The Limitations of Consumer Data Agency Under the California Consumer Privacy Act

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

Ashwin Pathi
Spring, 2021

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

Signature  Date 5/6/2021
Ashwin Pathi

Approved  Date _5/4/2021____
Hannah Rogers, Department of Engineering and Society

Abstract

Privacy and data protection have been highly contentious topics in the technology industry for the past decade. The complications that arise from data mishandling become more severe as companies begin to gather more data on consumers. Because of the increasing interest in privacy as a right, and with developments in similar legislation such as the GDPR, the United States recently drafted the California Consumer Privacy Act (CCPA) to manage data privacy and management concerns for consumers. Though the CCPA is a net benefit to consumers, the CCPA suffers from a lack of direction and fails at its original goal of giving consumers more agency over their data because of the differing privacy views of companies, consumers, and privacy advocates. Through the use of social construction of technology, I will study how the interpretive flexibility of privacy by various groups led to the current CCPA legislation, and how it can be improved going forward. Based on my analysis, I outlined several potential improvements relating to data formatting and consumer education. These changes were proposed with the aim of preventing a similar legislative situation from occurring in the future.

The Limitations of Consumer Data Agency Under the California Consumer Privacy Act

INTRODUCTION

According to Forbes, nearly 90% of all internet data has been collected in the past two years. What is more alarming is that most of this data has been collected without user consent or knowledge (Fertik, 2020). Data collection companies such as Google and Facebook have profited billions of dollars from user data, and this number is projected to increase in the future as the internet of things and wearable devices increase in usage (Feiner, 2019). However, in the wake of recent data scandals, consumers have begun to question the data collection practices of these companies (Confessore, 2018a), and as a result of these scandals the California Consumer Privacy Act (CCPA) was passed in 2018. This act acknowledged the leverage that technology companies had over consumers on their data, and sought to level the playing field of consumer and technology company data interactions.

The issue of data agency and privacy is one that spans both technical and social boundaries. Technologically, data collection agencies must think of ways to preserve user data and ensure that data is not accessed fraudulently, and that data is transmitted and accessible in readily available formats. Apart from technological aspects of data collection, improving data agency also grapples with issues of technological literacy as well as political communication. The CCPA largely focuses on general technical legislations, and outlines the processes and restrictions that companies need to put in place in order to be compliant. Though this may be the case, companies have interpreted the legislation in the CCPA differently. Discrepancies have arisen because the CCPA suffers from a lack of direction and fails at its original goal of giving consumers more agency over their data because of the differing privacy views of companies, consumers, and privacy advocates. Ultimately, this paper serves to answer how different groups

imbued their meaning onto data privacy, eventually resulting in the first iteration of the CCPA, and how subsequent iterations can be improved further to account for the differences in perspectives between the various social groups involved in the data collection process.

INTRODUCTION TO THE SOCIAL CONSTRUCTION OF TECHNOLOGY

One of the biggest reasons that the CCPA was not fully focused on consumers was because of how various social groups framed the issue of privacy. Social Construction of Technology (SCOT) aptly fits this analysis as it deals with the interactions between technological artifacts and various social groups (Bijker & Pinch, 1984). SCOT primarily focuses on the concepts of relevant social groups, interpretive flexibility, and closure and stabilization. Of these topics, the concept of interpretive flexibility is especially relevant to the CCPA, as it states that various social groups ascribe different meanings to a particular technology (Bijker & Pinch, 1984). In this case, various social groups in the legislative process of the CCPA attributed different meanings to the concept of privacy and data agency. Furthermore, SCOT analysis focuses on how the relationships between groups and the relevant technology shape the final form of the artifact, and often reveal further problems, potential solutions, and new insights into how a technology was formed and shaped. Ultimately, a better understanding of the CCPA's current legislation can be understood by analyzing the interpretive flexibility of privacy by consumers, legislators, corporations, and privacy advocates, and formulating the interactions between these groups through a SCOT lens. This analysis will also shed light onto how certain ideas became more dominant in the CCPA.

