U.S. House of Representatives Committee on the Judiciary
Hearing on
"Digital Dragnets: Examining the Government's Access to Your Personal Data"
July 19, 2022

Statement of Sarah Lamdan
Professor of Law

City University of New York School of Law*
2 Court Square
Long Island City, New York

*Affiliation for identification purposes only

To Chairman Nadler, Ranking Member Jordan, and Distinguished Members of the Committee:

Thank you for the opportunity to appear and speak about the government's use of personal data. I am a Professor of Law at the City University of New York School of Law, where I teach classes on administrative law and information law topics including government transparency and data privacy. I hold a J.D. from the University of Kansas and a masters degree in legal information management from Emporia State's School of Library and Information Management. I am a fellow at NYU School of Law's Engelberg Center on Innovation Law & Policy, a co-chair of the Invest in Open Infrastructure Community Oversight Council, and a Senior Fellow at the Scholarly Publishing and Academic Resources Coalition. Before entering academia, I worked as a research analyst for several international law firms and helped build legal research platforms and systems.

I have been researching the role of data analytics companies in data surveillance and other government data programs since 2017, focusing on the companies that comprise the backbone of federal, state, and local governments' modern surveillance infrastructures. My forthcoming book on the topic will be published by Stanford University Press this November. My research leads me to support more oversight in the government's partnerships with private data companies, as the government increasingly relies on voluntary surveillance tools and systems. Today's government data programs lack traditional due process safeguards, skirting the Fourth Amendment's warrant requirements. The are also opaque, skirting notice and public participation requirements included in the Privacy Act of 1974 and prone to mission creep and erroneous results.

In my testimony today, I will first describe how personal data is gathered and used in government surveillance. I will then provide some guidelines from the Fair Information Practice Principles that balance the need for robust national security and public safety regimes with the preservation of civil liberties and personal privacy ideals. I suggest that future legislation consider these Principles.

The Problems Inherent in Passive, Voluntary Surveillance

Surveillance has always been a part of U.S. national security and public safety regimes, but the nature of our surveillance processes has changed dramatically over the years, especially with technological advancements in data collection and the development of data analytics systems. Traditional surveillance was compelled and targeted—traffic stops, particularized warrants and subpoenas, and other searches done on a person-by-person basis. These "hard," compelled surveillance methods are being replaced by "soft," voluntary surveillance systems that are data-driven, like automated license plate readers snapping photos of every car on a particular road or algorithms sifting through people's social media posts to assess their levels of risk.

People call these datafied surveillance systems "voluntary," but most of us don't truly volunteer to be a part of them. We may technically consent to driving on a public road, we might click "I agree" to access an online service, or we may opt to live and work in buildings that require keycard access. But these choices are illusory. We must make them in order to participate in daily life. We trade our privacy for access to goods, services, and public participation. Most Americans don't want their data to be collected but they feel that, nowadays, it is impossible to avoid.[1] Every move we make online, including through the apps on our phones, connected home electronics, and wearable devices, generates data that can be collected, bundled, shared, and sold. Even when companies promise that they will anonymize your data, that data can easily be re-identified.[2]

Our personal data is being collected by companies that license access to robust dossiers to the government and other major decision-making institutions. Those companies partner with "designer" data companies that specialize in biometric and geospatial data products, as well as companies that build predictive policing and risk products. Collectively, these companies' products comprise much of our government's modern policing, surveillance, and personal data systems. Because they are not considered state actors, and because they merely license access to data products instead of selling datasets to government agencies outright, the companies help agencies skirt the due process and public notice requirements that apply to in-house agency data collections.

While commercial data brokering (collecting and sorting data to sell us things) is invasive, governmental and institutional data products and services can have far more serious consequences—they can lead to police intervention, criminal and legal penalties, and other enforcement outcomes. All data brokering raises privacy concerns ripe for legislation, but institutional data brokering, especially to law enforcement agencies, is the most urgently in need of oversight. When data systems pick surveillance targets, there's usually no notice that your data could be used to implicate you in a potential crime, nor is there an opportunity to consent to data surveillance or correct erroneous data.

