CONSUMERS, BIG DATA, AND ONLINE TRACKING IN THE RETAIL INDUSTRY

A CASE STUDY OF WALMART

THE CENTER FOR MEDIA JUSTICE

colorofchange.org

Sum Of + Us

NOVEMBER 2013

ABOUT US

Center for Media Justice

CenterforMediaJustice.org

Founded in 2008, the Center for Media Justice is a national movement building intermediary to strengthen the communications effectiveness of grassroots racial justice sectors, and sustain a powerful local-to-local movement for media rights and access. Our mission is to create media and cultural conditions that strengthen movements for racial justice, economic equity, and human rights.

ColorOfChange

ColorOfChange.org

ColorOfChange.org exists to strengthen Black America's political voice. Our goal is to empower our members - Black Americans and our allies - to make government more responsive to the concerns of Black Americans and to bring about positive political and social change for everyone.

Sum Of Us

SumOfUs.org

SumOfUs.org is a new world-wide movement for a better global economy.

Here's what we stand for:

- ✔ Governments that answer to citizens – not corporations
- ✔ Fair treatment of workers and the right of every human being to make a living, safely and ethically, for themselves and their family;
- ✔ The right of ordinary consumers to products that are produced and marketed ethically, sustainably and transparently;
- ✔ The right of communities to manage and protect their own environment and natural resources;
- ✔ Business models that put people and the planet first instead of being driven by shortsighted greed.

About Ian Davey and Technolegis:

Ian Davey is the CEO and Principal Consultant of Technolegis. Mr. Davey holds a master's from Princeton and a bachelor's from UVa, both in computer science. Prior to founding Technolegis, he was a Ph.D. candidate at Princeton's Center for Information Technology Policy studying under Ed Felten, and performed research into privacy, security, and Bitcoin economics. He has worked closely with both the patent litigation consulting firm Elysium Digital and the Federal Trade Commission's Bureau of Consumer Protection, and contributed to the blog Freedom to Tinker.

ACKNOWLEDGEMENTS

Many in the privacy-rights community have been generous with their time and have helped us to broaden our understanding of the issues covered in this report. We would particularly like to thank Jeff Chester of the Center for Digital Democracy for his time and support.

CONTENTS

EXECUTIVE SUMMARY

Walmart—the biggest retailer in the world—is tracking you.

This report is an examination of the many ways in which retailers track consumers online, using Walmart as a case study. To our knowledge, this report is the first comprehensive analysis of Walmart's efforts to gather "Big Data" or massive amounts of information about consumers, analyze that information in complex ways, and use the results of that analysis to track consumers on and offline.

This report is also an effort to help American consumers understand what information Walmart is gathering, how they are gathering it, and what we can do about it. We believe that all Americans should be aware of Walmart's online agenda and what it means for our families and communities. People of color and other marginalized communities should pay special attention to these types of predictive uses of massive data, as they magnify the risks for potential discrimination, including physical surveillance.

"Our ability to pull data together is unmatched."

—Walmart U.S. CEO Bill Simon, September 2013

Many retailers, including Walmart, have particularly targeted mobile technologies to track consumers and market products. Recent studies have shown that people of color are more likely than white consumers to use smart phones and are more likely to make purchases on them. Experts have warned of the potential for "data redlining" impacting low-income communities and communities of color, where discrimination occurs through information gathered or predicted through giant data sets.

We believe that consumers have a right to control their personal data and how it is used. The first step in giving consumers this control is helping them understand the myriad ways in which their data is being collected and used by companies like Walmart.

We know that Walmart:

- Could have exhaustive consumer data on more than 145 million Americans – more than 60 percent of U.S. adults. The company refers to having "petabytes" of data on consumers.
- Walmart shares consumers' online data with more than 50 third parties. This information includes every page and product viewed, unique identifiers for users and their devices, system information such as device type and operating system version, and location information. One third party who receives consumer information from Walmart boasts about having data on 80 percent of U.S. email addresses.
- Maintains the ability to track consumer movement in stores using in-store Wi-Fi.
- Is working to connect data from consumers' social media activity to transactions.

Consumers should understand that:

- Retailers like Walmart are using a broad and ever-expanding toolbox to gather consumer data online.
- Walmart's privacy policy gives the company extraordinarily broad latitude in how it gathers, stores, and utilizes information on consumers.
- Walmart has more than 100 lobbyists working for it and has relentlessly lobbied Congress on online privacy and advertising for years.

We are calling on Walmart and other online retailers to:

- ***Be transparent.***
  - Explain, in a clear and accessible fashion, how Walmart and their many third party partners are using the consumer information they collect.
  - Implement common sense limits to the company's ability to profile users, similar to many of those recently adopted by the European Union Parliament.
- ***Give us choices.***
  - Give users the right to have their data deleted and allow consumers to comprehensively "opt-out" of future online tracking. This includes honoring Do Not Track signals.
- ***Be fair.***
  - Explain how the company's use of predictive intelligence shapes marketing and other business practices, and what safeguards are in place to ensure that it does not result in digital redlining or other forms of discrimination.

INTRODUCTION

This report examines the many ways in which consumers' information is gathered online and then used to profile and advertise to them. We also draw on and refer to recent literature about the use of "big data" to support predictive data analysis and how this analysis can lead to potential discrimination. The report is centered on a technical analysis and case study of Walmart, the nation's largest retailer. Walmart is of particular relevance to communities of color as the company has shown significant growth in the proportion of people of color who shop at Walmart[1] and is doubling "multicultural ad spending."[2]

To our knowledge, this report is the first independent analysis of Walmart's efforts online to gather consumer information and to use that information to shape the company's marketing efforts. Although Walmart has badly trailed online behemoth Amazon thus far in e-commerce, the company is investing billions in an effort to catch up. Walmart has been open about a particular focus on "big data" technologies and "real-time predictive intelligence" that allow the company to gather massive amounts of information on American shoppers and analyze it to shape business decisions.[3]

We estimate that Walmart currently has data on more than 145 million Americans and, thanks in part to sustained lobbying by Walmart and other online marketers, they are able to continue to gather more data, in new ways, every day with minimal oversight. We also know that the types of big data analyses being used by Walmart and others to develop profiles of consumers based in part on predictions often operate in ways that are not effectively governed by existing privacy law. In the absence of an effective regulatory framework, Walmart's privacy policy gives the company extremely broad latitude to gather and process information on consumers.

Further, Walmart has repeatedly faced allegations of and been investigated for violating the law, including discrimination against workers, violation of consumer product safety regulations at home and abroad, and disregard of laws against foreign corruption. Given the already weak framework for regulatory oversight of consumer privacy, we have reason to be cautious about trusting Walmart with our data and that of hundreds of millions of Americans.

We believe that all Americans should be wary of Walmart's online agenda and what it means for our families and our society. This report is an effort to help American consumers understand what information Walmart is gathering, how they are gathering it, and what we can do about it. We believe that communities of color have reason to pay special attention as these technologies and regulatory oversight evolve. While we believe this report is a critical first step in examining the ways in which Walmart is compiling and using information on Americans, it is important to acknowledge that there is much that Walmart doesn't fully disclose or that isn't visible to outside observers and, therefore, much that we don't know. This unknown area is expanded by Walmart's use of predictive intelligence to model and predict consumer behavior. Privacy rights advocates, academics, government regulators, and others are still working to understand all the ways in which these types of evolving technologies will impact our lives.

