Online Browsing: Can, Should, and May Companies Combine Online and Offline Data to Learn About You?

Michelle Geronimo

Follow this and additional works at: https://repository.uchastings.edu/hastings_science_technology_law_journal

Part of the Science and Technology Law Commons

Recommended Citation
Available at: https://repository.uchastings.edu/hastings_science_technology_law_journal/vol9/iss2/4

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Science and Technology Law Journal by an authorized editor of UC Hastings Scholarship Repository. For more information, please contact wangangela@uchastings.edu.
Online Browsing: Can, Should, and May Companies Combine Online and Offline Data to Learn About You?

by MICHELLE GERONIMO*

Table of Contents

I. Introduction ....................................................................................... 211
II. How Companies Combine Online and Offline Data ........................ 212
   A. Online Data and Programmatic Advertising .......................... 213
   B. Data Brokers and Onboarding ............................................. 215
   C. What Sort of Data Gets Combined ...................................... 217
   D. Methods of Collecting and Combining Data ...................... 218
III. A Business Perspective: Why to Invest in Modern Advertising
    Methods ..................................................................................... 220
    A. Benefits to Companies and Consumers ............................. 220
    B. Privacy Harms to Consider ............................................... 222
IV. Legal Requirements Under U.S. Law ............................................. 224
    A. Current Regulations .......................................................... 224
    B. Proposed Regulations ....................................................... 229
    C. European Regulations ........................................................ 230
V. Conclusion ..................................................................................... 231

I. Introduction

Imagine that you are online and you visit five different clothing websites and then a news website. You might see a clothing advertisement on the news website, even though you are reading about the presidential election. Or, imagine that you are at the mall shopping for a birthday gift for your father, a Star Wars fan. One month after his birthday, you might get

---

* J.D. candidate, University of California, Hastings College of the Law, expected graduation date May 2017; B.S., University of San Francisco, 2013. The author would like to thank Professor Lothar Determann for his guidance.
ads about Star Wars served on your computer when you sign into your Facebook account.

There are many different privacy concerns surrounding this growing trend — often, the methods of collecting consumer information is not visible to the consumer. Moreover, sensitive information such as that regarding health, finances, or family members could potentially be used for unanticipated purposes. Modern advertising technology allows companies to know their consumers very well and create tailored advertisements. This can create benefits for both businesses and consumers, including saving time and money for businesses, and less advertisements and more focus on relevant interests for consumers.

Can, may, and should companies be allowed to combine offline and online data for purposes of targeted advertising and analytics under U.S. privacy law? This article will address this question in three parts. The first section will describe the “can” — how companies combine offline and online data, the types of data combined, and how such data is collected. The second section will address the “should” — benefits to companies and consumers, existing privacy harms, and possible scenarios of harm. Finally, the third section will discuss the “may” — legal requirements that apply under current U.S. law, and any additional protections that may be necessary.

II. How Companies Combine Online and Offline Data

Facebook has at least 1.23 billion monthly active users, 945 million mobile users, and 757 million daily users.¹ Currently, Facebook purchases offline data from at least six different data companies.² These companies collect information about consumers through a variety of sources, including store loyalty cards, mailing lists, public records information, and browser cookies.³ Data such as gender, economic status, purchasing habits, and even home ownership status are also collected, stored, and sold for advertising


purposes.Advertisers and advertising platforms are thus able to glean a clear sense of what consumers like, where they shop, and what they buy.

A. Online Data and Programmatic Advertising

It is useful to first define online behavioral advertising (“OBA”). What distinguishes OBA from other more traditional marketing tactics is the fact that the former presents users with advertisements that are calculated to reflect their interests, based on the various websites that these users visit. In a nutshell, OBA allows companies to match ads to a consumer’s interests, determined over time. Instead of the activities of a specific individual, OBA relies on anonymous and aggregated data to deliver ads to a computer based on the computer’s browser activity.