DATA AND THE CALIFORNIA CONSUMER PRIVACY ACT

The Importance of Data

Data is a nebulous concept, especially to consumers. Many technology companies claim to be *data driven* or use *big data* in their analyses and processes, but do not offer a concrete definition as to what their data consists of, or what consumer data really is. Erevelles, Fukawa, and Swayne (2016) define *big data* using the "Three V's: volume, velocity, and variety." In other words, big data is characterized by large, diverse datasets that are gathered quickly and through various means. Though the CCPA deals with big data, it deals specifically with a subset of data that is pertinent to consumers. This subset of data will be referred to as *personal data. The Harvard Business Review* lists several examples of types of personal data, including but not limited to: friends lists on certain apps, location data, browsing history, and chat history (Forbath et al., 2015). The CCPA focuses mainly on the protection and availability of these types of personal data, and the term "personal data" is used numerous times throughout the legislation (California Consumer Privacy Act, 2018). However, the definition of personal data varies from corporation to corporation, leading to ambiguity as to what data should be available to consumers. For example, vague definitions such as "personal data" and "households" manifest themselves in the CCPA's language, and for intentional reasons. This will be explored further using SCOT and various case analyses.

Another key concept in the analysis of the CCPA is *data agency.* Kennedy, Poell, and Dijck loosely describe data agency as the ability for a group or agency to understand what happens to data, how data-driven operations affect society, and the consequences of data analysis (Kennedy et al., 2015). The discrepancies in the agency between social groups is what ultimately led to the asymmetry in the legislative outcome of the CCPA, as different social groups have different amounts of power and influence over the data that they give or own.

The California Consumer Privacy Act

Consumer privacy has become an increasingly relevant topic in the modern era of big data and data analytics. Companies have begun amassing terabytes of data on customers with the intention of selling or leveraging their data to provide meaningful services to various customers. At its core, privacy has two different meanings to technology companies and consumers. One of the turning points for consumer's viewpoints on data privacy occurred in the wake of the Facebook and Cambridge Analytica Scandal. In the scandal, Cambridge Analytica, a data analytics firm in the United Kingdom, used improperly obtained Facebook user data to sway US elections. After it was revealed that nearly 80 million users had their data collected without their consent by Cambridge Analytica, many consumers began to resent technology companies and their handling of personal data (Confessore, 2018b). In the eyes of consumers and the government, it was Facebook's responsibility to vet data usage, and to protect user's privacy (Confessore, 2018b). However, data privacy pundits argued that Facebook instead prioritized profits over the protection of their user's data (Cyphers et al., 2018). The scandal showed that companies view privacy is a feature that users have to pay for, and a selling point. As Shoshana Zuboff (2015) puts it, consumers are ultimately data points to extract value from in the eyes of technology firms (p.79).

With the advent of accessible encryption technology, ad-blockers, and other privacy enhancing technologies, consumers are more aware of their ability to keep information confidential. However, consumers still lack control over how their data is used by other companies. Because of the stark power difference between consumers and companies, legislation has been put into place to even the playing field. In 2016, Europe implemented the General Data Protection Regulation (GDPR) to give consumers more agency in the data collection process. The United States followed suit recently with the California Consumer Privacy Act of 2018

(CCPA), which enables users to view and delete data collected on them, opt out of data collection, and not be discriminated against by companies for accessing or deleting data (California Consumer Privacy Act, 2018).