1 http://www.themediaaudit.com/press/archived-newsletters/the-media-audit-fyi/november-2011/mega-retailers-wal-mart-grocery-shoppers-less-affluent-younger
2 http://adage.com/article/hispanic-marketing/walmart-s-tony-rogers-100-growth-multicultural/238051/
3 http://walmartlabs.blogspot.com/2013/06/we-predict-big-data-will-move-much.html

This report begins with an overview of the many ways in which retailers in general and Walmart specifically are gathering massive amounts of information on American consumers. We then highlight some of the ways in which evolving uses of big data present a real threat of increased discrimination and therefore represent a particular risk for communities of color and low-income communities. Next, we highlight some of the specific strategies being used by Walmart to gather information on consumers. The following section provides a brief overview of the relative lack of effective regulation in the United States in the areas of consumer privacy and big data. Lastly, we include the results of a technical analysis of consumer information gathered by Walmart's website and mobile apps. The analysis was conducted by Ian Davey, from the firm Technolegis.

CONSUMERS ARE INCREASINGLY BEING TRACKED ONLINE INDUSTRY-WIDE

A personalized browsing experience is now common online. We see it when Amazon suggests products based on our shopping histories or shares the preferences of other customers who have purchased products we are considering. We see it when an item we were thinking about buying online follows us across the internet as ads on other, seemingly unrelated, websites or even from one device to the next.

Online and in-person, our privacy is increasingly being eroded as companies collect, analyze, and monetize consumers' actions and preferences. A law professor writing in the *New York Times* in November 2012 demonstrated this by clearing his computer and creating two different identities for himself in two different browsers.[4] In one he created a Democratic version of himself and in the other, a Republican version, each supposedly manifested in the sites he browsed and the things he searched. A few days later, he opened the same webpage in each browser and saw two different ads: one for Mitt Romney in one browser and one for a masters in human resources management in the other. All this is thanks to data brokers who assemble profiles of consumers and segment them into various groups which then determine the ads shown.

In writing this report, we have relied on the work of experts in the field of online privacy. In particular, we have drawn from an October 2013 paper by the researchers Kate Crawford and Jason Schultz entitled, "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms." The paper shaped our understanding of what the authors call "predictive privacy harms" and the inability of current regulatory structures to effectively oversee emerging big data analytics practices. The authors emphasize that "not only are these data sets sizeable, they often contain very intimate aspects of individual lives"[5]

Crawford and Schultz highlight in particular the real potential for discrimination to flow from the use of predictive data analysis: "Big data presents substantial privacy concerns – risks of bias or discrimination based on the inappropriate generation of personal data."[6] Later, we highlight some additional risks faced by communities of color and low income communities.

It is becoming ever more difficult to avoid being tracked, and many of the effects of tracking may go undetected. The internet was supposed to put the world at our fingertips, but that just may not be the case anymore. As a *Wall Street Journal* story put it last winter:

> *But the idea of an unbiased, impersonal Internet is fast giving way to an online world that, in reality, is increasingly tailored and targeted. Websites are adopting techniques to glean information about visitors to their sites, in real time, and then deliver different versions of the Web to different people. Prices change, products get swapped out, wording is modified, and there is little way for the typical website user to spot it when it happens.*[7]

4 http://www.nytimes.com/2012/12/02/magazine/who-do-online-advertisers-think-you-are.html
5 Kate Crawford and Jason Schultz, Big Data and Due Process; Toward a Framework to Redress Predictive Privacy Harms, Boston College Law Review, Vol. 55, No. 1, 2014 Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2325784
6 Ibid
7 http://online.wsj.com/news/articles/SB10001424052702304682504579157780178992984?tesla=y

And while some tracking may be harmless and even help to optimize a consumer's experience, many are increasingly wary of having unaccountable, largely unregulated companies like Walmart control so much of their personal data with no clear limits on how this data will be used.

Moving beyond cookies

Cookies, little files that track users' browsing habits and allow advertisers to key in on their target consumers, have been following internet users for years. They are a key tool for the $120 billion a year online advertising industry. However, they don't work well on mobile devices and they cannot identify the same user across multiple devices. To make up for cookies' shortcomings, advertisers are developing other methods, including authenticated tracking, browser fingerprinting, cross-device tracking, and more. According to the *San Francisco Chronicle*,

> *Each method carries a unique set of implications but the common theme is this: These techniques for tracking are generally more persistent, exacting a higher cost for a clean start. Rather than simply deleting cookies, people might have to throw away a device, forsake a social network or delete a rich archive of e-mails.*[8]

In recent months, Google, Facebook, and Microsoft have all announced their own efforts to track users as they navigate the internet. Google and Microsoft each announced that they are working on developing unique identifiers for users that the companies can share with advertisers, while Facebook is taking tracking into its own hands by using its own cookies to help advertisers target their Facebook ads.9 Of course, these companies and others already know who users are through the authentication process of logging in to use their sites and products, making tracking these users even easier.

Walmart proudly tracks and targets customers individually. In a November 2011 blog post about their practices, Walmart's Silicon Valley team wrote:

> *The targeting team within @WalmartLabs ingests just about every clickable action on Walmart.com: what individuals buy online and in stores, trends on Twitter, local weather deviations, and other local external events such as the San Francisco Giants winning the World Series. We capture these events and intelligently tease out meaningful patterns so our millions of Walmart.com customers have a shopping experience that is individually personalized.*[10]

While a "personalized shopper experience" may sound non-threatening, considerable sensitive information can be and has been gleaned or modeled from this information. A 2013 study from researchers at the University of Cambridge found that from Facebook likes "can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender."[11]

8 http://www.sfgate.com/technology/dotcommentary/article/Online-privacy-concerns-growing-4860821.php
9 http://online.wsj.com/news/articles/SB10001424052702304682504579157780178992984?tesla=y
10 http://walmartlabs.blogspot.com/2012/11/targeting-walmartlabs.html
11 Michal Kosinskia, David Stillwell, and Thore Graepel *Private traits and attributes are predictable from digital records of human behavior* Available at: http://www.pnas.org/content/110/15/5802.full

Data brokers offer consumer profiles, whether or not they're accurate

Data brokers help fuel ads online by assembling profiles of consumers and tracking their actions. Last year, the *New York Times* reported on one such data broker, Arkansas-based Acxiom. In an investor presentation, Acxiom highlighted how the company helps engage consumers with example car shopper "Becky," who is "37, married, 2 children, high value, lives in NY." The company knows her:

> *Service history & preferences*
>
> *Preferred media, channel & cadence*
>
> *Transaction & response behavior*
>
> *Garage data*
>
> *Social influence*[12]

Acxiom says it also "score[s] her future behavior" and has taken stock of the fact that she "responded favorably to safety-themed messages, configured a vehicle on our site, but hasn't requested a quote or indicated interest in visiting a dealer." Their response? An individually targeted car advertisement as she surfs the internet.[13]

According to the *Times*, as of last year Acxiom's database contained about 190 million individuals and 126 million US households, in addition to the databases it maintained for about half of the Fortune 100 companies, and Axciom was not even the country's largest data broker. Additionally, Acxiom "worked with the government after the September 2001 terrorist attacks, providing information about 11 of the 19 hijackers."[14]

This doesn't mean the firm's information on consumers is entirely accurate, however. Recently, Acxiom launched a website where individuals can look up some aspects of their profiles. One researcher logged in and found a slew of details—"Acxiom believes I have three children, own my sister's since-sold car, made just 14 purchases in the past two years; and I'm Christian and I'm into motorcycling"—none of which were accurate.[15]

Crawford and Schultz note that predictive modeling can be damaging to consumers, whether or not the information is correct.

