The Commoditization of User Data in Web Services:

A critique of the tech industry's data practices

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Harrison Li

Spring 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Harrison Li

STS advisor: Hannah Rogers PhD, Department of Engineering and Society

Abstract

Tech companies collect data from customers interacting with their services to use for business development. This behavior sets a bad precedent because it enables malicious tracking and surveillance of individuals. It is evident that the right of privacy is being violated through the cases of Steve Jackson Games vs United States Secret Service, Consumer Scoring Systems, and a recent class action lawsuit against Zoom Video Communications Inc.. These cases and their implications will be further explored to demonstrate the need for stricter privacy laws.

Introduction

Recall a recent impulse purchase of an item online. What cascade of events led to that purchase? Part of being a responsible consumer is to reflect on influences. The customer has final say whether to spend their money or not, but customers can be persuaded. It is difficult for a customer to realize the extent to which a business can influence them. This is because the business already knows about the customer before they visit their website through the commoditization of consumer data. Businesses plan and engineer their services around the customer to optimize purchases. After all, the end goal of a business is to provide value while maximizing profit, so they must be tactical. How far can businesses take data tactics before ethics are violated?

User experience is a broad term with many definitions. The definitions range from "traditional usability to beauty, hedonic, affective or experiential aspects of technology use" (Hassenzahl & Tractinsky, 2006). User experience has become sensationalized in the tech industry because it plays a critical role in attracting a loyal customer-base. For this paper, user experience will serve as the framework behind the motivation for companies to use customer

data. Businesses must identify their ideal customer when designing products and marketplaces because knowing the customer is the first step to selling a product.

How do large tech companies like Amazon come to learn about their customers? They collect and/or purchase it. The Law Insider derives the top definition of "customer data" using 304 legal documents: "any content, materials, data and information that Authorized Users enter into the production system of a Cloud Service or that Customer derives from its use of and stores in the Cloud Service" (Law Insider, n.d.). It is the customers themselves who are providing the data to businesses. When users create accounts for a website, the fine print in the terms and conditions can allow the site to do as they please with their data. This data can be sold to data brokers to be circulated indefinitely. Data privacy is becoming a serious issue for this reason. Internet users must have the right to control their data.

This paper will examine Amazon and their privacy policies to determine how and for what reason data is being collected and used. The paper will then detail the worries behind the use of personal customer data by tech advocacy groups such as the Electronic Frontier Foundation and the World Privacy Forum. The legal side of data privacy will be explored through a class action lawsuit against video communication company, Zoom. Finally, a common argument against data privacy will be showcased and refuted. The findings from this research will improve the reader's data awareness and seek to promote reflection about the state of technology use.

Amazon

Amazon began as an online bookstore founded by Jeff Bezos during the Dot Com Bubble of the late 1990's. The company name, Amazon, is that of the notoriously vast Amazon River in

South America. Rivers bring fish, transportation, and civilization. This river name associates the business with longevity and prosperity. Now, Amazon is a corporate conglomerate dominating in a variety of industries. Amazon subsidiaries include Whole Foods, Audible, Amazon Web Services, and Amazon Prime Video. Timeliness is valued in this fast paced world of globalized technology. A benefit Amazon Prime members receive is expedited shipping times. The value Amazon Marketplace brings to consumers around the globe is the quick access to a wide variety of products. It is clear why consumers have flocked to Amazon. This section will delve into Amazon's spanning reach and their potential sources of customer data.

Online business practices like affiliate marketing and online subscription services are practiced religiously by Amazon. Affiliate marketing is a form of advertising Amazon leverages (Warrier et al., 2017). To become an affiliate of Amazon means to advertise products for a commission. This is a symbiotic relationship for both parties because Amazon receives purchases, customer and affiliate data, and the affiliate gets a portion of the sales profit. Amazon's goal in affiliate marketing is for the consumer to purchase through their service. Amazon retains the information of all these consumers. This is a positive feedback loop.

The Amazon Marketplace has a direct outreach to the general public. Amazon's ecommerce platform is its data gold mine. This is because Amazon is a global corporation. In 2021, Amazon offered their Prime membership services to 17 countries (Warrier et. al, 2017). The majority of these countries are in the developed western world with China, India, and Japan in Asia (Warrier et. al, 2017). Amazon is widely used even in countries where Prime is not offered (Warrier et. al, 2017). Of 130 surveyed Malaysians, Indians, and Indonesians, 118 or 90% know of Amazon.com (Warrier et. al, 2017). Of the 118 participants who know of Amazon,

only 34 have never used their services with 43 shopping once a month or more (Warrier et. al, 2017).

