Nate Roberts

PHI 3370 - Summer 2018

Term Paper

Stewardship of a Vanishing Right

Protecting Privacy During Ubiquitous Intrusion

The right to privacy is one that the United States often purports to take seriously, and with good reason. We are, after all, a nation founded in disobedience to a ruling state, an act unthinkable in a world where nothing could be kept secret. For all our nominal defense of privacy, however, the average citizen's daily experience of it has degraded badly over the last few decades, and at a pace that is only accelerating. Advances in computer and network technology, along with substantial improvements to data storage capacities, have made it easier to collect and store more information about more people than at any other time in history. Search engines log date, time, IP address, and query contents of every search they perform; online retailers retain the name, address, credit card number, and complete order history of every customer they get; social media platforms induce us to voluntarily submit our likes, our dislikes, our families, friends, and acquaintances, details about our relationships and jobs and homes. And those are just the tip of the Big Data iceberg.

Data are the precious resource of the 21st century, and entire industries have warped or emerged around their harvesting and use. And as with the early days of other types of resources, we find ourselves in a kind of 'wild-west' free-for-all, with every person free to grab as much of it as they can carry, and sell it to whomever will pay. In the case of every other harvestable resource, a determination has eventually been made that its collection and usage have potential side effects or consequences significant enough to warrant careful scrutiny. The idea of 'data as resource' is still young, but its integration into world economies has been rapid, and the establishment of some form of oversight will only become more difficult as this integration continues. Federal and state governments, leery (perhaps justifiably) of over-regulating, have been hesitant to step in and establish policies to protect their constituents from potential fraud and abuse, but the need to do so has never been more pressing.

I submit that information privacy is a right worth protecting, and that if those collecting that information will not do it, it is incumbent upon our government to undertake that protection, whether by extending the authority of an existing organization like the Federal Trade Commission, or by the creation of a new regulatory body, empowered to oversee data collection as a resource, and to hold companies and individuals accountable for the way they use these data. To that end,

there are several questions that deserve thorough consideration - What is privacy, how and when do we relinquish it, and what makes it worth protecting?

From a legal standpoint, it can be difficult to say exactly what privacy is. We colloquially think of it as one of our fundamental rights in the U.S., but our constitution never mentions it explicitly. Our modern conception of privacy as a right is predicated on the fourth amendment -

> "The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."

It specified the privacy of an individual's physical belongings and documents, but can be read as intending to protect the privacy of thoughts, words, and actions as well. Perhaps the most notable expansion in this direction came from Louis Brandeis and Samuel Warren, who in 1890 published *The Right to Privacy* in the Harvard Law Review, promulgating the notion that the protection intended by the fourth amendment "extends beyond traditional notions of property, and even beyond the 'products and processes of the mind', [...] advanc[ing] the value of privacy per se by arguing for a 'right to be let alone'" (O'Connor & Lange, p.19).

Professors and authors Sara Baase and Timothy Henry offer a simpler definition of privacy in their book, A *Gift of Fire*. They suggest that it consists of three basic elements: "Freedom from intrusion - being left alone; Control of information about oneself; [and] Freedom from surveillance (from being followed, tracked, watched, and eavesdropped upon)" (Baase & Henry, p.52). The collection of personal information from the internet seems to lend itself readily to the second of these, the idea of 'control of information about oneself', but it is just as surely involved in all aspects of privacy. One might rightly feel intruded upon, for instance, when subjected to unexpected "targeted advertisements", ads based on information one had not intended to disclose to the sending agency. Likewise, given that our online data includes, among other things, a list of 'places' we have been (literally, in the case of location-tracking services), it is difficult to argue that collection of this information does not amount to a kind of surveillance.

To be sure, privacy is never absolute in a society that relies on the connection of its citizens. There must necessarily be times when other people know about our comings and goings, or the things we say, and in view of modern technology, there are many cases where we voluntarily give companies information in the name of increased convenience, or for services that would otherwise be impossible; as the Federal Trade Commission itself notes, big data "can guide the development of new products and services, predict the preferences of individuals, help tailor services and opportunities, and guide individualized marketing" (*Big Data*). It is also the case that everyone has a unique relationship to privacy, and a particular amount of it they need to feel comfortable. A 2011 study by Alan Gerber and associates found that approximately 20% of the population at large are what they call 'privacy unconcerned', "not valuing their own privacy and having a difficult time understanding why anyone would care about privacy" (Strahilevitz, p.2026). The remaining 80% of people do not, however, typically want everything shared with everyone, and to that end, it must be possible to establish how and when it is appropriate to collect, use, or distribute information about other people.