The primary technology used by the OBA industry is the tracking cookie. Ad providers can track users through the physical storage of a file — such as a cookie — on a consumer’s computer. While tracking cookies are a somewhat controversial tool used by the behavioral advertising industry, standard cookies are relatively innocuous and arguably essential to the modern Internet’s functionality. For example, dictionary.com places 234 tracking files on a user’s computer in a single visit. Not all websites engage in quite this much tracking — Facebook, in comparison, places a mere four tracking files on users’ computers per visit. Once an ad provider places a cookie onto the user’s computer, the ad provider can then recognize at any time when the same user navigates to a webpage wherever the ad provider may operate. By reaccessing the cookie at each new webpage that the user visits, the ad provider is able to build a profile of sorts based on the user’s online behavior. Some of the Internet’s most visited websites allow multiple ad-networks to place tracking cookies on users’ computers, a practice

9. Id.
10. Id.
11. Id.
that can result in as many as 200 separate cookies being placed on a user’s computer in a single visit.12

While some older versions of cookies might have expiration dates after which they no longer function, persistent cookies can remain active until actually deleted by the user.13 Recently, many advertisement providers have been implementing harder to erase tracking technologies, such as flash cookies, tracking beacons, biometric profiling and deep packet inspection.14 Such technologies can track users outside of the controls built into standard web browsers, essentially depriving users of their already limited ability to control their privacy settings.15

While companies now have the closest and clearest look at consumers to date, thanks to cookies collecting consumer data online, this is often merely a general outline lacking detail.16 The latest development in digital advertising is programmatic advertising, which seeks to close this gap. Programmatic marketing is a way to target what types of audience segments advertisers wish to show their advertising to.17 Other benefits include the ability to limit the ads to times of day and frequency and which publishers the ads will show on.18 The programmatic process of media buying, marketing, and advertising is an algorithmic purchase and sale of advertising space in real time, without relying on the human touch, manual insertions, or manual trading.19 Software is used to automate the buying, placement, and optimization of media inventory using a bidding system.20

---

13. Id.
18. Id.
19. Id.
20. Id.
B. Data Brokers and Onboarding

While cookies provide a means to target consumers based on real-time behaviors, they only capture a disparate and limited amount of data: only about 50 data points per cookie, on average.\(^{21}\) In addition, the life of a cookie is between 7 and 14 days, which constrains scale.\(^{22}\) Targeting solely with cookie data gives a limited and fleeting picture of the consumer. Offline data, by contrast, provides a reliable and stable data resource with a longer life span. With coverage on more than 90 percent of U.S. adults, offline data sources aggregate thousands of data points, resulting in massive reach and breadth of data that, when properly analyzed, can identify great prospects that online signals may have missed.\(^{23}\)

Commercial entities’ data processing and profiling of individuals is not a unique phenomenon to cyberspace. However, neither online data collection nor profiling is functionally identical to offline data collection, or more “primitive” targeting techniques like cookie tracking.\(^{24}\) As such, online data processing and behavioral advertising often raise several unique problems. It is becoming increasingly popular for companies to want to take and connect consumer data from customer scenarios, such as in-store transactions or customer-service call interactions, to the web to retain or entice customers to come back and buy again after time has passed without making a purchase.\(^{25}\) Consequently, onboarding can now allow customer data can be matched to email addresses for digitized direct-mail outreach, connected to site-optimization systems to customize web pages based on offline purchases, or dumped into demand-side platforms (“DSP”s) and ad-targeting systems to aim digital ads at customers.\(^{26}\) Today, there are only a handful of companies that operate behind-the-scenes, actually enabling the transfer of data gathered offline to the digital realm. In the industry, this offline-to-online data process is known as “onboarding.”\(^{27}\)


\(^{22}\) Eichelberger, *supra* note 10.

\(^{23}\) Lafond, *supra* note 22.


\(^{26}\) Id.

\(^{27}\) Id.
Twitter and Facebook helped spark interest in the onboarding data trend by partnering with Datalogix, another onboarding firm.28 These social sites aim to convince advertisers that digital ads drive in-store transactions. Datalogix and Twitter find the correlation between tweets and purchases by matching email addresses that have been scrambled to ensure users’ anonymity.29 Twitter users, when registering for their accounts, provide an email address, and DataLogix collects these emails through loyalty programs.30 In like fashion, Facebook has been beta-testing features that will allow brands to match data gathered through shopper loyalty programs to individual Facebook profiles for ad targeting on the platform.31 Retailers offer loyalty cards, and when shoppers register, they are asked for their email or phone numbers — Facebook users sign into the site using one or the other, and a match between two corresponding data points is detected to enable delivery of an ad.32 Loyalty-program members and Facebook users through email addresses and phone numbers are matched anonymously, which allows for the targeting function.33