> *When functioning well, big data techniques can predict your PII [personally identifiable information]; when not, you can be mislabeled with inaccurate PII. Either way, such processes create a model of possible personal information and associate it with an individual. Harms can result both when the model is accurate and when it is incorrect.*[16]

12 http://www.docstoc.com/docs/65163921/Acxiom-Corporation-Data-Demand-Respect
13 http://www.docstoc.com/docs/65163921/Acxiom-Corporation-Data-Demand-Respect
14 http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?_r=3&pagewanted=all&
15 http://www.npr.org/2013/09/05/219128324/data-marketing-critics-check-out-whats-written-about-them
16 Kate Crawford and Jason Schultz, Big Data and Due Process; Toward a Framework to Redress Predictive Privacy Harms, Boston College Law Review, Vol. 55, No. 1, 2014 Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2325784

The increasing prevalence of tracking consumers via mobile devices

Walmart says that more than half of its customers own smartphones, and among shoppers younger than 35, this figure is closer to three-quarters. Smartphone-carrying customers are important to Walmart; the company estimates that each month they make four more trips and spend 77% more in stores. Mobile users account for a third of Walmart.com traffic most of the year, and as much as 40% during the holidays.[17]

Walmart's focus on mobile technologies for gathering information should be of particular interest to communities of color, which have adopted smartphone technology at faster rates than white communities.[18] A 2013 report from the Pew Research Center found that 64% of African-Americans and 60% of Hispanics owned smart phones, compared to just 53% of whites.[19]

Walmart has decided it needs to reach these mobile customers. As Gibu Thomas, Walmart's global head of Mobile, explained: "Our goal is to create shopping tools that become second nature to the customer, providing assistance with every part of the retail experience from pre-store planning to in-store shopping and decision making to checking out," Thomas said. "The future of retail is the history of retailing. It's about a personalized experience for each shopper delivered through the smartphone."[20]

Mobile devices offer Walmart and other companies a sort of portal between what might have once been considered our separate online and offline worlds. These devices are associated with us through our phone numbers and email addresses, our individual settings and browsing habits, and unique device identifiers.

Smartphones come with a unique identifier called a MAC address; among other things, it allows a wireless router to distinguish between devices. Importantly, this unique identifier is not considered personally identifiable information, as the term is used in US privacy law and many corporate privacy policies. At the end of 2012, Nordstrom began experimenting with using Wi-Fi signals from customers' smartphones to track their movements throughout the store.[21] A backlash from customers prompted the company to end the experiment in May, but Walmart is quietly granting itself access to the same practices. The company's Wi-Fi Terms of Use explicitly state:

> *[W]e want you to know that, when you access or use [wireless networks and associated services], Walmart will receive information that may identify you or the device you are using. That information may include MAC address, IP address, and unique device identifier such as a name or number assigned to the device. If you were required to login in order to access or use the Service, then registration information we obtain from you may include your name, phone number and email address.*
>
> *We also will receive the real-time location of your device while it is accessing the Service.*[22]

17 http://reviews.cnet.com/8301-12261_7-57585710-10356022/walmart-exec-mobile-can-revive-personal-touch-for-shoppers/
18 http://colorlines.com/archives/2011/12/the_new_digital_divide_two_separate_but_unequal_internets.html
19 http://www.pewinternet.org/Reports/2013/Smartphone-Ownership-2013/Findings.aspx
20 http://www.qas.com/data-quality-news/investing_in_big_data_can_help_improve_corporate_e_commerce_initiatives_9599.htm
21 http://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html
22 http://corporate.walmart.com/privacy-security/wi-fi-terms-of-use

In-store video tracking

Several companies, including Retail Next and Brickstream use video to understand how customers shop. Brickstream offers a camera that "separates adults from children, and counts people in different parts of a store to determine which aisles are popular and how many cash registers to open."[23] This is nothing compared to London-based Realeyes which "analyzes facial cues for responses to online ads, monitors shoppers' so-called happiness levels in stores and their reactions at the register."[24] Walmart's privacy policy specifically gives the company permission to use video cameras to collect information on shoppers.[25]

More in-store tracking to come

In the near future, consumers may experience even greater integration of tracking and targeting technology in retail stores. In 2015, snack maker Mondalez International plans to launch "smart shelves" using Microsoft Kinect technology. The displays would be located next to check outs, and according to the *Wall Street Journal*, "will use sensor technology to identify the age and sex of the would-be snacker, analytics to determine what type of guilty pleasure best appeals and a video display to deliver custom advertisements."[26]

Walmart's latest additions to Scan & Go, its in-store self-checkout app, offer some insight into the company's data collection and targeting. If customers access the store's Wi-Fi network to use the app, their movements can be tracked as outlined above. Additionally, at the beginning of August, Walmart announced that it would offer mobile coupons to Scan & Go users. That same update to the app added electronic receipts and the ability to get a digital copy of qualified DVD or BluRay purchases (through the company's VUDU video on demand service). Industry blog StorefrontBacktalk described it as "another nice bit of camouflage CRM [customer relationship management]-database-boosting."[27]

Meanwhile, Walmart is reportedly exploring development of an in-store social network. According to a 2011 PSFK blog post, "Using their mobile phones to interact, customers would be able to ask other shoppers in the store questions or provide answers, seek and give opinions on products, and alert them to deals and special offers."[28]

The potential for price discrimination

A *Wall Street Journal* investigation into web pricing at Staples revealed one reason why consumers should be wary of corporate tracking. The paper found that Staples' website displayed different prices to different consumers, after estimating their locations relative to rival retailers. Varying prices by geography is not necessarily illegal, but it can correlate with discrimination along other lines as well: "the Journal's testing also showed that areas that tended to see the discounted prices had a higher average income than areas that tended to see higher prices."[29]

---

23 http://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html
24 http://www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html
25 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
26 http://blogs.wsj.com/cio/2013/10/11/snackmaker-modernizes-the-impulse-buy-with-sensors-analytics/?mod=wsj_streaming_stream&cb=logged0.1563186061378965
27 http://storefrontbacktalk.com/payment-systems/walmart-adds-to-its-scan-go%E2%80%94and-reminds-us-how-to-make-mobile-work/
28 http://www.psfk.com/2011/11/walmart-envisions-in-store-social-network-for-shoppers.html
29 http://online.wsj.com/news/articles/SB10001424127887323777204578189391813881534

On geographic price variation, the *Journal* concludes, "But using geography as a pricing tool can also reinforce patterns that e-commerce had promised to erase: prices that are higher in areas with less competition, including rural or poor areas. It diminishes the Internet's role as an equalizer."[30] Further, consumers currently have little to no way of knowing how the ads and prices they see were determined, a problem identified in the *New York Times* story about audience segmentation: "Unlike a marketplace where individuals haggle with sellers on equal terms, the new world of price discrimination is one where it's hard to escape your consumer profile, and you won't even know if companies are offering discounts to higher-status customers in the first place."[31]

Potential for predictive modeling from big data and "unintended" consequences

An often-cited example of consumer tracking and retailer data mining is a story reported in the *New York Times* in early 2012: Target realized a man's high school aged daughter was pregnant before he did, and the company sent her promotional mailings accordingly.[32] Evidently, it was only after receiving them that he found out he was going to be a grandfather; it was not because his daughter had previously chosen to disclose that information.

What is perhaps most disturbing about this example is that the young woman in question had not, in fact, disclosed that she was pregnant to Target. Instead, Target's "predictive analysis 'guessed' that a customer was pregnant and disclosed her name to their marketing department, manufacturing PII."[33] Crawford and Schultz call this type of harm a "predictive privacy harm" and emphasize the real threat of this kind of analysis to lead to discriminatory outcomes.

"Big data processes can generate a model of what has a high probability of being PII [personally identifiable information], essentially imagining your data for you."

