested content. Consumers are also not in a position to understand how the telemetry information is used; the use of this data quite privacy invasive.

---

[2] In-store Bluetooth can even locate customers within aisles and target ads accordingly ("You've been drinking a lot of Merlot lately; isn't it time you try some Cabernet?"). See, generally, JOSEPH TUROW, THE AISLES HAVE EYES; HOW RETAILERS TRACK YOUR SHOPPING, STRIP YOUR PRIVACY, AND DEFINE YOUR POWER (Yale University Press, 2017).

Let’s examine how the online ad industry works, as that drives this data collection system.

*How the Online Ad Industry Works: A Brief Summary*

In the early days of the web, online advertisers relied on cookies, short text files placed on a user's device that provide long-term records of user activity. Smartphones, which are in a user’s pocket almost all the time, provide quite valuable information for advertisers, who wanted to be able track the users over all their activities and locations.

The online ad industry seeks to accomplish two activities: deliver ads (and get paid for doing so) and target ads to individuals. The latter involves profiling consumers, and that is the aspect of the online ad industry that I focus on here.

The way the online ad industry does this is through the use of a device identifier, a string of 0’s and 1’s that identify the phone. This identifier allows tracking across different apps. When a consumer opens an app, the identifier is sent, along with other data (such as the user’s current IP address, perhaps GPS location if the user allows this) from the user’s phone to the app provider. The app provider combines this with other information it already has about the user, then shares all this with data brokers, companies whose business it is to collect consumer personal information and then sell shares of that data.

Originally the device identifiers were permanent. As a result, data brokers could develop a very detailed history of the phone user’s activities. And although the device identifiers only identified the device—not the user—data brokers could combine device activity with location information; this allowed them to identify the user. Thus, the unique device identifiers enabled data brokers to compile a detailed dossier of a user’s locations and activities.

To counter that level of privacy invasiveness, Apple and then Android began using an ad ID that could be reset. Depending on how often the user reset the new ad ID, the user could prevent some of the online tracking. Apple then went one step further by requiring users to expressly opt-in to ad tracking. Android partially followed suit; users can drop out of ad tracking, but users have to opt out (that is, opting out is not the default on Android devices the way it is on iPhones).

The value of an ad ID is that it simplifies the ability of data brokers to amass a larger and more precise profile of the user—and thus better target the ad. But there are multiple other ways to identify a user if the ad ID is unavailable. I look now at what parts of the system can collect what types of information.

*Who Collects What Information from Consumers—and What They Might Learn from It*

Telephone service providers collect Call Detail Records (CDRs): what number called which number when and for how long. Internet Service Providers (ISPs), which are now far more vertically integrated than a decade ago, can collect web browsing data (and some do), can share real-time location information with third parties (and some do), and can amass large amounts of personal data about users. App providers can learn about user behavior both from the user's

interaction with the app but also from telemetry information. The sensors on mobile devices are intended to increase the device's functionality—orientation of a page, enable mapping applications, and the like—but these devices can also provide information off device that enable tracking a user's location even if they have GPS tracking off, let a company know if a user's battery is low (and thus change how a video is displayed to preserve battery power or perhaps encourage a service to charge more since the user's phone is about to die). Large Internet companies that provide many services (e.g., search, mapping applications, friend connections, etc.) and earn their money from advertising, collect massive amounts of personal data about users.

In 2015, researchers at Stanford showed that just from using Call Detail Records and publicly available information, they could determine someone had a multiple sclerosis relapse, was having cardiac arrhythmia problems, was interested in buying an automatic rifle, planning to start a marijuana-growing venture, or having an abortion.[3] Under U.S. law, CDRs are not released except as required by law or with a customer's approval.[4] Such protections are critical; for example, by studying CDRs and discerning unusual patterns (in particular a very small group of people who called only each other), Hezbollah was able to uncover a Beirut-based CIA agent and his informers.[5]

CDRs also have purposes beyond billing. Service providers use CDRs for fraud detection and to project future service needs.

Smartphones are more than just phones; they are computers from which we can access the Internet. We use smartphones this way when we employ an app, send an IP-text (e.g.,WhatsApp, Google Messages, iMessage), communicate over VoIP, etc.

ISPs are not bound by the same privacy requirements on the use of customer transactional information—who is accessing what site when—as telephone service providers are regarding CDRs. The growing integration of ISPs with other services (video streaming, connected wearables, etc.), means that ISPs are no longer simply companies that deliver data to consumers; they are companies that develop and provide multiple forms of content and services to users. Thus, many ISPs are able not only to track users across their accesses to different websites and from different physical locations; they are also able to develop profiles about these customers as a result of the other services the ISPs offer.[6]

In some cases, ISPs also collect data about users' app usage, enabling them to further profile the users.[7] For example, a person using the Grindr app is more likely to be a member of the LGBQT

---

[3] Jonathan Mayer, Patrick Mutchler, and John C. Mitchell, *Evaluating the privacy properties of telephone metadata*, PNAS 113 (20), 5536.