Marketers have long used consumer profiles, developed from online search behavior, to learn where consumers go online. Despite the continued growth of online sales, consumers continue to spend time and money in brick and mortar stores. In this physical world where connected devices such as smartphones, tablets, Wi-Fi routers, and Bluetooth beacons are just about everywhere, real-time data that reveals the comings and goings of consumers locally can provide powerful insights that marketers can use to optimize the local marketplace.34 The rapid growth of mobile advertising and the Internet of Things (“IoT”) has resulted in a dramatic spike in demand for local data among digital advertisers and marketers.35

29. Id.
30. Id.
32. Id.
33. Id.
35. Id.
Offline first-party data provides valuable information to brands that are seeking to gain a more precise understanding of customer preferences and purchasing habits. Jivox, a leading platform for data-driven advertising and marketing, recently announced a partnership with LiveRamp, a leader in data onboarding and connectivity services, to help brands anonymize and integrate a broad spectrum of offline first- and third-party customer data into their online marketing and advertising, thus enabling significantly increased levels of personalization and campaign performance. Through such partnerships, companies can now safely anonymize and integrate their offline data with online first-party data in addition to a wide variety of other third-party data, contextual information and environmental data — thus gaining a complete 360-degree view of each customer.

C. What Sort of Data Gets Combined

Facebook uses both online and offline data to target Facebook users for specific ads. Such targeted ads have the ability to gather information about users from their offline activity. One way this is accomplished is through Facebook’s “like” feature, where users’ likes become subtle ads. What a user and his or her friends “like” helps in determining the ad content seen by the user’s friends list. In addition, basic user demographics, including getting engaged or taking a vacation, also contributes to targeted ads. Most tech-savvy Facebook users are not unaware of Facebook’s use of profiles to target ads. What many users do not know, however, is how the combination of both online and offline data influences the ads seen online.

Through partnerships with data companies, Facebook uses what its users buy in brick and mortar stores to influence and track the ads they see. To do this, Facebook combines the information they already have with information from data collection companies like Datalogix, Acxiom, Epsilon, and BlueKai. These companies function to collect information about people through sources such as store loyalty cards, mailing lists, public records information, and browser cookies. For example, if you buy socks at

39. Ford, supra note 16.
40. Id.
Target and use your Target card to get a discount, that information is cataloged and saved by a company like Datalogix. Data gathered by such collection companies include race, gender, economic status, and buying habits. This data is then typically sold to advertisers or corporations.

D. Methods of Collecting and Combining Data

It is easy to see how companies and even consumers can benefit from the combination of offline and online data. However, not much was known about the actual process of collecting this data, let alone combining it, until fairly recently. In December 2012, the Federal Trade Commission (“FTC”) initiated a study of data broker practices. To further increase transparency, the Commission’s order requested detailed information regarding data brokers’ practices, including the nature and sources of consumer data they collect; how they use, maintain, and disseminate the data; and the extent to which the data brokers allow consumers to access and correct data about them or to opt out of having their personal information sold or shared.

Some data brokers obtain information by combing through social media sites, where individuals have not set their privacy settings to restrict access to their information, and social media sites have given the data brokers access to such information. According to data brokers, some social media sites restrict third parties from collecting data from their sites in an automated way. For example, Facebook only allows specified search engines to crawl its site, and its Terms of Service bar scraping, or the copying of the information on Facebook’s website, without Facebook’s written permission.

The data brokers in the FTC’s study collect information from sources in numerous ways. First, some data brokers collect publicly available web-

41. Id.
43. Id. at ii.
44. Id. at 13.
45. Id.
46. Automated Data Collection Terms, FACEBOOK, https://www.facebook.com/apps/site_scraping_tox_terms.php?hc_location=ufi (last updated Apr. 15, 2010); see also Massimo Chieruzzi, Facebook just killed the most effective way to create Custom Audiences . . . and it’s good news for you!, ADEPRESSO (Sept. 12, 2014), https://adespresso.com/academy/blog/facebook-scraping-custom-audiences/.
based data through web crawlers, which are programs that capture content across the Internet and transmit it back to the data broker’s servers.47 The data brokers use software to determine which websites to crawl, how often, and what data points to collect from each website.48 Second, some data brokers buy or acquire printed information, such as telephone directories or local government records, and either scan these documents into an electronic format or have data entry professionals manually create an electronic record.49 Third, some data brokers arrange for batch processing of information.50 For example, some data brokers acquire data from their sources through a daily feed.51 Finally, data brokers may arrange for their sources to make available to them an Application Programming Interface (“API”) through which to process the data.52