The CCPA was originally drafted by Alastair Mactaggart, after he realized that the US has no centralized policy on data protection and privacy like the GDPR (Confessore, 2018a). Mactaggart was a real estate agent in San Francisco, where he interacted with many up-and-coming engineers from prominent software firms. After talking with a Google employee about the data they collected on users, Mactaggart became interested in the extent to which technology companies compile consumer data (Confessore, 2018a). After some research, Mactaggart realized that many of the tracking services that these companies provide under the guise of free services are in fact, products to other businesses who use that data for targeted advertisements (Confessore, 2018a). As Mactaggart noted, technology companies and the government have been partners for the last few decades, with the most recent technology legal case being related to Microsoft in the early 1990's on antitrust grounds. In his article about the history of the CCPA, author Nicholas Confessore also claims that technology companies use all the data and money that they have to leverage their political ideals through the government. On these grounds, Mactaggart created a ballot initiative, which was eventually modified to be part of the first iteration of the CCPA (Confessore, 2018a). Though the first iteration of the CCPA was well intentioned, the differing viewpoints of the various stakeholders during the legislation process led to social and technical issues for both corporations and consumers.

SOCIAL CONSTRUCTION ANALYSIS

Three primary social groups were considered in the SCOT analysis - consumers, technology corporations, and privacy lobbyists. Technology corporations and consumers were

chosen as relevant social groups because the CCPA directly deals with interactions between users of data collection products, and the creators of those products. Privacy lobbyists were chosen as the final group because they represent an informed and opposing opinion to technology corporations, and played a large role in the legislative process. This contrasts with the slightly uninformed, but generally privacy aware views of the consumer group. Overall, the three groups represent a combination of informed, uninformed, and at times opposing viewpoints of the major parties effected by the CCPA.

The CCPA was originally drafted to give consumers increased agency over their personal data, and as such, consumers were the main focus of the language of the CCPA. Consumers interact with privacy technologies and their data in numerous ways, by virtue of living in a connected world with social media, smartphones, and other smart devices. According to prior research (Anderson & Vogels, 2019), only one in five Americans read company privacy policies, and of those, only one in ten fully comprehend the language in those policies. Additionally, nearly 64% of adults believe that companies might mishandle their personal data if it is used for external reasons (Anderson & Vogels, 2019). These data points show that there is a gap between what corporations expect consumers to know and understand based on their privacy policies, and how many consumers actually understand those policies.

Furthermore, the percentage of consumers who mistrust corporations shows that despite their efforts to outline data collection policies, consumers remain weary of how their data is being used. With recent events such as the Cambridge Analytical Scandal and more recently Whatsapp's privacy policy change, consumers have grown to distrust the way that corporations handle personal data. Though consumers are aware of privacy issues, they are less likely to act on those issues as other, more involved groups. This is likely because of an information

asymmetry between consumers and corporations, and an overall lack of information regarding privacy defences and best practices (Acquisti & Grossklags, 2005). Thus, the CCPA is a much needed first step in order to let consumers manage their personal data as they see fit. Everyday consumers seem to view data privacy and protection as important issues, and view the broad and sweeping definitions and rights granted in the CCPA as useful.

Privacy advocate groups form the next social group that influenced the CCPA legislative process. This social group differs from the consumer group in a few key ways - namely, their fervor in protecting privacy rights, and their increased understanding of the technological practices behind data collection and usage, and consists of groups such as the Electronic Frontier Foundation (EFF). This social group, like the consumer group, also believes in the right to privacy, though to a greater extent. This can be seen in articles by the EFF, where they refer to data privacy as an innate human right, and argue that the CCPA and similar legislation is a way to enforce these rights (O'Brien, 2020). The views of this group are represented in the CCPA through many of its policies, especially the policies pertinent to the deletion, access, and opt-out of data collection. Furthermore, the initial draft of the CCPA was created by privacy advocates, leading to a more extreme approach in its initial policies. Ultimately, this social group views privacy as a fundamental right, and likely views the CCPA as a step in the right direction.