---

[1] Angela Chen, *Most Americans Think They're Being Constantly Tracked—and That There's Nothing They Can Do*, MIT TECHNOLOGY REV., Nov. 15, 2019, https://www.technologyreview.com/2019/11/15/238341/privacy-pew-research-data-collection-big-tech-facebook-google-apple/.

[2] Natasha Lomas, *Researchers Spotlight the Lie of 'Anonymous' Data*, TECHCRUNCH, Jul. 24, 2019, https://techcrunch.com/2019/07/24/researchers-spotlight-the-lie-of-anonymous-data/ (discussing studies finding that no "anonymized" data is safe from re-identification).

The Burgeoning Data Surveillance Industry

One major difference between the compelled surveillance tools the government used in the past, and the datafied, "voluntary" surveillance systems used today is that the data surveillance tools are not developed and deployed in-house. They are created by companies that sell and license their products to government agencies. Instead of relying on human intelligence tools (conducting stakeouts, pursuing sources, questioning witnesses, and other human-led interactions), agencies are paying third parties to supply data and data analytics products. An entire industry of predictive policing services and personal data providers cater to government agencies that want to track and sort people by running personal data through analytics systems (algorithms, machine learning, and other data-crunching technologies). Government agencies don't just buy access to our personal information, they also pay for predictions about who might commit crimes or pose risks in the future.[3]

When data companies started partnering with government agencies, law professor Chris Hoofnagle called their products *Big Brother's Little Helpers* because they transform intelligence gathering from a suspect-focused search into constant, intrusive surveillance of all of us.[4] Rather than focusing on particular suspects, data policing tools are dragnets, sifting through all of our data to draw up lists of suspects and other surveillance targets.[5] They're sold as "risk" products, because they rank us in order of how risky we are perceived to be—how likely we are to commit a crime, default on a loan, commit fraud—by running everyone's data through "predictive" algorithms and other data analytics systems. They can notify law enforcement the moment our data changes (if we get a traffic ticket, move to a new location, associate with certain people), flagging changes in our "risk" levels.

Modern data products are far more invasive and fast-moving than traditional, human intelligence-based surveillance.[6] With LexisNexis's Lumen, a police officer can snap a picture of someone on the street with their phone and run the photo through an app that compares that picture against databases full of mug shots collected by law enforcement agencies.[7] Companies like Palantir and PredPol generate "heat lists" with the click of a button. The lists rank who is

---

[3] These companies don't sell data to the government outright, instead, they sell "data as a service" or "software as a service," which is a service model where data or software (or both) are licensed to customers, but not sold outright. The company maintains control of the intellectual property and provides access to end users on their proprietary platform, through an app, or some other streaming or limited access point.

[4] Chris Hoofnagle, *Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT'L L. 595 (2003).

[5] Even data searches that use warrants are creating surveillance dragnets. Instead of focusing on one person, geofence and reverse keyword warrants allow law enforcement to compel companies to turn over IDs for everyone who used a digital device or searched for certain terms in a particular location. *See* Johana Bhuiyan, *The New Warrant: How U.S. Police Mine Google For Your Location and Search History*, GUARDIAN, Sept. 16, 2021, https://www.theguardian.com/us-news/2021/sep/16/geofence-warrants-reverse-search-warrants-police-google.

[6] *See, e.g.*, Sarah Brayne, *The Emergence of Big Data Policing*, U. of TEX. AUSTIN POPULATION RES. CTR. (Aug. 2017), https://repositories.lib.utexas.edu/bitstream/handle/2152/62430/prc-brief-2-11-brayne-policing.pdf; ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT (2017).