One popular set of guidelines for making such determinations has been around for quite a while - "Fair Information Practices/Principles", or FIPs. There are different variations on the specific language, but they include suggestions like "Inform people when you collect information about them, what you collect, and how you use it; collect only the data needed [... and] keep data only as long as needed; Maintain accuracy of data [...]; Protect security of data [...]" (Baase & Henry, p.62). In 2000, the FTC produced a report recommending privacy protection legislation based on such a set of FIPs, focusing on principles of "Notice, Choice, Access, Security, and Enforcement" (*Privacy Online*, p.37-38). The gist of FIPs seems to be that, in addition to what information is collected about them and how it will be used, consumers deserve to be protected from misuse of their information and potential security breaches, and to have some say in how that information is maintained and distributed. Such principles are widely accepted as being in the best interests of individuals by organizations such as the Electronic Frontier Foundation, the Organisation for Economic Co-operation and Development, and the Privacy Rights Clearinghouse.

Europe has had significant successes using these and similar principles to inform its proactive data protection policies and regulations. When the European Union passed its Data Protection Directive in 1995, its member states were instructed to formulate and enact privacy legislation, "requiring, among other things, transparency and proportionality in data processing operations, reasonable consumer access to information stored about them, and prohibitions on transferring data to other countries with 'inadequate' legal protections" (O'Conner & Lange, p.22). The importance of privacy has also been addressed by the European Court of Human Rights, which declares in the eighth article of its charter that "Everyone has the right to respect for his private and family life, his home and his correspondence," (*European Convention on Human Rights*), and by the Court of Justice of the European Union, which has (among other things) established for EU nations the "Right to Be Forgotten" by search engines and databases (O'Connor & Lange, p.23).

Despite Europe's strong example, and however wide the acceptance and endorsement of FIPs among rights advocacy groups and government agencies like the FTC, legal implementation of such policies in the U.S. remains largely piecemeal. As noted by Lior Jacob Strahilevitz in the Harvard Law Review, "American privacy regulations arise reactively, to the extent that they arise at all." (Strahilevitz, p.2041) Some broader laws have been passed covering citizens' privacy in specific areas, such as the Health Insurance Portability and Accountability Act (HIPAA) of 1996, or the Children's Online Privacy Protection Act (COPPA) of 1998. These rarely offer protections outside the scope of a particular field or group of people, however, and are not always effectively enforced. Perhaps with the maintenance of optimum freedom in mind, the U.S. has a deep-seated inclination to take every legal privacy challenge on its own merits, and to wait for each new technology to show its potential negative impact before deciding how to treat it. Consider, for instance, GPS (Global Positioning System) technology, which allows us to identify locations nearly anywhere in the world, and which today forms the backbone of myriad location-based services. Commercial GPS tracking devices began entering the American market in the late 1990s, but it wasn't until a case of flagrant police abuse of such a device reached the U.S. Supreme Court that we established a federal-level basis for regarding such technology as having the potential to violate the fourth amendment. *United*

States v. Jones, in 2012, brought before the court the case of a man who had a GPS tracking device placed on his car by the police, which tracked him for 28 days, assembling a staggering volume of location and movement data. The court unanimously found that Jones' rights had indeed been violated. According to a concurring opinion by Justice Sonia Sotomayor,

> "GPS monitoring - by making available at a relatively low cost such a substantial quantum of intimate information about any person whom the Government, in its unfettered discretion, chooses to track - may 'alter the relationship between citizen and government in a way that is inimical to democratic society. More fundamentally, it may be necessary to reconsider the premise that an individual has no reasonable expectation of privacy in information voluntarily disclosed to third parties. This approach is ill-suited to the digital age." (O'Connor & Lange, p.21).

This last point of Sotomayor's is worth further consideration on its own - the idea that perhaps, in the face of the realities of a world inextricably enmeshed with data-centric technologies, a person's "reasonable expectation of privacy" may need to extend further down the chain of transmission than it might have a century ago. The way individual pieces of information are used in today's Big Data schema, aggregated and collated and linked to other types of information, would likely be nearly, if not entirely, unrecognizable to someone with access to that same information in the not-so-distant past. If you were a regular customer at your local grocery, say, in 1918, you would likely have no expectation that the contents of your weekly shop were private. The clerk would know what you were buying, but outside of informing their decisions on item stock, that information would likely be inconsequential and quickly forgotten. For the average U.S. citizen in 2018, however, their local grocery is most likely part of a national (or international) chain, connected to a computer system that records every purchase made by every customer, storing that data indefinitely. That, on its own, might be unnerving to a 1918 shopper, but it is only the beginning of the story. That store chain will then take all those purchase records and compare them with other customers' records, allowing them to identify patterns and make data-based predictions on a heretofore unimagined scale. They might even share those records with other kinds of companies, enabling even further degrees of analysis.