—Crawford and Schultz 2013

Existing law doesn't stop these kinds of revelations of personal information by advertisers, particularly since it is not technically PII. Similarly, most company's privacy policies do nothing to prohibit these types of revalations. In fact, there is no reason these ads couldn't be more predatory. For example, one writer suggested that firms could determine that someone has (or has a high likelihood of having) a particular serious illness and use that information to market unnecessary insurance policies to them.[34]

30 http://online.wsj.com/news/articles/SB10001424127887323777204578189391813881534
31 http://www.nytimes.com/2012/12/02/magazine/who-do-online-advertisers-think-you-are.html
32 http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=1&pagewanted=all
33 Kate Crawford and Jason Schultz, Big Data and Due Process; Toward a Framework to Redress Predictive Privacy Harms, Boston College Law Review, Vol. 55, No. 1, 2014 Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2325784
34 http://www.motherjones.com/politics/2013/10/future-of-privacy-nsa-snowden

Further, as *Mother Jones* reports, "there's nothing that prevents the government from buying up all this information and combining it with its programs into an even bigger surveillance octopus."[35] It is not unlike a former Bush-era program called Total Information Awareness, which had this goal.[36]

More recently, revelations about the National Security Agency's tracking has shed light on how corporate data collection and government surveillance are in fact not separate realities. The NSA and FBI had access to the contents of messages and conversations taking place through Google, Facebook, Apple, Yahoo, Microsoft, and Skype.[37] And we now know that the NSA collects records on the phone calls of the majority of Americans, thanks to an arrangement with Verizon, AT&T, and Sprint.[38] Even more recently we found out that the CIA buys phone records from AT&T for over $10 million a year,[39] which was certainly news to most AT&T customers—and anyone who's ever gotten a call from an AT&T customer.

35 http://www.motherjones.com/politics/2013/10/future-of-privacy-nsa-snowden
36 http://www.motherjones.com/politics/2013/10/future-of-privacy-nsa-snowden
37 http://online.wsj.com/news/articles/SB10001424127887324299104578529112289298922
38 http://online.wsj.com/news/articles/SB10001424127887324299104578529112289298922
39 http://www.nytimes.com/2013/11/07/us/cia-is-said-to-pay-att-for-call-data.html

HIGHLIGHTING RISKS FOR COMMUNITIES OF COLOR AND LOW-INCOME COMMUNITIES

We believe that people of color, who are also more likely to be low-income than the US population as a whole, face particular risks as a result of evolving uses of big data in general and Walmart's use of these practices in particular.

Speaking of privacy rights in general, Professor Christopher Slogobin of Vanderbilt Law School writes of a "poverty exception" to privacy rights generally and concludes that "relative wealth plays a major role in determining whether a person's interests are protected by the [Fourth] Amendment ."[40]

The growth of big data presents more specific risks for marginalized groups. Above, we discussed what Crawford and Schultz describe as "risks of bias or discrimination based on the inappropriate generation of personal data."[41] Essentially, as companies like Walmart increasingly use data, both real and predicted, to put people into categories, the risk grows that some groups will fall disproportionately into categories which receive less favorable treatment. Professor Joseph Jerome of the Northwestern University explains:

> *Most of the biggest concerns we have about big data—discrimination, profiling, tracking, exclusion—threaten the self-determination and personal autonomy of the poor more than any other class. Even assuming they can be informed about the value of their privacy, the poor are not in a position to pay for their privacy or to value it over a pricing discount, even if this places them into an ill-favored category.*[42]

Among the more disturbing aspects of this potential discrimination is the complete lack of transparency and accountability. Whereas many forms of discrimination have historically been "testable," discrimination by big data will likely be very difficult to detect and—even when reasonably suspected—to prove. As MIT researcher Kate Crawford said at a recent conference, "It's not that big data is effectively discriminating—it is, we know that it is. It's that you will never actually know what those discriminations are."[43]

When a company like Walmart gathers massive quantities of consumer data, uses that data to predict other "information" and uses these complex profiles to target shoppers, discrimination may well be the result, even when no discriminatory intent exists.

As the arena of big data and predictive privacy harms evolves, we believe consumers in general and communities of color in particular, have reason to be concerned. This is particularly true in the case of companies like Walmart—which has made clear its massive big data ambitions, has clearly spelled out its intent to market heavily to communities of color, and has, thus far, provided little transparency to allow advocates and the public to understand how their information is being used, or abused.

---

40 Christopher Slobogin, *The Poverty Exception to the Fourth Amendment*, University of Florida Law Review, Vol. 55, 2003 Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=364580

41 Ibid

42 Joseph W. Jerome, Buying and Selling Privacy: Big Data's Different Burdens and Benefits, 66 Stan. L. Rev. Online 47 (2013).

43 http://finance.yahoo.com/blogs/the-exchange/big-data-could-create-era-big-discrimination-191444085.html

WALMART'S EFFORTS TO KEEP CONSUMER ONLINE TRACKING ARE LARGELY UNREGULATED IN THE U.S.

The use of online tracking and big data analysis in e-commerce is largely unregulated in the United States. The increasing reliance on predictive modeling makes effective regulation even more difficult. Crawford and Schultz note that "By analyzing primarily metadata, such as a set of predictive or aggregated findings without displaying or distributing the originating data, big data approaches often operate outside of current privacy protections, effectively marginalizing regulatory schema."[45]

To date, no US legislation has dealt with the policies that permit government agencies and corporations to abuse personal data.

—Terms and Conditions May Apply[44]

And Walmart appears committed to keeping this arena largely unregulated. Walmart has reported lobbying on some variation of "privacy, online advertising and data protection," "privacy and online behavioral advertising legislation," or "privacy issues related to e-commerce" every quarter for the past five years. The company has spent nearly $34 million lobbying the federal government since it first reported broaching this topic in the fourth quarter of 2008, through the third quarter of 2013.

Walmart also interfaces with the government on this issue in other ways. Although she said little of substance, in 2009 Walmart sent its then-Chief Privacy Officer, Zoe Strickland, to speak on the first panel of the Federal Trade Commission's Exploring Privacy roundtable series.[46] More recently, Walmart was invited to speak before a closed-door meeting of the House of Representatives privacy task force in September.[47] Walmart also receives cover from its membership in the Retail Industry Leaders Association. That group also actively lobbies on online privacy[48] and is much more forthright about its preference for a "self-regulatory model" (as opposed to an actual enforceable framework).[49]

44 http://vimeo.com/ondemand/termsandconditions

45 Crawford, Kate and Schultz, Jason, Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms (October 1, 2013). Boston College Law Review, Vol. 55, No. 1, 2014; NYU School of Law, Public Law Research Paper No. 13-64; NYU Law and Economics Research Paper No. 13-36. Available at SSRN: http://ssrn.com/abstract=2325784

46 http://www.ftc.gov/bcp/workshops/privacyroundtables/

47 http://www.politico.com/morningtech/0913/morningtech11764.html

48 http://www.opensecrets.org/lobby/clientissues_spec.php?id=D000026646&year=2012&spec=TEC

49 http://www.rila.org/governmentaffairs/public-policy/Pages/Privacy.aspx

Today, companies like Walmart know more about individual consumers' habits, preferences, likes, and dislikes than ever before. In 2013, Professor Ryan Calo from the University of Washington Law School wrote:

> *[T]he digitization of commerce dramatically alters the capacity of firms to influence consumers at a personal level. A specific set of emerging technologies and techniques will empower corporations to discover and exploit the limits of each, individual consumer's ability to pursue his or her own self-interest. Firms will increasingly be able to trigger irrationality or vulnerability in consumers—leading to actual and perceived harms that challenge the limits of consumer protection law, but which regulators can scarcely ignore.*[50]

Yet current law is woefully inadequate when it comes to protecting consumers from this new range of strategies being employed by Walmart and other online retailers.