One common method of combining user data is a system called “hashing.” Facebook and Datalogix match people and their purchases in a Rube Goldbergian system, that strips out personally identifiable information from their databases and the major retailers.53 The system works by creating three separate data sets. First, Datalogix, Facebook, and retailers “hash” their databases: they turn the names, addresses and other personally identifiable data for each person in its logs into long strings of numbers.54 Second, Datalogix compares its hashed data with Facebook’s to find matches.55 Each match indicates a potential test subject — someone on Facebook who is also part of Datalogix’s database.56 Datalogix runs a similar process with retailers’ transaction data. Finally, Datalogix compares

51. Id.
52. Id.
53. Farhad Manjoo, Facebook Followed You to the Supermarket, SLATE, (Mar. 20, 2013), http://www.slate.com/articles/technology/technology/2013/03/facebook_advertisement_studies_their_ads_are_more_like_tv_ads_than_google.single.html
54. Id.
55. Id.
56. Id.
the Facebook data and the retail data. None of the databases will include any personally identifiable data.\textsuperscript{57}

\textbf{III. A Business Perspective: Why to Invest in Modern Advertising Models}

\textbf{A. Benefits to Companies and Consumers}

It is easy to see why onboarding has become a growing trend among companies. U.S. advertising growth through 2018 is said to average 4.7 percent, compounded annually to an estimated $680 billion.\textsuperscript{58} In addition, Internet advertising in the U.S. will grow by 9 percent to $66 billion, with mobile and video Internet advertising being the fastest growing at 22.1 percent by 2018.\textsuperscript{59} The past few years have seen numerous private equity firms buying ad tech companies, with many investors trying to get into private ad tech firms in anticipation of an IPO.\textsuperscript{60} Additionally, the International Advertising Bureau estimated that by 2018, programmatic advertising spending will grow from 28 percent in 2013 to over 80 percent of marketing spending.\textsuperscript{61}

Some of the biggest names in digital advertising and data have already been acting in light of the onboarding trend. Recently, Google purchased Adometry, a company that uses in-store transaction information to measure digital marketing campaigns, and AOL acquired Convertro (Adometry’s competitor) for approximately $101 million soon after.\textsuperscript{62} More recently, data management and services provider Acxiom acquired data middleman LiveRamp for $310 million.\textsuperscript{63} In the midst of these advertising and data giants, the tech world is consumed by the war between Facebook and Google — two huge sites constantly battling one another for users, engineers, and advertising clients. Yet Facebook’s studies suggest that

\textsuperscript{57} Id.


\textsuperscript{59} Id.


\textsuperscript{62} Angin, supra note 2.

\textsuperscript{63} Id.
people react to ads on Facebook in the same way they respond to ads on television.64 Facebook has been selling itself as an ideal venue for brand advertising for several years now, because it has long seemed like a relatively poor environment for direct-response ads.65

The vast majority of the advertising world is structured around demand generation. These ads are not trying to get consumers to take action right away — instead, they attempt to plant a lingering idea in the consumer’s head.66 In the market context, a common purpose of onboarding and offline and online data analytics is to draw inferences about consumers’ likely choices. Companies may decide to adopt big data analytics to better understand consumers, potentially by using data to attribute to an individual the qualities of those who appear statistically similar, that is, those who made similar decisions in similar situations in the past.67 Companies would then be able to use data on its customers’ past purchases, web searches, shopping habits, and prices paid to create a statistical model of consumer purchases at different prices.68 With such a model, a retailer could then compare a prospective consumer’s characteristics or past purchases, web searches, and location information to predict how likely the consumer is to purchase a product at various price points.69

Another data onboarding trend aims to enhance brick and mortar shopping experiences and drive sales through real-time in-store personalization.70 In-store personalization revolves around connecting online and offline shopping experiences to customer data.71 The collection of data from in-store activities harmonizes data between and across channels, providing connected consumers with a seamless shopping experience. Furthermore, data onboarding allows marketers to tailor campaigns to specifically include or exclude customer segments. Such data can establish who may not necessarily need to receive certain advertisements, and exclude

64. Manjoo, supra note 54.
65. Id.
66. Id.
68. Id. at 5.
69. Id.
71. Id.
them from certain ad campaigns, thus preventing “brand noise” and the waste of ad dollars.\textsuperscript{72}

Programmatic media buying allows an owner or brand to tailor a specific message and creative to the right person, at the right time, in the right context — using audience insight from the brand around the kind of audience they want to target.\textsuperscript{73} Doing so delivers far more precision and personalization, resulting in more efficiently targeted campaigns in comparison to traditional digital advertising, which can be much less targeted and instead based on sheer volume. This way, brands pay for more highly effective ads, delivered to the right people at the right time.\textsuperscript{74} And if, as some consumers insist, ads really do not work on them and that they never buy things because of ads, it is theoretically possible that advertising systems would be able to figure this out and work around those consumers’ preferences.