Consider a somewhat infamous case from a few years ago, when the Target store chain decided to leverage its vast repositories of customer data to attempt to capture the business of

pregnant women before their competition even knew those women were pregnant. They hired a statistician, Andrew Poole, to create algorithms for analyzing purchase records for patterns leading up to the birth of a child. At this, he was remarkably successful, and Target was able to send advertisements and coupons for baby products to women who not only had not informed Target that they were pregnant, but in some cases, had not even informed their own friends or family. Needless to say, this proved somewhat controversial. (Duhigg)

Such "targeted advertising" is illustrative of the animating principle behind Sotomayor's comment - things can be done using data about us that could not be done when existing privacy laws were written. Huge amounts of information are collected from all areas of our lives, and can give analysts in different fields insights they would otherwise never have had. We're constantly called upon to give up personal information when using online services, be it shopping, staying in touch with friends, registering to vote, applying for a job, looking for an apartment, accessing medical information, any of a seemingly endless list of possibilities. It is often the case that the surrender of such information can be considered a fair exchange - after all, companies can use it to do things we want, such as provide no-cost services, or reduce operating expense, or provide more personalized services. To the extent that our information is used in a manner with which we agree when we are asked for it, the whole endeavor fits comfortably into most any ethical framework.

That notion of agreement, though, can present additional challenges with operations of significant scale. As attorney John Pavolotsky observes when considering Fair Information Practices, "Notice and consent are problematic, particularly so for Big Data." (Pavolotsky, p.220) As he notes, most companies and organizations these days have a privacy policy outlining what data they collect and what they use it for, but "many consumers do not read privacy policies, which are often inscrutable to lawyers and non-lawyers alike". (Pavolotsky, p.220) Even when an entire privacy policy is read and understood, it is always possible that the company or organization will come up with a new use for the data they've collected, one that isn't mentioned in the policy. Can we trust that an entity we share information with will never use it in a way with which we disagree? And, critically, what happens if they share those data with another party? The source of the data may

never even know that it has been passed along, much less to whom, or how that third party intends to use it.

So-called "secondary use" of data, the application of collected data to purposes other than those initially specified, is at once both the best thing about modern Big Data collection, and the worst. The unexpected uses for vast stores of information about people, the patterns they can reveal that we might not even have thought to look for, can provide tremendous insights about human behavior, health, financial trends, and so on. For all that good, though, any secondary use of personal information is by definition non-consensual. An individual cannot have agreed to such uses, because they cannot have been aware of them. Does the idea of consent remain meaningful in a situation where thousands, perhaps millions, of people would be called on to reevaluate decisions about use of their information ex post facto? That is to say, is it worth the effort involved to consider the will of people who have already chosen to surrender information? In his *On the Prospects of Collective Informed Consent*, regarding the possibility of obtaining the "collective informed consent" (hereafter CIC) of all those affected when making civic, commercial, and technical privacy decisions, Finnish philosophy professor Jukka Varelius concludes that it is. He identifies three principal criticisms of collective informed consent as a policy tool, and offers for each a refutation of of its standing to dismiss consent concerns.

The first issue raised against the possibility of collective informed consent, Varelius says, is the "Problem of Restricted Participation" - that, when applying individual consent to larger civic or corporate endeavors, "such procedures are too restrictive, as they allow the affected parties to participate only at the last stage of the decision-making process." (Varelius, p.37) In the context of data privacy, we may think of this as a concern that individuals might only get a single choice regarding any particular use of their data - 'in or out'. They had no input into the methods by which these secondary uses were developed, no opportunity to consent to the algorithms to be used or questions to be asked. If an organization seeks individuals' consent only on the tiniest sliver of their enterprise, it may be that those individuals never had a meaningful opportunity to consent to begin with. In answer to this, Varelius observes that any project of a scale that might require CIC is likely

to affect many more people than just its direct beneficiaries. The use of medical data in an examination of cancer trends, for instance, would not be of direct interest to every member of the public, but would require broad collections of data from people both with and without cancer to provide meaningful results. Thus, we may suppose that not inviting participation from everyone at earlier stages of decision-making on such a project would not necessarily negatively impact their interest or ability in deciding whether their data could be a part of that project. Varelius acknowledges the logistical challenge of securing CIC at large scales, supposing that "in practice, it will probably often be more cost-efficient to allow a limited group of persons to formulate opinions to be subjected to collective informed consent," and while this has the potential to lead to some disagreement about whether consent has been given, "when they are appropriately chosen and their work is made transparent, there would not seem to be anything morally problematic" in the use of such a representative group to determine consent. (Varelius, p.38)