The Federal Trade Commission's May 2012 report, *Protecting Consumer Privacy in an Era of Rapid Change*, found that:

> *First, the collection and commercial use of consumer data in today's society is ubiquitous and often invisible to consumers. Second, consumers generally lack full understanding of the nature and extent of this data collection and use and, therefore, are unable to make informed choices about it.*[52]

But, while regulators have not ignored the issue of online tracking, no comprehensive legislation has yet been passed on the subject.

Even if Congress were not in a state of virtual gridlock, crafting legislation that protects consumers from online tracking would be complicated, given rapid technological innovations that are making more and more complex strategies for tracking consumers possible.

US privacy legislation proposed, but not law

In recent years, several members of Congress have introduced legislation that would increase regulation of the ways in which Walmart and other firms track consumers online and through mobile devices, and how they use the information they collect. However—and perhaps the result of extensive lobbying on the issue by Walmart and other giants of e-commerce—the proposed legislation has not become law.

The consumer of the future will be increasingly mediated, and the firm of the future increasingly empowered to capitalize on that mediation in ways fair and suspect. A firm with the resources and inclination will be in a position to surface and exploit how consumers tend to deviate from rational decision making on a previously unimaginable scale. Firms will increasingly be in the position to create suckers, rather than waiting for one to be born everyone minute.[51]

–Professor Ryan Calo
University of Washington Law School

50 Ryan Calo, Digital Market Manipulation University of Washington School of Law, Legal Studies Research Paper No. 2013-27, Available at: http://data.over-blog-kiwi.com/0/55/29/63/201309/ob_376c62_digital-market-manipulation.pdf

51 Calo, M. Ryan, Digital Market Manipulation (August 15, 2013). George Washington Law Review, Forthcoming; University of Washington School of Law Research Paper No. 2013-27. Available at SSRN: http://ssrn.com/abstract=2309703 or http://dx.doi.org/10.2139/ssrn.2309703

52 http://www.ftc.gov/os/2012/03/120326privacyreport.pdf

Absent federal action, several states are taking on privacy legislation

Several states, including California, Colorado, Connecticut, and Delaware, have adopted laws related to consumer privacy on their own.[53] In 2013, California passed three separate bills related to online privacy which were signed into law by the state's governor.

The bills include AB 370, which requires companies to disclose in their privacy policy how the operator responds to browser "do not track" signals or other "mechanisms that provide consumers a choice regarding the collection of personally identifiable information that the operator collects about individual consumers who use or visit its Web site or online service and 3rd parties with whom the operator shares the information."[54] The law goes into effect on January 1, 2014. Walmart does not currently disclose this information in its privacy policy.[55]

California also enacted a law, SB 568, to require companies to create an "eraser button" for teens, with which can allow teens to delete their own postings.[56]

Europe taking the lead, with heavy opposition from US firms

Lawmakers at the European Union have been deliberating on significantly expanded legal protection for consumers online for several years. The legislation was approved by a key committee of the European Parliament in late October 2013 and will now move to approval for the EU Plenary, and then by the 28 member governments.

If enacted, the proposed law will:

- Limit companies' ability to profile users, including a requirement that firms fully explain their use of personal data in detail to customers and a mandate that they seek consent prior to collecting the information;
- Create a right for users to ask companies to fully erase their personal data;
- Require firms to designate "data protection officers" to ensure that the law is followed.

The European privacy law was heavily opposed by leading US tech firms, as well as the by the United States Commerce Department.57 Walmart operates stores in the UK under the name Asda, and the company also sells its George clothing line online in 24 countries, including several in Europe. Lawmakers have stated that they hope to conclude the process before end of their term in May 2014.

53 http://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-related-to-Internet-privacy.aspx
54 http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201320140AB370
55 http://corporate.walmart.com/privacy-security/walmart-privacy-policy#FullPolicy
56 http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201320140SB568
57 http://www.nytimes.com/2013/01/26/technology/eu-privacy-proposal-lays-bare-differences-with-us.html

WALMART'S E-COMMERCE GOALS: EVERY PERSON, EVERY PRODUCT, CONNECTED

This spring, Walmart's CEO of Global E-commerce, Neil Ashe, explained Walmart's technological aspirations at an investor conference: "We're building a global technology platform whose goals are as simple, frankly, as they are audacious. We want to know what every product in the world is. We want to know who every person in the world is. And we want to have the ability to connect them together in a transaction."[58]

Walmart investors and other business observers would be justified in any skepticism about such grandiose pronouncements. The company has struggled for more than a decade in e-commerce, with a revolving door of e-commerce leadership and a string of questionable acquisitions. *Retailing Today* called the gap in e-commerce between Walmart and industry leader Amazon "wide and growing."[59] And, as the publication *Walmart Supplier News* noted recently, "the company discloses few substantive details about the performance of its e-commerce business."[60]

Walmart CEO Mike Duke has said that his biggest regret is not moving faster in e-commerce.[61] The company has a history of failed attempts at e-commerce, which CFO Charles Holley acknowledged in early 2013: "We also know that we have not been good at the technology part of this, but we're catching up quickly with our search engine and how we market on the Internet. We have a lot of work to do to make sure we're efficient and getting the products to the customers, but we feel like that we have the tools to go do that."[62]

Regardless of whether it will turn its big data ambitions into real investment returns, it is clear that Walmart's *ambition* in e-commerce is real. Walmart US CEO Bill Simon told the audience at a September 2013 Goldman Sachs conference that Walmart's knowledge of consumers is unrivaled, thanks in part to data from its Sam's Club membership program. He said, "If you take that data and you correlate it with traceable tender that exists in Wal-Mart Stores and then the identified data that comes through Walmart.com and then the trend data that comes through the rest of the business and working with our suppliers, our ability to pull data together is unmatched."[63]

Playing catch up through acquisitions

Making up for lost time, Walmart has used its deep pockets to pursue a strategy that centers on acquisitions and acqui-hires (acquiring a company, at least in part, to "acquire" the talent of the founders)[64] to attempt to improve the company's competitiveness in e-commerce. A look at these purchases shows Walmart's interest in collecting information about consumers through social media and reaching them on their smart phones. Below, we describe some of Walmart's most important recent e-commerce acquisitions and initiatives.

58 Evan Clark. "Walmart Sets Aggressive Digital Plan." *Women's Wear Daily.* May 2, 2013.
59 http://www.retailingtoday.com/article/why-walmart-can%E2%80%99t-be-compared-amazon?ad=walmart-news-now
60 http://dig.retailingtoday.com/2013-summer-wsn/
61 http://www.businessinsider.com/walmart-ceo-shares-his-biggest-regret-2012-12
62 http://www.internetretailer.com/2013/03/19/wal-mart-outlines-its-e-commerce-priorities
63 http://az204679.vo.msecnd.net/media/documents/goldman-sachs-event-edited-transcript_130233998648598533.pdf
64 http://www.npr.org/blogs/alltechconsidered/2012/09/25/161573307/employee-shopping-acqui-hire-is-the-new-normal-in-silicon-valley

Building a "Social Genome" with Kosmix

In April 2011, Walmart purchased Kosmix, a social media start-up focused on e-commerce. Before the acquisition, Kosmix developed software to search and analyze connections in real-time data streams (such as Twitter) to deliver personalized insights to users. Kosmix was working to build a giant knowledge base called the "Social Genome" to capture information and relationships about people, events, topics, products, locations, and organizations.[65] Walmart turned the company into @WalmartLabs, the Silicon Valley arm of the world's largest retailer. The purchase was rumored to cost roughly $300 million.[66]

@WalmartLabs describes the Social Genome as "a vast, constantly changing, up-to-date knowledge base with hundreds of millions of entities and relationships. We then use the Social Genome to perform semantic analysis of social media and to power a broad array of e-commerce applications."[67]

