The Market for User Data

Olivier Sylvain
Fordham University School of Law, sylvain@law.fordham.edu
The Market for User Data

Cover Page Footnote
Professor at Fordham University School of Law and Director of the McGannon Center for Communications Research.

This introduction is available in Fordham Intellectual Property, Media and Entertainment Law Journal: https://ir.lawnet.fordham.edu/iplj/vol29/iss4/1
The Market for User Data

Olivier Sylvain*

Policymakers are today far more alert than ever before to the myriad ways in which tech companies collect and distribute consumers’ data with third-party data brokers and advertisers. We can attribute this new awareness to at least two major news stories from the past six or so years. The first came in 2013, when Edward Snowden, the former National Security Agency contractor, leaked highly classified materials that revealed the ways in which United States national security officials, with the indispensable cooperation of U.S. telecommunications companies, systematically monitored telephone conversations and electronic communications of U.S. citizens and foreign nationals.\(^1\) The story triggered a series of rebukes from civil rights groups, consumer advocates, and foreign leaders around the world. It is not clear whether or the extent to which the NSA or other government agencies have terminated those programs since Snowden’s revelation.\(^2\)

The second came in early 2018, when another whistleblower revealed to journalists that researchers to whom Facebook had allowed to collect and study dozens of millions of users’ personal data, in turn, shared those troves of personal data with Cambridge Analytica, a political consultancy firm.\(^3\) Cambridge Analytica had

---

* Professor at Fordham University School of Law and Director of the McGannon Center for Communications Research.


promoted their access to this data to peddle "psychographic targeting" to political campaigns, including that of Donald Trump in 2016. This more recent revelation has exposed Facebook to what will likely be the largest fine imposed by the Federal Trade Commission ("FTC") in history.

These stories were about the extraordinary misuse and abuse of consumer data by powerful tech companies. But they also are large-scale demonstrations of the Private-Sector Ecosystem of User Data—the theme of the fall 2018 symposium to which this volume of the Fordham IPLJ is committed. At a minimum, these recent episodes, along with others, have dramatically raised our collective awareness about the ways in which consumer data has become the lifeblood of the networked information environment.

Two decades ago, scholars and writers wondered whether online tech companies would ever find a sustainable business model. It appears, however, that, even at that time, some savvy entrepreneurs were on to something. DoubleClick, which is now owned by Alphabet, for example, had already developed techniques to track users' web browsing activity across their hundreds of affiliated sites. There, of course, was nothing novel in the idea of an advertising-based business model; advertising has defined the political economy of the media and communications industry at least since the nineteenth century. But, at the turn of the century, it was not evident to anyone but just a relatively few scholars and entrepreneurs in the start-up world that targeted behavioral advertising would have purchase in the networked information economy.

---


Two decades later, companies have been refining the advertising-based business model. They have been developing ever more powerful algorithmic processes for harvesting, trading on, and exploiting personal consumer data for advertisers. These computational techniques have empowered firms to collect extraordinary amounts of consumer data, anticipate consumer preferences based on that data, and microtarget advertising to individual users based on those predictions. It is all a marketer’s dream. And, if click-through rates are to be believed, consumers are sold.

WHEN ALGORITHMS GO AWRY

But sometimes these algorithms make mistakes. While these errors are often innocuous, they are occasionally tone-deaf, as when Virginia Eubanks, who wrote a marvelous book about the ways in which algorithmic decisionmaking processes can be disastrous for the most vulnerable among us, received an advertisement for her own book. Sometimes their mistakes are in very poor taste and offensive, as when Facebook created an advertisement out of a violently misogynistic Instagram post originally sent to a prominent female tech reporter. Or when a social media algorithm distributed

---

11 See @PopTechWorks, Twitter (May 14, 2010, 1:11 PM), https://twitter.com/PopTechWorks/status/112393396654464 [https://perma.cc/A3VH-YQAF].
an advertisement for gay conversion therapy to members of the LGBTQ community.13

Sometimes companies purposefully design their advertisements to target audiences in ways that, while lawful and rational, are unseemly. For example, in 2012, Orbitz, the travel fare aggregator, relied on consumer data to steer Mac users to pricier hotels on the finding that such users spend 30% more on hotels than PC users.14 Price discrimination is not illegal, for the most part, but selective marketing techniques like these consumers are hardly harmless to those who are systematically chosen to pay more than others or, for that matter, to those who are never exposed to fancier lodging. Much more recently, we learned that Netflix tested advertising about movies based on their various audiences’ race and gender by, for example, emphasizing black characters to black audiences even when those characters play minor roles.15

Every now and again, algorithmic microtargeting enables or encourages violations of law. Just this past spring, for example, Facebook agreed to settle a series of lawsuits that alleged that its Ad Manager generated marketing classifications that made it possible for advertisers to discriminate against people on the basis of protected categories like race, gender, and age in violation of civil rights laws.16 The social media company had harvested and analyzed

---


its consumers’ data to generate the unlawful categories. It is
impossible that the vast majority of users wanted data about them to
be used in this way.