[7] *See, e.g.,* Elise Schmelzer, *How Colorado Law Enforcement Quietly Expanded Its Use of Facial Recognition*, DENVER POST, Sept. 27, 2020, https://www.denverpost.com/2020/09/27/facial-recognition-colorado-police/; JPrivate, *Cops Use Lexis Nexis Facial Recognition to Identify Your Family and Friends*, TENTH AMEND. CTR. BLOG, May 20, 2019, https://blog.tenthamendmentcenter.com/2019/05/cops-use-lexis-nexis-facial-recognition-to-identify-your-family-and-friends/.

most likely to commit a crime based on their social media histories, geography, and even the weather.[8] These data tools are usually not subject to warrant and other due process requirements, even though predictive policing products and data surveillance can have the same outcomes for their subjects as other types of searches and seizures—they are considered "programmatic" and "suspicionless" police methodologies, not searches or seizures conducted by state actors but passive data-sorting and assessment done in non-governmental data systems.[9]

There are several types of companies that act in concert to build surveillance infrastructure for law enforcement, intelligence, and other government agencies:

- **Data brokers** like LexisNexis and Thomson Reuters provide huge data dossiers containing billions of datapoints that fuel the predictive data analytics systems and partner with the designer data firms to make their products more robust.[10]
- **Data analytics systems** like Palantir,[11] PredPol,[12] and CopLink[13] predict whether someone will commit crime, default on loan, etc.
- **"Designer data" companies** specialize in specific types of data, especially biometric data including Clearview AI[14] "faceprints" and DNA, or geospatial and geolocation data including Vigilant[15] license plate readers.

These companies work in harmony to make personal data-based government programs more invasive. GPS data alone doesn't do much more than mark where someone has been, but when you combine GPS data with someone's social media posts, home address, and marriage and criminal records, it's far more revealing.[16] Similarly, without data, predictive policing products are empty algorithms with no data to crunch. That's why DNA companies,[17] license plate reader companies,[18] and predictive policing software companies like Palantir and CopLink[19] partner

[8] Ali Winston, *Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technologies*, Feb. 27, 2018, VERGE, https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd.
[9] Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 917 (2021).
[10] Both LexisNexis and Thomson Reuters's data brokering services are part of giant information companies. LexisNexis is part of RELX, the same company that sells Elsevier academic information products and analytics, the Lexis legal research service, and other data and information products. Thomson Reuters also provides the Westlaw legal research platform, owns Reuters news agency, and sells other data and information products.
[11] *Gotham*, PALANTIR, https://www.palantir.com/platforms/gotham/ (last visited Feb. 7, 2022).
[12] PREDPOL, https://www.predpol.com/ (last visited July 10, 2022).
[13] *Advanced Crime Analytics Platform*, COPLINK, https://forensiclogic.com/coplink/.
[14] CLEARVIEW AI, https://www.clearview.ai/ (last visited Feb. 7, 2022) [https://perma.cc/3BHZ-ACNG].
[15] *Vigilant PlateSearch License Plate Recognition Software*, MOTOROLA SOLUTIONS, https://www.motorolasolutions.com/en_us/video-security-access-control/license-plate-recognition-camera-systems/vigilant-platesearch-lpr-analytics-software.html.
[16] IAN GOLDBERG, DAVID WAGNER & ERIC BREWER, PROCEEDINGS, IEEE COMPCON (1997), https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=584660.
[17] *See* Adam Stone, *LexisNexis, Bode Technology Team to Accelerate DNA-based Investigations*, WASH. EXEC. (Dec. 16, 2019), https://washingtonexec.com/2019/12/lexisnexis-bode-technology-team-to-accelerate-dna-based-investigations/#.YACST15Om8W.
[18] *See* Russell Brandom, *Ice Is About to Start Tracking License Plates Across The US*, VERGE (Jan. 26, 2018), https://www.theverge.com/2018/1/26/16932350/ice-immigration-customs-license-plate-recognition-contract-vigilant-solutions.
[19] *See Forensic Logic Launches COPLINK X, The Next-Generation Information Network for Law Enforcement*, PR NEWSWIRE (Jul. 16, 2019), https://www.prnewswire.com/news-releases/forensic-logic-launches-coplink-x-the-next-

with data brokers like Oracle, Experian, LexisNexis, and Thomson Reuters. Data is the lifeblood of our modern policing systems, flowing through all of the algorithms, machine learning, and designer data products, making them work.