[4] Federal Communications Commission, *Protecting Your Privacy: Phone and Cable Records*, December 30, 2019, https://www.fcc.gov/consumers/guides/protecting-your-privacy.

[5] Matthew Cole, OPSEC Failure of Spies, Black Hat USA 2013 (December 3, 2013), https://www.youtube.com/watch?v=BwGsr3SzCZc, 22:45-25:10.

[6] Federal Trade Commission, *A Look at What ISPs Know About You: Examining the Privacy Practices of Six Major Internet Service Providers: An FTC Staff Report*, October 21, 2021.

[7] Federal Trade Commission, *A Look at What ISPs Know About You: Examining the Privacy Practices of Six Major Internet Service Providers: An FTC Staff Report*, October 21, 2021, iii.

community; someone using a period-tracking app is more likely to be trying to get pregnant—or trying to avoid doing so.

ISPs are also able—and do—sell profile information to third parties, including data brokers.

Apps collect information about their users. This might be information that users knowingly supply, such as terms to a search engine or locations to a mapping application. Or it might be information that users are unaware of sharing, such as information supplied in the communications packet header or telemetry information. Packet headers include the addressing information that enable ISPs to deliver communications. These packet headers contain more information than the CDRs of old.[8] The IP address of the sender, for example, may enable the app to determine user location even if the app is not permitted to collect GPS data. To be able to properly display content, apps also learn the device and OS manufacturer of the user's mobile device; this information may help in profiling the user.

Apps can collect telemetry information, some of which can be very revealing. Accelerometer and gyroscope sensors are used for device functionality, including display orientation. Magnetometer sensors on a mobile device enable mapping applications. But combining information from these three sensors can yield perhaps surprising results. If an app provider knows a consumer's initial location and has access to telemetry information from a device's accelerometer, gyroscope, and magnetometer sensors, then the app will be able to track a user's location quite precisely.

There are other ways apps can acquire location information. When a device—a laptop, a smartphone—searches for a nearby Wi-Fi, the device discovers a number of network names, including BSSiD, the access point to that network. An app that collects local BSSIDs can use this information to check a public database—of which there are several—to determine a user's location.

Large Internet companies whose business is built on advertising collect massive amounts of information about users. Single sign-on services, such as being logged into Google or Facebook, results in massive capture of a user's actions. Because such companies typically share data across their multiple services, over time they develop extremely detailed profiles of their users including where they live, who they live with, how they spend time, when they deviate from that pattern (and potentially whether this is due to loss of a job, travel, birth of a child, etc.). Some of the information is available through data that consumers knowingly supply, while other data is determined through the types of surmising described above.

This list of the types of information that service providers, ISPs, apps, and large Internet companies can collect is a sampling rather than a comprehensive list, but it begins to describe a space in which the private sector has amassed detailed portraits on the vast majority of Americans. The free digital content and services we use—search, mapping applications, and the like—is built upon an online advertising industry that targets users according to their interests and activities. Big data enables this business.

---

[8] Steven M. Bellovin, Matt Blaze, Susan Landau, & Stephanie Pell, *It's Too Complicated: How the Internet Upends Katz, Smith, and Electronic Surveillance Law,* Harvard Journal of Law and Technology, Vol. 30, No. 1 (2017).

## Who Might Access this Data

These individual data collectors augment their information about users through accessing information from data brokers, companies that collect from the entities above as well as from various public sources, including government records. The result is that many of these entities have developed highly detailed dossiers on consumers: where individuals go, who they go with, what they like to do, even how they are feeling. Such data is a business asset; like any asset, it needs to be protected. How is the data protected, and from whom?

Note that location information is particularly private information: four location data points are often enough to completely identify an individual (at least during periods when people work outside the home). Location information can reveal whether someone is attending a drug treatment program, visiting an abortion clinic, spending nights at a different apartment than their home address, or, has, in the past, been homeless. In short, location information not only identifies an individual, it identifies the actions a person takes. In that sense, location history is highly personal. But through the methods I described above, this information is available to telephone service providers (whenever the phone is on), ISPs (whenever a consumer is using the Internet, which includes whenever a user is employing an app), app providers (at the time the user is employing an app), and many large Internet companies.