Corporations, on the other hand, display an immense amount of leverage on how consumer data is handled, and have fundamentally different perceptions of privacy than consumers do. Corporations provide software and services to consumers, and they also store and use consumer data. For example, companies such as Google claim to have several billion devices running their software, with each device sending multiple streams of user telemetry data (Popper, 2017). To corporations, total privacy and data opt outs are antithetical to the business model that

they operate under. For example, to LinkedIn, each user was worth roughly $260 at the time of its acquisition by Microsoft (Short & Todd, 2017). Similar metrics can be compiled for companies such as Facebook and Google, whose main products derive from the analysis, collection, and subsequent sale of user data to third parties. Evidence of this view can be seen in lobbying documents and change proposals to the CCPA. For example, Dr. Lea Kissner, a former privacy executive at Google, suggested changes to the CCPA that involve allowing record keeping of personal information for the improvement of privacy enhancing systems, and to only provide *accessible* information to consumers (Part 5 of 7—Written Comments Received During 45-Day Comment Period—California Consumer Privacy Act (CCPA), 2019). This example, among numerous other examples in the open comments by corporations, show how user data is inextricably linked with technology companies, and how paradoxically, corporations need access to private user data to improve their privacy systems. Overall, this social group views privacy as something that can be improved on, but also views some of the more restrictive and ambiguous portions of the CCPA as encroachments on their business models. This has led to companies taking advantage of current measures outlined in the CCPA, ultimately resulting in less consumer agency over their data.

From an analysis of the imperative flexibility of all three of the aforementioned social groups, it is evident that each one of them views data privacy as an important aspect of modern technology. Though the extent to which each group believes in total privacy varies, the general importance and concept of data privacy has reached rhetorical closure between the three groups. However, the means through which data privacy should be preserved has not been established by each group, leading to issues with the current CCPA.

ANALYSIS

A SCOT analysis of the relevant social groups in relation to data privacy and the CCPA show that each different group has their own view on the role of data privacy, and the meaning of privacy as a whole. As a result, the measures passed in the CCPA suffered from a lack of closure of the perspectives of the relevant social groups, leading to unclear legislation that failed at its original purpose of giving consumers more data agency.

The initial suggestion for many of the ideas in the CCPA originally came from privacy advocates, who wanted strong sweeping reforms to data privacy. By virtue of the interpretive flexibility that privacy advocates inscribed onto data privacy, many of the initial regulations in the CCPA were driven by a "privacy as a right" perspective, and offered broad, sweeping, and ambiguous policies that were sometimes infeasible to implement for corporations. Following this, corporations, with their strong lobbying power, attempted to add various clauses to weaken the definitions of privacy and personal data that were in the CCPA (Tsukuyama, 2019). Ultimately, these two groups laid the foundations for the clauses present in the CCPA, and were the dominant forces in the passage and language of the legislation. This led to legislation that was broad enough to appease privacy advocates, but also ambiguous enough such that corporations could still exploit loopholes in the law. In essence, the CCPA lacks sufficient practical guidelines to enable privacy and data agency for consumers because of the strong interests of well-informed privacy advocates, and compromising corporations.

In his *Washington Post* article, journalist Greg Bensinger, (2020) describes scenarios where corporations take advantage of the CCPA. For example, the CCPA states that downloaded personal information should "…be in a portable, and to the extent technically feasible, readily usable format that allows the customer to transmit this information to another entity without hindrance" (California Consumer Privacy Act, 2018). However, the act itself does not provide a

standard format for exported data. Bensinger (2020) explains that companies take advantage of this lack of restrictions on data formatting by giving end users hard to parse files. While the files themselves are portable, they are inaccessible to a normal consumer, and often require technical knowledge to understand. This in turn reduces the average consumer's ability to understand what data is being collected, thus shifting the burden of data management from the companies onto the customer.

Another example of broad legislation that the CCPA introduces is the concept of personal data. The CCPA describes personal data as any information that can reasonably be attributed to a customer or household (California Consumer Privacy Act, 2018). However, many companies have interpreted the concept of personal data differently. Singer (2019) states that various companies give completely disparate sets of personal data to consumers because the CCPA does not provide a codified guideline on what data companies should disclose to users. This ultimately affects the end user, as companies may withhold data that is considered to be personal data.

Both examples of broad legislation focus on an underlying right to privacy, but fail to adequately address the practical implications of increasing data privacy agency for consumers.