Together, these companies have become de facto primary factfinders in many law enforcement investigations.[20] They create mosaics of our lives by piecing together billions of datapoints about us to "form an ever-evolving, 360-degree view" of our lives, revealing where we go, who we know, and what we do each day.[21] Companies like LexisNexis and Thomson Reuters market their services to governments as offering "a holistic, singular view of your citizens" by linking personal data from over 10,000 sources to our personal identifiers, and updating this data in real time. Human intelligence is limited—people can only collect so much information on their fellow humans. But personal data dossiers like the ones LexisNexis and Thomson Reuters sell contain more information than humans could ever gather on their own. These companies have been called "shopping malls for information," offering an array of data types for a broad spectrum of customers.[22] They have data on millions of people, including over two-thirds of U.S. residents. Their dossiers likely know more about you than your family and friends do. Even if you try to opt-out of data collection by avoiding social media, the companies create "shadow profiles" about you based on the data your friends, family, and associates trail behind them when they go online.[23]

The data companies collect and retain more data than government agencies can. Unlike the government, private data companies don't have to limit their data use to certain purposes, nor do they expunge their data as part of mandated records management practices. They save our data indefinitely without deleting it, layering new data on top of old.[24] The companies also use personal identifiers to link data to our dossiers. In 1970's, people worried that universal identifiers, such as Social Security numbers, would be used to create invasive "master files" detailing our personal lives. Computer science experts urged Congress to limit the use of such identifiers to prevent the government from using personal dossiers and across agencies for all sorts of undefined purposes without public notice.[25] Fifty years later, the government licenses

---

generation-information-network-for-law-enforcement-300885164.html. *See also, Thomson Reuters and Palantir Technologies Enter Exclusive Agreement to Create Next-Generation Analytics Platform for Financial Client*, THOMSON REUTERS (Apr. 12, 2010), Internet Archive, https://web.archive.org/web/20120508172659/http://thomsonreuters.com/content/press_room/financial/2010_04_12_palantir_technologies_agreement.

[20] *See* CENTER FOR DEMOCRACY & TECHNOLOGY, LEGAL LOOPHOLES AND DATA FOR DOLLARS (2021), https://cdt.org/wp-content/uploads/2021/12/2021-12-08-Legal-Loopholes-and-Data-for-Dollars-Report-final.pdf.

[21] David E. Pozen, *The Mosaic Theory, National Security, and the Freedom of Information Act*, 115 YALE L. J. 628, 628–79 (2005); McKenzie Funk, *How ICE Picks Its Targets in the Surveillance Age*, N.Y. TIMES (Oct. 2, 2019), https://www.nytimes.com/2019/10/02/magazine/ice-surveillance-deportation.html.

[22] Alice Holbrook, *When LexisNexis Makes a Mistake, You Pay For It*, NEWSWEEK, Sept. 26, 2019, https://www.newsweek.com/2019/10/04/lexisnexis-mistake-data-insurance-costs-1460831.html; *ThreatMetrix For Government*, LEXISNEXIS RISK SOLUTIONS, https://perma.cc/NF7Q-BKWJ (last visited on Jul. 14, 2022).

[23] Andrew Quodling, *Shadow Profiles - Facebook Knows About You, Even if You're Not on Facebook*, CONVERSATION (Apr. 13, 2018, 2:41 AM), https://theconversation.com/shadow-profiles-facebook-knows-about-you-even-if-youre-not-on-facebook-94804.

[24] Even when companies *do* claim to expunge their data, they could be erasing the raw data but maintaining the analysis they've derived from that data.