As already stated, the telephone service providers are required by law to protect CDRs. There are notable fines for violations of this policy, but that does not mean that call detail records are fully secure. We carry our phones with us; that in itself, creates risks.

Cell site simulators, towers that simulate a cell tower, capture mobile phone metadata for phones in the local area; this includes the phone number, device identifier, and, if the phone is making a call, the number being called. In 2018, the Department of Homeland Security reported that a government investigation observed "anomalous activity that appeared consistent with [cell site simulator] technology within the [National Capital Region], including locations in proximity to potentially sensitive facilities like the White House."[9]

ISPs also maintain records of our connections; those records can provide a more revealing and intimate portrait of our activities and interests. A recent Federal Trade Commission (FTC) study of six major ISPs found that some companies collect a vast amount of personal information, including location information, app usage history, web browsing activity, TV and video streaming information, and, in the case of home IoT devices, information about users' homes, including "dwelling type, security activity and events, lighting type and energy usage, temperature readings, and alarm start and end times."[10] The FTC noted that some ISPs store data for longer than they need in order to perform required services[11] (long-term storage of sensitive

---

[9] Christopher C. Krebs, Senior Official Performing the Duties of the Under Secretary, National Protection and Programs Directorate, Department of Homeland Security, Letter to Senator Ron Wyden, May 22, 2018, https://www.wyden.senate.gov/imo/media/doc/Krebs%20letter%20to%20Wyden%20after%20May%20meeting.pdf.
[10] Federal Trade Commission, *A Look at What ISPs Know About You: Examining the Privacy Practices of Six Major Internet Service Providers: An FTC Staff Report*, October 21, 2021, 17.
[11] Ibid, 18.

data—and this ISP data is sensitive—creates a security risk). The agency also discovered that some ISPs track users across devices; thus, if a teenager were to look up LGBQT+ information on their phone, ads related to this browsing could appear on the family device.

ISPs share aggregated location information with customers, e.g., they might share information of the sort, "35% of visitors to a particular store are Hispanic-Latino with household incomes between $40K-$74.5K."[12] In some circumstances, ISPs also provide precise location information, including for provision of emergency services, fraud prevention, workforce management, and law enforcement. After news reports of precise location information being provided without customer consent to other entities, including car salesmen and bail bondsmen, the Federal Communications Commission held four major wireless carriers liable for a $200 million fine for selling access to customer location information without adequately protecting the information of individual users.[13]

If an ISP wants to show that it has displayed the contractually required number of ads, the ISP may send cookie or advertising identifiers to the advertising company. Some ISPs prohibit the advertising company from re-identifying the data (and thus discovering personal information about the user).[14] But the FTC noted that, "There is a trend in the ISP industry to offer real-time location data about specific subscribers to the ISPs' third-party customers."[15]

Apps are able to obtain much personal information; this can include precise location information. If the app, for example, receives information from data brokers providing locations for IP addresses, the app can learn the user's location. And the IP address sometimes enables an app to do cross-device tracking, that is, identify a laptop using that IP address—or the home computer—as belonging to the same user.

The large Internet companies that rely on advertising collect a vast amount of information about users and are able to develop very detailed user profiles. This information, along with their software, is their biggest asset; losing public trust about securing user private data would severely harm their businesses. Thus, for example, after some highly publicized episodes in the late 2000s and early 2010s, Google made it extremely difficult for employees to access data about individual users, and Google's security practices aimed at preventing outsiders from accessing user data at Google presents an exemplary model.

Data brokers traffic in user data. The scale of data they collect is massive. In 2014, the FTC studied the data broker market; at that time—and the market has grown enormously since then. One of the nine data brokers studied had 3000 data segments—different categories of consumers with similar characteristic (e.g., women who are active in sports and who are between 21 and 45)—on almost all U.S. consumers.[16] These data segments are used to analyze behaviors (e.g.,

---

[12] Ibid., 24.
[13] Press Release, Fed. Commc'ns. Comm'n, FCC Proposes over $200M in Fines for Wireless Location Data Violations (Feb. 28, 2020), https://www.fcc.gov/document/fcc-proposes-over-200m-fines-wireless-location-data-violations.
[14] Federal Trade Commission, *A Look at What ISPs Know About You: Examining the Privacy Practices of Six Major Internet Service Providers: An FTC Staff Report*, October 21, 2021, 26.
[15] Ibid, iii.
[16] Federal Trade Commission, *Data Brokers: A Call for Transparency and Accountability*, May 2014, iv.

which groups of consumers are most likely to return purchases?) and improve marketing. Data sources include government databases, with data coming from property records, professional licenses, voter registration information, motor vehicle records, driving records, and court records; from other publicly available sources, such as social media; and from commercial data sources such as information about transactions, self-reported information from product warranties and registrations.[17] Sometimes the information is used for identity verification and fraud detection, but the major aspect is for aspects related to marketing (direct marketing, online marketing, and marketing analytics).