Amazon's Privacy Notice

How do companies collect data? Amazon publishes a privacy notice on their website that details how they collect and use customer data. Amazon classifies data they use into three categories: information provided by the user, automatic information, and information provided by other parties (Amazon.com, Inc, 2021). Note that these categories are decreasing in specificity.

It is known for certain what kind of information users provide when they use Amazon's service. Search history, order history, delivery addresses, contact information, and payment information are all collected from the user and processed by Amazon (Amazon.com, Inc, 2021). The user can request data from Amazon in all or one of fifteen categories. These categories include addresses, payment methods, subscriptions, search history, Alexa and Echo Devices, Kindle, Fire TV, Fire Tablets, advertising data, Amazon Drive, Apps and More, Amazon Music, Prime Video, and Customer Service logs (Amazon.com, Inc, 2022). All of these categories are provided upfront by the user or obtained through device monitoring.

"Automatic information" is less understandable to the average Amazon customer. Amazon explains the term, "Like many websites, we use "cookies" and other unique identifiers, and we obtain certain types of information when your web browser or device accesses Amazon Services and other content served by or on behalf of Amazon on other websites" (Amazon.com, Inc, 2021). This means in addition to providing Amazon data, services owned and working for Amazon are able to share and access a user's data through cookies.

A web service is not invasive because it uses cookies. Web cookies have a bad reputation because they can be exploited by people up to no good. In reality, cookies are useful on the internet. Websites assign cookies to hold onto for later. These cookies can store information about user preferences, most recent page, video timestamp, etc.. Web cookies can be seen as a method of communication instead of malware. Amazon storing and retrieving cookies can be interpreted to be an act of intrusion, but has its justifications. Amazon features like account preferences, shopping cart, and auto-login all rely on cookies (Amazon.com, Inc, 2021). Amazon also states that cookies are used in fraud prevention and security improvements (Amazon.com, Inc, 2021).

The final method Amazon uses to obtain user data is through "other sources". This is the most generic and concerning category of the three because there is no restriction in the term "other sources". Amazon states on their site that they receive credit history reports, address changes, page view history from partnering companies and subsidiaries, and search results (Amazon.com, Inc, 2021). These data exchanges are alerting to the average internet user because the internet is thought to be anonymous and compartmentalized. Knowing the way information flows into Amazon breaks that expectation and leaves the consumer wondering what their data is used for. It is unnerving to discover Amazon knowing a new address before updating the setting on the website.

Amazon's Purposes for Collecting Data

Why is Amazon interested in stockpiling customer data? To understand the reasoning behind this, the user experience framework will be used. Businesses optimize their user experience to attract and retain customers. The top three purposes of collecting user data are for the "purchase and delivery of products and services", to "provide, troubleshoot, and improve

Amazon services", and enable "recommendations and personalization" (Amazon.com, Inc, 2021). The first two purposes are trivial, but the last purpose is a moneymaker.

Stores, online and physical, have tangible and intangible attributes (Findlay, 2002). The tangible aspects of a store consist of store color theming, layout, and product selection. The intangible aspect of a store is its image to the customers (Findlay, 2002). This is the personality of the store. The personality of the store must match the customer at the moment. This is because customers see themselves through the image of the stores they shop at. Customers will not shop at stores they find uncomforting or unrelatable. Therefore, a successful store will have an ideal customer and craft their image to suit that customer. With previous search and order history data, Amazon is able to create a profile for each customer. Specific products can be pushed higher in search results based on current interests. By knowing what the customer likes, Amazon is able to strategically choose the products that will be most likely to be bought. These products can then be displayed on the homepage for high visibility. There is no inherent issue with targeted advertising because it brings customers closer to the products they enjoy. The data privacy issue stems from the methods of obtaining the data used in targeted advertising and the security issues that circulation of data may bring.