\section*{B. Privacy Harms to Consider}

With strong sales, increased traffic, and customers visiting more frequently, it is not difficult to see the advantages of investing in digital technology over digital advertising.\textsuperscript{75} Though companies offer consumers choices about data collection, these companies may still use big data to draw inferences about consumers who opt to restrict collection of their data.\textsuperscript{76} Information about similarly situated individuals who chose not to share their data can still be inferred by using big data algorithms.\textsuperscript{77} Certainly, this is something most consumers may not be aware of, and seemingly goes against the purpose of the opt-in and opt-out models.

It may be comforting to know that, from an advertiser’s point of view, the flow of information does not actually reveal personal details. Thanks to anonymizing methods such as hashing, advertisers only learn how many potential customers might see an ad. However, a privacy harm that cannot be ignored relates to the use of personal data for purposes other than what

\begin{itemize}
\item \textsuperscript{72} Scott Hoffman, \textit{Why Data Onboarding is the Secret Weapon for Targeted Video Ad Campaigns}, LIVE\textsc{r}\textsc{a}M (Sept. 3, 2014), http://liveramp.com/blog/why-data-onboarding-is-the-secret-weapon-for-targeted-video-ad-campaigns/.
\item \textsuperscript{73} O’\textsc{sullivan}, \textit{supra} note 17.
\item \textsuperscript{74} Id.
\item \textsuperscript{76} Id.
\item \textsuperscript{77} Id.
\end{itemize}
consumers might intend. Concerns have been raised that when big data is used to target ads, particularly for financial products, low-income consumers who may otherwise be eligible for better offers may never receive them. According to the FTC’s studies, the concern about the potential exposure of characteristics that people may view as sensitive has also been raised. For example, one study combined data on Facebook “likes” and limited survey information to determine that researchers could accurately predict a male user’s sexual orientation 88 percent of the time; a user’s ethnic origin 95 percent of time; and whether a user was Christian or Muslim (82 percent), a Democrat or Republican (85 percent), or used alcohol, drugs, or cigarettes (between 65 percent and 75 percent). As mentioned earlier, unlike Facebook’s internal advertising system that uses information already provided to present ads, partnerships with real world data collection agencies enable advertisers to see what users are buying at stores offline. More relevant ads come at the price of privacy and security. With all this data out there, it they live, what they like, and even if they pregnant.

One example of an advertising privacy concern recently caused Facebook to launch a new automated system. Until February 2017, one of Facebook’s advertising features allowed advertisers to exclude specific groups it referred to as “Ethnic Affinities.” This effectively created ads that excluded users based on race, gender, and other sensitive factors. When such an ad is placed to target Facebook users who are searching for houses, and is set to exclude individuals with certain “ethnic affinities,” this becomes a problem. It is illegal “to make, print, or publish, or cause to be made, printed, or published any notice, statement, or advertisement, with respect to the sale or rental of a dwelling that indicates any preference, limitation, or discrimination based on race, color, religion, sex, handicap, familial status,

78. See Data Broker FTC Report, supra note 43.
or national origin.”84 As a result, Facebook has since recently updated its advertising policies to educate, enforce, and prohibit advertisers from discriminating against users “based on personal attributes such as race, ethnicity, color, national origin, religion, age, sex, sexual orientation, gender identity, family status, disability, medical or genetic condition.”85

Despite consumers’ privacy concerns, business analytics have made it so that consumers can expect retailers to know what they want without even telling them or without them even realizing.86 Consumers now know that modern retailers have that capability — and a majority of consumers have become less concerned about their privacy with respect to the data they provide. Thus, to meet consumers’ growing expectations, businesses must take advantage of online and in-store transactions and other interactions.