Microtargeting techniques present challenges as much as they
provide opportunities. So much of the algorithmic outputs depend
on the data on which their designers “train” them. If the algorithms
do not learn from their masters to be law-abiding, those algorithms
will of course break the law.

The same might be said about other classes of user information,
including biometric data which can, on the one hand, be inputs
through which algorithms might keep us safer and create new
efficiencies (See, for example, fingerprinting on iPhones.) But, in
the wrong hands, biometric data also can enable discrimination
against classes of people. This is to say nothing of the variety of
ways in which the technology is easily susceptible to abuse by
governments and private actors. This is why facial recognition
technology may very well be too difficult to administer in
democracies with longstanding constitutional commitments to
procedural and substantive fairness.

17 See generally AI Now, DISCRIMINATING SYSTEMS: GENDER, RACE, AND POWER IN AI
FGA4].
18 See Elizabeth Joh, Want to See My Genes? Get a Warrant, N.Y. Times (June 11,
Times (June 9, 2019), https://www.nytimes.com/2019/06/09/opinion/facial-recognition-
police-new-york-city.html [https://perma.cc/QM34-Y5E3].
19 See Joy Buolamwini, Artificial Intelligence Has a Problem with Gender and Racial
Bias. Here’s How to Solve It., TIME (Feb. 7, 2019), http://time.com/5520558/artificial-
intelligence-racial-gender-bias/ [https://perma.cc/C9UX-BGDT]; Steve Lohr, Facial
Recognition Is Accurate if You’re a White Guy, N.Y. Times (Feb. 9, 2018),
intelligence.html [https://perma.cc/4GAP-K5KH].
20 Evan Selinger & Woodrow Hartzog, Amazon Needs to Stop Providing Facial
Recognition Tech for the Government, MEDIUM (Jun. 21, 2018), https://medium.com/@story/amazon-needs-to-stop-providing-facial-recognition-tech-for-the-government-795741a0a16a [https://perma.cc/RXL3-DPK3]. See also Kate Conger,
Times (May 15, 2019), https://www.nytimes.com/2019/05/15/technology/facial-
recognition-san-francisco-ban.html [https://perma.cc/SKDW-ECWL].
WHERE AND HOW DOES LAW COME IN?

In the past couple of years, in light of mounting public concern, legislatures and regulators around the world have announced new protections for consumers. To be sure, consumers have had legal recourse since the nineteenth century. But, today, policymakers have sought reforms in recognition of the massive scale and distinctive nature of the private-sector ecosystem of consumer data.

The European Union has been at the forefront of this effort, enacting the comprehensive General Data Protection Regulation ("GDPR"). It has only been a year since that law has been in effect, but we can already identify the important trends. Under the GDPR, companies must give consumers (what the regulation calls "data subjects") access to the data that they have about them, and ensure meaningful user consent to process that information. Consumers also have the right to withdraw this consent "at any time," as well as the right to "rectification" and "erasure." (These are concepts that had been set out in the now defunct EU Data Protection Directive and elaborated in the European Court of Justice's pre-GDPR opinions.) The GDPR provides, moreover, that companies may not share a user's data to third parties for any purposes that are

---

24 GDPR, Ch. 2, Art 7.
25 Id.
26 GDPR, Ch. 16.
27 GDPR, Ch. 17.
“incompatible” with the ones for which that user shared their data to begin with. The GDPR imposes significant fines for violations of its terms. Questions remain about the new law’s extraterritorial scope, but it nevertheless is a significant reform, at least because U.S.-based companies that do business in Europe have been forced to change their consumer data management practices.

Federal lawmakers in the United States have not mustered a national consensus on anything that resembles the GDPR. Instead, here, federal privacy law has been sectoral for decades, with statutes addressing specific actors in delineated legislative fields through, for example, the Federal Educational Rights and Privacy Act, the Health Insurance Portability and Accountability Act, the Financial Credit Reporting Act, and the Video Privacy Protection Act.