[25] Arlen J. Large, *Congress Finishes Work on 'Privacy' Bill But Measure Has a Number of Loopholes*, *in* SOURCE BOOK, 1237 (1976). People were concerned that universal identifiers like social security numbers would become the

access to the kind of dossiers that the experts warned about, and uses them in the ways that 1970s-era data experts feared.[26]

Data companies also make it difficult for the public to correct errors in the data that the government uses. The Privacy Act of 1974 enables people to fix errors in their personal data, and to contest the use of their personal data in government datasets.[27] Private data companies don't provide the same correction and contestation rights. The lack of correction mechanisms is especially problematic because the companies also struggle to verify their data products. They receive a glut of data in real time from thousands of disparate sources which they cannot effectively vet without great financial and time expenditures.[28] Instead of preemptively vetting their data collections, the companies attach disclaimers to their datasets.[29] The companies place the onus of correcting errors on consumers. If people want to fix their data, they must contact the downstream data providers and request a fix. But, since the data companies don't list their data sources anywhere, correcting errors is described as a task "few people would have the time or patience to embark upon."[30] Because it is difficult, if not impossible, to correct data in these systems, erroneous personal data becomes part of the decision-making process.

In addition to data errors, data analytics systems themselves are notoriously inaccurate.[31] Predictive analytics systems are limited by the assumptions and practices of the humans that create them.[32] One information system expert calls predictive algorithms "as mythical as the crystal ball."[33] Data analytics system errors are hard to catch. Just as we can't easily see inside a car engine without taking it apart, it's nearly impossible to assess how the companies' systems work from the outside.

---

basis for "master files" where the government gathers our data into massive dossiers composed of merged, unrelated files that would be used across agencies and in various data analytics schemes and used to match people based on various data points. *See* HEW REPORT, *supra* note 3, at 20.

[26] RELX, LexisNexis's parent company, calls their universal identifiers "LexIDs," and the company uses its "linking technology" to enrich our LexIDs with data and connections. James Burton, *LexID Data Technology: What is It and What Does it Do?,* LEXISNEXIS RISK SOLUTIONS, https://blogs.lexisnexis.com/insurance-insights/2016/11/lexid-linking-technology-what-is-it-and-what-does-it-do/ (last visited on Jul. 11, 2022).

[27] The Privacy Act of 1974, 5 U.S.C. § 522a.

[28] A 2019 Newsweek article called *When LexisNexis Makes a Mistake, You Pay for It* describes how the company's datasets errors, including switching data between people with similar names, blocks the public from their bank accounts, insurance, and other services they require. Holbrook, *supra* note 23.

[29] For example, Thomson Reuters's disclaimer to consumers states that the company doesn't "warrant the comprehensiveness, completeness, accuracy, or adequacy" of their data. Shea Swauger, "My request finally came in! It's 41 pages long. Here's the cover letter they sent," TWITTER, Dec. 13, 2019, https://perma.cc/24HF-F54X ("The nature of the information and the collections processes self-limit the ability of any aggregator to independently verify and/or validate any of the database contents").

[30] David Lazarus, *Shadowy Data Brokers Make the Most of Their Invisibility Cloak*, LA TIMES, Nov. 19, 2019, https://www.latimes.com/business/story/2019-11-05/column-data-brokers.

[31] *See, e.g.* Jason Koebler, *Detroit Police Chief: Facial Recognition Software Misidentifies 96% of the Time*, VICE, Jun. 29, 2020, https://www.vice.com/en/article/dyzykz/detroit-police-chief-facial-recognition-software-misidentifies-96-of-the-time.

[32] Cathy O'Neil, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016); Safiya Umoja Noble, ALGORITHMS OF OPPRESSION: HOW SEARCH ENGINES REINFORCE RACISM (2018).

[33] Uri Gal, *Predictive Algorithms are No Better at Telling the Future than a Crystal Ball*, CONVERSATION, Feb. 11, 2018, https://theconversation.com/predictive-algorithms-are-no-better-at-telling-the-future-than-a-crystal-ball-91329.