IV. Legal Requirements Under U.S. Law

The regulatory framework for both online and offline privacy is currently in flux. Although modeled to be technologically neutral and applicable across industries, regulations are strained by a sea of change in terms of innovation and breakthroughs. This is most necessary in the online environment, which was merely in its early stages when the regulatory frameworks were put in place. This led the United States to introduce new regulatory and self-regulatory frameworks applicable to online behavioral tracking.87

A. Current Regulations

The FTC is a strong watchdog in this domain, based on its broad authority to regulate unfair or deceptive acts or practices pursuant to Section 5 of the Federal Trade Commission Act.88 In doing so, the FTC relied on a “notice and choice” model, where companies operating online are required to post detailed

84. 42 U.S.C. § 3604 (1988); see also Title VII of the Civil Rights Act of 1964 (prohibits the “printing or publication of notices or advertisements indicating prohibited preference, limitation, specification or discrimination”), available at https://www.eeoc.gov/laws/statutes/titlevii.cfm.


privacy policies describing their information collection and use practices, thereby enabling users to make informed choices. Failure to adhere to these obligations under a privacy policy could constitute “deceptive acts or practices” actionable by the FTC. The FTC stated in a recent Preliminary Report that “the notice-and-choice model, as implemented, has led to long, incomprehensible privacy policies that consumers typically do not read, let alone understand.” This view is reiterated by the Department of Commerce, which stated: “From the consumer perspective, the current system of notice-and-choice does not appear to provide adequately transparent descriptions of personal data use, which may leave consumers with doubts (or even misunderstandings) about how companies handle personal data and inhibit their exercise of informed choices.” The problem with notice and choice starts with the lack of transparency. Privacy policies are most often long documents drafted in dense legalese, and tend to be read more as liability disclaimers instead of as the protection of user rights. As such, users simply do not read privacy policies, even if they are shortened or relatively interactive.

In the U.S., there is no single, comprehensive federal law that regulates the collection and use of personal data. Instead, the U.S. has a sort of patchwork system of federal and state laws, as well as regulations that can sometimes overlap or contradict one another. In addition, there are many guidelines, developed by governmental agencies and industry groups, that are part of self-regulatory guidelines and frameworks considered to be best practices. The FTC is the primary federal privacy regulator in the U.S. Section 5 of the FTC Act the FTC’s primary enforcement tool: it is a general consumer protection law that prohibits “unfair or deceptive acts or practices in

90. Tene & Polonetsky, supra note 88.
91. FED. TRADE COMM’N, supra note 89 at iii.
93. Data Broker FTC Report, supra, note 43.
94. Id.
95. Tene & Polonetsky, supra note 89.
97. Id.
98. Id.
or affecting commerce.” However, the FTC Act does not actually regulate categories of data. Instead, it prohibits unfair or deceptive acts or practices that affect consumers’ personal information.

There are only a handful of laws that specifically target the practice of electronic marketing, and the relevant laws are specific to the marketing channels in question. Generally, violations of these privacy laws lead to civil penalties. The main exceptions are the laws directed at surveillance activities and computer crimes. Violations of the federal Electronic Communications Privacy Act (“ECPA”) or the Computer Fraud and Abuse Act (“CFAA”) can lead to criminal sanctions and civil liability. Commercial e-mail is regulated at the federal level by the Controlling the Assault of Non-Solicited Pornography and Marketing Act of 2003 ("CAN-SPAM"). There are also state laws regulating commercial email, but these laws are generally preempted by CAN-SPAM.

Generally, U.S. privacy laws apply to all processing of Personally Identifiable Information ("PII"). There are no formal designations of “controllers” and “processors” under U.S. law as there are in the laws of other jurisdictions. There are, however, specific laws that set forth different obligations based on whether an organization would be considered a data owner or a service provider. The most prominent example of this distinction is found in the U.S. state breach notification laws. Pursuant to these laws, it is


101. Id.


106. Id.

generally the case that the owner of the PII is responsible for notifying affected individuals of a breach, whereas a service provider is responsible for informing the data owner that it has suffered a breach affecting the data owner’s data.108 Once a data owner has been notified of a breach by a service provider, the data owner, not the service provider, then must notify affected individuals.109

US privacy laws generally do not limit the retention of PII to certain specified grounds. There are, however, laws that may indirectly affect an organization’s ability to retain PII. For example, organizations that are collecting personal information online from California residents must comply with the California Online Privacy Protection Act.110 Pursuant to this law, and general consumer expectations in the US, the organization must provide a privacy notice detailing the PII the company collects and how it is used.111 If the organization uses the PII in materially different ways than those set forth in the privacy notice without providing notice and obtaining consent for such uses from the relevant consumers, these uses would likely be considered a deceptive trade practice under federal and state unfair competition laws.112

