

Ethics of Data Collection

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

Daniel Patel
Fall 2019, Spring 2020

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 **Daniel Patel** Date 4/27/2020
Daniel Patel

Approved _____ Date _____
Richard Jacques, Department of Engineering and Society

Abstract

In this paper I delve into the question of if Google's actions related to their data collection should be considered moral. This problem is approached through the lenses of Deontological ethics derived from a formulation of Kant's Categorical Imperative and the STS framework. In the process Google is compared to a form of business that the reader may be more familiar with thinking about to draw out a proper understanding of their actions, and, in so doing, come to a conclusion about how a rational entity should understand and judge their actions.

Introduction:

In the modern era, there has become a feeding frenzy surrounding a new, previously untapped natural resource. By name, this resource is personal information. Due to the relevance and still growing impact of personal data collection in modernity, it is fitting to address the means by which it is collected and used. To partake in this task, this paper addresses the specific data collection accomplished by Google, a company that acquires everything from a user's location to the contents of their emails ("Privacy Policy – Privacy & Terms – Google," n.d.). To this end, the paper views this shift to an economy of information through the lens of Technological Momentum to argue that both it is imperative to address these issues quickly, and that doing so remains efficacious. To couple with the framework of Technological Momentum, a deontological moral framework based off the Kingdom of Ends formulation of Kant's Categorical Imperative is employed in a policy analysis driven exploration of the aforementioned company. These methods and frameworks join to form a vantage point from which the moral quality of a modern company can be addressed. The question answered by this paper is whether Google's data collection and distribution policies should be considered moral.

Part 1: What is Collected

To answer the question about the inherent moral value of a set of actions, the quality and extent of the actions must first be flashed out. Because of this, it is vital to look into the actual information collected by Google before any judgment can be rightly made. To list the major areas of data acquisition: Google collects the phone numbers of people who call in to their help lines, every URL request made to Google by a user as well as the date and time thereof, the apps downloaded on android users phones as well as other device information such as location from IP and GPS, contacts, and readings from many built in sensors, videos watched, interaction with

ads, and comments on YouTube, information about the voice of a user with Google Assistant, purchase activity and billing information through Google Pay, every site a user visits on chrome, as well as many other data points about their users (“Privacy Policy – Privacy & Terms – Google,” n.d.; “Google Maps Platform Terms of Service | Google Maps Platform | Google Cloud,” n.d.; “Data Transparency | Google Safety Center,” n.d.; Android – Android Enterprise Terms,” n.d. -a). Similarly, Google acquires the rights to process, redistribute, collect usage statistics, and provide to third parties this personal information (“Google Play Terms of Service,” n.d.; “Terms and Conditions,” n.d.; “Terms of Service – YouTube,” n.d.; “YouTube Data Processing Terms – YouTube,” n.d.).

Part 2: Employed Frameworks

In addition to the facts of the question, a proper inspection of morality will always employ a moral framework, and in the paper, the framework used in this paper is be a branch of deontological ethics derived from the Kingdom of Ends Formulation of Kant’s Categorical Imperative. The Kingdom of Ends Formulation of The Categorical Imperative gives a binary test for the morality of an action: that one must “act in accordance with the maxims of a member giving universal laws for a merely possible kingdom of ends” (Johnson & Cureton, 2019). In Kantian Philosophy, a Kingdom of Ends can be thought of as an ideal state where a population of rational beings are ruled by their peers (Bennet, n.d.). Restated, this defines morality as acting in such a way that you would not be in violation of the laws a perfect legislature would use to govern itself and its equals. From this framework, a moral test arises: Is the action something a rational entity should allow to happen to themselves. According to Kant this formulation is a consistent test of morality equivalent to all other formulations of his Categorical Imperative (“The Categorical Imperative,” n.d.).

To couple with the moral framework above, the topic of data collection should be viewed through the lens of Technological Momentum. Technological Momentum is an STS framework that espouses technologies are much easier for society to shape in their infancy than in maturity, and therefore, both decisions and inactions regarding emerging fields can have large effects and repercussions.

Part 3: Arguments Against Google's Actions

Using these frameworks, to argue Google's actions are immoral one might put forward a line of thinking similar to the following: If morality can be boiled down to whether a rational actor should allow an action to be taken toward themselves, It is preeminent to ask the question of whether a rational actor should be fine with a large corporation recording massive amounts of information about them.

It seems reasonable to assume that the data such an actor would find reasonable to collect is data that cannot cause harm to the actor in the hands of the collecting agent. There is likely also some set of data about the actor that they would prefer not be collect pertaining to their personal life, but this only need be accounted for in the case that the first set of data described does not cause a violation of the moral framework. With regards, then, to the type of data collected, the actor would seemingly only allow data that cannot cause harm by the collector to the actor and only allow the information collected to be disseminated to parties which also could not cause harm to the actor.

Given this understanding of the nature of data a rational actor would allow to be collected about himself and the uses he would allow thereof, the morality of the actions of Google can now be inspected. Google states that, "We use the information we collect to customize our services

for you, including providing recommendations, personalized content, and customized search results” (“Privacy Policy – Privacy & Terms – Google,” n.d.). This statement implies that applications such as YouTube are programmed in such a way that a user will be constantly fed items that pertain to their interests, increasing usage time, and therefore decreasing time that could otherwise be spent in more useful ways. In addition to this Google is able to better serve advertisements to customers, allowing them to increase the ways in which advertisers can target specific demographics with specific products. In doing this, whether or not a user specifically turns off personalized ads, the net result is the data about preferences according to characteristics of people and their spending habits being disseminated to companies attempting to provide a product to a specific subset of the population. While in many cases this only serves to increase economic efficiency, it does, in some cases, serve to harm parts of the population. For an example of this, the reader can look to ads for things such as online gambling, an activity that on average hurts the user, and in the extreme case, can singly handedly deplete the savings of certain individuals predisposed to lacking self-control with respect to gambling habits. In instances such as this, it is clear that any information that would allow an online gambling company to target such a person should never end up in their hands. It follows that any company whose product can be linked to decreased wellbeing of a subset of its users should not be allowed access to data which would allow them to target such users according to this moral framework.