The government outsources its enforcement, intelligence, and other work to these companies and their products despite their known imperfections. LexisNexis claims to have data contracts with 70 percent of local agencies and almost 80 percent of federal agencies;[34] 2,100 police departments and 955 sheriff departments;[35] and the company has a $16.8 million contract to provide data services to the U.S. Immigration and Customs Enforcement.[36] Thomson Reuters similarly supplies data brokering services to federal and local law enforcement agencies, the Department of Defense, the Department of Justice, and intelligence agencies.[37] Even non-surveillance focused agencies like the U.S. Postal Service and the IRS have started working with data companies to "assess threats" and track fraud.[38]

In some situations, data companies have replicated portions of the government's surveillance infrastructure beyond the scope of government oversight. For instance, after 9/11, federal and state officials joined forces to create a network of government-run fusion centers to share information. Civil rights experts decried the massive data-sharing networks, saying that they posed serious risks to our civil liberties. Today, LexisNexis runs its own private, third-party data center where thousands of law enforcement agencies consolidate and share their data.[39] The company's Public Safety Data Exchange compiles federal, state, and local law enforcement data, links it to our personal data dossiers, and makes it available to customers in products with names like "Accurint Virtual Crime Center."[40] This fusion-center-like product is advised by former FBI, secret service, and metropolitan police department employees. LexisNexis's Data Exchange may not be a government surveillance program, but it certainly feels like one with its government customers and ex-law enforcement leadership. The government's fusion centers are subject to oversight and public scrutiny, but the private data centers operate without transparency or government supervision.

---

[34] "*LexisNexis Special Services* Inc. (LNSSI) was founded to help government agencies create actionable intelligence and deliver data-driven decisions." *Industries We Serve*, LEXISNEXIS, https://www.lexisnexisspecialservices.com/who-we-are/industries/ (last visited Nov. 14, 2021) [https://perma.cc/LT87-PZCT].

[35] ACCURINT, https://www.accurint.com/hr.html (last visited Nov. 11, 2021) [https://perma.cc/A2E9-GP3A].

[36] Sam Biddle, *LexisNexis to Provide Giant Database of Personal Information to ICE*, THE INTERCEPT (Apr. 2, 2021), https://theintercept.com/2021/04/02/ice-database-surveillance-lexisnexis/.

[37] Chris Mills Rodrigo, *Majority of Independent Shareholders Vote to Review Thomson Reuters' ICE Contracts*, HILL, Jun. 9, 2021, https://thehill.com/policy/technology/557591-majority-of-independent-shareholders-vote-to-review-thomson-reuters-ice; "Who We Serve," Thomson Reuters Special Services, LLC. https://www.trssllc.com/government-2/ [https://perma.cc/9LHW-J7RJ] (last visited November 14, 2021).

[38] *EPIC v. U.S. Postal Service et. al*, No. 21-2156 (D.D.C. filed Aug. 12, 2021); Lee Fang, *IRS, Department of Homeland Security Contracted Firm That Sells Location Data Harvested From Dating Apps*, INTERCEPT, Feb. 8, 2022, https://theintercept.com/2022/02/18/location-data-tracking-irs-dhs-digital-envoy/.

[39] The Public Safety Data Exchange (PSDEX) is a contributory database of more than 1,300 law enforcement agencies across the U.S. that was created by LexisNexis Risk Solutions. LexisNexis describes its "Accurint Virtual Crime Center" as "linking billions of public records with agency-provided data." The Crime Center links to PSDEX and brings together disconnected data to provide a more comprehensive view of people's identities so that law enforcement agencies can better target investigations, identify patterns, predict upcoming events and deploy resources more efficiently. *Prevent and Solve More Crimes with Data-Driven Insights*, LEXISNEXIS, https://risk.lexisnexis.com/law-enforcement-and-public-safety/information-data-sharing (last visited July 10, 2022).

[40] *Id*.

Without oversight and supervision, it is hard to figure out exactly how the government uses data analytics products. Parsing the thousands of local, state, and federal government data contracts is nearly impossible. This obscurity seems intentional. Data broker experts say that the details around data contracts are "purposefully dense and dull." The companies and institutions involved make "the most interesting stuff the most impenetrable" to prevent the public from discovering just how our personal information is being used by powerful decision-making entities.[41] Sometimes, the intent for obscurity is more plain—LexisNexis includes clauses in government contracts that prohibit the agencies from discussing their partnerships.[42] These clauses flout freedom of information laws.