The Network Advertising Initiative (“NAI”) Code broadly defines online behavioral advertising as any process used whereby data are collected across multiple web domains owned or operated by different entities to categorize likely consumer interest segments for use in advertising online.113 The Code applies to third-party OBA as well as to first-party advertising and contextual advertising.114 Under the NAI Code, a member must provide robust notice on its website that is clear and conspicuous and includes six specified types of information: (1) the OBA, multisite advertising and/or ad delivery and reporting activities undertaken by the member company; (2) the types of data that are collected by the member company; (3) how such data will be used by the member company, including transfer, if any, of data to a third party; (4)


109. Id.


111. Id.

112. Id.; see also Tene & Polonetsky, supra note 88.


the types of PII and non-PII that will be merged by the member company, if any, and how any merged data will be used, including or transferred to a third party; (5) an easy to use procedure for exercising choice to opt out or opt in with respect to such data use for OBA; and, (6) the approximate length of time that data used for OBA, multisite advertising and/or ad delivery and reporting will be retained by the member company. 115

Introduced in 2013, the NAI Code of Conduct is a set of self-regulatory principles that require NAI member companies to provide notice and choice with respect to Interest-Based Advertising and Ad Delivery and Reporting activities.116 The NAI code requires a level of consumer choice, with the level of choice being commensurate with the increased privacy implications of the data to be used.117 Opt in consent is required to collect sensitive data — whether it is PII or not.118 PII is defined to include name, address, telephone number, email address, financial account number, government-issued identifier, and any other data used or intended to be used to identify, contact or precisely locate a person.119 Furthermore, the NAI Code does provide some guidance on what is sensitive data and requires opt in consent to collect sensitive data (whether it is PII or not).120 Sensitive data is defined to include social security numbers or other government-issued identifiers, insurance plan numbers, financial account numbers, information that describes the precise real-time geographic location of an individual and precise information about past, present, or potential future health or medical conditions or treatments, including genetic, genomic and family medical history.121

Furthermore, the NAI Code defines sensitive data to include “precise real-time geographic location of an individual” and recognizes that the “precise location of an individual (such as can be ascertained through GPS-enabled devices) may well be of great use to enable highly-personalized targeted advertising, particularly in the mobile marketing range.”122 In regard to the privacy gap concerning whether the creation and use of some profiles

115. Id.
116. Id.
117. Id.
118. Id.
120. Id.
122. Network Advertising Initiative, supra note 114.
for market segmentation purposes may be so sensitive that heightened regulation is needed, the NAI Code prohibits the use of PII or non-PII to create an OBA segment specifically targeting children under the age of 13 without verifiable parental consent.\(^{123}\) It also specifies that OBA segments may only be used for marketing purposes, thus limiting secondary uses of marketing segments.\(^{124}\) However, the NAI Code does not impose limits on the creation or use of other sensitive marketing segments and does not limit assignment of consumers to those segments based on their profiles.\(^{125}\) Nor does the NAI Code expressly restrict members from engaging in unfair or discriminatory profiling.\(^{126}\)

**B. Proposed Regulations**

In 2009, the FTC released a report discussing the potential benefits of behavioral advertising to consumers, including the free online content that advertising generally supports and personalization that many consumers appear to value.\(^{127}\) The FTC’s Behavioral Advertising Principles apply to the tracking of a consumer’s activities online over time, including the consumer’s searches, web pages’ visits, and viewed content to deliver advertising targeted to the individual consumer’s interests.\(^{128}\) However, to keep up with the rapid growth of privacy needs, a number of federal privacy bills were introduced in 2015, including S. 1158 (Consumer Privacy Protection Act), H.R. 2092 (Student Digital Privacy and Parental Rights Act), and S. 668 (Data Broker Accountability and Transparency Act). The Consumer Privacy Protection Act establishes a federal security breach notification law and provides protection for many types of data including social security numbers, financial account information, online usernames and passwords, unique biometric data (including fingerprints), information about a person’s physical and mental health, information about a person’s geo-location, and access to private digital photographs and videos.\(^{129}\)

---

123. Id.
124. Id.
125. Id.
126. Id; see also King, supra note 115.
The Student Digital Privacy and Parental Rights Act would prohibit operators of websites, applications and other online services from selling students’ personal information to third parties and using or disclosing students’ personal information to tailor advertising to them.\(^{130}\) The bill would also give parents access to information held about their children and allow them to correct it, delete information about their children that schools do not need to retain, and to download any material their children have created.\(^{131}\)