The second common objection to CIC is "the identification problem", or the supposition that it would be onerous and impractical to identify each and every affected party when seeking consent (Varelius, p.38). This concern seems, superficially, less applicable to the realm of data privacy, given that all data collected has at least the possibility of a digital record tying it to its origins. Even if we presume, though, that a company cannot identify the sources of all the data it wishes to reuse, or if we suppose we wished to identify only those within a data collection who might be likely to object to such reuse, we are not absolved of the responsibility to consider their desires. "Even if it were, in principle, impossible to identify all of the affected parties," Varelius says, "we could still hold accounting for their interests as an ideal that we should try to achieve to the extent that we can." (Varelius, p.39) Here again, he promulgates the notion of a representative body for groups too large to be practically consulted on an individual basis; one might again object that we could never be completely certain that a representative gave voice to each member of the group, but as Varelius puts it, "If the spokesmen know the parties they are to represent well, do not have conflicting interests, etc., referring to their views can result in the best possible approximation of the interests they are to represent." (Varelius, p.39)

The third potential strike Varelius notes against CIC is "the Veto Power problem," which is that having to ask permission from every individual in a large group can make it impossible to get things done. Considering the example of the group wishing to use medical records to study cancer again, they might feel their progress severely hampered by having to spend time and resources contacting thousands of patients to request further consent for use of their records, only to have some turn them down. This concern represents only one possible interpretation of consent, however, conflating it with control. "Plausibly, the main point of informed consent is to protect individuals' interests. That as such does not imply veto power in all cases in which an individual's interests are at stake." (Varelius, p.40) It is possible to require a good-faith effort on the part of organizations to secure the consent of individuals before using collected data, without necessarily holding Big Data operations hostage to the preferences of individuals. More, as in the case of so many broad endeavors, "if it were then necessary to act against some persons' interests, as it assumedly would be in concrete cases, those persons would be compensated to the extent that their legitimate interests were violated." (Varelius, p.40) It is conceivable today that an individual citizen could bring suit against a company or organization for flagrant misuse of their data, though such a proposition would be costly and arcane; perhaps a hypothetical regulatory body could offer a streamlined process by which people could seek a redress of grievances from such violations of their interests.

At this point, it may be appropriate to address a common question posed by the privacy-unconcerned: What interests are being violated, besides privacy for privacy's sake? Why might individuals object so strenuously to the reuse of seemingly insignificant data? It might seem petty or paranoid to worry about something like a database including how much television a person is watching, or what they most frequently buy at the grocery store, but even relatively banal data can have significant implications outside of their original context, and the types of data being collected are by no means all so frivolous. The problem of decontextualized and recontextualized data in particular is exacerbated by an entity which, I contend, poses a greater threat to data privacy than any organization which gathers data with a purpose in mind, and that is the Data Broker. This

is an individual or corporation whose business is to buy access to databases wholesale, including purchase histories, internet usage histories, legal documents, and more, aggregating them and reselling them to other customers. Consider just two examples which illustrate the risks of incautiously-used data, especially when channeled across unrelated industries by data brokers.

One such example relates to criminal record expungements. They are an important part of the American legal system, providing protection for those convicted of crimes, who have served their sentences in good faith, from the well-documented human prejudice against them based only on their history. Persons with criminal convictions in their past can face substantial hardships when attempting to secure employment or housing, which is why nearly every state in the country offers some process by which a person's official record can have such convictions expunged, sealed, or vacated (Wayne, p.257). Background checks which include a person's criminal record are an important part of processes such as applications for some kinds of job, and have been for many years, but the process of performing such a background check used to be significantly longer and more arduous. It is arguable that having to contact one's local government record-keeper and obtain physical documents may have put constraints on a potential employer's desire to run a background check which simply do not exist in a world where companies can promise complete records over the internet in a matter of minutes. Beyond this increased ease of access, background checks performed using brokered data have a serious problem of reliability, in that they may not reflect the most current state of affairs. "Data brokers maintain proprietary databases, and most are not required to update their records. [...] Because of this lack of regulation in updating their records, expunged convictions are stored and released from these proprietary databases as if they had never been expunged. [...] Data brokers [also] often omit information and reword or misinterpret language from the original court documents." (Wayne, p.259) The release of officially expunged convictions this way essentially nullifies a part of the legal process.