The Electronic Frontier Foundation

The Electronic Frontier Foundation (EFF) is a nonprofit organization that seeks to uphold civil liberties over the internet through "impact litigation, policy analysis, grassroots activism, and technology development" (EFF, n.d.). The EFF was founded as a result of an unconstitutional 1990 search and seizure by the United States Secret Service in the case Steve Jackson Games Inc. vs. United States Secret Service (EFF, n.d.).

Steve Jackson Games Inc. vs. United States Secret Service

The United States Secret Service saw that a security threat existed in a file that was being circulated and was conducting computer searches in homes and offices (EFF, n.d.). Steve Jackson Games was one of these offices (EFF, n.d.). Computers were taken from developers and not returned for some time, so the company suffered (EFF, n.d.). The Secret Service determined that the company was no threat and returned the computers (EFF, n.d.). Steve Jackson found on the returned computers that, "all of the electronic mail that had been stored on the company's electronic bulletin board computer, where non-employee users had dialed in and sent personal messages to one another, had been individually accessed and deleted" (EFF, n.d.). Jackson sued the Secret Service and the court determined that electronic messages are to be treated the same as phone calls and must be recorded before being tampered with (EFF, n.d.).

The Secret Service, with a search warrant, was able to raid a company office and read/delete the personal messages of non-search related individuals. How is Steve Jackson's case any different from Google and Amazon monitoring their customers every move, search, purchase, and life update without search warrants? Government security agencies are similar to large tech companies because both want to peer into the lives of their people. If the US government contracts out defense projects to defense companies, then the same can be done for information intelligence. Data in technology is not compartmentalized. Tech companies are extensions to the US government because the agencies utilize products created by contracted tech companies.

Online Behavioral Advertising

Target based advertising is portrayed to be predatory by the EFF because shady markets are able to use data insights to advertise to those who are most likely to fall victim. Tech

companies can do three things when it comes to their customer data: track, profile, and target (Cyphers & Schwartz, 2022). These three actions together constitute online behavioral advertising. According to the EFF, online behavioral advertising is, “almost single-handedly responsible for the worst privacy problems on the internet today” (Cyphers & Schwartz, 2022). This is because tracking, profiling, and targeting customers enables data to spread uncontrollably to places where it is not originally intended for use. Online behavioral advertising has led to the “development of technology so that our devices spy on us by default” (Cyphers & Schwartz, 2022). Cell phones and computers are now equipped with GPS, cameras, and microphones accessible to active and background applications. Home security cameras and assistants like Amazon’s Echo Dot upload video and sound logs to the cloud. Smart watches are recording sleep times, vital signs, and health habits. All this data is able to fall into the wrong hands. The monitoring from George Orwell’s *1984* is not science fiction. The EFF reminds internet users that the companies that participate in online behavioral advertising can transact data with any organization including hedge funds, credit card companies, and even the government (Cyphers & Schwartz, 2022). Data is not compartmentalized.

Data is vulnerable even within company boundaries. The EFF notes that internal privacy concerns exist in companies (Cyphers & Schwartz, 2022). Personal data can be accessed by employees with sufficient permissions. Employees are people. People are not perfect and can be susceptible to corruption and ulterior motives. Workplace ethics are seriously violated if a data engineer were to abuse their position of power. It is concerning for an employee to be able to dig up personal information and use that information to cause harm. This is why data privacy must be upheld as a principle in future technology.

The World Privacy Forum

The World Privacy Forum (WPF) is a front-running nonprofit research organization. Their mission is to research, extract, and share the information necessary to protect global privacy and autonomy (WPF, n.d.). WPF research areas span from privacy in the global medical industries to government organizations to financial markets (WPF, n.d.).

Consumer Scores

Credit score reports in the United States are generated by regulated credit bureaus and determine the trustworthiness of an individual for leases, loans, and purchases. Consumers are able to request and view their scores from credit bureaus. Furthermore, scores are controllable and merit based because on time payments have positive effects and vice versa. A 2014 WPF report exposes the concept of covert consumer scores in online retailing (Dixon & Gellman, 2014). Consumer scores are scores created and updated by companies to rank the trustworthiness of their customers (Dixon & Gellman, 2014). While functionally identical, a consumer score is different from a credit score because it is invisible to the consumer (Dixon & Gellman, 2014). Credit scores are public information and the methods to increase a credit score are known. Consumer scores are a mystery because there is no reason for a company to reveal their ranking methods. For this reason, the vast majority of consumers are not aware of the existence of scoring systems. The factors that affect a specific company's consumer score are unknown. This is akin to completing a school project without a rubric.