Harm to consumers from such data aggregation can be severe. Because this harm results from a complex set of technical interactions, how the damage is occurring is usually not readily apparent to the people against whom it happens. They thus lack the means to remedy the situation.

To begin with, despite the massive data collection on the U.S. public, the data broker market exists largely without consumer awareness—and thus without transparency to the public. An example of the type of harms that can arise is discussed in the August 29th FTC complaint against Kochava Inc., a data broker. The violation was that Kochava was tracking and selling precise user locations including to highly sensitive places, such as medical facilities, reproductive health centers, shelters for the homeless and for those suffering from domestic abuse.[18]

Through June of this year, Kochava provided a free sample of data for prospective customers that included seven-days-worth of precise location data from over 61 million unique mobile devices.[19] The company had a form that these prospective customers were to fill out that included company name and intended use, but the FTC determined that "business" was a sufficient use. In other words, anyone could obtain a free sample of the data of 61 million devices.

The situation was even worse than that Kochava's data was not anonymized; it included a Mobile Advertising ID (MAID). As the FTC noted, combining location data with a MAID would enable identifying the device's user. The free sample Kochava provided included a device that had visited a reproductive health clinic and then a single-family home, in other words, Kochava's free data enabled public tracking and identification of a person in a very sensitive situation.

Such a situation is unhealthy and quite dangerous for the American public.

## Protecting Users and Protecting the US

For several decades now the federal government has pursued a policy of privacy self regulation, with ex post facto actions that have done little to actually protect privacy. Recently some U.S.

[17] Ibid., 13-14.
[18] Federal Trade Commission, Plaintiff v. Kochava, Inc., Corporation, defendant, Case No. 2-22-cv-377, Complaint for Permanent Injunction and Other Relief (August 29, 2022), 1-2.
[19] Ibid, 4-5.

states have stepped in with privacy laws. A better solution would be a federal law, for federal action would provide needed uniformity. The current bill in front of Congress, *American Data Privacy and Protection Act*,[20] is a valuable step forward.[21] This present version, however, permits transferring data to third parties with "the affirmative express consent of the individual."[22] As I have briefly noted in my testimony, consumers are not actually in a position to effectively provide informed consent for uses of metadata and telemetry. The bill's solution to this issue provides categories of "sensitive covered data"[23] and gives the FTC rulemaking ability to extend the definition of sensitive covered data to other categories as needed.[24] I suggest that a better solution would be to limit the use of communications metadata and software and device telemetry to the following purposes:

1. delivery and display of content;
2. ensuring the system is working properly (e.g., for debugging purposes);
3. investigating fraud;
4. ensuring security, including device and user identification done for security purposes;
5. modeling to provide for future services;
6. during publicly declared public health emergencies, providing information on the movement of people in aggregate; this latter use for a very limited time only; or
7. conducting a public or peer-reviewed research project that (i) is in the public interest; and (ii) adheres to all relevant laws and regulations governing such research.[25]

Such an addition to the present bill would take a strong bill and make it even stronger.

Let me end by citing Chris Inglis and Harry Kresja, who wrote in February 2022 that, "With greater certainty over the direction of the United States' data security and privacy environment, U.S. firms would also find it easier to work with the data regimes of like-minded partners. Such collaboration would enable deeper interoperability and commercial exchange with countries such as Japan or those in the European Union that have already begun laying the foundations of twenty-first-century data law … The resulting international ties would help constrain the spread

---

[20] Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf (last viewed September 3, 2022).
[21] Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf (last viewed September 3, 2022), § 101 (b).
[22] Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf (last viewed September 3, 2022), § 102 (3).
[23] Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf (last viewed September 3, 2022), § 2 (28) (A).
[24] Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf (last viewed September 3, 2022), § 2 (28) (B).
[25] This last criteria is already in American Data Privacy and Protection Act, Amendment in the Nature of a Substitute to H.R. 8152, https://docs.house.gov/meetings/IF/IF00/20220720/115041/BILLS-117-8152-P000034-Amdt-1.pdf, § 101 (b)(10)(A).

of Beijing and Moscow's surveillance technologies."[26] By making Americans' personal data more private, such data becomes more secure. This, in turn, strengthens national security. That is a win-win for both individuals and society.

[26] Chris Inglis and Harry Kresja, The Cyber Social Contract: How to Rebuild Trust in a Digital World, Foreign Affairs, February. 21, 2022.