DISCUSSION

One reason why corporate and privacy advocate viewpoints won over consumer viewpoints may be due to the discrepancy between caring about privacy and acting on privacy infraction by consumers. Consumers often care about data privacy, but do not actively protect their data due to their concerns in a phenomenon otherwise known as the *privacy paradox*. Barth and De Jong argue that one of the reasons that there is a discrepancy between consumer beliefs and actions is because of a knowledge deficiency between the average consumer and corporations (Barth & De Jong, 2017). This knowledge deficiency is not present between

technology-savvy privacy advocates and corporations, which leads to an underrepresentation of the concepts and needs of consumers. Much of this gap can be solved by consumer education on how data aggregation technologies work, how data is used, and what avenues data is collected by. Through fixes like these, consumers would become more aware of the rights that the CCPA gives them, and would make them more inclined to exercise those rights.

Additionally, many of the outcomes of the CCPA legislative process, such as ambiguous language, can be solved by clearer and more explicit definitions of terms, language and specifications that websites should follow. For instance, to increase accessibility for consumers requesting personal data from a website, a specific industry standard data format such as JSON or HTML should be mandated. This way, consumers would only need to know how to operate and open one type of file or format in order to access their data for a wide variety of companies. Furthermore, these file formats are widely supported and can be transferred to and from various services easily. To further protect consumer rights, if a company does not provide a clearly and easily transferrable file format, consumers should be able to file a legal complaint and directly receive their data while penalizing the company. Doing so would greatly increase consumer data literacy, while still appeasing corporations and privacy advocates by packaging data in clear, established, and easily transferable formats.

Similarly, stipulations such as the mandatory availability of data collection opt-outs should be standardized. A Consumer Reports study showed that only 57% of companies confirmed a user's prompt to opt out of data collection services (Fahs et al., 2021). Since opt-out confirmations are not explicitly stated in the CCPA, one fix for this issue would be to mandate opt-out notifications, and having a standard format for what information should be requested. This would allow consumers to easily have available information for each opt-out request, and

allow for acknowledgement from companies that a user's request was processed or not. This would also increase the transparency of the data collection process to consumers, which would also be a benefit to corporations, who would be vying to keep consumers engaged with their product and platform.

CONCLUSION

The CCPA is the US's first large step towards unified privacy legislation, and represents a new direction for US privacy policy in the wake of legislation such as the GDPR. However, the CCPA in its current state is not without its flaws. A SCOT analysis revealed that though each social group had varying viewpoints on privacy, each also held a central point that data privacy is a necessity in the modern era of technology. However, the overwhelming majority of interests came from well informed privacy advocacy groups, and corporate lobbyists. This led to data privacy reforms that pushed consumer privacy rights forward, but also lacked practical dimensions on how consumers would interface with their new rights.

This analysis of the CCPA legislative process through SCOT would enable future policymakers and engineers to realize the critical role that consumers play in the legislative and technological process, and to cater consumer-centric solutions around consumer needs. Further research in the subject area could analyze subsequent CCPA additions to see if consumer considerations are taken more into account. More generally, additional studies could also analyze how consumer focused technological legislation has been handled in the past, and compare that to the current iteration of the CCPA. As society becomes more data driven, it is essential that consumers, corporations, and legislators understand the social and technological dimensions of data privacy, and enable consumers to own the data that they provide.