The WPF's worries are not in the scoring systems themselves but in the lack of transparency consumer score systems have. The WPF states their concern is in, "the factors used in new consumer scores, which may include readily commercially available information about

race, ethnicity, religion, gender, marital status, and consumer-reported health information" (Dixon & Gellman, 2014). Using commercial data in consumer scores opens up a world of injustices. Are customers from impoverished areas lower in the rankings? Will the system have a bias against any specific group? Additionally, purchased data can be outdated and not reflect current consumer habits. This would lead to an unfair evaluation. By keeping consumer scoring systems private, companies allow certain biases to exist unchecked by government regulation.

Zoom Class Action Lawsuit

Transparency is the key to earning loyal customers. Users feel taken advantage of when companies take actions that violate their trust. Zoom is a video communications service that gained popularity during the Covid-19 lockdowns in 2020. A class action lawsuit was filed against Zoom Video Communications Inc. citing violations of user privacy, and a flood of legal violations during the period of March 2016 to July 2021 (Epiq, 2022). This is Zoom's policy on their website regarding the sharing of data with third parties: "Zoom provides personal data to third parties only with consent or in one of the following circumstances (subject to your prior consent where required under applicable law)" (Zoom Video Communications Inc., 2021). The primary concern of the Plaintiff was that Zoom failed to uphold their own privacy policy (Epiq, 2022). The Plaintiff includes the unauthorized use of third party tools from Google and Facebook to share data with third parties as part of their allegation (Epiq, 2022). The Plaintiff and Zoom agreed to a settlement of $85 million dollars paid out to affected users of Zoom who claimed their share (Epiq, 2022).

Reaching a settlement means that the case did not reach a conclusion. Any and all legal backed investigations halted once this settlement was reached. This can be dangerous because it

sets the precedent that companies can pay their way out of accusations of privacy misconduct instead of risking exposure to illegal actions. This settlement incentivizes unethical and illegal activities like selling data that outpay the penalty fines. While it is possible that Zoom was in the legal right and settled for convenience, people felt like their privacy was being compromised by a company that promised better. In the end, there was enough evidence that a class action lawsuit was brought to light and settled indicating some level of wrongdoing.

Refuting the Nothing to Hide Argument

A common counterargument against data privacy advocates is that innocent citizens should not be worried about surveillance if they have nothing to hide. The argument stems from government surveillance programs sponsored by the United States National Security Agency in response to the threat of domestic terrorism (Abdo, 2013). The agency conducts searches and seizures of personal data to determine domestic threat risk (Abdo, 2013). The Nothing to Hide argument makes logical sense because data collected from law abiding citizens is benign. A lawful citizen will always be innocent.

Supporters of the Nothing to Hide Argument are downplaying the value of privacy in their personal lives. Alex Abdo of the American Civil Liberties Union asserts that it is wrong to assume that all who desire privacy are criminals (Abdo, 2013). "Privacy is a fundamental part of a dignified life", states Abdo (Abdo, 2013). Doors have locks. Windows have blinds. The home is a place of privacy. What is the difference between a neighbor peering through the window and the NSA peering through the security camera? Many homeowners would find it unsettling for a neighbor to peek through a window unsolicited. The neighbor can be reported and told off, but the NSA cannot. One does not need to see the perpetrator for there to be an invasion of privacy.

Edward Snowden, an ex-NSA contractor, writes, "Arguing that you don't care about the right to privacy because you have nothing to hide is no different than saying you don't care about free speech because you have nothing to say" (Snowden, 2015). Free speech is tied to privacy. The relationship is more apparent in oppressive governments where protests are punished. The First Amendment in the US Constitution protects the individual from the US government when they protest against it (Library of Congress, n.d.). Citizens of oppressive governments do not have freedom of speech. Therefore, all criticism of the government is best done in private under no surveillance. Data privacy is important because it can be the difference between freedom and oppression. Any free government shall not be able to survey and punish individuals with opposing views. Constant surveillance will push a society towards oppression because the people worry about who their words and actions will reach.