In January 2013, Walmart reported that its social media analytics project "operates on top of a searchable index of 60 billion social documents and helps merchants at Walmart monitor sentiments and popular interests real-time, or inquire into trends in the past. One can also see geographical variations of social sentiments and buzz levels. There are also tools that marry search trends on walmart.com, sales trends in our brick-and-mortar stores and social buzz all in one place, to help make correlations."[68]

Walmart's own search engine

As CFO Charles Holley indicated, Walmart now prides itself on having its own, homegrown search engine, called Polaris. According to *Internet Retailer*, Polaris now offers tailored results: "The search engine's algorithm now presents results based on the items consumers have recently bought on Walmart.com, the products shoppers are posting about and sharing on Facebook, the products that have received positive reviews, as well as the items that other consumers have searched for and clicked on."[69]

Scan & Go: Self-checkout app that gives Walmart mountains of new consumer data

Currently in the testing phase, Walmart launched its Scan & Go program as part of the Walmart app on iPhones and Android devices in August 2012. Customers use their phones to scan items as they shop and then pay at self-checkout as they leave the store.[70] As of November 20, 2013, Scan & Go was available at about 200 Walmart stores in the following markets: Bentonville, Atlanta, Denver, Phoenix, Omaha, Dallas, Austin, Oklahoma City, Tulsa, Wyoming, Bozeman, Seattle, San Jose, and Portland (OR).[71]

65 http://www.informationweek.in/Storage/11-12-19/How_Walmart_plans_to_use_Big_Data.aspx?utm_source=newsletter&utm_medium=email&utm_campaign=191211
66 http://allthingsd.com/20110418/exclusive-wal-mart-paid-300-million-plus-for-kosmix/
67 http://www.walmartlabs.com/social/social-genome/
68 http://walmartlabs.blogspot.com/2013/01/the-walmartlabs-social-media-analytics.html
69 http://www.internetretailer.com/2012/08/30/wal-mart-factors-popularity-site-search-results
70 http://wm5.walmart.com/uploadedFiles/Landing_Experiences/2012/Scan_and_Go/Walmart-ScanAndGo-faq-JUL132013.pdf
71 http://wm5.walmart.com/uploadedFiles/Landing_Experiences/2012/Scan_and_Go/scan-and-go-stores-032113.pdf

The Scan & Go technology gives Walmart extraordinary insight into shoppers' behavior. Not only do Walmart's apps include a geolocation feature, but they also offer users coupons while they navigate the store, as reported in the media and borne out by our own in-store test. This is a first step for the company in learning how customers respond to ads while they shop. StorefrontBacktalk sees Scan & Go's potential for tracking this way:

> *But mobile CRM [customer relationship management] is potentially much more exhaustive than plastic CRM. Not only does it include every item scanned, but it would know every item scanned and then put back (deleted). It would know how long that item had been in the cart and exactly where customers were when the decision was apparently made. Did some signage change their mind? Was it a different product? Was it when the customer was near a free sample area? And if it was mobile signage, the system could look up to determine the exact ad being delivered at that exact moment.*[72]

Despite the new initiatives and acquisitions, Walmart still has a long way to go, though. The company has reported that it expects online sales of $10 billion this year, less than 2% of overall sales and just 16% of Amazon's $61 billion in online sales last year.[73]

72 http://storefrontbacktalk.com/securityfraud/walmarts-crm-gateway-mobile-checkout/
73 http://online.wsj.com/article/SB10001424127887323566804578553301017702818.html

WALMART'S PRIVACY POLICY GIVES THE COMPANY BROAD LATITUDE FOR TRACKING CONSUMERS ONLINE

Like any large public company, Walmart has to please both investors and customers at the same time. Walmart's statements regarding online tracking and consumer privacy are a perfect example of the challenges of pleasing both these constituencies – who have very different interests.

Walmart says one thing to investors…

Walmart executives regularly try to impress the media and investors with the quantity of information the company collects about its customers. As noted earlier, Walmart US CEO Bill Simon recently bragged to investors that Walmart's ability to use big data to track consumers was "unmatched." Of the social media data the company collects and analyzes, @WalmartLabs commented on its blog, "Data volume is formidably huge. We are talking about petabytes here. Real-time social data processing requires sophisticated data stores and blazingly fast algorithms."[74] To put this in perspective, a petabyte is the same as 20 million four-drawer filing cabinets filled with text.[75]

In a speech earlier this year, Walmart's Director of Social Strategy, Umang Shah, confirmed that Walmart tracks customers using social media. He noted that Walmart does "try to tie some of the online social activity to transactions. Five years ago, tracking a sale from a social engagement was really hard and it may not even have been possible. Now, it is easy—or very, very possible."[76]

…AND ANOTHER TO CONSUMERS

Most consumers will surely never read Walmart's entire, nearly 3,200-word privacy policy. But, they might see that the company begins with a very public-friendly message: "Walmart recognizes the importance of our customers' privacy. We believe that privacy is more than an issue of compliance – it is one of trust. We strive to manage your personal information based on our basic belief of respect for the individual."[77]

WALMART'S PRIVACY POLICY LETS IT TRACK CONSUMERS IN MANY WAYS

Walmart's privacy policy is written in such a way to give the company extremely broad latitude in tracking consumers online. Specifically, the policy makes clear that Walmart can:

- Share your information with third parties, with the extremely broad limitation that it can be shared with all "service providers or suppliers that help with our business operations or joint products" or "when necessary to protect the safety, property, or other rights of Walmart, customers, or associates."[78]
- Track consumers through cookies, "internet protocol address, your device operating system and browser type, the address of a referring website, if any, and the path you take through our websites." Additionally, Walmart uses web beacons so that it can "know if a certain page was visited, an email

74 http://walmartlabs.blogspot.com/2013/01/the-walmartlabs-social-media-analytics.html
75 http://mozy.com/blog/misc/how-much-is-a-petabyte/
76 https://vimeo.com/66613816
77 Walmart Privacy Policy. Updated September 17, 2013; accessed November 4, 2013. http://corporate.walmart.com/privacy-security/walmart-privacy-policy
78 http://corporate.walmart.com/privacy-security/walmart-privacy-policy

was opened, or if ad banners on our website and other sites were effective."[79]

- Use data obtained when consumers create gift registries or mobile shopping lists[80]
- Use information obtained through the application for a hunting or fishing license[81]
- Combine information that Walmart collects with information from "other sources." The policy does not limit these sources and the company does specify that they collect information from "consumer reporting agencies or other service providers if you obtain certain financial products."[82]
- Automatically collect information from mobile devices, including "the type of mobile device you use, the temporary or persistent unique device identifiers (sometimes called UDID) placed by us or our service providers, the unique identifier assigned by Walmart to your device, the IP address of your mobile device, your mobile operating system, the type of mobile Internet browsers you use, and information about the way you use our mobile applications." And, in devices that have Walmart apps installed, Walmart collects location information unless users opt out.[83]

WALMART CUSTOMERS HAVE NO RIGHT TO HAVE THEIR INFORMATION DELETED

One of the tenets of the proposed European consumer privacy law is the right of consumers to have their information deleted. By contrast, Walmart does not appear to offer consumers any ability to have information about them held by Walmart deleted.