From this idea, it can be concluded that since Google both partakes in targeted advertising as well as sells data which it collects to companies that partake in such activities, it is nearly certain that, in at least one instance, Google provides a means by which a party whose data is collected ends up being harmed by the data, and therefore, that, with regards to the chosen

framework, Google's actions concerning data cannot be considered wholly moral, and thus, must be considered immoral.

Part 4: In Defense of Google

The argument above however, though seemingly compelling, neglects basic facets of the question that lead it to conclusions being misguided. Primarily, it is not as though Google just collects data about people, but rather about its users. In addition to this, at any time a user can turn off targeted advertising, see everything Google has collected about them, and have Google delete this information. These facts are important because they would necessarily inform the rational entity's decisions about what is acceptable. To make an analogy, consider a person buying an item from a company who cares deeply about customer service. While this person would likely not rationally hand their credit card to a stranger who they suspect will make a charge against it, here they will do so because it is in exchange for something they desire. In addition, if the object they acquire in this transaction is faulty they will likely be entitled to a refund. No one questions the morality of this situation, yet, in practice it is actually more onerous on the customer. With the customer in the imaginary store, they will likely have to return the item to get a refund, or, even if this isn't required, they still must make a complaint about what they received and await a response. In the case of Google's data collection, the user exchanges their data for the use of Google's services. In this exchange, if they feel that they didn't want Google to keep the "payment" of information they could go to their account and request for Google to delete the information it has collected about them, despite them still being able to use and enjoy the product. In business this would be equivalent to a company saying that they will give you a refund at any point, no questions asked, whether or not you like the product, and you can keep the item.

With this in mind, it seems perfectly reasonable to think that a rational entity would allow these actions to take place, and, if they believe the services were worthwhile, would likely make use of them. In addition to this when viewing the question with the idea of Technological Momentum in mind, if Google's practices are deemed to be unconscionable, and regulation is put in place to curtail some of their activities, while it may serve as a protection in its primary area of focus, it would likely serve a more costly role of further monopolizing the market. Google, and Alphabet Inc. its parent company are well established tech giants able to devote the resources to make changes in compliance with regulations while smaller startups will be more greatly affected. If small companies are forced to spend a larger portion of their revenue on ensuring regulation compliance, it lessens their profit, and therefore greatly reduces their chance at becoming viable. Similarly, any red tape put around the sharing of data collected by large companies could have undue effect on future collaborations in machine learning. Specifically machine learning requires large datasets to be able to complete tasks with reasonable accuracy, and if it becomes more difficult to share data, companies with data collection infrastructure will be the only ones easily able to create machine learning architectures capable of doing tasks that require a large amount of accuracy as they will be the only ones with the ability to collect it and will be made even more wary of sharing it than they otherwise would be due to the regulations.

Conclusion:

Given these ideas it seems that both a rational entity would allow for Google's actions, making them moral by the employed framework, and, due to technological momentum, any regulation meant to stop some of their actions could result in rather unfortunate consequences. Although Google collects a large amount of information on their users, they are transparent about what

they collect, how they use it, and what you can do about it, making the use of their products an exchange by consenting parties of amoral goods and therefore seemingly perfectly ethical.

References

Android – Android Enterprise Terms. (n.d.-a). Retrieved October 1, 2019, from Android website:

<https://www.android.com/enterprise/terms/>

Android – Android Enterprise Terms. (n.d.-b). Retrieved October 1, 2019, from Android

website: <https://www.android.com/enterprise/data-protection/terms/>

Bennett, J. (n.d.). Groundwork for the Metaphysic of Morals. *Immanuel Kant*, 53.

Data transparency | Google Safety Center. (n.d.). Retrieved October 1, 2019, from

<https://safety.google/privacy/data/>

Google Maps Platform Terms of Service | Google Maps Platform | Google Cloud. (n.d.).

Retrieved October 1, 2019, from <https://cloud.google.com/maps-platform/terms/>

Google Play Terms of Service. (n.d.). Retrieved October 1, 2019, from

<https://play.google.com/about/play-terms/index.html>

Johnson, R., & Cureton, A. (2019). Kant's Moral Philosophy. In E. N. Zalta (Ed.), *The Stanford*

Encyclopedia of Philosophy (Spring 2019). Retrieved from

<https://plato.stanford.edu/archives/spr2019/entries/kant-moral/>

Privacy Policy – Privacy & Terms – Google. (n.d.). Retrieved October 1, 2019, from

<https://policies.google.com/privacy?hl=en#infocollect>

Terms and conditions. (n.d.). Retrieved October 1, 2019, from Android Developers website:

<https://developer.android.com/studio/terms>

Terms of Service—YouTube. (n.d.). Retrieved October 1, 2019, from

<https://www.youtube.com/static?template=terms>

The Categorical Imperative. (n.d.). Retrieved October 1, 2019, from

<https://www3.nd.edu/~rbarger/categorical-imperative.html>

YouTube Data Processing Terms—YouTube. (n.d.). Retrieved October 1, 2019, from

https://www.youtube.com/t/terms_dataprocessing