Conclusion

There is nothing wrong or unethical about a company using customer data to improve their user experience. Both businesses and customers desire better products after all. What data privacy advocates from the EFF, WPF, and ACLU are addressing in their concerns is the lack of transparency in the companies and organizations that abuse the data and rights of their constituents. While a privacy policy seems comforting to have, it is unhelpful that Amazon posts a privacy notice leaving consumers with more questions than answers. Internet users would like to know exactly where data pertaining to them is sourced from, how it is used, and where it is going. Existence of hidden consumer score ranking systems only makes one paranoid of their online behavior.

It is the organizations like the EFF, WPF, and ACLU that are pushing governments to update regulations to reflect the dynamic tech industry. Companies like Zoom must be held legally responsible should they violate privacy laws and the privacy promises made to customers. Data policy is a factor all consumers should consider when deciding which business to support. This is because privacy on the internet is as fundamental as free speech in the United States.

References

Abdo, A. (2013, August 2). *You may have 'nothing to hide' but you still have something to fear*. American Civil Liberties Union. Retrieved May 2, 2022, from https://www.aclu.org/blog/national-security/secrecy/you-may-have-nothing-hide-you-still-have-something-fear

Amazon.com, Inc. (2021, February 12). *Amazon.com Privacy Notice*. Amazon. Retrieved May 2, 2022, from https://www.amazon.com/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8MHFRNJ

Amazon.com, Inc. (2022). *Request My Data*. Amazon. Retrieved May 2, 2022, from https://www.amazon.com/gp/privacycentral/dsar/preview.html

Cyphers, B., & Schwartz, A. (2022, March 21). *Ban online behavioral advertising*. Electronic Frontier Foundation. Retrieved May 2, 2022, from https://www.eff.org/deeplinks/2022/03/ban-online-behavioral-advertising

Dixon, P., & Gellman, B. (2014, April 2). *WPF report - The scoring of america: How secret consumer scores threaten your privacy and your future*. World Privacy Forum. Retrieved May 2, 2022, from https://www.worldprivacyforum.org/2014/04/wpf-report-the-scoring-of-america-how-secret-consumer-scores-threaten-your-privacy-and-your-future/

EFF. (n.d.). *About EFF*. Electronic Frontier Foundation. Retrieved May 2, 2022, from https://www.eff.org/about

Epiq. (2022). *In re: Zoom video communications, Inc.. Privacy Litigation*. Zoom Video Communications Litigation - Frequently Asked Questions. Retrieved May 2, 2022, from https://www.zoommeetingsclassaction.com/Home/FAQ

Findlay, A. M., & Sparks, L. (2002). *Retailing: Critical concepts*. Routledge.

Hassenzahl, M. (2006). User Experience - A Research Agenda. In N. Tractinsky (Ed.), *Behaviour & Information Technology* (2nd ed., Vol. 25, pp. 91–97). essay, Taylor & Francis.

Law Insider. (n.d.). *Customer data definition*. Law Insider. Retrieved May 2, 2022, from https://www.lawinsider.com/dictionary/customer-data

Library of Congress. (n.d.). *Constitution of the United States - First Amendment*. Congress. Retrieved May 2, 2022, from https://constitution.congress.gov/constitution/amendment-1/

Snowden, E. (2015, May 21). *Just days left to kill mass surveillance under Section 215 of the Patriot Act. We are Edward Snowden and the ACLU's Jameel Jaffer. AUA.* Reddit. Retrieved May 2, 2022, from https://www.reddit.com/r/IAmA/comments/36ru89/just_days_left_to_kill_mass_surveillance_under/crglgh2/

Warrier, U., Singh, P., Jien, C. W., Kee, D. M., Yi, G. Z., Jiann, T. W., Liang, T. Y., SB, G., Nair, S., Nair, R. K., Lokhande, S. D., & Ganatra, V. (2021). Factors that lead Amazon.com to a successful online shopping platform. *International Journal of Tourism and Hospitality in Asia Pacific*, *4*(1), 7–17. https://doi.org/10.32535/ijthap.v4i1.1017

WPF. (n.d.). *About Us*. World Privacy Forum. Retrieved May 2, 2022, from https://www.worldprivacyforum.org/about-us/

Zoom Video Communications Inc. (2021, December 15). *Zoom Privacy Statement*. Zoom. Retrieved May 2, 2022, from https://explore.zoom.us/en/privacy/