In the case of mobile apps, it is even more difficult for a consumer to end Walmart tracking. Even after a consumer uninstalls the Walmart app, information from Walmart remains on the device. From the company's privacy policy: "If you uninstall the mobile application from your device, the Walmart unique identifier associated with your device will continue to be stored. If you re-install the application on the same device, Walmart will be able to re-associate this identifier to your previous transactions."[84]

Even in the case of children who access Walmart's site, Walmart's policy does not offer any sort of guarantee that their information will be deleted. Walmart's online sites do not typically distinguish between children and adults for purposes of online tracking, although their policy contains the apparently toothless claim that the company does not "knowingly collect personal information from children under the age of 13 without prior parental consent." When this information is collected, the company's privacy policy puts the onus on parents to proactively contact the company if they would like to have the information deleted "and we will work to delete it"—a claim that seems significantly different from "we will delete it."[85]

79 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
80 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
81 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
82 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
83 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
84 http://corporate.walmart.com/privacy-security/walmart-privacy-policy
85 http://corporate.walmart.com/privacy-security/walmart-privacy-policy

Results of A Technical Analysis of Consumer Tracking on Walmart's Websites and Apps

Analysis Conducted by Ian Davey, CEO and Principal Consultant of Technolegis

In general, there is a significant amount of data collection occurring within Walmart's smartphone apps and on its websites, both by Walmart and by over fifty other parties. The data collected includes unique identifiers (both those generated by the party collecting the data and those belonging to the device being used), system information such as device type and operating system version, location information, and pages and screens the user has viewed. There are a number of reasons for collecting this data:

- Virtually all collected information is useful for targeting of advertisements on other sites and in other apps, as well as for suggesting additional products a user might be interested in.
- In Walmart's case, information about the device, location, and pages viewed is necessary for a functioning app or website; Walmart's servers must know what products a user wants to view in order to supply information on those products, and must know the type of device the user has in order to determine whether to supply a mobile or desktop version of its website. Walmart must also know where the user is located in order to find nearby stores or determine if the app should switch to in-store mode.
- Relatedly, information about the device as well as pages and screens viewed is also useful for the developers of a website or app in order to identify what aspects to improve in the future; if users spend more time searching than in Walmart's photo center, for example, the developers may then focus more of their resources on improving the former and less on improving the latter.

While the second use is necessary for a functioning app or website and the third is helpful for making improvements, the first is the most concerning for consumer privacy. A first-party such as Walmart can use this data to sort consumers into highly specific market segments, where consumers in different segments will see different advertisements or receive different coupons.[1] While we found no concrete evidence that Walmart in particular is performing price discrimination by segment through targeting coupons or special deals, trends in the retail industry suggest this may become commonplace in the near future.[2]

A website or app publisher may wish to bring in third-party trackers for a number of reasons. One fairly benign use for a third-party is for outsourcing developer analytics; however, our analysis shows that Walmart already collects its own analytics itself. Other uses may involve selling user behavioral data, such as which users view which products. With this data, the trackers can build a record for each user containing which websites the user visits, which pages on those websites the user views, and conclusions they draw about the user's interests and/or lifestyle. They can then use these profiles to target advertising and/or sell the profiles to other such interested parties. While we cannot say for certain what benefit Walmart is getting by allowing third-parties onto its websites and into its apps, most of the third-parties are heavily involved in the online advertising and data brokerage industries, and are likely either providing Walmart with extra revenue outright or advertising space on other sites.

1 www.nytimes.com/2012/12/02/magazine/who-do-online-advertisers-think-you-are.html
2 www.nytimes.com/2013/07/15/business/attention-shopper-stores-are-tracking-your-cell.html

When a website or app publisher wishes to incorporate third-party tracking, whether for outsourcing developer analytics, selling user behavioral data, or supplying advertising space, the developer will often incorporate scripts or libraries produced by the third-party into the website or app. These extra pieces of code will then provide all of the data and tools necessary for the tracker to operate.

There are a number of such tools used for online tracking. The most well known method involves a small file on the user's device called a cookie. Usually a cookie will contain a unique identifier, which a tracker uses to look up the user's profile on its servers. Trackers can use cookies in conjunction with page- or screen-specific hidden images known as either beacons or clear images; when a browser or app loads a beacon or clear image, it sends the unique identifiers in the cookie to the tracker, signaling to the tracker that the profile associated with that identifier has loaded an image specific to a certain page or screen, and thus that particular user has visited that particular page. Additionally, trackers may engage in a practice known as fingerprinting, where they exploit the variation between browser and smartphone configurations to construct a unique "fingerprint" for each user or device.[3,4]

While cookies and beacons work well in browsers, the isolation mechanisms between apps on a mobile device make it impossible to track between apps using cookies. To get around this, trackers will also associate the profiles on their servers with various unique identifiers belonging to the device itself, such as the mobile equipment identifier (MEID) or the network hardware address (commonly known as the MAC address).

To determine exactly which parties are collecting which information, we recorded all network traffic entering and leaving the device using a proxy[5] while each of Walmart's apps were running and Walmart's mobile website was loaded in the browser. We also used browser built-in developer tools to record all traffic related to Walmart's desktop site while it was open. We tested Walmart for iPhone and Walmart for iPad on a second-generation iPad running iOS 6.1.2, Walmart for Android and the mobile site on a rooted Motorola Droid 2 running Android 2.3.4, and the desktop site on Google Chrome 30.0.1599 running on Apple OS X 10.8.5. All app versions were those available in their respective app stores on October 23, 2013, and the website testing was performed November 11-12 and 16-17, 2013.

Before discussing the results of the analysis it is important to note that while it is possible to determine what data goes to what party, it is far more difficult to see what each party does with its data afterward through technical means. Any conclusions to this end are drawn from each party's privacy policies and/or marketing literature.

ANALYSIS OF WALMART'S DESKTOP WEBSITE

There is a substantial amount of tracking on Walmart's website, not just by Walmart itself but also by a considerable number of third-parties. While browsing, Walmart's servers receive requests for every image, page, and product viewed, though this is necessary for the website to operate. However, there are other aspects of the site that suggest first- and third-party tracking. Walmart's site contains a significant number of beacons, many of which (but not all) reference third-parties in the query strings of their URLs. Additionally, the domain omniture.walmart.com receives every type of page visited as well as every

3 https://panopticlick.eff.org/browser-uniqueness.pdf
4 http://online.wsj.com/news/articles/SB10001424052748704679204575646704100959546
5 We used Charles Proxy version 3.8.3, available at http://www.charlesproxy.com/

search performed, category or product viewed, the list of items currently in the cart on the cart management and checkout pages, and a list of plugins installed in the browser; this list can be used to determine whether the user can view different types of content (e.g., Flash, Silverlight), or it can be used for browser fingerprinting. This domain is most likely related to the Adobe Systems subsidiary Omniture, which specializes in using data-mining analytics for marketing.[6]

Google is one major third-party present on Walmart's website. One well-known subsidiary, the ad targeting and delivery service DoubleClick, sets a tracking cookie with a unique identifier and receives information about every search query the user performs and category or product the user views. The same information is sent to Google AdSense via the domain pagead2.googlesyndication.com.[7]

Rubicon Project appears to be a company that auctions online advertising space to individual advertisers. Our analysis shows that Rubicon collects the URLs for every product and category viewed and search performed, and shares unique user identifiers with fourteen more third-parties, some of which respond with their own cookies and/or beacons. Furthermore, Rubicon's scripts fetch cookies and/or beacons from thirty-eight additional third-parties. Though none of these additional third-parties receive any direct information about what the user is viewing on Walmart's site, they do receive numbers which Rubicon uses to identify the site (in either the URL query string or the HTTP referral header), and they may be able to use these numbers in conjunction with their own tracking cookies to uniquely identify users either directly or through more Rubicon referrals on other sites. Presumably Rubicon itself uses this data to gauge overall consumer trends and price advertising space accordingly, and its privacy policy also explicitly states that it uses the data collected to individually target advertising.[8]

Advertising.com, a division of AOL, discloses in its privacy policy that it collects website usage information and that the data it collects and receives from its partners "may be combined with information collected from other sources."[9] It also discloses that it uses the data it collects to target advertising and to measure the advertisements' effectiveness.[10] On Walmart's site, Advertising.com collects the user's search queries and products viewed. In addition, AOL receives Rubicon Project's unique identifier via its subsidiary Adap.tv, and fetches cookies from four additional third-parties.