A more broadly applicable example is the determination of health insurance rates. Independent of the conversation over rising insurance and health care costs in this country, we already expect insurers to take multiple factors into account when setting our rates - age, medical

history, general health - but in a world of endless data on tap, insurance companies have begun evaluating some unexpected criteria, too, including many factors which are not protected by federal HIPAA statutes. Data brokers are able to provide insurers with data from the entire spectrum of modern American life, which those insurers are happy to plug into new algorithms attempting to determine a person's current health, and by extension, their level of risk. They may attempt to make assessments of your health based on social media posts, cell phone records, financial records, how safe your neighborhood is, what kind of food you buy, how much time you spend watching television, even whether you own a FitBit. "Patient advocates warn that using unverified, error-prone "lifestyle" data to make medical assumptions could lead insurers to improperly price plans — for instance raising rates based on false information — or discriminate against anyone tagged as high cost." (Allen) Such assumptions are inherently fragile, but as long as insurance companies operate as for-profit entities, they are unlikely to ever pass up an opportunity to improve their margins.

We can safely say, then, that there exist some circumstances in which even a perfectly average citizen with "nothing to hide" would be well-served by robust data privacy regulations, and even more so those already disadvantaged by life circumstances. We have observed that secondary uses of data can be beneficial, but can also be dangerous, and in either case run contrary to the core principle of control over one's own information. We can say, too, that privacy is worth protecting, and that it can indeed be protected - many of the U.S.'s international allies already offer hearty proactive protections for identifying information about their citizens. We know that privacy protections based on consent, while challenging, are not impossible, and are worth striving for. We have seen that the U.S. government has attempted to adapt some of its systems to deal with the rapidly changing technological landscape, but that they simply are not keeping up with the needs of the day. I offer then, by way of conclusion, that more is required - a system of regulation designed with Big Data in mind, with the mandate and authority to hold companies accountable for how they collect and use information. The free-market approach this country has adopted so far has led to a mercenary disregard for the lives attached to individual bits of data, and as a country, we should

take it as a moral obligation to protect one another from the predators and pitfalls of the new information landscape. At the very least, we should protect ourselves from the damaging side-effects of data mining as strenuously as we protect our environment from other resource-harvesting operations. I cannot claim to have developed an expertise in this regard, but I do take solace in having found my own call for regulation echoed by academics, scholars, industry professionals, and even the FTC themselves.

Works Cited

Allen, Marshall. "Health Insurers Are Vacuuming Up Details About You - And It Could Raise Your Rates". *ProPublica*, 17. July, 2018. https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates. Accessed 20. July, 2018.

Baase, Sara & Henry, Timothy. *A Gift of Fire: Social, Legal, and Ethical Issues for Computing Technology*. Pearson Education, Inc., New York, 2018.

Big Data: A Tool for Inclusion or Exclusion? U.S. Federal Trade Commission, January 2016. https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf, accessed 19. July, 2018.

Duhigg, Charles. "How Companies Learn Your Secrets". *New York Times Magazine*, 16. February, 2012. https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html, accessed 21. July, 2018.

European Convention on Human Rights. European Court of Human Rights, Council of Europe. Ratified November 1950, amended June 2010.
Text from https://www.echr.coe.int/Documents/Convention_ENG.pdf, accessed 21. July, 2018.

O'Connor, Nuala & Lange, Alethea. "Privacy in the Digital Age." *Great Decisions*, 2015, pp. 17-28. Article courtesy of JSTOR, http://www.jstor.org/stable/44214790, accessed 17. July, 2018.

Pavolotsky, John. "Privacy in the Age of Big Data." *The Business Lawyer*, vol.69, no.1, November 2013, pp.217-225. Article courtesy of JSTOR, http://www.jstor.org/stable/43665655, accessed 11. July, 2018.

Privacy Online: Fair Information Practices in the Electronic Marketplace. U.S. Federal Trade Commission, May 2000.
https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-fair-information-practices-electronic-marketplace-federal-trade-commission-report/privacy2000.pdf, accessed 19. July, 2018.

Strahilevitz, Lior Jacob. "Towards a Positive Theory of Privacy Law." *Harvard Law Review*, vol.126, no.7, May 2013, pp.2010-2042. Article courtesy of JSTOR, http://www.jstor.org/stable/23415064, accessed 17. July, 2018.

Varelius, Jukka. "On the Prospects of Collective Informed Consent." *Journal of Applied Philosophy*, vol.25, no.1, 2008, pp.35-44. Article courtesy of JSTOR, http://www.jstor.org/stable/24354974, accessed 17. July, 2018.

Wayne, Logan Danielle. "The Data-Broker Threat." *The Journal of Criminal Law and Criminology*, vol.102, no.1, 2012, pp.253-282. Article courtesy of JSTOR, http://www.jstor.org/stable/23145791, accessed 17. July, 2018.