REFERENCES

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Security and Privacy Magazine, 3(1), 26–33. https://doi.org/10.1109/MSP.2005.22

Anderson , M., & Vogels, E. (2019, October 9). Americans and digital knowledge. Retrieved April 5, 2021, from Pew Research Center: Internet, Science & Tech website: https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/

Barth, S., & de Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. Telematics and Informatics, 34(7), 1038–1058. https://doi.org/10.1016/j.tele.2017.04.013

Bensinger, G. (2020, January 21). So far, under California's new privacy law, firms are disclosing too little data-Or far too much. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2020/01/21/ccpa-transparency/

Bijker, W. E., & Pinch, T. J. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. Social Studies of Science, 14(3), 399–441. https://doi.org/10.1177/030631284014003004

California Consumer Privacy Act, Cal. Civil Code §§ 1798.100 – 1798.199. (2018). https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=CIV&division=3.&title=1.81.5.&part=4.&chapter=&article=

Confessore, N. (2018a). The unlikely activists who took on Silicon Valley—And won. *The New York Times*. Retrieved from https://www.nytimes.com/2018/08/14/ magazine/facebook-google-privacy-data.html

Confessore, N. (2018b). Cambridge Analytica and Facebook: The scandal and the fallout so far. *The New York Times*. Retrieved from https://www.nytimes.com/2018/04/04/ us/politics/cambridge-analytica-scandal-fallout.html

Cyphers, B., Gebhart, G., & Schwartz, A. (2018). Data privacy scandals and public policy picking up speed: 2018 in review. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2018/12/data-privacy-scandals-and-public-policy-picking-speed-2018-year-review

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. Journal of Business Research, 69(2), 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

Fahs, G., Mahoney, M., & Marti, D. (2021, February). The State of Authorized Agent Opt Outs Under the California Consumer Privacy Act. Retrieved April 6, 2021, from Consumer Reports website: https://advocacy.consumerreports.org/press_release/consumer-reports-study-finds-authorized-agents-can-empower-people-to-exercise-their-digital-privacy-rights-in-california

Feiner, L. (2019, August). Facebook and Google's dominance in online ads is starting to show some cracks. Retrieved from https://www.cnbc.com/2019/08/02/facebook-and-googles-ad-dominance-is-showing-more-cracks.html

Fertik, M. (2020, January). CCPA is a win for consumers, but businesses must now step up on cx. Retrieved from https://www.forbes.com/sites/michaelfertik/2020/01/27/ccpa-is-a-win-for-consumers-but-businesses-must-now-step-up-on-cx/

Forbath, T., Morey, T., & Schoop, A. (2015, May 1). Customer data: Designing for transparency and trust. *Harvard Business Review*. Retrieved from https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust

Kennedy, H., Poell, T., & van Dijck, J. (2015). Data and agency. *Big Data & Society*, *2*(2), 205395171562156. https://doi.org/10.1177/2053951715621569

O'Brien, D. (2020, January 28). Data privacy or data protection day? It's a human right, either way. Retrieved April 5, 2021, from Electronic Frontier Foundation website: https://www.eff.org/deeplinks/2020/01/data-privacy-or-data-protection-day-its-human-right-either-way

Part 5 of 7—Written Comments Received During 45-Day Comment Period—California Consumer Privacy Act (CCPA). (2019). Retrieved from California Office of the Attorney General website: https://www.oag.ca.gov/sites/all/files/agweb/pdfs/privacy/ccpa-comments-45day-pt5.pdf

Popper, B. (2017, May 17). Google announces over 2 billion monthly active devices on Android. Retrieved April 5, 2021, from The Verge website: https://www.theverge.com/2017/5/17/15654454/android-reaches-2-billion-monthly-active-users

Short, J., & Todd, S. (2017). What's your data worth? Retrieved April 5, 2021, from MIT Sloan Management Review website: https://sloanreview.mit.edu/article/whats-your-data-worth/

Singer, N. (2019, December 29). What does California's new data privacy law mean? Nobody agrees. *The New York Times.* Retrieved from https://www.nytimes.com/2019/12/29/technology/california-privacy-law.html

Tsukayama, H. (2019, July 10). California's senate judiciary committee blocks efforts to weaken california's privacy law. Retrieved April 6, 2021, from Electronic Frontier Foundation

website: https://www.eff.org/deeplinks/2019/07/californias-senate-judiciary-committee-blocks-efforts-weaken-californias-privacy

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89. https://doi.org/10.1057/jit.2015.5