There are several other less-pervasive third-parties:

- Scorecard Research specializes in market research via voluntary user surveys and a tracking technology it refers to as web tagging.[11] We found that it collects the URLs for every search the user performs and product or category the user views, and sends them to domains belonging to Scorecard.
- Criteo, which collects behavioral data and uses it to target advertising on other sites,[12] keeps a unique identifier in a cookie. Criteo also provides this identifier to DoubleClick and Rubicon Project.

6 http://www.omniture.com/
7 http://www.donottrackplus.com/trackers/googlesyndication.com.php
8 http://www.rubiconproject.com/
9 http://www.advertising.com/
10 http://advertising.aol.com/privacy/aol-advertising
11 http://www.scorecardresearch.com/Home.aspx
12 http://www.criteo.com/

- Mixpo Inc., a video advertising provider,[13] receives a unique identifier of its own when the user visits Walmart's home page.
- Triggit, which specializes in targeting ads on Facebook's news feed,[14] sets a cookie with a unique identifier that is later sent to Facebook.
- ChoiceStream, which manages targeted online advertising campaigns,[15] sets a number of persistent cookies containing unique identifiers when the user first loads Walmart's site.
- VisualIQ sets a cookie with a unique identifier through its advertising effectiveness measurement service AudienceIQ.[16]
- BazaarVoice performs market research on social data to help its clients increase sales and build brand loyalty,[17] and on Walmart's site receives the user's search queries and information about the products the user views after searching.
- HookLogic, which specializes in advertising on e-commerce sites,[18] sets a cookie with a unique identifier and receives information about every product and category the user views via the HTTP referrer field.
- Online marketing campaign manager ValueClick,[19] through its subsidiary FastClick, receives information about some of the products the user views, along with a unique identifier.

While many of the above have a limited presence on Walmart's website, they still learn that the user has visited it, along with any other sites they partner with, and can use this data to build a record for the user and target advertising.

Finally, RapLeaf, one of the fourteen third-parties exchanging unique identifiers with Rubicon, deserves a special mention. According to its website, its databases contain 80% of U.S. email addresses, which it pairs with demographic and user interest data for targeted marketing purposes. While its privacy policy states that it does not sell or distribute the actual email addresses it has, it does share the rest of the data with other parties.[20] If there ever were a data breach at RapLeaf, the attacker would then have a large repository of email addresses ready for sending spam.

ANALYSIS OF WALMART'S MOBILE WEBSITE

The mobile website does not appear to have any third-party tracking. However, there is some first-party data collection in place. While most network traffic relates to the core functionality of the website, various data is transmitted to omniture.walmart.com and analytics.mobile.walmart.com. Omniture receives notifications about every type of page visited (e.g. home page, product details page, and cart), as well as the searches the user performs while on the site. The domain analytics.mobile.walmart.com, which is likely Walmart's in-house mobile data-mining system, also receives this information, as well as information about individual products the user views or adds to the cart.

13 http://multiscreen.mixpo.com/
14 http://triggit.com/facebook-retargeting
15 http://choicestream.com/
16 http://www.visualiq.com/products/audience-iq
17 http://www.bazaarvoice.com/
18 http://hooklogic.com/
19 http://www.valueclick.com/
20 http://www.rapleaf.com/

ANALYSIS OF WALMART FOR IPHONE

While running, the iPhone app makes a number of requests to servers controlled by Walmart. Many of these are simply requests for images and data related to products the user is viewing, and contain no more information than the device type and, where applicable, identifiers for the products the user wishes to view. However, the app also sends requests to the domains omniture.walmart.com and analytics.mobile.walmart.com. As with the websites, Omniture receives notifications about every screen visited and product or category the user views, paired with a generated unique identifier and a timestamp. The domain analytics.mobile.walmart.com receives notifications about screens, products, and categories viewed along with its own generated unique identifiers. When creating or logging into an account, analytics.mobile.walmart.com also receives the username (which consists of the user's email address) and whether the login was successful.

Additionally, a number of third-parties receive information from the app. The mobile analytics company UrbanAirship[21] receives data at two domains. The domain device-api.urbanairship.com receives a generated unique identifier and a list of indicators (referred to as tags) that appear if the user has used various features of the app such as setting a store (including which store) or searching for a product, and sets a cookie with a generated username and password. The domain combine.urbanairship.com receives this username with notifications about what screens the user views and the time the user views them. UrbanAirship's privacy policies,[22] marketing literature, and developer documentation[23] suggest the company is using this data to report app usage back to Walmart.

The mobile analytics company and recent Criteo acquisition AdXTracking collects general actions within the app such as adding an item to the cart or searching for a product, and each request includes not only app- and developer-specific identifiers but also the device's MAC address and global unique device identifier (UDID). In the iPhone version at least, there is no collection of what specific products the user is viewing. The sparse privacy policies AdXTracking makes available leave ambiguous whether they apply only on their website and only if the website's visitor has clicked an opt-out button. Criteo's website, however, discloses more information about online tracking more precisely, but the policies there seem only to apply to tracking on webpages, and they did not respond to a request for information on mobile tracking.

ANALYSIS OF WALMART FOR IPAD

While Walmart for iPad exhibits some of the same tracking behavior as Walmart for iPhone, there appears to be less. The app still makes functionality-related requests to Walmart's servers and sends the same data about screens, categories, and products viewed and login information to analytics.mobile.walmart.com, but omniture.walmart.com receives nothing.

Third-party tracking remains unchanged. UrbanAirship still receives tags describing features used and notifications of individual in-app actions. AdXTracking receives the same notifications about user actions with the device's MAC address and UDID.

21 http://urbanairship.com/
22 http://urbanairship.com/legal
23 http://docs.urbanairship.com/

ANALYSIS OF WALMART FOR ANDROID

The first-party tracking in Walmart for Android is lighter than that in the iOS apps. Walmart still receives the category, search, and product information necessary for the app to function, but analytics.mobile.walmart.com receives nothing. However, when configuring shipping addresses at checkout while logged in, notifications indicating the user is doing so are sent to omniture-ssl.walmart.com.

However, third-party tracking is more substantial. While UrbanAirship still collects the same set of information that it does from the iOS apps, AdXTracking collects more. In addition to the notifications about user actions and UDID (referred to on Android as Secure ID or Android ID), it collects the identifier for every product the user views or adds to the cart, as well as every search the user performs.

It is also worth noting that all three apps use a fraud-detection service called ieSnare, developed by iovation, Inc., at their checkout screens. ieSnare uniquely identifies mobile devices using its DevicePrint system and assigns each one a reputation in a centralized database referred to as the Device Reputation Authority.[24] Supposedly, if a device's reputation were too low, the app would not allow it to complete a checkout. While iovation's privacy policy currently states that it only shares its database with its customers (i.e., Walmart's checkout system), it also states that it retains the data (which contains device properties, IP addresses, and location data) indefinitely and "may modify this privacy policy from time to time;"[25] thus these practices are one such policy change or data breach away from a situation very harmful to consumer privacy.

24 http://docs.apwg.org/sponsors_technical_papers/iovation_whitepaper.pdf
25 https://www.iovation.com/privacy/


THE CENTER FOR
MEDIA
JUSTICE


colorofchange.org


Sum
Of
+ Us