It should not be this hard to learn more about how the government is using data about our private lives. Some surveillance particulars should be secret—laws exempt disclosure about information that could harm national security or interfere with public safety and ongoing investigations. But the public has a right to know how their data is being used and what the limits of that use are.[43] The public also has a right to know how their tax dollars are being spent.[44] National security and law enforcement exemptions to transparency requirements are supposed to be construed narrowly. They are not meant to be broad prohibitions against explaining how our private data is being used by the government. A lack of transparency leaves the public to sift through ancillary agency records or endure FOIA litigation to wrest the records from agencies' files.

Lack of Oversight and Transparency Subjects the Public to Surveillance That Laws were Meant to Prevent

The dearth of laws that apply to data companies put the industry beyond the scope of the due process protections that are built into our laws and administrative procedures. The companies are also not considered state actors bound by constitutional obligations even though they have become de facto "arms of the government."[45] Some legal experts posit that government agencies do their work through these companies to "buy their way around" due process requirements.[46] Because data brokers are not required to provide notice or obtain

[41] Charlie Warzel, *The Internet's Original Sin: Shoshana Wodinsky Explains Bad Ads*, GALAXY BRAIN, Sept. 23, 2021, https://warzel.substack.com/p/the-internets-original-sin.

[42] A contract between LexisNexis and ICE for data services includes the following clause: "*Customer will not name LN or refer to its use of the LN Services in any press releases, advertisements, promotional or marketing materials, or make any other third-party disclosures regarding LN or Customer's use of the LN Services*." Devin Coldewey, *Records Show ICE Uses LexisNexis to Check Millions, Far More Than Previously Thought*, TECHCRUNCH, Jun. 9, 2022, https://techcrunch.com/2022/06/09/records-show-ice-uses-lexisnexis-to-check-millions-far-more-than-previously-thought/.

[43] The Privacy Act of 1974 guarantees those rights, but in its current form, it does not always apply to the data companies.

[44] Access to information about government contracts is a given under freedom of information laws. The Department of Justice calls government contracts "public contracts" that taxpayers have a right to know about. *FOIA Update: Disclosure of Prices*, DOJ FOIA UPDATE, Vol. II, No. 2 (1981), https://www.justice.gov/oip/blog/foia-update-disclosure-prices.

[45] Hoofnagle, *supra* note 5, at 595.

[46] Gilad Edelman, *Can the Government Buy Its Way Around the Fourth Amendment*?, WIRED (Feb. 11, 2020), https://www.wired.com/story/can-government-buy-way-around-fourth-amendment/.

warrants or subpoenas before they get, search, and share personal data, they can advertise their services as ones that have "no need for a court order."[47]

Without legal safeguards, there is no limit on what kinds of surveillance products agencies license, and how they use them. When the FBI started using ChoicePoint data products[48] the agency's general counsel encouraged employees to "use ChoicePoint to your heart's content."[49] Government employees have taken advantage of this boundless access to data companies' products. In the 1990's the U.S. Marshals ran around twenty thousand data broker dossier searches a month.[50] A recent FOIA request to ICE showed that ICE employees performed at least 1.2 million LexisNexis searches over a 7-month period.[51]

Most agencies don't monitor how their employees are using data services. When the concept of large-scale data collection was first introduced to Congress, senators warned that a glut of personal data "creates a temptation to use it for improper purposes."[52] This warning has proven prescient—in 2013, the Minnesota Police Department found that over half of its eleven-thousand-person police force made "questionable" searches on their data services.[53] Without proper oversight or auditing of those searches, we don't know how our massive data dossiers are being used by these agencies.[54]

Balancing National Security, Public Safety, and Civil Liberties: General Recommendations

There is already a law on the books addressing the kinds of concerns raised by our voluntary surveillance schemes. The Privacy Act of 1974 was meant to prevent "dragnet behavior" in government data practices.[55] It is based on recommendations drafted by information science, privacy, and technology experts called the Fair Information Practice Principles.[56] Several are especially useful in a law enforcement and intelligence programs, including:

---

[47] Hoofnagle, *supra* note 5, at 621 (quoting eBay's director of Law Enforcement and Compliance Department regarding how the company was framing its privacy policy to cater to law enforcement searches).
[48] ChoicePoint was acquired by LexisNexis in 2008, and ChoicePoint's CLEAR products became part of Thomson Reuters.
[49] FBI Office of the General Counsel Routing Slip, September 16, 2001 (Obtained from the FBI via FOIA request, p. 5 of this document: https://epic.org/wp-content/uploads/privacy/choicepoint/cpfbic.pdf).
[50] Hoofnagle, *supra* note 5, at 600.
[51] Coldewey, *supra* note 44.
[52] SUPP. DETAILED RPTS. ON INTELLIGENCE ACTIVITIES & RTS. OF AMERICANS, 94th Cong., 2d sess., 1976, S. Rep. 94-755, p. 778.
[53] Sadie Gurman, *Across US, Police Officers Abuse Confidential Databases,* ASSOC. PRESS, Sept. 28, 2016, https://apnews.com/article/699236946e3140659fff8a2362e16f43.
[54] Coldewey, *supra* note 44.
[55] THE PRIVACY ACT OF 1974, Pub. L. No. 93-579, 88 Stat. 1896 (1974). I discuss the history, goals, and limitations of the Privacy Act of 1974 more thoroughly in my forthcoming symposium essay, Sarah Lamdan, *Revisiting the Privacy Act of 1974 for Big Data Policing*, Geo. L. & Tech. J. (Forthcoming, 2022). Dragnet behavior is discussed in RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS, REPORT OF THE SECRETARY'S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS (1973) at 15 [hereinafter *HEW Report*].
[56] The *HEW Report* describes the Principles. The Center for Democracy and Technology has discussed the principles and made similar recommendations to the Federal Trade Commission. See REFOCUSING THE FTC'S ROLE IN PRIVACY PROTECTION, Nov. 6, 2009, available at https://www.ftc.gov/sites/default/files/documents/public_comments/privacy-roundtables-comment-project-no.p095416-544506-00026/544506-00026.pdf. The Principles have been widely accepted as best practices for both government and commercial data operations. The Federal Trade Commission applies them to address online data

- Guaranteeing that the government will not keep, license, or otherwise obtain/use systems of personal data records whose very existence is secret;
- Ensuring that people can determine what records pertaining to them are collected, maintained, used, or disseminated by an agency;
- Requiring agencies to procure consent before records pertaining to an individual collected for one purpose could be used for other incompatible purposes;
- Affording individuals a right of access to records pertaining to them and to have them corrected if inaccurate; and
- Instructing agencies to collect, license, or use such records only for lawful and authorized purposes and safeguard them appropriately.[57]

According to the Principles, the public should be notified about how and why their data is being collected and used. The notice should set specific purposes and timelines for data programs. The programs should be audited regularly to be sure that they are being implemented for their intended purposes.

Programs that use personal data should be transparent and provide for public comment. People should be able to consent to the collection of personal data, and to see and correct their datasets, even when those datasets are being provided by a third party. There should be processes to ensure that, when the government implements data programs, it is using accurate data and unbiased data analytics systems that properly serve their assigned purposes.

As the government continues to work with data companies to build its surveillance infrastructure, we must balance the need for robust national security and public, and the benefits of quick and easy data services, with the privacy and civil rights of the American public. The Privacy Act, and the principles at its foundation, offer models that help achieve that balance and should be at the core of laws about the government's personal data programs.

---

privacy issues, and they are at the foundations of California's Consumer Privacy Act. *See* Federal Trade Commission, Self-Regulation and Privacy Online: A Report to Congress, Jul. 1, 1999, https://www.ftc.gov/system/files/documents/reports/self-regulation-privacy-onlinea-federal-trade-commission-report-congress/1999self-regulationreport.pdf; Ronald R. Raether, Jr. et al., TROUTMAN PEPPER, *Data Processing Obligations: Virginia Consumer Data Protection Act*, Troutman Pepper, Mar. 25, 2021, https://www.troutman.com/insights/virginia-consumer-data-protection-act-series-data-processing-obligations.html.

[57] *HEW Report*, *supra* note 59, at xx–xxi.