For the past several decades, the FTC has relied on its organic statute to justify a wide range of enforcement actions and guidance for industry’s administration of consumer data. It has exercised this authority in fits and starts, largely because Congress substantially curtailed the agency’s ability to promulgate “notice and comment” rules in the area. In spite of the substantial hurdles imposed by Congress in 1975 the agency proposed this past March to fortify existing rules for the protection of “the privacy and security of customer information held by financial institutions.” It has also been in talks with Facebook following its investigation of

---

29 GDPR, Ch. 2, Art. 5.
30 GDPR, Ch. 8, Art. 83.
34 Scholars have argued that Congress should increase the scope of the FTC’s authority. See, e.g., Daniel J. Solove & Woodrow Hartzog, The FTC and the New Common Law of Privacy, 114 COLUM. L. REV. 583 (2014).
the Cambridge Analytica scandal. The size and nature of their sanction against the social media giant will reveal how far the agency is willing to go to vindicate its position as the primary federal protector of consumer data.

State and city governments have stepped up in the absence of a comprehensive federal data protection law. Illinois became a leader among the states when, in 2008, its legislature passed the Biometric Information Privacy Act ("BIPA"). As the title suggests, BIPA imposes strict duties on the collection and storage of consumers' biometric data. It requires, among other things, that companies obtain consent from consumers to collect or disclose such information, destroy that data after a certain period of time, and securely store any information the company holds. California's Consumer Privacy Protection Act ("CCPA"), which becomes effective in 2020, however, is probably the most sweeping state law to date, adopting many of the same protections set out in the GDPR, including the right of consumers to access the data that companies have about them, the right of consumers to have that data deleted, and the right to block the selling of the data. San Francisco, for its part, is among the few local governments that have altogether banned government agencies from employing facial recognition technology. New York is considering a similar ban on the use of facial recognition technology in public schools.

---


Legislators and government regulators have not been the only reformers. The courts have played an important role in delineating the extent to which private companies today may share consumer data with government officials. Consider the Supreme Court’s 2018 opinion in *Carpenter v. United States.*\(^{42}\) There, the Court held that law enforcement officials may only request and obtain subscribers’ historical mobile phone location data from telecommunications companies on a showing of probable cause.\(^{43}\)

Generally, under the Fourth Amendment, law enforcement officials must have probable cause to search “persons, houses, papers, and effects.”\(^{44}\) But, under the third-party doctrine, the Supreme Court has held that this probable cause standard does not apply to law enforcement requests for business records about consumer activity from banks and telecommunications providers.\(^{45}\) The logic for this rule is relatively straightforward: the records that banks and telephone companies generate and collect about their consumers are an indispensable incident of the services they provide. Consumers do not have a reasonable expectation of privacy in those business records, the Court has explained, because they voluntarily give their data to banks and telephone companies in order to enjoy the services provided.\(^{46}\)

In *Carpenter,* the Supreme Court decided not to extend the third-party doctrine to historical mobile phone location data. In the aggregate, it explained, location information reveals an unprecedented amount about subscribers that far exceeds the kinds of information that justified the third-party doctrine.\(^{47}\) The Court did not do away with the doctrine. It only refused to extend it to mobile phone location, suggesting that the rule might be applied to other networked mobile technologies as well.\(^{48}\)

---


\(^{43}\) *Id.* at 2221.

\(^{44}\) U.S. CONST. amend. IV.


\(^{47}\) 138 S. Ct. at 2220.

\(^{48}\) *Id.* The Court considered whether the collection of location information for seven or more days required a warrant. It did not answer whether its decision would extend to
The private-sector ecosystem of user data today

The GDPR, BIPA, the CCPA, and the Carpenter decision are good indications that policymakers and courts today are adapting current laws to meet the challenges posed by today's networked information economy. Longstanding consumer protection norms like due process and "notice and consent," for example, will continue to play a role, although there is growing evidence that the latter in particular is not especially protective of users. The question policymakers will have to answer is: what should count as consent, when so much remains unknown to consumers and regulators? In this vein, scholars have made the case for regulatory conventions like transparency and mandated impact assessments. At least for now, these concerns—transparency and accountability—seem to be mobilizing voters and legislators. We might suppose that, no matter how we come out of this important constitutive moment, policymakers ought to foster trust between consumers and the commercial entities that hold and manage their personal data. But this is hardly an inevitable or necessarily optimal way of conceiving of data protection in light of the seductive commercial incentives to trade on access to consumer data.
Anyway, for consumers and data protection policymakers, the stakes seem especially high today because we are only now beginning to really comprehend the scale and pervasive integration of the market for user data. Techniques for the collection and distribution of user data define practically all of our experiences today—online and offline. It is not really until consumers and policymakers have a far better understanding of this scope that many of us can rest easy. The *Fordham IPLJ* symposium is one step in that direction.