The Data Broker Accountability and Transparency Act would require data brokers to establish procedures to ensure the accuracy of the personal information they collect, assemble, or maintain; and any other information that specifically identifies an individual, (unless the information only identifies an individual’s name or address).\(^{132}\) It would also require data brokers to provide individuals a cost-free method to review their personal or identifying information; allow individuals to dispute the accuracy of their personal information with a written request that the data broker make a correction.\(^{133}\)

C. European Regulation

In Europe, consumers are far more concerned with how their personal information is used. As such, there is a legal framework in place that applies to online behavioral tracking, consisting of the European Data Protection Directive, which regulates the collection, processing, storage and transfer of personal data, as well as the European e-Privacy Directive, which regulates data privacy on communication networks.\(^{134}\) The Data Protection Directive sets forth basic principles such as notice, consent, proportionality, purpose limitation, and retention periods — which apply to not only online but also


\(^{134}\) Id.
to offline data collection and use. 135 On the other hand, the e-Privacy Directive protects, among other things, the confidentiality of communications, spam, traffic and location data, and specifically addresses the use of cookies 136.

Even if the default settings of a browser were designed to reject all cookies and if then the user changed the settings to the effect that cookies should be generally accepted, one could not assume the existence of a valid informed consent. 137 Although the modification of the browser settings could be deemed to be an indication of wishes, this indication would neither be made in respect of the individual case. 138

V. Conclusion

From phones and tablets to connected cars, smart homes and connected medical devices, people are giving out information like never before. 139 While European consumers, who are much more concerned with data privacy, have indicated an uneasiness with the advances of offline and online data combination, most U.S. consumers do not find this phenomenon alarming. However, there are those who do find the trend off-putting. Much of this can be attributed to the consumer’s lack of trust, with regard to how companies use their data and what personal data is being shared. Thus, trust is a competitive advantage, but how can companies earn it? First, they must deliver to consumers a sort of value for the data they gather. 140 If a company’s service relies upon high-value data, such as medical records or financial history, it needs to offer high-value services to customers in return. 141 Conversely, lower value data (such as self-reported demographic information), requires less returned value. 142 Furthermore, companies need to give control to their customers. Even if they do not make use of this control, the feeling of ownership will allow customers to feel more comfortable sharing their data. 143

135. Id.
136. Id.
137. Id.
138. Id.
140. Tene & Polonetsky, supra note 89.
141. Id.
142. Id.
143. Id.
Much of the FTC’s findings point to a fundamental lack of transparency about data broker industry practices.\textsuperscript{144} Data brokers acquire a vast array of detailed and specific information about consumers, and in turn analyze it to make inferences about consumers — some of which may be considered quite sensitive — and the information is then shared with clients in a range of industries.\textsuperscript{145} Though there are a number of benefits, much of this activity does take place without consumers’ knowledge. In light of the FTC’s findings, the Commission unanimously recommended that Congress should enact legislation that would enable consumers to learn of the existence and activities of data brokers, in addition to providing consumers with reasonable access to information about them, which held by these entities.\textsuperscript{146} Thus, the challenge for companies is not whether they should use big data; the reality of today’s marketplace is that big data now fuels the creation of innovative products and systems that consumers and companies quickly are coming to rely upon and expect.\textsuperscript{147} Rather, the challenge is how companies can use big data in a way that benefits them and society, while minimizing legal and ethical risks.\textsuperscript{148}

Thus, companies should strive to educate their customers. Very few companies do this today, fearing that customers will stop sharing once they find out how extensively their data is gathered.\textsuperscript{149} However, companies that consider customer data as an asset, and blindly gather all the data they can, will be thwarted by the lack of consumer trust. Consumers feel uncomfortable with brands reflecting their true selves back to them, so brands must learn to keep their consumers’ secrets.\textsuperscript{150} The future of advertising depends on an informed, transparent dialogue with consumers. A world of personalized advertising should be more relevant to people, and lead the way to fewer, but more premium, useful, and interesting ads. If we start an honest dialogue, guard data seriously, and have opt-in services, we could one day access a benefit that works for all.

\textsuperscript{144} Data Broker FTC Report, \textit{supra} note 44, at vii.
\textsuperscript{145} \textit{Id}.
\textsuperscript{146} \textit{Id}.
\textsuperscript{147} Big Data FTC Report, \textit{supra} note 69, at 12.
\textsuperscript{148} \textit{Id}.
\textsuperscript{149} \textit{Id}.